В.С. Шипачев

КУРС ВЫСШЕЙ МАТЕМАТИКИ

В.С. Шипачев

КУРС ВЫСШЕЙ МАТЕМАТИКИ

Учебник для вузов

4-е издание, исправленное

Под редакцией академика
А.Н. Тихонова

Рекомендовано
Научно-методическим советом
Министерства образования и науки
Российской Федерации
в качестве учебника для студентов вузов

Москва
ОНИКС

УДК 510(075.8)
ББК 22.1я73
Ш63

Рецензенты:
зав. кафедрой автоматизации научных исследований факультета ВМК МГУ им. М.В. Ломоносова чл.-корр. РАН, проф. *Д.П. Костомаров;*
зав. кафедрой математики физического факультета МГУ им. М.В. Ломоносова д-р физ.-мат. наук, проф. *В.Ф. Бутузов*

Шипачев В.С.

Ш63 Курс высшей математики: Учебник для вузов / В.С. Шипачев; Под ред. А.Н. Тихонова. — 4-е изд., испр. — М.: Издательство Оникс, 2009. — 608 с.: ил.

ISBN 978-5-488-02067-2

В учебнике излагается материал по важным разделам высшей математики, таким, как сведения из теории множеств и теории вещественного числа, теория пределов последовательностей и функций, основы дифференциального и интегрального исчисления функций одной и нескольких переменных, элементы аналитической геометрии на плоскости и в пространстве, некоторые вопросы линейной и векторной алгебры, теории рядов и теории обыкновенных дифференциальных уравнений. Существенное внимание уделено решению типовых примеров и задач теоретического и прикладного характера.

Предназначен для студентов самых различных специальностей университетов и технических вузов.

УДК 510(075.8)
ББК 22.1я73

Учебное издание

Шипачев Виктор Семенович

КУРС ВЫСШЕЙ МАТЕМАТИКИ

Учебник для вузов

Заведующая редакцией *Л.В. Дудник*
Редактор *Т.И. Балашова*
Дизайн обложки *А.Л. Чириков*
Технический редактор *Л.А. Данкова*
Корректор *И.И. Иванова*

Подписано в печать 01.08.2008. Формат 60x90 1/16. Гарнитура «Таймс».
Усл. печ. л. 38. Тираж 5000 экз. Заказ № 5903.

Общероссийский классификатор продукции
ОК-005-93, том 2; 953005 – литература учебная

ООО «Издательство Оникс». 127422, Москва, ул. Б. Почтовая, д. 7, стр. 1
Отдел реализации: тел. (499) 619-02-20, 619-31-88. Интернет-магазин: www.onyx.ru

Издание осуществлено при техническом содействии ООО «Издательство АСТ»

ОАО «Владимирская книжная типография».
600000, г. Владимир, Октябрьский проспект, д. 7.
Качество печати соответствует качеству предоставленных диапозитивов

ISBN 978-5-488-02067-2

ПРЕДИСЛОВИЕ

Настоящий «Курс высшей математики» является расширенным изложением лекций, которые автор читал студентам геологического факультета Московского государственного университета им. М.В. Ломоносова. Учебник может быть использован студентами факультетов нематематического профиля, на которых различные разделы высшей математики объединены в один курс.

Автор стремился изложить материал по возможности полно, строго и доступно и ставил своей целью не просто сообщить читателю те или иные сведения по высшей математике, а развить у него математическое мышление, расширить кругозор и привить ему математическую культуру; показать внутреннюю связь математических понятий, т. е. представить математику в ее развитии.

Учебник содержит краткие сведения из теории множеств, теорию пределов, элементы аналитической геометрии на плоскости и в пространстве, основы дифференциального и интегрального исчислений функций одной и многих переменных, понятия о матрицах, об определителях, о рядах и теорию дифференциальных уравнений.

Основу курса составляет математический анализ, включающий в себя дифференциальное и интегральное исчисления, о которых упоминалось выше. В нем изучается важнейшее понятие высшей математики — понятие функции. С ним читатель уже знаком из курса элементарной математики. Однако полное и систематическое изучение проводится именно в высшей математике. Учитывая современные тенденции в преподавании математики, понятие функции определяется через понятие множества.

Опыт показал, что для многих начинающих значительную трудность представляет решение задач. Поэтому в учебнике главное внимание уделено решению типовых примеров и задач. Однако прежде чем начать их решать, надо сначала изучить нужный раздел высшей математики и добиться полной ясности в понимании соответствующих понятий и теорем. Для практических занятий следует использовать учебное пособие автора «Задачник по высшей математике».

Автор выражает глубокую благодарность академику А.Н. Тихонову за ценные советы и большую помощь; академику РАН В.А. Садовничему, членам-корреспондентам РАН А.В. Бицадзе, Л.Д. Кудрявцеву и Д.П. Костомарову; профессорам А.Г. Свешникову, Ш.А. Алимову и В.Ф. Бутузову за просмотр рукописи и сделанные ими весьма полезные замечания.

Автор

Ни одно человеческое исследование не может называться истинной наукой, если оно не прошло через математические доказательства.

Леонардо да Винчи *

ВВЕДЕНИЕ

Математика — самая древняя и в то же время самая юная из наук. Она стала складываться где-то во втором тысячелетии до н. э., когда потребности торговли, землемерия и мореплавания заставили упорядочить приемы счета и измерения, начало которых уходит еще в более глубокую древность. Уже строители египетских пирамид владели математическими знаниями.

В Древней Греции начиная с VI в. до н. э. математика приобретает статус самостоятельной науки, окончательное оформление которой было осуществлено в III в. Евклидом в его бессмертных «Началах». По этой книге или по ее более доступным изложениям изучали геометрию в течение более двух тысяч лет. Отсюда видно, что математика значительно отличается от всех других наук. Теоретические представления Аристотеля в области физики сейчас кажутся несколько наивными, они стали достоянием истории науки, хотя в свое время в них были обобщены все имевшиеся знания об окружающем мире. Теорема же Пифагора действительна и поныне и составляет одну из основ геометрии.

Раз сложившись в качестве науки, математика не переставала развиваться, разрабатывать новые методы, открывать новые области, совершенствовать свою символику, свой научный аппарат. Возникновение физики нового времени было связано с непосредственным применением математики Кеплером и Галилеем для изучения небесных и земных явлений. Великий поворот-

* Леонардо да Винчи. Избранные естественнонаучные произведения. М.: Изд-во АН СССР, 1955.

ный пункт в истории математики наступил в XVII в., когда Декарт создал аналитическую геометрию, а Ньютон и Лейбниц — дифференциальное интегральное исчисление. Эти открытия в огромной степени увеличили возможности математики как для ее собственного развития, так и для содействия успехам других наук, таких, как физика и астрономия.

Бурное развитие математики, последовавшее за этими открытиями, привело на рубеже XIX—XX вв. к новой революции, связанной, в частности, с признанием правомерности неевклидовых геометрий (Лобачевского, Римана, Бойяи) и созданием теории множеств Кантором. С тех пор математика продолжает развиваться, поражая воображение многообразием своих специальных областей, новизной и необычностью используемых представлений и понятий, неожиданным своеобразием своих методов, особенностями своего языка. Процесс дифференциации наук охватил и математику, приведя к возникновению внутри нее множества отраслей.

Одновременно с развитием методов и отраслей математики происходило и ее внедрение в другие науки, шел процесс так называемой математизации науки. В результате логики развития самой науки математика превратилась в метод научного исследования. Если в период классической физики математика служила преимущественно средством обработки экспериментальных данных, установления точного количественного отношения между открытыми физическими явлениями и процессами, то уже к концу XIX в. математические вычисления стали предварять физические гипотезы и открытия. Математика не только обрабатывала показания приборов и результаты экспериментов, но она стала идти впереди них, создавая такие математические модели, реальный физический смысл которых, еще не вполне ясный, предстояло открыть.

Именно это явление нередко получало неправильное истолкование, отчасти со стороны самих ученых, но особенно со стороны идеалистических философов, и было зафиксировано в известном афоризме «материя исчезла, остались одни уравнения».

Суть процесса, о котором идет речь, состоит в том, что физик, пользуясь математическими методами, может проникать в еще не исследованные области физического мира, пока недоступные физическим методам, открывать в них математические закономерности, создавать математические модели неизвестных еще физических процессов и тем самым толкать и направлять мысль экспериментатора. В наше время физик-теоретик — это прежде всего математик. «Математика для физика, — говорит крупный американский ученый Ф. Дж. Дайсон, — это не только инструмент, с помощью которого он может количественно описать любое явление, но и главный источник представлений и принципов, на основе которых зарождаются новые теории» *.

Выразительным примером роли математического мышления для физических открытий может служить общая теория относительности Эйнштейна. Вся теория была завершена до экспериментальной проверки, а когда такая проверка была произведена, ее результаты практически совпали с предсказанными теорией.

Не менее убедительный пример дает история возникновения квантовой механики, которая была первоначально построена чисто математическим путем, конечно, на основе некоторых известных, но не объясненных в то время физических данных, в результате гениального, но чисто «умозрительного скачка математического воображения» (Ф. Дайсон). Она не только была немедленно подтверждена соответствующими экспериментами, но явилась источником и стимулом дальнейшего развития физики микромира со всеми ее впечатляющими результатами, роль которых в XX в. слишком хорошо известна, чтобы о ней стоило говорить снова.

У всех, кто изучал историю математики и ее применения в науках о природе и в технике и размышлял об отношении математики к объективному миру, будь это сами математики, физики или философы, неизбежно возникал вопрос о чудесной способности математики давать правильное описание или отображение физических процессов, поведения физической вселенной.

---

* Дайсон Ф. Дж. Математика в физических науках.— В кн.: Математика в современном мире. М., 1967. С. 112.

Ученые, стоявшие у колыбели современной науки, такие, как Кеплер и Галилей, пораженные достигнутыми ими результатами, считали, что книга природы написана ее божественным творцом на языке математики, так что ученому остается только прочитать эти записи. С тех пор высказывалось множество предположений о природе математики и ее познавательной способности, но ни одно из них не получило всеобщего признания.

В XX в. одним из западных философов Витгенштейном была высказана поразительная мысль, что вся математика есть не что иное, как совокупность тавтологий, а математические доказательства представляют собой тавтологические преобразования. Эта теория объясняла абсолютную достоверность математики и ее универсальную применимость. Но она была бессильна объяснить способность математики открывать новое в мире, т. е. ту ее способность, которая является важнейшей для развития науки и дает возможность все более широкого применения математики в специальных науках.

Следует подчеркнуть, что математика оперирует не только абстракциями (это делают все науки), но абстракциями весьма высокой степени. Даже любое из самых обычных натуральных чисел, например 4, есть абстракция, отвлекающаяся от всех специфических особенностей каких-либо четырех предметов (деревьев, ножек стола, углов дома и т. д.), характеризуя лишь класс, имеющий 4 члена. Понятие же натурального числа — это абстракция еще более высокая, поскольку оно представляет собой класс всех классов, имеющих не менее одного члена.

Сила математики именно в ее способности создавать все более высокие абстракции, оперировать ими и изучать их особенности и закономерности. Именно поэтому она может применяться к различным другим наукам и помимо физики, по мере того как они сами становятся теоретическими, т. е. начинают создавать достаточно высокие абстракции и пользоваться ими.

Немецкий философ XVIII в. Иммануил Кант считал, что наука тем более заслуживает названия науки, чем больше в ней математики. В то время математика была неотъемлемым элементом лишь механики, физики и астрономии. В наше время науки настолько повысили свой теоретический уровень, а математика настолько разнообразила и усовершенствовала свои методы, что их слияние оказалось не только возможным, но и абсолютно необходимым как для развития этих наук, так и для самой математики.

Естественно, что процесс математизации не в одинаковой степени затронул все науки. Огромным успехом является применение математических методов, помимо наук о неживой природе, также к исследованиям в области биологии. Это оказалось возможным главным образом благодаря проникновению биологии во внутриклеточные процессы и анализу их на молекулярном уровне. В качестве примера здесь можно назвать хотя бы исследование функционирования и построение моделей некоторых функций нейрона и изучение проблем наследственности и расшифровки генетического кода.

Но роль математики в биологии этим не исчерпывается. Как говорит Э. Мур, «особая ценность математики для биологии состоит не в применении ее как аппарата исследований, а в возможности абстрактно подойти к решению фундаментальных проблем и обнаружить связи между принципиально различными явлениями и процессами»*.

В общественных науках, которые больше всего были изолированы от математики, если не считать применения статистических методов к исследованию некоторых социальных процессов и явлений, можно назвать такие весьма различные области, как проблемы демографии и проблемы структурной лингвистики, где применение математики дало превосходные результаты. Но, пожалуй, наиболее значительным научным достижением было внедрение математических методов в экономическую науку и в руководство экономическими процессами. В наше время научное управление этими процессами может быть осуществлено только на основе применения точных экономических методов ко всем сторонам и сферам народного хозяйства — от прогнозирования размещения полезных ископаемых до изучения спроса на товары широкого потребления и бытовые услуги, от изучения потребности в рабочей силе до плани-

---

* Мур Э. Математика в биологических исследованиях // Математика в современном мире. М., 1967. С. 129.

рования транспортных артерий, пассажирских перевозок и экспериментов по искусственному воздействию на атмосферные явления.

Короче говоря, жизнь современного человека невозможна без математики.

О какой, однако, математике идет здесь речь: о так называемой «чистой» или прикладной? Но это традиционное разграничение в настоящее время становится все более и более условным и утрачивает свой первоначальный смысл. Даже наиболее абстрактные отделы «чистой» математики могут, оказывается, получить конкретное приложение в самых неожиданных областях науки и техники. В то же время потребности в решении теоретических и практических проблем стимулируют разработку новых абстрактных методов и отраслей математической науки.

В наше время в связи с возросшей ролью математики в современной науке и технике будущие экономисты, биологи, геологи, социологи и т. д. нуждаются в серьезной математической подготовке. Этим определяется место математики в системе высшего образования. Современный научный работник или инженер должен не только знать основы математики, но и хорошо владеть всеми новейшими математическими методами исследования, которые могут применяться в области его деятельности. Смежные науки используют различный объем математических знаний и ставят новые задачи в изучении самой математики. Можно с уверенностью сказать, что изучение математики способствует усвоению самого современного стиля научного мышления и является условием его применения в конкретных науках.

Часть I
АНАЛИЗ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

Глава I. ВВОДНЫЕ ЗАМЕЧАНИЯ

§ 1. Множества. Обозначения

Понятие множества является одним из основных в математике. Это понятие принадлежит к числу так называемых первичных, неопределяемых через более простые. Слова: «совокупность», «семейство», «система», «набор», «объединение» и т. п. — являются синонимами слова «множество». Примерами множеств могут служить множество студентов данного университета; совокупность тех из них, кто сдал вступительные экзамены без троек; семейство звезд Большой Медведицы; набор трех уравнений с тремя неизвестными; множество всех целых чисел и т. д. Из приведенных примеров следует, что множество может содержать конечное и бесконечное число объектов произвольной природы.

Объекты, из которых состоит множество, называются его *элементами*, или *точками*. Множества чаще всего обозначаются большими, а их элементы — малыми буквами. Если x — элемент множества X, то пишут $x \in X$ (x принадлежит X). Если x не является элементом множества X, то пишут $x \notin X$ (x не принадлежит X). Если $x_1, \ldots, x_n$ — некоторые элементы, то запись $X = \{x_1, \ldots, x_n\}$ означает, что множество X состоит из элементов $x_1, \ldots, x_n$. Аналогичный смысл имеет запись $X = \{x_1, x_2, x_3, \ldots\}$.

Пусть X и Y — два множества. Если X и Y состоят из одних и тех же элементов, то говорят, что они совпадают, и пишут $X = Y$. Если в X нет элементов, не принадлежащих Y, то говорят, что X содержится в Y или что X есть *подмножество* множества Y. В этом случае пишут $X \subset Y$ или $Y \supset X$ (X содержится в Y или Y содержит X).

Множество, не содержащее ни одного элемента, называется *пустым* множеством и обозначается символом $\varnothing$.

Пустое множество является подмножеством любого множества.

Множество с установленным порядком расположения элементов называют упорядоченным. Упорядоченное множество в отличие от просто множества записывают внутри круглых скобок. Например, из двух элементов x_1 и x_2 можно составить два упорядоченных множества: $(x_1; x_2)$ и $(x_2; x_1)$.

В дальнейшем нам придется иметь дело с различными множествами вещественных чисел*. Всюду, где это не может привести к недоразумению, для краткости вещественные числа будем называть просто числами.

Пусть X — множество, а $P(x)$ — какое-то свойство. Тогда запись $\{x|P(x)\}$ обозначает совокупность тех элементов множества X, которые обладают свойством $P(x)$. Например, $\{x|x^2-3x+2=0\}$ есть совокупность вещественных корней уравнения $x^2-3x+2=0$, т. е. это множество из двух элементов: 1 и 2; $\{x|3<x<7\}$ есть множество всех чисел, удовлетворяющих неравенствам $3<x<7$; $\{x|x>7$ и $x<3\}=\varnothing$, т. е. пустое множество.

Пусть a и b — два числа и $a<b$. Тогда пользуются следующими обозначениями:

$$\{x|a\leqslant x\leqslant b\}=[a, b];\ \{x|a< x\leqslant b\}=(a, b];$$

$$\{x|a\leqslant x< b\}=[a, b);$$

$$\{x|a< x< b\}=(a, b);\ \{x|a\leqslant x\}=[a, +\infty);$$

$$\{x|a< x\}=(a, +\infty);$$

$$\{x|x\leqslant b\}=(-\infty, b];\ \{x|x< b\}=(-\infty, b);$$

$$\{x|-\infty< x+\infty\}=(-\infty, +\infty).$$

Все эти множества называются *промежутками*, причем $[a, b]$ называется *отрезком* (или *сегментом*), $[a, b)$, $(a, b]$, $[a, +\infty)$ и $(-\infty, b]$ — *полуинтервалами*, а (a, b), $(a, +\infty)$, $(-\infty, b)$ и $(-\infty, +\infty)$ — *интервалами*. Промежутки $[a, b]$ $(a, b]$, $[a, b)$ и (a, b) называются *конечными*, a и b называются их концами. Остальные промежутки называются *бесконечными*.

---

* Вместо термина *вещественные числа* часто употребляют термин *действительные числа*.

Множество $(-\infty, +\infty)$ всех чисел называется также *числовой прямой*,* или *числовой осью*, а любое число — *точкой* этой прямой. Пусть a — произвольная точка прямой и δ — положительное число. Тогда любой интервал $(a-\delta,\ a+\delta)$ называется δ-окрестностью точки a.

Если $x_1, \ldots, x_n$ — произвольные числа, то запись $x = \max\{x_1, \ldots, x_n\}$ $(x=\min\{x_1, \ldots, x_n\})$ означает, что число x максимальное (минимальное) из чисел $x_1, \ldots, x_n$.

§ 2. Вещественные числа и их основные свойства

Напомним, что множество вещественных чисел разбивается на два множества — *рациональные* и *иррациональные* числа. Любое вещественное число представляется бесконечной десятичной дробью. *Рациональным* числом называется число, которое можно представить в виде p/q, где p и q — целые числа, причем $q \neq 0$. *Иррациональным* числом называется всякое вещественное число, которое не является рациональным. Как известно, всякое рациональное число p/q является либо целым, либо представляется конечной или периодической бесконечной десятичной дробью. Иррациональное же число представляется непериодической бесконечной десятичной дробью. Например, рациональные числа 3/4 и 1/3 представляются следующими десятичными дробями: $3/4 = 0{,}75$; $1/3 = 0{,}333\ldots$; иррациональные числа $\sqrt{2}$ и π представляются непериодическими бесконечными десятичными дробями: $\sqrt{2} = 1{,}41421356\ldots$, $\pi = 3{,}14159\ldots$.

Перечислим основные свойства вещественных чисел.

I. Сложение и умножение вещественных чисел. Для любой пары a и b вещественных чисел определены, и притом единственным образом, два вещественные числа $a+b$ и $a \cdot b$, называемые их *суммой* и *произведением*, обладающие следующими свойствами. Каковы бы ни были числа a, b и c:

1) $a+b=b+a$ (переместительное свойство);
2) $a+(b+c)=(a+b)+c$ (сочетательное свойство);
3) $a \cdot b = b \cdot a$ (переместительное свойство);
4) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (сочетательное свойство);
5) $(a+b) \cdot c = ac + bc$ (распределительное свойство);
6) существует единственное число 0 такое, что $a+0=a$ для любого числа a;

---

* **Числовой прямой называется прямая, на которой выбраны точка,** являющаяся началом отсчета, масштабный отрезок и положительное направление.

7) для любого числа a существует такое число $-a$, что $a+(-a)=0$;
8) существует единственное число $1 \neq 0$ такое, что для любого числа a имеет место равенство $a \cdot 1 = a$;
9) для любого числа $a \neq 0$ существует такое число a^{-1}, что $a \cdot a^{-1} = 1$, число a^{-1} обозначается также символом $\frac{1}{a}$.

II. Неравенства между вещественными числами. Для любых двух различных вещественных чисел a и b установлено одно из отношений $a > b$ или $b > a$ (a больше b или b больше a).

Отношение $>$ обладает следующими свойствами. Каковы бы ни были числа a, b и c:

1) если $a > b$ и $b > c$, то $a > c$;
2) если $a > b$, то $a + c > b + c$;
3) если $a > 0$ и $b > 0$, то $a \cdot b > 0$.

Вместо $a > b$ пишут также $b < a$ (b меньше a). Запись $a \geqslant b$ (или, что то же, $b \leqslant a$) означает, что либо $a = b$, либо $a > b$. Соотношения $a < b$, $a \leqslant b$, $a > b$, $a \geqslant b$ называются *неравенствами*. Неравенства $a < b$ и $a > b$ называются *строгими неравенствами*.

III. Непрерывность вещественных чисел. Вещественные числа обладают свойством *непрерывности*.

Пусть X и Y — два множества, состоящие из вещественных чисел. Тогда, если для любых чисел $x \in X$ и $y \in Y$ выполняется неравенство $x \leqslant y$, то существует хотя бы одно число c такое, что для любых чисел x и y выполняются неравенства $x \leqslant c \leqslant y$.

Свойство непрерывности имеет простой геометрический смысл. В самом деле, если возьмем числовую прямую, то на ней каждая точка $x \in X$ будет расположена левее каждой точки $y \in Y$. Поэтому множество X расположено целиком левее множества Y. Согласно свойству непрерывности между множествами X и Y есть точка c, «отделяющая одно множество от другого» (рис. 1). При этом точка c может принадлежать как

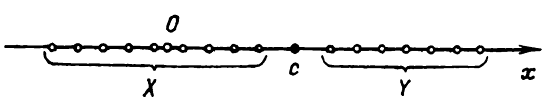


Рис. 1

множеству X, так и множеству Y, а также не принадлежать ни одному из них. Таким образом, числовая прямая является как бы сплошной линией без «дырок». В каком бы месте мы ни разрезали прямую на две части, разрез пройдет через одну из точек прямой.

Следует заметить, что свойством непрерывности обладает множество всех вещественных чисел, но им не обладает множество только рациональных чисел. Действительно, пусть множество X состоит из рациональных чисел x, для которых выполняется неравенство $x<\sqrt{2}$, а множество Y состоит из рациональных чисел y, для которых выполняется неравенство $y>\sqrt{2}$. Тогда, очевидно, для любого $x\in X$ и любого $y\in Y$ выполняется неравенство $x\leqslant y$, однако не существует рационального числа c такого, чтобы выполнялись неравенства $x\leqslant c\leqslant y$. В самом деле, таким числом могло бы быть только $\sqrt{2}$, которое, как известно, не является рациональным.

Перечисленные свойства вещественных чисел можно принять за аксиомы.

Мы ограничились здесь перечислением наиболее характерных свойств вещественных чисел, считая, что этих сведений достаточно для изучения последующего материала *.

§ 3. Грани числовых множеств

Говорят, что множество X ограничено сверху (снизу), если существует число c такое, что для любого $x\in X$ выполнено $x\leqslant c$ $(c\leqslant x)$. Число c в этом случае называется *верхней* (*нижней*) гранью множества X.

Множество, ограниченное и сверху и снизу, называется *ограниченным.*

Примеры. Любой конечный промежуток ($[a, b]$, $[a, b)$, $(a, b]$, (a, b)) ограничен. Интервал $(a, +\infty)$ есть множество, ограниченное снизу, но не ограниченное сверху; а интервал $(-\infty, +\infty)$ есть множество, не ограниченное ни сверху, ни снизу.

Очевидно, что любое ограниченное сверху (снизу) множество X имеет бесконечно много верхних (нижних) граней, образующих множество чисел, ограничивающих X сверху (снизу). В самом деле, если число c является

* Желающим ознакомиться с теорией вещественных чисел более полно рекомендуем книгу: Кудрявцев Л. Д. «Курс математического анализа». М., 2006. Т. 1.

верхней (нижней) гранью множества X, то любое число c', большее (меньшее) числа c, также является верхней (нижней) гранью множества X, так как из справедливости неравенства $x \leqslant c$ $(c \leqslant x)$ будет следовать, что $x \leqslant c'$ $(c' \leqslant x)$.

Естественно, возникает вопрос о существовании наименьшего из чисел ограниченного сверху множества и наибольшего из чисел ограниченного снизу множества.

Наименьшее из чисел, ограничивающих множество X сверху, называется *точной верхней гранью** множества X и обозначается символом $\sup X$**, а наибольшее из чисел, ограничивающих множество X снизу, называется точной нижней гранью этого множества и обозначается символом $\inf X$***.

Примеры. Пусть $X = (a, b)$. Тогда b является точной верхней гранью множества X, а a — точной нижней гранью, т. е. $b = \sup X$, $a = \inf X$. Пусть $X = (a, +\infty)$. Тогда $a = \inf X$, а точной верхней грани данное множество не имеет.

Точная верхняя грань $\sup X$ обладает следующим важным свойством. Как бы мало ни было число $\varepsilon > 0$, найдется число $x \in X$ такое, что $x > \sup X - \varepsilon$. Если бы такого числа x не нашлось, то число $\sup X - \varepsilon$ было бы также верхней гранью, и тогда число $\sup X$ не было бы точной верхней гранью. Другими словами, данное свойство выражает тот факт, что число $\sup X$ является наименьшим среди чисел, ограничивающих множество X сверху, и уменьшено быть не может.

Аналогичным свойством обладает и точная нижняя грань — как бы мало ни было число $\varepsilon > 0$, найдется число $x \in X$ такое, что $x < \inf X + \varepsilon$.

Возникает вопрос, при каких условиях числовое множество имеет точную верхнюю (нижнюю) грань. Ответ дает следующая весьма важная

Теорема 1.1. *Любое непустое ограниченное сверху (снизу) числовое множество имеет точную верхнюю (нижнюю) грань.*

Доказательство. Пусть X — непустое множество, ограниченное сверху. Тогда множество Y чисел, ог-

---

* Наряду с термином *точная верхняя грань* (*точная нижняя грань*) употребляют термины *верхняя грань* (*нижняя грань*).

** supremum (*лат.*) — наивысшее.

*** infimum (*лат.*) — наинизшее.

раничивающих X сверху, непусто. Из определения верхней грани следует, что для любого $x \in X$ и любого $y \in Y$ имеет место неравенство $x \leqslant y$. В силу свойства непрерывности существует такое число c, что для любых x и y выполняются неравенства

$$x \leqslant c \leqslant y. \tag{1}$$

Из первого из неравенств (1) следует, что число c ограничивает множество X сверху, т. е. является верхней гранью, а из второго — что оно — наименьшее из таких чисел,* т. е. является точной верхней гранью, причем может как принадлежать, так и не принадлежать множеству X.

Случай существования точной нижней грани у непустого ограниченного снизу множества рассматривается аналогично** ■

Если множество X не ограничено сверху (снизу), условимся писать: $\sup X = +\infty$ $(\inf X = -\infty)$.

§ 4. Абсолютная величина числа

Понятием абсолютной величины числа и неравенствами, связанными с абсолютными величинами, нам в дальнейшем придется очень часто пользоваться.

Определение. Абсолютной величиной, или модулем, числа x называется само число x, если $x \geqslant 0$, или число $-x$, если $x < 0$.

Абсолютная величина числа x обозначается символом $|x|$. Таким образом:

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

Из определения вытекает ряд свойств абсолютной величины числа.

$|x| \geqslant 0$. Действительно:

1) если $x \geqslant 0$, то $|x| = x \geqslant 0$;

2) если $x < 0$, то $|x| = -x$; но $-x > 0$, так как $x < 0$, т. е. $|x| > 0$.

Из 1) и 2) получаем, что $|x| \geqslant 0$.

---

* Так как $c \leqslant y$ для всех $y \in Y$.

** Здесь и в дальнейшем символ ■ будет означать конец доказательства.

II. $|x|=|-x|$. Действительно:

1) если $x \geqslant 0$, то $-x \leqslant 0$, и тогда $|-x|=-(-x)=$ $=x=|x|$;

2) если $x<0$, то $-x>0$, и тогда $|-x|=-x=|x|$, так как $x<0$.

Из 1) и 2) получаем: $|x|=|-x|$.

III. $-|x| \leqslant x \leqslant |x|$. Действительно:

1) если $x \geqslant 0$, то $|x|=x$ и подавно $|x| \geqslant -x$, т. е. $-|x| \leqslant x$;

2) если $x<0$, то $x<-x$, так как $2x<0$, $x+x<0$, $x<-x$, т. е. $x<|x|$.

Из 1) и 2) получаем, что $-|x| \leqslant x \leqslant |x|$.

Следующие три свойства в силу важности докажем в виде теорем.

Теорема 1.2. *Пусть ε — положительное число. Тогда неравенства $|x| \leqslant \varepsilon$ и $-\varepsilon \leqslant x \leqslant \varepsilon$ равносильны.*

Доказательство. Пусть $|x| \leqslant \varepsilon$. Тогда:

1) если $x \geqslant 0$, то $|x|=x$, т. е. $x \leqslant \varepsilon$, откуда $0 \leqslant x \leqslant \varepsilon$;

2) если $x<0$, то $|x|=-x$, т. е. $-x \leqslant \varepsilon$, откуда $-\varepsilon \leqslant$ $\leqslant x \leqslant 0$.

Объединяя 1) и 2), при любом x получаем: $-\varepsilon \leqslant x \leqslant \varepsilon$.

Пусть справедливо неравенство $-\varepsilon \leqslant x \leqslant \varepsilon$. Это значит, что одновременно выполняются неравенства $x \leqslant \varepsilon$ и $x \geqslant -\varepsilon$. Из последнего имеем: $-x \leqslant \varepsilon$. Так как по определению $|x|$ есть либо x, либо $-x$, то $|x| \leqslant \varepsilon$ ■

Теорема 1.3. *Абсолютная величина суммы двух чисел не больше суммы абсолютных величин этих чисел, т. е. $|x+y| \leqslant |x|+|y|$.*

Доказательство. Пусть x и y — любые числа. В силу свойства III для них справедливы неравенства

$$-|x| \leqslant x \leqslant |x| \quad \text{и} \quad -|y| \leqslant y \leqslant |y|.$$

Складывая их почленно, получим

$$-(|x|+|y|) \leqslant x+y \leqslant (|x|+|y|).$$

По теореме 1.2 это двойное неравенство равносильно неравенству $|x+y| \leqslant |x|+|y|$ ■

Заметим, что $|x-y| \leqslant |x|+|y|$.

Теорема 1.4. *Абсолютная величина разности двух чисел не меньше разности абсолютных величин этих чисел, т. е. $|x-y| \geqslant |x|-|y|$.*

Доказательство. Для любых чисел x и y имеем

$$x=y+(x-y).$$

По теореме 1.3 справедливо неравенство

$$|x| = |y + (x-y)| \leqslant |y| + |x-y|.$$

Отсюда получаем: $|x-y| \geqslant |x| - |y|$ ■

И в заключение отметим, что каковы бы ни были два числа x и y, имеют место легко проверяемые соотношения:

$$|x \cdot y| = |x| \cdot |y| \text{ и } \left|\frac{x}{y}\right| = \frac{|x|}{|y|}, \text{ если } y \neq 0.$$

Глава II. ТЕОРИЯ ПРЕДЕЛОВ

Понятие предела является фундаментальным понятием математического анализа. Начало изучению понятия предела положено в элементарной математике, где с помощью предельных переходов определяется длина окружности, объем цилиндра, конуса и т. д. Оно также использовано при определении суммы бесконечно убывающей геометрической прогрессии. Операция предельного перехода является одной из основных операций анализа. В настоящей главе рассматривается простейшая форма операции предельного перехода, основанная на понятии предела так называемой числовой последовательности. Понятие предела числовой последовательности позволит нам в дальнейшем определить и другие более сложные формы операции предельного перехода.

§ 1. Числовые последовательности

1. Числовые последовательности и арифметические действия над ними. С числовыми последовательностями мы уже встречались в средней школе. Примерами таких последовательностей могут служить: 1) последовательность всех элементов арифметической и геометрической прогрессии, 2) последовательность периметров правильных n-угольников, вписанных в данную окружность, 3) последовательность $x_1 = 1$, $x_2 = 1{,}4$, $x_3 = 1{,}41$ ... приближенных значений $\sqrt{2}$. В этом пункте мы уточним и расширим понятие числовой последовательности.

Определение. Пусть каждому числу n натурального ряда чисел

$$1, 2, 3, \ldots, n, \ldots$$

ставится в соответствие вещественное число x_n, т. е. заданы некоторые вещественные числа, определенным образом перенумерованные. Тогда множество занумерованных вещественных чисел

$$x_1,\ x_2,\ x_3,\ \ldots,\ x_n,\ \ldots \tag{1}$$

называется *числовой последовательностью*, или просто *последовательностью*.

Другими словами, числовую последовательность можно определить как множество пар чисел $(n; x_n)$.

Числа x_n будем называть *элементами*, или *членами*, последовательности (1), а число x_n — *общим*, или *n-м*, *членом* данной последовательности. Сокращенно последовательность (1) будем обозначать символом $\{x_n\}$. Так, например, символ $\left\{\frac{1}{n}\right\}$ будет обозначать последовательность $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$

Последовательность считается заданной, если указан способ получения любого ее элемента. Например, если $x_n = 1 + (-1)^n$, то последовательность запишется в виде: 0, 2, 0, 2, ... Обращая дробь $\frac{1}{3}$ в десятичную с точностью до одного, двух, трех и т. д. знаков после запятой, также получаем последовательность:

$$x_1 = 0{,}3;\ x_2 = 0{,}33;\ x_3 = 0{,}333,\ \ldots;\ x_n = 0{,}333\ldots3,\ \ldots$$

Необходимо заметить, что последовательность отличается от числового множества тем, что мы можем одновременно определить как сам элемент последовательности, так и его порядковый номер.

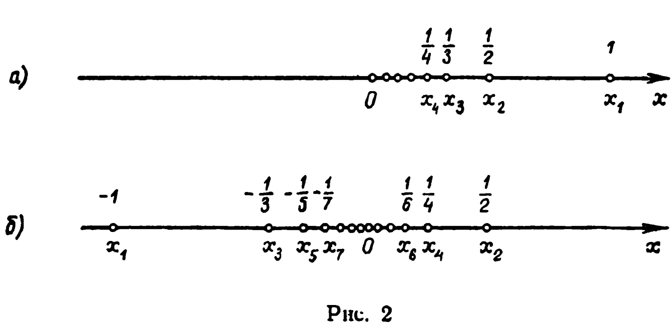


Рис. 2

Геометрически последовательность изображается на числовой прямой в виде последовательности точек, координаты которых равны соответствующим элементам последовательности. На рис. 2 изображены соответственно последовательности

$$\{x_n\}=\left\{\frac{1}{n}\right\} \text{ и } \{x_n\}=\left\{\frac{(-1)^n}{n}\right\}.$$

Введем понятие арифметических действий над числовыми последовательностями. Пусть даны произвольные последовательности $x_1, x_2, \ldots, x_n, \ldots$ и $y_1, y_2, \ldots, y_n, \ldots$ Произведением последовательности $x_1, x_2, \ldots, x_n \ldots$ на число m назовем последовательность

$$mx_1, mx_2, \ldots, mx_n, \ldots$$

Суммой данных последовательностей назовем последовательность

$$x_1+y_1,\ x_2+y_2,\ \ldots,\ x_n+y_n,\ \ldots,$$

разностью — последовательность

$$x_1-y_1,\ x_2-y_2,\ \ldots,\ x_n-y_n,\ \ldots,$$

произведением — последовательность

$$x_1\cdot y_1,\ x_2\cdot y_2,\ \ldots,\ x_n\cdot y_n,\ \ldots,$$

частным — последовательность

$$\frac{x_1}{y_1},\ \frac{x_2}{y_2},\ \ldots,\ \frac{x_n}{y_n},\ \ldots,$$

если все члены последовательности, на которую мы делим, отличны от нуля.

Указанные действия над последовательностями символически записываются так:

$$m\{x_n\}=\{mx_n\},$$

$$\{x_n\}+\{y_n\}=\{x_n+y_n\},$$

$$\{x_n\}-\{y_n\}=\{x_n-y_n\},$$

$$\{x_n\}\cdot\{y_n\}=\{x_n\cdot y_n\},$$

$$\frac{\{x_n\}}{\{y_n\}}=\left\{\frac{x_n}{y_n}\right\},\quad y_n\neq 0^*.$$

---

* $y_n\neq 0$ означает, что значения y_n отличны от нуля при любом n.

2. Ограниченные и неограниченные последовательности

Определение 1. Последовательность $\{x_n\}$ называется ограниченной сверху (снизу), если существует число M (m) такое, что любой элемент x_n этой последовательности удовлетворяет неравенству $x_n \leqslant M$ ($x_n \geqslant m$).

Определение 2. Последовательность $\{x_n\}$ называется ограниченной, если она ограничена и сверху и снизу, т. е. существуют числа m и M такие, что любой элемент x_n этой последовательности удовлетворяет неравенствам: $m \leqslant x_n \leqslant M$.

Обозначим $A = \max\{|m|, |M|\}$. Тогда условие ограниченности последовательности можно записать в виде $|x_n| \leqslant A$.

Определение 3. Последовательность $\{x_n\}$ называется неограниченной, если для любого положительного числа A существует элемент x_n этой последовательности, удовлетворяющий неравенству $|x_n| > A$.

Из данных определений следует, что в случае ограниченности последовательности сверху все ее элементы принадлежат промежутку $(-\infty, M]$, в случае ограниченности снизу — промежутку $[m, +\infty)$, а в случае ограниченности и сверху и снизу — промежутку $[m, M]$. Неограниченная последовательность может быть ограничена сверху (снизу). Рассмотрим несколько примеров:

1) Последовательность $\{x_n\} = \{n\}$ или, что то же, $1, 2, 3, \ldots, n, \ldots$ ограничена снизу, но не ограничена сверху.

2) Последовательность $\{x_n\} = \{-n\}$ или, что то же, $-1, -2, -3, \ldots, -n, \ldots$ ограничена сверху, но не ограничена снизу.

3) Последовательность $\{x_n\} = \left\{\frac{1}{n}\right\}$ или, что то же, $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ ограничена, так как любой элемент x_n этой последовательности удовлетворяет неравенствам: $0 < x_n \leqslant 1$ ($m = 0$, $M = 1$).

4) Последовательность $\{x_n\} = \{(-1)^n \cdot n\}$ или, что то же, $-1, 2, -3, 4, -5, \ldots, (-1)^n n, \ldots$ неограниченная. В самом деле, каково бы ни было число A, среди элементов x_n этой последовательности найдутся элементы, для которых будет выполняться неравенство $|x_n| > A$.

3. Бесконечно большие и бесконечно малые последовательности

Определение 1. Последовательность $\{x_n\}$ называется бесконечно большой, если для любого положительного числа A (сколь бы большим мы его ни взяли) существует номер N такой, что при $n>N$* выполняется неравенство $|x_n|>A$.

Замечание. Очевидно, что любая бесконечно большая последовательность является неограниченной. Однако неограниченная последовательность может и не быть бесконечно большой последовательностью. Например, неограниченная последовательность 1, 2, 1, 3, ..., ... 1, n, 1, $n+1$, ... не является бесконечно большой, поскольку при $A>1$ неравенство $|x_n|>A$ не имеет места для всех элементов x_n с нечетными номерами.

Определение 2. Последовательность $\{\alpha_n\}$ называется бесконечно малой, если для любого положительного числа ε (сколь бы малым мы его ни взяли) существует номер N такой, что при $n>N$ выполняется неравенство $|\alpha_n|<\varepsilon$.

Для более четкого понимания данных определений рассмотрим следующие примеры:

Пример 1. Докажем, что последовательность $\{n\}$ является бесконечно большой, пользуясь только определением 1.

Возьмем любое число $A>0$. Из неравенства $|x_n|=$ $=|n|>A$ получаем $n>A$. Если теперь взять $N\geqslant A$, то для всех $n>N$ будет выполняться неравенство $|x_n|>A$. Так как число A может быть сколь угодно большим, то согласно определению 1 последовательность $\{n\}$ будет бесконечно большой.

Пример 2. Докажем, что последовательность $\left\{\frac{1}{n}\right\}$ является бесконечно малой, пользуясь только определением 2.

Возьмем любое $\varepsilon>0$. Из неравенства $|\alpha_n|=\left|\frac{1}{n}\right|<\varepsilon$ получаем $n>1/\varepsilon$. Если теперь взять $N=\left[\frac{1}{\varepsilon}\right]$**, то для

* «При $n>N$» означает «для всех элементов последовательности с номерами $n>N$».

** Символ $[x]$ обозначает наибольшее целое число, не превосходящее x. Например, $[1]=1$, $[3,1]=3$, $[0,7]=0$, $[-0,6]=-1$, $[-172,9]=-173$ и т. д.

всех $n>N$ будет выполняться неравенство $|\alpha_n|<\varepsilon$. При $\varepsilon=\frac{1}{10}$ получим: $N=[10]=10$, при $\varepsilon=4/15$ получим: $N=\left[\frac{15}{4}\right]=3$ и т. д. Так как число ε может быть сколь угодно малым, то согласно определению 2 последовательность $\left\{\frac{1}{n}\right\}$ есть бесконечно малая.

Теперь докажем теорему, устанавливающую связь между бесконечно большими и бесконечно малыми последовательностями.

Теорема 2.1. *Если $\{x_n\}$ — бесконечно большая последовательность и все ее члены отличны от нуля, $x_n\neq 0$, то $\{\alpha_n\}=\left\{\frac{1}{x_n}\right\}$ есть бесконечно малая, и обратно, если $\{\alpha_n\}$ — бесконечно малая последовательность, $\alpha_n\neq 0$, то $\{x_n\}=\left\{\frac{1}{\alpha_n}\right\}$ есть бесконечно большая.*

Доказательство. Пусть $\{x_n\}$ — бесконечно большая последовательность. Возьмем любое $\varepsilon>0$ и положим $A=1/\varepsilon$. Согласно определению 1 для этого A существует номер N такой, что при $n>N$ будет $|x_n|>A$. Тогда $|\alpha_n|=\left|\frac{1}{x_n}\right|=\frac{1}{|x_n|}<\frac{1}{A}=\varepsilon$, т. е. $|\alpha_n|<\varepsilon$ для всех $n>N$. А это значит, что последовательность $\{1/x_n\}$ есть бесконечно малая.

Доказательство второй части теоремы проводится аналогично ■

4. Основные свойства бесконечно малых последовательностей

Теорема 2.2. *Алгебраическая сумма двух бесконечно малых последовательностей есть бесконечно малая последовательность.*

Доказательство. Пусть $\{\alpha_n\}$ и $\{\beta_n\}$ — бесконечно малые последовательности. Требуется доказать, что последовательность $\{\alpha_n\pm\beta_n\}$ бесконечно малая. Пусть ε — произвольное положительное число, N_1 — номер, начиная с которого $|\alpha_n|<\frac{\varepsilon}{2}$, а N_2 — номер, начиная с которого $|\beta_n|<\frac{\varepsilon}{2}$ (такие номера N_1 и N_2 найдутся по определению бесконечно малой последовательности). Возьмем $N=\max\{N_1, N_2\}$, тогда при $n>N$ будут одновременно выполняться два неравенства: $|\alpha_n|<\frac{\varepsilon}{2}$, $|\beta_n|<\frac{\varepsilon}{2}$. Следовательно, при $n>N$

$$|\alpha_n\pm\beta_n|\leqslant|\alpha_n|+|\beta_n|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

Это значит, что последовательность $\{\alpha_n\pm\beta_n\}$ бесконечно малая ■

Следствие. Алгебраическая сумма любого конечного числа бесконечно малых последовательностей есть бесконечно малая последовательность.

Теорема 2.3. *Произведение двух бесконечно малых последовательностей есть бесконечно малая последовательность.*

Доказательство. Пусть $\{\alpha_n\}$ и $\{\beta_n\}$ — бесконечно малые последовательности. Требуется доказать, что $\{\alpha_n\cdot\beta_n\}$ — бесконечно малая последовательность. Так как $\{\alpha_n\}$ — бесконечно малая, то для любого $\varepsilon>0$ существует номер N_1 такой, что $|\alpha_n|<\varepsilon$ при $n>N_1$, а так как $\{\beta_n\}$ тоже бесконечно малая, то для $\varepsilon=1$ существует номер N_2 такой, что $|\beta_n|<1$ при $n>N_2$. Возьмем $N=\max\{N_1, N_2\}$, тогда при $n>N$ будут выполняться оба неравенства. Следовательно, при $n>N$

$$|\alpha_n\cdot\beta_n|=|\alpha_n|\cdot|\beta_n|<\varepsilon\cdot 1=\varepsilon.$$

Это означает, что $\{\alpha_n\cdot\beta_n\}$ есть бесконечно малая ■

Следствие. Произведение любого конечного числа бесконечно малых последовательностей есть бесконечно малая последовательность.

Замечание. Частное двух бесконечно малых последовательностей может быть любой последовательностью и может не иметь смысла. Например, если $\alpha_n=\frac{1}{n}$, $\beta_n=\frac{1}{n}$, то все элементы $\left\{\frac{\alpha_n}{\beta_n}\right\}$ равны единице и данная последовательность является ограниченной. Если $\alpha_n=\frac{1}{n}$, $\beta_n=\frac{1}{n^2}$, то $\left\{\frac{\alpha_n}{\beta_n}\right\}$ бесконечно большая, и, наоборот, если $\alpha_n=\frac{1}{n^2}$, а $\beta_n=\frac{1}{n}$, то $\left\{\frac{\alpha_n}{\beta_n}\right\}$ бесконечно малая. Если начиная с некоторого номера элементы $\{\beta_n\}$ равны нулю, то $\left\{\frac{\alpha_n}{\beta_n}\right\}$ не имеет смысла.

Теорема 2.4. *Произведение ограниченной последовательности на бесконечно малую есть бесконечно малая последовательность.*

Доказательство. Пусть $\{x_n\}$ — ограниченная, а $\{\alpha_n\}$ — бесконечно малая последовательности. Требуется доказать, что $\{x_n \cdot \alpha_n\}$ — бесконечно малая. Так как $\{x_n\}$ ограничена, то существует число $A>0$ такое, что любой элемент x_n удовлетворяет неравенству $|x_n| \leqslant A$. Возьмем любое $\varepsilon>0$. Поскольку $\{\alpha_n\}$ — бесконечно малая, то для положительного числа $\frac{\varepsilon}{A}$ существует номер N такой, что при $n>N$ выполняется неравенство $|\alpha_n| < \frac{\varepsilon}{A}$. Тогда при $n>N$

$$|x_n \cdot \alpha_n| = |x_n| \cdot |\alpha_n| < A \cdot \frac{\varepsilon}{A} = \varepsilon.$$

Это означает, что последовательность $\{x_n \cdot \alpha_n\}$ бесконечно малая ■

Следствие. Произведение бесконечно малой последовательности на число есть бесконечно малая последовательность.

Переходим теперь к одному из важнейших в математическом анализе понятий — понятию предела числовой последовательности.

§ 2. Сходящиеся последовательности

1. Понятие сходящейся последовательности

Определение. Число a называется *пределом* числовой последовательности $\{x_n\}$, если для любого положительного числа ε существует номер N такой, что для всех членов последовательности с номерами $n>N$ выполняется неравенство

$$|x_n - a| < \varepsilon. \quad (1)$$

При этом последовательность $\{x_n\}$ называется *сходящейся.*

Если последовательность $\{x_n\}$ сходится и имеет своим пределом число a, то символически это записывается так:

$$\lim_{n\to\infty} x_n = a^*, \text{ или } x_n \to a \text{ при } n \to \infty. \quad (2)$$

* limes (*лат.*) — предел.

Рассмотрим

Пример. Показать, что $\left\{\frac{n}{n+1}\right\}$ имеет своим пределом число 1, пользуясь только определением.

Возьмем любое $\varepsilon>0$. Так как $|x_n-1|=\left|\frac{n}{n+1}-1\right|=\frac{1}{n+1}$, то для нахождения значений n, удовлетворяющих неравенству $|x_n-1|<\varepsilon$, достаточно решить неравенство $\frac{1}{n+1}<\varepsilon$, откуда получим: $n>\frac{1-\varepsilon}{\varepsilon}$. Следовательно, за N можно взять целую часть числа $\frac{1-\varepsilon}{\varepsilon}$, т. е. $N=\left[\frac{1-\varepsilon}{\varepsilon}\right]$ Тогда неравенство $|x_n-1|<\varepsilon$ будет выполняться при всех $n>N$. Поскольку ε брали произвольно, то этим и доказано, что 1 есть предел $\left\{\frac{n}{n+1}\right\}$. В частности, если, например, $\varepsilon=0{,}01$, то $N=\left[\frac{1-0{,}01}{0{,}01}\right]=99$ и при $n>N=99$ будет $|x_n-1|<0{,}01$; если $\varepsilon=1/2$, то $N=\left[\frac{1-1/2}{1/2}\right]=1$ и при $n>N=1$ будет $|x_n-1|<\frac{1}{2}$ и т. д. Проверим последний случай. Возьмем $n=2$, тогда

$$|x_2-1|=\left|\frac{2}{2+1}-1\right|=\left|\frac{2-3}{3}\right|=\left|-\frac{1}{3}\right|=\frac{1}{3}<\frac{1}{2}.$$

Пользуясь (2), полученный результат можем записать в виде:

$$\lim_{n\to\infty}\frac{n}{n+1}=1, \text{ или } \frac{n}{n+1}\to 1 \text{ при } n\to\infty.$$

По определению a в данном примере равно 1.

Замечание 1. Пусть $\{x_n\}$ сходится и имеет своим пределом некоторое число a. Тогда разность $\{x_n-a\}=\{\alpha_n\}$ будет бесконечно малой последовательностью, так как для любого $\varepsilon>0$ существует номер N такой, что при $n>N$ выполняется неравенство $|\alpha_n|=|x_n-a|<\varepsilon$. Следовательно, любой элемент x_n сходящейся последовательности, имеющий пределом число a, можно представить в виде

$$x_n=a+\alpha_n, \tag{3}$$

где α_n — элемент бесконечно малой последовательности $\{\alpha_n\}$. Очевидно, справедливо и обратное: если x_n можно представить в виде $x_n = a + \alpha_n$, где $\{\alpha_n\}$ — бесконечно малая последовательность, то $\lim\limits_{n\to\infty} x_n = a$. Представление (3) будет нами часто использоваться при доказательствах многих теорем.

Замечание 2. Предел числовой последовательности имеет вполне четкое геометрическое истолкование. Неравенство (1) равносильно неравенствам

$$-\varepsilon < x_n - a < \varepsilon, \text{ или } a - \varepsilon < x_n < a + \varepsilon,$$

которые означают, что элемент x_n находится в ε-окрестности числа a (рис. 3).

Поэтому определение предела последовательности можно сформулировать и следующим образом: число a является пределом последовательности $\{x_n\}$, если, какую бы окрестность $(a - \varepsilon, a + \varepsilon)$ мы ни задали, существует номер N такой, что все элементы x_n с номерами $n > N$ попадут в заданную окрестность.

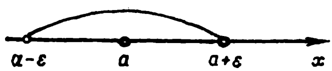


Рис. 3

2. Основные свойства сходящихся последовательностей. Перед тем, как перейти к доказательству следующей теоремы, докажем лемму.

Лемма 2.1. *Если все элементы бесконечно малой последовательности $\{\alpha_n\}$ равны одному и тому же числу c, то $c = 0$.*

Доказательство. Предположим обратное, что $c \neq 0$. Положим $\varepsilon = \frac{|c|}{2}$. Тогда по определению бесконечно малой последовательности существует номер N такой, что при $n > N$ выполняется неравенство $|\alpha_n| < \varepsilon$. Так как $\alpha_n = c$, а $\varepsilon = \frac{|c|}{2}$, то последнее неравенство можно переписать в виде: $|c| < \frac{|c|}{2}$, откуда $1 < \frac{1}{2}$. Полученное противоречие показывает, что предположение $c \neq 0$ не может иметь места ■

Теорема 2.5. *Сходящаяся последовательность имеет только один предел.*

Доказательство. Предположим обратное, что сходящаяся последовательность $\{x_n\}$ имеет два предела

a и b. Тогда по формуле (3) для элементов x_n получим

$$x_n=a+\alpha_n \quad \text{и} \quad x_n=b+\beta_n,$$

где α_n и β_n — элементы бесконечно малых последовательностей $\{\alpha_n\}$ и $\{\beta_n\}$. Вычитая написанные соотношения одно из другого, найдем: $\alpha_n-\beta_n=b-a$. Так как все элементы бесконечно малой последовательности $\{\alpha_n-\beta_n\}$ имеют одно и то же постоянное значение $b-a$, то по доказанной лемме $b-a=0$, т. е. $b=a$ ■

Теорема 2.6. *Сходящаяся последовательность ограничена.*

Доказательство. Пусть $\{x_n\}$ — сходящаяся последовательность и число a — ее предел. Пусть, далее, ε — произвольное положительное число и N — номер, начиная с которого выполняется $|x_n-a|<\varepsilon$. Тогда

$$|x_n|=|(x_n-a)+a|\leqslant|x_n-a|+|a|<|a|+\varepsilon$$

для всех $n>N$. Возьмем $A=\max\{|a|+\varepsilon, |x_1|, |x_2|, \ldots, \ldots, |x_N|\}$. Очевидно, $|x_n|\leqslant A$ для всех номеров n, что и означает ограниченность последовательности $\{x_n\}$ ■

Замечание. Ограниченная последовательность может и не быть сходящейся.

Так, например, последовательность $-1, 1, -1, \ldots, (-1)^n, \ldots$ ограничена, но не сходящаяся. Будем вести рассуждения от противного. Предположим, что данная последовательность имеет пределом число a. Это значит, что для любого $\varepsilon>0$, в частности и для $\varepsilon=1/2$, существует номер N такой, что при $n>N$ $|x_n-a|<\frac{1}{2}$. Так как x_n принимает попеременно значения 1 и -1, то можем записать $|1-a|<\frac{1}{2}$ и $|(-1)-a|<\frac{1}{2}$. Тогда получим

$$2=|1-a+a-(-1)|\leqslant$$
$$\leqslant|1-a|+|a-(-1)|<\frac{1}{2}+\frac{1}{2}=1,$$

т. е. $2<1$, чего быть не может.

Докажем следующие основные теоремы.

Теорема 2.7. *Алгебраическая сумма двух сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$ есть сходящаяся* *последовательность, предел которой равен сумме пределов последовательностей $\{x_n\}$ и $\{y_n\}$.*

Доказательство. Пусть a и b — соответственно пределы $\{x_n\}$ и $\{y_n\}$. Тогда по формуле (3)

$$x_n = a + \alpha_n,\quad y_n = b + \beta_n,$$

где $\{\alpha_n\}$ и $\{\beta_n\}$ — бесконечно малые последовательности. Следовательно,

$$(x_n \pm y_n) - (a \pm b) = \alpha_n \pm \beta_n.$$

По теореме 2.2 последовательность $\{\alpha_n \pm \beta_n\}$ бесконечно малая. Таким образом, последовательность $\{(x_n \pm y_n) - (a \pm b)\}$ также бесконечно малая и поэтому последовательность $\{x_n \pm y_n\}$ сходится и имеет своим пределом число $a \pm b$ ■

Теорема 2.8. *Произведение сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$ есть сходящаяся последовательность, предел которой равен произведению пределов последовательностей $\{x_n\}$ и $\{y_n\}$.*

Доказательство. Пусть a и b — соответственно пределы $\{x_n\}$ и $\{y_n\}$. Тогда по формуле (3)

$$x_n = a + \alpha_n,\quad y_n = b + \beta_n,$$

где $\{\alpha_n\}$ и $\{\beta_n\}$ — бесконечно малые последовательности. Следовательно,

$$x_n y_n - ab = a\beta_n + b\alpha_n + \alpha_n\beta_n.$$

Согласно теоремам 2.2 — 2.3 последовательность $\{a\beta_n + b\alpha_n + \alpha_n\beta_n\}$ бесконечно малая. Таким образом, последовательность $\{x_n y_n - ab\}$ также бесконечно малая и поэтому последовательность $\{x_n y_n\}$ сходится и имеет своим пределом число ab ■

Теорема 2.9. *Частное двух сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$ при условии, что предел $\{y_n\}$ отличен от нуля*, есть сходящаяся последовательность, предел которой равен частному пределов последовательностей $\{x_n\}$ и $\{y_n\}$.*

---

* В силу условия $\lim\limits_{n\to\infty} y_n \neq 0$ элементы y_n начиная с некоторого номера N не обращаются в нуль, поэтому частное $\left\{\frac{x_n}{y_n}\right\}$ имеет смысл для всех $n > N$.

Доказательство. Если a и b — пределы последовательностей $\{x_n\}$ и $\{y_n\}$ соответственно, то для доказательства теоремы достаточно установить, что $\left\{\frac{x_n}{y_n} - \frac{a}{b}\right\}$ есть бесконечно малая последовательность. Представим $\frac{x_n}{y_n} - \frac{a}{b}$, используя формулу (3), в виде

$$\frac{x_n}{y_n} - \frac{a}{b} = \frac{bx_n - ay_n}{by_n} = \frac{b(a+\alpha_n) - a(b+\beta_n)}{by_n} = \frac{1}{y_n}\left(\alpha_n - \frac{a}{b}\beta_n\right).$$

В силу свойств бесконечно малых последовательностей множитель $\alpha_n - \frac{a}{b}\beta_n$ есть бесконечно малая последовательность. Покажем, что $\left\{\frac{1}{y_n}\right\}$ есть ограниченная последовательность. Так как $y_n \to b$, $b \neq 0$, при $n \to \infty$, то для $\varepsilon = |b|/2$ найдется номер N такой, что для всех $n > N$ будет $|y_n - b| < \frac{|b|}{2}$. Тогда

$$|y_n| = |b - (b - y_n)| \geqslant$$
$$\geqslant |b| - |y_n - b| > |b| - \frac{|b|}{2} = \frac{|b|}{2},$$

откуда $|y_n| > \frac{|b|}{2}$, и, следовательно, $\left|\frac{1}{y_n}\right| < \frac{2}{|b|}$ для всех $n > N$, что и означает ограниченность последовательности $\left\{\frac{1}{y_n}\right\}$

По теореме 2.4 последовательность $\left\{\frac{1}{y_n}\left(\alpha_n - \frac{a}{b}\beta_n\right)\right\}$ бесконечно малая, поэтому последовательность $\left\{\frac{x_n}{y_n} - \frac{a}{b}\right\}$ также бесконечно малая. Следовательно, последовательность $\left\{\frac{x_n}{y_n}\right\}$ сходится и имеет своим пределом число a/b ■

Теоремы, доказанные в этом пункте, имеют очень большое не только теоретическое, но и практическое значение.

3. Предельный переход в неравенствах.

Теорема 2.10. *Если элементы сходящейся последовательности $\{x_n\}$, начиная с некоторого номера, удовлетворяют неравенству $x_n \geqslant b$ ($x_n \leqslant b$), то и предел a этой последовательности удовлетворяет неравенству $a \geqslant b$ ($a \leqslant b$).*

Доказательство. Пусть все элементы x_n, начиная с некоторого номера, удовлетворяют неравенству $x_n \geqslant b$. Требуется доказать неравенство $a \geqslant b$. Предположим обратное, что $a < b$.

Поскольку a — предел $\{x_n\}$, то для $\varepsilon = b - a$ существует номер N такой, что при $n > N$ выполняется неравенство $|x_n - a| < b - a$, которое равносильно следующим двум неравенствам: $-(b-a) < x_n - a < b - a$. Из правого неравенства мы получаем $x_n < b$, а это противоречит условию теоремы. Случай $x_n \leqslant b$ рассматривается аналогично ■

Следствие 1. Если элементы x_n и y_n сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$, начиная с некоторого номера, удовлетворяют неравенству $x_n \leqslant y_n$, то их пределы удовлетворяют неравенству $\lim\limits_{n\to\infty} x_n \leqslant \lim\limits_{n\to\infty} y_n$.

В самом деле, начиная с некоторого номера элементы $\{y_n - x_n\}$ неотрицательны, а поэтому неотрицателен и ее предел

$$\lim_{n\to\infty}(y_n - x_n) = \lim_{n\to\infty} y_n - \lim_{n\to\infty} x_n \geqslant 0.$$

Отсюда следует, что

$$\lim_{n\to\infty} x_n \leqslant \lim_{n\to\infty} y_n.$$

Следствие 2. Если все элементы сходящейся последовательности $\{x_n\}$ находятся на отрезке $[a, b]$, то и ее предел c также находится на этом отрезке.

В самом деле, так как $a \leqslant x_n \leqslant b$, то $a \leqslant c \leqslant b$.

Следующая теорема играет важную роль в различных приложениях.

Теорема 2.11. *Пусть даны три последовательности $\{x_n\}$, $\{y_n\}$ и $\{z_n\}$, связанные неравенствами $x_n \leqslant y_n \leqslant z_n$ для всех n. Тогда, если $\{x_n\}$ и $\{z_n\}$ имеют один и тот же предел a, то $\{y_n\}$ также имеет предел a.*

Доказательство. Возьмем любое $\varepsilon>0$. По этому ε для $\{x_n\}$ найдется номер N_1 такой, что $|x_n-a|<\varepsilon$ для всех $n>N_1$, т. е.

$$\underline{a-\varepsilon<x_n}<a+\varepsilon. \qquad (4)$$

По этому же ε для $\{z_n\}$ найдется номер N_2 такой, что $|z_n-a|<\varepsilon$ для всех $n>N_2$, т. е.

$$a-\varepsilon<\underline{z_n<a+\varepsilon}. \qquad (5)$$

Возьмем $N=\max\{N_1, N_2\}$. Тогда для $n>N$ будут выполняться одновременно неравенства (4) и (5). Используя подчеркнутые их части, а также неравенства, данные в условии теоремы, можем записать:

$$a-\varepsilon<x_n\leqslant y_n\leqslant z_n<a+\varepsilon \text{ для } n>N.$$

Отсюда

$$a-\varepsilon<y_n<a+\varepsilon, \text{ или } |y_n-a|<\varepsilon \text{ для } n>N.$$

Последнее означает, что пределом $\{y_n\}$ является a ■

§ 3. Монотонные последовательности

1. Определение и признак сходимости монотонных последовательностей

Определение. Последовательность $\{x_n\}$ называется *возрастающей*, если $x_1<x_2<x_3<\ldots<x_n<x_{n+1}<\ldots$; *неубывающей*, если $x_1\leqslant x_2\leqslant x_3\leqslant\ldots\leqslant x_n\leqslant x_{n+1}\leqslant\ldots$; *убывающей*, если $x_1>x_2>x_3>\ldots>x_n>x_{n+1}>\ldots$; *невозрастающей*, если $x_1\geqslant x_2\geqslant x_3\geqslant\ldots\geqslant x_n\geqslant x_{n+1}\geqslant\ldots$

Все такие последовательности объединяются под общим названием *монотонных*. Возрастающие и убывающие последовательности называются также *строго монотонными*.

Рассмотрим примеры монотонных последовательностей.

1) Последовательность $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ убывающая и ограниченная.

2) Последовательность $1, 1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \ldots, \frac{1}{n}, \frac{1}{n}, \ldots$ невозрастающая и ограниченная.

3) Последовательность 1, 2, 3, ..., n, ... возрастающая и неограниченная.
4) Последовательность 1, 1, 2, 2, 3, 3, ..., n, n, ... неубывающая и неограниченная.
5) Последовательность $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \frac{n}{n+1}, \ldots$ возрастающая и ограниченная.

Отметим, что монотонные последовательности ограничены по крайней мере с одной стороны: неубывающие последовательности — снизу ($x_n \geqslant x_1$ для всех n), невозрастающие — сверху ($x_n \leqslant x_1$ для всех n). Оказывается, что если монотонная последовательность ограничена с обеих сторон, т. е. просто ограничена, то она сходится. Немонотонные последовательности этим свойством не обладают. Например, немонотонная последовательность $\{(-1)^n\}$ ограничена, но не сходится (см. замечание к теореме 2.6).

Имеет место следующая основная теорема о монотонных последовательностях.

Теорема 2.12. *Монотонная ограниченная последовательность имеет предел.*

Доказательство. Рассмотрим случай монотонно неубывающей последовательности.

Пусть $x_1 \leqslant x_2 \leqslant x_3 \leqslant \ldots \leqslant x_n \leqslant x_{n+1} \leqslant \ldots$ и существует число M такое, что все элементы x_n не больше, чем M, т. е. $x_n \leqslant M$. Рассмотрим числовое множество X, состоящее из элементов данной последовательности. По условию это множество ограничено сверху и непусто. Поэтому в силу теоремы 1.1 множество X имеет точную верхнюю грань. Обозначим ее через a и покажем, что a будет пределом для данной последовательности.

Поскольку a — точная верхняя грань множества элементов последовательности $\{x_n\}$, то в силу ее свойства для любого $\varepsilon > 0$ найдется номер N такой, что $x_N > a - \varepsilon$. Так как $\{x_n\}$ — неубывающая последовательность, то при $n > N$ будет $x_n > a - \varepsilon$. С другой стороны, по определению верхней грани $x_n \leqslant a < a + \varepsilon$ для всех n. Таким образом, для $n > N$ получаем неравенства $a - \varepsilon < x_n < a + \varepsilon$, из которых вытекает неравенство $|x_n - a| < \varepsilon$. Последнее и означает, что число a — предел последовательности $\{x_n\}$.

Случай монотонно невозрастающей последовательности разбирается аналогично ■

З а м е ч а н и е. Условие ограниченности монотонной последовательности представляет собой необходимое и достаточное условие ее сходимости.

В самом деле, если монотонная последовательность ограничена, то в силу теоремы она сходится; если же монотонная последовательность сходится, то в силу теоремы 2.6 она ограничена.

2. Число e. Рассмотрим последовательность $\{x_n\}$ с общим членом $x_n = \left(1+\frac{1}{n}\right)^n$:

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

и докажем, что она сходится. Для этого достаточно показать, что последовательность $\{x_n\}$ — возрастающая и ограниченная сверху. Применив формулу бинома Ньютона (см. гл. VI, § 3, п. 4), получим

$$x_n = 1 + n\frac{1}{n} + \frac{n(n-1)}{2!}\cdot\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\cdot\frac{1}{n^3} + \ldots$$
$$\ldots + \frac{n(n-1)(n-2)\ldots[n-(n-1)]}{n!}\cdot\frac{1}{n^n}.$$

Представим это выражение в следующей форме:

$$x_n = 2 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots$$
$$\ldots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right). \quad (1)$$

Аналогичным образом запишем $(n+1)$-й элемент:

$$x_{n+1} = 2 + \frac{1}{2!}\left(1-\frac{1}{n+1}\right) +$$
$$+ \frac{1}{3!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right) + \ldots$$
$$\ldots + \frac{1}{(n+1)!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\ldots\left(1-\frac{n}{n+1}\right).$$

Прежде всего заметим, что $\left(1-\frac{k}{n}\right) < \left(1-\frac{k}{n+1}\right)$ при $0<k<n$, т. е. каждое слагаемое в выражении x_{n+1} больше соответствующего слагаемого в выражении x_n,

и, кроме того, у x_{n+1} по сравнению с x_n добавляется еще одно положительное слагаемое. Поэтому

$$x_n < x_{n+1},$$

т. е. последовательность $\{x_n\}$ *возрастающая.*

Для доказательства ограниченности данной последовательности сверху заметим, что каждое выражение в круглых скобках в соотношении (1) меньше единицы. Учитывая также, что $\frac{1}{n!} < \frac{1}{2^{n-1}}$ при $n > 2$, получим

$$x_n < 2 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} <$$

$$< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}}.$$

Применим формулу для суммы геометрической прогрессии в последнем выражении, тогда

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

Таким образом, мы доказали, что последовательность $\{(1+1/n)^n\}$ монотонно возрастающая и ограничена сверху. По теореме 2.12 она имеет предел. Этот предел называют числом e. Следовательно, по определению

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

Отметим, что число e имеет большое значение во многих вопросах теории и практики. В настоящем параграфе мы дали только определение числа e. Дальше мы покажем способ вычисления этого числа с любой степенью точности.

Здесь мы лишь отметим, что поскольку $x_n < 3$ и из (1) непосредственно очевидно, что $2 < x_n$, то число e заключено в пределах

$$2 \leqslant e \leqslant 3.$$

Изучение теории пределов мы закончим доказательством теоремы, которую далее будем неоднократно использовать при доказательстве других важных теорем.

§ 4. Теорема о вложенных отрезках

Пусть дана последовательность отрезков $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$, ... таких, что каждый последующий содержится в предыдущем: $[a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \supset \ldots$, т. е.

$$a_n \leqslant a_{n+1} < b_{n+1} \leqslant b_n \text{ для всех } n, \qquad (1)$$

и пусть $\lim\limits_{n\to\infty}(b_n - a_n) = 0$. Будем называть ее *последовательностью вложенных отрезков*. Имеет место следующая

Теорема 2.13. *Для любой последовательности вложенных отрезков существует единственная точка c, принадлежащая всем отрезкам этой последовательности, т. е. такая, что $a_n \leqslant c \leqslant b_n$.*

Доказательство. Из неравенств (1) следует, что левые концы отрезков образуют монотонно неубывающую последовательность:

$$a_1 \leqslant a_2 \leqslant a_3 \leqslant \ldots \leqslant a_n \leqslant a_{n+1} \leqslant \ldots, \qquad (2)$$

а правые концы — монотонно невозрастающую последовательность:

$$b_1 \geqslant b_2 \geqslant b_3 \geqslant \ldots \geqslant b_n \geqslant b_{n+1} \geqslant \ldots \qquad (3)$$

При этом последовательность (2) ограничена сверху, а (3) ограничена снизу, так как $a_n \leqslant b_1$, а $b_n \geqslant a_1$ для любого n. Следовательно, на основании теоремы 2.12 эти последовательности имеют пределы. Пусть $\lim\limits_{n\to\infty} a_n = c'$, $\lim\limits_{n\to\infty} b_n = c''$. Тогда из условия

$$\lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = c'' - c' = 0$$

следует, что $c' = c''$, т. е. последовательности $\{a_n\}$ и $\{b_n\}$ имеют общий предел. Обозначая этот предел через c, получаем, что для любого n справедливы неравенства $a_n \leqslant c \leqslant b_n$, т. е. точка c принадлежит всем отрезкам последовательности (1).

Покажем теперь, что точка c является единственной. Допустим, что существует еще точка c_1 ($c_1 \neq c$), принадлежащая всем отрезкам последовательности (1).

Тогда для любого n должно выполняться неравенство $b_n - a_n \geqslant |c_1 - c|$ и, следовательно, $\lim\limits_{n\to\infty}(b_n - a_n) \geqslant |c_1 - c| \neq 0$, что противоречит условию теоремы ■

Замечание. Теорема будет неверна, если вместо отрезков рассматривать интервалы. Например, для последовательности вложенных интервалов

$$(0, 1) \supset \left(0, \frac{1}{2}\right) \supset \left(0, \frac{1}{4}\right) \supset \ldots \supset \left(0, \frac{1}{2^n}\right) \supset \ldots \qquad (4)$$

не существует точки, принадлежащей всем интервалам. В самом деле, какую бы точку c на интервале $(0, 1)$ мы ни взяли, всегда найдется номер N такой, что при $n>N$ будет $\frac{1}{2^n}<c$ и точка c не будет принадлежать интервалам последовательности (4), начиная с $\left(0, \frac{1}{2^N}\right)$. Точка 0 также не принадлежит им, так как является общим левым концом всех интервалов.

Для дальнейшего изложения курса нам понадобятся некоторые сведения из аналитической геометрии. Поэтому следующая глава посвящена этому разделу математики.

Глава III. АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ НА ПЛОСКОСТИ

Аналитическая геометрия есть область математики, изучающая геометрические образы алгебраическими методами. Еще в XVII в. французским математиком Декартом был разработан и впервые применен так называемый «метод координат», давший возможность переводить геометрические понятия на алгебраический язык.

В основе «метода координат» лежат некоторые построения — система координат. Таких систем существует довольно много. Мы познакомимся с прямоугольной декартовой и полярной системами координат.

§ 1. Прямоугольная система координат

Две взаимно перпендикулярные числовые оси Ox и Oy, имеющие общее начало O и одинаковую масштабную единицу (рис. 4), образуют *прямоугольную систему координат* на плоскости.

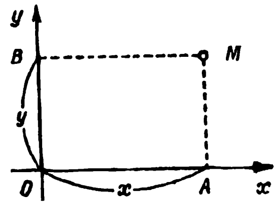


Рис. 4

Ось Ox называется *осью абсцисс*, ось Oy — *осью ординат*, а обе оси вместе —*осями координат*. Точка O пересечения осей называется *началом координат*. Плоскость, в которой расположены оси Ox и Oy, называется *координатной плоскостью* и обозначается Oxy.

Пусть M — произвольная точка плоскости. Опустим из нее перпендикуляры MA и MB на оси Ox и Oy, которые на осях отсекут отрезки OA и OB*. При этом предполагается, что при выбранной единице измерения каждому отрезку сопоставляется вещественное число, называемое его *величиной*.

Прямоугольными координатами x и y точки M будем называть соответственно величины отрезков OA и OB: $x=OA$, $y=OB$, взятых с соответствующими знаками: если точка A лежит на оси Ox правее (левее) точки O, то величине отрезка OA приписывается знак $+$ (знак $-$); если точка B лежит на оси Oy выше (ниже) точки O, то величине отрезка OB приписывается знак $+$ (знак $-$).

Тот факт, что точка M имеет координаты x и y, символически обозначают так: $M(x; y)$. Начало координат имеет координаты $(0; 0)$. Координаты x и y точки M называются соответственно ее *абсциссой* и *ординатой*.

Таким образом, при выбранной системе координат каждой точке M плоскости соответствует единственная пара чисел $(x; y)$ ** — ее прямоугольные координаты и,

---

* Здесь и в дальнейшем символами OA, OB будем обозначать величины отрезков, а символами $|OA|$, $|OB|$ — их длины.

** Напомним, что здесь речь идет об *упорядоченной паре чисел*, т. е. о наборе из двух чисел, в котором указано, какое число является первым, а какое — вторым. Если $x \neq y$, то пары $(x; y)$ и $(y; x)$ различны, так как в первой из них первым числом является x, а во второй — y.

обратно, каждой паре чисел (x; y) соответствует, и притом одна, точка M на плоскости Oxy.

Итак, прямоугольная система координат на плоскости устанавливает взаимно-однозначное соответствие между множеством всех точек плоскости и множеством пар чисел, которое дает возможность при решении геометрических задач применять алгебраические методы.

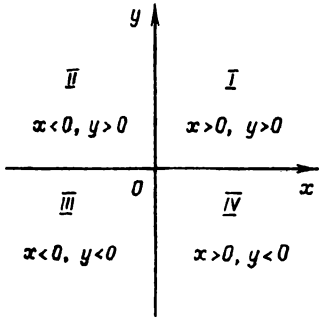


Рис. 5

Координатные оси разбивают плоскость на четыре части, их называют *четвертями, квадрантами* или *координатными углами* и нумеруют по определенному правилу (I, II, III, IV) (рис. 5). На рис. 5 указана расстановка знаков координат точек в зависимости от их расположения в той или иной четверти.

§ 2. Простейшие задачи аналитической геометрии на плоскости

Рассмотрим некоторые простейшие задачи на применение прямоугольных координат на плоскости.

1. Расстояние между двумя точками

Теорема 3.1. *Для любых точек $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$ плоскости расстояние d между ними определяется формулой*

$$d=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}. \qquad (1)$$

Доказательство. Опустим из точек M_1 и M_2 перпендикуляры на оси Oy и Ox и обозначим точки их пересечения соответственно через B и A. Кроме того, через K обозначим точку пересечения прямых M_1B и M_2A (рис. 6). Очевидно, $M_1K=M_1B+BK$, $M_1B=-x_1$, $BK=x_2$, $M_1K=x_2-x_1$; $M_2K=M_2A+AK$, $M_2A=y_2$, $AK=$ $=-y_1$; $M_2K=y_2-y_1$. Так как треугольник M_1M_2K — прямоугольный, то по теореме Пифагора

$$d=\sqrt{(M_1K)^2+(M_2K)^2}=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}\blacksquare$$

2. Площадь треугольника

Теорема 3.2. *Каковы бы ни были три точки $A(x_1; y_1)$, $B(x_2; y_2)$ и $C(x_3; y_3)$, не лежащие на одной*

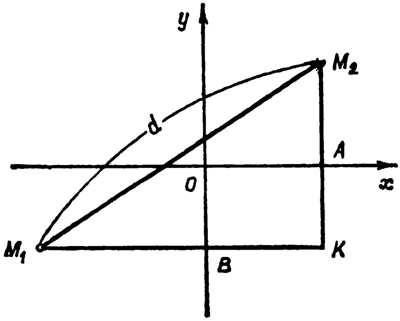


Рис. 6

прямой, площадь S треугольника ABC определяется формулой

$$S=\frac{1}{2}\left|\left[(x_2-x_1)(y_3-y_1)-(x_3-x_1)(y_2-y_1)\right]\right|. \qquad (2)$$

Доказательство. Обозначим через φ угол между отрезками AB и AC, через d и d' — длины этих от-

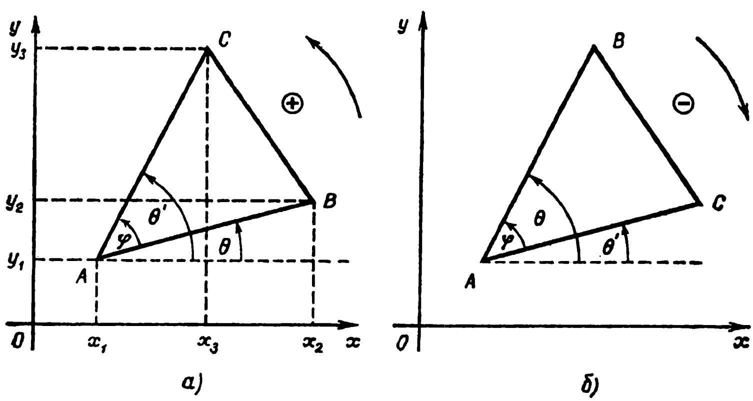


Рис. 7

резков, а через θ и θ' соответственно — углы между отрезками AB, AC с осью Ox (рис. 7). Как известно из элементарной геометрии, площадь треугольника равна половине произведения двух его сторон на синус угла между ними, т. е.

$$S = \frac{1}{2} dd' \sin \varphi. \qquad (3)$$

Если кратчайший поворот отрезка AB к отрезку AC на угол φ положителен (т. е. совершается против часовой стрелки), то $\varphi=\theta'-\theta$ (см. рис. 7,*а*), если же кратчайший поворот AB к AC отрицателен, то $\varphi=\theta-\theta'$ (см. рис. 7,*б*), т. е. $\varphi=\pm(\theta'-\theta)$. Подставляя φ в формулу (3), получаем

$$S = \pm \frac{1}{2} dd' \sin(\theta' - \theta) =$$

$$= \pm \frac{1}{2} dd' (\sin\theta' \cos\theta - \cos\theta' \sin\theta) =$$

$$= \pm \frac{1}{2} (d\cos\theta\, d' \sin\theta' - d' \cos\theta'\, d\sin\theta). \qquad (4)$$

Если из точек A, B и C опустим перпендикуляры на оси координат, то будут иметь место следующие выражения:

$$d\cos\theta = x_2 - x_1, \quad d\sin\theta = y_2 - y_1,$$

$$d'\cos\theta' = x_3 - x_1, \quad d'\sin\theta' = y_3 - y_1.$$

Подставляя эти выражения в формулу (4) и учитывая, что $S>0$, получим

$$S = \frac{1}{2} \left| [(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)] \right| \blacksquare$$

Пример. Даны точки $A(1; 1)$, $B(6; 4)$, $C(8; 2)$. Найти площадь треугольника ABC. По формуле (2)

$$S = \frac{1}{2} \left| [(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)] \right| =$$

$$= \frac{1}{2} \left| [-16] \right| = 8.$$

Следовательно $S=8$.

3. Деление отрезка в данном отношении. Пусть на плоскости даны две произвольные точки M_1 и M_2. Проведем через данные точки прямую u и назначим на ней положительное направление, тем самым мы сделаем ее осью. Пусть, далее, M — любая точка оси u, отличная от точки M_2 (рис. 8).

Число λ, определяемое равенством

$$\lambda = \frac{M_1M}{MM_2} \quad (M \neq M_2),$$

называется отношением, в котором точка M делит отрезок M_1M_2.

Если при этом точка M находится между точками M_1 и M_2, то числа M_1M, MM_2 положительны и λ есть число положительное. Если же точка M находится вне отрезка $[M_1M_2]$, то в этом случае одно из чисел M_1M, MM_2 положительно, другое отрицательно, а так как M_1M и MM_2 — величины разных знаков, то λ в данном случае есть число отрицательное.

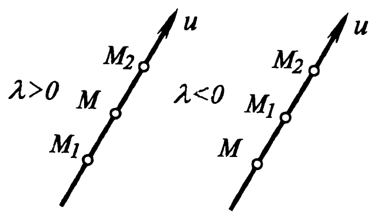


Рис. 8

Решение задачи о делении отрезка в данном отношении дает следующая теорема.

Теорема 3.3. *Если точка* $M(x; y)$ *делит отрезок* M_1M_2 *в отношении* $\lambda \neq -1$, *то координаты этой точки выражаются формулами*

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}; \quad y = \frac{y_1 + \lambda y_2}{1 + \lambda}, \tag{5}$$

где $(x_1; y_1)$ — *координаты точки* M_1, $(x_2; y_2)$ — *координаты точки* M_2.

Доказательство. Опустим перпендикуляры из точек M_1, M, M_2 на ось Ox и обозначим точки пересечения соответственно через P_1, P и P_2 (рис. 9). На основании известной теоремы элементарной геометрии о пропорциональности отрезков прямых, заключенных между параллельными прямыми, имеем

$$\frac{P_1P}{PP_2} = \frac{M_1M}{MM_2} = \lambda,$$

но

$$P_1P = x - x_1, \quad PP_2 = x_2 - x.$$

Отсюда $\frac{x - x_1}{x_2 - x} = \lambda$, т. е. $x = \frac{x_1 + \lambda x_2}{1 + \lambda}$. Получили первую из формул (5). Вторая получается совершенно аналогично ■

Следствие. Если точки $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$ — две произвольные точки и точка $M(x; y)$ — середина отрезка M_1M_2, т. е. $M_1M = MM_2$, то

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}.$$

Иными словами, каждая координата середины отрезка равна полусумме соответствующих координат.

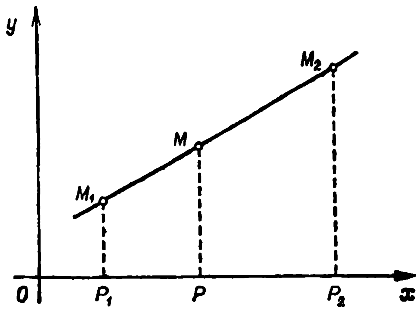


Рис. 9

Пример 1. Даны точки $M_1(1; 1)$ и $M_2(7; 4)$. Найти точку $M(x, y)$, которая в два раза ближе к M_1, чем к M_2, и находится между M_1 и M_2.

Решение. Искомая точка делит отрезок M_1M_2 в отношении $\lambda = 1/2$. Применяя формулы (5), находим координаты этой точки: $x = 3$, $y = 2$.

Пример 2. Даны точки $M_1(1; 1)$ и $M_2(7; 4)$. Найти точку $M(x; y)$, которая в два раза ближе к M_1, чем к M_2, и находится вне отрезка, ограниченного точками M_1 и M_2.

Решение. Искомая точка делит отрезок M_1M_2 в отношении $\lambda = -1/2$. Применяя формулы (5), находим координаты этой точки: $x = -5$, $y = -2$.

§ 3. Полярные координаты

Наиболее важной после прямоугольной системы координат является полярная система координат. Эта система состоит из некоторой точки O, называемой *полю-*

сом, и проходящего через нее луча Ox, называемого *полярной осью*. Кроме того, задана единица масштаба для измерения длин отрезков.

Пусть заданы полюс, полярная ось и M — произвольная точка плоскости. Обозначим через ρ расстояние точки M от точки O, через φ — угол, на который нужно повернуть против часовой стрелки полярную ось для совмещения с лучом OM (рис. 10).

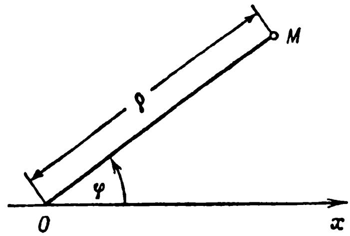


Рис. 10

Полярными координатами точки M (относительно заданной системы) называются числа ρ и φ. При этом число ρ называется *первой координатой*, или *полярным радиусом*, число φ — *второй координатой*, или *полярным углом*.

Точка M с полярными координатами ρ и φ обозначается так: $M(\rho; \varphi)$.

Что касается значений, принимаемых полярными координатами, то обычно считают, что ρ и φ изменяются в следующих границах:

$$0 \leqslant \rho < +\infty, \quad 0 \leqslant \varphi < 2\pi.$$

Однако в ряде случаев приходится рассматривать углы, большие 2π, а также отрицательные углы, т. е. углы, отсчитываемые от полярной оси по часовой стрелке.

Установим связь между полярными координатами точки и ее прямоугольными координатами. При этом будем предполагать, что начало прямоугольной системы координат находится в полюсе, а ось абсцисс совпадает с полярной осью. Пусть точка M имеет прямоугольные координаты x и y и полярные координаты ρ и φ (рис. 11). Очевидно,

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi. \qquad (1)$$

Формулы (1) выражают прямоугольные координаты через полярные. Выражения полярных координат через прямоугольные следуют из формул (1):

$$\rho = \sqrt{x^2 + y^2}, \quad \operatorname{tg} \varphi = \frac{y}{x}. \qquad (2)$$

Необходимо заметить, что формула $\operatorname{tg}\varphi=\frac{y}{x}$ определяет два значения полярного угла φ, так как φ изменяется от 0 до 2π. Из этих двух значений угла φ выбирается то, при котором удовлетворяются равенства (1).

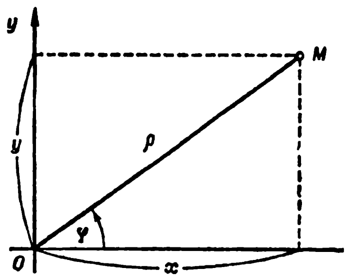


Рис. 11

Пример. Даны прямоугольные координаты точки (2; 2). Найти ее полярные координаты, считая, что полюс совмещен с началом прямоугольной системы, а полярная ось совпадает с положительной полуосью абсцисс.

Решение. По формулам (2) имеем:

$$\rho=2\sqrt{2},\quad \operatorname{tg}\varphi=1.$$

Согласно второму из этих равенств $\varphi=\frac{\pi}{4}$. Итак,

$$\rho=2\sqrt{2},\quad \varphi=\frac{\pi}{4}.$$

§ 4. Преобразование прямоугольных координат

При решении многих задач в аналитической геометрии наряду с данной прямоугольной системой координат приходится вводить и другие прямоугольные системы координат. При этом, естественно, изменяются как координаты точек, так и уравнения кривых. Возникает задача: зная координаты точки в одной системе координат, найти координаты этой же точки в другой системе координат. Для этой цели служат *формулы преобразования координат*.

Мы рассмотрим два вида преобразований прямоугольных координат:

1) *параллельный сдвиг осей*, когда меняется положение начала координат, а направления осей остаются прежними;
2) *поворот осей координат*, когда обе оси поворачиваются в одну сторону на один и тот же угол, а начало координат не перемещается.

1. Параллельный сдвиг осей. Пусть произвольная точка M плоскости имеет относительно прямоугольной системы координат Oxy координаты $(x; y)$. Перенесем

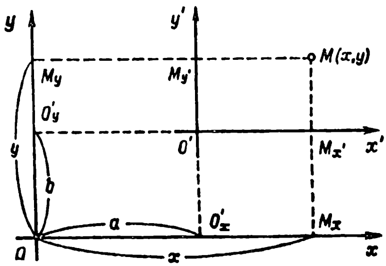


Рис. 12

начало координат в точку $O'(a; b)$, где a и b — координаты нового начала относительно старой системы координат Oxy. Новые оси $O'x'$ и $O'y'$ направим соответственно параллельно старым осям Ox и Oy. Обозначим новые координаты точки M в системе $O'x'y'$ через (x', y'). Выведем формулы, выражающие связь между новыми и старыми координатами точки M. Для этого проведем $MM_x \perp Ox$, $O'O'_x \perp Ox$ и $MM_y \perp Oy$, $O'O'_y \perp Oy$ и введем соответствующие обозначения (рис. 12). Тогда

$$x = OM_x = OO'_x + O'_x M_x = OO'_x + O'M_{x'} = a + x',$$

$$y = OM_y = OO'_y + O'_y M_y = OO'_y + O'M_{y'} = b + y'.$$

Итак,

$$x = x' + a, \quad y = y' + b,$$

или

$$x' = x - a, \quad y' = y - b. \tag{1}$$

Это и есть искомые формулы.

2. Поворот осей координат. Повернем систему координат Oxy вокруг начала координат O на угол α в положение $Ox'y'$. Тогда произвольная точка $M(x; y)$ плоскости в системе координат $Ox'y'$ будет иметь координаты $(x'; y')$, при этом $x=OM_x$, $y=OM_y$, $x'=OM_{x'}$, $y'=OM_{y'}$ (рис. 13). Выведем формулы, устанавливающие

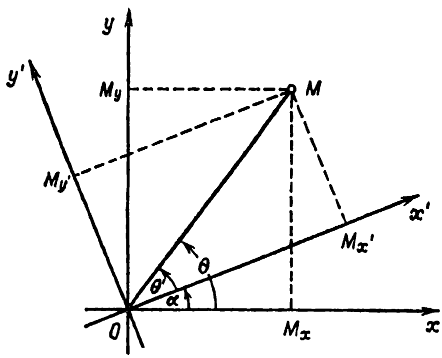


Рис. 13

связь между старыми и новыми координатами точки M. Для этого обозначим через $(\rho; \theta)$ полярные координаты точки M, считая полярной осью Ox, а через $(\rho; \theta')$ — полярные координаты той же точки M, считая полярной осью Ox', причем в каждом случае $\rho=|OM|$. Очевидно, что $\theta=\theta'+\alpha$. Далее, согласно формулам (1) из § 3

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

и аналогично

$$x'=\rho\cos\theta',\quad y'=\rho\sin\theta'.$$

Таким образом,

$$x=\rho\cos\theta=\rho\cos(\theta'+\alpha)=\rho(\cos\theta'\cos\alpha-\sin\theta'\sin\alpha)=$$
$$=\rho\cos\theta'\cos\alpha-\rho\sin\theta'\sin\alpha=x'\cos\alpha-y'\sin\alpha;$$
$$y=\rho\sin\theta=\rho\sin(\theta'+\alpha)=\rho(\cos\theta'\sin\alpha+\sin\theta'\cos\alpha)=$$
$$=\rho\cos\theta'\sin\alpha+\rho\sin\theta'\cos\alpha=x'\sin\alpha+y'\cos\alpha.$$

Итак,

$$\begin{aligned} x&=x'\cos\alpha-y'\sin\alpha,\\ y&=x'\sin\alpha+y'\cos\alpha, \end{aligned}\tag{2}$$

или

$$x'=x\cos\alpha+y\sin\alpha,$$
$$y'=-x\sin\alpha+y\sin\alpha.$$

Это и есть искомые формулы.

Пример. Определить координаты точки $M(3; 5)$ относительно новой координатной системы $O'x'y'$, начало которой находится в точке $(-2; 1)$ и оси которой параллельны осям координатной системы Oxy.

Решение. По формуле (1) имеем

$$x'=3+2=5,\ y'=5-1=4,$$

т. е. координаты точки M будут $(5; 4)$.

§ 5. Линии и их уравнения

Пусть на плоскости заданы: прямоугольная система координат и некоторая линия L (рис. 14). Рассмотрим соотношение вида

$$F(x, y)=0, \qquad (1)$$

связывающее переменные величины x и y. Равенство вида (1) будем называть уравнением с двумя переменными x, y, если это равенство справедливо не для всех пар чисел x и y.

Примеры уравнений: $2x+3y=0$, $x^2+y^2-25=0$, $\sin x+\sin y-1=0$ и т. д.

Если (1) справедливо для всех пар чисел x и y, то оно называется тождеством.

Примеры тождеств: $(x+y)^2-x^2-2xy-y^2=0$, $(x+y)(x-y)-x^2+y^2=0$ и т. д.

Важнейшим понятием аналитической геометрии является понятие уравнения линии.

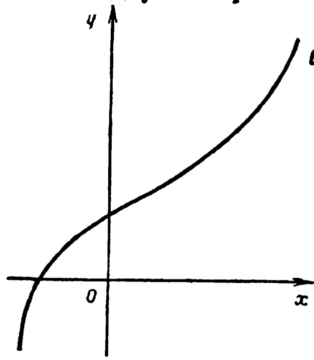


Рис. 14

Определение. Уравнение (1) называется *уравнением линии* L (относительно заданной системы координат), если этому уравнению удовлетворяют координаты x и y любой точки, лежащей на линии L, и не удовлетворяют кординаты никакой точки, не лежащей на линии L.

Из определения следует, что сама линия L представляет собой множество всех тех точек плоскости, координаты которых удовлетворяют этому уравнению.

Поскольку величины x и y рассматриваются как координаты переменной точки M, их называют текущими координатами.

Если (1) является уравнением линии L, то мы будем говорить, что (1) определяет линию L.

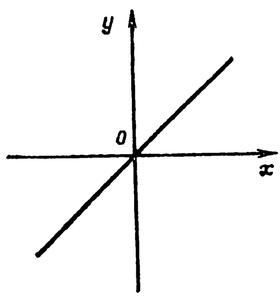


Рис. 15

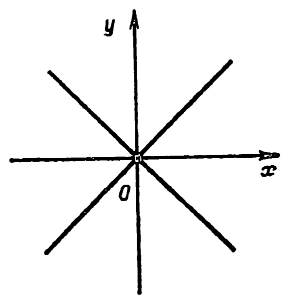


Рис. 16

Понятие уравнения линии дает возможность сводить геометрические задачи к алгебраическим. Например, задача нахождения точки пересечения двух линий, определяемых уравнениями $x+y=0$ и $x^2+y^2=1$, сводится к алгебраической задаче совместного решения этих уравнений.

Рассмотрим несколько простейших примеров определения линий уравнениями.

1) $x-y=0$. Представив уравнение в виде $y=x$, заключаем, что множество точек, координаты которых удовлетворяют данному уравнению, есть биссектриса I и III координатных углов. Это и есть линия, определенная уравнением $x-y=0$ (рис. 15).

2) $x^2-y^2=0$. Представив уравнение в виде $(x-y)\times$ $\times(x+y)=0$, заключаем, что множество точек, координаты которых удовлетворяют данному уравнению, есть две прямые (рис. 16).

3) $x^2+y^2=0$. Множество точек, координаты которых удовлетворяют данному уравнению, состоит из одной

точки $(0, 0)$. В данном случае уравнение определяет, как говорят, вырожденную линию.

4) $x^2+y^2+1=0$. Так как при любых x и y числа x^2 и y^2 неотрицательны, то $x^2+y^2+1>0$. Значит, нет ни одной точки, координаты которой удовлетворяют данному уравнению, т. е. никакого геометрического образа на плоскости данное уравнение не определяет.

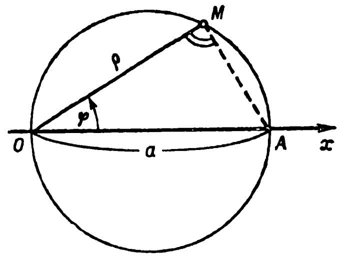


Рис. 17

5) $\rho=a\cos\varphi$, где a — положительное число, переменные ρ и φ — полярные координаты. Обозначим через M точку с полярными координатами $(\rho; \varphi)$, через A — точку с полярными координатами $(a;\ 0)$. Если $\rho=a\cos\varphi$, то угол OMA — прямой, и обратно. Следовательно, множество точек, полярные координаты которых удовлетворяют данному уравнению, есть окружность с диаметром OA (рис. 17).

6) $\rho=a\varphi$, где a — положительное число, ρ и φ — полярные координаты. Обозначим через M точку с полярными координатами $(\rho; \varphi)$. Если $\varphi=0$, то и $\rho=0$, если φ возрастает, начиная от нуля, то ρ будет возрастать пропорционально φ. Точка $M(\rho; \varphi)$, таким образом, исходя из полюса, движется вокруг него с ростом φ (в положительном направлении), одновременно удаляясь от него. Множество точек, полярные координаты которых удовлетворяют уравнению $\rho=a\varphi$, называется *спиралью Архимеда* (рис. 18).

Если точка M совершает один полный оборот вокруг полюса, то φ возрастает на 2π, а ρ — на $2a\pi$, т. е. спираль рассекает любую прямую, проходящую через полюс, на равные отрезки (не считая отрезков, примыкающих к полюсу), которые имеют постоянную длину $2a\pi$.

Теперь рассмотрим обратную задачу: по заданному множеству точек, т. е. заданной линии L, найти ее уравнение $F(x, y)=0$.

Пример. Вывести уравнение множества точек, каждая из которых отстоит от точки $C(\alpha;\ \beta)$ на расстояние R.

Решение. Обозначим буквой M переменную точку, буквами x, y — ее текущие координаты, тогда $CM=R$. По формуле (1) § 2 имеем

$$R=\sqrt{(x-\alpha)^2+(y-\beta)^2}.$$

Возводя обе части равенства в квадрат, мы получаем уравнение окружности с центром в точке $C(\alpha, \beta)$ и радиусом R (рис. 19):

$$(x-\alpha)^2+(y-\beta)^2=R^2.$$

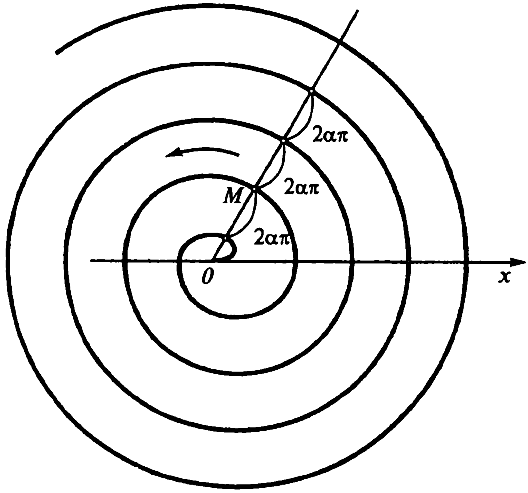


Рис. 18

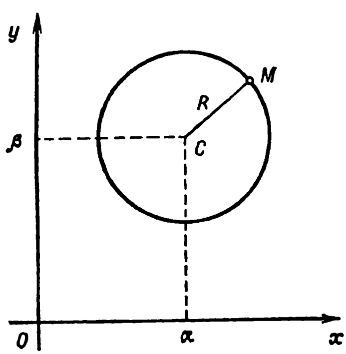


Рис. 19

Оно встречается во многих геометрических задачах. Полагая в нем $\alpha=0$, $\beta=0$, получим уравнение окружности с центром в начале координат:

$$x^2+y^2=R^2.$$

§ 6. Линии первого порядка

1. Уравнение прямой с угловым коэффициентом. Пусть дана некоторая прямая, не перпендикулярная оси Ox. Назовем углом наклона данной прямой к оси Ox угол α, на который нужно повернуть ось Ox, чтобы ее направление совпало с одним из направлений прямой.

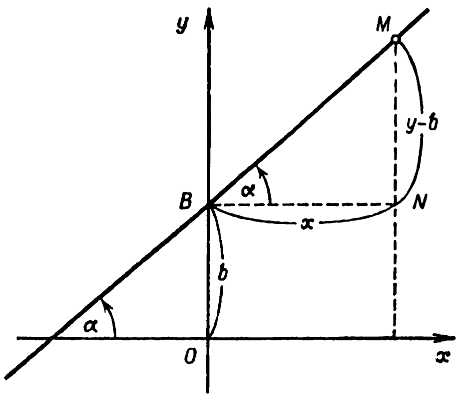


Рис. 20

Угол α может иметь множество различных значений, которые отличаются друг от друга на величину $\pm n\pi$, где n — натуральное число. Чаще всего в качестве угла наклона берут наименьшее положительное значение угла α, на который нужно повернуть против часовой стрелки ось Ox, чтобы ее направление совпало с одним из направлений прямой (рис. 20). В таком случае $0\leqslant\alpha<\pi$.

Тангенс угла наклона прямой к оси Ox называется *угловым коэффициентом* этой прямой и обозначается через k:

$$k=\operatorname{tg}\alpha. \qquad (1)$$

Из формулы (1), в частности, следует, что если $\alpha=0$, то $k=0$, т. е. прямая параллельна оси Ox. Если $\alpha=\pi/2$, то $k=\operatorname{tg}\alpha$ теряет смысл. В таком случае говорят, что угловой коэффициент «обращается в бесконечность».

Выведем уравнение данной прямой, полагая известными ее угловой коэффициент k и величину b отрезка OB, который она отсекает на оси Oy (см. рис. 20).

Обозначим через M переменную точку с текущими координатами x и y, кроме того, введем в рассмотрение точку $B(0; b)$, в которой прямая пересекает ось Oy. Если провести прямые BN и MN, параллельные осям, то образуется прямоугольный треугольник BNM. Ясно, что

$$\frac{MN}{BN}=\operatorname{tg}\alpha,\ BN=x,\ MN=y-b.$$

Отсюда и учитывая формулу (1), получаем

$$\frac{y-b}{x}=k. \tag{2}$$

Так как точка $M(x; y)$ лежит на данной прямой, то выражение (2) является уравнением данной прямой, которое после преобразования примет вид

$$y=kx+b. \tag{3}$$

Уравнение (3) называют *уравнением прямой с угловым коэффициентом*.

Итак, любая прямая, не перпендикулярная к оси Ox, определяется уравнением вида (3). Верно и обратное, любое уравнение вида (3) определяет прямую, которая имеет угловой коэффициент k и отсекает на оси Oy отрезок величины b. Рассмотрим

Пример. Построить прямую по уравнению

$$y=\frac{3}{4}x+2.$$

Решение. Отложим на оси Oy отрезок $OB=2$ (рис. 21); проведем через точку B параллельно оси Ox отрезок $BN=4$ и через точку N параллельно оси Oy отрезок $NM=3$. После этого, соединяя точки B и M, получим искомую прямую, которая имеет данный угловой коэффициент $k=\frac{3}{4}$ и отсекает на оси отрезок $b=2$.

2. Уравнение прямой, проходящей через данную точку, с данным угловым коэффициентом. В ряде случаев встречается необходимость составить уравнение прямой, зная одну ее точку $M_1(x_1; y_1)$ и угловой коэффициент k. Так как искомая прямая проходит через точку $M_1(x_1; y_1)$ (рис. 22), то координаты этой точки должны удовлетворять уравнению (3): $y_1 = kx_1 + b$.

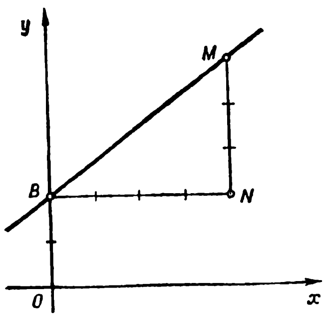


Рис. 21

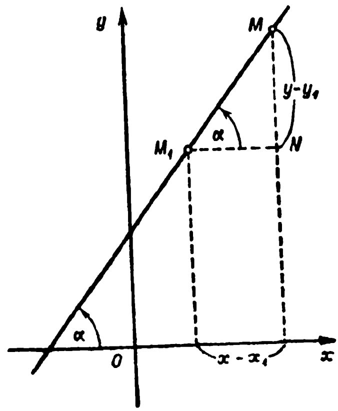


Рис. 22

Определяя b из последнего равенства и подставляя его в уравнение (3), получим искомое уравнение прямой:

$$y - y_1 = k(x - x_1). \qquad (4)$$

3. Уравнение прямой, проходящей через две данные точки. Пусть даны две точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$.

Приняв в (4) точку $M(x, y)$ за $M_2(x_2, y_2)$, получим

$$y_2 - y_1 = k(x_2 - x_1).$$

Определяя k из последнего равенства и подставляя его в уравнение (4), получим искомое уравнение прямой:

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1).$$

Это уравнение, если $y_1 \neq y_2$, можно записать в виде

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \qquad (5)$$

Если $y_1=y_2$, то уравнение искомой прямой $y=y_1$.

Замечание. Если $x_1=x_2$, то прямая, проходящая через точки M_1 и M_2, параллельна оси Oy, ее уравнение имеет вид $x=x_1$.

Пример. Составить уравнение прямой, проходящей через точки $M_1(3;\ 1)$ и $M_2(5;\ 4)$.

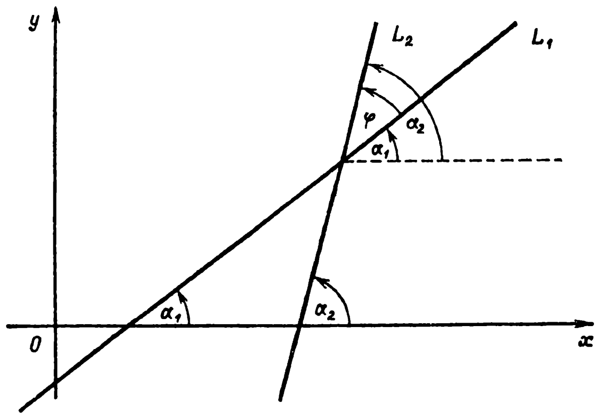


Рис. 23

Решение. Подставляя данные координаты в соотношение (5), получим

$$\frac{x-3}{2}=\frac{y-1}{3}, \text{ или } 3x-2y-7=0.$$

4. Угол между двумя прямыми. Возьмем две прямые L_1 и L_2: уравнение L_1 пусть будет $y=k_1x+b_1$, где $k_1=\operatorname{tg}\alpha_1$, а уравнение L_2 — $y=k_2x+b_2$, где $k_2=\operatorname{tg}\alpha_2$ (рис. 23). Пусть φ — угол между прямыми L_1 и L_2: $0\leqslant\varphi<\pi$.

Из геометрических соображений устанавливаем зависимость между углами α_1, α_2, φ: $\alpha_2=\alpha_1+\varphi$, или $\varphi=\alpha_2-\alpha_1$, отсюда

$$\operatorname{tg}\varphi=\operatorname{tg}(\alpha_2-\alpha_1)=\frac{\operatorname{tg}\alpha_2-\operatorname{tg}\alpha_1}{1+\operatorname{tg}\alpha_1\operatorname{tg}\alpha_2},$$

или

$$\operatorname{tg}\varphi=\frac{k_2-k_1}{1+k_2k_1}. \qquad (6)$$

Формула (6) определяет один из углов между прямыми. Второй угол равен $\pi-\varphi$.

Пример. Даны прямые $y=2x+3$, $y=-3x+2$. Найти угол между ними.

Решение. Пусть $k_1=2$, $k_2=-3$, тогда по формуле (6) находим:

$$\operatorname{tg}\varphi=\frac{-3-2}{1+(-3)\cdot 2}=\frac{-5}{-5}=1.$$

Таким образом, один из углов, которые составляют данные прямые, равен 45°.

5. Условия параллельности и перпендикулярности двух прямых. Если прямые L_1 и L_2 параллельны, то $\varphi=0$ и $\operatorname{tg}\varphi=0$. В этом случае числитель в правой части формулы (6) равен нулю:

$$k_2-k_1=0,$$

т. е.

$$k_2=k_1.$$

Отсюда заключаем, что условием параллельности двух прямых является равенство их угловых коэффициентов.

Если прямые L_1 и L_2 перпендикулярны, т. е. $\varphi=\pi/2$, то

$$\alpha_2=\frac{\pi}{2}+\alpha_1,\ \operatorname{tg}\alpha_2=\operatorname{tg}\left(\frac{\pi}{2}+\alpha_1\right)=-\operatorname{ctg}\alpha_1=-\frac{1}{\operatorname{tg}\alpha_1},$$

т. е.

$$k_2=-1/k_1.$$

Отсюда заключаем, что условием перпендикулярности двух прямых является то, что их угловые коэффициенты обратны по величине и противоположны по знаку.

6. Общее уравнение прямой

Теорема 3.4. *В прямоугольных координатах каждая прямая определяется уравнением первой степени*

$$Ax+By+C=0 \tag{7}$$

и обратно, уравнение (7) при произвольных коэффициентах A, B, C (A и B не равны нулю одновременно) определяет некоторую прямую.

Доказательство. Сначала докажем первое утверждение. Если прямая не перпендикулярна Ox, то, как было показано в п. 1, она определяется уравнением первой степени: $y=kx+b$ (см. (3)), т. е. уравнением вида (7), где $A=k$, $B=-1$ и $C=b$. Если прямая перпендикулярна Ox, то все ее точки имеют одинаковые абсциссы, равные величине отрезка, отсекаемого прямой на оси Ox (рис. 24), т. е. $x=a$, что также является уравнением первой степени вида (7), где $A=1$, $B=0$, $C=-a$. Тем самым первое утверждение доказано.

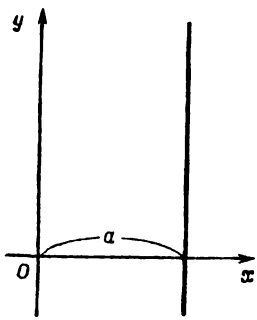


Рис. 24

Докажем обратное утверждение. Пусть дано уравнение (7).

Если $B\neq 0$, то (7) можно записать в виде $y=-\dfrac{A}{B}x-\dfrac{C}{B}$. Полагая $k=-\dfrac{A}{B}$, $b=-\dfrac{C}{B}$, получим уравнение $y=kx+b$, т. е. уравнение (3), которое определяет прямую.

Если $B=0$, то $A\neq 0$ и (7) примет вид $x=-C/A$. Обозначая $-C/A$ через a, получим $x=a$, т. е. уравнение прямой, перпендикулярной к оси Ox ■

Линии, определяемые уравнением первой степени, называются *линиями первого порядка*. Таким образом, каждая прямая есть линия первого порядка и, обратно, каждая линия первого порядка есть прямая.

Уравнение вида $Ax+By+C=0$ называется *общим уравнением прямой* (или *полным уравнением прямой*). При различных численных значениях A, B, C оно определяет всевозможные прямые.

7. Неполное уравнение первой степени. Уравнение прямой «в отрезках». Рассмотрим три частных случая, когда уравнение $Ax+By+C=0$ является неполным:

1) $C=0$; уравнение имеет вид $Ax+By=0$ и определяет прямую, проходящую через начало координат.

2) $B=0$ $(A\neq 0)$; уравнение имеет вид $Ax+C=0$ и определяет прямую, параллельную оси Oy. Как было показано в теореме 3.4, это уравнение приводится к виду $x=a$, где $a=-\dfrac{C}{A}$, a есть величина отрезка, который отсекает прямая на оси Ox (см. рис. 24). В частности, если $a=0$, то прямая совпадает с осью Oy. Таким

образом, уравнение

$$x=0$$

определяет ось ординат.

3) $A=0$ $(B\neq 0)$; уравнение имеет вид $By+C=0$ и определяет прямую, параллельную оси Ox. Это устанавливается аналогично предыдущему случаю. Если положить $-\frac{C}{B}=b$, то уравнение примет вид $y=b$, где b — величина отрезка, который отсекает прямая на оси Oy (рис. 25). В частности, если $b=0$, то прямая совпадает с осью Ox. Таким образом, уравнение

$$y=0$$

определяет ось абсцисс.

Пусть теперь дано уравнение $Ax+By+C=0$ при условии, что ни один из коэффициентов A, B, C не равен нулю. Преобразуем его к виду

$$\frac{x}{-\frac{C}{A}}+\frac{y}{-\frac{C}{B}}=1.$$

Вводя обозначения $a=-\frac{C}{A}$, $b=-\frac{C}{B}$, получим

$$\frac{x}{a}+\frac{y}{b}=1. \tag{8}$$

Уравнение (8) называется *уравнением прямой «в отрезках»*. Эта форма уравнения удобна для геометрического построения прямой.

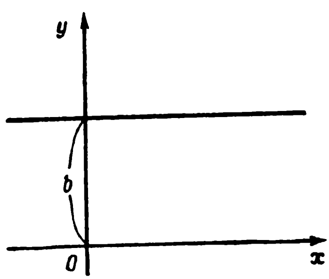


Рис. 25

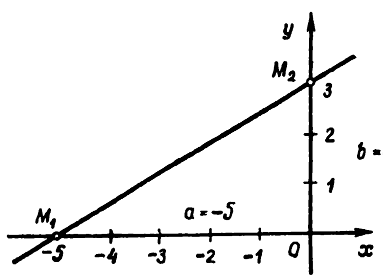


Рис. 26

Пример. Дана прямая $3x-5y+15=0$. Составить для этой прямой уравнение «в отрезках» и построить прямую.

Решение. Для данной прямой уравнение «в отрезках» имеет вид

$$\frac{x}{-5}+\frac{y}{3}=1.$$

Мы можем построить эту прямую, если отложим на координатных осях Oy и Ox отрезки, величины которых соответственно равны $a=-5$, $b=3$, и далее проведем прямую через точки $M_1(-5; 0)$ и $M_2(0; 3)$ (рис. 26).

8. **Нормальное уравнение прямой. Расстояние от точки до прямой.** Пусть дана некоторая прямая L. Проведем через начало координат прямую n, перпендикулярную к данной, которую назовем *нормалью*; буквой N пометим точку, в которой она пересекает прямую L (рис. 27). На нормали введем положительное направление от точки O к точке N. Таким образом, нормаль станет осью.

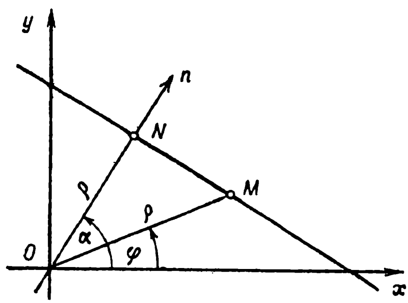


Рис. 27

Обозначим через α угол, на который нужно повернуть против часовой стрелки ось Ox до совмещения ее положительного направления с направлением нормали, через p — длину отрезка ON. Тем самым $0\leqslant\alpha<2\pi$, $p\geqslant 0$. Выведем уравнение данной прямой, считая известными числа α и p. С этой целью возьмем на прямой произвольную точку M с полярными координатами $(\rho; \varphi)$, где O — полюс, Ox — полярная ось. Из прямоугольного треугольника ONM имеем

$$p=\rho\cos(\alpha-\varphi)=\rho(\cos\alpha\cos\varphi+\sin\alpha\sin\varphi).$$

Это выражение можно переписать в виде

$$\rho \cos \varphi \cos \alpha + \rho \sin \varphi \sin \alpha - p = 0. \qquad (9)$$

Так как точки, не лежащие на данной прямой L, не удовлетворяют уравнению (9), то (9) есть уравнение прямой в полярных координатах. По формулам, связывающим прямоугольные координаты с полярными, имеем $\rho \cos \varphi = x$ и $\rho \sin \varphi = y$. Следовательно, уравнение (9) в прямоугольной системе координат примет вид

$$x \cos \alpha + y \sin \alpha - p = 0. \qquad (10)$$

Уравнение (10) называется *нормальным уравнением* прямой L.

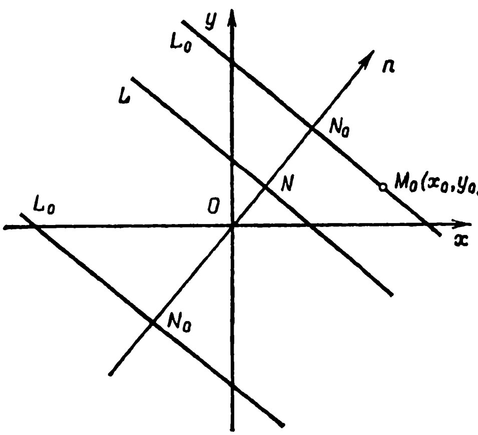


Рис. 28

С помощью нормального уравнения прямой можно определить расстояние от данной точки до прямой на плоскости.

Пусть L — некоторая прямая, заданная в нормальном виде: $x \cos \alpha + y \sin \alpha - p = 0$, и пусть $M_0(x_0; y_0)$ — точка, лежащая вне этой прямой. Требуется определить расстояние d от точки M_0 до прямой L.

Через точку M_0 проведем прямую L_0 параллельно прямой L. Пусть N_0 — точка пересечения L_0 с нормалью, p_0 — длина ON_0 (рис. 28).

Если точка N_0 лежит по ту же сторону от точки O, что и точка N, то нормальное уравнение прямой L_0

имеет вид $x\cos\alpha+y\sin\alpha-p_0=0$. Так как точка $M_0(x_0;\ y_0)\in L_0$, то $x_0\cos\alpha+y_0\sin\alpha-p_0=0$, откуда $p_0=x_0\cos\alpha+y_0\sin\alpha$. В этом случае

$$d=|p_0-p|=|x_0\cos\alpha+y_0\sin\alpha-p|.$$

Если же точка N_0 лежит по другую сторону от точки O, то нормальное уравнение прямой L_0 имеет вид $x\cos(\pi+\alpha)+y\sin(\pi+\alpha)-p_0=0$. В этом случае $p_0=x_0\cos(\pi+\alpha)+y_0\sin(\pi+\alpha)=-x_0\cos\alpha-y_0\sin\alpha$ и $d=|p_0+p|=|-x_0\cos\alpha-y_0\sin\alpha+p|=|x_0\cos\alpha+y_0\sin\alpha-p|$.

Таким образом, для любого взаимного расположения точки M_0 и прямой L получаем формулу

$$d=|x_0\cos\alpha+y_0\sin\alpha-p|. \qquad (11)$$

Из формулы (11) следует, что для вычисления расстояния d от точки M_0 до прямой L нужно в левую часть нормального уравнения прямой L подставить вместо текущих координат координаты точки M_0 и полученное число взять по модулю.

Теперь покажем, как привести общее уравнение прямой к нормальному виду. Пусть

$$Ax+By+C=0 \qquad (12)$$

— общее уравнение некоторой прямой, а

$$x\cos\alpha+y\sin\alpha-p=0 \qquad (13)$$

— ее нормальное уравнение.

Так как уравнения (12) и (13) определяют одну и ту же прямую, то их коэффициенты пропорциональны. Это означает, что, умножив все члены уравнения (12) на некоторое число μ, мы получим уравнение

$$\mu Ax+\mu By+\mu C=0,$$

которое при надлежащем подборе μ должно обратиться в уравнение (13), т. е. мы будем иметь:

$$\mu A=\cos\alpha,\ \mu B=\sin\alpha,\ \mu C=-p. \qquad (14)$$

Чтобы найти множитель μ, возведем первые два из этих равенств в квадрат и сложим:

$$\mu^2(A^2+B^2)=\cos^2\alpha+\sin^2\alpha=1.$$

Отсюда

$$\mu = \pm \frac{1}{\sqrt{A^2 + B^2}}. \tag{15}$$

Число μ называется ***нормирующим множителем***. Знак нормирующего множителя определяется с помощью третьего из равенств (14). Согласно этому равенству μC есть число отрицательное. Следовательно, в формуле (15) знак берется противоположным знаку C. Если $C=0$, то знак нормирующего множителя можно выбрать по желанию.

Итак, для приведения общего уравнения прямой к нормальному виду надо найти значение нормирующего множителя μ, а затем все члены уравнения умножить на μ.

Пример. Даны прямая $3x-4y+10=0$ и точка $M(4; 3)$. Найти расстояние d от точки M до данной прямой.

Решение. Приведем данное уравнение к нормальному виду. С этой целью находим нормирующий множитель по формуле (15):

$$\mu = -\frac{1}{\sqrt{3^2+4^2}} = -\frac{1}{5}.$$

Умножая данное уравнение на μ, получим искомое нормальное уравнение

$$-\frac{1}{5}(3x - 4y + 10) = 0.$$

По формуле (11) подставляем в левую часть этого уравнения координаты точки M и находим расстояние

$$d = \left| -\frac{1}{5}(3 \cdot 4 - 4 \cdot 3 + 10) \right| = |-2| = 2.$$

§ 7. Линии второго порядка

Здесь мы рассмотрим три вида линий: эллипс, гиперболу и параболу, уравнения которых в прямоугольной системе координат являются уравнениями второй степени. Такие линии называются *линиями второго порядка.*

1. Эллипс

Определение. Эллипсом называется множество всех точек плоскости, для которых сумма расстояний от двух данных точек, называемых фокусами, есть величина постоянная, большая чем расстояние между фокусами.

Для вывода уравнения эллипса введем на плоскости прямоугольную систему координат так, чтобы фокусы эллипса лежали на оси абсцисс, а начало координат делило бы расстояние между фокусами пополам. Выведем уравнение эллипса в выбранной системе координат.

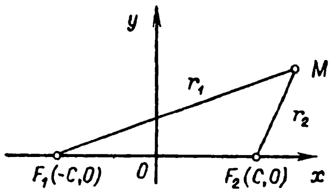


Рис. 29

Обозначим фокусы эллипса через F_1 и F_2. Пусть M — произвольная точка эллипса. Расстояние $|F_1F_2|$ между фокусами обозначим через $2c$, сумму расстояний от точки M до фокусов — через $2a$. Так как по определению $|F_1M|+|F_2M|>|F_1F_2|$, то $2a>2c$, т. е. $a>c$.

Обозначим, далее, через r_1 и r_2 расстояния от точки M до фокусов ($r_1=F_1M$, $r_2=F_2M$) (рис. 29). Числа r_1 и r_2 называются *фокальными радиусами* точки M. Из определения следует, что точка $M(x; y)$ будет находиться на данном эллипсе в том и только в том случае, когда

$$r_1+r_2=2a. \tag{1}$$

Чтобы получить искомое уравнение эллипса, нужно в равенстве (1) заменить переменные r_1 и r_2 их выражениями через координаты x, y. Так как F_1 и F_2 расположены на оси Ox симметрично относительно начала координат, то они имеют соответственно координаты $(-c, 0)$ и $(c, 0)$; приняв это во внимание и применяя формулу (1) § 2, находим:

$$r_1=\sqrt{(x+c)^2+y^2}, \quad r_2=\sqrt{(x-c)^2+y^2}. \tag{2}$$

Подставляя значения r_1 и r_2 в уравнение (1), получим

$$\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2}=2a. \tag{3}$$

Уравнение (3) является *уравнением эллипса*, которому будут удовлетворять координаты каждой точки эллипса. Однако для пользования оно неудобно, поэтому уравнение эллипса приводится обычно к более простому виду. Для этого перенесем второй корень уравнения (3) в правую часть уравнения, после чего возведем обе части равенства в квадрат:

$$(x+c)^2+y^2=4a^2-4a\sqrt{(x-c)^2+y^2}+(x-c)^2+y^2,$$

или

$$a\sqrt{(x-c)^2+y^2}=a^2-cx. \tag{4}$$

Снова возведем обе части равенства в квадрат:

$$a^2x^2-2a^2cx+a^2c^2+a^2y^2=a^4-2a^2cx+c^2x^2,$$

откуда

$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2). \tag{5}$$

Введем в рассмотрение новую величину

$$b=\sqrt{a^2-c^2}, \tag{6}$$

геометрический смысл которой будет раскрыт несколько позднее. Так как по условию $a>c$, то $a^2-c^2>0$ и, следовательно, b — число положительное. Из равенства (6) имеем

$$b^2=a^2-c^2,$$

вследствие чего уравнение (5) можно переписать в виде

$$b^2x^2+a^2y^2=a^2b^2.$$

Разделив обе части на a^2b^2, получим окончательно

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{7}$$

Так как уравнение (7) получено из уравнения (3), то координаты любой точки эллипса, удовлетворяющие уравнению (3), будут удовлетворять и уравнению (7). Однако при упрощении уравнения (3) обе его части дважды возводились в квадрат и могли появиться «лишние» корни, вследствие чего уравнение (7) может быть неравносильным уравнению (3). Поэтому надо убедиться в том, что любая точка, координаты которой удовлетворяют уравнению (7), удовлетворяет и уравнению (3), т. е. точка принадлежит эллипсу. Для этого, очевидно, достаточно показать, что величины r_1 и r_2 для любой точки, координаты которой удовлетворяют уравнению (7), удовлетворяют соотношению (1). Действительно, пусть координаты x и y некоторой точки удовлетворяют уравнению (7). Тогда, подставляя в выражения (2) для r_1 значение $y^2 = b^2\left(1 - \frac{x^2}{a^2}\right)$, полученное из (7), после несложных преобразований, найдем, что $r_1 = \sqrt{\left(a + \frac{c}{a}x\right)^2}$. Так как $|x| \leqslant a$ (следует из (7)) и $\frac{c}{a} < 1$, то $a + \frac{c}{a}x > 0$ и $r_1 = a + \frac{c}{a}x$.

Аналогично найдем, что $r_2 = a - \frac{c}{a}x$. Складывая почленно эти равенства, получаем соотношение (1), что и требовалось. Таким образом, любая точка, координаты которой удовлетворяют уравнению (7), принадлежит эллипсу, и наоборот, т. е. уравнение (7) есть уравнение эллипса. Уравнение (7) называется *каноническим* (или *простейшим*) уравнением эллипса. Таким образом, эллипс — линия второго порядка.

Займемся теперь исследованием формы эллипса по его каноническому уравнению (7).

Заметим, что уравнение (7) содержит члены только с четными степенями текущих координат x и y, поэтому эллипс симметричен как относительно оси Ox, так и оси Oy. В силу сказанного, мы будем знать форму всего эллипса, если установим вид той его части, которая лежит в первом координатном угле. Для этого разрешим уравнение (7) относительно y:

$$y = \frac{b}{a}\sqrt{a^2 - x^2}. \tag{8}$$

Из равенства (8) вытекают следующие утверждения.

1) При $x=0$ имеем $y=b$. Следовательно, точка $(0, b)$ лежит на эллипсе. Обозначим ее через B.
2) При возрастании x от 0 до a y уменьшается.
3) Если $x=a$, то $y=0$. Следовательно, точка $(a, 0)$ лежит на эллипсе, обозначим ее через A.

4) При $x>a$ получаем мнимые значения y. Следовательно, точек эллипса, у которых $x>a$, не существует.

Итак, частью эллипса, расположенной в первом координатном угле, является дуга BA *.

Производя симметрию относительно координатных осей, получим весь эллипс (рис. 30).

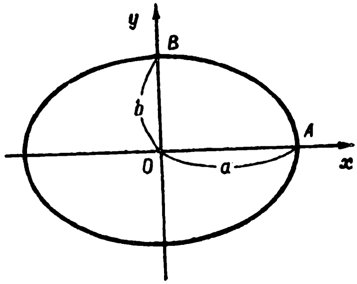


Рис. 30

Замечание. Если $a=b$, то уравнение (7) примет вид $x^2+y^2=a^2$. Это уравнение окружности (окружность — частный случай эллипса). Заметим, что эллипс можно получить из уравнения окружности радиуса a, если сжать ее в a/b раз вдоль оси Oy. При таком сжатии точка $(x;\ y)$ перейдет в точку $(x;\ y_1)$, где $y_1 = y\dfrac{b}{a}$. Подставляя $y = y_1\dfrac{a}{b}$ в уравнение окружности, получим уравнение эллипса: $\dfrac{x^2}{a^2}+\dfrac{(y_1)^2}{b^2}=1$.

Оси симметрии эллипса называются его *осями*, а точка пересечения осей — *центром эллипса*. Точки, в которых эллипс пересекает оси, называются его *вершинами*. Так как на основании равенства (6) $a \geqslant b$, то $2a$ есть длина *большой оси* симметрии эллипса, $2b$ — *малой оси*. Следовательно, числа a и b являются длинами соответственно *большой* и *малой* полуосей эллипса.

Введем еще одну величину, характеризующую форму эллипса.

Определение. *Эксцентриситетом* эллипса называется отношение фокусного расстояния $2c$ к длине $2a$ его большой оси.

Эксцентриситет обычно обозначают ε: $\varepsilon = \dfrac{c}{a}$. Так

---

* В гл. VI будет введено понятие направления выпуклости графика функции $y=f(x)$ и будет показано, что дуга BA направлена выпуклостью вверх.

как $c<a$, то $0 \leqslant \varepsilon < 1$, т. е. эксцентриситет эллипса меньше единицы. Заметив, что $c^2=a^2-b^2$, находим

$$\varepsilon^2 = \frac{c^2}{a^2} = \frac{a^2-b^2}{a^2} = 1 - \left(\frac{b}{a}\right)^2,$$

отсюда

$$\frac{b}{a} = \sqrt{1-\varepsilon^2}.$$

Из последнего равенства легко получить геометрическое истолкование эксцентриситета эллипса. При очень малом ε числа a и b почти равны, т. е. эллипс близок к окружности. Если же ε близко к единице, то число b весьма мало по сравнению с числом a и эллипс сильно вытянут вдоль большей оси. Таким образом, эксцентриситет эллипса характеризует меру вытянутости эллипса.

Как известно, планеты и некоторые кометы движутся по эллиптическим траекториям. Оказывается, эксцентриситеты планетных орбит весьма малы, а кометных — велики, т. е. близки к единице. Таким образом, планеты движутся почти по окружности, а кометы то приближаются к Солнцу (Солнце находится в одном из фокусов их орбиты), то удаляются от него.

2. Гипербола.

Определение. Гиперболой называется множество всех точек плоскости, для которых модуль разности расстояний от двух данных точек, называемых фокусами, есть величина постоянная, меньшая чем расстояние между фокусами.

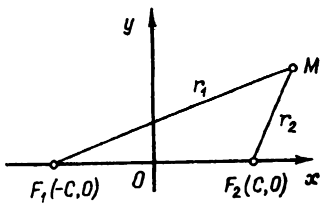


Рис. 31

Для вывода уравнения гиперболы введем на плоскости прямоугольную систему координат так, чтобы фокусы гиперболы лежали на оси абсцисс, а начало координат делило бы расстояние между фокусами пополам. Выведем уравнение гиперболы в выбранной системе координат.

Обозначим фокусы гиперболы через F_1 и F_2. Пусть точка M — произвольная точка гиперболы. Расстояние $|F_1F_2|$ между фокусами обозначим через $2c$, а модуль разности расстояний от точки M до фокусов — через $2a$. Так как по определению $||F_1M|-|F_2M||<$

$<|F_1F_2|$, то $2a<2c$, т. е. $a<c$. Числа $|F_1M|$ и $|F_2M|$ называются *фокальными радиусами* точки M и обозначаются r_1 и r_2 (рис. 31). Из определения следует, что точка $M(x;\ y)$ будет находиться на данной гиперболе в том и только в том случае, когда $|r_1-r_2|=2a$. Отсюда

$$r_1-r_2=\pm 2a. \tag{9}$$

Чтобы получить искомое уравнение гиперболы, нужно в равенстве (9) заменить переменные r_1 и r_2 их выражениями через координаты. Так как фокусы F_1 и F_2 расположены на оси Ox симметрично относительно начала координат, то они имеют соответственно координаты $(-c;\ 0)$ и $(c;\ 0)$. По формуле (1) § 2 находим

$$r_1=\sqrt{(x+c)^2+y^2},\quad r_2=\sqrt{(x-c)^2+y^2}. \tag{10}$$

Подставляя значения r_1 и r_2 в уравнение (9), получим

$$\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}=\pm 2a. \tag{11}$$

Уравнение (11) является *уравнением гиперболы*, которому будут удовлетворять координаты каждой точки гиперболы. Упростим это уравнение аналогично тому, как мы упрощали уравнение (3) для эллипса. Перенесем второй корень в правую часть уравнения, после чего возведем обе части равенства в квадрат:

$$(x+c)^2+y^2=4a^2\pm 4a\sqrt{(x-c)^2+y^2}+(x-c)^2+y^2,$$

или

$$cx-a^2=\pm a\sqrt{(x-c)^2+y^2}. \tag{12}$$

Снова возводим в квадрат:

$$c^2x^2-2a^2cx+a^4=a^2x^2-2a^2cx+a^2c^2+a^2y^2,$$

откуда

$$(c^2-a^2)x^2-a^2y^2=a^2(c^2-a^2). \tag{13}$$

Введем в рассмотрение новую величину

$$b=\sqrt{c^2-a^2}, \tag{14}$$

геометрический смысл которой будет раскрыт несколько позднее. Так как $c>a$, то $c^2-a^2>0$ и b — число положительное. Из равенства (14) имеем

$$b^2 = c^2 - a^2.$$

При этом уравнение (13) принимает вид

$$b^2x^2 - a^2y^2 = a^2b^2,$$

или

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \qquad (15)$$

Это и есть *каноническое уравнение* гиперболы.

Так же как для эллипса, аналогичными рассуждениями можно показать равносильность уравнений (15) и (11). Предлагается это сделать самостоятельно.

Исследуем форму гиперболы по уравнению (15).

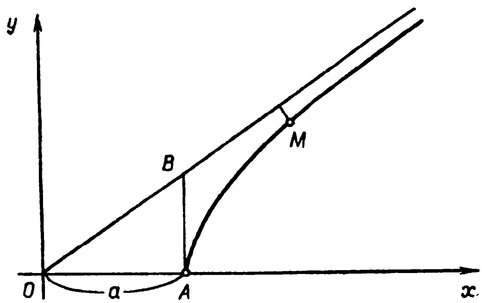


Рис. 32

Так как уравнение (15) содержит члены только с четными степенями текущих координат x и y, то по аналогии с эллипсом достаточно рассмотреть лишь часть гиперболы, лежащую в первом координатном угле. Разрешим уравнение (15) относительно y:

$$y = \frac{b}{a}\sqrt{x^2 - a^2}. \qquad (16)$$

Из равенства (16) вытекают следующие утверждения.

1) Если $0 \leqslant x < a$, то y получает мнимые значения, т. е. точек гиперболы с абсциссами $0 \leqslant x < a$ нет.
2) Если $x = a$, то $y = 0$, т. е. точка $(a;\ 0)$ принадлежит гиперболе. Обозначим ее через A.

3) Если $x>a$, то $y>0$, причем при возрастании x возрастает и y, т. е. при $x\to+\infty$ также и $y\to+\infty$. Переменная точка $M(x;\ y)$, описывающая график, движется все время «вправо» и «вверх», имея своим начальным положением точку $A(a,\ 0)$ (рис. 32). Здесь необходимо уточнить, как именно точка M «уходит в бесконечность». С этой целью мы наряду с уравнением (16) рассмотрим еще уравнение

$$y=\frac{b}{a}x, \tag{17}$$

которое, как мы уже знаем, определяет прямую с угловым коэффициентом $k=b/a$, проходящую через начало координат. Часть этой прямой, расположенная в первом координатном угле, изображена на рис. 33. Для построения ее был использован прямоугольный треугольник OAB с катетами $OA=a$ и $AB=b$.

Покажем, что точка M, уходя в бесконечность, неограниченно приближается к прямой (17), которая называется *асимптотой* гиперболы *.

Возьмем произвольное значение x $(x\geqslant a)$ и рассмотрим две точки $M(x;\ y)$ и $N(x;\ Y)$, где

$$y=\frac{b}{a}\sqrt{x^2-a^2} \text{ и } Y=\frac{b}{a}x.$$

Поскольку обе точки имеют одну и ту же абсциссу x, прямая, соединяющая точки M и N, перпендикулярна к оси Ox (см. рис. 33).

Подсчитаем длину отрезка MN. Прежде всего заметим, что при $a\leqslant x$

$$Y=\frac{b}{a}x=\frac{b}{a}\sqrt{x^2}>\frac{b}{a}\sqrt{x^2-a^2}=y.$$

Следовательно, $Y>y$, а это означает, что при одной и той же абсциссе точка гиперболы лежит под соответ-

---

* В гл. VI будет дано определение асимптоты графика функции $y=f(x)$ и показано, что прямая $y=\frac{b}{a}x$ является асимптотой гиперболы. Там же будет рассмотрен вопрос о направлении выпуклости гиперболы.

ствующей точкой асимптоты. Тогда

$$MN = Y - y = \frac{b}{a}x - \frac{b}{a}\sqrt{x^2 - a^2} = \frac{b}{a}(x - \sqrt{x^2 - a^2}) =$$

$$= \frac{b}{a}\,\frac{(x - \sqrt{x^2 - a^2})(x + \sqrt{x^2 - a^2})}{x + \sqrt{x^2 - a^2}} = \frac{ab}{x + \sqrt{x^2 - a^2}}.$$

Из полученного выражения следует, что дробь при $x \to +\infty$ стремится к нулю, так как знаменатель растет, а числитель есть постоянная величина ab. Следовательно, $|MN| = Y - y$ стремится к нулю при $x \to +\infty$.

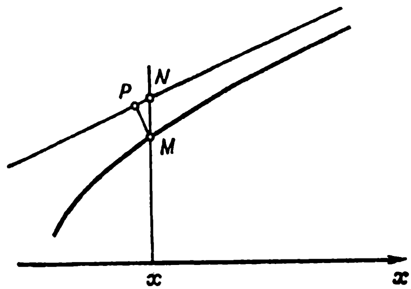


Рис. 33

Обозначим через P основание перпендикуляра, опущенного из точки M на прямую (17), $|MP|$ — расстояние от точки M до этой прямой. Очевидно, $|MP| < |MN|$, а так как $|MN| \to 0$, то и $|MP| \to 0$, что мы и хотели показать. Аналогичное рассуждение можно провести в любом координатном угле.

Итак, часть рассматриваемой гиперболы, лежащая в первом координатном угле, исходит из точки $A(a;\ 0)$ и идет бесконечно «направо» и «вверх», асимптотически приближаясь к прямой $y = \frac{b}{a}x$ (см. рис. 32).

Общий вид целой гиперболы теперь можно легко установить при помощи симметрии относительно координатных осей (рис. 34). Из рисунка видно, что гипербола имеет две асимптоты:

$$y = \frac{b}{a}x \text{ и } y = -\frac{b}{a}x,$$

первая из этих прямых нам уже знакома, вторая представляет собой ее симметричное отражение относительно оси Ox (или оси Oy).

Точки пересечения гиперболы с осью называются ее *вершинами* (они на рисунке обозначены буквами A' и A). Отрезки длиной $2a$ и $2b$, соединяющие середины противоположных сторон прямоугольника $BB'C'C$, называются *осями* гиперболы, а прямоугольник $BB'C'C$

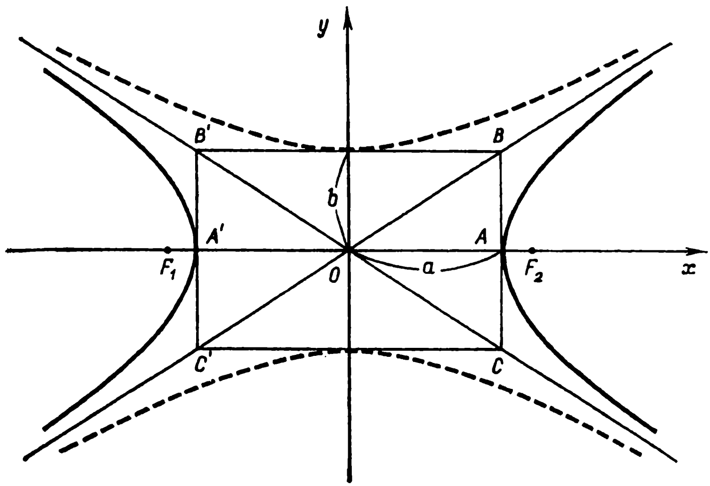


Рис. 34

называется *основным прямоугольником* гиперболы. Соответственно этому говорят, что уравнение (7) определяет гиперболу с полуосями a и b.

В случае канонического задания гиперболы координатные оси являются *осями симметрии* гиперболы, а начало координат — *центром симметрии.*

Уравнение вида

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$$

путем перестановки букв x и y, a и b сводится к уравнению (15). Отсюда ясно, что оно определяет гиперболу, расположенную так, как показано на рис. 34 пунктирными линиями; вершины ее лежат на оси Oy. Данное уравнение также называется каноническим уравнением гиперболы.

Гипербола с равными полуосями ($a=b$) называется *равносторонней*, и ее каноническое уравнение имеет вид

$$x^2-y^2=a^2.$$

Так как основной прямоугольник равносторонней гиперболы является квадратом, то асимптоты равносторонней гиперболы перпендикулярны друг к другу.

Определение. *Эксцентриситетом* гиперболы называется отношение расстояния между фокусами к расстоянию между вершинами: $\varepsilon=c/a$. Так как $c>a$, то $\varepsilon>1$, т. е. эксцентриситет гиперболы больше единицы. Заметив, что $c^2=a^2+b^2$, находим

$$\varepsilon^2=\frac{c^2}{a^2}=\frac{a^2+b^2}{a^2}=1+\left(\frac{b}{a}\right)^2,$$

отсюда

$$\frac{b}{a}=\sqrt{\varepsilon^2-1}$$

Из последнего равенства легко получить геометрическое истолкование эксцентриситета гиперболы. Чем меньше эксцентриситет, т. е. чем ближе он к единице, тем меньше отношение b/a, а это значит, что основной прямоугольник более вытянут в направлении оси, соединяющей вершины. Таким образом, эксцентриситет гиперболы характеризует форму ее основного прямоугольника, а значит, и форму самой гиперболы.

В случае равносторонней гиперболы ($a=b$) $\varepsilon=\sqrt{2}$.

3. Директрисы эллипса и гиперболы

Определение 1. Две прямые, перпендикулярные к большой оси эллипса и расположенные симметрично относительно центра на расстоянии $\dfrac{a}{\varepsilon}$ от него, называются *директрисами* эллипса.

Уравнения директрис эллипса, заданного каноническим уравнением (7); имеют вид

$$x=-\frac{a}{\varepsilon} \text{ и } x=\frac{a}{\varepsilon}.$$

Так как для эллипса $\varepsilon<1$, то $\dfrac{a}{\varepsilon}>a$. Отсюда следует, что правая директриса расположена правее правой вершины эллипса, а левая — левее его левой вершины (рис. 35).

Определение 2. Две прямые, перпендикулярные к той оси гиперболы, которая ее пересекает, и расположенные симметрично относительно центра на расстоянии a/ε от него, называются *директрисами* гиперболы.

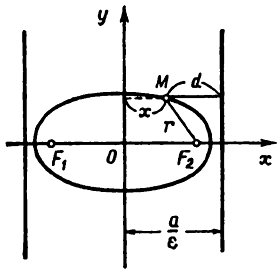


Рис. 35

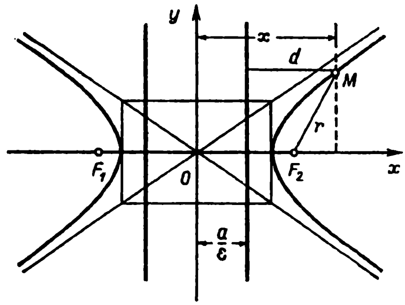


Рис. 36

Уравнения директрис гиперболы, заданной каноническим уравнением (15), имеют вид

$$x = -\frac{a}{\varepsilon} \text{ и } x = \frac{a}{\varepsilon}.$$

Так как для гиперболы $\varepsilon > 1$, то $\frac{a}{\varepsilon} < a$. Отсюда следует, что правая директриса расположена между центром и правой вершиной гиперболы, а левая — между центром и левой вершиной (рис. 36).

С помощью понятий директрисы и эксцентриситета можно сформулировать общее свойство, присущее эллипсу и гиперболе. Имеют место следующие две теоремы.

Теорема 3.5. *Если r — расстояние от произвольной точки M эллипса до какого-нибудь фокуса, d — расстояние от той же точки до соответствующей этому фокусу директрисы, то отношение r/d есть постоянная величина, равная эксцентриситету эллипса.*

Доказательство. Предположим для определенности, что речь идет о правом фокусе F_2 и правой директрисе. Пусть $M(x; y)$ — произвольная точка эллипса (см. рис. 35). Расстояние от точки M до правой директрисы выражается равенством

$$d=\frac{a}{\varepsilon}-x, \tag{18}$$

которое легко усматривается из чертежа. Из равенств (2) и (4) мы имеем

$$r=r_2=\sqrt{(x-c)^2+y^2}=a-\frac{c}{a}x.$$

Полагая здесь $\frac{c}{a}=\varepsilon$, получаем формулу расстояния от точки M до правого фокуса:

$$r=a-\varepsilon x. \tag{19}$$

Из соотношений (18) и (19) имеем

$$\frac{r}{d}=\frac{a-\varepsilon x}{\frac{a}{\varepsilon}-x}=\frac{(a-\varepsilon x)\varepsilon}{a-\varepsilon x}=\varepsilon \blacksquare$$

Теорема 3.6. *Если r — расстояние от произвольной точки M гиперболы до какого-нибудь фокуса, d — расстояние от той же точки до соответствующей этому фокусу директрисы, то отношение r/d есть величина постоянная, равная эксцентриситету гиперболы.*

Доказательство. Предположим для определенности, что речь идет о правом фокусе F_2 и правой директрисе. Пусть $M(x;\ y)$ — произвольная точка гиперболы (см. рис. 36). Нам придется рассмотреть два случая.

1) Точка M находится на правой ветви гиперболы. Тогда расстояние от точки M до правой директрисы выражается равенством

$$d=x-\frac{a}{\varepsilon}, \tag{20}$$

которое легко усматривается из чертежа. Из равенств (10) и (12) мы имеем

$$r=r_2=\sqrt{(x-c)^2+y^2}=\frac{c}{a}x-a.$$

Полагая здесь $c/a=\varepsilon$, получаем формулу расстояния от точки M до правого фокуса:

$$r=\varepsilon x-a. \tag{21}$$

Из соотношений (20) и (21) имеем

$$\frac{r}{d}=\frac{\varepsilon x-a}{x-\frac{a}{\varepsilon}}=\frac{(\varepsilon x-a)\,\varepsilon}{\varepsilon x-a}=\varepsilon.$$

2) Точка M находится на левой ветви гиперболы. Тогда расстояние от точки M до правой директрисы выражается равенством

$$d=-x+\frac{a}{\varepsilon}. \tag{22}$$

Из равенств (10) и (12)

$$r=r_1=\sqrt{(x+c)^2+y^2}=-\left(\frac{c}{a}x-a\right).$$

Полагая здесь $\frac{c}{a}=\varepsilon$, получаем формулу расстояния от точки M до правого фокуса:

$$r=-(\varepsilon x-a). \tag{23}$$

Из соотношений (22) и (23) имеем

$$\frac{r}{d}=\frac{-(\varepsilon x-a)}{-x+\frac{a}{\varepsilon}}=\frac{(-\varepsilon x+a)\,\varepsilon}{(-\varepsilon x+a)}=\varepsilon \blacksquare$$

Выявленное свойство эллипса и гиперболы можно положить в основу общего определения этих линий: множество точек, отношение расстояний которых до фокуса и до соответствующей директрисы есть величина постоянная, равная ε, есть эллипс, если $\varepsilon<1$, есть гипербола, если $\varepsilon>1$.

Естественно возникает вопрос, что представляет собой множество точек, определенное аналогичным образом, но при условии $\varepsilon=1$. Оказывается, это есть новая линия второго порядка, называемая параболой.

4. Парабола

Определение. Параболой называется множество всех точек плоскости, каждая из которых находится на одинаковом расстоянии от данной точки, называемой фокусом, и от данной прямой, называемой директрисой и не проходящей через фокус.

Для вывода уравнения параболы введем на плоскости прямоугольную систему координат так, чтобы ось абсцисс прошла через фокус перпендикулярно к директрисе, и будем считать ее положительным направлением направление от директрисы к фокусу; начало координат расположим посредине между фокусом и директрисой. Выведем уравнение параболы в выбранной системе координат.

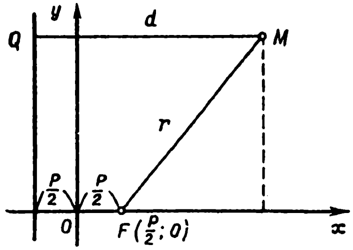


Рис. 37

Пусть $M(x;\ y)$ — произвольная точка параболы. Обозначим через r расстояние от точки M до фокуса $F(r=|FM|)$, через d — расстояние от точки M до директрисы, а через p — расстояние от фокуса до директрисы (рис. 37). Величину p называют *параметром* параболы, геометрический смысл его будет раскрыт несколько позже. Точка M находится на данной параболе в том и только в том случае, когда

$$r=d. \tag{24}$$

Чтобы получить искомое уравнение, нужно в равенстве (24) заменить переменные r и d их выражениями через текущие координаты x и y. Фокус F имеет координаты $(p/2;\ 0)$; приняв это во внимание и применяя формулу (1) § 2, находим

$$r=\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}\,. \tag{25}$$

Обозначим через Q основание перпендикуляра, опущенного из точки M на директрису. Очевидно, точка Q имеет координаты $(-p/2;\ y)$; отсюда и из формулы (1) § 2 получаем

$$d=|MQ|=\sqrt{\left(x+\frac{p}{2}\right)^2+(y-y)^2}=x+\frac{p}{2}. \tag{26}$$

Заменяя в равенстве (24) r и d их выражениями (25) и (26), найдем

$$\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}=x+\frac{p}{2}. \tag{27}$$

Это и есть искомое уравнение параболы, которому будут удовлетворять координаты каждой точки параболы.

Приведем уравнение параболы к более удобному виду, для чего возведем обе части равенства (27) в квадрат:

$$x^2 - px + \frac{p^2}{4} + y^2 = x^2 + px + \frac{p^2}{4},$$

или

$$y^2 = 2px. \qquad (28)$$

Теперь проверим, не приобрело ли уравнение (28) после возведения в квадрат обеих частей равенства (27) «лишних» корней. Для этого достаточно показать, что для любой точки, координаты x и y которой удовлетворяют уравнению (28), выполнено соотношение (24). Действительно, из уравнения (28), вытекает, что $x \geqslant 0$, поэтому для точек с неотрицательными абсциссами $d = \frac{p}{2} + x$. Подставляя значение y^2 из (28) в выражение (25) для r и учитывая, что $x \geqslant 0$, получим $r = \frac{p}{2} + x$, т. е. величины r и d равны, что и требовалось доказать. Таким образом, уравнению (28) удовлетворяют координаты точек данной параболы, и только они, т. е. уравнение (28) есть уравнение этой параболы.

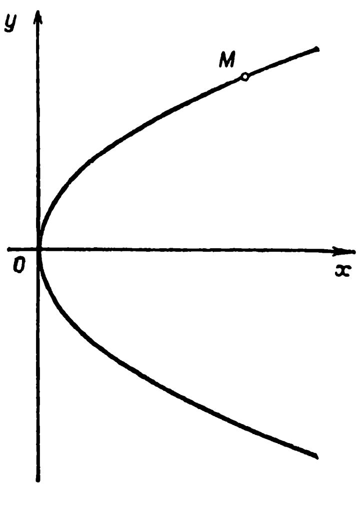


Рис. 38

Уравнение (28) называется *каноническим уравнением параболы*. Это есть уравнение второй степени. Таким образом, парабола есть линия второго порядка.

Исследуем теперь форму параболы по ее уравнению (28).

Так как уравнение (28) включает y только в четной степени, то парабола симметрична относительно оси Ox. Поэтому нам достаточно изучить лишь часть ее, лежащую в верхней полуплоскости. Для этого разрешим уравнение (28) относительно y:

$$y = \sqrt{2px}. \tag{29}$$

Из равенства (29) вытекают следующие утверждения.

1) Если $x<0$, то уравнение (29) дает мнимые значения y. Следовательно, левее оси Oy ни одной точки параболы нет.
2) Если $x=0$, то $y=0$. Таким образом, начало координат лежит на параболе и является самой «левой» ее точкой.
3) При возрастании x возрастает и y, причем если $x \to +\infty$, то и $y \to +\infty$.

Таким образом, переменная точка $M(x;\ y)$, описывающая рассматриваемую часть параболы, исходит из начала координат и движется «вправо» и «вверх», причем удаление точки M как от оси Oy, так и от оси Ox является бесконечным.

Теперь, когда мы установили форму части параболы при ее симметричном отражении относительно оси Ox, мы получим всю параболу (рис. 38), заданную уравнением (28).

Точка O называется *вершиной* параболы, ось Ox — *осью* параболы. Число p, т. е. параметр параболы, как мы знаем, выражает расстояние от фокуса до директрисы. Естественно поставить вопрос: как влияет параметр параболы на ее форму? С этой целью возьмем какое-нибудь определенное значение абсциссы, например $x=1$, и найдем из уравнения (28) соответствующие значения ординаты: $y = \pm\sqrt{2p}$. Мы получаем на параболе две точки $M_1(1;\ +\sqrt{2p})$ $M_2(1;\ -\sqrt{2p})$, симметричные относительно ее оси; расстояние между ними равно $2\sqrt{2p}$, т. е. тем больше, чем больше p. Следовательно, параметр p характеризует «ширину» области, ограниченной параболой. В этом и заключается геометрический смысл параметра p.

Парабола, уравнение которой $y^2=-2px$, $p>0$, расположена слева от оси ординат (рис. 39, *а*). Вершина этой параболы лежит в начале координат, осью симметрии является ось Ox.

По аналогии с предыдущим мы можем утверждать, что уравнение $x^2=2py$, $p>0$, является уравнением параболы, вершина которой лежит в начале координат, осью симметрии является ось Oy (рис. 39, *б*). Эта парабола лежит выше оси абсцисс. Уравнение же $x^2 = -2py$, $p>0$, также является уравнением параболы, лежащей ниже оси Ox, с вершиной в начале координат (рис. 39, *в*).

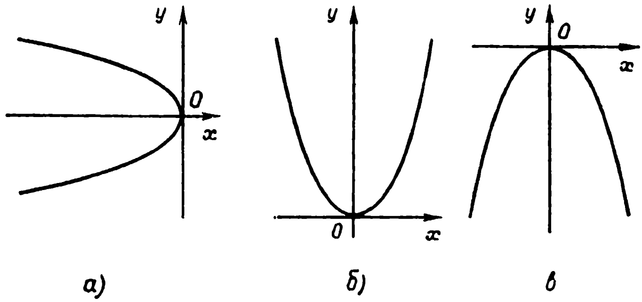


Рис. 39

§ 8. Общее уравнение линии второго порядка

Важной задачей аналитической геометрии является исследование общего уравнения линии второго порядка и приведение его к простейшим (каноническим) формам.

Общее уравнение линии второго порядка имеет следующий вид:

$$Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0, \qquad (1)$$

где коэффициенты A, $2B$, C, $2D$, $2E$ и F — любые числа и, кроме того, числа A, B и C не равны нулю одновременно.

1. Приведение общего уравнения линии второго порядка к простейшему виду

Лемма 3.1. *Пусть в прямоугольной системе координат Oxy задано уравнение* (1). *Тогда при помощи параллельного переноса начала координат в некоторую точку* (x_0; y_0) *и последующего поворота осей координат на некоторый угол α уравнение (1) приводится к виду*

$$A'x''^2 + C'y''^2 + F' = 0, \qquad (2)$$

где A', C', F' — некоторые числа, (x''; y'') — координаты точки в новой системе координат.

Доказательство. Пусть прямоугольная система координат $O'x'y'$ получена параллельным переносом начала координат Oxy в точку $(x_0;\ y_0)$. Тогда старые координаты x, y будут связаны с новыми x', y' формулами

$$x=x'+x_0, \qquad y=y'+y_0$$

(см. (1), § 4). В новых координатах уравнение (1) примет вид

$$Ax'^2+2Bx'y'+Cy'^2+2D'x'+2E'y'+F'=0, \qquad (3)$$

где

$$D'=Ax_0+By_0+D=0,$$

$$E'=Bx_0+Cy_0+E=0,$$

$$F'=Ax_0^2+2Bx_0y_0+Cy_0^2+2Dx_0+2Ey_0+F.$$

В преобразованном уравнении (1) коэффициенты D' и E' обратятся в нуль, если мы подберем координаты точки $(x_0;\ y_0)$ так, чтобы выполнялись равенства

$$\begin{cases} Ax_0+By_0+D=0, \\ Bx_0+Cy_0+E=0. \end{cases} \qquad (4)$$

Если пара чисел x_0, y_0 представляет собой решение системы (4), то уравнение (3) можно записать в виде

$$Ax'^2+2Bx'y'+Cy'^2+F'=0. \qquad (5)$$

Пусть теперь прямоугольная система координат $O'x''y''$ получена поворотом системы $O'x'y'$ на угол α. Тогда координаты x', y' будут связаны с координатами x'', y'' формулами

$$x'=x''\cos\alpha-y''\sin\alpha, \qquad y'=x''\sin\alpha+y''\cos\alpha$$

(см. (2), § 4). В системе координат $O'x''y''$ уравнение (5) примет вид

$$A'x''^2+2B'x''y''+C'y''^2+F'=0, \qquad (6)$$

где

$$A'=A\cos^2\alpha+2B\cos\alpha\sin\alpha+C\sin^2\alpha,$$

$$B'=-A\sin\alpha\cos\alpha+B(\cos^2\alpha-\sin^2\alpha)+C\sin\alpha\cos\alpha,$$

$$C'=A\sin^2\alpha-2B\cos\alpha\sin\alpha+C\cos^2\alpha.$$

Выберем угол α так, чтобы коэффициент B' в (6) обратился в нуль, это приведет к уравнению $2B\cos 2\alpha = (A-C)\sin 2\alpha$. Если $A=C$, то $\cos 2\alpha=0$ и можно положить $\alpha=\pi/4$. Если $A\neq C$, то выбираем $\alpha = \frac{1}{2}\operatorname{arctg}\frac{2B}{A-C}$ и уравнение (6) примет вид

$$A'x'^2 + C'y'^2 + F' = 0,$$

т. е. получили уравнение (2) ■

Замечание. Уравнения (4) называются *уравнениями центра линии второго порядка*, а точка $(x_0;\ y_0)$, где x_0 и y_0 — решения системы (4), называется *центром* этой линии. Заметим, что необходимым и достаточным условием существования единственного решения системы (4) является неравенство нулю выражения $AC-B^2$, называемого *определителем* системы (ч. II, гл. X, § 2).

2. Инвариантность выражения $AC-B^2$. Классификация линий второго порядка. При упрощении общего уравнения линии второго порядка коэффициенты при старших членах A, B и C при параллельном переносе осей координат, как следует из леммы 3.1, не меняются, но они меняются при повороте осей координат. Однако выражение $AC-B^2$ остается неизменным как при переносе, так и при повороте осей, т. е. не зависит от преобразования координат. Действительно, при параллельном переносе этот факт очевиден (см. (1) и (5)), проверим его при повороте осей. Для этого воспользуемся выражениями коэффициентов A', B' и C' уравнения (6). Имеем

$$A'C' - B'^2 = (A\cos^2\alpha + 2B\sin\alpha\cos\alpha + C\sin^2\alpha)\times$$

$$\times(A\sin^2\alpha - 2B\sin\alpha\cos\alpha + C\cos^2\alpha) -$$

$$-[(C-A)\sin\alpha\cos\alpha + B(\cos^2\alpha - \sin^2\alpha)]^2.$$

Раскрыв скобки и сделав приведение подобных членов, получим

$$A'C' - B'^2 = AC(\cos^2\alpha + \sin^2\alpha)^2 -$$

$$-B^2(\cos^2\alpha + \sin^2\alpha)^2 = AC - B^2,$$

что и требовалось.

Величина $AC-B^2$ называется *инвариантом* общего уравнения линии второго порядка. Она имеет важное значение в исследовании линий второго порядка.

В зависимости от знака величины $AC-B^2$ линии второго порядка разделяются на следующие три типа:

1°) эллиптический тип, если $AC-B^2>0$;

2°) гиперболический тип, если $AC-B^2<0$;

3°) параболический тип, если $AC-B^2=0$.

Перейдем к рассмотрению линий различных типов.

1°. Эллиптический тип. Поскольку $AC-B^2>0$, то согласно лемме 3.1 общее уравнение линии второго порядка может быть приведено к виду

$$Ax^2+Cy^2+F=0.$$

Возможны следующие случаи.

а) $A>0$, $C>0$ ($A<0$, $C<0$ сводится к $A>0$, $C>0$ путем умножения уравнения на -1) и $F<0$. Перенесем F в правую часть и разделим на него. Уравнение примет вид

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,$$

где $a^2=-F/A$, $b^2=-F/C$. Сопоставляя полученное уравнение с уравнением эллипса (см. (7), § 7), заключаем, что оно является *каноническим уравнением эллипса.*

б) $A>0$, $C>0$ и $F>0$. Тогда аналогично предыдущему уравнение приводится к виду

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=-1.$$

Этому уравнению не удовлетворяют координаты никакой точки плоскости. В этом случае уравнение называется *уравнением мнимого эллипса.*

в) $F=0$. Уравнение имеет вид

$$a^2x^2+c^2y^2=0.$$

Ему удовлетворяют координаты лишь одной точки $x=0$, $y=0$. Такое уравнение назовем *уравнением пары мнимых пересекающихся прямых.*

2°. Гиперболический тип. Поскольку $AC-B^2<0$, то согласно лемме 3.1 общее уравнение линии второго порядка приводится к виду

$$Ax^2+Cy^2+F=0.$$

Возможны следующие случаи.

а) $A>0$, $C<0$ ($A<0$, $C>0$ сводится к $A>0$, $C<0$ путем умножения уравнения на -1) и $F\neq 0$. Пусть $F<0$. Перенесем F в правую часть и разделим на него. Уравнение примет вид

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1,$$

где $a^2=-F/A$, $b^2=F/C$. Сопоставляя с уравнением гиперболы (см. (15), § 7), заключаем, что полученное уравнение является *каноническим уравнением гиперболы.*

б) $A>0$, $C<0$ и $F=0$. Уравнение примет вид

$$a^2x^2-c^2y^2=0, \qquad \text{или} \quad (ax-cy)(ax+cy)=0.$$

Последнему уравнению удовлетворяют лишь координаты точек плоскости, расположенных на прямых $(ax-cy)=0$ и $(ax+cy)=0$, пересекающихся в начале координат, и мы имеем, таким образом, *пару пересекающихся прямых.*

3°. Параболический тип. Если $AC-B^2=0$, то путем поворота осей координат общее уравнение линии второго порядка может быть приведено к виду

$$Ax^2+Cy^2+2Ey+2Dx+F=0. \tag{7}$$

Здесь $AC=0$, и, следовательно, один из коэффициентов A или C равен нулю.

Пусть $A=0$, $C\neq 0$. Представим уравнение (7) в виде

$$C\left[y^2+\frac{2E}{C}y+\left(\frac{E}{C}\right)^2\right]+2Dx+F-\frac{E}{C}^2=0,$$

или

$$C\left(y+\frac{E}{C}\right)^2+2Dx+F^*=0.$$

Если сделаем перенос начала координат в соответствии с формулами перехода $x'=x$, $y'=y+E/C$ параллельно оси Oy в точку $(0; -E/C)$, то получим уравнение вида

$$Cy'^2+2Dx'+F^*=0.$$

Возможны следующие случаи.

а) $D \neq 0$. Преобразуем уравнение

$$Cy'^2 + 2D\left(x' + \frac{E^*}{2D}\right) = 0.$$

Если теперь сделаем перенос начала координат в соответствии с формулами перехода $x'' = x' + \frac{F^*}{2D}$, $y'' = y'$ параллельно оси Ox' в точку $(-F^*/2D;\ 0)$, то приведем уравнение к виду

$$Cy''^2 + 2Dx'' = 0, \text{ или } y''^2 = 2px'',$$

где $p = -D/C$. Сопоставляя последнее уравнение с уравнением параболы (см. (28), § 7), заключаем, что оно является *каноническим уравнением параболы.*

б) $D=0$. Уравнение имеет вид

$$Cy'^2 + F^* = 0.$$

Если C и F^* имеют разные знаки, то $y'^2 - \frac{F^*}{C} = (y'-a)(y'+a) = 0$ и уравнение представляет собой *пару параллельных прямых.* Если C и F^* имеют одинаковые знаки, то $y'^2 + \frac{F^*}{C} = y'^2 + a = 0$ и этому уравнению не удовлетворяют координаты никакой точки плоскости. Уравнение определяет *пару мнимых параллельных прямых.* Наконец, если $F^*=0$, то $y'^2 = 0$ представляет собой ось Ox', уравнение которой $y'=0$. Это уравнение можно рассматривать как предельный случай при $F^* \to 0$, т. е. *как пару совпавших прямых.*

На этом заканчивается исследование общего уравнения линии второго порядка. Полученные результаты сформулируем в виде теоремы.

Теорема 3.7. *Пусть в прямоугольной системе координат задано общее уравнение линии второго порядка*

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0.$$

Тогда существует такая прямоугольная система координат, в которой это уравнение принимает один из следующих девяти канонических видов: 1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

(*эллипс*); 2) $\frac{x^2}{a^2}+\frac{y^2}{b^2}=-1$ (*мнимый эллипс*); 3) a^2x^2+ $+c^2y^2=0$ (*пара мнимых пересекающихся прямых*); 4) $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ (*гипербола*); 5) $a^2x^2-c^2y^2=0$ (*пара пересекающихся прямых*); 6) $y^2=2px$ (*парабола*); 7) $y^2-a^2=0$ (*пара параллельных прямых*); 8) $y^2+a=0$ (*пара мнимых параллельных прямых*); 9) $y^2=0$ (*пара совпавших прямых*).

Глава IV. ФУНКЦИИ ОДНОЙ ПЕРЕМЕННОЙ

С этой главы мы начинаем изучение важнейшего понятия математического анализа — понятия функции. С помощью теории пределов мы введем понятие предела функции, а также познакомимся с понятием непрерывности функции.

§ 1. Понятие функции

1. Определение функции

Определение. Пусть X и Y — некоторые множества. *Функцией* f называется множество пар чисел $(x;\ y)$ таких, что $x\in X$, $y\in Y$ и каждое x входит в одну и только одну пару этого множества. В этом случае говорят, что числу x поставлено в соответствие число y и пишут $y=f(x)$. При этом y называют *зависимой переменной*, x — *независимой переменной* (или *аргументом*), множество X — *областью определения* (или *существования*) функции, а множество Y — *множеством значений* функции.

Кроме буквы f для обозначения функций употребляются и другие буквы, например: $y=y(x)$, $y=g(x)$, $y=\varphi(x)$, $y=A(x)$, $y=F(x)$ и т. д. Другими буквами могут обозначаться y и x. Иногда зависимую переменную y также называют функцией.

Пусть на некотором множестве X определена функция $f(x)$, тогда значение этой функции, соответствующее некоторому значению аргумента x_0, обозначается $f(x_0)$. Например, если $f(x)=x^2$, то $f(3)=9$, $f(-2)=4$ и т. д.

Функция, все значения которой равны между собой, называется *постоянной*. Постоянная функция часто обозначается буквой C. Про функцию $f(x)$, определенную на некотором множестве X, говорят, что она *ограничена*, если есть число $M>0$ такое, что для любого $x \in X$ будет $|f(x)| \leqslant M$. Например, функция $f(x) = \sin x$ ограничена на всей числовой прямой, так как $|\sin x| \leqslant 1$ при любом x.

На плоскости функция изображается в виде *графика* — множества точек $(x; y)$, координаты которых связаны соотношением $y=f(x)$, называемым *уравнением* графика.

2. Способы задания функций. Задать функцию — значит указать, как по каждому значению аргумента x находить ему соответствующее значение функции $f(x)$. Существуют три основных способа задания функций: *аналитический, табличный* и *графический*.

1°. Аналитический способ состоит в том, что зависимость между переменными величинами определяется с помощью формулы, указывающей, какие действия нужно выполнить, чтобы получить значение функции, соответствующее данному значению аргумента. Обратимся к примерам.

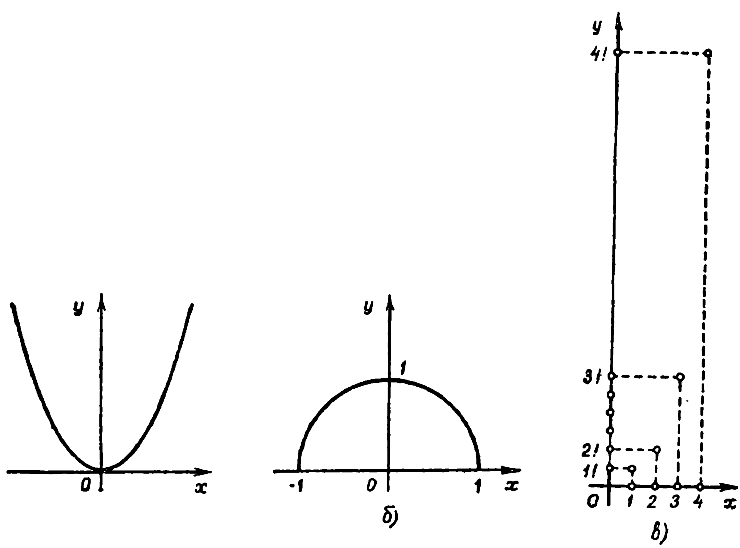


Рис. 40

1) Формула $y=x^2$ задает функцию, область определения которой есть интервал $(-\infty, +\infty)$, а множество значений — полуинтервал $[0, +\infty)$ (рис. 40, *а*).

2) Формула $y=\sqrt{1-x^2}$ задает функцию, областью определения которой является отрезок $[-1, 1]$, а множеством значений — отрезок $[0, 1]$ (рис. 40, *б*).

3) Формула $y=n!$ ставит в соответствие каждому натуральному числу (т. е. целому положительному числу) n число y. Например, если $n=3$, то $y=3!$ Таким образом, формула $y=n!$ задает функцию, область определения которой есть $\{1, 2, 3, \ldots, n, \ldots\}$, а множество значений — $\{1!, 2!, 3!, \ldots, n!, \ldots\}$ (рис. 40, *в*).

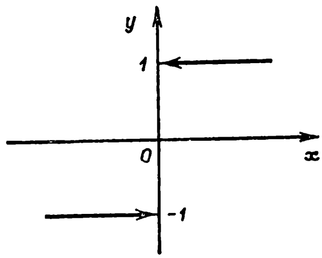


Рис. 41

4) Формула

$$y=\operatorname{sgn} x=\begin{cases} +1, & \text{если } x>0, \\ 0, & \text{если } x=0, \\ -1, & \text{если } x<0 \end{cases}$$

(термин sgn происходит от латинского слова signum — знак), задает функцию, определенную на всей прямой $(-\infty, +\infty)$, а множество всех ее значений состоит из трех точек: -1, 0 и $+1$ (рис. 41).

5) Функция Дирихле*

$$y=\begin{cases} 1, & \text{если } x \text{ — рациональное число,} \\ 0, & \text{если } x \text{ — иррациональное число.} \end{cases}$$

Эта формула задает функцию, определенную на всей прямой $(-\infty, +\infty)$, а множество ее значений состоит из двух точек: 0 и 1.

2°. Табличный способ. Рассмотрим таблицу.

x	0	0,1	0,2	3	0,6	4	0,8	1,5	2
y	-1	10	1	2	-3	-2	0,5	5	7

Подставим в соответствие каждому x, написанному в

* Дирихле Петер Густав Лежен (1805—1859) — немецкий математик.

первой строке таблицы, число y, стоящее во второй строке под этим числом x, и про получившуюся функцию будем говорить, что она задана таблицей. Областью определения указанной функции является множество, состоящее из девяти чисел x, перечисленных в первой строке таблицы, а множеством ее значений — множество, состоящее из девяти чисел y, перечисленных во второй строке таблицы.

Заметим, что с помощью таблицы можно задать функцию только при конечном числе значений аргумента.

Таблицами часто пользуются для задания функций, так всем известны, например, таблицы тригонометрических функций, таблицы логарифмов и многие другие.

3°. Графический способ задания функции обычно используется в практике физических измерений, когда соответствие между переменными x и y задается посредством графика. Во многих случаях графики чертятся с помощью самопишущих приборов. Так, например, для измерения атмосферного давления на различных высотах пользуются специальным самопишущим прибором — барографом, который на движущейся ленте записывает в виде кривой линии изменение давления в зависимости от высоты.

3. Классификация функций. Постоянная функция $f(x)=C$, $C=\text{const}$, степенная функция x^α (α — любое число), показательная функция a^x $(0<a\neq 1)$, логарифмическая функция $\log_a x$ $(0<a\neq 1)$, тригонометрические функции: $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$ и обратные тригонометрические функции: $\arcsin x$, $\arccos x$, $\operatorname{arctg} x$, $\operatorname{arcctg} x$ носят название *простейших элементарных функций*.

Простейшие элементарные функции играют важную роль в раскрытии основных понятий анализа, они составляют базу для изучения более сложных функций. Начало изучения этих функций положено в элементарной математике. Сведения о них мы будем постепенно пополнять новыми, полученными в результате дальнейших наших исследований.

Все функции, получаемые посредством *конечного* числа арифметических действий над простейшими элементарными функциями, а также получаемые путем суперпозиции (или наложения) этих функций, составляют так называемый класс *элементарных функций*. Например:

$$f(x) = |x| \ (|x| = \sqrt{x^2}),$$

$$f(x) = \ln|\sin 3x| - e^{\operatorname{arctg}\sqrt{x}} \text{ и т. д.}$$

Имеет место следующая классификация элементарных функций.

1) Функции вида

$$P(x) = a_0x^m + a_1x^{m-1} + \ldots + a_{m-1}x + a_m,$$

где $m \geqslant 0$ — целое число, a_0, a_1, ..., a_m — любые числа — коэффициенты ($a_0 \neq 0$), называется *целой рациональной функцией*, или *алгебраическим многочленом степени m*.

2) Функция, представляющая собой отношение двух целых рациональных функций

$$R(x) = \frac{a_0x^m + a_1x^{m-1} + \ldots + a_{m-1}x + a_m}{b_0x^n + b_1x^{n-1} + \ldots + b_{n-1}x + b_n},$$

называется *дробно-рациональной функцией*.

Совокупность целых рациональных и дробно-рациональных функций образует класс *рациональных функций*.

3) Функция, полученная путем конечного числа суперпозиций и четырех арифметических действий над степенными функциями как с целыми, так и с дробными показателями, называется *иррациональной функцией*. Например:

$$f(x) = \sqrt{x},\ f(x) = x + \sqrt{x},\ f(x) = \sqrt{\frac{5x^2 + 4x - 7}{3x^2 - 8x + 4}} +$$

$$+ (\sqrt[5]{x} + x)^3 \text{ и т. д.}$$

Заметим, что класс рациональных функций содержится в классе иррациональных функций.

4) Всякая неиррациональная функция называется *трансцендентной функцией*. Например: $f(x) = \sin x$, $f(x) = \sin x + x$ и т. д.

§ 2. Предел функции

Пусть функция $f(x)$ определена на некотором промежутке X и пусть точка $x_0 \in X$ или $x_0 \notin X$. Возьмем из X последовательность точек, отличных от x_0.

$$x_1,\ x_2,\ \ldots,\ x_n,\ \ldots, \tag{1}$$

сходящуюся к x_0*. Значения функции в точках этой последовательности также образуют числовую последовательность:

$$f(x_1),\ f(x_2),\ f(x_3),\ \ldots,\ f(x_n),\ \ldots, \tag{2}$$

по отношению к которой можно ставить вопрос о существовании предела.

Определение 1. Число A называется *пределом* функции $f(x)$ в точке $x=x_0$, если для *любой* сходящейся к x_0 последовательности (1) значений аргумента x, отличных от x_0, соответствующая последовательность (2) значений функции сходится к числу A.

Символически это записывается так: $\lim\limits_{x\to x_0} f(x)=A$.

Отметим, что функция $f(x)$ может иметь в точке x_0 только один предел. Это вытекает из того, что последовательность $\{f(x_n)\}$ может иметь только один предел.

Рассмотрим примеры:

1) Функция $f(x)=C=\text{const}$ имеет предел в каждой точке x_0 числовой прямой. В самом деле, если (1) — любая последовательность, сходящаяся к x_0, то последовательность (2) имеет вид $C, C, \ldots, C, \ldots$, т. е. если $x_n \to x_0$, то $f(x_n)=C$. Отсюда заключаем, что $f(x_n)\to C$ при $n\to\infty$ или $\lim\limits_{x\to x_0} f(x)=C$.

2) Функция $f(x)=x$ имеет в любой точке x_0 числовой прямой предел, равный x_0. В этом случае последовательности (1) и (2) тождественны, т. е. если $x_n\to x_0$, то $f(x_n)=x_n$. Следовательно, $f(x_n)\to x_0$ при $n\to\infty$ или

$$\lim_{x\to x_0} f(x)=\lim_{x\to x_0} x=f(x_0)=x_0.$$

3) Функция Дирихле, значения которой в рациональных точках равны единице, а в иррациональных — нулю, не имеет предела ни в одной точке x_0 числовой прямой. Действительно, для сходящейся к точке x_0 по-

---

* Предполагается, что такая последовательность существует.

следовательности рациональных значений аргумента предел соответствующей последовательности значений функции равен единице, а для сходящейся к точке x_0 последовательности иррациональных значений аргумента предел соответствующей последовательности значений функции равен нулю.

Существует и другое определение предела функции.

Определение 2. Число A называется пределом функции $f(x)$ в точке $x=x_0$, если для любого числа $\varepsilon>0$ существует число $\delta>0$ такое, что для всех $x\in X$, $x\neq x_0$, удовлетворяющих неравенству $|x-x_0|<\delta$, выполняется неравенство $|f(x)-A|<\varepsilon$.

Первое определение основано на понятии предела числовой последовательности, поэтому его часто называют определением «на языке последовательностей». Определение же второе использует понятие ε-окрестности и δ-окрестности, и потому его называют определением «на языке ε—δ».

Теорема 4.1. *Первое и второе определения предела функции эквивалентны.*

Доказательство. Пусть A есть предел $f(x)$ в точке x_0 по первому определению. Покажем, что A есть предел по второму определению. Предположим обратное, т. е. A не является пределом этой функции по второму определению. Это значит, что не для любого $\varepsilon>0$ можно найти такое $\delta>0$, чтобы из неравенства $0<|x-x_0|<\delta$ следовало неравенство $|f(x)-A|<\varepsilon$, т. е. существует такое $\varepsilon=\varepsilon_0$, для которого, какое бы δ мы ни взяли, найдется хоть одна точка $x\neq x_0$ такая, что $|x-x_0|<\delta$, но $|f(x)-A|\geqslant\varepsilon_0$. Будем выбирать в качестве δ последовательно числа:

$$1,\ 1/2,\ 1/3,\ \ldots,\ 1/n,\ \ldots,$$

Тогда

для $\delta=1$ в X существует $x_1\neq x_0$, что $|x_1-x_0|<1$, а $|f(x_1)-A|\geqslant\varepsilon_0$;

для $\delta=1/2$ в X существует $x_2\neq x_0$, что $|x_2-x_0|<1/2$, а $|f(x_2)-A|\geqslant\varepsilon_0$;

для $\delta=1/3$ в X существует $x_3\neq x_0$, что $|x_3-x_0|<1/3$, а $|f(x_3)-A|\geqslant\varepsilon_0$;

. . . .

для $\delta=1/n$ в X существует $x_n\neq x_0$, что $|x_n-x_0|<1/n$, а $|f(x_n)-A|\geqslant\varepsilon_0$.

. . . .

В результате получается последовательность точек, отличных от x_0:

$$x_1,\ x_2,\ x_3,\ \ldots,\ x_n,\ \ldots,$$

сходящаяся к точке x_0, так как $|x_n - x_0| < 1/n \to 0$ при $n \to \infty$. Тогда, согласно первому определению, соответствующая последовательность $\{f(x_n)\}$ значений функции сходится к числу A. Следовательно, для ε_0 найдется номер N такой, что для всех $n > N$ $|f(x_n) - A| < \varepsilon_0$. Но этого быть не может, так как для всех x_n выполняется неравенство $|f(x_n) - A| \geqslant \varepsilon_0$. Полученное противоречие и доказывает, что число A есть предел функции $f(x)$ в точке x_0 по второму определению.

Пусть теперь A есть предел $f(x)$ в точке x_0 по второму определению. Это значит, что для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что из неравенства $0 < |x - x_0| < \delta$ будет следовать неравенство $|f(x) - A| < \varepsilon$. Покажем, что A есть предел $f(x)$ по первому определению. Возьмем любую последовательность точек:

$$x_1,\ x_2,\ x_3,\ \ldots,\ x_n,\ \ldots,$$

сходящуюся к точке $x_0 (x_n \neq x_0)$. Тогда для указанного значения $\delta > 0$, зависящего от ε, найдется такое N, что при $n > N$ будет $|x_n - x_0| < \delta$. Но вместе с этим будет выполняться и неравенство $|f(x_n) - A| < \varepsilon$. А так как ε выбиралось произвольно, то это и означает, что $f(x_n) \to A$ для любой последовательности $\{x_n\}$, сходящейся к точке x_0, т. е. число A будет пределом $f(x)$ в точке x_0 по первому определению ■

После того как мы установили эквивалентность обоих определений предела функции, можно пользоваться любым из них, в зависимости от того, какое более удобно при решении той или иной задачи.

Заметим, что определение предела функции «на языке последовательностей» называют также определением предела функции по Гейне*, а определение предела функции «на языке ε—δ» — определением предела функции по Коши**.

В дальнейшем мы будем использовать понятия од-

---

* Гейне Генрих Эдуард (1821—1881) — немецкий математик.
** Коши Огюстен Луи (1789—1857) — французский математик.

носторонних пределов функции, которые определяются следующим образом.

Определение 3. Число A называется *правым (левым) пределом* функции $f(x)$ в точке x_0, если для любой сходящейся к x_0 последовательности (1), элементы x_n которой больше (меньше) x_0, соответствующая последовательность (2) сходится к A.

Символическая запись: $\lim\limits_{x\to x_0+} f(x) = A$ ($\lim\limits_{x\to x_0-} f(x) = A$).

В качестве примера рассмотрим функцию $f(x) = \operatorname{sgn} x$*.

Эта функция имеет в нуле правый и левый пределы: $\lim\limits_{x\to 0+} \operatorname{sgn} x = 1$, $\lim\limits_{x\to 0-} \operatorname{sgn} x = -1$. В самом деле, если (1) — любая сходящаяся к нулю последовательность значений аргумента этой функции, элементы x_n которой больше нуля ($x_n > 0$), то $\operatorname{sgn} x_n = 1$ и $\lim\limits_{x\to 0+} \operatorname{sgn} x = 1$. Аналогично устанавливается, что $\lim\limits_{x\to 0-} \operatorname{sgn} x = -1$.

Можно дать равносильное определение односторонних пределов функции «на языке ε—δ»: число A называется правым (левым) пределом функции $f(x)$ в точке x_0, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех x, удовлетворяющих неравенствам $x_0 < x < x_0 + \delta$ ($x_0 - \delta < x < x_0$), выполняется неравенство $|f(x) - A| < \varepsilon$.

Замечание. Если в точке x_0 правый и левый пределы функции равны, то в точке x_0 существует предел этой функции, равный указанным односторонним пределам.

В самом деле, пусть $\lim\limits_{x\to x_0-} f(x) = \lim\limits_{x\to x_0+} f(x) = A$. Тогда, согласно только что данному определению предела функции справа и слева, для любого $\varepsilon > 0$ существуют такие $\delta_1 > 0$ и $\delta_2 > 0$, что для всех x, удовлетворяющих неравенствам $x_0 - \delta_1 < x < x_0$, и для всех x, удовлетворяющих неравенствам $x_0 < x < x_0 + \delta_2$, выполняется неравенство $|f(x) - A| < \varepsilon$. Если взять $\delta = \min\{\delta_1, \delta_2\}$, то очевидно, что для всех x, удовлетворяющих неравенству $|x - x_0| < \delta$, $x \neq x_0$, будет выполняться неравенство $|f(x) - A| < \varepsilon$. А это и означает, что $\lim\limits_{x\to x_0} f(x) = A$.

---

* Определение функции $f(x) = \operatorname{sgn} x$ дано в п. 2, § 1.

Верно и обратное: если в точке x_0 предел функции существует, то в этой точке существуют также правый и левый пределы функции, равные пределу функции в точке x_0.

Определение 4. Число A называется *пределом* функции $f(x)$ при $x \to \infty$, если для любой бесконечно большой последовательности (1) значений аргумента соответствующая последовательность (2) значений функции сходится к A.

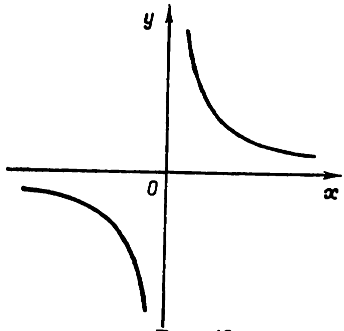


Рис. 42

Символическая запись: $\lim\limits_{x \to \infty} f(x) = A$.

Определение 5. Число A называется *пределом* функции $f(x)$ при $x \to +\infty$ ($x \to -\infty$), если для любой бесконечно большой последовательности (1) значений аргумента, элементы x_n которой положительны (отрицательны), соответствующая последовательность (2) значений функции сходится к A.

Символическая запись: $\lim\limits_{x \to +\infty} f(x) = A$ ($\lim\limits_{x \to -\infty} f(x) = A$).

Рассмотрим пример.

Пусть $f(x) = \dfrac{1}{x}$. Эта функция имеет предел при $x \to \infty$, равный нулю. Действительно, если (1) — бесконечно большая последовательность значений аргумента, то последовательность (2) значений функции $1/x_1$, $1/x_2$, ..., $1/x_n$, ... по теореме 2.1 бесконечно малая и поэтому имеет предел, равный нулю, т. е. $\lim\limits_{x \to \infty} \dfrac{1}{x} = 0$ $(A = 0)$ (рис. 42).

Заметим, что определения 4 и 5 даны «на языке последовательностей». Можно дать равносильные определения «на языке ε—δ». Рекомендуем читателю сделать это самостоятельно.

§ 3. Теоремы о пределах функций

Определение предела функции, данное «на языке последовательностей», дает возможность перенести доказанные выше теоремы о пределах последовательностей на функции. Покажем это на примере двух теорем.

Теорема 4.2. *Пусть функции $f(x)$ и $g(x)$ имеют в точке x_0 пределы B и C. Тогда функции $f(x)\pm g(x)$, $f(x)\cdot g(x)$ и $\frac{f(x)}{g(x)}$ $(C\neq 0)$ имеют в точке x_0 пределы, равные соответственно $B\pm C$, BC и B/C.*

Доказательство. Пусть $\{x_n\}$ $(x_n\neq x_0)$ — произвольная сходящаяся к x_0 последовательность значений аргумента функций $f(x)$ и $g(x)$. Соответствующие последовательности $\{f(x_n)\}$ и $\{g(x_n)\}$ значений этих функций имеют пределы B и C. Но тогда, в силу теорем 2.7—2.9, последовательности $\{f(x_n)\pm g(x_n)\}$, $\{f(x_n)g(x_n)\}$ и $\left\{\frac{f(x_n)}{g(x_n)}\right\}$ имеют пределы, соответственно равные $B\pm\pm C$, BC и B/C. В силу определения 1 предела функции это означает, что

$$\lim_{x\to x_0}[f(x)\pm g(x)]=B\pm C,\ \lim_{x\to x_0}[f(x)g(x)]=BC,$$

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}=B/C \blacksquare$$

Теорема 4.3. *Пусть функции $f(x)$, $g(x)$ и $h(x)$ определены в некоторой окрестности точки x_0, за исключением, может быть, самой точки x_0, и функции $f(x)$, $h(x)$ имеют в точке x_0 предел, равный A, т. е. $\lim\limits_{x\to x_0}f(x)=\lim\limits_{x\to x_0}h(x)=A$. Пусть, кроме этого, выполняются неравенства $f(x)\leqslant g(x)\leqslant h(x)$, тогда*

$$\lim_{x\to x_0}g(x)=A.$$

Доказательство. Пусть $\{x_n\}$ $(x_n\neq x_0)$ — произвольная сходящаяся к x_0 последовательность значений аргумента функций $f(x)$ и $h(x)$. Соответствующие последовательности $\{f(x_n)\}$ и $\{h(x_n)\}$ значений этих функций имеют предел A, т. е. $f(x_n)\to A$, $h(x_n)\to A$ при $n\to\infty$. Используя неравенства, данные в условии теоремы, можем написать:

$$f(x_n)\leqslant g(x_n)\leqslant h(x_n).$$

Отсюда по теореме 2.11 следует, что $g(x_n)\to A$. В силу определения 1 предела функции это означает, что $\lim\limits_{x\to x_0}g(x)=A$ $\blacksquare$

Замечание. Теоремы 4.2 и 4.3 верны также и в случае, когда x_0 является одним из символов ∞, $+\infty$ и $-\infty$.

§ 4. Два замечательных предела

1. $\lim\limits_{x\to 0}\frac{\sin x}{x}=1$ **(первый замечательный предел).** Докажем данное равенство.

Функция $f(x)=\frac{\sin x}{x}$ четная, так как $f(-x)=\frac{\sin(-x)}{-x}=\frac{\sin x}{x}=f(x)$. По условию $x\to 0$, следовательно, достаточно рассмотреть значения x, удовлетворяющие неравенствам $0<x<\frac{\pi}{2}$.

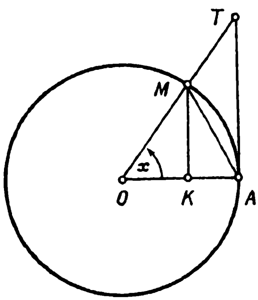


Рис. 43

Возьмем дугу окружности радиуса $R=1$ и угол, радианная мера которого равна x (рис. 43). Тогда

$$OA=1,\quad \sin x=MK,\quad \operatorname{tg} x=AT. \tag{1}$$

Очевидно, что площадь треугольника OAM меньше площади сектора OAM, которая меньше площади треугольника OAT или, что то же самое, $\frac{1}{2}OA\cdot MK<\frac{1}{2}OA\cdot \overset{\frown}{AM}<\frac{1}{2}OA\times$ $\times AT$. Принимая во внимание равенства (1), последнее соотношение можем записать:

$$\frac{1}{2}\sin x<\frac{1}{2}x<\frac{1}{2}\operatorname{tg} x,$$

откуда получаем

$$\sin x<x<\operatorname{tg} x. \tag{2}$$

Разделив эти неравенства на $\sin x$, получим $1>\frac{\sin x}{x}>\cos x$, откуда находим $0<1-\frac{\sin x}{x}<1-\cos x$. Так как $\sin\frac{x}{2}<$

< 1, то $\sin^2\frac{x}{2} < \sin\frac{x}{2}$. Следовательно, с учетом первого неравенства (2) для всех x, удовлетворяющих неравенствам $0 < x < \frac{\pi}{2}$, получим $1-\cos x = 2\sin^2\frac{x}{2} < 2\sin\frac{x}{2} < < 2\frac{x}{2} = x$. Так что $0 < 1 - \frac{\sin x}{x} < x$ при $0 < x < \frac{\pi}{2}$.

Возьмем любое $\varepsilon > 0$ и $\delta = \min\left\{\varepsilon, \frac{\pi}{2}\right\}$. Тогда для всех x, удовлетворяющих неравенствам $0 < x < \delta$, будет выполняться неравенство $x < \varepsilon$, а потому

$$0 < 1 - \frac{\sin x}{x} < \varepsilon, \quad \text{или} \quad \left|1 - \frac{\sin x}{x}\right| < \varepsilon. \qquad (3)$$

А это по второму определению предела функции означает, что 1 есть предел функции $\frac{\sin x}{x}$ в точке $x=0$, т. е. $\lim\limits_{x\to 0}\frac{\sin x}{x} = 1$ ■

С помощью первого замечательного предела вычисляются другие пределы.

Пример 1. Найти $\lim\limits_{x\to 0}\frac{1-\cos x}{x}$.

Решение. Знаменатель дроби при $x\to 0$ стремится к нулю. Поэтому теорема 4.2 здесь неприменима. Для нахождения предела преобразуем данную дробь:

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

$$= \lim_{x\to 0} = \frac{\sin\frac{x}{2}}{\frac{x}{2}}\lim_{x\to 0}\sin\frac{x}{2} = 1\cdot 0 = 0.$$

Пример 2. Найти $\lim\limits_{x\to 0}\frac{\operatorname{tg} x}{x}$.

Решение. Имеем

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

Пример 3. Найти $\lim\limits_{x\to 0}\frac{5x}{\sin 4x}$.

Решение. Имеем

$$\lim_{x\to 0}\frac{5x}{\sin 4x}=\lim_{x\to 0}\frac{\frac{5}{4}}{\frac{\sin 4x}{4x}}=\frac{\frac{5}{4}}{\lim\limits_{x\to 0}\frac{\sin 4x}{4x}}=\frac{\frac{5}{4}}{1}=1{,}25.$$

2. $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e$ **(второй замечательный предел).** Как известно, $\lim\limits_{x\to\infty}\left(1+\frac{1}{n}\right)^n=e$ (см. п. 2, § 3, гл. II). Докажем, что $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e$. Действительно, пусть $x>1$. Положим $n=[x]$, тогда $x=n+\alpha$, где n — натуральное число, а α удовлетворяет условию $0\leqslant\alpha<1$. Так как $n\leqslant x<n+1$, $\frac{1}{n+1}<\frac{1}{x}\leqslant\frac{1}{n}$, то

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

Как нетрудно видеть, при $x\to+\infty$ $(n\to\infty)$

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

и

$$\lim_{n\to\infty}\left(1+\frac{1}{n+1}\right)^n=\frac{\lim\limits_{n\to\infty}\left(1+\frac{1}{n+1}\right)^{n+1}}{\lim\limits_{n\to\infty}\left(1+\frac{1}{n+1}\right)}=\frac{e}{1}=e,$$

откуда по теореме 4.3 получим

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

Пусть теперь $x<-1$, положим $x=-y$. Тогда

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

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

при $x\to-\infty$.

Объединяя оба случая, окончательно получим

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

Второй замечательный предел имеет широкое применение. С его помощью находятся многие другие пределы.

Пример 4. Найти $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}}$.

Решение. Сделаем замену переменной, положив $1/x=\alpha$. Тогда очевидно, что $\alpha\to\infty$ при $x\to 0$. Поэтому

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

Пример 5. Найти $\lim\limits_{x\to\infty}\left(1+\frac{3}{x}\right)^x$

Решение. Положим $x=3t$. Тогда при $x\to\infty$ и $t\to\infty$. Следовательно,

$$\lim_{x\to\infty}\left(1+\frac{3}{x}\right)^x=\lim_{t\to\infty}\left(1+\frac{1}{t}\right)^{3t}=$$

$$=\lim_{x\to\infty}\left[\left(1+\frac{1}{t}\right)^t\left(1+\frac{1}{t}\right)^t\left(1+\frac{1}{t}\right)^t\right]=$$

$$=\lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t\lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t\lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t=e\cdot e\cdot e=e^2.$$

Пример 6. Найти $\lim\limits_{x\to 0}\frac{\log_a(1+x)}{x}$

Решение. Для нахождения предела преобразуем данную дробь:

$$\lim_{x\to 0}\frac{\log_a(1+x)}{x}=\lim_{x\to 0}\left[\frac{1}{x}\log_a(1+x)\right]=$$

$$=\lim_{x\to 0}\log_a(1+x)^{\frac{1}{x}}=\log_a\left[\lim_{x\to 0}(1+x)^{\frac{1}{x}}\right].$$

Но $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}}=e$ (см. пример 4). Поэтому

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

В частности, при $a=e \lim\limits_{x\to 0}\frac{\ln(1+x)}{x}=1.$

§ 5. Бесконечно малые и бесконечно большие функции

Определение 1. Функция $f(x)$ называется *бесконечно малой функцией* (или *просто бесконечно малой*) в точке $x=x_0$ (или при $x \to x_0$), если $\lim\limits_{x\to x_0} f(x)=0$.

Аналогично определяются бесконечно малые функции при $x \to \infty$, $x \to +\infty$, $x \to -\infty$, $x \to x_0-$ и $x \to x_0+$.

Так же как бесконечно малые последовательности, бесконечно малые функции играют существенную роль в том, что общее понятие предела функции может быть сведено к понятию бесконечно малой. Имеет место следующая

Теорема 4.4. *Для выполнения равенства* $\lim\limits_{x\to x_0} f(x)=A$ *необходимо и достаточно, чтобы функция*

$$\alpha(x)=f(x)-A$$

была бесконечно малой при $x \to x_0$.

Доказательство. *Необходимость.* Пусть $\lim\limits_{x\to x_0} f(x)=A$. Рассмотрим разность $f(x)-A=\alpha(x)$ и покажем, что $\alpha(x)$ — бесконечно малая функция при $x \to x_0$. Действительно, пределы каждой из функций $f(x)$ и A при $x \to x_0$ равны A, и поэтому в силу теоремы 4.2

$$\lim_{x\to x_0}\alpha(x)=\lim_{x\to x_0}[f(x)-A]=\lim_{x\to x_0} f(x)-\lim_{x\to x_0} A=0.$$

Достаточность. Пусть $f(x)-A=\alpha(x)$, где $\alpha(x)$ — бесконечно малая функция при $x \to x_0$. Покажем, что $\lim\limits_{x\to x_0} f(x)=A$. Так как $f(x)=A+\alpha(x)$, то

$$\lim_{x\to x_0} f(x)=\lim_{x\to x_0}[A+\alpha(x)]=\lim_{x\to x_0} A+\lim_{x\to x_0}\alpha(x)=A+0=A \blacksquare$$

Из теоремы 4.4 мы получаем специальное представление для функции, имеющей в точке $x=x_0$ предел, равный A:

$$f(x)=A+\alpha(x), \quad \text{где } \lim_{x\to x_0}\alpha(x)=0. \qquad (1)$$

При этом обычно говорят, что функция $f(x)$ в окрестности точки x_0 отличается от A на бесконечно малую функцию.

Бесконечно малые функции обладают теми же свойствами, что и бесконечно малые последовательности. Справедлива

Теорема 4.5. *Алгебраическая сумма и произведение конечного числа бесконечно малых функций при* $x \to x_0$, *а также произведение бесконечно малой функции на ограниченную функцию являются бесконечно малыми функциями при* $x \to x_0$.

Эта теорема непосредственно вытекает из первого определения предела функции и теорем 2.2—2.4.

Все сказанное о бесконечно малых функциях при $x \to x_0$ справедливо и для бесконечно малых функций при $x \to \infty$, $x \to +\infty$, $x \to -\infty$, $x \to x_0-$ и $x \to x_0+$.

Определение 2. Функция $f(x)$ называется *бесконечно большой функцией* (или *просто бесконечно большой*) в точке $x = x_0$ (или при $x \to x_0$), если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех $x \in X$, $x \neq x_0$, удовлетворяющих неравенству $|x - x_0| < \delta$, выполняется неравенство $|f(x)| > \varepsilon$.

В этом случае пишут $\lim\limits_{x \to x_0} f(x) = \infty$.

Если же выполняется неравенство $f(x) > \varepsilon$ ($f(x) < -\varepsilon$), то пишут $\lim\limits_{x \to x_0} f(x) = +\infty$ ($\lim\limits_{x \to x_0} f(x) = -\infty$).

По аналогии с конечными односторонними пределами определяются и односторонние бесконечные пределы:

$$\lim_{x \to x_0+} f(x) = +\infty, \quad \lim_{x \to x_0+} f(x) = -\infty;$$

$$\lim_{x \to x_0-} f(x) = +\infty, \quad \lim_{x \to x_0-} f(x) = -\infty.$$

Так, например, говорят, что $\lim\limits_{x \to x_0+} f(x) = +\infty$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех $x \in X$, $x \neq x_0$, удовлетворяющих неравенствам $x_0 < x < x_0 + \delta$, выполняется неравенство $f(x) > \varepsilon$.

То же самое «на языке последовательностей»: $\lim\limits_{x \to x_0+} f(x) = +\infty$, если для любой сходящейся к x_0 последовательности $\{x_n\}$ значений аргумента x, элементы x_n которой больше x_0, соответствующая последовательность $\{f(x_n)\}$ значений функций является бесконечно большой положительного знака.

Точное определение других подобных пределов рекомендуем читателю сделать самостоятельно.

Аналогично определяются бесконечно большие функции при $x \to \infty$, $x \to +\infty$ и $x \to -\infty$. Так, например, функция $f(x)$ называется бесконечно большой при $x \to \infty$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех x, удовлетворяющих неравенству $|x| > \delta$, выполняется неравенство $|f(x)| > \varepsilon$ и пишут $\lim\limits_{x \to \infty} f(x) = \infty$.

Если же выполняется неравенство $f(x) > \varepsilon$ ($f(x) < -\varepsilon$), то пишут $\lim\limits_{x \to \infty} f(x) = +\infty$ ($\lim\limits_{x \to \infty} f(x) = -\infty$).

Предполагается самостоятельно сформулировать определение бесконечно большой функции при $x \to +\infty$ и $x \to -\infty$.

В заключение заметим, что между бесконечно малыми и бесконечно большими функциями существует аналогичная связь, как и между соответствующими последовательностями, т. е. функция, обратная бесконечно малой, является бесконечно большой и наоборот.

§ 6. Сравнение бесконечно малых и бесконечно больших функций

Мы уже знаем, что сумма, разность и произведение бесконечно малых функций являются бесконечно малыми функциями. Этого, вообще говоря, нельзя сказать о частном: деление одной бесконечно малой на другую может привести к различным результатам. Так, например, если $\alpha(x) = x$ и $\beta(x) = 2x$, то

$$\lim_{x \to 0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to 0} \frac{x}{2x} = \frac{1}{2}.$$

Если же $\alpha(x) = x$, $\beta(x) = x^2$, то

$$\lim_{x \to 0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to 0} \frac{1}{x} = \infty, \quad \lim_{x \to 0} \frac{\beta(x)}{\alpha(x)} = \lim_{x \to 0} x = 0.$$

Существуют правила сравнения бесконечно малых функций.

Пусть при $x \to x_0$ функции $\alpha(x)$ и $\beta(x)$ являются бесконечно малыми. Тогда:

1) если $\lim\limits_{x \to x_0} \dfrac{\alpha(x)}{\beta(x)} = 0$, то $\alpha(x)$ называется беско-

нечно малой более высокого порядка, чем $\beta(x)$ (имеет более высокий порядок малости);

2) если $\lim\limits_{x\to x_0} \frac{\alpha(x)}{\beta(x)} = A \neq 0$ (A — число), то $\alpha(x)$ и $\beta(x)$ называются бесконечно малыми одного порядка (имеют «одинаковую скорость» стремления к нулю);

3) если $\lim\limits_{x\to x_0} \frac{\alpha(x)}{\beta(x)} = 1$, то $\alpha(x)$ и $\beta(x)$ называются эквивалентными бесконечно малыми (эквивалентность записывается так: $\alpha(x) \sim \beta(x)$).

В отдельных случаях оказывается недостаточно знать, что одна из двух бесконечно малых является бесконечно малой более высокого порядка. Нужно еще оценить, как высок этот порядок. Поэтому вводится следующее правило:

4) если $\lim\limits_{x\to x_0} \frac{\alpha(x)}{\beta^n(x)} = A \neq 0$, то $\alpha(x)$ называется бесконечно малой n-го порядка относительно $\beta(x)$.

Рассмотрим примеры.

1) Функции $\sin x$ и x являются при $x\to 0$ эквивалентными бесконечно малыми, так как $\lim\limits_{x\to 0} \frac{\sin x}{x} = 1$.

2) Функции $\sin 3x$ и $\sin x$ являются при $x\to 0$ бесконечно малыми одного порядка, так как

$$\lim_{x\to 0}\frac{\sin 3x}{\sin x} = \lim_{x\to 0}\frac{\frac{3\sin 3x}{3x}}{\frac{\sin x}{x}} = 3\lim_{x\to 0}\frac{\sin 3x}{3x}\lim_{x\to 0}\frac{x}{\sin x} = 3.$$

3) Функция $\alpha(x) = 1 - \cos x$ является при $x\to 0$ бесконечно малой второго порядка малости по отношению к бесконечно малой x, так как

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

При сравнении бесконечно малых функций часто употребляют символ o («о» малое). Если функция $\alpha(x)$ — бесконечно малая в точке x_0 более высокого порядка, чем бесконечно малая в этой же точке $\beta(x)$, то это условно записывается так:

$$\alpha(x)=o(\beta(x)).$$

Заметим также, что если функции $\alpha(x)$ и $\beta(x)$ — бесконечно малые в точке x_0, то функция $\alpha(x)\cdot\beta(x)$ имеет более высокий порядок малости, чем каждый из сомножителей. В самом деле,

$$\lim_{x\to x_0}\frac{\alpha(x)\beta(x)}{\beta(x)}=\lim_{x\to x_0}\alpha(x)=0$$

и поэтому $\alpha(x)\beta(x)=o(\beta(x))$, $\alpha(x)\beta(x)=o(\alpha(x))$.

Для бесконечно больших функций имеют место аналогичные правила сравнения.

Рассмотрим несколько примеров.

1) Функции $\alpha(x)=\dfrac{1+x}{x}$ и $\beta(x)=\dfrac{1}{x}$ являются при $x\to 0$ эквивалентными бесконечно большими, так как

$$\lim_{x\to 0}\frac{\alpha(x)}{\beta(x)}=\lim_{x\to 0}(1+x)=1,$$

а при $x\to\infty$ функция $\alpha(x)$ является бесконечно большой более высокого порядка, чем $\beta(x)$ (имеет более высокий порядок роста), так как $\lim\limits_{x\to\infty}\dfrac{\alpha(x)}{\beta(x)}=\infty$.

2) Функция $\alpha(x)=x^2+4$ является при $x\to\infty$ бесконечно большой более низкого порядка, чем $\beta(x)=x^3-2$ (имеет менее высокий порядок роста), так как

$$\lim_{x\to\infty}\frac{\alpha(x)}{\beta(x)}=\lim_{x\to\infty}\frac{x^2+4}{x^3-2}=\lim_{x\to\infty}\frac{1+\dfrac{4}{x^2}}{x-\dfrac{2}{x^2}}=\lim_{x\to\infty}\frac{1}{x}=0.$$

3) Бесконечно большие функции $\alpha(x)=2x^2+1$ и $\beta(x)=$ $=x^2-1$ имеют при $x\to\infty$ одинаковый порядок роста, так как

$$\lim_{x\to\infty}\frac{2x^2+1}{x^2-1}=\lim_{x\to\infty}\frac{2+\dfrac{1}{x^2}}{1-\dfrac{1}{x^2}}=2.$$

4) Функция $\alpha(x)=x^4+x+1$ является при $x\to\infty$ бесконечно большой второго порядка по отношению к беско-

нечно большой $\beta(x)=x^2+1$, так как

$$\lim_{x\to\infty}\frac{x^4+x+1}{(x^2+1)^2}=\lim_{x\to\infty}\frac{x^4+x+1}{x^4+2x^2+1}=\lim_{x\to\infty}\frac{1+\frac{1}{x^3}+\frac{1}{x^4}}{1+\frac{2}{x^2}+\frac{1}{x^4}}=1.$$

§ 7. Понятие непрерывности функции

Понятие непрерывности функции является одним из основных понятий математического анализа, так как непрерывные функции обладают целым рядом важных свойств, которых лишены функции разрывные. Непрерывные функции составляют основной для анализа класс функций.

1. Определение непрерывности функции. Пусть на некотором промежутке X определена функция $f(x)$ и точка x_0 принадлежит этому промежутку *.

Определение 1. Функция $f(x)$ называется *непрерывной* в точке x_0, если для любой последовательности значений аргумента x: $x_1, x_2, x_3, \ldots, x_n, \ldots$, сходящейся к x_0, последовательность соответствующих значений функции: $f(x_1), f(x_2), f(x_3), \ldots, f(x_n), \ldots$ сходятся к $f(x_0)$.

Другими словами, функция $f(x)$ непрерывна в точке x_0, если предел функции в этой точке и ее значение равны, т. е.

$$\lim_{x\to x_0} f(x)=f(x_0). \tag{1}$$

Так как $\lim\limits_{x\to x_0} x=x_0$, то соотношению (1) можно придать следующую форму:

$$\lim_{x\to x_0} f(x)=f(\lim_{x\to x_0} x),$$

т. е. для непрерывной функции знаки функции и предела можно переставлять.

По аналогии с определением предела функции можно сформулировать определение непрерывности функции и «на языке ε—δ».

* Заметим, что этого не требовалось, когда мы рассматривали предел функции $f(x)$ в точке x_0. В этом заключено отличие понятия непрерывности функции от понятия ее предела!

Определение 2. Функция $f(x)$ называется *непрерывной* в точке x_0, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что для всех $x \in X$, удовлетворяющих неравенству $|x-x_0|<\delta$, выполняется неравенство $|f(x)-f(x_0)|<\varepsilon$.

Эквивалентность этих определений очевидна.

Если $\lim\limits_{x\to x_0+} f(x)=f(x_0)$ $\left(\lim\limits_{x\to x_0-} f(x)=f(x_0)\right)$, то функцию $f(x)$ называют *непрерывной* в точке x_0 *справа* (*слева*). Если функция $f(x)$ непрерывна в точке x_0 и слева и справа, то она непрерывна в этой точке. В самом деле,

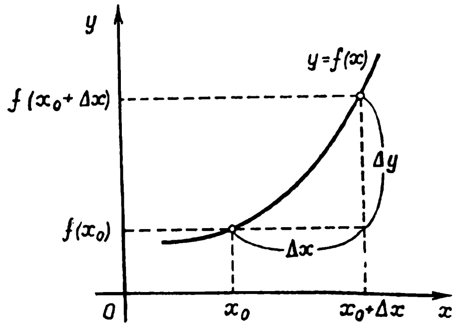


Рис. 44

в силу замечания § 2 настоящей главы, в этом случае предел функции в точке x_0 равен ее значению в этой точке.

Наконец, дадим еще одно определение непрерывности функции, которое по существу является перефразировкой первого определения.

Перенося в равенстве (1) $f(x_0)$ в левую часть, внося $f(x_0)$ под знак предела и замечая, что условия $x\to x_0$ и $x-x_0\to 0$ равносильны, получим

$$\lim_{x-x_0\to 0}[f(x)-f(x_0)]=0. \tag{2}$$

Разность $x-x_0$ называется *приращением аргумента* x и обозначается обычно Δx, а разность $f(x)-f(x_0)$ — *приращением функции* в точке x_0, вызванным приращением аргумента Δx, и обозначается Δy. Таким образом,

$$\Delta x=x-x_0, \qquad \Delta y=f(x_0+\Delta x)-f(x_0),$$

геометрический смысл которых очевиден (рис. 44). Равенство (2) в новых обозначениях примет вид

$$\lim_{\Delta x \to 0} \Delta y = 0. \qquad (3)$$

Соотношение (3) и является еще одним определением непрерывности функции, которое словами можно сформулировать так:

Определение 3. Функция $f(x)$ называется *непрерывной* в точке x_0, если ее приращение в этой точке является бесконечно малой функцией при $\Delta x \to 0$.

Последнее определение для практического использования бывает иногда более удобным, и им мы будем также пользоваться.

2. Арифметические действия над непрерывными функциями

Теорема 4.6. *Пусть функции $f(x)$ и $g(x)$ непрерывны в точке x_0. Тогда функции $f(x) \pm g(x)$, $f(x) \cdot g(x)$ и $f(x)/g(x)$ также непрерывны в этой точке (последняя при $g(x_0) \neq 0$).*

Доказательство. Так как непрерывные в точке x_0 функции $f(x)$ и $g(x)$ имеют в этой точке пределы, равные $f(x_0)$ и $g(x_0)$, то по теореме 4.2 пределы функций $f(x) \pm g(x)$, $f(x)g(x)$ и $f(x)/g(x)$ существуют и равны соответственно $f(x_0) \pm g(x_0)$, $f(x_0)g(x_0)$, $f(x_0)/g(x_0)$. Но эти величины как раз и равны значениям соответствующих функций в точке x_0. Следовательно, по определению 1 функции $f(x) \pm g(x)$, $f(x)g(x)$ и $f(x)/g(x)$ непрерывны в точке x_0 ■

§ 8. Непрерывность некоторых элементарных функций

Одним из важных свойств элементарных функций является их непрерывность в каждой точке области их определения. На примере некоторых функций мы и проверим данный факт, пользуясь непосредственно определениями непрерывности функций в точке.

1. Непрерывность рациональных функций. Простейшим примером функции, непрерывной в любой точке x_0 числовой прямой, может служить постоянная функция $f(x) = C$. В самом деле, в этом случае $\lim_{x \to x_0} f(x) = C$

(см. пример 1, § 2), т. е. постоянная функция непрерывна в каждой точке прямой.

Непрерывна также в каждой точке x_0 числовой прямой функция $f(x)=x$, так как $\lim\limits_{x \to x_0} x = x_0$ (см. пример 2, § 2), т. е. предел функции в точке x_0 равен ее значению в этой точке.

Из сказанного и теоремы 4.6 следует, что в любой точке x_0 функции

$$x^2=x \cdot x, \quad x^3=x^2 \cdot x, \quad x^4=x^3 \cdot x, \; ..., \; x^n=x^{n-1} \cdot x$$

($n>0$ — целое число) непрерывны. Как мы знаем, функция $f(x)=x^n$ называется степенной функцией, а функция вида

$$P(x) = C_0x^n + C_1x^{n-1} + C_2x^{n-2} + \ldots + C_{n-1}x + C_n,$$

где $n \geqslant 0$ — целое число, $C_0, C_1, C_2, ..., C_n$ — постоянные числа, называется алгебраическим многочленом.

Каждое из слагаемых

$$C_0x^n, \; C_1x^{n-1}, \; C_2x^{n-2}, \ldots, C_n$$

есть произведение двух непрерывных функций (постоянной и степенной). По теореме 4.6 оно непрерывно в любой точке x. Многочлен $P(x)$ есть, таким образом, сумма функций, непрерывных в любой точке x, и, следовательно, он непрерывен в любой точке x.

Дробно-рациональная функция, т. е. функция, имеющая вид

$$R(x) = \frac{P(x)}{Q(x)},$$

где $P(x)$ и $Q(x)$ — алгебраические многочлены, непрерывна во всех таких точках x, в которых ее знаменатель не равен нулю, т. е. во всех точках, кроме корней знаменателя, как частное непрерывных функций.

Например, функция $R(x) = \dfrac{3x^2+7x-1}{x^2-1}$ непрерывна во всех точках x, отличных от $+1$ и -1.

2. Непрерывность тригонометрических функций. Рассмотрим тригонометрические функции:

$$\sin x, \quad \cos x, \quad \operatorname{tg} x, \quad \operatorname{ctg} x, \quad \sec x, \quad \operatorname{cosec} x.$$

Покажем, что функция $\sin x$ непрерывна в любой точке x. Воспользуемся определением 3 непрерывности функции. Давая аргументу x приращение Δx, получим приращение функции

$$\Delta y = \sin(x+\Delta x) - \sin x,$$

или

$$\Delta y = 2\cos\left(x+\frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}.$$

Переходя к пределу в левой и правой частях равенства при $\Delta x \to 0$, получаем

$$\lim_{\Delta x\to 0}\Delta y = 2\lim_{\Delta x\to 0}\left[\cos\left(x+\frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}\right] = 0,$$

так как $\left|\cos\left(x+\frac{\Delta x}{2}\right)\right| \leqslant 1$, а

$$\lim_{\Delta x\to 0}\sin\frac{\Delta x}{2} = \lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\cdot\frac{\Delta x}{2} = \frac{1}{2}\lim_{\Delta x\to 0} 1\cdot\Delta x = 0$$

(произведение ограниченной величины на бесконечно малую есть бесконечно малая).

Таким образом, функция $\sin x$ непрерывна в любой точке x.

Непрерывность функции $\cos x$ в любой точке x доказывается совершенно аналогично.

Из непрерывности функций $\sin x$ и $\cos x$ по теореме 4.6 о непрерывности частного следует непрерывность функций $\operatorname{tg} x = \frac{\sin x}{\cos x}$ и $\sec x = \frac{1}{\cos x}$ во всех точках, где $\cos x \neq 0$, т. е. кроме точек $x = \frac{\pi}{2} + n\pi$, и функций $\operatorname{ctg} x = \frac{\cos x}{\sin x}$ и $\operatorname{cosec} x = \frac{1}{\sin x}$ во всех точках, кроме $x = n\pi$ $(n = 0, \pm 1, \pm 2, \ldots)$.

3. Непрерывность функции $f(x) = |x|$. Функция $f(x) = |x|$, график которой изображен на рис. 45, определена и непрерывна на всей числовой прямой. В самом деле, на интервале $(0, +\infty)$ она непрерывна, так как при $x > 0$ $f(x) = x$ (см. п. 1). На интервале $(-\infty, 0)$ $f(x)$ также непрерывна, так как при $x < 0$ она принимает вид $f(x)=-x$ и ее можно представить как произведение двух непрерывных функций (-1) и x и применить теорему 4.6 о непрерывности произведения. Чтобы установить непрерывность функции $|x|$ в точке $x=0$, вычислим односторонние пределы функции в этой точке:

$$\lim_{x\to 0-}|x| = \lim_{x\to 0-}(-x) = -\lim_{x\to 0-} x = 0; \quad \lim_{x\to 0+}|x| = \lim_{x\to 0+} x = 0,$$

т. е. пределы функции слева и справа совпадают и равны значению функции в точке $x=0$. А это означает, что функция непрерывна и в точке $x=0$, т. е. на всей прямой.

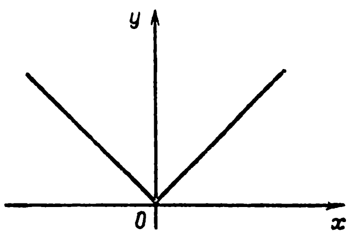


Рис. 45

Таким образом, мы убедились, что рассмотренные функции непрерывны в области их определения, а на основании теоремы 4.6 о непрерывности суммы, разности, произведения и частного можно утверждать, что функции, получаемые из них путем конечного числа арифметических действий, являются также непрерывными функциями в областях их определения.

Про функцию $f(x)$ будем говорить, что она непрерывна в интервале (a, b), если она непрерывна в каждой точке этого интервала; непрерывна на отрезке $[a, b]$, если она непрерывна в интервале (a, b) и непрерывна в точке a справа, а в точке b слева, т. е.

$$\lim_{x\to a+} f(x) = f(a), \text{ а } \lim_{x\to b-} f(x) = f(b).$$

§ 9. Классификация точек разрыва функции

1. Определение и классификация точек разрыва функции

Определение. Точка x_0 называется *точкой разрыва* функции $f(x)$, если $f(x)$ в точке x_0 не является непрерывной.

Разрывы функций можно классифицировать следующим образом.

1°. Разрыв 1-го рода. Точка x_0 называется точкой разрыва 1-го рода, если в этой точке функция $f(x)$ имеет конечные, но не равные друг другу правый и левый пределы:

$$\lim_{x \to x_0+} f(x) \neq \lim_{x \to x_0-} f(x).$$

Пример. Для функции $f(x) = \operatorname{sgn} x$ точка $x=0$ является точкой разрыва 1-го рода (см. рис. 41), так как

$$\lim_{x \to 0+} \operatorname{sgn} x = 1, \quad \lim_{x \to 0-} \operatorname{sgn} x = -1.$$

2°. Разрыв 2-го рода. Точка x_0 называется точкой разрыва 2-го рода, если в этой точке функция $f(x)$ не имеет по крайней мере одного из односторонних пределов или если хотя бы один из односторонних пределов бесконечен.

Пример. Для функции $f(x) = \frac{1}{x}$ точка $x=0$ является точкой разрыва 2-го рода (см. рис. 42), так как

$$\lim_{x \to 0+} \frac{1}{x} = +\infty, \quad \lim_{x \to 0-} \frac{1}{x} = -\infty.$$

2. Кусочно-непрерывные функции. Функция $f(x)$ называется *кусочно-непрерывной* на отрезке $[a, b]$, если она непрерывна во всех внутренних точках $[a, b]$, за исключением, быть может, конечного числа точек, в которых имеет разрыв 1-го рода и, кроме того, имеет односторонние пределы в точках a и b.

Функция называется кусочно-непрерывной на интервале или бесконечной прямой, если она кусочно-непрерывна на любом принадлежащем им отрезке.

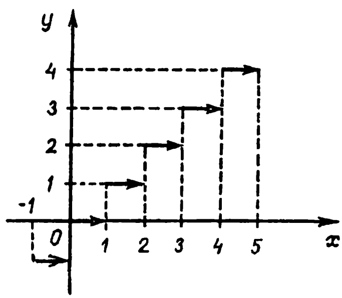


Рис. 46

Пример. Функция $f(x) = [x]$ кусочно-непрерывна как на любом отрезке, так и на бесконечной прямой. Напомним, что символ $[x]$ обозначает целую часть числа x. График функции $f(x) = [x]$ изображен на рис. 46, функция $[x]$ в точках $x = n$ ($n = 0, \pm 1, \pm 2, \ldots$) непрерывна справа и разрывна слева. Во всех же других точках она непрерывна как справа, так и слева.

§ 10. Основные свойства непрерывных функций

1. Теорема об устойчивости знака непрерывной функции

Теорема 4.7. *Пусть функция $f(x)$ определена в некоторой точке x_0, непрерывна в точке x_0 и $f(x_0) \neq 0$. Тогда существует $\delta > 0$ такое, что для всех $x \in (x_0 - \delta, x_0 + \delta)$ $f(x)$ имеет тот же знак, что $f(x_0)$.*

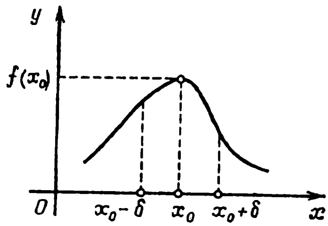


Рис. 47

Доказательство. Пусть $f(x_0) > 0$ (рис. 47). Тогда в силу второго определения непрерывности функции для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $|f(x) - f(x_0)| < \varepsilon$ выполняется для всех x, удовлетворяющих условию $|x - x_0| < \delta$, или, что то же самое, выполняются неравенства

$$f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon \tag{1}$$

для всех $x \in (x_0 - \delta, x_0 + \delta)$. Возьмем $\varepsilon = f(x_0)$. Тогда из левого неравенства (1) получим: $f(x) > 0$ для всех $x \in (x_0 - \delta, x_0 + \delta)$, что и требовалось.

Если же $f(x_0) < 0$, то рассмотрим функцию $-f(x)$. Так как $-f(x_0) > 0$, то по доказанному существует δ-окрестность точки x_0, в которой $-f(x) > 0$, и, следовательно, $f(x) < 0$ ■

2. Прохождение непрерывной функции через любое промежуточное значение

1°. Теорема о прохождении непрерывной функции через нулевое значение при смене знаков

Теорема 4.8 (первая теорема Больцано — Коши*). *Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и на концах имеет значения разных знаков. Тогда существует точка $c \in (a, b)$*, в которой $f(c) = 0$.

Доказательство. Пусть для определенности $f(a) < 0$ и $f(b) > 0$ (рис. 48). Разделим $[a, b]$ пополам. Если значение функции в середине отрезка $[a, b]$ рав-

---

* Больцано Бернард (1781—1848) — чешский математик.

но нулю, то теорема доказана. В противном случае выберем тот из полученных отрезков, на концах которого функция имеет разные знаки: $[a_1, b_1]$. Разделим $[a_1, b_1]$ пополам и выберем тот отрезок, на концах которого $f(x)$ имеет разные знаки: $[a_2, b_2]$ и т. д. Продолжая этот процесс неограниченно, получим последовательность

$$[a, b]\supset[a_1, b_1]\supset[a_2, b_2]\supset\ldots\supset[a_n, b_n]\supset\ldots$$

вложенных отрезков, причем $b_n-a_n=b-a/2^n\to 0$ при $n\to\infty$.

По теореме 2.13 о вложенных отрезках существует точка c, принадлежащая всем отрезкам. Докажем, что $f(c)=0$.

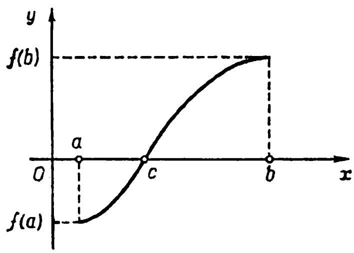


Рис. 48

Действительно, если допустить, что $f(c)>0$, то по теореме 4.7 об устойчивости знака существует окрестность точки c, в которой $f(x)>0$. В эту окрестность при достаточно большом n попадает отрезок $[a_n, b_n]$, в котором, следовательно, будет $f(x)>0$, а это противоречит выбору последовательности вложенных отрезков. Аналогично доказывается, что $f(c)$, не может быть меньше нуля. Остается принять, что $f(c)=0$, причем точка $c\in(a, b)$■.

Доказанная теорема имеет простой геометрический смысл: непрерывная кривая при переходе с одной полуплоскости по отношению к оси на другую пересекает эту ось.

2°. Теорема о прохождении непрерывной функции через любое промежуточное значение

Теорема 4.9 (вторая теорема Больцано — Коши). *Пусть $f(x)$ непрерывна на отрезке $[a, b]$, причем $f(a)=A$, $f(b)=B$. Пусть, далее, C — любое число, заключенное между A и B. Тогда на $[a, b]$ найдется точка c такая, что $f(c)=C$.*

Другими словами, непрерывная функция при пере-

ходе от одного значения к другому принимает и все промежуточные значения.

Доказательство. Пусть для определенности $A<B$ и $A<C<B$ (рис. 49). Рассмотрим вспомогательную функцию

$$\varphi(x)=f(x)-C.$$

Эта функция непрерывна на отрезке $[a, b]$ (как разность непрерывных функций) и принимает на концах этого отрезка значения разных знаков

$$\varphi(a)=f(a)-C=A-C<0,$$
$$\varphi(b)=f(b)-C=B-C>0.$$

По теореме 4.8 существует точка $c\in(a, b)$, такая, что $\varphi(c)=f(c)-C=0$, т. е. $f(c)-C=0$. Отсюда $f(c)=C$ ■

Следствие. *Если функция $f(x)$ определена и непрерывна на некотором промежутке X, то множество ее значений Y тоже представляет некоторый промежуток.*

Доказательство. Пусть $m=\inf\limits_X Y$, $M=\sup\limits_X Y$, где m и M — числа, называемые соответственно точной нижней и верхней гранями функции (при этом не исключено, что $m=-\infty$ и $M=+\infty$ *). Возьмем любое y из Y и выберем два значения y_1 и y_2 функции $f(x)$, так, чтобы

$$m\leqslant y_1<y<y_2\leqslant M.$$

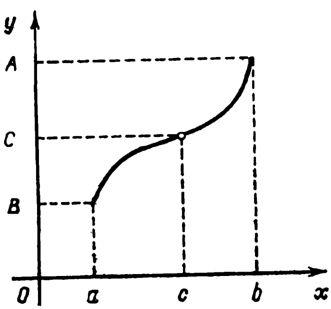


Рис. 49

Существование таких значений функции $f(x)$ следует из определения точных граней. Тогда по теореме 4.9 о промежуточных значениях непрерывной функции существует точка x такая, что $f(x)=y$. Следовательно, множество Y представляет собой некоторый промежуток (конечный или бесконечный) с концами m и M, которые в зависимости от конкретного случая могут ему принадлежать или нет ■

---

* Напомним, что если Y не ограничено сверху (снизу), то мы условились в § 3, гл. I полагать $\sup\limits_X Y=+\infty$ ($\inf\limits_X Y=-\infty$).

3. Теорема об ограниченности непрерывной функции на отрезке. Дадим следующее определение.

Определение. Функция $f(x)$ называется *ограниченной* на отрезке $[a, b]$, если существует число $M>0$ такое, что для всех $x\in[a, b]$ выполняется неравенство

$$|f(x)|\leqslant M, \quad \text{или} \quad -M\leqslant f(x)\leqslant M,$$

т. е. график $f(x)$ не выходит из полосы, ограниченной прямыми $y=M$ и $y=-M$ (рис. 50).

Теорема 4.10 (первая теорема Вейерштрасса*). *Если функция $f(x)$ определена и непрерывна на отрезке $[a, b]$, то она ограничена на этом отрезке.*

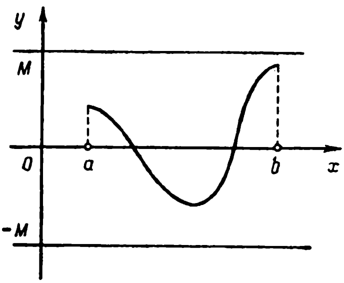


Рис. 50

Для доказательства теоремы докажем сначала лемму.

Лемма. *Функция $f(x)$, непрерывная в точке x_0, ограничена в некоторой ее окрестности.*

Доказательство. Возьмем $\varepsilon=1$. Тогда по второму определению непрерывности функции в точке для данного ε существует $\delta>0$ такое, что для всех $x\in(x_0-\delta, x_0+\delta)$ будет выполняться $|f(x)-f(x_0)|<1$. Так как

$$|f(x)|=|(f(x)-f(x_0))+f(x_0)|\leqslant$$
$$\leqslant|f(x)-f(x_0)|+|f(x_0)|<1+|f(x_0)|,$$

то $|f(x)|<M$, где $M=1+|f(x_0)|$. Отсюда заключаем, что функция $f(x)$ ограничена в δ-окрестности точки x_0 ■

Теперь перейдем к доказательству теоремы.

Доказательство теоремы. Предположим обратное, т. е. допустим, что функция $f(x)$ не ограничена на отрезке $[a, b]$. Разделим отрезок $[a, b]$ пополам, тогда, по крайней мере, в одном полученном отрезке функция $f(x)$ не ограничена (в противном случае она была бы ограничена на $[a, b]$). Обозначим его через $[a_1, b_1]$. Разделим $[a_1, b_1]$ пополам и обозначим через $[a_2, b_2]$ тот отрезок, на котором $f(x)$ не ограничена и т. д. Продолжая этот процесс неограниченно, получим последовательность

* Вейерштрасс Карл (1815—1897) — немецкий математик.

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \supset \ldots$$

вложенных отрезков, на каждом из которых $f(x)$ не ограничена.

По теореме 2.13 о вложенных отрезках существует точка c, принадлежащая всем отрезкам. Функция $f(x)$ по условию определена и непрерывна в точке c, а следовательно, по доказанной лемме в некоторой окрестности точки c она ограничена. При достаточно большом n в эту окрестность попадает отрезок $[a_n, b_n]$, в котором $f(x)$ будет также ограничена, что противоречит выбору последовательности вложенных отрезков. Полученное противоречие доказывает теорему ■

Замечание. Теорема становится неверной, если в ней отрезок $[a, b]$ заменить интервалом (a, b). Так, например, функция $f(x) = \frac{1}{x}$ непрерывна на $(0, 1)$, но не ограничена, так как $\lim\limits_{x \to 0+} \frac{1}{x} = +\infty$. Доказательство не проходит в том месте, где мы утверждаем, что в точке x_0 функция непрерывна. Для интервала x_0 может совпасть с его концом и $f(x_0)$ будет не определена.

4. Теорема о достижении функцией, непрерывной на отрезке, своих точных граней. В том случае, когда точные грани функции являются значениями функции, то говорят, что функция достигает своих точных граней. Однако, как известно (см. (1), теорема 1.1, гл. I), не всякому множеству принадлежат его точные грани. Следующий пример показывает, что точные грани функции не всегда являются *достижимыми.*

Пусть на отрезке $[0, b]$, $b \geqslant 1$, определена функция $f(x) = x - [x]$, график которой изображен на рис. 51. Множество ее значений $[0, 1)$. Функция ограничена и сверху и снизу, имеет на данном отрезке точную верхнюю грань, равную 1, и точную нижнюю грань, рав-

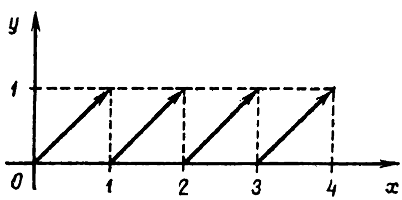


Рис. 51

ную 0. Однако функция не принимает значения, равного точной верхней грани. Следовательно, можно сказать, что функция достигает своей точной нижней и не достигает своей точной верхней грани.

Возникает вопрос, при каком условии функция достигает своих точных граней. Ответ дает следующая теорема.

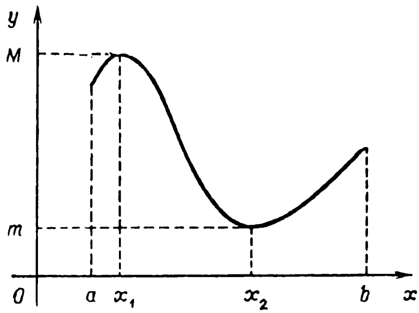


Рис. 52

Теорема 4.11 (вторая теорема Вейерштрасса). *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она достигает на этом отрезке своих точных граней, т. е. существуют точки x_1, $x_2 \in [a, b]$ такие, что*

$$f(x_1) = M = \sup_{[a,\, b]} Y,\quad f(x_2) = m = \inf_{[a,b]} Y \text{ (рис. 52).}$$

Доказательство. Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то по теореме 4.10 она будет ограниченной. Следовательно, согласно теореме 1.1 существуют точная верхняя M и точная нижняя m грани множества значений функции Y.

Покажем, что $f(x)$ достигает M, т. е. существует такая точка $x_1 \in [a, b]$, что $f(x_1) = M$. Будем рассуждать от противного. Пусть функция $f(x)$ не принимает ни в одной точке $[a, b]$ значения, равного M. Тогда для всех $x \in [a, b]$ справедливо неравенство $f(x) < M$. Рассмотрим на $[a, b]$ вспомогательную, всюду положительную функцию

$$F(x) = \frac{1}{M - f(x)}.$$

По теореме 4.6 о непрерывности частного функция $F(x)$ непрерывна как частное двух непрерывных функций. В таком случае, согласно теореме 4.10, $F(x)$ ограничена, т. е. найдется положительное число μ такое, что для всех $x \in [a, b]$

$$F(x) = \frac{1}{M - f(x)} \leqslant \mu,$$

откуда

$$f(x) \leqslant M - \frac{1}{\mu}.$$

Получилось, что число $M - \frac{1}{\mu}$ меньшее, чем M, является точной верхней гранью для Y. Но это противоречит тому, что число M является точной верхней, т. е. наименьшей, гранью функции $f(x)$ на отрезке $[a, b]$. Полученное противоречие и доказывает, что существует точка $x_1 \in [a, b]$, в которой $f(x_1) = M$.

Аналогично доказывается, что $f(x)$ достигает своей точной нижней грани m ■

Замечание 1. После того как доказано, что функция $f(x)$, непрерывная на отрезке $[a, b]$, достигает на этом отрезке своих точных верхней M и нижней m граней, мы можем назвать точную верхнюю грань *максимальным значением*, а точную нижнюю грань *минимальным значением* функции $f(x)$ на этом отрезке и сформулировать теорему 4.11 в виде: непрерывная на отрезке функция имеет на этом отрезке максимальное и минимальное значения.

Замечание 2. Разность между наибольшим и наименьшим значениями функции $f(x)$ на отрезке $[a, b]$ называется колебанием непрерывной функции на этом отрезке и обозначается $\omega = M - m$, где

$$M = \sup_{[a,b]} Y, \quad m = \inf_{[a,b]} Y.$$

5. Понятие равномерной непрерывности функции. К числу других свойств функции, непрерывной на отрезке, относится очень важное свойство, называемое *равномерной непрерывностью*. Оно широко используется при доказательстве ряда фундаментальных теорем.

Пусть $f(x)$ — функция, непрерывная на некотором промежутке X, и пусть точка $x_0 \in X$. Так как $f(x)$ непрерывна в точке x_0, то по второму определению непрерывности для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что $|f(x) - f(x_0)| < \varepsilon$ при $|x - x_0| < \delta$. Ясно, что δ зависит от ε, но δ зависит также и от x_0. При изменении x_0 в пределах рассматриваемого промежутка (при постоянном ε) число δ будет различным. Чем круче идет график функции $f(x)$ в окрестности точки x_0, тем меньше будет δ, соответствующее этой точке (рис. 53).

Таким образом, при заданном ε для каждой точки x рассматриваемого промежутка определяется соответствующее ей δ. Если бы точек было *конечное* число, то

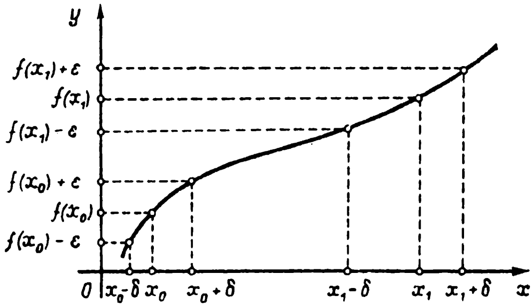


Рис. 53

из конечного множества чисел δ можно было бы выбрать наименьшее δ, которое зависело бы только от ε и не зависело от x. Для бесконечного числа точек этого, «вообще говоря», сделать нельзя, так как этим точкам соответствует бесконечное множество чисел δ, среди которых могут найтись и сколь угодно малые.

Возникает вопрос, существуют ли непрерывные функции, определенные на некоторых промежутках, для которых по любому $\varepsilon>0$ находилось бы $\delta>0$, не зависящее от x, т. е. δ было бы общим для всех x из рассматриваемого промежутка. Этот вопрос и приводит к так называемому *понятию равномерной непрерывности* функции. Ниже будет показано, что при некоторых условиях такие функции существуют.

Дадим следующее

Определение. Функция $f(x)$ называется *равномерно-непрерывной* на некотором промежутке X, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что для любых двух точек $x', x''\in X$, удовлетворяющих неравенству $|x''-x'|<\delta$, выполняется неравенство

$$|f(x'')-f(x')|<\varepsilon.$$

По самому определению δ зависит только от ε и является общим для всех x', x'' промежутка X.

Надо сказать, что понятие равномерной непрерывности функции относится к наиболее сложным и трудным для понимания понятиям анализа.

Из определения очевидно, что равномерно-непрерывная функция на X является непрерывной на этом промежутке.

Следующая теорема устанавливает условие, при котором непрерывная функция является и равномерно-непрерывной.

6. Теорема о равномерной непрерывности функции.

Теорема 4.12 (теорема Кантора*). *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она и равномерно-непрерывна на нем.*

Доказательство. Сначала докажем, что если $f(x)$ непрерывна на $[a, b]$, то для любого $\varepsilon > 0$ можно разбить $[a, b]$ на конечное число непересекающихся отрезков, которые или не имеют общих точек, или имеют только одну общую граничную точку и на каждом из которых для любых двух точек x', x'' будет выполняться неравенство $|f(x'') - f(x')| < \varepsilon$.

Предположим обратное, т. е. допустим, что существует $\varepsilon > 0$, для которого такое разбиение $[a, b]$ невозможно. Разделим $[a, b]$ пополам и выберем тот из полученных двух отрезков, для которого такое разбиение невозможно. Обозначим его $[a_1, b_1]$. Разделим теперь $[a_1, b_1]$ пополам и выберем тот из полученных двух отрезков, для которого такое разбиение невозможно, и т. д. Продолжая этот процесс неограниченно, получим последовательность вложенных отрезков

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \supset \ldots,$$

обладающих тем свойством, что ни один из них нельзя разбить на конечное число непересекающихся отрезков, на каждом из которых для любых двух точек x' и x'' будет выполняться неравенство $|f(x'') - f(x')| < \varepsilon$. По теореме 2.13 о вложенных отрезках существует точка c, принадлежащая всем отрезкам. Так как $f(x)$ непрерывна в точке c, то для рассматриваемого ε найдется δ такое, что $|f(x) - f(c)| < \frac{\varepsilon}{2}$ для любого x из δ-окре-

* Кантор Георг (1845 — 1918) — немецкий математик, основатель современной теории множеств.

стности точки c. Тогда для любых двух точек x' и x'' δ-окрестности будет выполняться неравенство

$$|f(x'')-f(x')|=|(f(x'')-f(c))+(f(c)-f(x'))|\leqslant$$

$$\leqslant|f(x'')-f(c)|+|f(c)-f(x')|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}<\varepsilon,$$

т. е.

$$|f(x'')-f(x')|<\varepsilon.$$

В δ-окрестность точки c при достаточно большом n попадет отрезок $[a_n, b_n]$, и, следовательно, для любых двух точек x' и x'' этого отрезка будет справедливо неравенство $|f(x'')-f(x')|<\varepsilon$, а это противоречит выбору последовательности вложенных отрезков.

Перейдем теперь непосредственно к доказательству теоремы. По только что доказанному для любого $\varepsilon>0$ существует разбиение $[a, b]$ на конечное число отрезков, в каждом из которых разность между любыми двумя значениями функции $f(x)$ будет по абсолютной величине меньше $\varepsilon/2$. Обозначим через δ длину наименьшего из отрезков разбиения и рассмотрим любые две точки x' и x'' отрезка $[a, b]$, отстоящие друг от друга меньше, чем на δ, т. е. $|x''-x'|<\delta$. Может представиться два случая: 1) либо эти точки принадлежат одному отрезку разбиения, 2) либо принадлежат двум соседним. В первом случае $|f(x'')-f(x')|<\frac{\varepsilon}{2}<\varepsilon$; во втором случае, обозначая через x_0 общую граничную точку соседних отрезков, получим:

$$|f(x'')-f(x')|=|(f(x'')-f(x_0))+(f(x_0)-f(x'))|\leqslant$$

$$\leqslant|f(x'')-f(x_0)|+|f(x_0)-f(x')|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

Таким образом, для любого $\varepsilon>0$ найдется $\delta>0$ такое, что для любых двух точек x' и x'' отрезка $[a, b]$, удовлетворяющих неравенству $|x''-x'|<\delta$, выполняется неравенство $|f(x'')-f(x')|<\varepsilon$, что и требовалось доказать ■

Следствие. *Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$. Тогда для любого $\varepsilon>0$ существует $\delta>0$ такое, что если $[a, b]$ произвольно разбить на конечное число отрезков с длинами, меньшими δ, то в* *каждом из них колебание ω функции $f(x)$ будет меньше ε.*

Доказательство. Действительно, по доказанной теореме $f(x)$ будет равномерно-непрерывной на $[a, b]$. Следовательно, для любого $\varepsilon>0$ найдется $\delta>0$ такое, что для любых точек x' и x'' отрезка $[a, b]$, удовлетворяющих неравенству $|x''-x'|<\delta$, выполняется неравенство $|f(x'')-f(x')|<\varepsilon$. Если отрезок $[a, b]$ произвольно разбить на конечное число отрезков с длинами, меньшими указанного δ, то, поскольку $f(x)$ непрерывна на

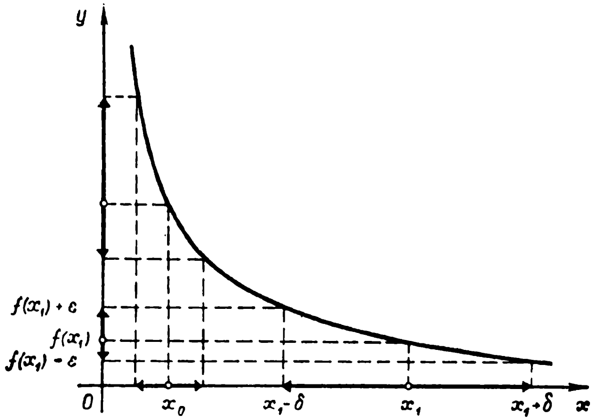


Рис. 54

$[a, b]$, на каждом из частичных отрезков можно указать такие точки x' и x'', что $f(x')=m$, а $f(x'')=M$, где m и M — точные нижняя и верхняя грани функции $f(x)$ на данном частичном отрезке. Так как $|x''-x'|<$ $<\delta$, то $|f(x'')-f(x')|<\varepsilon$. Но $f(x'')-f(x')=M-m=\omega$. Поэтому $\omega<\varepsilon$ ■

Замечание. Теорема становится неверной, если в ней отрезок $[a, b]$ заменить интервалом или полуинтервалом.

Пример. Рассмотрим функцию $f(x)=\frac{1}{x}$ определенную на интервале $(0, 1)$. Данная функция непрерывна на интервале $(0, 1)$, но не является равномерно-непрерывной на нем. Это следует из того, что для любого фиксированного $\varepsilon>0$, какое бы $\delta>0$ мы ни взяли, всегда найдутся точки x' и x'', достаточно близкие к нулю, расстояние между которыми меньше δ, а модуль разности $|f(x'')-f(x')|$ больше ε (рис. 54).

§ 11. Понятие сложной функции

Определение. Если на некотором промежутке X определена функция $z=\varphi(x)$ с множеством значений Z, а на множестве Z определена функция $y=f(z)$, то $y=f[\varphi(x)]$ называется *сложной функцией** от x, а переменная z — *промежуточной переменной* сложной функции.

Пример. Функция $y=\sin x^2$ — сложная функция, определенная на всей числовой прямой, так как $y=f(z)=\sin z$, $z=\varphi(x)=x^2$.

Докажем следующую теорему.

Теорема 4.13. *Пусть функция $z=\varphi(x)$ непрерывна в точке x_0, а функция $y=f(z)$ непрерывна в точке $z_0=\varphi(x_0)$. Тогда сложная функция $y=f[\varphi(x)]$ непрерывна в точке x_0.*

Доказательство. Возьмем из X любую последовательность точек

$$x_1,\ x_2,\ x_3,\ \ldots,\ x_n,\ \ldots,$$

сходящуюся к точке x_0. Тогда в силу непрерывности функции $z=\varphi(x)$ в точке x_0 имеем: $\lim\limits_{n\to\infty} z_n=\lim\limits_{n\to\infty}\varphi(x_n)=\varphi(x_0)=z_0$, т. е. соответствующая последовательность точек

$$z_1,\ z_2,\ z_3,\ \ldots,\ z_n,\ \ldots,$$

сходится к точке z_0. В силу же непрерывности функции $f(z)$ в точке z_0 получаем: $\lim\limits_{n\to\infty} f(z_n)=f(z_0)$, т. е. $\lim\limits_{n\to\infty} f[\varphi(x_n)]=f[\varphi(x_0)]$. Следовательно, предел функции $f[\varphi(x)]$ в точке x_0 равен ее значению в этой точке, что и доказывает непрерывность сложной функции $f[\varphi(x)]$ в точке x_0 ■

Пример. Доказать непрерывность функции $y=\sin x^2$ в точке $x=0$.

---

* Наряду с термином *сложная функция* существует равнозначный термин *композиция* (или *суперпозиция*) *функций*.

Решение. Так как функция $z=x^2$ непрерывна в точке $x=0$, а функция $y=\sin z$ непрерывна в точке $z=0$, то по доказанной теореме сложная функция $y=\sin x^2$ непрерывна в точке $x=0$.

§ 12. Понятие обратной функции

1. Определение обратной функции. По аналогии с числовыми последовательностями (см. § 3, гл. II) будем говорить, что функция $f(x)$ *неубывающая* (*невозрастающая*) на множестве X, если для любых $x_1, x_2 \in X$, удовлетворяющих условию $x_1 < x_2$, справедливо неравенство $f(x_1) \leqslant f(x_2)$ $(f(x_1) \geqslant f(x_2))$.

Неубывающие и невозрастающие функции объединяются общим названием *монотонные функции.*

Если для любых $x_1, x_2 \in X$, удовлетворяющих условию $x_1 < x_2$, справедливо неравенство $f(x_1) < f(x_2)$ $(f(x_1) > f(x_2))$, то функция $f(x)$ называется *возрастающей* (*убывающей*) на множестве X. Возрастающие и убывающие функции называются также *строго монотонными.*

Примеры:

1) функция $f(x) = \operatorname{sgn} x$ является неубывающей на всей числовой прямой;
2) функция $f(x) = x$ является возрастающей на всей числовой прямой.

Введем теперь понятие обратной функции.

Определение. Пусть X и Y — некоторые множества и пусть задана функция f, т. е. множество пар чисел $(x; y)$ $(x \in X, y \in Y)$, в котором каждое число x входит в одну и только одну пару. Если в каждой паре этого множества числа x и y поменять местами, то получим множество пар чисел $(y; x)$, которое называется *обратной функцией* φ к функции f.

Обратная функция чаще обозначается символом $x = \varphi(y)$.

Отметим, что обратная функция, вообще говоря, не является функцией, так как каждое число y может входить не только в одну, но и в несколько пар. Так, например, для функции $y=x$ обратная функция $x=y$ однозначна (каждое число y входит в одну пару), для функции $y=x^2$ обратная функция $x = \pm\sqrt{y}$ двузначна (каждое число y входит в две пары), а обратная функция $x=\arcsin y$ для функции $y=\sin x$ многозначна (каждое число y входит в бесконечное число пар). Геометрически данный факт очевиден.

Из определения следует, что множество значений Y функции f является областью определения обратной функции φ, а область определения X функции f — множеством значений обратной функции φ. Пусть, например, функция $y=f(x)$ определена на отрезке $[a, b]$ и

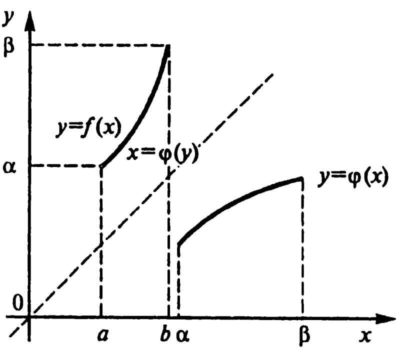


Рис. 55

отрезок $[\alpha, \beta]$ является множеством ее значений. Тогда по определению на $[\alpha, \beta]$ определена обратная функция $x=\varphi(y)$, множеством значений которой является $[a, b]$ (рис. 55).

Таким образом, функция $y=f(x)$ и обратная функция $x=\varphi(y)$ имеют один и тот же график. Так, например, функция $y=5x$ и обратная функция $x=\frac{1}{5}y$ изображаются графически одной прямой.

Если x и y поменять местами, т. е. повернуть плоскость Oxy вокруг биссектрисы первого координатного угла на 180°, то новое положение графика обратной функции $x=\varphi(y)$ будет графиком обратной функции $y=\varphi(x)$ (см. рис. 55).

2. Теорема о непрерывности обратной функции

Теорема 4.14. *Пусть функция $y=f(x)$ определена, строго монотонна и непрерывна на некотором промежутке X. Тогда на соответствующем промежутке Y*

обратная функция $x=\varphi(y)$ однозначна, строго монотонна и непрерывна.

Доказательство. Пусть для определенности функция $f(x)$ строго возрастает, т. е. для любых $x_1, x_2 \in X$, удовлетворяющих условию $x_1 < x_2$, выполняется неравенство $y_1 < y_2$ ($y_1 = f(x_1)$ и $y_2 = f(x_2)$) (рис. 56).

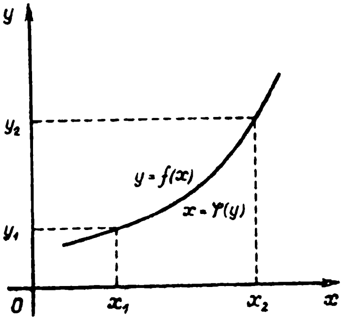


Рис. 56

Однозначность обратной функции $x=\varphi(y)$ следует из строгого возрастания функции $y=f(x)$ на X: $y_1 = f(x_1) \neq f(x_2) = y_2$ при $x_1 \neq x_2$, т. е. каждому $y \in Y$ соответствует единственное значение $x \in X$.

Докажем теперь, что обратная функция $x=\varphi(y)$ строго возрастает на Y. Действительно, если $y_1 < y_2$, то и $x_1 < x_2$ ($x_1 = \varphi(y_1)$ и $x_2 = \varphi(y_2)$), ибо если $x_1 \geqslant x_2$, то из возрастания $f(x)$ следовало бы $y_1 \geqslant y_2$, что противоречило бы предположению $y_1 < y_2$. Следовательно, $x_1 < x_2$. Таким образом, факт строгой монотонности обратной функции $x=\varphi(y)$ установлен.

И наконец, покажем, что обратная функция $x=\varphi(y)$ непрерывна на Y. Пусть точка $y_0 \in Y$, $x_0 = \varphi(y_0)$ и $m = \inf\limits_X Y$, $M = \sup\limits_X Y$ (рис. 57). Кроме этого, пусть сначала

$$m < y_0 < M,$$

т. е. $y_0 \in (m, M)$. Тогда в силу строгого возрастания $\varphi(y)$ будет

$$\varphi(m) < x_0 < \varphi(M).$$

Возьмем $\varepsilon > 0$ таким, чтобы

$$\varphi(m) \leqslant x_0 - \varepsilon < x_0 < x_0 + \varepsilon \leqslant \varphi(M), \tag{1}$$

и положим $y_1 = f(x_0 - \varepsilon)$ и $y_2 = f(x_0 + \varepsilon)$. Тогда в силу строгого возрастания $f(x)$ и неравенств (1) получим

$$m \leqslant y_1 < y_0 < y_2 \leqslant M.$$

Возьмем теперь $\delta > 0$ таким, чтобы

$$y_1 \leqslant y_0 - \delta < y_0 < y_0 + \delta \leqslant y_2,$$

откуда, принимая во внимание, что Y есть некоторый промежуток (см. следствие из теоремы 4.9), получим

$$y_0-\delta<y<y_0+\delta$$

и тем более

$$y_1<y<y_2. \tag{2}$$

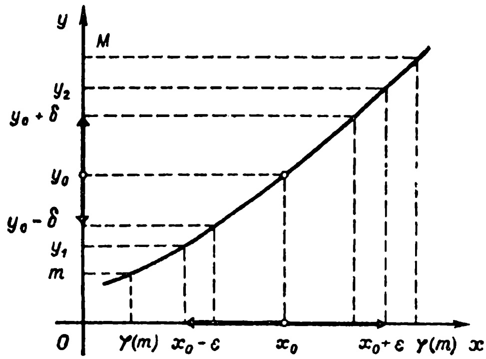


Рис. 57

Из неравенств (2) в силу строгого возрастания $\varphi(y)$ следует, что

$$\varphi(y_1)<\varphi(y)<\varphi(y_2),$$

или, с учетом того, что $\varphi(y_1)=x_0-\varepsilon=\varphi(y_0)-\varepsilon$ и $\varphi(y_2)=$ $=x_0+\varepsilon=\varphi(y_0)+\varepsilon$,

$$\varphi(y_0)-\varepsilon<\varphi(y)<\varphi(y_0)+\varepsilon.$$

Таким образом, доказано, что для любого достаточно малого $\varepsilon>0$ существует $\delta>0$ такое, что для всех $y\in(m, M)$, удовлетворяющих неравенству $|y-y_0|<\delta$, выполняется неравенство $|\varphi(y)-\varphi(y_0)|<\varepsilon$, т. е. обратная функция $\varphi(y)$ непрерывна в точке y_0. Но y_0 — любая точка интервала (m, M). Значит, $\varphi(y)$ непрерывна на (m, M).

Для доказательства непрерывности $x=\varphi(y)$ на $[m, M]$ достаточно положить $y_0=m$ или $y_0=M$ и аналогичными рассуждениями показать непрерывность $\varphi(y)$ справа в точке m и слева в точке M.

Итак, факт непрерывности обратной функции $x=\varphi(y)$ на Y доказан.

В случае строгого убывания функции $f(x)$ доказательство теоремы аналогичное ■

Замечание. Необходимо отметить, что если обратная функция $x=\varphi(y)$ однозначна, то, очевидно, функция $y=f(x)$ является обратной для функции $x=\varphi(y)$. Такие функции называют также *взаимно-обратными*.

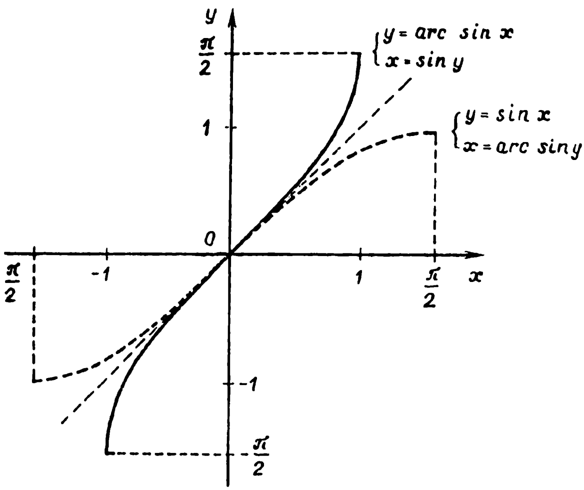


Рис. 58

Рассмотрим конкретный пример. Функция $y=\sin x$ на отрезке $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ строго возрастает, непрерывна и имеет в качестве множества значений отрезок $[-1, 1]$. По теореме 4.14 на $[-1, 1]$ существует непрерывная возрастающая обратная функция с множеством значений $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Эту обратную функцию обозначают $x=\arcsin y$, график ее совпадает с графиком функции $y=\sin x$, рассматриваемой только при $-\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2}$ (рис. 58). Если теперь x и y поменять местами, т. е. если рассматривать функцию $y=\arcsin x$, то график примет вид, указанный сплошной линией.

Аналогично можно рассмотреть и другие обратные функции.

Глава V. ДИФФЕРЕНЦИРОВАНИЕ

§ 1. Понятие производной

1. Определение производной. Пусть на некотором промежутке X определена функция $f(x)$. Возьмем любое значение $x_0 \in X$ и зададим аргументу x в точке x_0 произвольное приращение Δx такое, что значение $x_0+\Delta x$ также принадлежит X. Это вызовет соответствующее приращение функции $\Delta y=f(x_0+\Delta x)-f(x_0)$.

Определение. Производной функции $f(x)$ в точке x_0 называется *предел* при $\Delta x \to 0$ отношения приращения функции в этой точке к приращению аргумента (при условии, что этот предел существует).

Для обозначения производных пользуются следующими символами: y', $y'(x)$, $f'(x)$.

Итак, по определению,

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}.$$

Если для некоторого значения x_0 выполняется условие

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=+\infty, \text{ или } \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=-\infty,$$

то говорят, что в точке x_0 существует *бесконечная производная*.

Если вместо x_0 подставлять различные значения x из X, то будем получать различные значения производной. Следовательно, производную можно рассматривать снова как функцию от x, также определенную на X.

Из определения производной вытекает и способ ее вычисления.

Пример. Найти производную функции $f(x)=x^2$ в точке $x=x_0$.

Решение. Давая аргументу x в точке x_0 приращение Δx, найдем соответствующее приращение функции:

$$\Delta y=f(x_0+\Delta x)-f(x_0)=(x_0+\Delta x)^2-x_0^2=x_0^2+$$
$$+2x_0\Delta x+(\Delta x)^2-x_0^2=2x_0\Delta x+(\Delta x)^2.$$

Составим отношение

$$\frac{\Delta y}{\Delta x} = \frac{2x_0\Delta x + (\Delta x)^2}{\Delta x}.$$

Найдем предел этого отношения при $\Delta x \to 0$:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{2x_0\Delta x + (\Delta x)^2}{\Delta x} = 2x_0.$$

Следовательно, производная функции $f(x) = x^2$ в точке x_0 равна числу $2x_0$, что в принятых обозначениях можно записать так: $f'(x_0) = 2x_0$.

2. Геометрический смысл производной. Пусть функция $f(x)$ определена и непрерывна на интервале (a, b). Пусть, далее, точка M на графике функции соответствует некоторому значению аргумента x_0, а точка P —

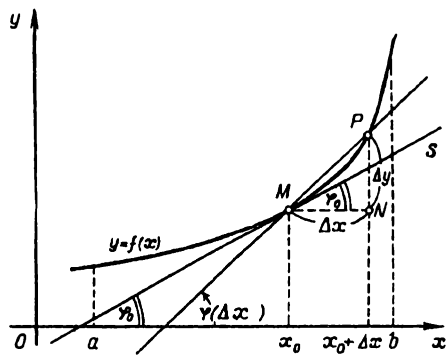


Рис. 59

значению $x_0 + \Delta x$, где Δx — приращение аргумента. Проведем через точки M и P прямую и назовем ее *секущей*. Обозначим через $\varphi(\Delta x)$ угол между секущей и осью Ox. Очевидно, что этот угол зависит от Δx. Касательной S к графику функции $f(x)$ в точке M будем называть предельное положение секущей MP при неограниченном приближении точки P по графику к точке M (или, что то же самое, при $\Delta x \to 0$). Из рис. 59 ясно, что

$$\operatorname{tg} \varphi(\Delta x) = \frac{PN}{MN} = \frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Так как при $\Delta x \to 0$ секущая MP переходит в касательную, то

$$\lim_{\Delta x \to 0} \operatorname{tg} \varphi(\Delta x) = \operatorname{tg} \varphi_0,$$

где φ_0 — угол, который образует касательная с осью Ox. С другой стороны,

$$\lim_{\Delta x \to 0} \operatorname{tg} \varphi(\Delta x) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0).$$

Следовательно, $f'(x_0) = \operatorname{tg} \varphi_0$. Таким образом, производная функции $f(x)$ в точке x_0 равна угловому коэффициенту касательной к графику функции $f(x)$ в точке $M(x_0; f(x_0))$.

3. Физический смысл производной. Предположим, что функция $y = f(x)$ описывает закон движения материальной точки M по прямой линии, т. е. зависимость пути y, пройденного точкой от начала отсчета за время x.

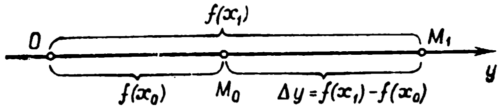


Рис. 60

Тогда за время x_0 пройден путь $y = f(x_0)$, а за время x_1 — путь $y = f(x_1)$. За промежуток времени $\Delta x = x_1 - x_0$ точка M пройдет отрезок пути $\Delta y = f(x_1) - f(x_0) = f(x_0 + \Delta x) - f(x_0)$ (рис. 60). Отношение $\frac{\Delta y}{\Delta x}$ называется средней скоростью движения ($V_{\text{ср}}$) за время Δx, а предел отношения $\Delta y / \Delta x$ при $\Delta x \to 0$ определяет мгновенную скорость точки в момент времени x ($V_{\text{мгн}}$).

Понятие скорости, заимствованное из физики, удобно при исследовании поведения произвольной функции. Какую бы зависимость ни отражала функция $y = f(x)$, отношение $\Delta y / \Delta x$ есть средняя скорость изменения y относительно изменения x, а $y'(x)$ есть мгновенная скорость изменения y при некотором $x = x_0$.

Значение производной состоит в том, что при изучении любых процессов и явлений природы с ее помощью можно оценить скорости изменения связанных между собой величин.

При изучении общего курса физики читатель встретится с многочисленными примерами приложения понятия производной.

4. Правая и левая производные. По аналогии с понятием правого и левого пределов функции вводятся понятия правой и левой производных функции $f(x)$ в точке x_0.

Определение. Правой (левой) производной функции $f(x)$ в точке x_0 называется *правое* (*левое*) *предельное значение* $\Delta y/\Delta x$ (при условии, что это предельное значение существует).

Обозначение:

$$f'_+(x_0) = \lim_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} \left(f'_-(x_0) = \lim_{\Delta x \to 0-} \frac{\Delta y}{\Delta x} \right).$$

Если функция $f(x)$ имеет в точке x_0 производную, то она имеет в этой точке и правую и левую производные, совпадающие между собой.

Вместе с тем существуют функции, имеющие в данной точке x_0 и правую и левую производные, но не имеющие производной в этой точке. Примером такой функции может служить функция $f(x) = |x|$. Эта функция имеет в точке $x=0$ правую производную, равную $f'_+(0) = \lim\limits_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} = 1$ (при $x \geqslant 0$ $\Delta y = \Delta x$) и левую производную, равную $f'_-(0) = \lim\limits_{\Delta x \to 0-} \frac{\Delta y}{\Delta x} = -1$ (при $x<0$ $\Delta y = -\Delta x$), но не имеет в точке $x=0$ производной, так как $f_+'(0) \neq f_-'(0)$.

§ 2. Понятие дифференцируемости функции

1. Понятие дифференцируемости функции в данной точке

Определение. Функция $f(x)$ называется *дифференцируемой* в точке x_0, если ее приращение Δy в этой точке можно представить в виде

$$\Delta y = A\,\Delta x + \alpha(\Delta x)\,\Delta x, \tag{1}$$

где A — некоторое число, не зависящее от Δx, а $\alpha(\Delta x)$ — функция аргумента Δx, являющаяся бесконечно малой при $\Delta x \to 0$, т. е. $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$, причем $\alpha(0) = 0$.

Выясним теперь связь между дифференцируемостью в точке и существованием производной в той же точке.

Теорема 5.1. *Для того чтобы функция $f(x)$ была дифференцируема в данной точке x_0, необходимо и достаточно, чтобы она имела в этой точке конечную производную.*

Доказательство. *Необходимость.* Пусть функция $f(x)$ дифференцируема в данной точке x_0, т. е. $\Delta y = A\,\Delta x + \alpha(\Delta x)\Delta x$. Тогда, предположив, что $\Delta x \neq 0$, и разделив равенство на Δx, получим

$$\frac{\Delta y}{\Delta x} = A + \alpha(\Delta x).$$

Переходя к пределу при $\Delta x \to 0$, имеем

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (A + \alpha(\Delta x)) = A = f'(x_0).$$

Отсюда следует, что производная в точке x_0 существует.

Достаточность. Пусть существует производная $f'(x_0)$, т. е. существует $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x} = f'(x_0) = A$. Тогда в силу специального представления (1) (см. § 5, гл. IV) функция $\alpha(\Delta x) = \dfrac{\Delta y}{\Delta x} - A$ является бесконечно малой при $\Delta x \to 0$, т. е.

$$\Delta y = A\Delta x + \alpha(\Delta x)\,\Delta x,$$

где $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$. Получили представление (1), тем самым доказано, что функция $f(x)$ дифференцируема в точке x_0 ■

Таким образом, для функций одной переменной дифференцируемость и существование производной — понятия равносильные. Поэтому операцию нахождения производной часто называют *дифференцированием*. Однако, забегая немного вперед, можно сказать, что для функций двух и более переменных дело обстоит иначе.

2. Связь между понятиями дифференцируемости и непрерывности.

Теорема 5.2. *Если функция $f(x)$ дифференцируема в данной точке x_0, то она и непрерывна в этой точке.*

Доказательство. Так как функция $f(x)$ дифференцируема в точке x_0, то ее приращение в этой точке может быть представлено соотношением (1). Тогда, переходя к пределу при $\Delta x \to 0$, получим

$$\lim_{\Delta x \to 0} \Delta y = A \lim_{\Delta x \to 0} \Delta x + \lim_{\Delta x \to 0} \alpha(\Delta x) \lim_{\Delta x \to 0} \Delta x = 0,$$

что и означает непрерывность функции $f(x)$ в точке x_0 в силу третьего определения непрерывности функции в точке ■

Замечание. Обратное утверждение неверно. Функция может быть непрерывной в точке, но не быть дифференцируемой, т. е. не иметь производной в этой точке.

Примером такой функции может служить функция $f(x) = |x|$. Эта функция, как известно, непрерывна в точке $x=0$, но она, как показано в п. 4 § 1, не имеет в этой точке производной, т. е. не является дифференцируемой.

Если функция $f(x)$ имеет производную в каждой точке некоторого промежутка (дифференцируема в каждой точке этого промежутка), то мы будем говорить, что функция $f(x)$ *имеет производную*, или что она *дифференцируема* на указанном промежутке.

§ 3. Понятие дифференциала

1. Определение и геометрический смысл дифференциала. Пусть функция $f(x)$ дифференцируема в точке x_0, т. е. приращение Δy можно записать в виде суммы двух слагаемых:

$$\Delta y = A\,\Delta x + \alpha(\Delta x)\,\Delta x,$$

где $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$. Первое слагаемое $A\,\Delta x$ является при $\Delta x \to 0$ бесконечно малой одного порядка с Δx (при $A \neq 0$). Оно линейно относительно Δx, второе слагаемое $\alpha(\Delta x)\Delta x$ при $\Delta x \to 0$ — бесконечно малая более высокого порядка, чем Δx $\left(\lim\limits_{\Delta x \to 0} \dfrac{\alpha(\Delta x)\,\Delta x}{\Delta x} = 0\right)$.

Таким образом, первое слагаемое (при $A \neq 0$) является главной частью приращения функции $f(x)$.

Определение. Дифференциалом функции $f(x)$ в точке x_0 называется *главная, линейная относительно*

Δx, *часть приращения* функции:

$$dy = A\,\Delta x. \tag{1}$$

Если $A=0$, то $A\,\Delta x=0$, и поэтому слагаемое $A\,\Delta x$ уже не будет главной частью приращения Δy, так как слагаемое $\alpha(\Delta x)\Delta x$, вообще говоря, отлично от нуля. Однако и в этом случае по определению полагаем дифференциал функции в точке x_0 равным $A\,\Delta x$, т. е. в этом случае $dy=0$.

Если учесть теорему 5.1, т. е. учесть, что $A=f'(x_0)$, то формулу (1) можно переписать в виде

$$dy = f'(x_0)\,\Delta x. \tag{2}$$

Пусть $f(x)=x$. Тогда по формуле (2)

$$dy = dx = (x_0)'\,\Delta x = \left(\lim_{\Delta x\to 0}\frac{(x_0+\Delta x)-x_0}{\Delta x}\right)\Delta x =$$

$$= \left(\lim_{\Delta x\to 0}\frac{\Delta x}{\Delta x}\right)\Delta x = 1\cdot\Delta x = \Delta x$$

(дифференциал независимой переменной равен приращению этой переменной), и соотношение (1) примет окончательный вид

$$dy = f'(x_0)\,dx. \tag{3}$$

Заметим, что с помощью равенства (3) производную $f'(x_0)$ можно определить как отношение дифференциала функции dy к дифференциалу аргумента dx, т. е.

$$f'(x_0) = \frac{dy}{dx}.$$

Дифференциал функции имеет определенный геометрический смысл.

Пусть точка M на графике $y=f(x)$ соответствует значению аргумента x_0, точка P — значению аргумента $x_0+\Delta x$, прямая MS — касательная к графику $y=f(x)$ в точке M, α — угол между касательной и осью Ox. Пусть, далее, $MN \parallel Ox$, $PN \parallel Oy$, Q — точка пересечения касательной MS с прямой PN (рис. 61). Тогда приращение функции Δy равно величине отрезка NP. В то же время из прямоугольного треугольника MNQ следует, что $NQ=\Delta x \operatorname{tg}\alpha=\Delta x\, f'(x_0)=dy$, т. е. дифференциал функции dy равен по величине отрезку NQ, когда аргумент x получил приращение Δx, равное по величине отрезку MN. Из геометрического построения видно, что величины отрезков NP и NQ различны.

Таким образом, дифференциал функции $f(x)$ в точке x_0 равен приращению ординаты касательной MS к графику этой функции в рассматриваемой точке, а приращение функции $\Delta y = NP$ есть приращение ординаты самой функции $f(x)$ в точке x_0, соответствующее приращению Δx.

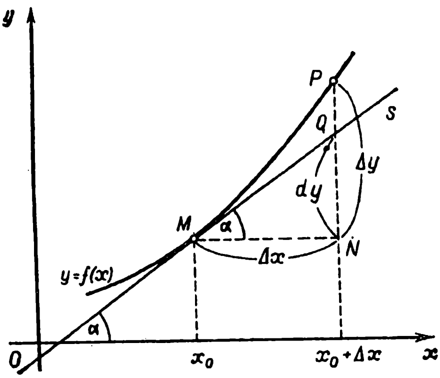


Рис. 61

2. Приближенные вычисления с помощью дифференциала. Из определения дифференциала следует, что он зависит линейно от Δx и является главной частью приращения функции Δy. Само же Δy зависит от Δx более сложно. Например, если $f(x) = x^3$, то $\Delta y = (x_0 + \Delta x)^3 - x_0^3 = 3x_0^2\Delta x + 3x_0(\Delta x)^2 + (\Delta x)^3$, в то время как

$$dy = (x_0^3)'\Delta x = \left(\lim_{\Delta x \to 0} \frac{(x_0 + \Delta x)^3 - x_0^3}{\Delta x}\right)\Delta x = 3x_0^2\Delta x.$$

Кроме того, для вычисления дифференциала можно воспользоваться равенством $dy = f'(x_0)dx$. Во многих задачах вместо приращения функции в данной точке рассматривают дифференциал функции в этой точке, полагая

$$\Delta y \approx dy,$$

если $f'(x_0) \neq 0$. При этом абсолютная погрешность равна $|\Delta y - dy|$, которая является бесконечно малой более высокого порядка, чем Δx.

Пример. Покажем, что если α очень мало, то можно пользоваться приближенной формулой

$$\sqrt{1+\alpha} \approx 1 + \frac{\alpha}{2}.$$

Действительно, возьмем функцию $f(x) = \sqrt{x}$. Тогда

$$\Delta y = \sqrt{x_0 + \Delta x} - \sqrt{x_0} \approx dy,$$

$$\sqrt{x_0 + \Delta x} - \sqrt{x_0} \approx (\sqrt{x_0})' \Delta x = \left[\lim_{\Delta x \to 0} \frac{\sqrt{x_0 + \Delta x} - \sqrt{x_0}}{\Delta x}\right] \Delta x =$$

$$= \left[\lim_{\Delta x \to 0} \frac{(x_0 + \Delta x) - x_0}{\Delta x (\sqrt{x_0 + \Delta x} + \sqrt{x_0})}\right] \Delta x = \frac{1}{2\sqrt{x_0}} \Delta x,$$

откуда, положив $x_0 = 1$, $\Delta x = \alpha$, получим

$$\sqrt{1+\alpha} \approx 1 + \frac{\alpha}{2}.$$

В частности, при $\alpha = 0{,}0003$ найдем: $\sqrt{1{,}0003} \approx 1{,}00015$.

Перейдем теперь к установлению правил дифференцирования и вычислению производных простейших элементарных функций. Заметим только, что при выводе формул и практическом вычислении производных обычно пишут не x_0, а просто x, но при этом x считается фиксированным.

§ 4. Правила дифференцирования суммы, разности, произведения и частного

Теорема 5.3. *Если функции $u = u(x)$ и $v = v(x)$ дифференцируемы в точке x, то сумма, разность, произведение и частное этих функций (частное при условии $v(x) \neq 0$) также дифференцируемы в этой точке и имеют место следующие формулы:*

$$1°.\ (u \pm v)' = u' \pm v', \quad 2°.\ (uv)' = u'v + uv',$$

$$3°.\ \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}. \tag{1}$$

Доказательство. Для вывода формул (1) мы воспользуемся определением производной, очевидным равенством $f(x+\Delta x)=f(x)+\Delta y$ и теоремой 4.2. Рассмотрим отдельно каждый случай.

1°. $$(u\pm v)'=\lim_{\Delta x\to 0}\frac{[u(x+\Delta x)\pm v(x+\Delta x)]-u[(x)\pm v(x)]}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}\left[\frac{u(x+\Delta x)-u(x)}{\Delta x}\pm\frac{v(x+\Delta x)-v(x)}{\Delta x}\right]=$$
$$=\lim_{\Delta x\to 0}\frac{u(x+\Delta x)-u(x)}{\Delta x}\pm\lim_{\Delta x\to 0}\frac{v(x+\Delta x)-v(x)}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}\pm\lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x}=u'\pm v'.$$

2°. $$(uv)'=\lim_{\Delta x\to 0}\frac{u(x+\Delta x)\,v(x+\Delta x)-u(x)\,v(x)}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}\frac{[u(x)+\Delta u][v(x)+\Delta v]-u(x)\,v(x)}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}\frac{u(x)\,v(x)+\Delta u v(x)+u(x)\,\Delta v+\Delta u\Delta v-u(x)\,v(x)}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}\left[v(x)\frac{\Delta u}{\Delta x}+u(x)\frac{\Delta v}{\Delta x}+\Delta v\frac{\Delta u}{\Delta x}\right]=$$
$$=v\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}+u\lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x}+\lim_{\Delta x\to 0}\Delta v\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=$$
$$=vu'+uv'+0\cdot u'=u'v+uv',$$

так как $\lim\limits_{\Delta x\to 0}\Delta v=0$, а множители u и v являются постоянными и не зависят от Δx.

3°. $$\left(\frac{u}{v}\right)'=\lim_{\Delta x\to 0}\frac{\dfrac{u(x+\Delta x)}{v(x+\Delta x)}-\dfrac{u(x)}{v(x)}}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}\frac{u(x+\Delta x)\,v(x)-u(x)\,v(x+\Delta x)}{\Delta x v(x+\Delta x)\,v(x)}=$$
$$=\lim_{\Delta x\to 0}\frac{[u(x)+\Delta u]\,v(x)-u(x)\,[v(x)+\Delta v]}{\Delta x v(x)\,[v(x)+\Delta v]}=$$
$$=\lim_{\Delta x\to 0}\frac{uv+\Delta uv-uv-u\Delta v}{\Delta x v(v+\Delta v)}=$$
$$=\lim_{\Delta x\to 0}\frac{v\dfrac{\Delta u}{\Delta x}-u\dfrac{\Delta v}{\Delta x}}{v^2+v\Delta v}=\frac{v\lim\limits_{\Delta x\to 0}\dfrac{\Delta u}{\Delta x}-u\lim\limits_{\Delta x\to 0}\dfrac{\Delta v}{\Delta x}}{v^2+v\lim\limits_{\Delta x\to 0}\Delta v}=\frac{u'v-uv'}{v^2}\ \blacksquare$$

§ 5. Вычисление производных постоянной, степенной, тригонометрических функций и логарифмической функции

В этом параграфе мы приступаем к вычислению производных простейших элементарных функций.

1. Производная постоянной функции. Производная функции $y=f(x)=C$, где C — постоянное число, выражается формулой

$$y'=0.$$

Доказательство. Для любых x и Δx будет $f(x+\Delta x)=C$ и $\Delta y=f(x+\Delta x)-f(x)=0$. Отсюда при любом $\Delta x\neq 0$ отношение $\frac{\Delta y}{\Delta x}=0$ и, следовательно,

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=0\ ■$$

2. Производная степенной функции. Производная функции $y=x^n$, показатель n которой является целым положительным числом, выражается формулой

$$y'=nx^{n-1}.$$

Доказательство. Используя формулу бинома Ньютона, можем записать

$$\Delta y=(x+\Delta x)^n-x^n=$$

$$=\left[x^n+nx^{n-1}\Delta x+\frac{n(n-1)}{2!}x^{n-2}(\Delta x)^2+\ldots+(\Delta x)^n\right]-x^n=$$

$$=nx^{n-1}\Delta x+\frac{n(n-1)}{2!}x^{n-2}(\Delta x)^2+\ldots+(\Delta x)^n.$$

Таким образом, при $\Delta x\neq 0$ будем иметь

$$\frac{\Delta y}{\Delta x}=nx^{n-1}+\frac{n(n-1)}{2!}x^{n-2}\Delta x+\ldots+(\Delta x)^{n-1}.$$

Так как

$$\lim_{\Delta x\to 0}\Delta x=0,\ \lim_{\Delta x\to 0}(\Delta x)^2=0,\ \ldots,\ \lim_{\Delta x\to 0}(\Delta x)^{n-1}=0,$$

то

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=nx^{n-1}\ ■$$

Замечание. Случай степенной функции, показатель которой является любым вещественным числом, мы рассмотрим в § 9 п. 2.

3. Производные тригонометрических функций

1°. $y=\sin x$, $y'=\cos x$.

Имеем

$$\Delta y = \sin(x+\Delta x)-\sin x = 2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right).$$

Таким образом, при $\Delta x\neq 0$

$$\frac{\Delta y}{\Delta x}=\frac{2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right)}{\Delta x}=\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\cos\left(x+\frac{\Delta x}{2}\right).$$

Так как $\lim\limits_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}=1$ (первый замечательный предел), а

$$\lim_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right)=\cos x$$

в силу непрерывности функции $\cos x$, то

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\cos x \blacksquare$$

2°. $y=\cos x$, $y'=-\sin x$.

Имеем

$$\Delta y=\cos(x+\Delta x)-\cos x=-2\sin\frac{\Delta x}{2}\sin\left(x+\frac{\Delta x}{2}\right).$$

Таким образом, при $\Delta x\neq 0$

$$\frac{\Delta y}{\Delta x}=-\frac{2\sin\frac{\Delta x}{2}\sin\left(x+\frac{\Delta x}{2}\right)}{\Delta x}=-\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\sin\left(x+\frac{\Delta x}{2}\right).$$

Так как

$$\lim_{\Delta x\to 0}\sin\left(x+\frac{\Delta x}{2}\right)=\sin x$$

в силу непрерывности функции $\sin x$, то

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=-\sin x \blacksquare$$

3°. $y=\operatorname{tg} x$, $y'=\dfrac{1}{\cos^2 x}\left(x\neq\dfrac{\pi}{2}+n\pi\right)$.

Так как $\operatorname{tg} x = \frac{\sin x}{\cos x}$, то по теореме 5.3 о производной частного получим

$$y' = \frac{(\sin x)' \cos x - \sin x (\cos x)'}{\cos^2 x} = \frac{\cos x \cos x - \sin x (-\sin x)}{\cos^2 x} =$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x},$$

следовательно,

$$y' = \frac{1}{\cos^2 x} \blacksquare$$

4°. $y = \operatorname{ctg} x$, $y' = -\frac{1}{\sin^2 x} (x \neq n\pi)$.

Так как $\operatorname{ctg} x = \frac{\cos x}{\sin x}$, то аналогично предыдущему

$$y' = \frac{(\cos x)' \sin x - \cos x (\sin x)'}{\sin^2 x} = \frac{(-\sin x) \sin x - \cos x \cos x}{\sin^2 x} =$$

$$= -\frac{\sin^2 x + \cos^2 x}{\sin^2 x},$$

следовательно,

$$y' = -\frac{1}{\sin^2 x} \blacksquare$$

4. Производная логарифмической функции. Производная функции $y = \log_a x (0 < a \neq 1)$ выражается формулой

$$y' = \frac{1}{x} \log_a e = \frac{1}{x \ln a}.$$

Доказательство. Имеем

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

Таким образом, при $\Delta x \neq 0$

$$\frac{\Delta y}{\Delta x} = \frac{1}{\Delta x} \log_a \left(1 + \frac{\Delta x}{x}\right) = \frac{1}{x} \cdot \frac{x}{\Delta x} \log_a \left(1 + \frac{\Delta x}{x}\right),$$

или

$$\frac{\Delta y}{\Delta x} = \frac{1}{x} \log_a \left[\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right].$$

Положив $\frac{x}{\Delta x} = h$, будем иметь

$$\lim_{\Delta x \to 0} \left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}} = \lim_{h \to \infty} \left(1 + \frac{1}{h}\right)^{h} = e$$

(второй замечательный предел), а так как логарифмическая функция является непрерывной, то

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{1}{x} \log_a \left[\lim_{\Delta x \to 0} \left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right] =$$

$$= \frac{1}{x} \log_a e = \frac{1}{x \ln a} \blacksquare$$

Следствие. Если $y = \log_e x = \ln x$, то $y' = (\ln x)' = \frac{1}{x}$.

§ 6. Теорема о производной обратной функции

Пусть функция $y=f(x)$ удовлетворяет условиям теоремы 4.14 об обратной функции и функция $x=\varphi(y)$ является для нее обратной. Тогда имеет место следующая

Теорема 5.4. *Если функция $y=f(x)$ имеет в точке x_0 производную $f'(x_0) \neq 0$, то обратная функция $x=\varphi(y)$ также имеет в соответствующей точке $y_0=f(x_0)$ производную, причем*

$$\varphi'(y_0) = \frac{1}{f'(x_0)}.$$

Доказательство. Дадим значению аргумента y_0 обратной функции $x=\varphi(y)$ некоторое приращение $\Delta y \neq 0$, которое, в свою очередь, вызовет приращение функции Δx, причем в силу строгого возрастания (или убывания) функции $\Delta x \neq 0$. Следовательно, можем записать:

$$\frac{\Delta x}{\Delta y} = \frac{1}{\frac{\Delta y}{\Delta x}}.$$

Переходя к пределу при $\Delta y \to 0$, получим

$$\varphi'(y_0) = \lim_{\Delta y \to 0} \frac{\Delta x}{\Delta y} = \frac{1}{\lim_{\Delta y \to 0} \frac{\Delta y}{\Delta x}}.$$

Поскольку функция $y=f(x)$ непрерывна, так как имеет производную $f'(x_0)$, то обратная функция $x=\varphi(y)$ также непрерывна, т. е.

$$\lim_{\Delta y\to 0} \Delta x = 0.$$

Значит,

$$\lim_{\Delta y\to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x} = f'(x_0)$$

и, следовательно,

$$\varphi'(y_0) = \frac{1}{f'(x_0)} \blacksquare \tag{1}$$

Доказанная теорема имеет простой геометрический смысл.

Рассмотрим в некоторой окрестности точки x_0 график функции $y=f(x)$ (или обратной функции $x=\varphi(y)$). Пусть точке x_0 на этом графике соответствует точка M (рис. 62). Очевидно, что производная $f'(x_0)$ равна тангенсу угла наклона α касательной, проходящей через точку M, к оси Ox. Производная обратной функции $\varphi'(y_0)$ равна тангенсу угла наклона β той же касательной к оси Oy. Поскольку углы α и β в сумме составляют $\pi/2$, то формула (1) выражает очевидный факт:

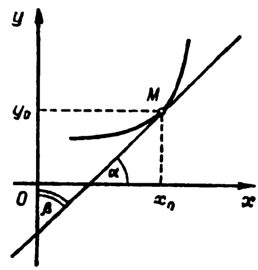


Рис. 62

$$\operatorname{tg}\beta = \frac{1}{\operatorname{ctg}\beta} = \frac{1}{\operatorname{ctg}\left(\frac{\pi}{2}-\alpha\right)} = \frac{1}{\operatorname{tg}\alpha}.$$

§ 7. Вычисление производных показательной функции и обратных тригонометрических функций

Опираясь на доказанную выше теорему 5.4, мы продолжим вычисление производных простейших элементарных функций.

1. Производная показательной функции. Производная функции $y=a^x (0<a\neq 1)$ выражается формулой

$$y'=a^x \ln a.$$

Доказательство. Показательная функция $y=a^x$ является обратной для логарифмической функции $x=\log_a y$. Таким образом,

$$x'(y)=\frac{1}{y}\log_a e,$$

а в силу теоремы 5.4 о производной обратной функции и известного из элементарной математики соотношения $\log_a b=\frac{1}{\log_b a}$, получим

$$y'(x)=\frac{1}{x'(y)}=\frac{y}{\log_a e}=a^x \ln a \blacksquare$$

Следствие. Если $y=e^x$, то $y'=(e^x)'=e^x$.

2. Производные обратных тригонометрических функций

1°. $y=\arcsin x$, $y'=\frac{1}{\sqrt{1-x^2}} \, (|x|<1)$.

Функция $y=\arcsin x$ является обратной для функции $x=\sin y$;

$$x'(y)=\cos y.$$

Отсюда по теореме 5.4 о производной обратной функции получим

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

Знак у корня положительный, потому что $\cos y$ положителен на интервале $-\frac{\pi}{2}<y<\frac{\pi}{2}$. Учитывая, что $\sin y=x$, окончательно получим

$$y'(x)=\frac{1}{\sqrt{1-x^2}} \blacksquare$$

2 . $y=\arccos x$.

$$y'(x)=-\frac{1}{\sqrt{1-x^2}}.$$

Доказательство аналогично предыдущему.

3°. $y=\operatorname{arc\,tg} x$, $y'=\dfrac{1}{1+x^2}$.

Функция $y=\operatorname{arc\,tg} x$ является обратной для функции $x=\operatorname{tg} y$;

$$x'(y)=\frac{1}{\cos^2 y}$$

(так как $-\frac{\pi}{2}<y<\frac{\pi}{2}$, то $x'(y)$ всегда существует и отлично от нуля). Отсюда

$$y'(x)=\frac{1}{x'(y)}=\cos^2 y.$$

Но $\dfrac{1}{\cos^2 y}=1+\operatorname{tg}^2 y=1+x^2$, следовательно,

$$y'(x)=\frac{1}{1+x^2} \blacksquare$$

4°. $y=\operatorname{arc\,ctg} x$.

$$y'=-\frac{1}{1+x^2}.$$

Доказательство аналогично предыдущему.

§ 8. Правило дифференцирования сложной функции

Теорема 5.5. *Если функция $x=\varphi(t)$ имеет производную в точке t_0, а функция $y=f(x)$ имеет производную в соответствующей точке $x_0=\varphi(t_0)$, то сложная функция $f[\varphi(t)]$ имеет производную в точке t_0 и имеет место следующая формула:*

$$y'(t_0)=f'(x_0)\,\varphi'(t_0). \tag{1}$$

Доказательство. Так как функция $y=f(x)$ предполагается дифференцируемой в точке x_0, то приращение этой функции может быть записано в виде

$$\Delta y=f'(x_0)\,\Delta x+\alpha(\Delta x)\,\Delta x, \tag{2}$$

где

$$\lim_{\Delta x\to 0}\alpha(\Delta x)=0.$$

Разделив равенство (2) на Δt, будем иметь

$$\frac{\Delta y}{\Delta t}=f'(x_0)\frac{\Delta x}{\Delta t}+\alpha(\Delta x)\frac{\Delta x}{\Delta t}. \tag{3}$$

Равенство (3) справедливо при любых достаточно малых Δx. Возьмем Δx равным приращению функции $x=\varphi(t)$, соответствующему приращению Δt аргумента t. Устремим в этом равенстве Δt к нулю ($\Delta t \to 0$). Так как по условию функция $x=\varphi(t)$ имеет в точке t_0 производную, то она непрерывна в этой точке. Следовательно, по третьему определению непрерывности функции в точке, $\Delta x \to 0$ при $\Delta t \to 0$. Но тогда устремится к нулю и $\alpha(\Delta x)$, т. е. будем иметь

$$\lim_{\Delta t \to 0} \left(\alpha(\Delta x) \frac{\Delta x}{\Delta t} \right) = \lim_{\Delta t \to 0} \alpha(\Delta x) \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = 0 \cdot \varphi'(t_0) = 0. \quad (4)$$

Из соотношения (4) следует существование предела всей правой части равенства (3) при $\Delta t \to 0$, равного $f'(x_0)\varphi'(t_0)$. Значит, существует предел при $\Delta t \to 0$ и левой части равенства (3), который, по определению производной, равен производной сложной функции $y=f[\varphi(t)]$. Тем самым доказана дифференцируемость сложной функции и установлена формула (1) ■

Замечание. В данной теореме мы рассматривали сложную функцию, где y зависит от t через промежуточную переменную x. Возможна и более сложная зависимость — с двумя, тремя и большим числом промежуточных переменных, но правило дифференцирования остается прежним.

Так, например, если $y=f(x)$, где $x=\varphi(u)$, а $u=\psi(v)$ и $v=\chi(t)$, то производную $y'(t)$ следует искать по формуле

$$y'(t)=y'(x)x'(u)u'(v)v'(t). \quad (5)$$

Рассмотрим примеры дифференцирования сложной функции.

Пример 1. Вычислить производную функции $y=e^{\operatorname{arc\,tg} x}$. Данную функцию можно представить в виде $y=e^u$, где $u=\operatorname{arc\,tg} x$. Тогда по формуле (1)

$$y'(x)=y'(u)u'(x)=e^u \frac{1}{1+x^2}.$$

Заменив u на $\operatorname{arc\,tg} x$, окончательно получим

$$y'=e^{\operatorname{arc\,tg} x} \frac{1}{1+x^2}.$$

Пример 2. Вычислить производную функции $y=\operatorname{tg}^2\sqrt{x^2+1}$. Данную функцию можно представить в виде $y=u^2$, где $u=\operatorname{tg} v$, а $v=\sqrt{w}$ и $w=x^2+1$. Используя формулу (5), получим

$$y'(x)=y'(u)\,u'(v)\,v'(w)\,w'(x)=(u^2)'(\operatorname{tg} v)'(\sqrt{w})'(x^2+1)'=$$
$$=2u\sec^2 v\,\frac{1}{2\sqrt{w}}\,2x=\frac{2x\operatorname{tg}\sqrt{x^2+1}\sec^2\sqrt{x^2+1}}{\sqrt{x^2+1}}.$$

§ 9. Логарифмическая производная. Производная степенной функции с любым вещественным показателем. Таблица производных простейших элементарных функций

1. Понятие логарифмической производной функции. Вычислим производную функции $y=\ln|x|$ $(x\neq 0)$. Так как $(\ln x)'=\frac{1}{x}$, а $(\ln(-x))'=\frac{(-x)'}{-x}=\frac{1}{x}$ (последнее равенство получено на основании правила дифференцирования сложной функции), то производная функции выражается следующей формулой:

$$y'=(\ln|x|)'=\frac{1}{x}. \tag{1}$$

С учетом полученной формулы (1) вычислим производную сложной функции $y=\ln|u|$, где $u=f(x)$ дифференцируемая функция. Имеем

$$y'=(\ln|u|)'=\frac{1}{u}u'=\frac{f'(x)}{f(x)},$$

или

$$(\ln|f(x)|)'=\frac{f'(x)}{f(x)}, \tag{2}$$

$(\ln|f(x)|)'$ и называется *логарифмической производной функции* $f(x)$. Для упрощения записи при логарифмическом дифференцировании знак модуля у функции $f(x)$ можно опустить.

В качестве примера с помощью логарифмической производной вычислим производную *показательно-степенной функции* $y=u(x)^{v(x)}$, где u и v — некоторые функции от $x (u>0)$, имеющие в данной точке производные $u'(x)$ и $v'(x)$.

Так как $\ln y = v(x) \ln u(x)$, то по формуле (2) получим

$$\frac{y'}{y} = [v(x)\ln u(x)]' = v'(x)\ln u(x) + v(x)\frac{u'(x)}{u(x)}.$$

Учитывая, что $y = u(x)^{v(x)}$, получаем следующую формулу для производной показательно-степенной функции:

$$y' = u(x)^{v(x)}\left[v'(x)\ln u(x) + v(x)\frac{u'(x)}{u(x)}\right]. \qquad (3)$$

Пример. Вычислить производную функции $y = x^x$.

Решение. Данную функцию можно представить в виде: $y = u(x)^{v(x)}$, где $u(x) = x$ и $v(x) = x$. Используя формулу (3), получим

$$y' = x^x\left[1\cdot\ln x + x\frac{1}{x}\right] = x^x(\ln x + 1).$$

Логарифмическая производная очень удобна при нахождении производной степенной функции с любым вещественным показателем.

2. Производная степенной функции с любым вещественным показателем. Производная функции $y = x^\alpha$ (α — любое вещественное число) выражается формулой

$$y' = \alpha x^{\alpha-1}.$$

Доказательство. Прологарифмировав равенство $y = x^\alpha$, получим

$$\ln y = \alpha \ln x.$$

По формуле (2) находим логарифмическую производную

$$\frac{y'}{y} = [\alpha\ln x]' = \frac{\alpha}{x}.$$

Отсюда, учитывая, что $y = x^\alpha$, получаем формулу для производной степенной функции:

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

Таким образом, нами вычислены производные всех простейших элементарных функций и мы можем составить следующую таблицу.

3. Таблица производных простейших элементарных функций

1°. $(C)' = 0$.

2°. $(x^\alpha)' = \alpha x^{\alpha-1}$, в частности $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$,

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}}.$$

3°. $(\log_a x)' = \frac{1}{x}\log_a e$, в частности $(\ln x)' = \frac{1}{x}$.

4°. $(a^x)' = a^x \ln a$, в частности $(e^x)' = e^x$.

5°. $(\sin x)' = \cos x$.

6°. $(\cos x)' = -\sin x$.

7°. $(\operatorname{tg} x)' = \frac{1}{\cos^2 x}$.

8°. $(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}$.

9°. $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$.

10. $(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$.

11°. $(\operatorname{arctg} x)' = \frac{1}{1+x^2}$.

12°. $(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}$.

Указанная таблица вместе с правилами дифференцирования суммы, разности, произведения, частного и правилом дифференцирования сложной функции составляет основу дифференциального исчисления.

Из правил и формул дифференцирования можно сделать один важный вывод: производная любой элементарной функции есть также функция элементарная. Таким образом, операция дифференцирования не выводит нас из класса элементарных функций.

§ 10. Производные и дифференциалы высших порядков

1. Понятие производной n-го порядка. Как уже отмечалось в § 1 данной главы, производная $f'(x)$ функции $y=f(x)$ сама является некоторой функцией аргумента x. Следовательно, по отношению к ней снова можно ставить вопрос о существовании и нахождении производной.

Назовем $f'(x)$ производной первого порядка.

Производная от производной некоторой функции называется *производной второго порядка* (или *второй производной*). Производная от второй производной называется *производной третьего порядка* (или *третьей производной*) и т. д. Производные, начиная со второй, называются производными высшего порядка и обозначаются:

$$y'',\ y''',\ y^{(4)},\ y^{(5)},\ \ldots,\ y^{(n)},\ \ldots,$$

или

$$y''_{x^2},\ y'''_{x^3},\ y^{(4)}_{x^4},\ y^{(5)}_{x^5},\ \ldots,\ y^{(n)}_{x^n},\ \ldots,$$

или

$$f''(x),\ f'''(x),\ f^{(4)}(x),\ f^{(5)}(x),\ \ldots,\ f^{(n)}(x),\ \ldots,$$

или также

$$\frac{d^2y}{dx^2},\ \frac{d^3y}{dx^3},\ \frac{d^4y}{dx^4},\ \frac{d^5y}{dx^5},\ \ldots,\ \frac{d^ny}{dx^n},\ \ldots$$

Производная n-го порядка есть производная от производной $(n-1)$-го порядка, т. е. $y^{(n)}=(y^{(n-1)})'$.

Производные высших порядков имеют широкое применение в физике. Здесь мы ограничимся физическим толкованием второй производной: $f''(x)$. Если функция $y=f(x)$ описывает закон движения материальной точки по прямой линии, то, как известно, первая производная $f'(x)$ есть мгновенная скорость точки в момент времени x, а вторая производная в таком случае равна скорости изменения скорости, т. е. равна ускорению движущейся точки в момент времени x.

2. n-е производные некоторых функций

1°. Вычислим n-ю производную степенной функции $y=x^\alpha\,(x>0)$ (α — любое вещественное число). Последовательно дифференцируя, будем иметь

$$y' = \alpha x^{\alpha-1},\quad y^{(2)} = \alpha(\alpha-1)x^{\alpha-2},\quad y^{(3)} = \alpha(\alpha-1)(\alpha-2)x^{\alpha-3}, \ldots,$$
$$\ldots, y^{(n)} = \alpha(\alpha-1)(\alpha-2)\ldots[\alpha-(n-1)]x^{\alpha-n}.$$

В частном случае, если $\alpha = m$, где m — натуральное число, получим

$$(x^m)^{(m)} = m(m-1)(m-2)\ldots[m-(m-1)]\cdot 1 = m!,$$
$$(x^m)^{(n)} = 0 \text{ при } n > m.$$

Нетрудно заметить, что, зная общий вид n-й производной, можно сразу записать производную любого порядка, не вычисляя при этом предшествующие производные.

Например, $(x^3)^{(3)} = 3\cdot 2\cdot 1 = 6$, $(x^3)^{(2)} = 3\cdot 2\cdot x = 6x$, а $(x^3)^{(4)} = 0$.

2°. Вычислим n-ю производную показательной функции $y = a^x$ $(0 < a \neq 1)$. Последовательно дифференцируя, будем иметь

$$y' = a^x \ln a,\ y^{(2)} = a^x(\ln a)^2,$$
$$y^{(3)} = a^x(\ln a)^3, \ldots, y^{(n)} = a^x(\ln a)^n.$$

В частности, если $y = e^x$, то

$$(e^x)^{(n)} = e^x$$

для любого n.

3°. Вычислим n-ю производную функции $y = \sin x$. Последовательно дифференцируя, будем иметь

$$y' = \cos x = \sin\left(x + \frac{\pi}{2}\right),$$

$$y^{(2)} = -\sin x = \sin(x + \pi) = \sin\left(x + 2\frac{\pi}{2}\right),$$

$$y^{(3)} = -\cos x = \sin\left(x + 3\frac{\pi}{2}\right), \ldots, y^{(n)} = \sin\left(x + n\frac{\pi}{2}\right).$$

Таким образом, производную любого порядка от $\sin x$ можно вычислять по формуле

$$(\sin x)^{(n)} = \sin\left(x + n\frac{\pi}{2}\right).$$

Например,

$$(\sin x)^{(10)} = \sin\left(x + 10\frac{\pi}{2}\right) = \sin(x + \pi) = -\sin x.$$

4°. Совершенно аналогично получается формула n-й производной функции $y=\cos x$:

$$(\cos x)^{(n)} = \cos\left(x + n\frac{\pi}{2}\right).$$

3. Формула Лейбница для n-й производной произведения двух функций. Пусть $y=uv$, где u и v — некоторые функции от переменной x, имеющие производные любого порядка. Тогда

$$y'=u'v+uv',$$

$$y''=u''v+u'v'+u'v'+uv''=u''v+2u'v'+uv'',$$

$$y'''=u'''v+u''v'+2u''v'+2u'v''+u'v''+uv''' =u'''v + +3u''v'+3u'v''+uv'''.$$

Мы видим, таким образом, что правые части разложений напоминают разложения различных степеней бинома $(a+b)^n$ по формуле Ньютона, где вместо показателей степени стоят числа, указывающие на порядок производных, а сами функции u и v можно рассматривать как «производные нулевого порядка» $u^{(0)}$ и $v^{(0)}$. Учитывая это, запишем по аналогии общий вид n-й производной произведения двух функций:

$$y^{(n)} = (u\cdot v)^{(n)} = u^{(n)}v + nu^{(n-1)}v' + \frac{n(n-1)}{2!}u^{(n-2)}v'' + \ldots +$$

$$+ \ldots + \frac{n(n-1)\ldots(n-k+1)}{k!}u^{(n-k)}v^{(k)} + \ldots + uv^{(n)}. \quad (1)$$

Формула (1) носит название *формулы Лейбница**. Докажем эту формулу методом математической индукции.

При $n=1$ эта формула принимает вид $(u\cdot v)'=$ $=u'v+uv'$, что совпадает с формулой дифференцирования произведения двух функций. Для $n=2$ и 3 она также проверена. Поэтому достаточно, предположив справедливость формулы (1) для некоторого n, доказать ее справедливость для $n+1$. С этой целью продифференцируем эту формулу, т. е. найдем $y^{(n+1)}=(y^{(n)})'$:

---

* Лейбниц Готфрид Вильгельм (1646—1716) — немецкий философ и математик.

$$y^{(n+1)} = u^{(n+1)}v + u^{(n)}v' + n\left[u^{(n)}v' + u^{(n-1)}v''\right] +$$

$$+ \frac{n(n-1)}{2!}\left[u^{(n-1)}v'' + u^{(n-2)}v'''\right] + \ldots +$$

$$+ \frac{n(n-1)\ldots(n-k+1)}{k!}\left[u^{(n-k+1)}v^{(k)} + u^{(n-k)}v^{(k+1)}\right] + \ldots$$

$$\ldots + u'v^{(n)} + uv^{(n+1)}.$$

Раскрывая скобки и приводя подобные члены, получим

$$y^{(n+1)} = u^{(n+1)}v + (n+1)u^{(n)}v' + \left(n + \frac{n(n-1)}{2!}\right)u^{(n-1)}v'' + \ldots$$

$$\ldots + \left[\frac{n(n-1)\ldots(n-k+2)}{(k-1)!} + \right.$$

$$\left. + \frac{n(n-1)\ldots(n-k+1)}{k!}\right]u^{(n-k+1)}v^{(k)} + \ldots$$

$$\ldots + (n+1)u'v^{(n)} + uv^{(n+1)}.$$

Но выражение, стоящее в квадратных скобках, мы можем представить как

$$\frac{n(n-1)\ldots(n-k+2)}{(k-1)!} + \frac{n(n-1)\ldots(n-k+1)}{k!} =$$

$$= \frac{n(n-1)(n-2)\ldots(n-k+2)(n-k+1)\times}{(k-1)!(n-k+1)\cdot(n-k)\times}$$

$$\frac{\times(n-k)(n-k-1)(n-k-2)\ldots\cdot 1}{\times(n-k-2)\ldots\cdot 1} +$$

$$+ \frac{n(n-1)\ldots(n-k+1)(n-k)(n-k-1)(n-k-2)\ldots\cdot 1}{k!(n-k)(n-k-1)(n-k-2)\ldots\cdot 1} =$$

$$= \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} =$$

$$= \frac{n!}{(k-1)!(n-k!)}\left(\frac{1}{n-k+1} + \frac{1}{k}\right) =$$

$$= \frac{n!}{(k-1)!(n-k)!} \cdot \frac{n+1}{k(n-k+1)} = \frac{(n+1)!}{k!(n-k+1)!} =$$

$$= \frac{(n+1)\cdot n(n-1)(n-2)\ldots(n-k+2)(n-k+1)\times}{k!(n-k+1)(n-k)\times}$$

$$\frac{\times (n-k)(n-k-1)(n-k-2)\ldots\cdot 1}{\times (n-k-1)(n-k-2)\ldots\cdot 1} =$$
$$= \frac{(n+1)\,n\,(n-1)(n-2)\ldots(n-k+2)}{k!}.$$

Тогда

$$y^{(n+1)} = u^{(n+1)}v + (n+1)\,u^{(n)}v' + \frac{(n+1)\,n}{2!}\,u^{(n-1)}v'' + \ldots$$
$$+ \frac{(n+1)\,n\,(n-1)\ldots(n-k+2)}{k!}\,u^{(n-k+1)}v^{(k)} + \ldots$$
$$\ldots + (n+1)\,u'\,v^{(n)} + uv^{(n+1)}.$$

Формула (1) доказана полностью ■

Рассмотрим примеры.

Пример 1. Вычислить пятую производную от функции $y=x^5e^x$.

Решение. Полагая $u=x^5$ и $v=e^x$, найдем

$$u'=5x^4,\ u''=20x^3,\ u'''=60x^2,\ u^{(4)}=120x,\ u^{(5)}=120;$$
$$v'=v''=v'''=v^{(4)}=v^{(5)}=e^x.$$

Подставляя при $n=5$ в формулу (1), получим

$$y^{(5)} = 120\cdot e^x + 5\cdot 120\,xe^x + \frac{5\cdot 4}{1\cdot 2}\,60\,x^2e^x +$$
$$+ \frac{5\cdot 4\cdot 3}{1\cdot 2\cdot 3}\,20x^3e^x + 5\cdot 5\,x^4e^x + x^5\,e^x =$$
$$= e^x\,(120 + 600\,x + 600\,x^2 + 200\,x^3 + 25\,x^4 + x^5).$$

Пример 2. Вычислить n-ю ($n\geqslant 2$) производную от функции $y=x^2\cos x$.

Решение. Полагая $u=\cos x$ и $v=x^2$, найдем

$$u^{(n)} = \cos\left(x + n\frac{\pi}{2}\right),\ v' = 2x,\ v'' = 2,$$
$$v''' = v^{(4)} = v^{(5)} = \ldots = 0.$$

Подставляя в формулу (1), получим

$$y^{(n)} = \cos\left(x + n\frac{\pi}{2}\right)x^2 + 2n\cos\left[x + (n-1)\frac{\pi}{2}\right]x +$$
$$+ \frac{2n\,(n-1)}{1\cdot 2}\cos\left[x + (n-2)\frac{\pi}{2}\right].$$

4. Дифференциалы высших порядков. Перейдем теперь к рассмотрению дифференциалов высших порядков. Для удобства мы будем вместо dy, dx писать равнозначные выражения δy, δx.

Пусть функция $f(x)$ дифференцируема в некоторой точке x. Тогда, как известно, ее дифференциал

$$dy=f'(x)\,dx,$$

который назовем дифференциалом первого порядка, является функцией двух переменных: аргумента x и переменной dx. Пусть функция $f'(x)$, в свою очередь, дифференцируема в некоторой точке x. Будем рассматривать dx в выражении для dy как постоянный множитель. Тогда функция dy представляет собой функцию только аргумента x и ее дифференциал имеет вид

$$\delta(dy)=\delta[f'(x)\,dx].$$

Дифференциал $\delta(dy)$ от дифференциала dy в некоторой точке x, взятый при $\delta x=dx$, называется дифференциалом второго порядка функции $f(x)$ и обозначается d^2y. В свою очередь, дифференциал $\delta(d^2y)$ от дифференциала d^2y, взятый при $\delta x=dx$, называется дифференциалом третьего порядка функции $f(x)$ и обозначается d^3y и т. д. При этом дифференциал независимой переменной dx при дифференцировании по x рассматривается как постоянная. Дифференциал $\delta(d^{(n-1)}y)$ от дифференциала $d^{n-1}y$, взятый при $\delta x=dx$, называется дифференциалом n-го порядка (или n-м дифференциалом) функции $f(x)$ и обозначается d^ny.

Таким образом, мы получаем

$$dy=y'dx,$$

$$d^2y=\delta(dy)=\delta(y'dx)=[y'dx]'\delta x=y''dx\delta x=y''(dx)^2,$$

$$d^3y=\delta(d^2y)=\delta[y''(dx)^2]=[y''(dx)^2]'\delta x= =y'''(dx)^2\delta x=y'''(dx)^3,$$

. . . .

$$d^ny=\delta(d^{n-1}y)=\delta[y^{n-1}(dx)^{n-1}]'\delta x=y^{(n)}(dx)^{n-1}\delta x= =y^{(n}(dx)^n.$$

Итак, для n-го дифференциала функции $f(x)$ справедлива формула

$$d^ny=y^{(n)}(dx)^n,\ n=1, 2, \ldots,$$

из которой следует, что для любого n справедливо равенство

$$y^{(n)}=\frac{d^ny}{(dx)^n} \text{ или } y^{(n)}=\frac{d^ny}{dx^n},$$

т. е. n-я производная функция $f(x)$ в некоторой точке x равна отношению n-го дифференциала этой функции в точке x к дифференциалу аргумента в степени n.

Пример 3. Вычислить дифференциал d^3y от функции $y=x^4-3x^2+4$.

Решение. Последовательно дифференцируя, получим

$$dy=y'dx=(4x^3-6x)dx,$$

$$d^2y=d(dy)=d[(4x^3-6x)dx] = [(4x^3-6x)dx]'dx = (12x^2-6)(dx)^2,$$

$$d^3y=d(d^2y)=d[(12x^2-6)(dx)^2] = [12x^2-6)(dx)^2]'dx = 24x(dx)^3.$$

§ 11. Параметрическое представление функции и ее дифференцирование

Пусть даны две функции

$$x=\varphi(t),\ y=\psi(t)$$

одной независимой переменной t, определенные и непрерывные в одном и том же промежутке. Если $x=\varphi(t)$ строго монотонна, то существует обратная к ней функция $t=\Phi(x)$, также непрерывная и строго монотонная. Поэтому y можно рассматривать как переменную, зависящую от переменной x через посредство переменной t, называемой *параметром.*

Положив

$$y=\psi[\Phi(x)],$$

мы видим, что функция $\psi[\Phi(x)]$ непрерывна. В этом случае говорят, что функция y от x задана *параметрически* при помощи переменной t. Рассмотрим примеры.

Пример 1. Пусть $x=R\cos t$, $y=R\sin t$ $(0\leqslant t\leqslant\pi)$.

Так как функция $x=R\cos t$ строго убывает для $0\leqslant t\leqslant\pi$, то, определив t из первого уравнения и подставив во второе, получим искомую функцию переменной x.

Еще проще придем к цели, если заметим, что

$$x^2+y^2=R^2(\cos^2 t+\sin^2 t)=R^2;$$

отсюда

$$y=\sqrt{R^2-x^2}.$$

Так как функция $y=R\sin t$ неотрицательна для $0\leqslant t\leqslant\pi$, то выбираем положительный знак у корня.

Взяв $\pi\leqslant t\leqslant 2\pi$, получим

$$y=-\sqrt{R^2-x^2}.$$

Мы видим, таким образом, что когда t меняется от 0 до 2π, то формулы $x=R\cos t$ и $y=R\sin t$ определяют две функции переменной x, графики которых образуют целую окружность.

Пример 2. Пусть $x=a\cos t$, $y=b\sin t$ $(0\leqslant t\leqslant 2\pi)$.

Нетрудно понять, что данные равенства являются параметрическими уравнениями эллипса, если вспомнить (см. замечание п. 1, § 7, гл. III), что эллипс получается из уравнения окружности радиуса a путем ее сжатия в a/b раз вдоль оси Oy. Из примера 1 следует, что параметрическими уравнениями окружности x^2+ $+y^2=a^2$ являются равенства $x=a\cos t$, $y=a\sin t$, $0\leqslant t\leqslant 2\pi$. Отсюда ясно, что параметрические уравнения эллипса получаются из параметрических уравнений окружности умножением ординаты y на b/a и имеют вид $x=a\cos t$, $y=b\sin t$, $0\leqslant t\leqslant 2\pi$. А можно поступить совсем просто. Исключая из этих уравнений параметр t (разрешая их относительно $\cos t$ и $\sin t$, возводя полученные равенства в квадрат и складывая), получим:

$$\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2=\cos^2 t+\sin^2 t=1, \text{ или } \frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

— уравнение эллипса.

Параметрическое представление функции имеет особо важное значение при изучении движения точки. Если точка движется на плоскости, то ее координаты x, y являются функциями времени t. Задав эти функции:

$$x=\varphi(t),\ y=\psi(t),$$

мы полностью определим движение точки. В каждом промежутке времени, в котором функция $\varphi(t)$ строго монотонна, можно, поступая как раньше, определить функцию $y=\psi[\Phi(x)]$, графиком которой будет кривая, описанная за этот промежуток времени движущейся точкой. В последнем примере функции описывали равномерное движение точки по эллипсу.

Предположим теперь, что функции $x=\varphi(t)$ и $y=\psi(t)$ имеют производные, причем $\varphi'(t)\neq 0$ на некотором промежутке. Из последнего неравенства вытекает, как мы увидим далее, строгая монотонность функции $x=\varphi(t)$ (см. теорему 6.7 гл. VI) и, следовательно, существование обратной функции $t=\Phi(x)$ в указанном промежутке. По теореме 5.4 о производной обратной функции эта функция имеет производную, выражающуюся по формуле

$$\Phi'(x)=\frac{1}{\varphi'(t)},$$

а по теореме 5.5 о производной сложной функции функция $y=\psi[\Phi(x)]$ имеет производную

$$y_x'=\psi'(\Phi(x))\Phi'(x),$$

следовательно,

$$y_x'=\frac{\psi'(t)}{\varphi'(t)}, \text{ или, короче, } y_x'=\frac{y_t'}{x_t'}.$$

Таким образом, мы доказали, что производная функции, представленной параметрически, выражается формулой

$$y_x'=\frac{y_t'}{x_t'}, \qquad (1)$$

если $\varphi'(t)\neq 0$.

Пример 1. Найти y'_x, если $x=R\cos t$, $y=R\sin t$ $(0<t<\pi)$.

Решение. По формуле (1) получаем

$$y_x'=\frac{R\cos t}{-R\sin t}=-\operatorname{ctg} t \quad (t\neq 0,\ \pi).$$

Пример 2. Найти y'_x, если $x=2t+t^2$, $y=t^2-2t^3$.

Решение. По формуле (1) получаем

$$y_x'=\frac{2t-6t^2}{2+2t}=\frac{2t(1-3t)}{2(1+t)}=\frac{t(1-3t)}{1+t}.$$

Пусть существуют вторые производные у функций $\varphi'(t)$ и $\psi'(t)$ в некоторой точке t. Тогда можно вычислить вторую производную функции, представленной параметрически. Заметим, что функция $y'_x=\frac{\psi'(t)}{\varphi'(t)}$, в свою очередь, задана параметрически функциями

$$y'_x=\frac{\psi'(t)}{\varphi'(t)}=\psi_1(t) \text{ и } x=\varphi(t).$$

Поэтому по формуле (1) получим

$$y''_{xx}=(y'_x)'_x=\frac{\psi'_1(t)}{\varphi'(t)}=\frac{\left(\frac{\psi'(t)}{\varphi'(t)}\right)'}{\varphi'(t)}=\frac{\frac{\psi''(t)\,\varphi'(t)-\varphi''(t)\,\psi'(t)}{(\varphi'(t))^2}}{\varphi'(t)}=$$

$$=\frac{\psi''(t)\,\varphi'(t)-\varphi''(t)\,\psi'(t)}{[\varphi'(t)]^3}.$$

Здесь нам пришлось применить правило дифференцирования частного.

Итак, мы получили

$$y''_{xx}=\frac{\psi''(t)\,\varphi'(t)-\varphi''(t)\,\psi'(t)}{[\varphi'(t)]^3},$$

или, короче,

$$y''_{xx}=\frac{y''_t x'_t-x''_t y'_t}{(x'_t)^3}. \qquad (2)$$

Аналогичным путем можно получить производную от y по x любого порядка.

Пример 3. Найти y''_{xx}, заданной параметрически функциями: $x=\cos t$, $y=\sin t$ $(0\leqslant t\leqslant\pi)$.

Решение. $y'_t=\cos t$, $y''_t=-\sin t$; $x'_t=-\sin t$, $x''_t=$ $=-\cos t$. Подставляя в формулу (2), находим

$$y''_{xx}=\frac{(-\sin t)(-\sin t)-(-\cos t)(\cos t)}{(-\sin t)^3}=$$

$$=\frac{\sin^2 t+\cos^2 t}{(-\sin t)^3}=-\frac{1}{\sin^3 t}.$$

Глава VI. ПРИМЕНЕНИЕ ДИФФЕРЕНЦИАЛЬНОГО ИСЧИСЛЕНИЯ К ИССЛЕДОВАНИЮ ФУНКЦИЙ

В предыдущей главе мы познакомились с дифференцированием функций. Теперь мы можем перейти к более глубокому изучению функций, вскрыть связи между отдельными свойствами функций и их производных. С помощью первой и второй производных изучим более точные и совершенные методы исследования функций и построения графиков, которые очень широко используются как в теории, так и в практике.

Данной главой мы завершим курс дифференциального исчисления.

§ 1. Основные теоремы дифференциального исчисления

1. Теорема 6.1 (теорема Ферма*). *Пусть функция $f(x)$ определена на интервале (a, b) и в некоторой точке x_0 этого интервала принимает наибольшее или наименьшее значение. Тогда, если в точке x_0 существует конечная производная, то она равна нулю, т. е. $f'(x_0)=0$.*

Доказательство. Пусть для определенности функция $f(x)$ в точке x_0 принимает наибольшее значение, т. е. $f(x)\leqslant f(x_0)$ для любого $x\in(a, b)$. Это значит, что для любой точки $x_0+\Delta x\in(a, b)$ $\Delta y=f(x_0+\Delta x)-f(x_0)\leqslant 0$. Тогда, если $\Delta x>0\,(x>x_0)$, то $\Delta y/\Delta x\leqslant 0$, и следовательно,

$$f'_+(x_0)=\lim_{\Delta x\to 0+}\frac{\Delta y}{\Delta x}\leqslant 0,$$

если же $\Delta x<0\,(x<x_0)$, то $\dfrac{\Delta y}{\Delta x}\geqslant 0$ и, следовательно,

$$f'_-(x_0)=\lim_{\Delta x\to 0-}\frac{\Delta y}{\Delta x}\geqslant 0,$$

т. е. правая производная неположительна, а левая неотрицательна. По условию $f'(x_0)$ существует. Это возможно только в случае, когда $f_+'(x_0)=f_-'(x_0)=0$. Но тогда и $f'(x_0)=0$.

* Ферма Пьер (1601—1665) — французский математик.

Аналогично рассматривается случай, когда в точке x_0 $f(x)$ принимает наименьшее значение ■

Геометрический смысл теоремы Ферма состоит в том, что если в точке x_0 $f(x)$ имеет наибольшее или наименьшее значение и $f'(x_0)=0$, то в точке $(x_0; f(x_0))$ касательная к графику функции $f(x)$ параллельна оси Ox (рис. 63).

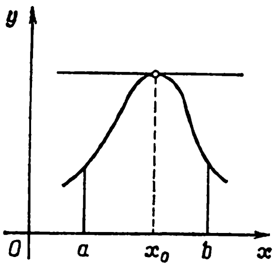


Рис. 63

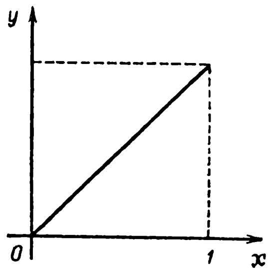


Рис. 64

Замечание. Теорема становится неверной, если функцию $f(x)$ рассматривать на отрезке $[a, b]$. Так, например, функция $f(x)=x$ на отрезке $[0,1]$ в точке $x=0$ принимает наименьшее, а в точке $x=1$ — наибольшее значения, однако как в той, так и другой точке производная в нуль не обращается, а равна единице (рис. 64).

Необходимо заметить, что теорема включает и те случаи, когда $f(x)=f(x_0)$ для нескольких значений или для всех значений $x\in(a, b)$.

II. Теорема 6.2 (теорема Ролля *). *Пусть на $[a, b]$ определена функция $f(x)$, причем: 1) $f(x)$ непрерывна на $[a, b]$, 2) $f(x)$ дифференцируема на (a, b), 3) $f(a)=f(b)$. Тогда существует точка $c\in(a, b)$, в которой $f'(c)=0$.*

Доказательство. Так как функция $f(x)$ непрерывна на $[a, b]$, то по второй теореме Вейерштрасса она достигает на этом отрезке своего максимального значения M и своего минимального значения m, т. е. существуют такие точки $x_1, x_2\in[a, b]$, в которых $f(x_1)=m$ и $f(x_2)=M$ и выполняются неравенства

$$m\leqslant f(x)\leqslant M.$$

* Ролль Мишель (1652—1719) — французский математик.

Могут представиться два случая:
1) $M=m$, 2) $m<M$.

В случае 1) $M=m=f(x)=\text{const}$. Поэтому производная $f'(x)$ равна нулю в любой точке $[a, b]$, и теорема доказана.

В случае 2) $m<M$, поскольку $f(a)=f(b)$, тогда хотя бы одно из двух значений m или M не принимается на концах отрезка $[a, b]$ т. е. существует точка $c\in(a, b)$, в которой $f(x)$ принимает наибольшее или наименьшее значение на интервале (a, b). В этом случае, поскольку $f(x)$ дифференцируема в точке c, из теоремы Ферма следует, что $f'(c)=0$ ■

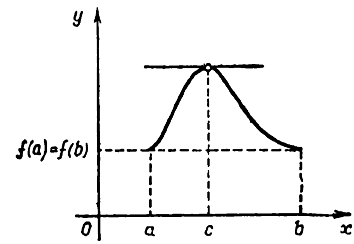


Рис. 65

Геометрически теорема Ролля означает, что у графика непрерывной на отрезке $[a, b]$ и дифференцируемой внутри него функции, принимающей на концах этого отрезка равные значения, существует точка $(c; f(c))$, в которой касательная параллельна оси Ox (рис. 65). В данном случае в точке c $f(x)$ принимает наибольшее значение M.

Необходимо отметить, что все три условия теоремы Ролля существенны. Чтобы убедиться в этом, достаточно привести примеры функций, для которых выполнялись бы два условия теоремы, а третье бы не выполнялось. Так, например, функция $f(x)=x$, $x\in[0, 1]$ (см. рис. 64) удовлетворяет условиям 1) и 2), но не удовлетворяет условию 3), и для нее не существует точки c такой, что $f'(c)=0$. Осталось проверить еще два случая. Предлагаем это сделать самостоятельно. Заметим только, что в математике существенность тех или иных условий доказываемых теорем проверяется построением соответствующих примеров, когда, разумеется, это возможно сделать.

III. Теорема 6.3 (теорема Лагранжа*). *Пусть на* $[a, b]$ *определена функция* $f(x)$, *причем:* 1) $f(x)$ *непрерывна на* $[a, b]$, 2) $f(x)$ *дифференцируема на* (a, b).

---

* Лагранж Жозеф-Луи (1736—1813) — французский математик.

Тогда существует точка $c \in (a, b)$ такая, что справедлива формула

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

Доказательство. Введем в рассмотрение на $[a, b]$ следующую вспомогательную функцию:

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

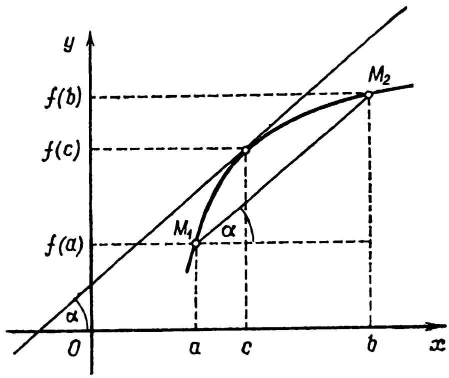


Рис. 66

Функция $F(x)$ удовлетворяет всем трем условиям теоремы Ролля:

1) $F(x)$ непрерывна на $[a, b]$ (как разность двух непрерывных функций $f(x)$ и линейной функции $f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$);

2) $F(x)$ дифференцируема на (a, b), т. е. внутри $[a, b]$ имеет производную, равную $F'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$;

3) $F(a) = 0$ и $F(b) = 0$, т. е. $F(a) = F(b)$.

Следовательно, по теореме Ролля существует точка $c \in (a, b)$ такая, что $F'(c) = 0$, т. е. $f'(c) - \frac{f(b)-f(a)}{b-a} = 0$.

Откуда получаем $\quad f'(c) = \frac{f(b)-f(a)}{b-a}$ ■

Выясним геометрический смысл теоремы Лагранжа (рис. 66). Величина $\frac{f(b)-f(a)}{b-a}$ есть угловой коэффициент секущей, проходящей через точки $M_1(a; f(a))$ и $M_2(b; f(b))$ кривой $y=f(x)$, а $f'(c)$ есть угловой коэффициент касательной к кривой $y=f(x)$ в точке $(c; f(c))$, т. е. существует точка c такая, что касательная к кривой в точке $(c; f(c))$ параллельна секущей M_1M_2. Таких точек может быть и несколько, но, по крайней мере, одна всегда существует.

Замечание 1. Равенство

$$f(b)-f(a)=f'(c)(b-a),\quad a<c<b, \tag{1}$$

называется *формулой Лагранжа*, или *формулой конечных приращений*.

Замечание 2. Так как точка c лежит между a и b, то можно записать

$$c=a+\theta(b-a),\text{ где } 0<\theta<1.$$

Здесь $\theta(b-a)$ является частью длины отрезка $[a, b]$. Учитывая это, формулу Лагранжа можно записать в виде

$$f(b)-f(a)=f'(a+\theta(b-a))(b-a),\text{ где } 0<\theta<1.$$

Замечание 3. Если положить $a=x$, $b=x+\Delta x$, то получим

$$f(x+\Delta x)-f(x)=f'(x+\theta\Delta x)\Delta x,\text{ где } 0<\theta<1.$$

Такая запись формулы Лагранжа часто бывает удобнее, чем запись в виде (1).

Как мы увидим в дальнейшем, теорема Лагранжа лежит в основе многих формул и теорем анализа.

IV. Теорема 6.4 (теорема Коши). *Пусть функции $f(x)$ и $g(x)$ непрерывны на $[a, b]$ и дифференцируемы на (a, b). Пусть, кроме того, $g'(x)\neq 0$. Тогда существует точка $c\in(a, b)$ такая, что справедлива формула*:

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(c)}{g'(c)}. \tag{2}$$

Доказательство. Покажем сначала, что $g(b)\neq g(a)$, т. е. что формула (2) имеет смысл. Действительно, если допустить, что $g(b)=g(a)$, то по теореме Ролля для функции $g(x)$ найдется точка $\xi\in(a, b)$, в которой $g'(\xi)=0$. А это противоречит условию $g'(x)\neq 0$ на (a, b).

Рассмотрим на $[a, b]$ вспомогательную функцию

$$F(x)=f(x)-f(a)-\frac{f(b)-f(a)}{g(b)-g(a)}[g(x)-g(a)].$$

Нетрудно заметить, что $F(x)$ на $[a, b]$ удовлетворяет условиям теоремы Ролля. В самом деле, $F(x)$ непрерывна на $[a, b]$, дифференцируема на (a, b), и, кроме того, подстановка $x=a$ и $x=b$ дает $F(a)=0$ и $F(b)=0$, т. е. $F(a)=F(b)$. По теореме Ролля для $F(x)$ существует точка c, $a<c<b$, такая, что $F'(c)=0$.

Так как

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

то

$$F'(c)=f'(c)-\frac{f(b)-f(a)}{g(b)-g(a)}g'(c)=0.$$

Учитывая, что $g'(c)\neq 0$, получаем формулу (2) ■

Замечание. Формула (2) называется *формулой Коши*, или *обобщенной формулой конечных приращений.*

§ 2. Раскрытие неопределенностей. Правило Лопиталя

1. Раскрытие неопределенности вида $\frac{0}{0}$. Будем говорить, что отношение двух функций $\frac{f(x)}{g(x)}$ при $x\to a$ есть неопределенность вида $\frac{0}{0}$, если

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

Раскрыть эту неопределенность — значит вычислить предел $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$, если он существует, или установить, что он не существует. Следующая теорема дает правило для раскрытия данной неопределенности.

Теорема 6.5 (теорема Лопиталя *). *Пусть функции $f(x)$ и $g(x)$ определены и дифференцируемы в некоторой окрестности точки a, за исключением, быть*

---

* Лопиталь Гильом Франсуа (1661—1704) — французский математик.

может, самой точки a. Пусть, далее, $\lim\limits_{x\to a} f(x)=\lim\limits_{x\to a} g(x)=0$ * *и $g'(x)\neq 0$ в указанной окрестности точки a. Тогда, если существует предел отношения производных $\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}$ (конечный или бесконечный), то существует и предел $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}$, причем справедлива формула*

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

Доказательство. Пусть $\{x_n\}$ — произвольная последовательность значений аргумента, сходящаяся к

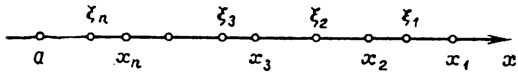


Рис. 67

точке a. Доопределим функции $f(x)$ и $g(x)$ в точке a, положив их равными нулю, т. е. $f(a)=g(a)=0$. Тогда, очевидно, $f(x)$ и $g(x)$ будут непрерывны на $[a, x_n]$, дифференцируемы на (a, x_n) и, по условию, $g'(x)\neq 0$. Таким образом, для $f(x)$ и $g(x)$ выполнены все условия теоремы Коши на $[a, x_n]$, т. е. внутри $[a, x_n]$ существует точка ξ_n такая, что

$$\frac{f(x_n)-f(a)}{g(x_n)-g(a)}=\frac{f'(\xi_n)}{g'(\xi_n)},\quad \xi_n\in(a, x_n).$$

По нашему доопределению $f(a)=g(a)=0$, следовательно,

$$\frac{f(x_n)}{g(x_n)}=\frac{f'(\xi_n)}{g'(\xi_n)},\quad \xi_n\in(a, x_n). \tag{1}$$

Пусть теперь в формуле (1) $n\to\infty$. Тогда, очевидно, $\xi_n\to a$ при $n\to\infty$ (рис. 67). Так как $\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}$ существует, то правая часть формулы (1) имеет предел при $n\to\infty$, равный данному, т. е. $\lim\limits_{n\to\infty}\dfrac{f'(\xi_n)}{g'(\xi_n)}=\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}$.

---

* Теорема остается справедливой и в случае, когда $x\to a-$ и $x\to a+$.

Следовательно, существует предел при $n \to \infty$ и левой части формулы (1), равный

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

а так как $\{x_n\}$ — произвольная последовательность значений аргумента, то окончательно получаем, что

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \quad \blacksquare$$

Доказанную теорему обычно называют *правилом Лопиталя.*

З а м е ч а н и е 1. Если производные $f'(x)$ и $g'(x)$ удовлетворяют тем же требованиям, что и сами функции $f(x)$ и $g(x)$, то правило Лопиталя можно *применить повторно.* При этом получим

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

З а м е ч а н и е 2. Теорема остается верной и в случае, когда $x \to \infty$, $x \to +\infty$, $x \to -\infty$. В самом деле, пусть, например, $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = 0$ и $\lim\limits_{x\to\infty} \dfrac{f'(x)}{g'(x)}$ существует (конечный или бесконечный). Сделаем подстановку $x = 1/t$, тогда $t \to 0$ при $x \to \infty$ и

$$f(x) = f\left(\frac{1}{t}\right) \to 0, \quad g(x) = g\left(\frac{1}{t}\right) \to 0 \text{ при } t \to 0,$$

т. е. функции $f\left(\dfrac{1}{t}\right)$, $g\left(\dfrac{1}{t}\right)$ удовлетворяют условиям доказанной теоремы.

Применяя к функциям $f\left(\dfrac{1}{t}\right)$ и $g\left(\dfrac{1}{t}\right)$ теорему 6.5 и правило дифференцирования сложной функции, получим

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{t\to 0} \frac{f\left(\frac{1}{t}\right)}{g\left(\frac{1}{t}\right)} = \lim_{t\to 0} \frac{f'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)}{g'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)} =$$

$$= \lim_{t\to 0} \frac{f'\left(\frac{1}{t}\right)}{g'\left(\frac{1}{t}\right)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}.$$

Рассмотрим примеры:

1) $$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\lim_{x\to 0}\frac{\sin x}{2x}=\frac{1}{2}\lim_{x\to 0}\frac{\sin x}{x}=$$
$$=\frac{1}{2}\cdot 1=\frac{1}{2};$$

2) $$\lim_{x\to 0}\frac{x-\sin x}{x^3}=\lim_{x\to 0}\frac{1-\cos x}{3x^2}=\lim_{x\to 0}\frac{\sin x}{6x}=$$
$$=\frac{1}{6}\lim_{x\to 0}\frac{\sin x}{x}=\frac{1}{6}\cdot 1=\frac{1}{6};$$

3) $$\lim_{x\to 0}\frac{e^x-1}{x}=\lim_{x\to 0}\frac{e^x}{1}=\lim_{x\to 0}e^x=1.$$

2. Раскрытие неопределенности вида $\frac{\infty}{\infty}$. Будем говорить, что отношение двух функций $\frac{f(x)}{g(x)}$ при $x\to a$ есть неопределенность вида $\frac{\infty}{\infty}$, если

$$\lim_{x\to a}f(x)=\lim_{x\to a}g(x)=\infty,\ +\infty \text{ или } -\infty.$$

Для раскрытия этой неопределенности справедливо утверждение, аналогичное теореме 6.5, а именно: если в формулировке теоремы заменить требование

$\lim\limits_{x\to a}f(x)=\lim\limits_{x\to a}g(x)=0$ на условие $\lim\limits_{x\to a}f(x)=\lim\limits_{x\to a}g(x)=\infty$,

то теорема останется справедливой.

Рассмотрим примеры:

1) $$\lim_{x\to+\infty}\frac{\ln x}{x}=\lim_{x\to+\infty}\frac{\frac{1}{x}}{1}=\lim_{x\to+\infty}\frac{1}{x}=0;$$

2) $$\lim_{x\to+\infty}\frac{x^n}{e^x}=\lim_{x\to+\infty}\frac{n\,x^{n-1}}{e^x}=\lim_{x\to+\infty}\frac{n\,(n-1)\,x^{n-2}}{e^x}=\ldots$$
$$\ldots=\lim_{x\to+\infty}\frac{n!}{e^x}=0;$$

3) $\lim\limits_{x\to 0+} \sqrt{x}\ln x = \lim\limits_{x\to 0+} \dfrac{\ln x}{x^{-1/2}} = \lim\limits_{x\to 0+} \dfrac{\frac{1}{x}}{\left(-\frac{1}{2}\right) x^{-3/2}} =$

$$= -2\lim_{x\to 0+} \sqrt{x} = 0.$$

3. Другие виды неопределенностей и их раскрытие. Неопределенности вида $0\cdot\infty$ и $\infty-\infty$ можно свести к неопределенностям $\frac{0}{0}$ и $\frac{\infty}{\infty}$. Покажем это на примерах.

Пример. 1. Найти $\lim\limits_{x\to 0+} x\ln x$.

Решение. Имеем неопределенность вида $0\cdot\infty$. Но $x\ln x = \dfrac{\ln x}{\frac{1}{x}}$, и мы получили неопределенность вида $\frac{\infty}{\infty}$. Применяя правило Лопиталя, получим

$$\lim_{x\to 0+} x\ln x = \lim_{x\to 0+} \frac{(\ln x)'}{\left(\frac{1}{x}\right)'} = \lim_{x\to 0+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim_{x\to 0+} x = 0.$$

Пример 2. Найти $\lim\limits_{x\to \frac{\pi}{2}} (\sec x - \operatorname{tg} x)$.

Решение. Имеем неопределенность вида $\infty-\infty$. Но $\sec x - \operatorname{tg} x = \dfrac{1}{\cos x} - \dfrac{\sin x}{\cos x} = \dfrac{1-\sin x}{\cos x}$, и мы получили при том же условии $x\to\frac{\pi}{2}$ неопределенность вида $\frac{0}{0}$. Применяя правило Лопиталя, получим

$$\lim_{x\to \frac{\pi}{2}} (\sec x - \operatorname{tg} x) = \lim_{x\to \frac{\pi}{2}} \left(\frac{1-\sin x}{\cos x}\right) =$$

$$= \lim_{x\to \frac{\pi}{2}} \frac{-\cos x}{-\sin x} = 0.$$

И наконец, рассмотрим неопределенности вида 0^0, 1^∞, ∞^0. Такие неопределенности возникают при рассмотрении функции $y = f(x)^{g(x)}$, если при $x \to a$ $f(x)$ стремится соответственно к 0, 1 и ∞, а $g(x)$ стремится соответственно к 0, ∞ и 0. Эти неопределенности с помощью очевидного тождества

$$f(x)^{g(x)} = e^{g(x)\ln f(x)}$$

преобразуются к неопределенности вида $0\cdot\infty$, с которой мы уже знакомы.

Пример 3. Найти $\lim\limits_{x\to 0+} x^x$.

Решение. Имеем неопределенность вида 0^0. Но $x^x = e^{x\ln x}$, и мы получили в показателе неопределенность вида $0\cdot\infty$, которая нами уже была рассмотрена (см. пример 1). Следовательно,

$$\lim_{x\to 0+} x^x = \lim_{x\to 0+} e^{x\ln x} = e^{\lim\limits_{x\to 0+} x\ln x} = e^0 = 1.$$

Пример 4. Найти $\lim\limits_{x\to 0} (1+x^2)^{\frac{1}{e^x-1-x}}$.

Решение. Имеем неопределенность вида 1^∞. Но

$$(1+x^2)^{\frac{1}{e^x-1-x}} = e^{\frac{\ln(1+x^2)}{e^x-1-x}},$$

и мы в показателе имеем неопределенность вида $\frac{0}{0}$. Применяя правило Лопиталя, получим

$$\lim_{x\to 0} \frac{\ln(1+x^2)}{e^x-1-x} = \lim_{x\to 0} \frac{\frac{2x}{1+x^2}}{e^x-1} = \lim_{x\to 0} \frac{2x}{(e^x-1)(1+x^2)} =$$

$$= \lim_{x\to 0} \frac{2}{e^x(1+x^2)+(e^x-1)\,2x} = \frac{2}{1} = 2.$$

Следовательно,

$$\lim_{x\to 0} (1+x^2)^{\frac{1}{e^x-1-x}} = e^{\lim\limits_{x\to 0} \frac{\ln(1+x^2)}{e^x-1-x}} = e^2.$$

§ 3. Формула Тейлора

Мы подошли к изучению одной из главных формул математического анализа, имеющей многочисленные применения как в анализе, так и в смежных дисциплинах.

1. Теорема 6.6 (теорема Тейлора *). *Пусть функция $f(x)$ имеет в точке a и некоторой ее окрестности производные порядка $n+1$* **. *Пусть x — любое значение аргумента из указанной окрестности. Тогда между точками a и x найдется точка ξ такая, что справедлива следующая формула:*

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

$$\dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}. \qquad (1)$$

Доказательство. Обозначим через $\varphi(x, a)$ многочлен относительно x порядка n, стоящий в правой части формулы (1), т. е. положим

$$\varphi(x, a) = f(a) + \frac{f'(a)}{1!}(x-a) +$$

$$+ \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

(он называется многочленом Тейлора порядка n для функции $f(x)$).

Далее обозначим через $R_{n+1}(x)$ разность

$$R_{n+1}(x) = f(x) - \varphi(x, a).$$

Теорема будет доказана, если мы установим, что

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \quad a < \xi < x.$$

Фиксируем любое значение x из указанной окрестности. Для определенности считаем $x > a$. Обозначим через t переменную величину, $a \leqslant t \leqslant x$, и рассмотрим на

---

* Тейлор Брук (1685—1731) — английский математик.

** Отсюда следует, что сама функция $f(x)$ и ее производные до порядка n непрерывны.

отрезке $[a, x]$ вспомогательную функцию вида

$$F(t) = f(x) - \varphi(x, t) - \frac{(x-t)^{n+1} R_{n+1}(x)}{(x-a)^{n+1}}. \qquad (2)$$

Функция $F(t)$ удовлетворяет на $[a, x]$ всем условиям теоремы Ролля.

1) Полагая в (2) $t=a$, будем иметь

$$F(a) = f(x) - \varphi(x, a) - R_{n+1}(x) = R_{n+1}(x) - R_{n+1}(x) = 0.$$

Равенство $F(x)=0$ вытекает из формулы (2), т. е.

$$F(x) = f(x) - f(x) - \frac{f'(x)}{1!}(x-x) - \frac{f''(x)}{2!}(x-x)^2 - \ldots$$

$$\ldots - \frac{f^{(n)}(x)}{n!}(x-x)^n - \frac{(x-x)^{n+1} R_{n+1}(x)}{(x-a)^{n+1}} = 0.$$

Таким образом, выполнение условия $F(a)=F(x)=0$ проверено.

2) Из формулы (2) и из условий, наложенных на функцию $f(x)$, вытекает, что $F(t)$ непрерывна на $[a, x]$ и дифференцируема, ибо $f(t)$ и ее производные до порядка n непрерывны на $[a, x]$, а $f^{(n+1)}(t)$ существует и конечна на этом отрезке.

На основании теоремы Ролля внутри $[a, x]$ существует точка ξ такая, что

$$F'(\xi) = 0. \qquad (3)$$

Вычислим производную $F'(t)$. Дифференцируя равенство (2) по t, будем иметь

$$F'(t) = -f'(t) + \frac{f'(t)}{1!} - \frac{f''(t)}{1!}(x-t) + \frac{f''(t)}{2!}2(x-t) -$$

$$- \frac{f'''(t)}{2!}(x-t)^2 + \ldots + \frac{f^{(n)}(t)}{n!}n(x-t)^{n-1} -$$

$$- \frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{(n+1)(x-t)^n R_{n+1}(x)}{(x-a)^{n+1}}.$$

Нетрудно заметить, что все члены, за исключением двух последних, взаимно уничтожаются. Таким образом,

$$F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{(n+1)(x-t)^n R_{n+1}(x)}{(x-a)^{n+1}}. \qquad (4)$$

Полагая в (4) $t=\xi$ и используя равенство (3), получим

$$F'(\xi) = -\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n + \frac{(n+1)(x-\xi)^n R_{n+1}(x)}{(x-a)^{n+1}} = 0,$$

откуда

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \blacksquare$$

Формула (1) называется *формулой Тейлора*, а выражение для $R_{n+1}(x)$ — *остаточным членом в форме Лагранжа*. Его можно переписать в ином виде. Так как точка $\xi \in (a, x)$, то найдется такое число θ из интервала $0<\theta<1$, что $\xi - a = \theta(x-a)$, т. е. $\xi = a + \theta(x-a)$ и остаточный член примет вид

$$R_{n+1}(x) = \frac{f^{(n+1)}[a+\theta(x-a)]}{(n+1)!}(x-a)^{n+1}, \quad 0<\theta<1.$$

Эта форма остаточного члена наиболее употребительна в приложениях.

2. Другая запись формулы Тейлора и остаточного члена. Часто записывают формулу Тейлора (1) в несколько ином виде. Положим в (1) $a=x_0$, $x-a=\Delta x$, $x=x_0+\Delta x$. Тогда получим

$$f(x_0+\Delta x) - f(x_0) = \frac{f'(x_0)}{1!}\Delta x + \frac{f''(x_0)}{2!}(\Delta x)^2 + \ldots +$$

$$+ \ldots + \frac{f^{(n)}(x_0)}{n!}(\Delta x)^n + \frac{f^{(n+1)}(x_0+\theta\Delta x)}{(n+1)!}(\Delta x)^{n+1}, \quad 0<\theta<1. \tag{5}$$

При $n=0$ из (5) получается формула Лагранжа

$$f(x_0+\Delta x) - f(x_0) = f'(x_0+\theta\Delta x)\Delta x.$$

Покажем, что остаточный член $R_{n+1}(x)$ является бесконечно малой более высокого порядка, чем $(x-a)^n$ при $x \to a$:

$$\lim_{x\to a}\frac{R_{n+1}(x)}{(x-a)^n} = \lim_{x\to a}\frac{f^{(n+1)}(\xi)}{(n+1)!}\cdot\frac{(x-a)^{n+1}}{(x-a)^n} =$$

$$= \lim_{x\to a}\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a) = 0,$$

так как $f^{n+1}(\xi)$ ограничена, а $(x-a)\to 0$ при $x\to a$. Таким образом,

$$R_{n+1}(x)=o[(x-a)^n] \quad \text{при} \quad x\to a. \qquad (6)$$

Формула (6) называется остаточным членом *в форме Пеано**.

3. Формула Маклорена. Принято называть формулой Маклорена** формулу Тейлора (1) при $a=0$:

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+R_{n+1}(x).$$

Остаточный член имеет вид

1) в форме Лагранжа: $R_{n+1}(x)=\dfrac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}$, $0<\theta<1$;

2) в форме Пеано: $R_{n+1}(x)=o(x^n)$.

4. Разложение некоторых элементарных функций по формуле Маклорена.

1) $f(x)=e^x$. Поскольку

$$f(x)=f'(x)=f''(x)=\ldots=f^{n+1}(x)=e^x,$$
$$f(0)=f'(0)=f''(0)=\ldots=f^{n+1}(0)=1,$$

то для любых x и n формула Маклорена имеет вид

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\frac{e^{\theta x}}{(n+1)!}x^{n+1},\ 0<\theta<1. \qquad (7)$$

Если заменить функцию e^x ее многочленом Тейлора степени n, то получим приближенное равенство

$$e^x\approx 1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}, \qquad (8)$$

абсолютная погрешность которого равна

$$|R_{n+1}(x)|=\frac{e^{\theta x}}{(n+1)!}|x|^{n+1},\ 0<\theta<1.$$

Если рассматривать функцию e^x для $-1\leqslant x\leqslant 1$, то

$$|R_{n+1}(x)|\leqslant\frac{e^{\theta x}}{(n+1)!}<\frac{3}{(n+1)!}.$$

---

* Пеано Джузеппе (1858—1932) — итальянский математик.

** Маклорен Колин (1698—1746) — английский математик.

Полагая в (8) $x=1$, получим приближенное значение числа e:

$$e \approx 1 + 1 + \frac{1}{2!} + \ldots + \frac{1}{n!} = 2{,}7182\ldots$$

При этом абсолютная погрешность меньше $\frac{3}{(n+1)!}$. Если требуется вычислить значение e с точностью до 0,001, то число n определяется из неравенства $\frac{3}{(n+1)!} < 0{,}001$, или $(n+1)! > 3000$. Следовательно, если взять $n=6$, то требуемое неравенство удовлетворяется.

Формулу (7) можно записать с остаточным членом в форме Пеано:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + o(x^n). \qquad (9)$$

2) $f(x) = \sin x$. Поскольку

$$f^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right),$$

$$f^{(n)}(0) = \begin{cases} 0 & \text{при } n \text{ четном,} \\ (-1)^{\frac{n-1}{2}} & \text{при } n \text{ нечетном,} \end{cases}$$

то формула Маклорена имеет вид

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \qquad (10)$$

$$\ldots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!} + o(x^{2n}).$$

3) $f(x) = \cos x$. Поскольку

$$f^{(n)}(x) = \cos\left(x + n\frac{\pi}{2}\right),$$

$$f^{(n)}(0) = \begin{cases} 0 & \text{при } n \text{ нечетном,} \\ (-1)^{\frac{n}{2}} & \text{при } n \text{ четном,} \end{cases}$$

то формула Маклорена имеет вид

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$$
$$\ldots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1}). \qquad (11)$$

В формуле (10) мы написали остаточный член в виде $o(x^{2n})$, а не в виде $o(x^{2n-1})$, так как следующий за последним член равен нулю (то же самое относится к формуле (11)).

4) $f(x) = (1+x)^{\alpha}$, где α — вещественное число. Поскольку

$$f^{(n)}(x) = \alpha(\alpha-1)\ldots(\alpha-n+1)(1+x)^{\alpha-n},$$
$$f^{(n)}(0) = \alpha(\alpha-1)\ldots(\alpha-n+1),$$

то формула Маклорена имеет вид

$$(1+x)^{\alpha} = 1 + \frac{\alpha}{1!} x + \frac{\alpha(\alpha-1)}{2!} x^2 + \ldots$$
$$\ldots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} x^n + R_{n+1}(x), \qquad (12)$$

где остаточный член в форме Лагранжа равен

$$R_{n+1}(x) = \frac{\alpha(\alpha-1)\ldots(\alpha-n)}{(n+1)!} (1+\theta x)^{\alpha-(n+1)} x^{n+1}, \quad 0<\theta<1.$$

В частном случае, когда $\alpha=n$ — натуральное число, $R_{n+1}(x)=0$, и мы получим известную из элементарной математики *формулу бинома Ньютона*

$$(1+x)^n = 1 + \frac{n}{1!} x + \frac{n(n-1)}{2!} x^2 + \ldots + x^n.$$

Проведенные выше разложения показывают, что с помощью формулы Маклорена функции можно с определенной степенью точности заменять многочленами, являющимися наиболее простыми элементарными функциями. Над многочленами просто выполняются арифметические действия и дифференцирование, многочлен непрерывен в любой точке и т. д. Формулы Тейлора и Маклорена заменяют многочленами и более сложные функции, не допуская при этом больших погрешностей.

Помимо этого эти формулы имеют широкий круг приложений. Мы ограничимся рассмотрением одного.

5. Использование формулы Маклорена для вычисления пределов. Формула Тейлора является мощным средством для вычисления пределов функций, с которыми часто приходится сталкиваться при том или ином исследовании функций.

Рассмотрим примеры.

1) Найти $\lim\limits_{x\to 0}\frac{\sin x - x}{x^3}$. По формуле (10), взятой при $n=3$, будем иметь

$$\lim_{x\to 0}\frac{x-\frac{x^3}{3!}+o(x^4)-x}{x^3}=\lim_{x\to 0}\frac{-\frac{1}{3!}+\frac{o(x^4)}{x^3}}{1}=$$

$$=-\frac{1}{3!}+\lim_{x\to 0}\frac{o(x^4)}{x^3}=-\frac{1}{3!}+0=-\frac{1}{6}.$$

2) Найти $\lim\limits_{x\to 0}\frac{e^{\frac{-x^2}{2}}-\cos x}{x^3\sin x}$. По формулам (9), (10) и (11) будем иметь

$$\lim_{x\to 0}\frac{e^{\frac{-x^2}{2}}-\cos x}{x^3\sin x}=\lim_{x\to 0}\frac{1-\frac{x^2}{2}+\frac{x^4}{8}+o(x^4)-1+\frac{x^2}{2}-\frac{x^4}{24}}{x^3(x+o(x))}=$$

$$=\lim_{x\to 0}\frac{\frac{x^4}{8}-\frac{x^4}{24}+o(x^4)}{x^4+o(x^4)}=\lim_{x\to 0}\frac{\frac{1}{8}-\frac{1}{24}+\alpha(x)}{1+\alpha(x)}=$$

$$=\frac{1}{8}-\frac{1}{24}=\frac{1}{12}.$$

(Здесь символом $\alpha(x)$ мы обозначим величину $\frac{o(x^4)}{x^4}$, являющуюся бесконечно малой при $x\to 0$.)

§ 4. Геометрическое исследование поведения функций

Основное содержание данного параграфа — исследование поведения функций и построение графиков.

Особое внимание следует обратить на практическую сторону вопроса.

1. Признак монотонности функции.

Теорема 6.7. *Если функция $f(x)$ дифференцируема на интервале (a, b) и $f'(x) \geqslant 0$ ($f'(x) \leqslant 0$) на (a, b), то функция $f(x)$ не убывает (не возрастает на (a, b).*

Доказательство. Для определенности рассмотрим случай $f'(x) \geqslant 0$. Пусть x_1 и x_2 — две произвольные точки из (a, b) и $x_1 < x_2$, тогда на отрезке $[x_1, x_2]$ выполняются все условия теоремы Лагранжа, согласно которой имеем

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1), \quad c \in (x_1, x_2).$$

По условию $f'(c) \geqslant 0$, $x_2 - x_1 > 0$, поэтому $f(x_2) - f(x_1) \geqslant 0$ или $f(x_2) \geqslant f(x_1)$, т. е. функция $f(x)$ не убывает.

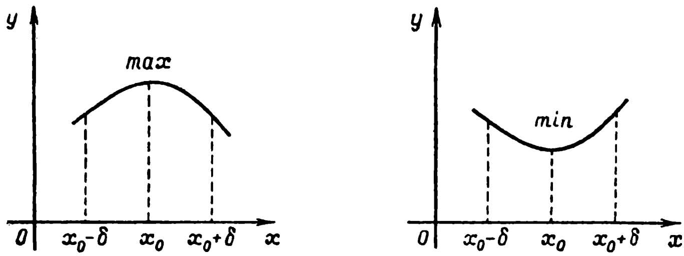


Рис. 68

Доказательство для случая $f'(x) \leqslant 0$ аналогично ■

Замечание. Из теоремы следует, что если $f'(x)$ на (a, b) строго >0 (<0), то $f(x)$ возрастает (убывает) на (a, b).

2. Отыскание точек локального экстремума функции.

Определение. Точка x_0 называется *точкой строгого локального максимума (минимума)* функции $f(x)$, если для всех x из некоторой δ-окрестности точки x_0 выполняется неравенство $f(x) < f(x_0)$ $(f(x) > f(x_0))$ при $x \neq x_0$ (рис. 68).

Локальный максимум (max) и локальный минимум (min) объединяются общим названием локальный *экстремум.*

Из определения следует, что понятие экстремума носит локальный характер в том смысле, что в случае экстремума неравенство $f(x)<f(x_0)$ ($f(x)>f(x_0)$) не обязано выполняться для всех значений x в области определения функции, а должно выполняться лишь в некоторой окрестности точки x_0. Очевидно, функция может иметь несколько локальных максимумов и несколько локальных минимумов, причем может так случиться, что иной максимум окажется меньше другого минимума.

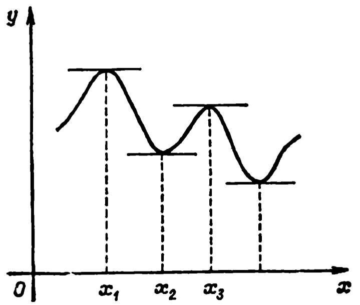


Рис. 69

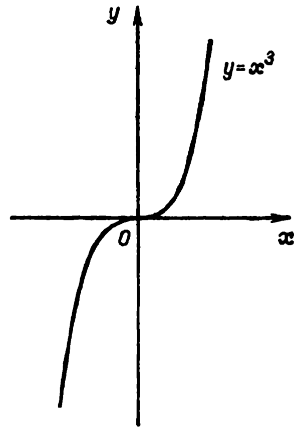


Рис. 70

Теорема 6.8 (необходимое условие локального экстремума). *Если функция $f(x)$ имеет в точке x_0 локальный экстремум и дифференцируема в этой точке, то* $f'(x_0)=0$.

Доказательство. Так как в точке x_0 функция $f(x)$ имеет локальный экстремум, то это означает, что существует такой интервал $(x_0-\delta,\ x_0+\delta)$, в котором значение $f(x_0)$ является наибольшим или наименьшим среди всех других значений этой функции. Тогда по теореме Ферма производная функции в точке x_0 должна быть равна нулю, т. е. $f'(x_0)=0$ ■

Теорема 6.8 имеет следующий геометрический смысл. Если точки x_1, x_2 и x_3 — точки локального экстремума и в соответствующих точках графика существуют касательные, то эти касательные параллельны оси Ox (рис. 69).

Иногда такие точки называют стационарными, мы же их будем называть *точками возможного экстремума*. Если точка x_0 — точка возможного экстремума, т. е. $f'(x_0)=0$, то она может и не быть точкой локального максимума или минимума. Например, $f(x)=x^3$, $f'(x)=3x^2=0$ при $x=0$, но тем не менее в точке $x=0$ нет локального экстремума (рис. 70). Поэтому мы их и назвали точками возможного экстремума, а условие $f'(x_0)=0$ является лишь *необходимым*. Установим достаточное условие существования локального экстремума. Этому посвящается следующая

Теорема 6.9 (достаточное условие локального экстремума). *Пусть функция $f(x)$ дифференцируема в некоторой окрестности точки x_0 $(x_0-\delta,\ x_0+\delta)$. Тогда, если для всех x из $(x_0-\delta,\ x_0)$ $f'(x)>0$ $(f'(x)<0)$, а для всех x из $(x_0,\ x_0+\delta)$ $f'(x)<0$ $(f'(x)>0)$, то в точке x_0 $f(x)$ имеет локальный максимум (локальный минимум), если же $f'(x)$ в окрестности точки x_0 имеет один и тот же знак, то в точке x_0 локального экстремума нет.*

Другими словами, если $f'(x)$ при переходе через точку x_0 меняет знак с + на —, то точка x_0 — точка локального максимума, если же $f'(x)$ в точке x_0 меняет знак с — на +, то точка x_0 — точка локального минимума, если же в точке x_0 $f'(x)$ знака не меняет, то в точке x_0 экстремума не существует.

Доказательство. Пусть $f'(x)$ при переходе через точку x_0 меняет знак с + на — и пусть $x\in(x_0-\delta, x_0)$. Применим формулу Лагранжа к $[x,\ x_0]$:

$$f(x_0)-f(x)=f'(c)(x_0-x),\quad c\in(x,\ x_0).$$

Так как на $(x_0-\delta,\ x_0)$ $f'(x)>0$, то $f'(c)>0$ и, кроме того, $x_0-x>0$, следовательно,

$$f(x_0)-f(x)>0,\ \text{или}\ f(x_0)>f(x). \tag{1}$$

Рассмотрим теперь интервал справа от точки x_0, т. е. $x\in(x_0,\ x_0+\delta)$. Применим формулу Лагранжа к $[x_0,\ x]$:

$$f(x)-f(x_0)=f'(c)(x-x_0),\quad c\in(x_0,\ x).$$

Так как на $(x_0,\ x_0+\delta)$ $f'(x)<0$, то $f'(c)<0$ и, кроме того, $x-x_0>0$, следовательно,

$$f(x)-f(x_0)<0,\ \text{или}\ f(x_0)>f(x). \tag{2}$$

Из неравенства (1) и (2) следует, что в рассматриваемой окрестности точки x_0 выполняется $f(x)<f(x_0)$, а это означает, что в точке x_0 $f(x)$ имеет локальный максимум.

Аналогично рассматривается случай перемены знака $f'(x)$ с — на +.

Осталось рассмотреть случай, когда $f'(x)$ знака не меняет.

Пусть $f'(x)>0$ в некоторой окрестности $(x_0-\delta, x_0+\delta)$, тогда по теореме 6.7 (по признаку монотонности) функция $f(x)$ не убывает на $(x_0-\delta, x_0+\delta)$, т. е. для любых $x<x_0$ выполняется $f(x)<f(x_0)$, а для любых $x>x_0$ выполняется $f(x)>f(x_0)$. А это означает, что точка x_0 не является точкой локального экстремума, так как при переходе через нее в данном случае не сохраняется знак разности $f(x)-f(x_0)$ в окрестности этой точки ■

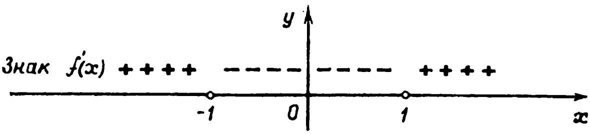


Рис. 71

Замечание. Теорема 6.9 остается справедливой, если функция $f(x)$ в самой точке x_0 не дифференцируема, а только непрерывна. Примером такой функции является $f(x)=|x|$, которая в точке $x=0$ непрерывна, но не дифференцируема.

В качестве примера рассмотрим вопрос об отыскании точек локального экстремума функции $f(x)=x^3-3x$. Находим производную: $f'(x)=3x^2-3=3(x^2-1)$. Решая уравнение $3(x^2-1)=0$, получаем две точки возможного экстремума: $x_1=-1$ и $x_2=1$. Дальнейшее исследование удобно вести, нарисовав вспомогательный чертеж (рис. 71). Отметив на нем точки $x_1=-1$ и $x_2=1$ и исследовав знак $f'(x)$ в окрестности этих точек, получаем: в точке $x_1=-1$ $f(x)$ имеет локальный максимум, а в точке $x_2=1$ — локальный минимум. Осталось найти y_{max} и y_{min}: $y_{max}=f(-1)=2$, $y_{min}=f(1)=-2$.

На рис. 71 ясны и интервалы монотонности: $(-\infty, -1)$, $(-1, 1)$ и $(1, +\infty)$, причем в первом и третьем из них функция возрастает, а во втором — убывает.

3. Направление выпуклости и точки перегиба графика функции. Пусть функция $f(x)$ дифференцируема в любой точке (a, b), тогда существует касательная к графику функции $f(x)$, проходящая через любую точку

$M(x; f(x))$ этого графика ($a<x<b$), причем эта касательная не параллельна оси Oy, ибо угловой коэффициент ее, равный $f'(x)$, конечен.

Определение 1. Будем говорить, что график функции $f(x)$ имеет на (a, b) выпуклость, направленную

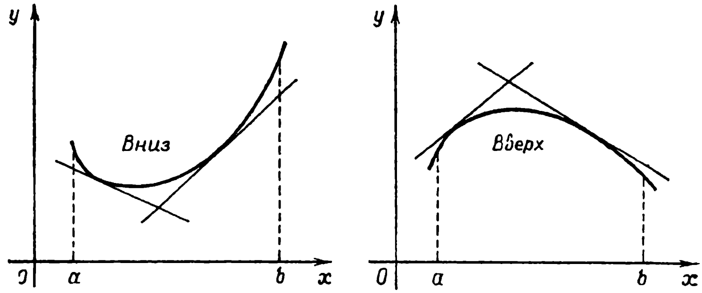


Рис. 72

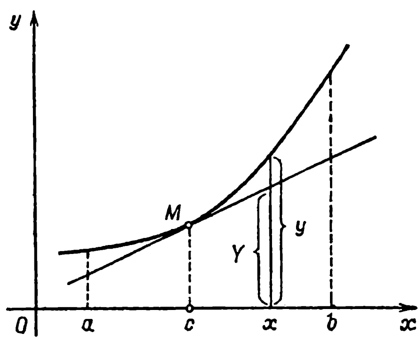


Рис. 73

вниз (*вверх*), если он расположен *не ниже* (*не выше*) любой касательной к графику функции на (a, b) (рис. 72).

Из определения следует, что на участке выпуклости касательные к графику функции не пересекаются с самим графиком и имеют с ним лишь точки касания.

Теорема 6.10. *Если функция $f(x)$ имеет на интервале (a, b) конечную $f''(x)$ и $f''(x)\geqslant 0$ ($f''(x)\leqslant 0$) во всех точках (a, b), то график функции $f(x)$ имеет выпуклость, направленную вниз (вверх) на (a, b).*

Доказательство. Для определенности рассмотрим случай $f''(x)\geqslant 0$ на (a, b). Обозначим через c любую точку (a, b) (рис. 73). Требуется доказать, что график $y=f(x)$ лежит не ниже касательной, проходящей через точку $M(c; f(c))$.

Запишем уравнение касательной, обозначая ее текущую ординату через Y:

$$Y-f(c)=f'(c)(x-c). \tag{3}$$

Разложим функцию $f(x)$ в окрестности точки c по формуле Тейлора с $n=1$. Получим:

$$y=f(x)=f(c)+\frac{f'(c)}{1!}(x-c)+$$

$$+\frac{f''(\xi)}{2!}(x-c)^2,\ \xi\in(c, x). \tag{4}$$

Поскольку по условию $f(x)$ имеет $f''(x)$ на (a, b), то согласно теореме Тейлора формула (4) справедлива для любого x из (a, b).

Сопоставляя текущие ординаты y и Y, соответствующие одной и той же абсциссе (см. рис. 73), с учетом формул (3) и (4) будем иметь

$$y-Y=\frac{f''(\xi)}{2!}(x-c)^2. \tag{5}$$

Так как по условию $f''(x)\geqslant 0$ на (a, b) всюду, то правая часть равенства (5) неотрицательна, т. е. для всех x из (a, b) $y-Y\geqslant 0$ или $y\geqslant Y$. Последнее неравенство и доказывает, что график $y=f(x)$ всюду в пределах (a, b) лежит не ниже касательной (3).

Аналогично доказывается теорема для случая $f''(x)\leqslant 0$ ■

Определение 2. Точка $M(x_0; f(x_0))$ называется *точкой перегиба графика* функции $f(x)$, если существует такая окрестность точки x_0, в пределах которой график функции $f(x)$ слева и справа от точки x_0 имеет разные направления выпуклости.

Очевидно, что в точке перегиба касательная пересекает график функции, так как с одной стороны от этой точки график лежит под касательной, а с другой — над нею, т. е. в окрестности точки перегиба график функции геометрически переходит с одной стороны касательной на другую и «перегибается» через нее. Отсюда и произошло название: «точка перегиба».

Теорема 6.11 (необходимое условие точки перегиба). *Пусть график функции $f(x)$ имеет перегиб в точке $M(x_0; f(x_0))$ и пусть функция $f(x)$ имеет в точке*

x_0 непрерывную вторую производную, тогда $f''(x)$ в точке x_0 обращается в нуль, т. е. $f''(x_0)=0$.

Доказательство. Предположим обратное, т. е. допустим, что $f''(x_0)\neq 0$. Тогда в силу непрерывности второй производной по теореме 4.7 об устойчивости знака непрерывной функции существует окрестность точки x_0, в которой $f''(x)<0$ (соответственно $f''(x)>0$), и, значит, согласно теореме 6.10 функция $f(x)$ имеет определенное направление выпуклости в этой окрестности,

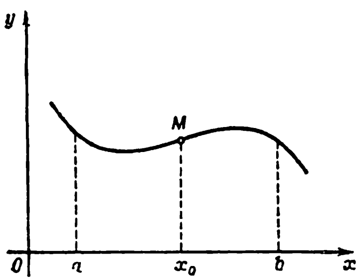


Рис. 74

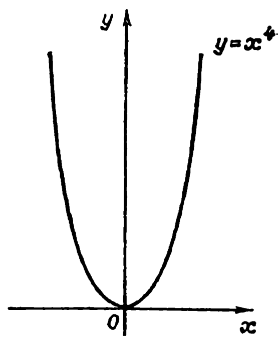


Рис. 75

а это противоречит наличию перегиба в точке $M(x_0; f(x_0))$ (рис. 74). Полученное противоречие доказывает теорему ■

Следует заметить, что не всякая точка $(x_0; f(x_0))$, для которой $f''(x_0)=0$, является точкой перегиба. Например, график функции $f(x)=x^4$ не имеет перегиба в точке $(0; 0)$, хотя $f''(x)=12x^2=0$ при $x=0$ (рис. 75). Поэтому равенство нулю второй производной является лишь *необходимым* условием перегиба. Такие точки $(x; f(x))$ графика, для которых $f''(x)=0$, будем называть *критическими*. Необходимо дополнительно исследовать вопрос о наличии перегиба в каждой критической точке, т. е. следует установить достаточное условие перегиба, к чему мы и переходим.

Теорема 6.12 (достаточное условие точки перегиба). *Пусть функция $f(x)$ имеет $f''(x)$ в некоторой окрестности точки x_0 и $f''(x)=0$. Тогда, если в пределах* *указанной окрестности $f''(x)$ имеет разные знаки слева и справа от точки x_0, то график $y=f(x)$ имеет перегиб в точке $M(x_0; f(x_0))$.*

Доказательство. Из того, что $f''(x)$ слева и справа от точки x_0 имеет разные знаки, и из теоремы 6.10 заключаем, что направление выпуклости слева и справа от точки x_0 различно, что говорит о наличии перегиба в точке $M(x_0; f(x_0))$ ■

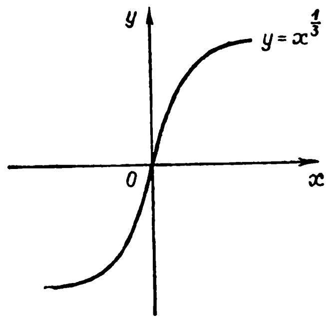


Рис. 76

Замечание. Теорема остается верной, если $f(x)$ имеет конечную $f''(x)$ в некоторой окрестности точки x_0, за исключением, быть может, самой точки x_0 (в ней $f''(x)$ может и не существовать). Тогда, если в пределах указанной окрестности $f''(x)$ имеет разные знаки слева и справа, то график функции $f(x)$ имеет перегиб в точке $M(x_0; f(x_0))$. Доказательство данного факта аналогично доказательству теоремы.

Рассмотрим пример: $f(x)=x^{\frac{1}{3}}$. Эта функция непрерывна на всей числовой прямой, но в точке $x=0$ уже первая производная $f'(x)=\frac{1}{3\sqrt[3]{x^2}}$ не является конечной. Однако график функции имеет перегиб в точке $(0; 0)$, так как вторая производная $f''(x)=-\frac{2}{9}\frac{1}{x^{5/3}}$ имеет слева и справа от точки $x=0$ разные знаки (рис. 76).

Итак, вопрос о направлении выпуклости и точках перегиба графика функции исследуется с помощью второй производной.

В качестве примера возьмем функцию $f(x)=x^3-3x$, которую мы уже начали рассматривать в п. 2, причем изучение знака второй производной можно вести на вспомогательном чертеже, изображенном на рис. 71. Находим вторую производную: $f''(x)=6x$. Из уравнения $6x=0$ получаем одну критическую точку: $(0; 0)$. Отметив точку $x=0$ на вспомогательном чертеже

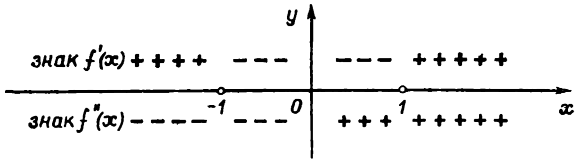


Рис. 77

(рис. 77) и исследовав знак $f''(x)$ в ее окрестности, получаем: слева от точки $x=0$ $f''(x)<0$ (выпуклость графика направлена вверх), а справа — $f''(x)>0$ (выпуклость графика направлена вниз), т. е. точка M (0; 0) является точкой перегиба графика рассматриваемой функции (рис. 78).

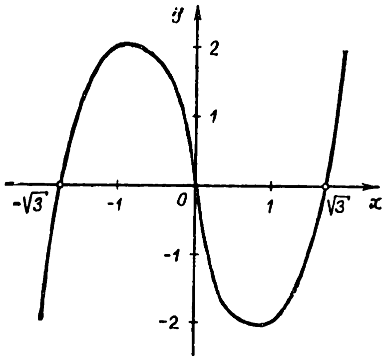


Рис. 78

Теперь покажем, что известная из аналитической геометрии кривая, называемая эллипсом (см. гл. III, § 6, п. 1), имеет на интервале $(0, a)$ выпуклость, направленную вверх. В самом деле, из уравнения эллипса находим: $y=\frac{b}{a}\sqrt{a^2-x^2}$. Найдем y' и y'': $y'=-\frac{b}{a}\times$ $\times\frac{x}{\sqrt{a^2-x^2}}$, $y''=-\frac{ba}{(a^2-x^2)^{3/2}}$. Из вида второй производной вытекает, что эта производная отрицательна на интервале $(0, a)$. Значит, данная кривая на всем интервале $(0, a)$ имеет выпуклость, направленную вверх (см. рис. 30).

Аналогичным образом можно показать, что другая известная кривая, называемая гиперболой, на интервале $(a, +\infty)$ имеет выпуклость, направленную вверх. Предлагаем читателю сделать это самостоятельно.

4. Асимптоты графика функции. При исследовании поведения функции на бесконечности, т. е. при $x \to +\infty$ и при $x \to -\infty$, или вблизи точек разрыва второго рода часто оказывается, что график функции сколь угодно близко приближается к той или иной прямой. Такие прямые называют *асимптотами*.

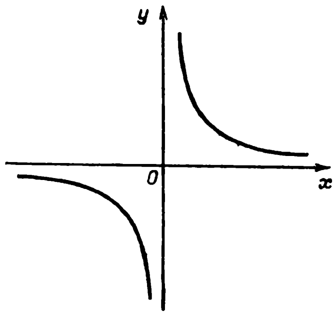


Рис. 79

Определение 1. Прямая линия называется *асимптотой* для кривой $y=f(x)$, если расстояние d от точки M, лежащей на кривой, до прямой стремится к нулю при удалении точки M от начала координат в бесконечность*. В таких случаях говорят, что кривая $y=f(x)$ асимптотически приближается к прямой.

Существует три вида асимптот: *вертикальные, горизонтальные* и *наклонные*.

Определение 2. Прямая $x=x_0$ называется *вертикальной асимптотой* графика функции $y=f(x)$, если хотя бы одно из предельных значений $\lim\limits_{x \to x_0+} f(x)$ или $\lim\limits_{x \to x_0-} f(x)$ равно $+\infty$ или $-\infty$.

Например, график функции $f(x)=\frac{1}{x}$ (рис. 79) имеет вертикальную асимптоту $x=0$, так как при $x \to 0+$ $y \to +\infty$ и при $x \to 0-$ $y \to -\infty$.

В этом случае расстояние точки $M(x; f(x))$ до прямой $x=x_0$

$$d=\sqrt{(x-x_0)^2+(f(x)-f(x))^2}=|x-x_0| \to 0 \text{ при } x \to x_0.$$

Определение 3. Если $\lim\limits_{\substack{x \to +\infty \\ (x \to -\infty)}} f(x)=A$, то пря-

---

* С понятием асимптоты мы уже встречались в аналитической геометрии на примере гиперболы (см. гл. III, § 6, п. 2).

мая $y=A$ называется *горизонтальной асимптотой* графика $y=f(x)$ при $x\to+\infty$ $(x\to-\infty)$.

Например, график рассмотренной выше функции $f(x)=\frac{1}{x}$ имеет горизонтальную асимптоту $y=0$, так как при $x\to+\infty$ и при $x\to-\infty$ $y\to 0$.

В этом случае расстояние точки $M(x;\,f(x))$ до прямой $y=A=0$ равно $d=\sqrt{(x-x)^2+(f(x)-A)^2}=|f(x)-A|\to 0$ при $x\to\infty$, так как $\lim\limits_{x\to\infty}|f(x)-A|=0$.

Определение 4. Если существуют такие числа k и b, что $\lim\limits_{\substack{x\to+\infty\\(x\to-\infty)}}\frac{f(x)}{x}=k$, $\lim\limits_{\substack{x\to+\infty\\(x\to-\infty)}}[f(x)-kx]=b$, то прямая $y=kx+b$ называется *наклонной асимптотой* графика функции $y=f(x)$ при $x\to+\infty$ $(x\to-\infty)$.

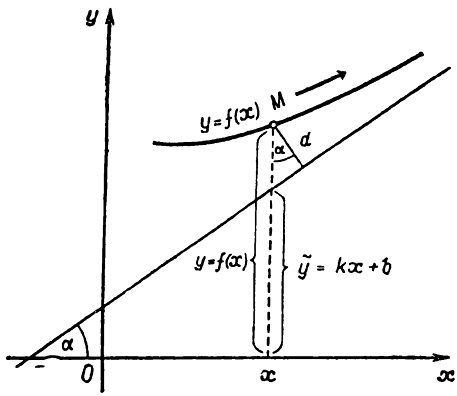


Рис. 80

Рассмотрим геометрический смысл наклонной асимптоты. Для определенности разберем случай, когда $x\to+\infty$ (случай $x\to-\infty$ рассматривается аналогично).

Пусть $M(x;\,f(x))$ — точка графика функции $y=f(x)$ и пусть далее, прямая $\tilde{y}=kx+b$ является асимптотой для графика функции $y=f(x)$. Обозначим ее текущую ординату через $\tilde{y}$ (рис. 80).

Это означает, что $d\to 0$ при $x\to+\infty$. Но $d=(y-\tilde{y})\cos\alpha$, где α — угол между асимптотой и осью

Ox, а следовательно, если $\lim\limits_{x\to+\infty} d=0$, то и $\lim\limits_{x\to+\infty}(y-\tilde{y})=0$ и наоборот ($\cos\alpha$ — постоянное число). Отсюда заключаем, что условие $d\to0$ равносильно условию $y-\tilde{y}\to0$, т. е. условию

$$\lim_{x\to+\infty}[y-(kx+b)]=0. \tag{6}$$

Таким образом, условие (6) является *необходимым* и *достаточным* для того, чтобы прямая $\tilde{y}=kx+b$ была асимптотой.

Определим теперь числа k и b. Из соотношения (6) следует, что выражение, стоящее в квадратных скобках, представляет собой бесконечно малую функцию при $x\to+\infty$: $f(x)-(kx+b)=\alpha(x)\to0$ при $x\to+\infty$. Разделив последнее равенство на x и перейдя к пределу, получаем

$$\lim_{x\to+\infty}\left[\frac{f(x)}{x}-k-\frac{b}{x}\right]=\lim_{x\to+\infty}\frac{\alpha(x)}{x}.$$

Так как $\lim\limits_{x\to+\infty}\frac{b}{x}=0$ и $\lim\limits_{x\to+\infty}\frac{\alpha(x)}{x}=0$, то $\lim\limits_{x\to+\infty}\left[\frac{f(x)}{x}-k\right]=0$, откуда

$$k=\lim_{x\to+\infty}\frac{f(x)}{x}. \tag{7}$$

Из соотношения (6), зная k, находим b:

$$b=\lim_{x\to+\infty}[f(x)-kx]. \tag{8}$$

Таким образом, если прямая $\tilde{y}=kx+b$ есть асимптота, то числа k и b находятся по формулам (7) и (8). Обратно, если оба предела (7) и (8) существуют и конечны, причем $\lim\limits_{x\to+\infty}\frac{f(x)}{x}\neq0$, то существует и наклонная асимптота.

Практически целесообразно искать асимптоты в следующем порядке:

1) вертикальные асимптоты;

2) горизонтальные асимптоты;

3) наклонные асимптоты.

Пример 1. Найти асимптоты для графика функции $y=\frac{x^2+2x-3}{x}$.

Решение. 1) Находим вертикальные асимптоты. Точка $x=0$ — точка разрыва второго рода. При $x \to 0-$ $y \to +\infty$; при $x \to 0+$ $y \to -\infty$. Следовательно, ось ординат $x=0$ есть вертикальная асимптота.

2) Находим горизонтальные асимптоты:

$$\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(\frac{x^2+2x-3}{x}\right) = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(x+2-\frac{3}{x}\right) = \underset{(-\infty)}{+\infty},$$

следовательно, горизонтальных асимптот нет.

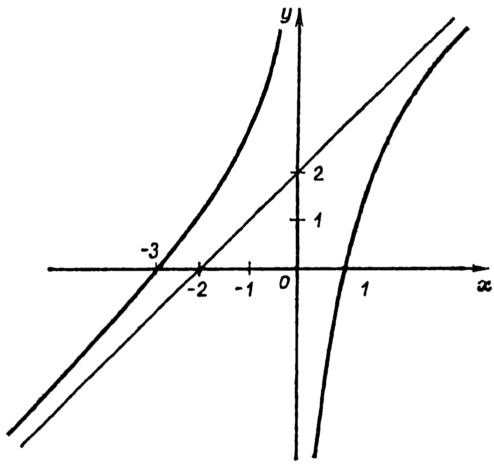


Рис. 81

3) Находим наклонные асимптоты:

$$k = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{f(x)}{x} = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(1+\frac{2}{x}-\frac{3}{x^2}\right) = 1,$$

$$b = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} [f(x)-kx] = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left[\frac{x^2+2x-3}{x}-x\right] =$$

$$= \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(\frac{2x-3}{x}\right) = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(2-\frac{3}{x}\right) = 2.$$

Следовательно, прямая $y=x+2$ является наклонной асимптотой графика заданной функции, причем оказалось, что k и b не зависят от того, стремится ли x к $+\infty$ или к $-\infty$. График функции изображен на рис. 81.

Пример 2. Показать, что гипербола $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ имеет своими асимптотами прямые $y=\pm\frac{b}{a}x$.

Решение. Так как $y=\pm\frac{b}{a}\sqrt{x^2-a^2}$, то

$$\lim_{x\to\infty}\frac{f(x)}{x}=\lim_{x\to\infty}\left[\pm\frac{b}{a}\sqrt{1-\left(\frac{a}{x}\right)^2}\right]=\pm\frac{b}{a};$$

$$\lim_{x\to\infty}[f(x)-kx]=\lim_{x\to\infty}\left[\pm\frac{b}{a}\sqrt{x^2-a^2}\mp\frac{b}{a}x\right]=$$

$$=\lim_{x\to\infty}\left[\pm\frac{b}{a}(\sqrt{x^2-a^2}-x)\right]=$$

$$=\pm\frac{b}{a}\lim_{x\to\infty}\frac{-a^2}{\sqrt{x^2-a^2}+x^2}=0.$$

Следовательно, прямые $y=\pm\frac{b}{a}x$ суть асимптоты для данной гиперболы.

5. Схема исследования графика функции. Изучение заданной функции и построение ее графика целесообразно проводить в следующем порядке.

1) Определить область существования функции.

2) Найти точки пересечения графика функции с осями координат.

3) Найти асимптоты.

4) Найти точки возможного экстремума.

5) Найти критические точки.

6) С помощью вспомогательного чертежа провести исследование знака первой и второй производных. Определить участки возрастания, убывания и направление выпуклости, найти точки экстремумов и точки перегиба.

7) Построить график.

При этом в начале исследования функции полезно установить, не является ли данная функция четной или нечетной, чтобы при построении использовать симметрию графика относительно оси ординат или начала координат.

В качестве примера построим график функции

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

по изложенной выше схеме.

1) Областью существования функции является множество вещественных чисел, кроме $x=1$, при котором обращается в нуль знаменатель.

2) Так как уравнение $x^2+1=0$ не имеет вещественных корней, график функции не имеет точек пересечения с осью Ox, но пересекает ось Oy в точке $(0; -1)$.

3) Выясним вопрос о существовании асимптот. Исследуем поведение функции вблизи точки разрыва. При $x \to 1-$ $y \to -\infty$; при $x \to 1+$ $y \to +\infty$. Таким образом, прямая $x=1$ является вертикальной асимптотой графика функции.

При $x \to +\infty$ $(x \to -\infty)$ $y \to +\infty$ $(y \to -\infty)$, т. е. горизонтальной асимптоты у графика нет. Далее, из существования пределов

$$k=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{f(x)}{x}=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{x^2+1}{x^2-x}=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{1+\frac{1}{x^2}}{1-\frac{1}{x^2}}=1,$$

$$b=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} [f(x)-kx]=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left[\frac{x^2+1}{x-1}-x\right]=$$

$$=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{1+x}{x-1}=\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{1+\frac{1}{x}}{1-\frac{1}{x}}=1$$

вытекает, что при $x \to +\infty$ и при $x \to -\infty$ график функции имеет наклонную асимптоту $y=x+1$.

4) Для нахождения точек возможного экстремума вычислим первую производную функции:

$$f'(x)=\frac{2x(x-1)-(x^2+1)}{(x-1)^2}=\frac{2x^2-2x-x^2-1}{(x-1)^2}=$$

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

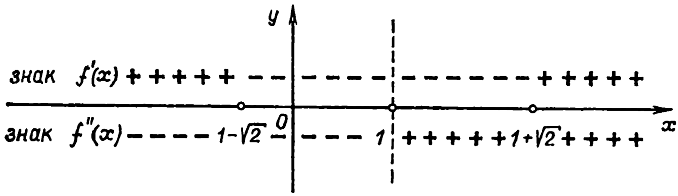
Решая уравнение $x^2-2x-1=0$, получаем две точки возможного экстремума: $x_1 = 1-\sqrt{2}$ и $x_2 = 1+\sqrt{2}$.

5) Для нахождения критических точек вычислим вторую производную:

$$f''(x) = \frac{(2x-2)(x-1)^2 - 2(x-1)(x^2-2x-1)}{(x-1)^4} = \frac{4}{(x-1)^3}.$$

Видим, что $f''(x)$ в нуль не обращается, следовательно, критических точек не существует.

6) Рисуем вспомогательный чертеж и исследуем на нем знак первой и второй производных.



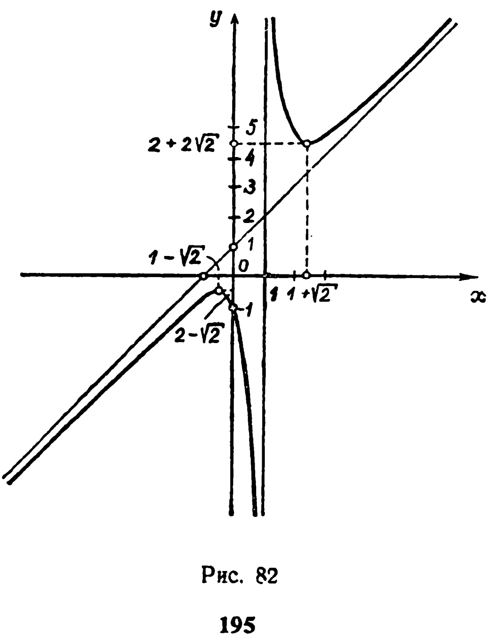


Рис. 82

Отсюда следует, что функция на $(-\infty, 1-\sqrt{2})$ возрастает, на $(1-\sqrt{2}, 1+\sqrt{2})$ убывает, а на $(1+\sqrt{2}, +\infty)$ снова возрастает. Точки укстремума: 1) максимум при $x=1-\sqrt{2}$, причем $f(1-\sqrt{2})=2-2\sqrt{2}$, 2) минимум при $x=1+\sqrt{2}$, причем $f(1+\sqrt{2})=2+2\sqrt{2}$. На $(-\infty, 1)$ направление выпуклости графика вверх, а на $(1, +\infty)$ — вниз, следовательно, точка $(1; 0)$ является точкой перегиба.

7) По полученным данным строим эскиз графика (рис. 82).

§ 5. Приближенные методы вычисления корней уравнений

Этим параграфом заканчивается изучение дифференциального исчисления функций одной переменной. В нем будет рассмотрен вопрос о приближенном вычислении одного из корней уравнения $f(x)=0$, где $f(x)$ — некоторая непрерывная функция.

Из элементарной математики мы знаем, как найти корни уравнения $f(x)=0$, если $f(x)$ есть линейная или квадратичная функция. Для функций более сложной природы обычно приходится прибегать к различным приближенным методам вычисления корней уравнений. Мы познакомимся с двумя методами: *методом «вилки»* и *методом касательных**.

1. Метод «вилки». Пусть уже известно, что интересующий нас корень $x=c$ уравнения $f(x)=0$ является внутренней точкой отрезка $[a, b]$, и других корней на $[a, b]$ нет. Относительно функции $f(x)$ мы предположим, что она непрерывна на $[a, b]$ и имеет на концах этого отрезка значения разных знаков. На практике обычно путем грубой прикидки находится такой отрезок.

Для определенности будем считать, что $f(a)<0$, $f(b)>0$. Разделим $[a, b]$ пополам и выберем тот из полученных отрезков, на концах которого $f(x)$ имеет разные знаки. Обозначим его $[a_1, b_1]$. (Если бы значение $f(x)$ в середине $[a, b]$ равно было бы нулю, то искомый корень был бы найден.) Разделим $[a_1, b_1]$ пополам и выберем тот из полученных отрезков, на концах которого $f(x)$ имеет разные знаки, и т. д. Продолжая этот процесс неограниченно, получим последовательность

* Этот метод называется также методом Ньютона.

вложенных отрезков

$$[a, b]\supset[a_1, b_1]\supset[a_2, b_2]\supset\ldots\supset[a_n, b_n]\supset\ldots,$$

обладающих тем свойством, что для любого n $f(a_n)<0$, $f(b_n)>0$. По теореме 2.13 о вложенных отрезках существует точка c, принадлежащая всем отрезкам и к которой сходится каждая из последовательностей $\{a_n\}$ и $\{b_n\}$.

Покажем, что c является искомым корнем, т. е. $f(c)=0$. Поскольку $f(x)$ непрерывна в точке c, то каждая из последовательностей $\{f(a_n)\}$ и $\{f(b_n)\}$ сходится к $f(c)$. Но тогда из условий $f(a_n)<0$, $f(b_n)>0$ в силу теоремы 2.10 получим, что одновременно справедливы неравенства $f(c)\leqslant 0$ и $f(c)\geqslant 0$, т. е. $f(c)=0$, что и требовалось.

Теперь легко сообразить, как получить интересующий нас корень $x=c$. За приближенное значение этого корня можно взять середину отрезка $[a_n, b_n]$, т. е. точку $\frac{a_n+b_n}{2}$. Поскольку длина $[a_n, b_n]$ равна $\frac{b-a}{2^n}$, то число $\frac{a_n+b_n}{2}$ отличается от точного значения корня не более чем на $\frac{b-a}{2^n}$. Таким образом, описанный выше метод отыскания корня уравнения $f(x)=0$ позволяет вычислить искомый корень c с любой точностью.

Изложенный метод удобен тем, что требует большого числа однотипных вычислительных операций. Поэтому им часто пользуются для проведения вычислений на современных быстродействующих вычислительных машинах.

Перейдем к рассмотрению другого метода.

2. Метод касательных. Метод касательных является одним из самых эффективных приближенных методов вычисления корней уравнения $f(x)=0$.

Пусть по-прежнему интересующий нас корень $x=c$ является внутренней точкой $[a, b]$. Предположим также, что на $[a, b]$ функция $f(x)$ имеет непрерывные производные $f'(x)$ и $f''(x)$ одного знака, а ее значения $f(a)$ и $f(b)$ — разных знаков. Постоянство знака $f'(x)$ указывает на то, что функция $f(x)$ на $[a, b]$ либо только возрастает, либо только убывает и в обоих случаях график функции $f(x)$ пересекает ось Ox только в одной точке, т. е. $x=c$ будет единственным корнем на $[a, b]$. Постоянство же знака $f''(x)$ указывает на то, что направление выпуклости графика функции $f(x)$ на этом отрезке не меняется.

Для определенности рассмотрим случай, когда $f'(x)>0$ и $f''(x)>0$ (рис. 83). Проведем через точку

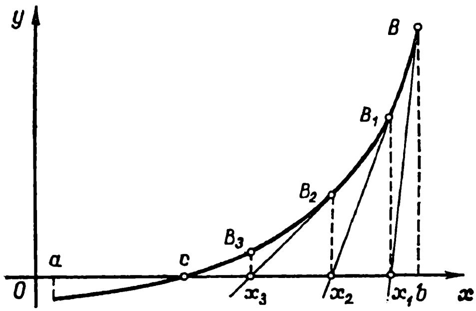


Рис. 83

$B(b;\ f(b))$ касательную к графику функции $f(x)$ и составим ее уравнение:

$$f(x)-f(b)=f'(b)\ (x-b).$$

Положив $f(x)=0$, найдем точку пересечения касательной с осью Ox:

$$x_1=b-\frac{f(b)}{f'(b)}.$$

Возьмем за первое приближенное значение корня точку x_1. Далее, приведем касательную через точку $B_1(x_1, f(x_1))$ и возьмем за второе приближенное значение корня точку x_2:

$$x_2=x_1-\frac{f(x_1)}{f'(x_1)}.$$

Продолжая этот процесс неограниченно, получим для любого n формулу

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}, \tag{1}$$

выражающую x_{n+1} через x_n. При этом получается последовательность приближенных значений корня c:

$$x_1>x_2>x_3>\ldots>x_n>x_{n+1}>\ldots>c. \tag{2}$$

Формула (1) является *основной* расчетной формулой метода касательных. Таким образом, метод касательных представляет собой метод последовательных приближений (или, как говорят, метод итераций), которые строятся при помощи формулы (1).

Покажем, что последовательность (2) сходится к искомому корню c, и дадим оценку погрешности, т. е. отклонения приближенного значения x_n от точного значения корня c. Действительно, последовательность (2) убывает и ограничена снизу числом c. Следовательно, по теореме 2.12 она имеет конечный предел $c' \geqslant c$. Далее, переходя к пределу в равенстве (1) с учетом непрерывности $f(x)$ и $f'(x)$, получим:

$$c' = c' - \frac{f(c')}{f'(c')},$$

откуда следует, что $f(c')=0$, т. е. c' есть корень уравнения $f(x)=0$. Но поскольку на $[a, b]$ имеется только один корень c, то $c'=c$.

Найдем теперь оценку отклонения n-го приближения x_n от точного значения корня c. Применяя к выражению $f(x_n)=f(x_n)-f(c)$ формулу Лагранжа, будем иметь $f(x_n)=(x_n-c)f'(\xi_n)$, где $c<\xi_n<x_n$. Отсюда получаем следующую оценку:

$$|x_n - c| \leqslant \frac{|f(x_n)|}{m}, \tag{3}$$

где m — наименьшее значение $|f'(x)|$ на отрезке $[a, b]$. Формула (3) позволяет оценить отклонение приближенного значения корня x_n от точного значения корня c через значение модуля функции $f(x)$ в точке x_n.

Осталось заметить, что мы рассмотрели случай, когда $f'(x)>0$ и $f''(x)>0$ на $[a, b]$. В зависимости от комбинации знаков $f'(x)$ и $f''(x)$ возможны еще три случая: 1) $f'(x)>0$ и $f''(x)<0$; 2) $f'(x)<0$ и $f''(x)>0$; 3) $f'(x)<0$ и $f''(x)<0$, в каждом из которых обоснование метода касательных проводится в полной аналогии с рассмотренным случаем.

Пример. Вычислить корень уравнения $x^2-5=0$ методом касательных.

Решение. Рассмотрим функцию $f(x)=x^2-5$. Эта функция непрерывна на всей прямой. Грубой прикидкой найдем отрезок, на концах которого $f(x)$ имеет значения разных знаков. Так как $f(1)=-4$, $f(2)=-1$, $f(3)=4$, то этим отрезком будет [2, 3], и внутри него находится искомый корень уравнения. Функция $f(x)$ имеет на этом отрезке непрерывные положительные производные $f'(x)=2x$ и $f''(x)=2$. Следовательно, касательную к графику функции $f(x)$ нужно проводить в точке (3; 4). Положив в формуле (1) $x_0=3$, получим первое приближение корня: $x_1 = 3 - \frac{4}{2\cdot 3} = 2\frac{1}{3}$. Положив теперь в формуле (1) $x_1 = 2\frac{1}{3}$, получим второе приближение корня: $x_2 = 2\frac{1}{3} - \frac{2}{21} = 2\frac{5}{21}$, и, наконец, положив $x_2 = 2\frac{5}{21}$ в формуле (1), получим третье приближение корня:

$$x_3 = 2\frac{2}{21} - \frac{2}{987} = 2{,}23607 \text{ и т. д.}$$

Подсчитаем, какова будет погрешность приближения, если остановиться на значении x_3. Воспользуемся формулой (3). Так как производная $f'(x)=2x$ есть на [2, 3] строго возрастающая функция, то наименьшим ее на этом отрезке значением будет $f'(2)=4$, т. е. $m=4$. Найдем $f(x_3)=(2{,}23607)^2-5=0{,}00001$. Тогда по формуле (3) имеем:

$$|x_3 - c| < \frac{0{,}00001}{4} = 0{,}0000025.$$

Если по условию конкретной задачи такая точность вычисления корня достаточна, то на этом можно остановиться, если нет, то процесс приближения можно продолжить до нужной точности.

Глава VII. ИНТЕГРИРОВАНИЕ

Интегральное исчисление непосредственно вытекает из дифференциального. Неопределенный интеграл наряду с производной является основным понятием математического анализа.

§ 1. Первообразная и неопределенный интеграл

1. Понятие первообразной функции. Одной из основных задач дифференциального исчисления является отыскание производной заданной функции. Разнообразные вопросы математического анализа с его многочисленными приложениями в геометрии, механике, физике и технике приводят к решению обратной задачи: по данной функции $f(x)$ найти такую функцию $F(x)$, производная которой равнялась бы функции $f(x)$, т. е. $(F(x))'=f(x)$.

Восстановление функции по известной производной этой функции составляет одну из основных задач интегрального исчисления.

Определение 1. Функция $F(x)$ называется *первообразной* для функции $f(x)$ на некотором промежутке X, если для всех значений x из этого промежутка выполняется равенство $(F(x))'=f(x)$.

Рассмотрим примеры:

1) функция $F(x)=\sin x$ является первообразной для функции $f(x)=\cos x$ на всей прямой, так как при любом значении x $(\sin x)'=\cos x$;
2) функция $F(x)=x^3$ является первообразной для функции $f(x)=3x^2$ на всей прямой, ибо в каждой точке x $(x^3)'=3x^2$;
3) функция $F(x)=\sqrt{1-x^2}$ является первообразной для функции $f(x)=-\dfrac{x}{\sqrt{1-x^2}}$ на интервале $(-1, +1)$, так как в любой точке x этого интервала $(\sqrt{1-x^2})'=-\dfrac{x}{\sqrt{1-x^2}}$.

Задача отыскания по данной функции $f(x)$ ее первообразной решается неоднозначно. Действительно, если $F(x)$ — первообразная для $f(x)$, т. е. $(F(x))'=f(x)$, то функция $F(x)+C$, где C — произвольная постоянная, также будет первообразной для $f(x)$, так как $[F(x)+C]'=f(x)$ для любого числа C. Например, для $f(x)=\cos x$ первообразной является не только $\sin x$, но и функция $\sin x+C$, так как $(\sin x+C)'=\cos x$.

Теперь покажем, что множество функций $F(x)+C$, где $F(x)$ — некоторая первообразная для функции $f(x)$, а C — произвольная постоянная, исчерпывает все первообразные для функции $f(x)$.

Лемма 7.1. *Функция, производная которой на некотором промежутке X равна нулю, постоянна.*

Доказательство. Пусть во всех точках промежутка X производная функции $f(x)$ равна нулю, т. е. $f'(x)=0$. Тогда для любых двух точек x_1, $x_2 \in X$ по теореме Лагранжа

$$f(x_2)-f(x_1)=f'(\xi)(x_2-x_1),\ x_1<\xi<x_2.$$

Так как $f'(\xi)=0$, то $f(x_2)=f(x_1)$. Это и означает, что значения функции во всех точках промежутка одинаковы, т. е. $f(x)=C$, где C — некоторое число ■

Теорема 7.1. *Если $F(x)$ — первообразная для функции $f(x)$ на некотором промежутке X, то любая другая первообразная для $f(x)$ на том же промежутке может быть представлена в виде $F(x)+C$*, где C — произвольная постоянная.

Доказательство. Пусть $\Phi(x)$ — любая другая первообразная для функции $f(x)$ на промежутке X, т. е. $(\Phi(x))'=f(x)$. Тогда для любого $x \in X$

$$[\Phi(x)-F(x)]'=(\Phi(x))'-(F(x))'=f(x)-f(x)=0,$$

а это значит (по лемме 7.1), что функция $\Phi(x)-F(x)$ постоянна, т. е. $\Phi(x)-F(x)=C$, где C — некоторое число. Следовательно, $\Phi(x)=F(x)+C$ ■

Из доказанной теоремы следует, что множество функций $F(x)+C$, где $F(x)$ одна из первообразных для функции $f(x)$, а C — произвольная постоянная, исчерпывает все семейство первообразных функций для $f(x)$.

2. Неопределенный интеграл

Определение 2. Если функция $F(x)$ — первообразная для функции $f(x)$, то множество функций $F(x)+C$, где C — произвольная постоянная, называется *неопределенным интегралом* от функции $f(x)$ и обозначается символом

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

При этом функция $f(x)$ называется *подынтегральной функцией*, $f(x)dx$ — *подынтегральным выражением*, а переменная x — *переменной интегрирования.*

Символ $\int f(x)dx$ обозначает, таким образом, совокупность всех первообразных для функции $f(x)$. Но иногда мы будем понимать его как любой элемент из этой совокупности, т. е. как какую-то из первообразных.

Восстановление функции по ее производной или, что то же, отыскание неопределенного интеграла по данной подынтегральной функции, называется *интегрированием* этой функции. Интегрирование представляет собой операцию, обратную дифференцированию. Для того чтобы проверить, правильно ли выполнено интегрирование, достаточно продифференцировать результат и получить при этом подынтегральную функцию.

В этой главе мы не будем заниматься вопросом о существовании первообразных (а значит, и неопределенных интегралов) для широких классов функций. Здесь мы лишь отметим, что в гл. VIII § 6 будет доказано, что любая непрерывная на отрезке функция имеет на этом отрезке первообразную (а следовательно, и неопределенный интеграл).

Примеры:

1) $\int 3x^2dx = x^3 + C$, так как $(x^3+C)' = 3x^2$;
2) $\int \cos x dx = \sin x + C$, так как $(\sin x + C)' = \cos x$;
3) $\int 1/x dx = \ln|x| + C$, так как $(\ln|x| + C)' = 1/x$;
4) $\int e^{-2x}dx = -1/2e^{-2x} + C$, так как $(-1/2e^{-2x} + C)' = e^{-2x}$

и т. д.

§ 2. Основные свойства неопределенного интеграла

Из определения неопределенного интеграла непосредственно вытекают следующие его свойства.

1. Производная неопределенного интеграла равна подынтегральной функции; дифференциал от неопределенного интеграла равен подынтегральному выражению, т. е.

$$(\int f(x)dx)' = f(x) \text{ и } d\int f(x)dx = f(x)dx.$$

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

$$\text{и } d\int f(x)dx = (\int f(x)dx)'dx = f(x)dx.$$

2. Неопределенный интеграл от дифференциала некоторой функции равен сумме этой функции и произвольной постоянной, т. е.

$$\int dF(x) = F(x) + C.$$

Действительно, так как $dF(x)=(F(x))'dx$, то $\int(F(x)'dx=F(x)+C$.

3. Постоянный множитель можно вынести из-под знака интеграла, т. е. если $k=\text{const}\neq 0$, то

$$\int kf(x)dx=k\int f(x)dx.$$

Действительно, пусть $F(x)$ — первообразная для функции $f(x)$, т. е. $(F(x))'=f(x)$. Тогда $kF(x)$ — первообразная для функции $kf(x)$: $(kF(x))'=k(F(x))'=kf(x)$. Из определения следует, что

$$k\int f(x)dx=k[F(x)+C]=kF(x)+C_1=\int kf(x)dx,$$

где $C_1=kC$.

4. Неопределенный интеграл от алгебраической суммы двух функций равен алгебраической сумме интегралов от этих функций в отдельности, т. е.

$$\int[f(x)\pm g(x)]dx=\int f(x)dx\pm\int g(x)dx.$$

Действительно, пусть $F(x)$ и $G(x)$ являются первообразными для функций $f(x)$ и $g(x)$:

$$(F(x))'=f(x),\ (G(x))'=g(x).$$

Тогда функции $F(x)\pm G(x)$ будут первообразными для функций $f(x)\pm g(x)$. Следовательно,

$$\int f(x)dx\pm\int g(x)dx=[F(x)+C_1]\pm[G(x)+C_2]=$$
$$=[F(x)\pm G(x)]+[C_1\pm C_2]=[F(x)\pm G(x)]+C=$$
$$=\int[f(x)\pm g(x)]dx.$$

Разумеется, это свойство справедливо для любого конечного числа слагаемых функций.

§ 3. Таблица основных интегралов

Приведем таблицу так называемых *основных* интегралов. Часть формул этой таблицы непосредственно следует из определения интегрирования как операции, обратной дифференцированию, и таблицы производных. Справедливость остальных формул легко проверить дифференцированием.

$$\text{I} \int x^{\alpha}dx = \frac{x^{\alpha+1}}{\alpha+1} + C \, (\alpha \neq -1),$$

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

$$\text{III} \int \frac{dx}{1+x^2} = \operatorname{arctg} x + C,$$

$$\text{IV} \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C,$$

$$\text{V} \int a^x dx = \frac{a^x}{\ln a} + C \quad (0 < a \neq 1)$$

$$\text{VI} \int e^x dx = e^x + C,$$

$$\text{VII} \int \sin x dx = -\cos x + C,$$

$$\text{VIII} \int \cos x dx = \sin x + C,$$

$$\text{IX} \int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C,$$

$$\text{X} \int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C,$$

$$\text{XI} \int \frac{dx}{x^2-a^2} = \frac{1}{2a} \ln\left|\frac{x-a}{x+a}\right| + C \quad (a \neq 0),$$

$$\text{XII} \int \frac{dx}{\sqrt{x^2+k}} = \ln|x + \sqrt{x^2+k}\,| + C,$$

$$\text{XIII} \int \frac{dx}{x^2+a^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C,$$

$$\text{XIV} \int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin \frac{x}{a} + C.$$

Интегралы, содержащиеся в этой таблице, принято называть *табличными.*

§ 4. Основные методы интегрирования

1. Непосредственное интегрирование. Вычисление интегралов путем непосредственного использования таблицы простейших интегралов и основных свойств неопределенных интегралов называется *непосредственным интегрированием.*

Пример 1.

$$\int\left(5\cos x+2-3x^2+\frac{1}{x}-\frac{4}{x^2+1}\right)dx=$$
$$=5\int\cos x dx+2\int dx-3\int x^2dx+\int\frac{dx}{x}-4\int\frac{dx}{x^2+1}=$$
$$=5\sin x-2x-x^3+\ln|x|-4\operatorname{arctg} x+C.$$

Пример 2.

$$\int\left(\sin\frac{x}{2}+\cos\frac{x}{2}\right)^2dx=\int\left(\sin^2\frac{x}{2}+2\sin\frac{x}{2}\cos\frac{x}{2}+\right.$$
$$\left.+\cos^2\frac{x}{2}\right)dx=\int(1+\sin x)\,dx=$$
$$=\int dx+\int\sin x\;dx=x-\cos x+C.$$

Пример 3.

$$\int\operatorname{tg}^2 x dx=\int(\sec^2 x-1)\,dx=\int\frac{dx}{\cos^2 x}-\int dx=\operatorname{tg} x-x+C.$$

2. Метод подстановки. Во многих случаях введение новой переменной интегрирования позволяет свести нахождение данного интеграла к нахождению табличного интеграла, т. е. перейти к непосредственному интегрированию. Такой метод называется методом *подстановки*, или методом *замены переменной*. Он основан на следующей теореме.

Теорема 7.2. *Пусть функция $x=\varphi(t)$ строго монотонна и дифференцируема на некотором промежутке T и пусть X — множество значений этой функции, на котором определена функция $f(x)$, т. е. на T определена сложная функция $f[\varphi(t)]$. Тогда, если на множестве X $f(x)$ имеет первообразную $F(x)$, то справедлива формула*

$$\int f(x)dx=\int f[\varphi(t)]\varphi'(t)dt. \qquad (1)$$

Доказательство. Так как первообразная $F(x)$ определена на том же множестве, что и функция $f(x)$, и существует сложная функция $f[\varphi(t)]$, то существует и сложная функция $F[\varphi(t)]$. Тогда по правилу дифференцирования сложной функции с учетом того, что $(F(x))'=f(x)$, получим

$$(F[\varphi(t)])'=(F[\varphi(t)])'_x\varphi'(t)=f[\varphi(t)]\varphi'(t),$$

т. е. функция $f[\varphi(t)]\cdot\varphi'(t)$ имеет первообразную $F[\varphi(t)]$ и, следовательно,

$$\int f[\varphi(t)]\cdot\varphi'(t)\,dt=F[\varphi(t)]+C.$$

Замечая, что $\int f(x)\,dx=F(x)+C=F[\varphi(t)]+C$, окончательно получим

$$\int f(x)\,dx=\int f[\varphi(t)]\varphi'(t)\,dt,$$

т. е. искомую формулу (1) ■

Формула (1) называется формулой *замены переменной* в неопределенном интеграле.

Пример 1. Вычислить интеграл $\int\frac{x^3}{(x-1)^2}dx.$

Решение. Положим $x-1=t$; следовательно, $x=t+1$. Отсюда $dx=dt$.
Тогда по формуле (1)

$$\int\frac{x^3}{(x-1)^2}dx=\int\frac{(t+1)^3}{t^2}dt=\int\left(t+3+\frac{3}{t}+\frac{1}{t^2}\right)dt=$$
$$=\frac{1}{2}t^2+3t+3\ln|t|-\frac{1}{t}+C.$$

Возвращаясь к переменной x, окончательно получим

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

Замечание. При замене переменной в неопределенном интеграле иногда более удобно задавать не x как функцию t, а, наоборот, задавать t как функцию от x. Теоретически оба способа равнозначны, так как из монотонности функции $x=\varphi(t)$ вытекает существование обратной функции $t=\psi(x)$.

Пример 2. Вычислить интеграл $\int\frac{x^4dx}{x^5+7}.$

Решение. Положим $x^5+7=t$, $dt=5x^4dx$; следовательно,

$$\int\frac{x^4dx}{x^5+7}=\frac{1}{5}\int\frac{dt}{t}=\frac{1}{5}\ln|t|+C,$$

так что

$$\int \frac{x^4 dx}{x^5+7} = \frac{1}{5} \ln |x^5+7| + C.$$

Необходимо заметить, что удачный выбор подстановки обычно представляет известные трудности. Для их успешного преодоления необходимо хорошо владеть техникой дифференцирования и твердо знать табличные интегралы.

Рассмотрим еще примеры.

Пример 3. Вычислить интеграл $\int \frac{dx}{\sqrt{x^2+a}}$.

Решение. Положим $\sqrt{x^2+a} + x = t$, откуда $\left(\frac{x}{\sqrt{x^2+a}} + 1\right) dx = dt$. Таким образом,

$$dx = \frac{\sqrt{x^2+a}}{\sqrt{x^2+a}+x}\, dt,$$

так что

$$\int \frac{dx}{\sqrt{x^2+a}} = \int \frac{dt}{t} = \ln |t| + C = \ln |\sqrt{x^2+a} + x| + C.$$

Пример 4. Вычислить интеграл $\int \sin^n x \cos x \, dx$.

Решение. Положим $t = \sin x$, $dt = \cos x \, dx$. Тогда

$$\int \sin^r x \cdot \cos x \; dx = \int t^n dt =$$

$$= \begin{cases} \frac{t^{n+1}}{n+1} + C = \frac{\sin^{n+1} x}{n+1} \text{ при } n \neq -1, \\ \ln \; t| + C = \ln |\sin x| + C \text{ при } n = -1. \end{cases}$$

Пример 5. Вычислить интеграл $\int \frac{xdx}{(x^2+1)^n}$, $n \neq 1$.

Решение. Положим $x^2+1=t$, $2xdx=dt$, следовательно,

$$\int \frac{xdx}{(x^2+1)^n} = \frac{1}{2} \int \frac{dt}{t^n} = -\frac{1}{2(n-1)} \cdot \frac{1}{t^{n-1}} + C = -$$

$$-\frac{1}{2(n-1)} \cdot \frac{1}{(x^2+1)^{n-1}} + C.$$

При $n=1$ аналогично получим

$$\int \frac{x\,dx}{x^2+1} = \frac{1}{2}\ln(x^2+1)+C.$$

3. Метод интегрирования по частям. Метод интегрирования по частям основан на использовании формулы дифференцирования произведения двух функций.

Теорема 7.3. *Пусть функции $u(x)$ и $v(x)$ определены и дифференцируемы на некотором промежутке X и пусть функция $u'(x)v(x)$ интегрируема на этом промежутке. Тогда на промежутке X функция $u(x)v'(x)$ также интегрируема и справедлива формула*

$$\int u(x)\,v'(x)\,dx = u(x)\,v(x) - \int v(x)\,u'(x)\,dx. \tag{2}$$

Доказательство. Из равенства

$$[u(x)v(x)]' = u'(x)v(x) + u(x)v'(x)$$

следует

$$u(x)v'(x) = [u(x)v(x)]' - u'(x)v(x).$$

Первообразной функции $[u(x)v(x)]'$ является функция $u(x)v(x)$. Функция $u'(x)v(x)$ интегрируема по условию теоремы. Следовательно, и функция $u(x)v'(x)$ интегрируема как разность интегрируемых функций. Интегрируя последнее равенство, получим формулу (2) ■

Формула (2) называется формулой *интегрирования по частям*. Так как $v'(x)\,dx = dv$, $u'(x)\,dx = du$, то ее можно записать в виде

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

Она позволяет свести вычисление $\int u\,dv$ к вычислению интеграла $\int v\,du$, который может оказаться более простым для интегрирования.

Пример 1.

$$\int \underbrace{\operatorname{arctg} x}_{u}\, \underbrace{dx}_{dv} = \left| \begin{matrix} u = \operatorname{arctg} x,\ du = \dfrac{dx}{1+x^2}, \\ dv = dx,\ \int dv = \int dx,\ v = x \end{matrix} \right| =$$

$$= \underbrace{x}_{v}\, \underbrace{\operatorname{arctg} x}_{u} - \int \underbrace{x}_{v}\, \underbrace{\frac{dx}{1+x^2}}_{du} = x \operatorname{arctg} x - \frac{1}{2}\int \frac{d(x^2+1)}{1+x^2} =$$

$$= x \operatorname{arctg} x - \frac{1}{2}\ln(1+x^2) + C.$$

Пример 2.

$$\int xe^x dx = \left| \begin{array}{l} u = x,\ du = dx \\ dv = e^x dx,\ \int dv = \int e^x dx,\ v = e^x \end{array} \right| =$$

$$= xe^x - \int e^x dx = xe^x - e^x + C.$$

Пример 3.

$$\int x \ln x dx = \left| \begin{array}{l} u = \ln x,\ du = \dfrac{dx}{x} \\ dv = xdx;\ \int dv = \int xdx,\ v = \dfrac{x^2}{2} \end{array} \right| =$$

$$= \frac{x^2}{2}\ln x - \int \frac{x}{2} dx = \frac{x^2}{2}\ln x - \frac{x^2}{4} + C.$$

В заключение вычислим интеграл

$$J_n = \int \frac{dx}{(x^2+1)^n}$$

(n — целое положительное число), который нам понадобится в следующем параграфе. При $n=1$ имеем

$$J_1 = \operatorname{arc\,tg} x + C.$$

Пусть $n>1$. Заменив в числителе 1 разностью $(x^2+1)-x^2$, получим

$$J_n = \int \frac{dx}{(x^2+1)^{n-1}} - \int \frac{x^2 dx}{(x^2+1)^n}.$$

Во втором интеграле положим

$$u=x,\ du=dx,$$

$$dv = \frac{xdx}{(x^2+1)^n},\ v = \int \frac{xdx}{(x^2+1)^n} = -\frac{1}{(2n-2)(x^2+1)^{n-1}}$$

(см. п. 2, пример 5), поэтому

$$\int \frac{x^2 dx}{(x^2+1)^n} = -\frac{x}{(2n-2)(x^2+1)^{n-1}} + \int \frac{dx}{(2n-2)(x^2+1)^{n-1}},$$

следовательно,

$$J_n = J_{n-1} + \frac{x}{(2n-2)(x^2+1)^{n-1}} - \frac{1}{2n-2} J_{n-1},$$

т. е.

$$J_n = \frac{x}{(2n-2)(x^2+1)^{n-1}} + \frac{2n-3}{2n-2} J_{n-1}.$$

Таким образом,

$$\int \frac{dx}{(x^2+1)^n} = \frac{x}{(2n-2)(x^2+1)^{n-1}} + \frac{2n-3}{2n-2} \int \frac{dx}{(x^2+1)^{n-1}} \quad (n > 1). \tag{3}$$

Формулы вида (3) называются *формулами приведения.*

Пример 4. Вычислить $\int \frac{dx}{(x^2+1)^3}$.

Решение. По формуле приведения (3) имеем

$$\int \frac{dx}{(x^2+1)^3} = \frac{x}{4(x^2+1)^2} + \frac{3}{4} \int \frac{dx}{(x^2+1)^2},$$

$$\int \frac{dx}{(x^2+1)^2} = \frac{x}{2(x^2+1)} + \frac{1}{2} \int \frac{dx}{x^2+1},$$

$$\int \frac{dx}{x^2+1} = \operatorname{arctg} x,$$

поэтому окончательно

$$\int \frac{dx}{(x^2+1)^2} = \frac{x}{4(x^2+1)^2} + \frac{3}{8(x^2+1)} + \frac{3}{8} \operatorname{arctg} x + C.$$

§ 5. Интегрирование рациональных функций

Важный класс функций, интегралы от которых всегда выражаются через элементарные функции, образуют рациональные функции, т. е. функции, которые можно представить в виде дроби

$$\frac{P(x)}{Q(x)},$$

где $P(x)$, $Q(x)$ — многочлены.

Если степень многочлена в числителе равна или больше степени многочлена в знаменателе, то, выполнив деление, получим

$$\frac{P(x)}{Q(x)} = W(x) + \frac{R(x)}{Q(x)}, \tag{1}$$

где $W(x)$ — некоторый многочлен, а $R(x)$ — многочлен степени ниже, чем $Q(x)$.

Примеры:

1) $\frac{x^5+x^3-x^2+1}{x^3-2x+1}=x^2+3-\frac{2x^2-6x+2}{x^3-2x+1}$;

2) $\frac{x^3+x+1}{x^2+1}=x+\frac{1}{x^2+1}$.

В высшей алгебре доказывается, что каждый многочлен $Q(x)$ может быть представлен в виде произведения

$$Q(x)=A(x-\alpha)(x-\beta)\dots(x-\gamma),$$

где A — коэффициент при старшей степени многочлена $Q(x)$, а α, $\beta, \dots, \gamma$ — корни уравнения $Q(x)=0$. Множители $(x-\alpha)(x-\beta)\dots(x-\gamma)$ называются *элементарными множителями*. Если среди них существуют совпадающие, то, собирая их, получим представление

$$Q(x)=A(x-\alpha)^r(x-\beta)^s\dots(x-\gamma)^t, \qquad (2)$$

где r, $s, \dots, t$ — целые числа, которые называются кратностями, соответствующими корням $\alpha, \beta, \dots, \gamma$, причем $r+s+\dots+t=n$, n означает степень многочлена $Q(x)$.

Среди корней представления (2) могут быть и комплексные. В алгебре доказывается, что если $\alpha=a+bi$ — r-кратный комплексный корень многочлена с вещественными коэффициентами, то он имеет также сопряженный с ним r-кратный корень $\bar{\alpha}=a-bi$. Другими словами, если в представление (2) входит множитель $(x-\alpha)^r$, то оно содержит также и множитель $(x-\bar{\alpha})^r$. Перемножим их:

$$(x-\alpha)^r(x-\bar{\alpha})^r=\{[x-(a+bi)][x-(a-bi)]\}^r=$$
$$=[x^2-x(a+bi)-x(a-bi)+a^2+b^2]^r=$$
$$=[x^2-2ax+a^2+b^2]^r=(x^2+2px+q)^r,$$

где $p=-a$, $q=a^2+b^2$, $p^2-q<0$.

Поступая аналогично с остальными комплексными корнями, представление (2) запишем в виде

$$Q(x)=A(x-\alpha)^r(x-\beta)^s\dots(x^2+2px+q)^t(x^2+2ux+v)^\eta\dots \qquad (3)$$

В высшей алгебре доказывается

Теорема. *Если рациональная функция* $\frac{R(x)}{Q(x)}$ *в соотношении (1) имеет степень многочлена в числителе меньше степени многочлена в знаменателе, а многочлен* $Q(x)$ *представлен в виде (3), то эту функцию можно единственным образом представить в виде*

$$\frac{R(x)}{Q(x)}=\frac{A_1}{(x-\alpha)}+\frac{A_2}{(x-\alpha)^2}+\ldots+\frac{A_r}{(x-\alpha)^r}+\ldots+$$
$$+\ldots+\frac{M_1x+N_1}{x^2+2px+q}+\frac{M_2x+N_2}{(x^2+2px+q)^2}+\ldots+$$
$$+\ldots+\frac{M_tx+N_t}{(x^2+2px+q)^t}+\ldots, \tag{4}$$

где $A_1, A_2, \ldots, A_r, \ldots, M_1, N_1, M_2, N_2, \ldots, M_t, N_t, \ldots$ — *некоторые числа.*

Разложение (4) называется *разложением рациональной функции на элементарные дроби.*

Равенство (4) имеет место для всех x, не являющихся вещественными корнями многочлена $Q(x)$.

Чтобы определить числа $A_1, A_2, \ldots, A_r, \ldots, M_1, N_1, \ldots, M_t, N_t, \ldots$, умножим обе части разложения (4) на $Q(x)$. Поскольку равенство между многочленом $R(x)$ и многочленом, который получится в правой части, будет справедливо для всех x, то коэффициенты, стоящие при равных степенях x, равны между собой. Получим, таким образом, ряд уравнений первой степени, из которых найдем неизвестные числа $A_1, A_2, \ldots, A_r, \ldots, M_1, N_1, \ldots, M_t, N_t, \ldots$

Пример. Разложить рациональную функцию $\frac{2x-1}{x^2-5x+6}$ на элементарные дроби.

Решение. Так как $x^2-5x+6=(x-3)(x-2)$, то по формуле (4) имеем

$$\frac{2x-1}{x^2-5x+6}=\frac{A}{x-3}+\frac{B}{x-2}.$$

Отсюда, умножая обе части равенства на x^2-5x+6, получаем

$$2x-1=A(x-2)+B(x-3);$$

поэтому

$$2x-1=(A+B)x-2A-3B.$$

Приравнивая коэффициенты при одинаковых степенях x, получаем уравнения первой степени:

$$\begin{cases} A+B=2, \\ 2A+3B=1, \end{cases}$$

откуда $A=5$, $B=-3$.

Таким образом,

$$\frac{2x-1}{x^2-5x+6}=\frac{5}{x-3}-\frac{3}{x-2}.$$

Из изложенного следует, что задача интегрирования рациональной функции (1) сводится к интегрированию рациональной функции $w(x)=a_0x^m+a_1x^{m-1}+\ldots+a_m$, интеграл от которой является табличным:

$$\int w(x)\,dx=a_0\frac{x^{m+1}}{m+1}+a_1\frac{x^m}{m}+\ldots+a_mx+C,$$

и интегрированию рациональной функции $\frac{R(x)}{Q(x)}$, что сводится к нахождению интегралов четырех типов:

I $\displaystyle\int\frac{A}{x-a}\,dx=A\ln|x-a|+C;$

II $\displaystyle\int\frac{A}{(x-a)^r}\,dx=-\frac{A}{(r-1)(x-a)^{r-1}}+C\;(r>1);$

III $\displaystyle\int\frac{Ax+B}{x^2+2px+q}\,dx;$

IV $\displaystyle\int\frac{Ax+B}{(x^2+2px+q)^r}\,dx\;(r>1).$

При этом многочлен $x^2+2px+q$ не имеет вещественных корней, так что $p^2-q<0$.

Займемся вычислением интеграла типа III, который принадлежит к числу интегралов, часто встречающихся на практике.

Выделим из трехчлена в знаменателе полный квадрат:

$$x^2+2px+q=(x+p)^2+q-p^2.$$

Это разложение подсказывает нам подстановку

$$x+p=t,\quad x=t-p,\quad dx=dt.$$

Положим далее $q-p^2=h>0$ и перейдем к переменной t. В результате интеграл преобразуется к виду

$$\int \frac{Ax+B}{x^2+2px+q}\,dx=\int \frac{At+B-Ap}{t^2+h}\,dt=$$

$$=\frac{1}{2}A\int \frac{2t\,dt}{t^2+h}+(B-Ap)\int \frac{dt}{t^2+h}.$$

Первый интеграл в правой части берется непосредственно:

$$\int \frac{2t\,dt}{t^2+h}=\ln|t^2+h|+C=\ln|x^2+2px+q|+C.$$

Второй интеграл вычисляется по формуле XIII таблицы основных интегралов.

Пример 1. Вычислить интеграл $\int \frac{6x+5}{x^2+4x+9}\,dx.$

Решение. Выделим в знаменателе полный квадрат: $x^2+4x+9=(x+2)^2+5$. Сделаем подстановку: $x+2=t$, $x=t-2$, $dx=dt$, которая дает

$$\int \frac{6x+5}{x^2+4x+9}\,dx=\int \frac{6x+5}{(x+2)^2+5}\,dx=\int \frac{6(t-2)+5}{t^2+5}\,dt=$$

$$=\int \frac{6t-7}{t^2+5}\,dt=3\int \frac{2t\,dt}{t^2+5}-7\int \frac{dt}{t^2+5}=$$

$$=3\ln(t^2+5)-\frac{7}{\sqrt{5}}\operatorname{arctg}\frac{t}{\sqrt{5}}+C.$$

Возвращаясь к переменной x, получим

$$\int \frac{6x+5}{x^2+4x+9}\,dx=3\ln(x^2+4x+9)-\frac{7}{\sqrt{5}}\operatorname{arctg}\frac{x+2}{\sqrt{5}}+C.$$

Теперь перейдем к вычислению интеграла типа IV: $\int \frac{Ax+B}{(x^2+2px+q)^r}\,dx$, $q-p^2>0$, $r>1$. Для этого введем новую переменную z по формуле

$$z=\frac{x+p}{\sqrt{q-p^2}},\ \text{откуда}$$

$$x=z\sqrt{q-p^2}-p,\quad dx=\sqrt{q-p^2}\,dz. \tag{5}$$

Далее,

$$z^2+1=\frac{(x+p)^2}{q-p^2}+1=\frac{x^2+2px+q}{q-p^2}. \tag{6}$$

Таким образом, пользуясь подстановкой (5) и принимая во внимание (6), получим

$$\int \frac{Ax+B}{(x^2+2px+q)^r}\,dx=\int \frac{A[z\sqrt{q-p^2}-p]+B}{(z^2+1)^r(q-p^2)^r}\sqrt{q-p^2}\,dz=$$

$$=\int \frac{Mz+N}{(z^2+1)^r}\,dz=M\int \frac{z\,dz}{(z^2+1)^r}+N\int \frac{dz}{(z^2+1)^r},$$

где M, N — постоянные числа, значения которых ясны из последнего равенства.

Ко второму интегралу можем применить формулу приведения (см. § 4, п. 3, формула (3)). Положив в первом интеграле $z^2+1=t$, получим

$$M\int \frac{z\,dz}{(z^2+1)^r}=\frac{M}{2}\int \frac{dt}{t^r}=M\begin{cases}\frac{1}{2}\ln|t|+C\ (r=1),\\ -\frac{1}{2(r-1)}\cdot\frac{1}{t^{r-1}}+C(r\neq 1)\end{cases}=$$

$$=M\begin{cases}\frac{1}{2}\ln(z^2+1)+C,\ r=1,\\ -\frac{1}{2(r-1)}\cdot\frac{1}{(z^2+1)^{r-1}}+C,\ r\neq 1.\end{cases}$$

Пример 2. Вычислить интеграл $\int \frac{5x+3}{(x^2-2x+5)^2}\,dx$.

Решение. Положим $z=\frac{x-1}{\sqrt{5-1}}=\frac{x-1}{2}$, откуда $x=1+2z$, $dx=2dz$, а $x^2-2x+5=4(z^2+1)$; следовательно

$$\int \frac{5x+3}{(x^2-2x+5)^2}\,dx=\int \frac{5(1+2z)+3}{4^2(z^2+1)^2}\,2dz=\int \frac{10z+8}{8(z^2+1)^2}\,dz=$$

$$=\frac{5}{4}\int \frac{z\,dz}{(z^2+1)^2}+\int \frac{dz}{(z^2+1)^2}.$$

Но

$$\int \frac{z\,dz}{(z^2+1)^2}=-\frac{1}{2}\cdot\frac{1}{z^2+1},$$

$$\int \frac{dz}{(z^2+1)^2} = \frac{z}{2(x^2+1)} + \frac{1}{2}\operatorname{arctg} z.$$

Таким образом,

$$\int \frac{5x+3}{(x^2-2x+5)^2}dx = -\frac{5}{8}\cdot\frac{1}{z^2+1} + \frac{z}{2(z^2+1)} + $$
$$+\frac{1}{2}\operatorname{arctg} z + C = \frac{4z-5}{8(z^2+1)} + \frac{1}{2}\operatorname{arctg} z + C.$$

Возвращаясь теперь к переменной x, получим

$$\int \frac{5x+3}{(x^2-2x+5)^2}dx = \frac{2x-7}{2(x^2-2x+5)} + \frac{1}{2}\operatorname{arctg}\left(\frac{x-1}{2}\right) + C.$$

Итак, мы установили, что интегрирование любой рациональной функции сводится к интегрированию многочлена и конечного числа элементарных дробей, интегралы от которых выражаются через рациональные функции, логарифмы и арктангенсы. Иными словами, любая рациональная функция интегрируется в элементарных функциях.

Глава VIII. ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

§ 1. Определение определенного интеграла

Пусть функция $y=f(x)$ определена на отрезке $[a, b]$, $a<b$. Разобьем этот отрезок на n произвольных частей точками

$$a=x_0<x_1<x_2<\ldots<x_{i-1}<x_i<\ldots<x_n=b.$$

Обозначим это разбиение через τ, а точки $x_0, x_1, \ldots, x_n$ будем называть точками разбиения τ. В каждом из полученных частичных отрезков $[x_{i-1}, x_i]$ выберем произвольно точку $\xi_i\,(x_{i-1}\leqslant\xi_i\leqslant x_i)$. Через Δx_i обозначим разность $x_i - x_{i-1}$, которую будем называть длиной частичного отрезка $[x_{i-1}, x_i]$. Составим сумму:

$$\sigma = f(\xi_1)\,\Delta x_1 + f(\xi_2)\,\Delta x_2 + \ldots + f(\xi_n)\,\Delta x_n = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i, \tag{1}$$

которую назовем *интегральной суммой* для функции $f(x)$ на $[a, b]$, соответствующей данному разбиению $[a, b]$ на частичные отрезки и данному выбору промежуточных точек ξ_i. Геометрический смысл суммы σ очевиден: это сумма площадей прямоугольников с основаниями Δx_1, Δx_2, ..., Δx и высотами $f(\xi_1)$, $f(\xi_2)$, ..., $f(\xi_n)$, если $f(x) \geqslant 0$ (рис. 84).

Обозначим через λ длину наибольшего частичного отрезка разбиения τ: $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\}$.

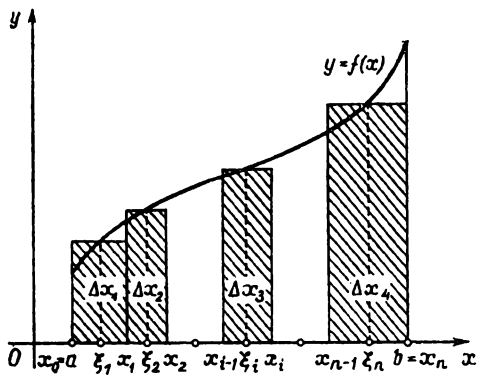


Рис. 84

Определение. Если существует конечный предел J интегральной суммы (1) при $\lambda \to 0$, не зависящей ни от способа разбиения отрезка $[a, b]$, ни от выбора точек ξ_i, то этот предел называется *определенным интегралом* от функции $f(x)$ по отрезку $[a, b]$ и обозначается следующим образом:

$$J = \int_a^b f(x)\, dx, \tag{2}$$

или

$$\int_a^b f(x)\, dx = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\, \Delta x_i,$$

В этом случае функция $f(x)$ называется *интегрируемой* на $[a, b]$. Числа a и b называются соответственно *нижним* и *верхним* пределами интегрирования, $f(x)$ — *подынтегральной функцией*, x — *переменной интегрирования*.

Здесь необходимо сделать ряд пояснений, так как мы имеем не совсем обычный предельный переход. Данное определение определенного интеграла по форме напоминает первое определение предела функции на языке последовательностей, где вместо функции стоит интегральная сумма (1), являющаяся переменной величиной, зависящей от λ. Действительно, пусть отрезок $[a, b]$ последовательно разбивается на части, сначала одним способом, затем — вторым, третьим и т. д. Причем длина наибольшего отрезка в каждом случае уменьшается ($\lambda \to 0$)*, когда n стремится к бесконечности. Таким образом, получаем последовательность разбиений $\{\tau_n\}$, у которой $\lim\limits_{n\to\infty} \lambda_n = 0$, и мы можем дать определение определенного интеграла на знакомом нам языке последовательностей: функция $f(x)$ называется интегрируемой на $[a, b]$, если для любой последовательности разбиений $\{\tau_n\}$, у которой $\lim\limits_{n\to\infty} \lambda_n = 0$, соответствующая последовательность интегральных сумм $\{\sigma_n\}$ стремится всегда к одному и тому же пределу $J = \lim\limits_{n\to\infty} \sigma_n$.

Можно дать определение определенного интеграла и на языке $\varepsilon - \delta$: число J называется определенным интегралом функции $f(x)$ на отрезке $[a, b]$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\lambda < \delta$ (т. е. отрезок разбит на части с длинами $\Delta x_i < \delta$) независимо от выбора точек ξ_i выполняется неравенство

$$\left| \sum_{i=1}^{n} f(\xi_i) \Delta x_i - J \right| < \varepsilon.$$

Доказательство эквивалентности обоих определений можно провести по аналогии эквивалентности двух определений предела функции. Определение на языке последовательностей дает возможность перенести основные понятия теории пределов и на этот новый вид предела.

Из определения определенного интеграла следует, что величина интеграла (2) зависит только от вида

* Вместо $\lambda \to 0$ было бы неправильно писать $n \to \infty$, так как можно привести пример, когда увеличение числа точек разбиения $[a, b]$ еще не обязательно означает, что все Δx_i неограниченно убывают; если же $\lambda \to 0$, то все $\Delta x_i \to 0$ и обязательно $n \to \infty$.

функций $f(x)$ и от чисел a и b. Следовательно, если заданы $f(x)$ и пределы интегрирования, то интеграл (2) определяется однозначно и представляет собой *некоторое число.*

§ 2. Условия существования определенного интеграла

1. Ограниченность интегрируемой функции

Теорема 8.1. (необходимое условие). *Если функция $f(x)$ интегрируема на отрезке $[a, b]$, то она ограничена на этом отрезке.*

Доказательство. Предположим обратное, т. е. допустим, что $f(x)$ не ограничена на $[a, b]$. Покажем, что в этом случае интегральную сумму σ можно за счет выбора точек $\xi_1, \xi_2, ..., \xi_n$ сделать сколь угодно большой при любом разбиении отрезка $[a, b]$.

Действительно, так как $f(x)$ не ограничена на $[a, b]$, то при любом разбиении отрезка $[a, b]$ она обладает этим свойством хотя бы на одном частичном отрезке разбиения, скажем на Δx_1. Выберем тогда на остальных отрезках $\Delta x_2, \Delta x_3, ..., \Delta x_n$ точки $\xi_2, \xi_3, ..., \xi_n$ произвольно и обозначим

$$\sigma' = f(\xi_2)\Delta x_2 + f(\xi_3)\Delta x_3 + ... + f(\xi_n)\Delta x_n.$$

После этого возьмем такое ξ_1 на Δx_1, чтобы

$$|f(\xi_1)| \geqslant \frac{|\sigma'| + M}{\Delta x_1},$$

где M — любое наперед заданное число. Это можно сделать ввиду неограниченности $f(x)$ на Δx_1. Имеем

$$|f(\xi_1)|\Delta x_1 \geqslant |\sigma'| + M$$

$$\text{и} \quad |\sigma| = |f(\xi_i)\Delta x_1 + \sigma'| \geqslant |f(\xi_1)|\Delta x_1 - |\sigma'| \geqslant M,$$

т. е. интегральная сумма σ будет по абсолютной величине больше любого наперед заданного числа. Поэтому интегральная сумма σ не будет иметь конечного предела, а это означает, что определенный интеграл от неограниченной функции не существует ■

Замечание. Обратная теорема неверна, т. е. условие ограниченности функции $f(x)$, являясь необходимым условием интегрируемости функции, не является вместе с тем достаточным. Поясним это утверждение на примере. Рассмотрим функцию Дирихле на отрезке [0, 1]:

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

Функция Дирихле, очевидно, ограничена. Однако она не интегрируема на [0, 1]. Покажем это. Если при любом разбиении [0, 1] выбрать точки $\xi_i\,(x_{i-1} \leqslant \xi_i \leqslant x_i)$ рациональными, то получим

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{n} 1 \cdot \Delta x_i = 1,$$

а если взять ξ_i иррациональными, то получим

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{n} 0 \cdot \Delta x_i = 0.$$

Таким образом, при разбиении на сколь угодно малые отрезки интегральная сумма может принимать как значение, равное 0, так и значение, равное 1. Поэтому интегральная сумма σ при $\lambda \to 0$ предела не имеет.

Таким образом, очевидно, что для существования определенного интеграла от некоторой функции $f(x)$ последняя, помимо ограниченности, должна обладать дополнительными свойствами, обеспечивающими ее интегрируемость. Но для установления этих свойств нам необходимо ввести понятия нижних и верхних сумм и изучить их.

2. Суммы Дарбу*. Пусть функция $f(x)$ ограничена на отрезке $[a, b]$ и τ — разбиение этого отрезка точками $a = x_0 < x_1, < \ldots < x_{i-1} < x_i < \ldots < x_n = b$. Обозначим через m_i и M_i соответственно точную нижнюю и точную верхнюю грани этой функции на отрезке $[x_{i-1}, x_i]$ и составим следующие суммы:

$$S = M_1 \Delta x_1 + M_2 \Delta x_2 + \ldots + M_n \Delta x_n = \sum_{i=1}^{n} M_i \Delta x_i,$$

$$s = m_1 \Delta x_1 + m_2 \Delta x_2 + \ldots + m_n \Delta x_n = \sum_{i=1}^{n} m_i \Delta x_i.$$

---

* Дарбу Гастон (1842—1917) — французский математик.

Эти суммы называются соответственно *верхней* и *нижней суммами*, или *верхней* и *нижней суммами Дарбу* функции $f(x)$ для данного разбиения τ отрезка $[a, b]$.

Из определения нижней и верхней граней имеем $m_i \leqslant f(\xi_i) \leqslant M_i$ при $\xi_i \in [x_{i-1}, x_i]$. Отсюда

$$s = \sum_{i=1}^{n} m_i \Delta x_i \leqslant \sigma = \sum_{i=1}^{n} f(\xi_i) \Delta x_i \leqslant \sum_{i=1}^{n} M_i \Delta x_i = S,$$

т. е. интегральная сумма и суммы Дарбу связаны неравенствами

$$s \leqslant \sigma \leqslant S. \tag{1}$$

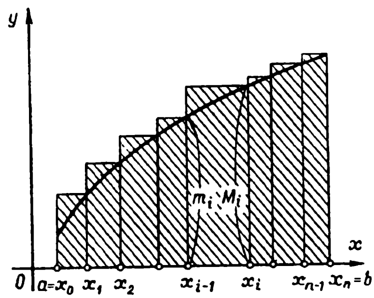


Рис. 85

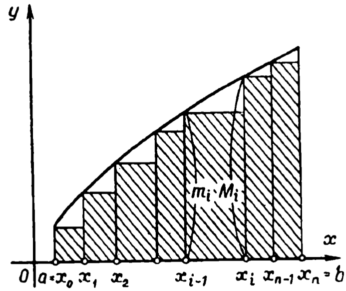


Рис. 86

Суммы Дарбу имеют простой геометрический смысл. Для простоты рассмотрим неотрицательную, непрерывную функцию $f(x)$ на $[a, b]$ и *криволинейную трапецию*, т. е. фигуру, ограниченную графиком функции $f(x)$, двумя ординатами, проведенными в точках a и b оси Ox, и осью Ox (рис. 85 и 86). Поскольку функция $f(x)$ непрерывна на $[a, b]$, то она будет непрерывной на $[x_{i-1}, x_i]$. По второй теореме Вейерштрасса $f(x)$ достигает на $[x_{i-1}, x_i]$ своих точных граней и, следовательно, числа m_i и M_i представляют собой наименьшее и наибольшее значения функции на $[x_{i-1}, x_i]$. Поэтому сумма S равна площади заштрихованной на рис. 85 ступенчатой фигуры, описанной около криволинейной трапеции, а сумма s равна площади заштрихованной на рис. 86 ступенчатой фигуры, вписанной в данную криволинейную трапецию.

Следует особо подчеркнуть, что суммы Дарбу зависят только от разбиения отрезка $[a, b]$, в то время как интегральная сумма σ зависит еще и от выбора точек ξ_i на частичных отрезках $[x_{i-1}, x_i]$. При фиксированном разбиении отрезка $[a, b]$ суммы s и S будут некоторыми числами, в то время как сумма σ является переменной величиной ввиду произвольности точек ξ_i.

3. **Свойства сумм Дарбу.** Суммы Дарбу обладают следующими свойствами.

Свойство I. Для любого фиксированного разбиения τ и для любого $\varepsilon>0$ точки ξ_i на отрезках $[x_{i-1}, x_i]$ можно выбрать так, что интегральная сумма σ будет удовлетворять неравенствам $0 \leqslant S - \sigma < \varepsilon$. Точки ξ_i можно выбрать также и таким образом, что интегральная сумма будет удовлетворять неравенствам $0 \leqslant \sigma - s < \varepsilon$.

Доказательство. Пусть τ — некоторое фиксированное разбиение отрезка $[a, b]$. Докажем, например, неравенства $0 \leqslant S - \sigma < \varepsilon$. По свойству точной верхней грани M_i для данного $\varepsilon>0$ на $[x_{i-1}, x_i]$ можно указать такую точку ξ_i, что

$$0 \leqslant M_i - f(\xi_i) < \frac{\varepsilon}{b-a}, \quad i = 1, 2, \ldots, n.$$

Умножая эти неравенства на Δx_i и затем складывая, получим

$$0 \leqslant S - \sigma < \varepsilon.$$

Аналогично устанавливаются неравенства $0 \leqslant \sigma - s < \varepsilon$.

Свойство II. От добавления к данному разбиению τ отрезка $[a, b]$ новых точек разбиения нижняя сумма Дарбу не уменьшается, а верхняя — не увеличивается.

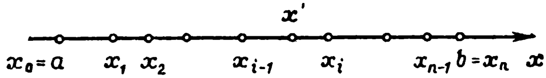


Рис. 87

Доказательство. Для доказательства этого свойства достаточно ограничиться добавлением к данному разбиению τ еще одной точки разбиения x', так как добавление нескольких точек разбиения можно провести, добавляя их по одной. Предположим, что эта новая точка x' попала на отрезок $[x_{i-1}, x_i]$ (рис. 87). Обозначим соответственно через s и s' нижние, а через S и S' верхние суммы Дарбу для данного разбиения τ и полученного из него добавлением x' разбиения τ'.

Проведем доказательство для нижних сумм Дарбу s и s'. Обозначим через m_i' и m_i'' точные нижние грани функции $f(x)$ на отрезках $[x_{i-1}, x']$ и $[x', x_i]$. В сумму s входит слагаемое $m_i \Delta x_i$, а в сумму s' вместо него входят два слагаемых:

$$m_i'(x'-x_{i-1})+m_i''(x_i-x').$$

Остальные слагаемые в суммах s и s' одинаковы. Так как $m_i' \geqslant m_i$, $m_i'' \geqslant m_i$ (точная нижняя грань на части $[x_{i-1}, x_i]$ не меньше точной нижней грани на всем $[x_{i-1}, x_i]$), то

$$m_i'(x'-x_{i-1}) + m_i''(x_i-x') \geqslant$$
$$\geqslant m_i(x'-x_{i-1}) + m_i(x_i-x') = m_i \Delta x_i,$$

а это означает, что $s' \geqslant s$.

Аналогично доказывается, что $S' \leqslant S$ ■

Свойство III. Нижняя сумма Дарбу для любого разбиения τ' не превосходит верхней суммы для любого другого разбиения τ''

Доказательство. Пусть s', S' и s'', S'' — соответственно нижние и верхние суммы Дарбу для разбиений τ' и τ''. Для доказательства этого свойства рассмотрим разбиение τ, состоящее из всех точек, входящих в разбиения τ' и τ''. Так как разбиение τ может быть получено из разбиения τ' добавлением к нему точек разбиения τ'', то по второму свойству с учетом очевидного неравенства $s \leqslant S$ получаем

$$s' \leqslant s \leqslant S \leqslant S'.$$

Но разбиение τ может быть также получено из разбиения τ'' добавлением к нему точек разбиения τ' Поэтому

$$s'' \leqslant s \leqslant S \leqslant S''.$$

Сравнивая установленные выше неравенства с последними, получаем $s' \leqslant S''$, $s'' \leqslant S'$ ■

Свойство IV. Множество $\{S\}$ верхних сумм Дарбу данной функции $f(x)$ для всевозможных разбиений отрезка $[a, b]$ ограничено снизу, а множество $\{s\}$ нижних сумм Дарбу ограничено сверху, причем точная верхняя грань множества $\{s\}$ не превосходит точную нижнюю грань множества $\{S\}$.

Доказательство. Это свойство непосредственно следует из свойства III. Действительно, множество всех верхних сумм Дарбу $\{S\}$ ограничено снизу, например, любой нижней суммой Дарбу s, а множество всех нижних сумм Дарбу $\{s\}$ ограничено сверху, например, любой верхней суммой Дарбу S. Тогда по теореме 1.1 множества $\{S\}$ и $\{s\}$, отвечающие всевозможным разбиениям отрезка $[a, b]$, имеют точные грани. Обозначим через $\bar{I}$ точную нижнюю грань множества $\{S\}$, а через $\underline{I}$ — точную верхнюю грань множества $\{s\}$:

$$\bar{I} = \inf\{S\}, \quad \underline{I} = \sup\{s\}.$$

Покажем, что $\underline{I} \leqslant \bar{I}$. Пусть $\underline{I} > \bar{I}$. Обозначим их разность через ε, так что $\underline{I} - \bar{I} = \varepsilon > 0$. Из свойства точных граней $\underline{I}$ и $\bar{I}$ вытекает, что существуют числа S' и s'', представляющие собой соответственно верхнюю и нижнюю суммы Дарбу некоторых разбиений τ' и τ'' отрезка $[a, b]$, такие, что $\bar{I} + \frac{\varepsilon}{2} > S'$ и $\underline{I} - \frac{\varepsilon}{2} < s''$. Вычитая второе неравенство из первого и учитывая, что $\underline{I} - \bar{I} = \varepsilon$, получим $s'' > S'$, что противоречит свойству III ■

4. Необходимое и достаточное условие интегрируемости. Имеет место следующая основная теорема.

Теорема 8.2. *Для того чтобы ограниченная на отрезке $[a, b]$ функция $f(x)$ была интегрируемой на этом отрезке, необходимо и достаточно, чтобы*

$$\lim_{\lambda \to 0} (S - s) = 0. \tag{2}$$

Условие (2) означает, что для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\lambda < \delta$ выполняется неравенство $|S - s| < \varepsilon$. Так как $s \leqslant S$, то последнее неравенство равносильно неравенству

$$S - s < \varepsilon. \tag{3}$$

Доказательство. *Необходимость.* Пусть функция $f(x)$ интегрируема на отрезке $[a, b]$, т. е. существует определенный интеграл $I = \int_a^b f(x)\,dx$. Это означает, что для любого $\varepsilon>0$ существует $\delta>0$ такое, что для любого разбиения τ, удовлетворяющего условию $\lambda<\delta$, независимо от выбора точек ξ_i, выполняется неравенство

$$|\sigma-J|<\frac{\varepsilon}{4}. \tag{4}$$

Зафиксируем одно такое разбиение τ. Тогда для данного разбиения τ в силу свойства I можно указать такие интегральные суммы σ' и σ'', что

$$S-\sigma'\leqslant\frac{\varepsilon}{4},\quad \sigma''-s\leqslant\frac{\varepsilon}{4}. \tag{5}$$

Из соотношения

$$S-s=(S-\sigma')+(\sigma'-J)+(J-\sigma'')+(\sigma''-s)$$

и неравенств (4) и (5) вытекает, что

$$S-s<\varepsilon,$$

а это и означает выполнение условия (3).

Достаточность. Пусть выполнено условие (3). В силу свойства IV $s\leqslant\underline{J}\leqslant\overline{J}\leqslant S$ для любых нижних и верхних сумм Дарбу, поэтому $0\leqslant\overline{J}-\underline{J}\leqslant S-s$, откуда согласно (2) следует, что $\overline{J}-\underline{J}=0$, т. е. $\overline{J}=\underline{J}$. Обозначая их общее значение через J, получим

$$s\leqslant J\leqslant S. \tag{6}$$

Если же интегральная сумма σ и суммы Дарбу s и S отвечают одному и тому же разбиению τ, то, как мы знаем (см. (1)),

$$s\leqslant\sigma\leqslant S. \tag{7}$$

Сопоставляя неравенства (6) и (7), получаем

$$|\sigma-J|\leqslant S-s. \tag{8}$$

По условию (3) $S-s<\varepsilon$. Но тогда из неравенства (8) следует, что и

$$|\sigma-J|<\varepsilon \text{ при } \lambda<\delta,$$

а это означает, что число J является пределом интегральной суммы σ, т. е. функция $f(x)$ интегрируема на отрезке $[a, b]$ ■

В дальнейшем нам понадобится иная форма записи необходимого и достаточного условия интегрируемости.

Если обозначить колебание $M_i - m_i$ функции $f(x)$ на отрезке $[x_{i-1}, x_i]$ через ω_i, то будем иметь

$$S - s = \sum_{i=1}^{n} M_i \Delta x_i - \sum_{i=1}^{n} m_i \Delta x_i =$$

$$= \sum_{i=1}^{n} (M_i - m_i) \Delta x_i = \sum_{i=1}^{n} \omega_i \Delta x_i.$$

Так как $M_i \geqslant m_i$ и $\Delta x_i > 0$, то каждое слагаемое в последней сумме неотрицательно, и условие существования определенного интеграла может быть переписано так:

$$\sum_{i=1}^{n} \omega_i \Delta x_i \leqslant \varepsilon \text{ при } \lambda < \delta.$$

В таком виде оно обычно и применяется.

§ 3. Интегрируемость непрерывных и некоторых разрывных функций

Докажем следующую *основную* теорему.

Теорема 8.3. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на нем.*

Доказательство. Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то по теореме Кантора она и равномерно непрерывна на нем. Пусть дано любое $\varepsilon > 0$. В силу следствия из теоремы Кантора для положительного числа $\dfrac{\varepsilon}{b-a}$ всегда найдется $\delta > 0$ такое, что при разбиении отрезка $[a, b]$ на частичные отрезки $[x_{i-1}, x_i]$ с длинами $\Delta x_i < \delta$ все колебания ω_i будут меньше $\dfrac{\varepsilon}{b-a}$. Отсюда

$$S - s = \sum_{i=1}^{n} \omega_i \Delta x_i < \frac{\varepsilon}{b-a} \sum_{i=1}^{n} \Delta x_i = \varepsilon.$$

Следовательно, для непрерывной на отрезке $[a, b]$ функции $f(x)$ выполнено достаточное условие интегрируемости, а из него вытекает существование определенного интеграла*. ■

---

* Более простое доказательство теоремы см. кн.: Шипачев В. С. Основы высшей математики. М.,2001.

Как следует из теоремы, условие непрерывности функции является достаточным условием интегрируемости функции. Но это еще не значит, что определенный интеграл существует только для непрерывных функций. Класс интегрируемых функций этим не ограничивается. Так, например, существует определенный интеграл от функций, имеющих конечное число точек разрыва. Поэтому в заключение сформулируем и докажем более общую теорему существования определенного интеграла.

Теорема 8.4. *Если функция $f(x)$ ограничена на отрезке $[a, b]$ и непрерывна на нем всюду, кроме конечного числа точек, то она интегрируема на этом отрезке.*

Доказательство. Для доказательства достаточно рассмотреть случай, когда между a и b имеется лишь одна точка разрыва x'. Пусть M и m — точные грани функции $f(x)$ на $[a, b]$, $\Omega = M - m$ — ее колебание на данном отрезке. Возьмем любое достаточно малое $\varepsilon > 0$ и рассмотрим отрезки $\left[a,\ x' - \frac{\varepsilon}{8\Omega}\right]$ и $\left[x' + \frac{\varepsilon}{8\Omega},\ b\right]$ (рис. 88). На каждом из этих отрезков $f(x)$ непрерывна и, следовательно, найдется $\delta' > 0$ такое, что при разбиении их на частичные отрезки $[x'_{i-1}, x'_i]$ с длинами $\Delta x_i' < \delta'$ все колебания ω_i' будут меньше $\frac{\varepsilon}{2(b-a)}$.

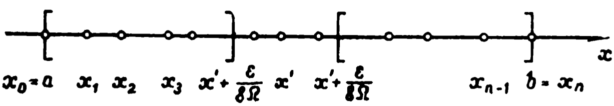


Рис. 88

Пусть $\delta = \min\left\{\delta', \frac{\varepsilon}{8\Omega}\right\}$. Рассмотрим теперь произвольное разбиение $[a, b]$ на частичные отрезки с длинами $\Delta x_i < \delta$ (см. рис. 88). Для этого разбиения сумму

$$S - s = \sum_{i=1}^{n} \omega_i \Delta x_i, \quad \omega_i = M_i - m_i$$

разобьем на два слагаемых: $\Sigma'\omega_i\Delta x_i + \Sigma''\omega_i\Delta x_i$, где в первую сумму входят частичные отрезки, лежащие целиком вне $\frac{\varepsilon}{8\Omega}$-окрестности точки x', а во вторую — частичные отрезки, либо заключенные целиком внутри $\frac{\varepsilon}{8\Omega}$-окрестности точки x', либо имеющие с ней общие точки.

Для первой суммы, как и в предыдущей теореме, будем иметь

$$\Sigma' \omega_i \Delta x_i < \frac{\varepsilon}{2}.$$

Что касается второй суммы, то заметим, что длины отрезков, целиком попавших внутрь $\frac{\varepsilon}{8\Omega}$-окрестности точки x', в сумме меньше или равны $\frac{\varepsilon}{4\Omega}$; число отрезков же лишь частично попавших может быть не больше двух, и сумма их длин меньше $2\delta < \frac{\varepsilon}{4\Omega}$. Следовательно,

$$\Sigma'' \omega_i \Delta x_i < \Omega \Sigma'' \Delta x_i < \Omega \frac{\varepsilon}{2\Omega} = \frac{\varepsilon}{2}.$$

Таким образом, окончательно имеем

$$S - s = \sum_{i=1}^{n} \omega_i \Delta x_i = \Sigma' \omega_i \Delta x_i + \Sigma'' \omega_i \Delta x_i < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Это и доказывает интегрируемость функции $f(x)$ на $[a, b]$ ■

Следствие. Кусочно-непрерывная на данном отрезке функция интегрируема на этом отрезке.

Замечание. Теперь мы можем ответить, почему функция Дирихле на отрезке $[0, 1]$ не интегрируема. Эта функция имеет бесконечное число точек разрыва на $[0, 1]$. В самом деле, пусть x_0 — любая точка отрезка $[0, 1]$. Будь она рациональная или иррациональная, в любой ее окрестности найдутся как рациональные, так и иррациональные точки. Следовательно, в любой окрестности точки x_0 функция будет иметь значения, равные 0 и 1. В таком случае не может существовать предела функции в точке x_0 ни слева, ни справа, т. е.

во всех точках [0, 1] функция Дирихле имеет разрыв второго рода. А так как отрезок [0, 1] содержит бесконечное множество точек, то и точек разрыва будет бесконечно.

§ 4. Основные свойства определенного интеграла

Докажем справедливость следующих свойств определенного интеграла.

1. Интеграл $\int\limits_a^b f(x)\,dx$ был введен для случая $a<b$. Обобщим понятие определенного интеграла на случай, когда $a=b$ и $a>b$. Если $a=b$, то по определению полагаем

$$\int\limits_a^a f(x)\,dx=0, \tag{1}$$

рассматривая эту формулу как естественное распространение понятия определенного интеграла на отрезок нулевой длины.

Если $a>b$, то также по определению полагаем

$$\int\limits_a^b f(x)\,dx=-\int\limits_b^a f(x)\,dx, \tag{2}$$

рассматривая формулу (2) как естественное распространение понятия определенного интеграла на случай, когда отрезок $[a, b]$ при $a<b$ пробегается в направлении от b к a. В этом случае точки разбиения x_i отрезка $[a, b]$ будут занумерованы в обратном порядке и в интегральной сумме все разности $\Delta x_i=x_i-x_{i-1}$ имеют отрицательный знак.

2. Каковы бы ни были числа a, b, c, всегда имеет место равенство

$$\int\limits_a^b f(x)\,dx=\int\limits_a^c f(x)\,dx+\int\limits_c^b f(x)\,dx \tag{3}$$

(здесь и в дальнейшем предполагается, что интегралы, входящие в доказываемые формулы, существуют).

Доказательство. Допустим сначала, что $a<c<b$. Так как предел интегральной суммы σ не зависит от

способа разбиения отрезка $[a, b]$, то будем разбивать так, чтобы точка c всегда была бы точкой разбиения $[a, b]$. Если, например, $c = x_m$, то σ можно разбить на две суммы:

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{m} f(\xi_i)\,\Delta x_i + \sum_{i=m+1}^{n} f(\xi_i)\,\Delta x_i.$$

Переходя в последнем равенстве к пределу при $\lambda \to 0$, мы и получим равенство (3).

Суть доказанного свойства состоит в том, что определенный интеграл по всему отрезку равен сумме интегралов по его частям.

Другое расположение a, b, c легко сводится к рассмотренному случаю. Пусть, например, $a < b < c$, тогда по доказанному будем иметь

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

откуда с учетом (2)

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

т. е. опять пришли к равенству (3) ■

3. Постоянный множитель можно выносить за знак определенного итеграла, т. е.

$$\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx. \qquad (4)$$

Доказательство. Действительно, для любого разбиения отрезка $[a, b]$ и любого выбора точек ξ_i

$$\sum_{i=1}^{n} kf(\xi_i)\,\Delta x_i = k\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i.$$

Переходя к пределу при $\lambda \to 0$, получаем

$$\int_a^b kf(x)\,dx = \lim_{\lambda\to 0}\sum_{i=1}^{n} kf(\xi_i)\,\Delta x_i = \lim_{\lambda\to 0} k\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i =$$

$$= k\lim_{\lambda\to 0}\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = k\int_a^b f(x)\,dx,$$

т. е. получили равенство (4) ■

4. Определенный интеграл от алгебраической суммы функций равен алгебраической сумме их интегралов, т. е.

$$\int_a^b [f(x) \pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$$

Доказательство. Действительно, для любого разбиения отрезка $[a, b]$ и любого выбора точек ξ_i

$$\sum_{i=1}^{n} [f(\xi_i) \pm g(\xi_i)]\,\Delta x_i = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \pm \sum_{i=1}^{n} g(\xi_i)\,\Delta x_i.$$

Так как

$$\lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \int_a^b f(x)\,dx \quad \text{и} \quad \lim_{\lambda \to 0} \sum_{i=1}^{n} g(\xi_i)\,\Delta x_i = \int_a^b g(x)\,dx,$$

то получаем

$$\int_a^b [f(x) \pm g(x)]\,dx = \lim_{\lambda \to 0} \sum_{i=1}^{n} [f(\xi_i) \pm g(\xi_i)]\,\Delta x_i =$$

$$= \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \pm \lim_{\lambda \to 0} \sum_{i=1}^{n} g(\xi_i)\,\Delta x_i =$$

$$= \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx \ \blacksquare$$

Замечание. Свойство 4 имеет место для любого конечного числа слагаемых.

§ 5. Оценки интегралов. Формула среднего значения

1. Оценки интегралов

1°. *Если всюду на отрезке* $[a, b]$ *функция* $f(x) \geqslant 0$, *то*

$$\int_a^b f(x)\,dx \geqslant 0.$$

Доказательство. В самом деле, любая интегральная сумма σ для функции $f(x)$ на $[a, b]$ неотрицательна, так как

$$f(\xi_i) \geqslant 0, \quad \Delta x_i = x_i - x_{i-1} > 0, \quad i = 1, 2, \ldots, n.$$

Переходя к пределу при $\lambda \to 0$ в неравенстве

$$\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \geqslant 0,$$

получаем

$$\int_a^b f(x)\,dx \geqslant 0 \quad \blacksquare$$

2°. *Если всюду на отрезке* $[a, b]$ $f(x) \leqslant g(x)$, *то*

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx. \tag{1}$$

Доказательство. Применяя оценку 1° к функции $g(x) - f(x) \geqslant 0$, получим

$$\int_a^b [g(x) - f(x)]\,dx \geqslant 0.$$

Но в силу свойства 4

$$\int_a^b [g(x) - f(x)]\,dx = \int_a^b g(x)\,dx - \int_a^b f(x)\,dx \geqslant 0,$$

откуда получаем неравенство (1) $\blacksquare$

3°. *Для функции* $f(x)$, *определенной на отрезке* $[a, b]$, *имеет место*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx. \tag{2}$$

Доказательство. Применив оценку 2° к очевидным неравенствам

$$-|f(x)| \leqslant f(x) \leqslant |f(x)|$$

и проинтегрировав их почленно, с учетом свойства 3 получим

$$-\int_a^b |f(x)|\,dx \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b |f(x)|\,dx,$$

а это равносильно неравенству (2) $\blacksquare$

Следствие. Если всюду на отрезке $[a, b]$, $a<b$, $|f(x)|\leqslant k$, то

$$\left|\int_a^b f(x)\,dx\right| \leqslant k(b-a).$$

Действительно, из неравенства $|f(x)|\leqslant k$ и оценок 2° и 3° следует

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx \leqslant \int_a^b k\,dx = k(b-a).$$

4°. *Если m и M — соответственно наименьшее и наибольшее значения функции $f(x)$ на отрезке $[a, b]$, $a<b$, то*

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a). \qquad (3)$$

Доказательство. По условию для любого $x \in [a, b]$ имеем

$$m \leqslant f(x) \leqslant M.$$

Применив оценку 2° к этим неравенствам и проинтегрировав их почленно, получим

$$m\int_a^b dx \leqslant \int_a^b f(x)\,dx \leqslant M\int_a^b dx;$$

отсюда, замечая, что

$$\int_a^b dx = \lim_{\lambda\to 0}\sum_{i=1}^{n} 1\cdot\Delta x_i = b-a,$$

получаем соотношение (3) ■

2. Формула среднего значения

Теорема 8.5 (о среднем). *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то на этом отрезке существует точка c такая, что*

$$\int_a^b f(x)\,dx = f(c)(b-a). \qquad (4)$$

Доказательство. Так как $f(x)$ непрерывна на $[a, b]$, то по второй теореме Вейерштрасса существуют числа m и M такие, что

$$\min_{[a,b]} f(x) = m \leqslant f(x) \leqslant M = \max_{[a,b]} f(x).$$

Отсюда в силу оценки 4° находим

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a),$$

откуда

$$m \leqslant \frac{\int_a^b f(x)\,dx}{b-a} \leqslant M.$$

Положим

$$\frac{\int_a^b f(x)\,dx}{b-a} = \mu \quad (m \leqslant \mu \leqslant M).$$

Так как μ заключено между наименьшим и наибольшим значениями непрерывной функции $f(x)$ на $[a, b]$ (рис. 89),

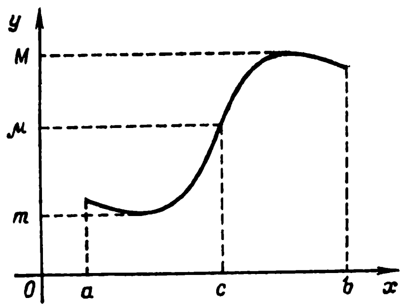


Рис. 89

то по теореме 4.9 о прохождении функции через любое промежуточное значение существует точка $c \in [a, b]$ такая, что $f(c) = \mu$, или

$$\frac{\int_a^b f(x)\,dx}{b-a} = f(c),$$

а это равносильно равенству (4) ■

Равенство (4) называется *формулой среднего значения*, а величину $f(c)$ принято называть средним значением функции $f(x)$ на отрезке $[a, b]$.

Замечание. Теорема о среднем имеет определенный геометрический смысл: величина определенного интеграла при $f(x) \geqslant 0$ равна площади прямоугольника, имеющего высоту $f(c)$ и основание $b - a$.

§ 6. Определенный интеграл с переменным верхним пределом

До сих пор мы рассматривали определенный интеграл с постоянными пределами интегрирования a и b. Если изменять, например, верхний предел, не выходя

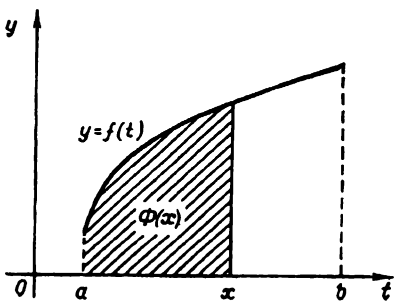


Рис. 90

из отрезка $[a, b]$, величина интеграла будет меняться. Другими словами, интеграл с переменным верхним пределом представляет собой функцию своего верхнего предела. Таким образом, если мы имеем интеграл

$$\int_a^x f(t)\, dt^* \ (a \leqslant x \leqslant b)$$

с постоянным нижним пределом a и переменным верхним пределом x, то величина этого интеграла будет функцией верхнего предела x. Обозначим эту функцию

---

* Для удобства переменную интегрирования мы обозначили буквой t, так как буквой x обозначен верхний предел интегрирования.

через $\Phi(x)$, т. е. положим

$$\Phi(x)=\int_a^x f(t)\,dt \tag{1}$$

(рис. 90) и назовем ее *определенным интегралом с переменным верхним пределом.* Геометрически функция $\Phi(x)$ представляет собой площадь заштрихованной криволинейной трапеции, если $f(x)>0$.

Теперь остановимся на теореме, которую следует считать одной из основных теорем математического анализа.

Теорема 8.6. *Производная определенного интеграла от непрерывной функции по переменному верхнему пределу существует и равна значению подынтегральной функции в верхнем пределе, т. е.*

$$\left(\int_a^x f(t)\,dt\right)'_x=f(x),\ \text{или}\ (\Phi(x))'=f(x). \tag{2}$$

Доказательство. Возьмем любое значение $x\in[a, b]$ и придадим ему приращение $\Delta x\neq 0$ такое, чтобы $x+\Delta x\in[a, b]$, т. е. $a\leqslant x+\Delta x\leqslant b$. Тогда функция $\Phi(x)$, определенная выражением (1), получит новое значение:

$$\Phi(x+\Delta x)=\int_a^{x+\Delta x} f(t)\,dt.$$

В силу свойства 2 определенного интеграла (см. § 4) будем иметь

$$\Phi(x+\Delta x)=\int_a^x f(t)\,dt+\int_x^{x+\Delta x} f(t)\,dt=\Phi(x)+\int_x^{x+\Delta x} f(t)\,dt.$$

Отсюда находим приращение функции $\Phi(x)$:

$$\Phi(x+\Delta x)-\Phi(x)=\int_x^{x+\Delta x} f(t)\,dt.$$

Применяя теорему 8.5, получим

$$\Phi(x+\Delta x)-\Phi(x)=f(c)\,\Delta x,$$

где c — число, заключенное между числами x и $x+\Delta x$. Разделим обе части равенства на Δx:

$$\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x}=f(c).$$

Если теперь $\Delta x \to 0$, то $c \to x$, и тогда, в силу непрерывности функции $f(x)$ на $[a, b]$, $f(c) \to f(x)$. Поэтому, переходя к пределу в последнем равенстве при $\Delta x \to 0$, получаем

$$\lim_{\Delta x \to 0}\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x}=\lim_{\Delta x \to 0} f(c)=\lim_{c \to x} f(c)=f(x),$$

или $(\Phi(x))'=f(x)$ ■

Таким образом, мы установили, что любая непрерывная на отрезке $[a, b]$ функция $f(x)$ имеет на этом отрезке первообразную.

§ 7. Формула Ньютона — Лейбница

Вычисление определенных интегралов методом, основанным на определении интеграла как предела интегральной суммы, как правило, связано с большими трудностями. Поэтому существует другой практически более удобный метод вычисления определенных интегралов, который, как мы увидим, основан на тесной связи, существующей между понятиями неопределенного и определенного интегралов.

Как мы установили выше, функция $f(x)$, непрерывная на отрезке $[a, b]$, имеет на этом отрезке первообразные, причем одной из них является функция

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

Пусть $F(x)$ — любая другая первообразная для функции $f(x)$ на том же отрезке $[a, b]$. Так как первообразные $\Phi(x)$ и $F(x)$ отличаются на постоянную (см. теорему 7.1), то имеет место равенство

$$\int_a^x f(t)\,dt=F(x)+C,\ a \leqslant x \leqslant b,$$

где C — некоторое число. Подставляя в это равенство значение $x=a$ и используя формулу (1) из § 4, будем иметь

$$\int_a^a f(t)\,dt = F(a) + C,\ 0 = F(a) + C,\ C = -F(a),$$

т. е. для любого $x \in [a, b]$ имеем

$$\int_a^x f(t)\,dt = F(x) - F(a).$$

Полагая $x=b$, получаем соотношение

$$\int_a^b f(x)\,dx = F(b) - F(a), \qquad (1)$$

являющееся основной формулой интегрального исчисления, которая называется формулой Ньютона—Лейбница.

Разность $F(b)-F(a)$ принято условно записывать в виде

$$F(x)\big|_a^b, \text{ или } [F(x)]_a^b,$$

и тогда формула (1) принимает вид

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

Необходимо еще раз подчеркнуть, что в формуле (1) в качестве $F(x)$ может быть любая первообразная функции $f(x)$ из семейства $F(x)+C$.

Итак, полученная нами формула (1), с одной стороны, устанавливает связь между определенным и неопределенным интегралами, с другой стороны, она дает простой метод вычисления определенного интеграла: определенный интеграл от непрерывной функции равен разности значений любой ее первообразной, вычисленных для верхнего и нижнего пределов интегрирования. Эта формула открывает широкие возможности для вычисления определенных интегралов, ибо задача вычисления определенного интеграла сводится к задаче вычисления неопределенного интеграла, которая достаточно полно нами изучена.

Рассмотрим примеры:

1) $\int\limits_a^b \sin x dx = -\cos x \,|_a^b = \cos a - \cos b,$

2) $\int\limits_0^2 (3x^2 - 1)\, dx = [x^3 - x]_0^2 = (2^3 - 2) - (0^3 - 0) = 6,$

3) $\int\limits_1^2 \frac{dx}{x} = \ln x \,|_1^2 = \ln 2 - \ln 1 = \ln 2,$

4) $\int\limits_{-1}^1 \frac{dx}{1+x^2} = \operatorname{arctg} x \,|_{-1}^1 = \operatorname{arctg} 1 - \operatorname{arctg}(-1) =$

$= \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2},$

5) $\int\limits_0^3 \frac{dx}{\sqrt{1+x^2}} = \ln (x + \sqrt{1+x^2}) \,|_0^3 = \ln (3 + \sqrt{10}).$

§ 8. Замена переменной в определенном интеграле

Теорема 8.7. *Пусть $f(x)$ — непрерывная функция на отрезке $[a, b]$. Тогда, если: 1) функция $x = \varphi(t)$ дифференцируема на $[\alpha, \beta]$ и $\varphi'(t)$ непрерывна на $[\alpha, \beta]$, 2) множеством значений функции $x = \varphi(t)$ является отрезок $[a, b]$, 3) $\varphi(\alpha) = a$ и $\varphi(\beta) = b$ (рис. 91), то справедлива формула*

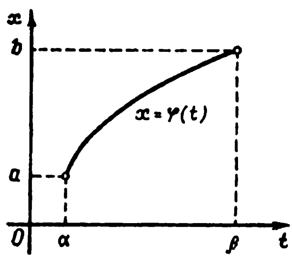


Рис. 91

$$\int\limits_a^b f(x)\, dx = \int\limits_\alpha^\beta f[\varphi(t)]\, \varphi'(t)\, dt. \qquad (1)$$

Доказательство. По формуле Ньютона—Лейбница

$$\int\limits_a^b f(x)\, dx = F(b) - F(a),$$

где $F(x)$ — какая-нибудь первообразная для функции $f(x)$ на $[a, b]$. С другой стороны, рассмотрим сложную функцию от переменной t: $\Phi(t) = F[\varphi(t)]$. Согласно правилу дифференцирования сложной функции находим

$$(\Phi(t))' = F'[\varphi(t)]\varphi'(t) = f[\varphi(t)]\varphi'(t).$$

Отсюда следует, что функция $\Phi(t)$ является первообразной для функции $f[\varphi(t)]\varphi'(t)$, непрерывной на $[\alpha, \beta]$, и поэтому согласно формуле Ньютона—Лейбница получаем

$$\int_{\alpha}^{\beta} f[\varphi(t)]\varphi'(t)\,dt = \Phi(\beta) - \Phi(\alpha) = F[\varphi(\beta)] - F[\varphi(\alpha)] =$$

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

Этим доказана справедливость формулы (1) ■

Формула (1) называется *формулой замены переменной* или *подстановки* в определенном интеграле.

Замечание 1. Если при вычислении неопределенного интеграла с помощью замены переменной мы должны были от новой переменной t возвращаться к старой переменной x, то при вычислении определенного интеграла этого можно не делать.

Пример 1. Вычислить $\int_0^1 \sqrt{1-x^2}\,dx$.

Решение. Рассмотрим подстановку $x = \sin t$, $0 \leqslant t \leqslant \frac{\pi}{2}$. Проверим законность такой подстановки.

Во-первых, $f(x) = \sqrt{1-x^2}$ непрерывна на $[0, 1]$; во-вторых, функция $x = \sin t$ дифференцируема на $\left[0, \frac{\pi}{2}\right]$ и $x'_t = \cos t$ непрерывна на $\left[0, \frac{\pi}{2}\right]$ и, в-третьих, при изменении t от 0 до $\frac{\pi}{2}$ функция $x = \sin t$ возрастает от 0 до 1, при этом $\varphi(0) = 0$ и $\varphi\left(\frac{\pi}{2}\right) = 1$. Таким образом, данная подстановка удовлетворяет всем условиям теоремы 8.7. Применяя формулу (1), получаем

$$\int_0^1 \sqrt{1-x^2}\,dx = \int_0^{\frac{\pi}{2}} \sqrt{1-\sin^2 t}\cos t\,dt =$$

$$= \int_0^{\frac{\pi}{2}} \cos^2 t\,dt = \frac{1}{2}\int_0^{\frac{\pi}{2}} (1+\cos 2t)\,dt = \frac{\pi}{4}.$$

Замечание 2. При использовании формулы (1) необходимо проверять выполнение перечисленных в теореме условий. Если эти условия нарушаются, замена переменной по указанной формуле может привести к неверному результату.

Пример 2. $\int_0^{\pi} dx = x\Big|_0^{\pi} = \pi.$

Однако то же выражение можно записать в виде

$$\int_0^{\pi} dx = \int_0^{\pi} \frac{dx}{\sin^2 x + \cos^2 x} = \int_0^{\pi} \frac{dx}{\cos^2 x(1+\operatorname{tg}^2 x)}.$$

Подстановка $\operatorname{tg} x = t$ формально приводит к следующему результату:

$$\int_0^{\pi} dx = \int_0^{\pi} \frac{d(\operatorname{tg} x)}{1+\operatorname{tg}^2 x} = \int_0^{0} \frac{dt}{1+t^2} = 0.$$

Как видим, пришли к абсурду. Это произошло потому, что функция $t = \operatorname{tg} x$ разрывна при $x = \frac{\pi}{2}$ и не удовлетворяет условиям теоремы 8.7.

§ 9. Формула интегрирования по частям в определенном интеграле

Теорема 8.8. *Если функции $u(x)$ и $v(x)$ имеют непрерывные производные на отрезке $[a, b]$, то имеет место формула*

$$\int_a^b u\,dv = uv\Big|_a^b - \int_a^b v\,du. \qquad (1)$$

Доказательство. Так как функция $u(x)v(x)$ является первообразной для функции $[u(x)v(x)]' = u(x)v'(x)+v(x)u'(x)$, то по формуле Ньютона—Лейбница:

$$\int_a^b [u(x)v'(x) + v(x)u'(x)]\,dx = [u(x)v(x)]\,\Big|_a^b.$$

Отсюда, используя свойство 4 определенного интеграла (см. § 4), получим

$$\int_a^b u(x)v'(x)\,dx + \int_a^b v(x)u'(x)\,dx = [u(x)v(x)]\,\Big|_a^b$$

или, что то же, в виде

$$\int_a^b udv + \int_a^b vdu = uv\,\Big|_a^b,$$

откуда и следует формула (1) ■

Формула (1) называется *формулой интегрирования по частям* в определенном интеграле.

Пример 1. Вычислить $\int_1^e \ln x dx$.

Решение. Положим $u = \ln x$, $dv = dx$; отсюда $du = \frac{dx}{x}$, $v = x$, и по формуле (1) находим

$$\int_1^e \ln x dx = x\ln x\,\Big|_1^e - \int_1^e x\,\frac{dx}{x} = [x\,\mathrm{lu}\,x - x]_1^e = 1.$$

Пример 2. Вычислить $\int_1^2 xe^x dx$.

Решение. Положим $u=x$, $dv=e^x dx$, отсюда $du=dx$, $v=e^x$, и по формуле (1) находим

$$\int_1^2 xe^x dx = xe^x\,\Big|_1^2 - \int_1^2 e^x dx = [e^x(x-1)]_1^2 = e^2.$$

Пример 3. Вычислить $\int\limits_0^1 \operatorname{arctg} x dx$.

Решение. Положим $u = \operatorname{arctg} x$, $dv = dx$; отсюда $du = \frac{dx}{1+x^2}$, $v = x$, и по формуле (1) находим

$$\int\limits_0^1 \operatorname{arctg} x dx = x \operatorname{arctg} x \Big|_0^1 - \int\limits_0^1 \frac{x dx}{1+x^2} =$$

$$= \left[x \operatorname{arctg} x - \frac{1}{2}\ln(1+x^2) \right]_0^1 = \frac{\pi}{4} - \ln\sqrt{2}$$

§ 10. Некоторые физические и геометрические приложения определенного интеграла

1. Площадь криволинейной трапеции. Пусть на плоскости Oxy дана фигура, ограниченная отрезком $[a, b]$ оси Ox, прямыми $x=a$, $x=b$ и графиком непрерывной

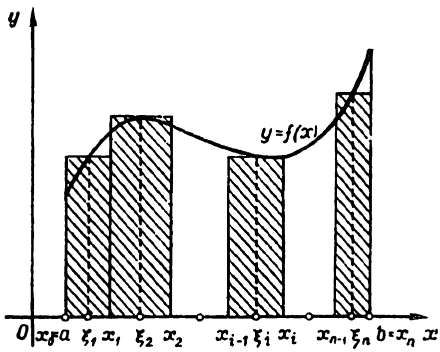


Рис. 92

и неотрицательной функции $y=f(x)$ на $[a, b]$. Такую фигуру, как мы знаем, называют криволинейной трапецией, площадь S* которой может быть вычислена по формуле

$$S = \int\limits_a^b f(x)\, dx. \qquad (1)$$

* О понятии площади произвольной плоской фигуры (а также объема тела и площади поверхности) см. кн.: Ильин В. А., Позняк Э. Г. Основы математического анализа. Т. 1. М., 2001.

Доказательство. Разобьем произвольно отрезок $[a, b]$ на n частей точками $a=x_0<x_1<x_2<\ldots<x_{i-1}<<x_i<\ldots<x_n=b$, выберем на каждом частичном отрезке $[x_{i-1}, x_i]$, $i=1, 2, \ldots, n$ произвольно точку ξ_i $(x_{i-1}\leqslant\xi_i\leqslant x_i)$ и рассмотрим ступенчатую фигуру (рис. 92). Ее площадь будем считать приближенно равной площади S криволинейной трапеции:

$$S \approx \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i,$$

где $\Delta x_i = x_i - x_{i-1}$. Таким образом, мы получили интегральную сумму σ для интеграла (1). Так как функция $f(x)$ непрерывна на $[a, b]$, то предел этой суммы существует при $\lambda = \max\limits_{1\leqslant i\leqslant n}\{\Delta x_i\} \to 0$ и площадь S криволинейной трапеции численно равна определенному интегралу от функции $f(x)$ на $[a, b]$:

$$S = \lim_{\lambda\to 0}\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \int_a^b f(x)\,dx \blacksquare$$

Итак, определенный интеграл от неотрицательной непрерывной функции $f(x)$ на $[a, b]$ численно равен площади криволинейной трапеции с основанием $[a, b]$, ограниченной сверху графиком функции $y=f(x)$. В этом заключается *геометрический смысл определенного интеграла.*

Пример 1. Найти площадь фигуры, ограниченной графиком функции $y=x^\alpha$, $\alpha>0$, прямой $x=1$ и осью Ox (рис. 93).

Решение. По формуле (1) имеем

$$S = \int_0^1 x^\alpha dx = \left.\frac{x^{\alpha+1}}{\alpha+1}\right|_0^1 = \frac{1}{\alpha+1}.$$

При этом, если $\alpha=1$, то $S=\frac{1}{2}$; если $\alpha=2$, то $S=\frac{1}{3}$ и т. д.

Пусть фигура ограничена снизу и сверху графиками функций $y_1=f_1(x)$, $y_2=f_2(x)$ и $f_1(x)\leqslant f_2(x)$, $a\leqslant x\leqslant b$ (рис. 94), где $f_1(x)$, $f_2(x)$ — две непрерывные функции. Если обе функции неотрицательны, то площадь данной фигуры является разностью площадей криволинейных трапеций, ограниченных сверху соответственно графиками функций $y_2=f_2(x)$, $y_1=f_1(x)$, прямыми $x=a$, $x=b$ и осью Ox. Следовательно, площадь S данной фигуры равна

$$S=\int_a^b f_2(x)\,dx-\int_a^b f_1(x)\,dx=\int_a^b [f_2(x)-f_1(x)]\,dx. \qquad (2)$$

Заметим, что формула (2) справедлива и тогда, когда $f_1(x)$ и $f_2(x)$ неположительны, так как в силу

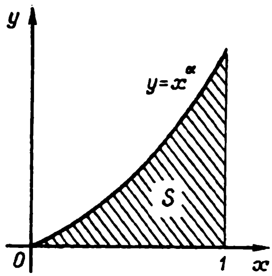


Рис. 93

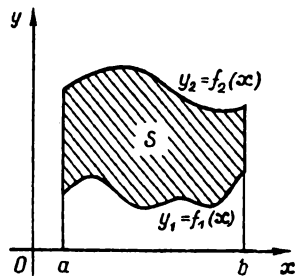


Рис. 94

их ограниченности существует число h такое, что функции

$$f_1^*(x)=f_1(x)+h,\ f_2^*(x)=f_2(x)+h$$

становятся неотрицательными и имеет место очевидное равенство

$$\int_a^b [f_2^*(x)-f_1^*(x)]\,dx=\int_a^b [f_2(x)-f_1(x)]\,dx.$$

Пример 2. Вычислить площадь фигуры, ограниченной графиками функций $y_1=f_1(x)=x$ и $y_2=f_2(x)=$ $=2-x^2$ (рис. 95).

Решение. Найдем абсциссы точек пересечения прямой с параболой. Решая систему

$$\begin{cases} y=x, \\ y=2-x^2, \end{cases}$$

получим $x_1=-2$, $x_2=1$ — это и будут пределы интегрирования. Для искомой площади фигуры согласно формуле (2) будем иметь

$$S = \int_{-2}^{1} [f_2(x) - f_1(x)]\,dx = \int_{-2}^{1} [(2 - x^2) - x]\,dx =$$

$$= \left[2x - \frac{x^3}{2} - \frac{x^2}{2}\right]_{-2}^{1} = \frac{9}{2}.$$

Замечание. При вычислении площади криволинейной трапеции в случае, когда верхняя граница за-

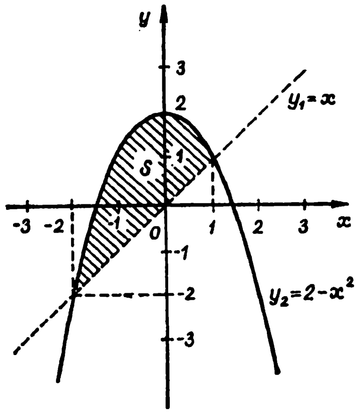


Рис. 95

дана параметрическими уравнениями $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leqslant t \leqslant \beta$, в формуле (1) надо сделать замену переменной, положив $x = \varphi(t)$, $dx = \varphi'(t)\,dt$. Тогда получим

$$S = \int_{\alpha}^{\beta} \psi(t)\,\varphi'(t)\,dt,$$

где α и β — значения параметра t, соответствующие значениям $x = a$ и $x = b$, т. е. $a = \varphi(\alpha)$, $b = \varphi(\beta)$.

Пример 3. Вычислить площадь фигуры, ограниченной эллипсом

$$\begin{cases} x = a\cos t, \\ y = b\sin t, \end{cases} \quad 0 \leqslant t \leqslant 2\pi.$$

Решение. Ввиду симметрии данной кривой относительно осей координат достаточно вычислить площадь части фигуры, находящейся в первой четверти

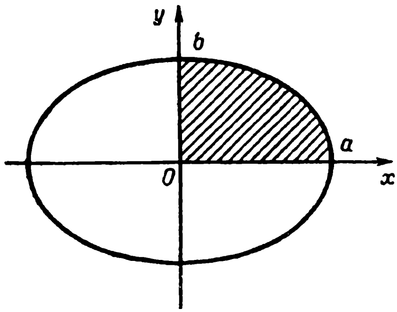


Рис. 96

(рис. 96). Следовательно, искомая площадь равна

$$S=4\int_{\pi/2}^{0} b\sin t\,(a\cos t)'\,dt=4ab\int_{0}^{\pi/2}\sin^2 t\,dt=$$

$$=2ab\int_{0}^{\pi/2}(1-\cos 2t)\,dt=2ab\left[t-\frac{1}{2}\sin 2t\right]_0^{\pi/2}=\pi ab.$$

2. Площадь криволинейного сектора. Пусть кривая AB задана в полярных координатах уравнением

$$\rho=\rho(\varphi),\quad \alpha\leqslant\varphi\leqslant\beta,$$

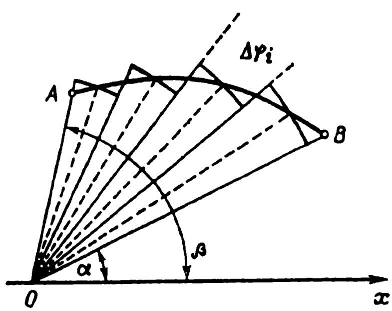


Рис. 97

причем функция $\rho(\varphi)$ непрерывна и неотрицательна на отрезке $[\alpha, \beta]$. Плоскую фигуру, ограниченную кривой AB и двумя полярными радиусами, составляющими с полярной осью углы α и β, мы будем называть криволинейным сектором (рис. 97). Площадь S криволинейного сектора может быть вычислена по формуле

$$S = \frac{1}{2} \int_{\alpha}^{\beta} \rho^2 (\varphi)\, d\varphi. \tag{3}$$

Доказательство. Разобьем произвольно отрезок $[\alpha, \beta]$ на n частей точками

$$\alpha = \varphi_0 < \varphi_1 < \varphi_2 < \ldots < \varphi_{i-1} < \varphi_i < \ldots < \varphi_n = \beta,$$

выберем на каждом частичном отрезке $[\varphi_{i-1}, \varphi_i]$, $i = 1, 2, \ldots, n$ произвольно точку $\xi_i (\varphi_{i-1} \leqslant \xi_i \leqslant \varphi_i)$ и построим круговые секторы с радиусами $\rho(\xi_i)$. В результате мы получим веерообразную фигуру, площадь которой будем считать приближенно равной площади S криволинейного сектора:

$$S \approx \frac{1}{2} \sum_{i=1}^{n} \rho^2 (\xi_i)\, \Delta\varphi_i,$$

где $\Delta\varphi_i = \varphi_i - \varphi_{i-1}$. Таким образом, получили интегральную сумму σ для интеграла (3). Так как функция $\rho^2(\varphi)$ непрерывна на отрезке $[\alpha, \beta]$, то предел этой суммы существует при $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta\varphi_i\} \to 0$ и площадь криволинейного сектора численно равна определенному интегралу от функции $\rho^2(\varphi)$ на $[\alpha, \beta]$:

$$S = \frac{1}{2} \lim_{\lambda \to 0} \sum_{i=1}^{n} \rho^2 (\xi_i)\, \Delta\varphi_i = \frac{1}{2} \int_{\alpha}^{\beta} \rho^2 (\varphi)\, d\varphi.$$

Отсюда следует справедливость формулы (3) ■

Пример 4. Вычислить площадь фигуры, ограниченной полярной осью и первым витком спирали Архимеда $\rho = a\varphi$, где a — положительное число (рис. 98).

Решение. Когда φ изменяется от 0 до 2π, то полярный радиус ρ описывает кривую, ограничивающую криволинейный сектор $OABC$. Поэтому по формуле (3) имеем

$$S_{OABC} = \frac{a^2}{2}\int_0^{2\pi}\varphi^2 d\varphi = \frac{a^2}{2}\cdot\frac{\varphi^3}{3}\Big|_0^{2\pi} = \frac{a^2}{2}\cdot\frac{8\pi^3}{3} = \frac{4}{3}\pi^3 a^2.$$

Заметим, что точка C отдалена от полюса на расстояние $\rho = 2\pi a$. Поэтому круг радиуса OC имеет площадь $\pi\cdot OC^2 = 4\pi^3 a^2 = 3\cdot\frac{4}{3}\pi^3 a^2 = 3S_{OABC}$, т. е. площадь фигуры, ограниченной полярной осью и первым витком спирали Архимеда, равна 1/3 площади круга с радиусом, равным наибольшему из полярных радиусов витка. К этому выводу пришел еще Архимед.

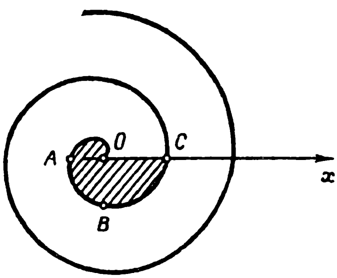


Рис. 98

3. Длина дуги кривой. Пусть плоская кривая AB задана уравнением $y = f(x)$, $a \leqslant x \leqslant b$, где $f(x)$ — непрерывная функция на отрезке $[a, b]$. Разобьем кривую AB на n произвольных частей точками $A = M_0$, M_1,

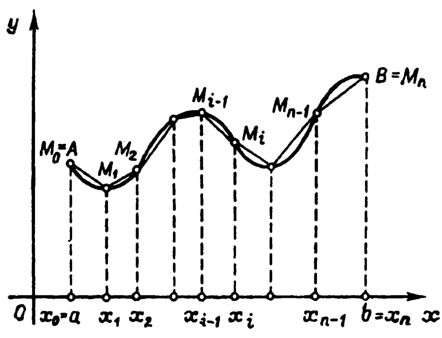


Рис. 99

M_2, ..., M_{i-1}, M_i, ..., $M_n = B$ в направлении от A к B. Соединив эти точки хордами, получим некоторую вписанную ломаную линию, периметр которой обозначим через P (рис. 99). Через l_i обозначим длину одного звена

$M_{i-1}M_i$ ломаной линии, а через μ — длину наибольшего из ее звеньев: $\mu = \max\limits_{1\leq i\leq n}\{l_i\}$.

Определение. Число L называется *пределом* периметров P при $\mu\to 0$, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что для всякой ломаной, у которой $\mu<\delta$, выполняется неравенство

$$|L-P|<\varepsilon.$$

Если существует конечный предел L периметра P вписанной в кривую ломаной линии при $\mu\to 0$, то этот предел называется длиной дуги $\overset{\frown}{AB}$:

$$L = \lim_{\mu\to 0} P.$$

Если функция $f(x)$ непрерывна вместе с $f'(x)$ на отрезке $[a, b]$, то длина дуги AB выражается формулой

$$L = \int_a^b \sqrt{1+f'^2(x)}\,dx. \tag{4}$$

Доказательство. Обозначим через x_i и $f(x_i)$ координаты точки M_i, так что для абсцисс этих точек получим: $a=x_0<x_1<x_2<\ldots<x_{i-1}<x_i<\ldots<x_n=b$. Тогда длина одного звена ломаной будет равна

$$l_i = \sqrt{(x_i-x_{i-1})^2+[f(x_i)-f(x_{i-1})]^2}.$$

По формуле Лагранжа

$$f(x_i)-f(x_{i-1})=f'(\xi_i)(x_i-x_{i-1}),\quad x_{i-1}<\xi_i<x_i.$$

Следовательно,

$$l_i = \sqrt{1+f'^2(\xi_i)}\,\Delta x_i,\quad \Delta x_i = x_i - x_{i-1}.$$

Таким образом, периметр всей ломаной равен

$$P = \sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \sqrt{1+f'^2(\xi_i)}\,\Delta x_i,$$

т. е. получили интегральную сумму σ для интеграла (4). Так как функция $\sqrt{1+f'^2(x)}$ непрерывна на $[a, b]$, то предел этой суммы при $\lambda = \max\limits_{1\leq i\leq n}\{\Delta x_i\}\to 0$ существует и численно равен соответствующему определенному интегралу. Так как $\lambda \leqslant \mu$, то $\lambda \to 0$ при $\mu \to 0$ *. Поэтому

$$\lim_{\mu \to 0} P = \lim_{\lambda \to 0} \sum_{i=1}^{n} \sqrt{1+f'^2(\xi_i)}\, \Delta x_i = \int_a^b \sqrt{1+f'^2(x)}\, dx = L \blacksquare$$

Пример 5. Вычислить длину дуги верхней ветви полукубической параболы $y=x^3/^2$ от $x=0$ до $x=5$ (рис. 100).

Решение. Из уравнения $y=x^{\frac{3}{2}}$ находим $y'=\frac{3}{2}x^{1/2}$ Следовательно, по формуле (4) получим

$$L=\int_0^5 \sqrt{1+y'^2}\, dx = \int_0^5 \sqrt{1+\frac{9x}{4}}\, dx =$$

$$= \frac{8}{27}\left(1+\frac{9x}{4}\right)^{3/2}\Bigg|_0^5 = \frac{335}{27}.$$

Замечание 1. При вычислении длины дуги в случае, когда кривая AB задана параметрическими уравнениями

$$x=\varphi(t),\ y=\psi(t),\ \alpha \leqslant t \leqslant \beta,$$

где α и β — значения параметра t, соответствующие значениям $x=a$ и $x=b$, т. е. $a=\varphi(\alpha)$, $b=\varphi(\beta)$, в формуле

$$L=\int_a^b \sqrt{1+y'^2(x)}\, dx$$

надо сделать замену переменной, положив $x=\varphi(t)$, $dx=\varphi'(t)dt$. Тогда получим

$$L=\int_a^b \sqrt{1+y'^2(x)}\, dx = \int_\alpha^\beta \sqrt{1+\left(\frac{\psi'(t)}{\varphi'(t)}\right)^2}\, \varphi'(t)\, dt =$$

$$= \int_\alpha^\beta \sqrt{\varphi'^2(t)+\psi'^2(t)}\, dt. \qquad (5)$$

---

* $l_i=\sqrt{(\Delta x_i)^2+(\Delta y_i)^2}$, откуда $|\Delta x_i| \leqslant l_i$.

Пример 6. Вычислить длину дуги одной арки циклоиды*: $x=a(t-\sin t)$, $y=a(1-\cos t)$, $0\leqslant t\leqslant 2\pi$ (рис. 101).

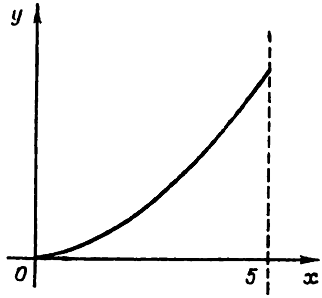


Рис. 100

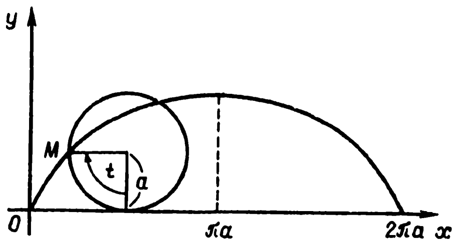


Рис. 101

Решение. Из уравнения циклоиды находим, что $\varphi'(t)=a(1-\cos t)$, $\psi'(t)=a\sin t$. Когда x пробегает отрезок $[0,\ 2\pi a]$, параметр t пробегает отрезок $[0,\ 2\pi]$. Следовательно, искомая длина дуги равна

$$L=\int_0^{2\pi a}\sqrt{1+y'^2(x)}\,dx=\int_0^{2\pi}\sqrt{\varphi'^2(t)+\psi'^2(t)}\,dt=$$

---

* Циклоида — плоская кривая, которую описывает точка M окружности радиуса a, катящейся без скольжения по прямой линии.

$$= \int_0^{2\pi} a\sqrt{(1-\cos t)^2+\sin^2 t}\,dt = 2a\int_0^{2\pi} \sin\frac{t}{2}\,dt =$$

$$= -4a\cos\frac{t}{2}\Big|_0^{2\pi} = 8a.$$

Замечание 2. При вычислении длины дуги в случае, когда кривая AB задана в полярных координатах уравнением $\rho=\rho(\varphi)$, $\alpha\leqslant\varphi\leqslant\beta$, где $\rho(\varphi)$ имеет непрерывную производную $\rho'(\varphi)$ на отрезке $[\alpha, \beta]$ и точкам A и B соответствуют значения α и β, переходя от полярных координат (см. гл. III, § 3, (1)) к прямоугольным, получим параметрическое задание кривой AB уравнениями $x=\rho\cos\varphi$, $y=\rho\sin\varphi$ с параметром φ. Тогда

$$x'(\varphi)=\rho'(\varphi)\cos\varphi-\rho\sin\varphi,\quad y'(\varphi)=\rho'(\varphi)\sin\varphi+\rho(\varphi)\cos\varphi$$

и формула (5) примет вид

$$L=\int_\alpha^\beta \sqrt{\rho(\varphi)+\rho'^2(\varphi)}\,d\varphi, \tag{6}$$

где α и β — значения параметра φ.

Пример 7. Вычислить длину первого витка архимедовой спирали $\rho=a\varphi$ (см. рис. 98).

Решение. Первый виток архимедовой спирали образуется при изменении полярного угла φ от 0 до 2π. Тогда по формуле (6) искомая длина дуги равна

$$L=\int_0^{2\pi}\sqrt{a^2\varphi^2+a^2}\,d\varphi = a\int_0^{2\pi}\sqrt{\varphi^2+1}\,d\varphi =$$

$$= a\left[\pi\sqrt{4\pi^2+1}+\frac{1}{2}\ln\left(2\pi+\sqrt{4\pi^2+1}\right)\right].$$

4. Объем тела вращения. Пусть функция $f(x)$ непрерывна и неотрицательна на отрезке $[a, b]$. Тогда тело, образованное вращением вокруг оси Ox криволинейной трапеции, ограниченной сверху графиком функции $f(x)$, имеет объем V, который может быть найден по формуле

$$V=\pi\int_a^b y^2dx. \tag{7}$$

Доказательство. Разобьем произвольно отрезок $[a, b]$ на n частей точками $a=x_0<x_1<x_2<\ldots<x_{i-1}<<x_i<\ldots<x_n=b$. На каждом частичном отрезке $[x_{i-1}, x_i]$ построим прямоугольник (рис. 102), который

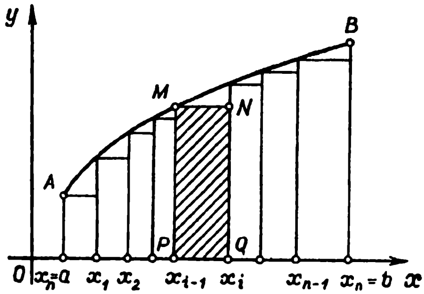


Рис. 102

при вращении вокруг оси Ox опишет цилиндр. Найдем объем i-того цилиндра, образованного вращением прямоугольника $PMNQ$:

$$v_i = \pi y_{i-1}^2 \Delta x_i,$$

где $\Delta x_i = x_i - x_{i-1}$.

Сумма объемов всех n цилиндров будет приближенно равна объему V данного тела вращения

$$V \approx \sum_{i=1}^{n} \pi y_{i-1}^2 \Delta x_i.$$

Таким образом, получили интегральную сумму σ для интеграла (7). Так как функция $y^2 = [f(x)]^2$ непрерывна на $[a, b]$, то предел этой суммы при $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\} \to 0$ существует и численно равен соответствующему определенному интегралу, т. е.

$$V = \lim_{\lambda \to 0} \sum_{i=1}^{n} \pi y_{i-1}^2 \Delta x_i = \pi \int_a^b y^2 dx \blacksquare$$

Пример 8. Вычислить объем тора. Тором называется тело, получающееся при вращении круга радиуса a вокруг оси, лежащей в его плоскости на расстоянии b от центра ($b \geqslant a$). Форму тора имеет, например, баранка.

Решение. Пусть круг вращается вокруг прямой AE (рис. 103). Тогда объем тора может быть рассмотрен как разность объемов вращения криволинейных трапеций $ABCDE$ и $ABLDE$ вокруг оси Ox.

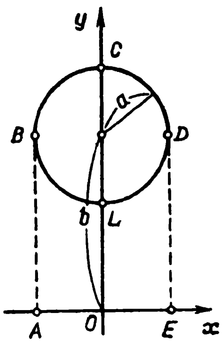


Рис. 103

Если оси координат выбрать, как показано на рис. 103, то уравнение окружности $LBCD$ будет иметь вид

$$x^2+(y-b)^2=a^2 \quad (b>a),$$

где a — радиус круга, причем уравнение кривой BCD

$$y=y_1(x)=b+\sqrt{a^2-x^2},$$

а уравнение кривой BLD

$$y=y_2(x)=b-\sqrt{a^2-x^2}.$$

Тогда объем тора по формуле (7) будет равен

$$V=2\pi\int_0^a y_1^2\,dx-2\pi\int_0^a y_2^2\,dx=2\pi\int_0^a (y_1^2-y_2^2)\,dx=$$

$$=2\pi\int_0^a [(b+\sqrt{a^2-x^2})^2-(b-\sqrt{a^2-x^2})^2]\,dx=$$

$$=8\pi b\int_0^a \sqrt{a^2-x^2}\,dx=2\pi^2a^2b.$$

5. Площадь поверхности вращения. Пусть функция $f(x)$ неотрицательна и непрерывна вместе со своей первой производной на отрезке $[a, b]$. Тогда поверхность, образованная вращением графика этой функции вокруг оси Ox, имеет площадь P, которая может быть вычислена по формуле

$$P=2\pi\int_a^b f(x)\sqrt{1+f'^2(x)}\,dx. \qquad (8)$$

Доказательство. Разобьем произвольно отрезок $[a, b]$ на n частей точками $a=x_0<x_1<x_2<\ldots<x_{i-1}<<x_i<\ldots<x_n=b$, и пусть $A_0, A_1, A_2, \ldots, A_{i-1}, A_i, \ldots, A_n$ — соответствующие точки графика функции $f(x)$. Построим ломаную $A_0, A_1, A_2, \ldots, A_n$ (рис. 104). При вращении этой ломаной вокруг оси Ox мы получим поверхность вращения, составленную из боковых поверхностей усе-

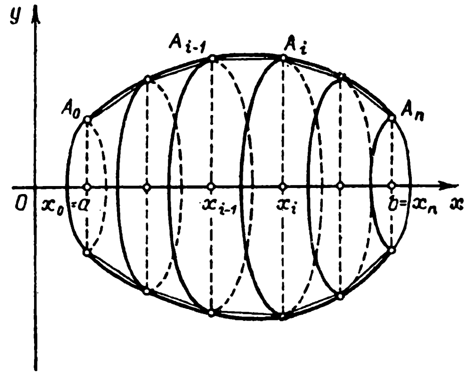


Рис. 104

ченных конусов (или цилиндров). Площадь боковой поверхности усеченного конуса (или цилиндра), образованного вращением i-того звена, равна $2\pi\frac{f(x_{i-1})+f(x_i)}{2}l_i$, где l_i — длина хорды $A_{i-1}A_i$ — равна

$$\sqrt{(x_i-x_{i-1})^2+[f(x_i)-f(x_{i-1})]^2}.$$

По формуле Лагранжа

$$f(x_i)-f(x_{i-1})=f'(\xi_i)(x_i-x_{i-1}),\quad x_{i-1}\leqslant\xi_i\leqslant x_i.$$

Полагая $x_i-x_{i-1}=\Delta x_i$, получим

$$l_i=\sqrt{1+f'^2(\xi_i)}\,\Delta x_i.$$

Значит, площадь поверхности вращения будет приближенно равна площади поверхности вращения ломаной

$$P\approx\sum_{i=1}^{n}2\pi\frac{f(x_{i-1})+f(x_i)}{2}\sqrt{1+f'^2(\xi_i)}\,\Delta x_i.$$

Разобьем эту сумму на две суммы:

$$P \approx 2\pi \sum_{i=1}^{n} f(\xi_i)\sqrt{1+f'^2(\xi_i)}\,\Delta x_i +$$

$$+\pi\left\{\sum_{i=1}^{n} [(f(x_{i-1})-f(\xi_i))+(f(x_i)-f(\xi_i))]\sqrt{1+f'^2(\xi_i)}\,\Delta x_i\right\}. \tag{9}$$

Первая сумма в правой части этого равенства является интегральной суммой σ для интеграла (8) и при $\lambda=\max\limits_{1\leqslant i\leqslant n}\{\Delta x_i\}\to 0$ в силу непрерывности функции $f(x)\sqrt{1+f'^2(x)}$ имеет своим пределом этот интеграл. Покажем, что выражение в фигурных скобках в правой части равенства (9) имеет при $\lambda\to 0$ предел, равный нулю. Действительно, так как функция $f(x)$ равномерно непрерывна на $[a, b]$, то по теореме Кантора для любого $\varepsilon>0$ существует $\delta>0$ такое, что при $\lambda<\delta$ выполняются неравенства $|f(x_{i-1})-f(\xi_i)|<\varepsilon$ и $|f(x_i)-f(\xi_i)|<\varepsilon$. Если через M обозначим максимальное значение функции $\sqrt{1+f'^2(x)}$ **на отрезке** $[a, b]$, то выражение в фигурных скобках оценивается следующим образом:

$$\left|\left\{\sum_{i=1}^{n} [(f(x_{i-1})-f(\xi_i))+(f(x_i)-f(\xi_i))]\sqrt{1+f'^2(\xi_i)}\,\Delta x_i\right\}\right|<$$

$$< 2M\varepsilon\sum_{i=1}^{n}\Delta x_i = 2M(b-a)\varepsilon.$$

Так как $2M(b-a)$ — постоянное число, а ε произвольно мало, то предел указанного выражения равен нулю при $\lambda\to 0$.

Таким образом, переходя в равенстве (9) к пределу при $\lambda\to 0$, получаем:

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

т. е. получили искомую формулу (8) ■

Замечание. Если поверхность получается путем вращения вокруг оси Ox кривой AB, заданной параметрическими уравнениями $x=\varphi(t)$, $y=\psi(t)$, $\alpha\leqslant t\leqslant\beta$, причем $\psi(t)\geqslant 0$, $\varphi(t)$ изменяется от a до b при изменении t от α до β, то, осуществляя в интеграле (8) замену переменной по формулам $x=\varphi(t)$, $y=\psi(t)$, получим

$$P=2\pi\int_{\alpha}^{\beta}\psi(t)\sqrt{\varphi'^2(t)+\psi'^2(t)}\,dt. \tag{10}$$

Наконец, если кривая задана уравнением в полярных координатах $\rho=\rho(\varphi)$, $\alpha\leqslant\varphi\leqslant\beta$, где $\rho(\varphi)$ имеет непрерывную производную на $[\alpha, \beta]$, то этот случай, как мы уже отмечали в п. 3, с помощью формул перехода $x=\rho(\varphi)\cos\varphi$, $y=\rho(\varphi)\sin\varphi$ приводится к параметрической форме задания кривой, и формула (10) принимает вид

$$P=2\pi\int_{\alpha}^{\beta}\rho\sin\varphi\sqrt{\rho^2+\rho'^2}\,d\varphi.$$

Пример 9. Вычислить площадь P поверхности шара радиуса R.

Решение. Здесь имеем

$$f(x)=\sqrt{R^2-x^2}\ (-R\leqslant x\leqslant R),\ f'(x)=-\frac{x}{\sqrt{R^2-x^2}}.$$

По формуле (8) получаем

$$P=2\pi\int_{-R}^{R}\sqrt{R^2-x^2}\sqrt{1+\frac{x^2}{R^2-x^2}}\,dx=$$

$$=2\pi\int_{-R}^{R}R\,dx=4\pi R^2.$$

Пример 10. Вычислить площадь P поверхности, полученной вращением циклоиды $x=a(t-\sin t)$, $y=a(1-\cos t)$, $0\leqslant t\leqslant 2\pi$ вокруг оси Ox (см. рис. 101).

Решение. По формуле (10) имеем

$$P=2\pi\int_{0}^{2\pi}a(1-\cos t)\sqrt{1+\left[\frac{a\sin t}{a(1-\cos t)}\right]^2}\,a(1-\cos t)dt=$$

$$= 2\sqrt{2}\,\pi a^2 \int\limits_0^{2\pi} (1-\cos t)^{3/2}\,dt = \frac{64}{3}\pi a^2.$$

6. Работа переменной силы. Из рассмотренных выше задач на геометрические приложения определенного интеграла видно, что для решения их применяется один и тот же вычислительный метод: составляется интегральная сумма, а затем предельным переходом получается нужный интеграл. Этим методом решается целый ряд задач механики, физики и техники. В качестве примера рассмотрим задачу определения работы переменной силы.

Пусть материальная точка перемещается под действием силы F, направленной вдоль оси Ox и имеющей переменную величину, зависящую от x. Требуется определить работу A, совершаемую силой F по перемещению материальной точки вдоль оси Ox из точки $x=a$ в точку $x=b\,(a<b)$. Функция $F(x)$ предполагается непрерывной на отрезке $[a, b]$ (рис. 105).

Разобьем произвольно отрезок $[a, b]$ на n частей точками

$$a=x_0<x_1<x_2<\ldots<x_{i-1}<x_i<\ldots<x_n=b.$$

Выберем на каждом частичном отрезке $[x_{i-1}, x_i]$ точку ξ_i. Сила, действующая на $[x_{i-1}, x_i]$, меняется от точки к точке. Но если длина отрезка мала, то ее величина в точках отрезка $[x_{i-1}, x_i]$ мало отличается от ее значения в любой точке $\xi_i\in[x_{i-1}, x_i]$, так как $F(x)$ непрерывна. Поэтому работу A_i, совершенную силой F на $[x_{i-1}, x_i]$, можно считать приближенно равной работе, совершаемой на том же участке постоянной силой величины $F(\xi_i)$, т. е.

$$A_i\approx F(\xi_i)\Delta x_i.$$

Проведя аналогичные рассуждения для каждого отрезка разбиения, получаем приближенное значение работы A силы F на всем отрезке $[a, b]$:

$$A\approx\sum_{i=1}^{n} F(\xi_i)\Delta x_i,$$

т. е. получили интегральную сумму σ для функции $F(x)$. Так как функция $F(x)$ непрерывна на $[a, b]$, то предел этой суммы при $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\} \to 0$ существует и численно равен соответствующему определенному интегралу

$$A = \lim_{\lambda \to 0} \sum_{i=1}^{n} F(\xi_i)\,\Delta x_i = \int_a^b F(x)\,dx.$$

Таким образом,

$$A = \int_a^b F(x)\,dx. \tag{11}$$

Пример 11. Определить работу A, необходимую для запуска ракеты весом P с поверхности Земли вертикально вверх на высоту h (рис. 106).

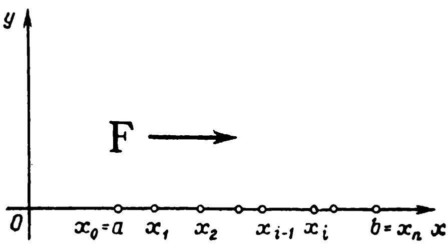


Рис. 105

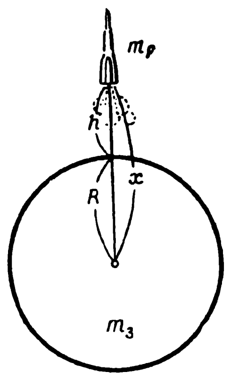


Рис. 106

Решение. Обозначим через F величину силы притяжения ракеты Землей. Пусть $m_{\text{р}}$ — масса ракеты, $m_{\text{з}}$ — масса Земли. Согласно закону Ньютона

$$F = k\frac{m_{\text{р}} m_{\text{з}}}{x^2},$$

где x — расстояние от ракеты до центра Земли. Полагая $km_{\text{р}}m_{\text{з}} = \gamma$, получим $F(x) = \frac{\gamma}{x^2}$, $R \leqslant x \leqslant h + R$, R — радиус Земли. При $x = R$ сила $F(R) = \frac{\gamma}{R^2} = P$, т. е равна весу ракеты P, откуда $\gamma = PR^2$ и $F(x) = \frac{PR^2}{x^2}$.

Таким образом, по формуле (11) получаем

$$A=\int_{R}^{R+h} F(x)\,dx = PR^2\int_{R}^{R+h}\frac{dx}{x^2} = -PR^2\frac{1}{x}\Bigg|_{R}^{R+h} = \frac{PRh}{R+h}.$$

§ 11. Несобственные интегралы

Вводя определенный интеграл как предел интегральных сумм, мы предполагали, что отрезок интегрирования конечный, а подынтегральная функция ограничена на этом отрезке. Если хотя бы одно из этих условий не выполнено, то данное выше определение определенного интеграла теряет смысл. Так, в случае бесконечного отрезка интегрирования нельзя разбить отрезок на n частей конечной длины, а в случае неограниченной функции интегральная сумма не имеет конечного предела. Однако и в этих случаях удается обобщить понятие определенного интеграла. В результате такого обобщения и появилось понятие несобственного интеграла.

1. Несобственные интегралы с бесконечными пределами интегрирования

Определение. Пусть функция $f(x)$ определена в промежутке $[a, +\infty)$ и интегрируема в любой его части $[a, R]$, т. е. существует определенный интеграл $\int_a^R f(x)\,dx$ при любом $R>a$. Тогда, если существует *конечный* предел

$$\lim_{R\to+\infty}\int_a^R f(x)\,dx, \tag{1}$$

то его называют *несобственным интегралом первого рода* и обозначают символом

$$\int_a^{+\infty} f(x)\,dx = \lim_{R\to+\infty}\int_a^R f(x)\,dx. \tag{2}$$

В этом случае говорят, что интеграл (2) существует или сходится. Если же предел (1) не существует или бесконечен, то говорят, что несобственный интеграл не существует или расходится.

Аналогично интегралу (2) вводится несобственный интеграл вида

$$\int_{-\infty}^{b} f(x)\,dx = \lim_{R\to-\infty} \int_{a}^{b} f(x)\,dx. \qquad (3)$$

Наконец, как сумму подобных интегралов можно определить несобственный интеграл с обоими бесконечными пределами, т. е. определить его равенством

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{+\infty} f(x)\,dx, \qquad (4)$$

где c — любое число; при условии существования обоих интегралов справа.

Легко установить геометрический смысл несобственного интеграла первого рода. Пусть $f(x) \geqslant 0$. Тогда, если определенный интеграл $\int_a^R f(x)\,dx$ выражает площадь области, ограниченной сверху графиком функции $f(x)$, снизу осью Ox, слева прямой $x=a$, справа прямой $x=R$, то естественно считать, что несобственный интеграл $\int_a^{+\infty} f(x)\,dx$ выражает конечную площадь бесконечной области, ограниченной снизу осью Ox, сверху графиком функции $f(x)$, слева прямой $x=a$ (рис. 107). Аналогичные рассуждения имеют место для интегралов (3) и (4).

Рассмотрим несколько примеров несобственных интегралов первого рода.

Пример 1.

$$\int_0^{+\infty} \frac{dx}{1+x^2} = \lim_{R\to+\infty} \int_0^R \frac{dx}{1+x^2} =$$

$$= \lim_{R\to+\infty} \operatorname{arctg} x \Big|_0^R = \lim_{R\to+\infty} \operatorname{arctg} R = \frac{\pi}{2},$$

т. е. данный интеграл сходится.

Пример 2.

$$\int_0^{+\infty} \cos x dx = \lim_{R\to+\infty} \int_0^R \cos x dx = \lim_{R\to+\infty} \sin x \Big|_0^R = \lim_{R\to+\infty} \sin R,$$

а предел функции $\sin R$ при $R \to +\infty$ не существует, следовательно, интеграл расходится.

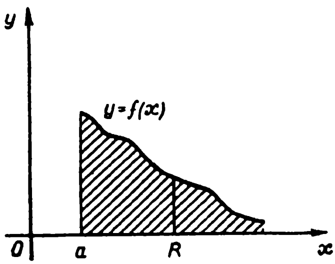


Рис. 107

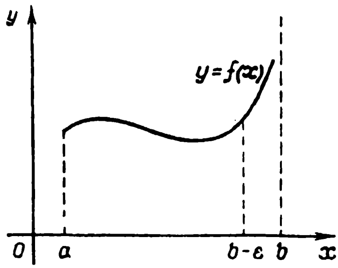


Рис. 108

Пример 3. $\int_{-\infty}^{+\infty} e^x dx = \int_{-\infty}^{0} e^x dx + \int_0^{+\infty} e^x dx$ расходится, так как

$$\int_0^{+\infty} e^x dx = \lim_{R\to+\infty} \int_0^R e^x dx = \lim_{R\to+\infty} e^x \Big|_0^R = \lim_{R\to+\infty} (e^R - 1) = \infty.$$

Пример 4. $\int_1^{+\infty} \frac{dx}{x^\alpha}$, α — некоторое число; 1) если $\alpha \neq 1$, то для любого $R > 0$

$$\lim_{R\to+\infty} \int_1^R \frac{dx}{x^\alpha} = \lim_{R\to+\infty} \frac{x^{1-\alpha}}{1-\alpha} \Big|_1^R =$$

$$= \lim_{R\to+\infty} \frac{R^{1-\alpha} - 1}{1-\alpha} = \begin{cases} \dfrac{-1}{1-\alpha}, & \alpha > 1, \\ \infty &, \alpha < 1; \end{cases}$$

2) если $\alpha = 1$, то для любого $R > 0$

$$\lim_{R\to+\infty} \int_1^R \frac{dx}{x} = \lim_{R\to+\infty} \ln x \Big|_1^R = \lim_{R\to+\infty} \ln R = \infty.$$

Таким образом, данный интеграл сходится при $\alpha > 1$ и расходится при $\alpha \leqslant 1$.

Заметим, что в рассмотренных примерах само вычисление несобственного интеграла основано на его определении.

2. Несобственные интегралы от неограниченных функций

Определение. Пусть функция $f(x)$ определена в промежутке $[a, b)$. Точку $x=b$ будем называть *особой*, если функция $f(x)$ не ограничена в окрестности этой точки, но ограничена на любом $[a, b-\varepsilon]$, заключенном в $[a, b)$ (рис. 108). Предполагается, что на любом $[a, b-\varepsilon]$, функция $f(x)$ интегрируема. Тогда, как бы ни было мало $\varepsilon>0$, если существует конечный предел

$$\lim_{\varepsilon\to 0}\int_a^{b-\varepsilon} f(x)\,dx, \qquad (5)$$

то его называют несобственным интегралом второго рода и обозначают

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0}\int_a^{b-\varepsilon} f(x)\,dx. \qquad (6)$$

В этом случае говорят, что интеграл (6) существует или сходится. Если же предел (5) не существует или бесконечен, то говорят, что интеграл (6) не существует или расходится.

Аналогично, если точка $x=a$ — особая точка, то несобственный интеграл в этом случае определяется так:

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0}\int_{a+\varepsilon}^{b} f(x)\,dx.$$

Если $f(x)$ не ограничена в окрестности какой-нибудь внутренней точки $c\in[a, b]$, то по определению полагают

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

при условии существования обоих интегралов справа.

Наконец, если a и b — особые точки, то в этом случае несобственный интеграл определяется как сумма

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

где c — любая точка из (a, b), если оба интеграла справа существуют.

Пример 5.

$$\int_0^1 \frac{dx}{x^\alpha} = \lim_{\varepsilon\to 0+} \int_\varepsilon^1 \frac{dx}{x^\alpha} = \lim_{\varepsilon\to 0+} \frac{x^{1-\alpha}}{1-\alpha}\Big|_\varepsilon^1 =$$

$$= \lim_{\varepsilon\to 0+} \frac{1-\varepsilon^{1-\alpha}}{1-\alpha} = \begin{cases} \infty, & \text{при } \alpha > 1, \\ \dfrac{1}{1-\alpha}, & \text{при } \alpha < 1, \end{cases}$$

а при $\alpha = 1$

$$\int_0^1 \frac{dx}{x} = \lim_{\varepsilon\to 0+} \int_\varepsilon^1 \frac{dx}{x} = \lim_{\varepsilon\to 0+} \ln x\Big|_\varepsilon^1 = \lim_{\varepsilon\to 0+} (-\ln\varepsilon) = \infty,$$

таким образом, данный интеграл сходится при $\alpha<1$ и расходится при $\alpha\geqslant 1$.

3. Признак сходимости несобственных интегралов. Рассмотрим вопрос сходимости интегралов с бесконечными пределами интегрирования вида $\int_a^{+\infty} f(x)dx$. Аналогичные рассуждения можно провести и для интегралов других видов.

Следующая теорема познакомит нас с одним из признаков сходимости несобственных интегралов.

Теорема 8.9 (признак сравнения несобственных интегралов). *Если функции $f(x)$ и $g(x)$ непрерывны на полуинтервале $[a, +\infty)$ и удовлетворяют на нем условию $0\leqslant f(x)\leqslant g(x)$, то из сходимости интеграла*

$$\int_a^{+\infty} g(x)\,dx \tag{7}$$

следует сходимость интеграла

$$\int_a^{+\infty} f(x)\,dx, \tag{8}$$

а из расходимости интеграла (8) следует расходимость интеграла (7).

Доказательство. Так как интеграл (7) по условию сходится, то согласно теореме 2.6 это означает, что существует число $M>0$ такое, что для любого $R>a$ выполняется неравенство $\Phi(R)=\int\limits_a^R g(x)\,dx\leqslant M$. Тогда, в силу данного условия и оценки 2° (см. § 5), $\int\limits_a^R f(x)\,dx\leqslant$ $\leqslant\int\limits_a^R g(x)dx\leqslant M$. А это означает, что функция $F(R)=$ $=\int\limits_a^R f(x)\,dx$ монотонно возрастает и ограничена, т. е. имеет конечный предел при $R\to+\infty$, и, следовательно, интеграл $\int\limits_a^{+\infty} f(x)\,dx$ сходится.

Если же (8) расходится, то, допустив сходимость интеграла (7), получим по только что доказанному сходимость интеграла (8), что противоречит условию. Окончательно получаем, что интеграл (7) также расходится ■

Замечание. Что касается сходимости несобственных интегралов второго рода, то теория этих интегралов аналогична теории несобственных интегралов первого рода. В частности, признак сравнения для таких интегралов можно сформулировать следующим образом: если функции $f(x)$ и $g(x)$ непрерывны на полуинтервале $(a, b]$ и для всех точек x в некоторой окрестности точки a выполняются условия $0\leqslant f(x)\leqslant g(x)$, то из сходимости интеграла $\int\limits_a^b g(x)dx$ следует сходимость интеграла $\int\limits_a^b f(x)dx$, а из расходимости $\int\limits_a^b f(x)dx$ следует расходимость $\int\limits_a^b g(x)dx$.

Пример 6. Исследовать сходимость $\int\limits_0^1 \frac{dx}{x\sqrt[3]{x}}$.

Решение. В точке $x=0$ подынтегральная функция обращается в бесконечность, т. е. данный интеграл принадлежит к несобственным интегралам второго рода, а точка $x=0$ является особой.

По признаку сравнения $0 < \frac{1}{x^{5/3}} < \frac{1}{x^{4/3}}$, т. е. $\int\limits_0^1 \frac{dx}{x^{5/3}} < \int\limits_0^1 \frac{dx}{x^{4/3}}$ (см. пример 5) $\alpha = \frac{5}{3} > 1$. Следовательно, данный интеграл расходится.

4. Пример использования несобственного интеграла. В качестве примера использования несобственного интеграла проведем подсчет второй космической скорости.

Вычислим начальную скорость ракеты, при которой она способна выйти из поля притяжения Земли и уйти в межпланетное пространство.

Ранее (см. § 10, п. 6, пример 11) с помощью определенного интеграла была подсчитана работа силы, необходимая для запуска ракеты весом P с поверхности Земли на высоту h, т. е. было получено

$$A = \int\limits_R^{R+h} F(x)\,dx = \frac{PRh}{R+h} = \frac{Ph}{1+\frac{h}{R}}$$

Откуда следует, что при $R \gg h$ $A \approx mgh$, где m — масса ракеты, g — ускорение свободного падения у поверхности Земли.

В силу закона сохранения сумма кинетической и потенциальной энергии ракеты во все моменты времени одна и та же. Следовательно, чтобы полностью освободить ракету от земного притяжения, необходимо, чтобы начальная кинетическая энергия была не меньше потенциальной энергии ракеты, бесконечно удаленной от поверхности Земли, т. е. надо подсчитать работу силы при $h \to \infty$:

$$\lim_{h\to\infty} A(h) = \lim_{h\to\infty} \int_R^{R+h} F(x)\,dx = \lim_{h\to\infty} \frac{PRh}{R+h} =$$

$$= \lim_{h\to\infty} \frac{PR}{1+\frac{R}{h}} = PR = mgR$$

(движение Земли и притяжение других планет при этом не учитывается). Отсюда начальная скорость ракеты v должна быть такова, чтобы

$$\frac{mv^2}{2} \geqslant mgR \text{ или } v \geqslant \sqrt{2gR} =$$

$$= \sqrt{2\cdot 10\cdot 6400000}\,\text{м/с} = 1{,}4\cdot 8000\,\text{м/с} = 11{,}2\,\text{км/с}.$$

Если скорость ракеты равна 11,2 км/с, то ее траектория движения представляет собой параболу. При начальной скорости большей 11,2 км/с траектория станет гиперболой.

§ 12. Приближенное вычисление определенных интегралов

При решении физических и технических задач встречаются определенные интегралы от функций, первообразные которых не выражаются через элементарные функции. Это привело к необходимости разработки так называемых *приближенных* формул вычисления определенных интегралов. Таких формул существует достаточно много, но мы познакомимся лишь с двумя: *формулой трапеций* и *формулой парабол.*

Пусть требуется вычислить интеграл

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

где $f(x)$ — непрерывная функция и $f(x) \geqslant 0$.

Разобьем отрезок $[a, b]$ на n равных частей точками $a = x_0 < x_1 < x_2 < \ldots < x_{k-1} < x_k < \ldots < x_n = b$ (рис. 109) и с помощью прямых $x = x_k$ построим n прямолинейных трапеций (эти трапеции заштрихованы на рис. 109), сумму площадей которых будем считать приближенно равной площади данной криволинейной трапеции, т. е. имеем

$$\int_a^b f(x)\,dx \approx \frac{f(x_0)+f(x_1)}{2}(x_1-x_0)+$$

$$+\frac{f(x_1)+f(x_2)}{2}(x_2-x_1)+\ldots+\frac{f(x_{n-1})+f(x_n)}{2}(x_n-x_{n-1})=$$

$$=\frac{b-a}{2n}\left\{f(a)+f(b)+2\sum_{k=1}^{n-1}f(x_k)\right\},$$

где соответственно $f(x_{k-1})$ и $f(x_k)$ — основания трапеций, а $x_k-x_{k-1}=\dfrac{b-a}{n}$ — их высоты.

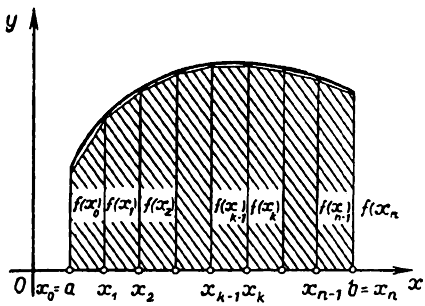


Рис. 109

Таким образом, получили приближенную формулу

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2n}\left\{f(a)+f(b)+2\sum_{k=1}^{n-1}f(x_k)\right\},$$

которая и называется *формулой трапеций*. Она будет тем более точной, чем больше будет число n.

Вычислим интеграл $\int_0^1 x^2dx$.

Точное значение этого интеграла находится легко:

$$\int_0^1 x^2dx=\frac{x^3}{3}\Big|_0^1=\frac{1}{3}=0{,}3333\ldots$$

Вычислим по формуле трапеций приближенное значение. Возьмем $n=5$. Тогда будем иметь

$$a=x_0=0,\ x_1=0{,}2,\ x_2=0{,}4,\ x_3=0{,}6,\ x_4=0{,}8,\ x_5=1=b$$

и соответственно

$$f(x_0)=0,\ f(x_1)=0{,}04,\ f(x_2)=0{,}16,\ f(x_3)=0{,}36,\ f(x_4)=$$
$$=0{,}64,\ f(x_5)=1.$$

Значит,

$$\int_0^1 x^2dx \approx \frac{1}{10}\{1+2[0{,}04+0{,}16+0{,}36+0{,}64]\}=\frac{3{,}4}{10}=0{,}34.$$

Поскольку точное значение интеграла есть 0,3333..., то абсолютная ошибка меньше 0,007. Во многих технических задачах эта точность достаточна.

Если увеличим число n, то, естественно, точность будет большей. Так, например, при $n=10$

$$\int_0^1 x^2dx \approx \frac{1}{20}\{1+2\cdot 2{,}85\}=\frac{6{,}7}{20}=0{,}335,$$

т. е. абсолютная ошибка будет меньше 0,002.

В более полных курсах доказывается, что если $f(x)$ имеет на $[a,\ b]$ непрерывную вторую производную, то абсолютная ошибка формулы трапеций не больше, чем

$$k\,\frac{(b-a)^3}{12n^2},$$

где k — наибольшее значение $|f''(x)|$ на отрезке $[a,\ b]$.

К сожалению, с ростом числа n увеличивается не только точность вычисления определенного интеграла, но и объем вычислительной работы. Однако здесь на помощь приходят электронно-вычислительные машины.

Вычислим приближенно интеграл

$$\int_0^1 \frac{dx}{1+x}$$

по формуле трапеций при $n=10$. Разобьем отрезок $[0,\ 1]$ на 10 равных частей точками $x_0=0$, $x_1=0{,}1$, ..., $x_9=0{,}9$, $x_{10}=1$. Вычислим приближенно значения функции $f(x)=1/(1+x)$ в этих точках:

$$f(0)=1{,}0000, \qquad f(0{,}5)=0{,}6667,$$
$$f(0{,}1)=0{,}9091, \qquad f(0{,}6)=0{,}6250,$$
$$f(0{,}2)=0{,}8333, \qquad f(0{,}7)=0{,}5882,$$

$$f(0,3)=0,7692, \qquad f(0,8)=0,5556,$$
$$f(0,4)=0,7143, \qquad f(0,9)=0,5263,$$
$$f(0,5)=0,6667 \qquad f(1)=0,5000.$$

По формуле трапеций получим

$$\int_0^1 \frac{dx}{1+x} \approx \frac{1}{10}\left(\frac{1,000+0,5000}{2}+0,9091+0,8333+\right.$$
$$+0,7692+0,7143+0,6667+0,6250+0,5882+0,5556+$$
$$\left.+0,5263\right)=0,69377 \approx 0,6938.$$

Произведем оценку погрешности полученного результата. Так как
$f(x)=1/(1+x)$, то $f'(x)=-1/(1+x)^2$, $f''(x)=2/(1+x)^3$. На отрезке $[0, 1]$ имеем $|f''(x)| \leqslant 2$. Следовательно, в качестве k можно взять число 2. Отсюда находим оценку погрешности:

$$2/(12 \cdot 10^2)=1/600<0,0017.$$

Для получения точного значения данного интеграла вычислим его по формуле Ньютона — Лейбница, получим

$$\int_0^1 \frac{dx}{1+x}=\ln(1+x)\Big|_0^1=\ln 2 \approx 0,69315.$$

Видим, что абсолютная ошибка результата, полученного по формуле трапеций, меньше 0,0007. Это говорит о том, что формула, выражающая величину погрешности формулы трапеций, дает довольно грубую оценку погрешности и имеет больше теоретическое, чем практическое, значение.

Идею, которая была использована при построении формулы трапеций, можно развивать дальше и использовать для получения более точных приближенных формул для вычисления определенного интеграла. Например, если будем заменять график подынтегральной функции на отдельных отрезках $[x_{k-1}, x_k]$ не хордами, как при выводе формулы трапеций, а дугами парабол с осями, параллельными оси Oy, то придем к формуле

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6n}\left[y_0+y_{2n}+2(y_2+y_4+\ldots+y_{2n-2})+\right.$$
$$\left.+4(y_1+y_3+\ldots+y_{2n-1})\right].$$

Эта формула называется *формулой парабол* или *формулой Симпсона**.

* Симпсон Томас (1710—1761) — английский математик.

В этой формуле, как и в формуле трапеций, предполагается, что отрезок $[a,\ b]$ разбит точками $x_0=a$, x_1, ..., $x_{2n}=b$, $x_0<x_1<\ ...\ <x_{2n}$ на $2n$ равных частей. Значение функции $f(x)$ в нечетных точках разбиения x_1, x_3, ..., x_{2n-1} входит с коэффициентом 4, в четных x_2, x_4, ..., x_{2n} — с коэффициентом 2 и в двух граничных $x_0=a$, $x_{2n}=b$ — с коэффициентом 1.

Вывод формулы Симпсона не содержит новых идей по сравнению с выводом формулы трапеций, но является более громоздким. Мы на нем останавливаться не будем.

Геометрический смысл формулы Симпсона очевиден: площадь криволинейной трапеции, лежащей под графиком функции $f(x)$ на отрезке $[a,\ b]$, приближенно заменяется суммой площадей фигур, лежащих под параболами.

В полных курсах доказывается, что если функция $f(x)$ имеет непрерывные производные до четвертого порядка включительно, то абсолютная величина погрешности формулы Симпсона не больше, чем

$$M(b-a)^5/2880n^4,$$

где M — наибольшие значения $|f_{(x)}^{(4)}|$ на отрезке $[a,\ b]$. Выше отмечалось, что погрешность трапеций оценивается числом

$$k(b-a)^3/12n^2 \qquad (k=\max_{[a,b]}|f''(x)|).$$

Так как n^4 растет быстрее, чем n^2, то погрешность формулы Симпсона убывает значительно быстрее, чем в случае формулы трапеций. Этим и объясняется, что формула Симпсона дает большую точность, чем формула трапеций.

Для сравнения точности приближенных формул вычислим снова интеграл

$$\int_0^1 \frac{dx}{1+x}$$

по формуле Симпсона при $n=4$. Разобьем отрезок $[0,\ 1]$ на четыре равные части точками $x_0=0$, $x_1=1/4$, $x_2=1/2$, $x_3=3/4$, $x_4=1$ и вычислим приближенно значения функции $f(x)=1/(1+x)$ в этих точках: $y_0=1{,}000$, $y_1=0{,}8000$,

$$y_2=0{,}6667,\ y_3=0{,}5714,\ y_4=0{,}5000.$$

По формуле Симпсона получим:

$$\int_0^1 \frac{dx}{1+x} \approx \frac{b-a}{6n} [y_0+y_4+2y_2+4(y_1+y_3)] =$$

$$=1-0/12[1{,}0000+0{,}5000+2\cdot 0{,}6667+4(0{,}8000+$$
$$+0{,}5314)] \approx 0{,}69325.$$

Произведем оценку погрешности полученного результата. Для подынтегральной функции $f(x)=1/(1+x)$ имеем: $f(x)^{(4)}=24(1+x)^5$, откуда следует, что на отрезке $[0, 1]$ $|f(x)^{(4)}| \leqslant 24$. Следовательно, в качестве M можно взять число 24. Отсюда находим оценку погрешности:

$$24/(2880\cdot 4^4) < 0{,}0004,$$

из которой видно, что оценка погрешности формулы Симпсона намного точнее оценки погрешности формулы трапеций.

Кроме того, абсолютная ошибка результата, полученного по формуле Симпсона, по сравнению с точным меньше 0,0001. Это говорит о том, что формула Симпсона значительно точнее формулы трапеций даже при меньшем количестве точек разбиения. Поэтому этой формулой для приближенного вычисления определенного интеграла пользуются чаще, чем формулой трапеций.

Как выше отмечалось, приближенные формулы для вычисления определенного интеграла применяются в тех случаях, когда первообразная аналитически заданной функции не выражается в элементарных функциях.

В качестве иллюстрации к сказанному вычислим интеграл

$$\int_0^1 e^{-x^2}\, dx\,*$$

по формуле Симпсона с точностью до 0,001.

Чтобы подобрать необходимое для обеспечения заданной точности число $2n$, найдем $f(x)^{(4)}$. Последовательно дифференцируя функцию $f(x)=e^{-x^2}$, получим

$$f(x)^{(4)}=4e^{-x^2}(4x^4-12x^2+3).$$

Так как на отрезке $[0, 1]$ $e^{-x^2} \leqslant 1$, $|4x^4-12x^2+3| \leqslant 5$, то $|f(x)^{(4)}| \leqslant 20$. Следовательно, в качестве M можно взять число 20. Используя формулу погрешности, получаем

$$20/2880n^4 < 1/1000,$$

* Рассматриваемый интеграл не выражается через элементарные функции, но имеет большое значение в статистической физике, теории теплопроводности и диффузии.

откуда

$$n^4 < 1000/144.$$

Для того чтобы выполнялось это неравенство, достаточно взять $n=2$, т. е. $2n=4$.

Разобьем теперь отрезок [0, 1] на четыре равные части точками $x_0=0$, $x_1=1/4$, $x_2=1/2$, $x_3=3/4$, $x_4=1$ и вычислим приближенно значения функции $f(x)e^{-x^2}$ в этих точках: $y_0=1{,}0000$, $y_1=0{,}9394$, $y_2=0{,}7788$, $y_3=0{,}5698$, $y_4=0{,}3679$. Применяя формулу Симпсона, получим

$$\int_0^1 e^{-x^2}\,dx \approx \frac{1}{12}[1{,}0000+0{,}3679+2\cdot 0{,}7788+$$
$$+4\,(0{,}9394+0{,}5698)\,] \approx 0{,}7469.$$

Таким образом,

$$\int_0^1 e^{x^2}\,dx \approx 0{,}747$$

с точностью до 0,001. Таким образом, разбив отрезок [0, 1] всего на четыре части и заменив рассматриваемый интеграл суммой, стоящей в правой части формулы Симпсона, мы вычислим этот интеграл с нужной нам точностью.

На этом заканчивается первая часть курса — анализ функций одной переменной.

Часть II

АНАЛИЗ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Естественным продолжением данного курса является изучение функций не одной, а двух и более переменных. Однако прежде чем перейти к рассмотрению таких функций, нам необходимо познакомиться с некоторыми понятиями аналитической геометрии в пространстве, так как, например, областью определения функции двух переменных является уже не прямая (или ее часть), а плоскость (или ее часть). Графиком функции двух переменных является некоторая поверхность в пространстве.

В этой части мы познакомимся также с такими понятиями высшей алгебры, как матрица и определитель, важными не только для изучения математического анализа, но имеющими широкое применение в других разделах математики.

Глава IX

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ В ПРОСТРАНСТВЕ

§ 1. Точка и координаты

Прямоугольная система координат в пространстве определяется заданием линейной единицы для измерения длин и трех пересекающихся в одной точке взаимно перпендикулярных осей: Ox, Oy и Oz.

Точка *O — начало координат,* *Ox — ось абсцисс,* *Oy — ось ординат,* *Oz — ось аппликат.*

Пусть M — произвольная точка пространства (рис. 1). Проведем через точку M три плоскости, перпендикулярные координатным осям Ox, Oy и Oz. Точки пересечения указанных плоскостей с осями соответственно обозначим через M_x, M_y и M_z. Прямоугольными координатами точки M называются числа

$$x=OM_x,\ y=OM_y,\ z=OM_z,$$

где x, y, z — величины отрезков OM_x, OM_y, OM_z, при этом x называется *абсциссой*, y — *ординатой*, а z — *аппликатой* точки M.

Таким образом, при выбранной системе координат каждой точке M пространства соответствует единственная упорядоченная тройка чисел $(x;\ y;\ z)$ — ее прямоугольные координаты и, обратно, каждой упорядоченной тройке чисел $(x;\ y;\ z)$ соответствует, и притом одна, точка M в пространстве.

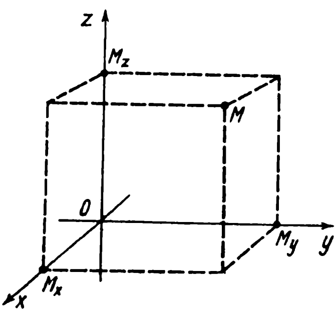


Рис. 1

Итак, прямоугольная система координат в пространстве устанавливает взаимно-однозначное соответствие между множеством всех точек пространства и множеством упорядоченных троек чисел.

Плоскости Oxy, Oyz, Oxz называются *координатными плоскостями*. Они делят все пространство на восемь частей, называемых *октантами*.

§ 2. Понятие вектора

1. Скалярные и векторные величины. В физике, механике и других прикладных науках существуют такие величины, которые полностью определяются заданием некоторого числа. Таковы, например, объем, масса, плотность, температура тела и др. Такие величины называются *скалярными*. В связи с этим числа иногда называют скалярами. Но есть и такие величины, которые определяются заданием не только числа, но и некоторого направления. Например, при движении тела мы должны указать не только величину скорости, с которой движется тело, но и направление движения. Точно так же, изучая действие какой-либо силы, необходимо указать не только величину этой силы, но и направление ее действия. Такие величины называются *векторными*. Для их описания было введено понятие вектора, оказавшееся полезным для математики.

2. Определение вектора. Любая упорядоченная пара точек A и B пространства определяет так называемый *направленный отрезок*, т. е. отрезок вместе с данным на нем направлением. Если точка A — первая, то ее называют *началом* направленного отрезка, а точку B — его *концом*. Направлением отрезка считается направление от начала к концу.

Рис. 2

Определение 1. *Направленный отрезок (или, что то же, упорядоченная пара точек) называется вектором.*

Другими словами, вектор — это есть прямолинейный отрезок, имеющий начало и конец.

Вектор обозначается символом $\overline{AB}$, причем первая буква обозначает начало вектора, а вторая — его конец. Вектор также обозначается и одной буквой с черточкой наверху, например $\bar{a}$. Направление вектора на рисунке указывается стрелкой (рис. 2).

Вектор, у которого начало и конец совпадают, называется *нулевым* и обозначается $\bar{0}$ или просто 0.

Расстояние между началом и концом вектора называется его *длиной* и обозначается $|\overline{AB}|$ или $|\bar{a}|$.

Векторы $\bar{a}$ и $\bar{b}$ называются *коллинеарными*, если они лежат на одной прямой или на параллельных прямых.

Нулевой вектор будем считать коллинеарным любому вектору, так как он не имеет определенного направления, длина его, разумеется, равна нулю, т. е. $|\bar{0}|=0$.

Теперь мы можем сформулировать важное понятие равенства двух векторов.

Определение 2. *Векторы $\bar{a}$ и $\bar{b}$ называются равными ($\bar{a}=\bar{b}$), если они коллинеарны, одинаково направлены и равны по длине.*

На рис. 3 изображены слева неравные, а справа равные векторы $\bar{a}$ и $\bar{b}$. Из определения равенства векторов следует, что вектор можно переносить параллельно самому себе, т. е. начало вектора может быть в любой точке пространства, но длина и направление фиксированы. Такие векторы называются *свободными*.

В механике и физике кроме свободных векторов рассматривают *скользящие* и *связанные* векторы. *Скользящими* называют такие векторы, которые не только равны, но и лежат на одной прямой. *Связанными* называются такие векторы, которые не только равны, но и имеют общее начало.

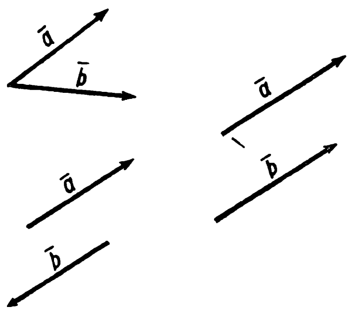


Рис. 3

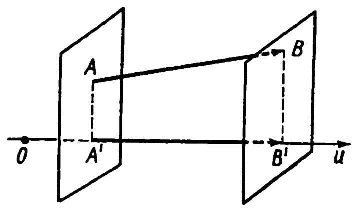


Рис. 4

В дальнейшем изложении будем пользоваться только свободными векторами, называя их просто векторами.

3. Проекция вектора на ось. Пусть в пространстве заданы числовая ось Ou и некоторый вектор $\overline{AB}$. Проведем через точки A и B плоскости, перпендикулярные к оси Ou. Обозначим через A' и B' точки пересечения этих плоскостей с осью (рис. 4).

Проекцией вектора $\overline{AB}$ на ось Ou называется

$$\begin{cases} |\overline{A'B'}|, & \text{если направление } \overline{A'B'} \text{ совпадает с направлением } Ou, \\ -|\overline{A'B'}|, & \text{если направление } \overline{A'B'} \text{ противоположно направлению } Ou. \end{cases} \qquad (1)$$

Проекция вектора $\overline{AB}$ на ось Ou обозначается $\text{Пр.}_u\overline{AB}$.

Имеет место

Теорема 9.1. *Проекция вектора $\overline{AB}$ на ось Ou равна длине вектора $\overline{AB}$, умноженной на косинус угла наклона вектора $\overline{AB}$ к оси Ou, т. е.*

$$\text{Пр.}_u \overline{AB} = |\overline{AB}| \cos \varphi, \qquad (2)$$

где через φ обозначен угол между вектором $\overline{AB}$ и осью (рис. 5).

Доказательство. Если $\varphi \leqslant \frac{\pi}{2}$ (рис. 5, *а*), то в силу (1) $\text{Пр.}_u\overline{AB} = |\overline{A'B'}| = |\overline{AB}| \cos \varphi$. Если $\varphi > \frac{\pi}{2}$ (рис. 5, *б*), то в силу (1) $\text{Пр.}_u\overline{AB} = -|\overline{A'B'}| =$

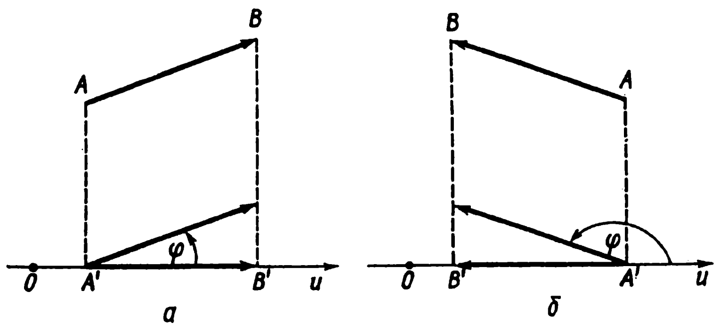


Рис. 5

$= -|\overline{AB}| \cos(\pi - \varphi) = |\overline{AB}| \cos \varphi$, т. е. для любого угла φ справедливо равенство (2) ■

Замечание 1. Пусть $\overline{A_1B_1} = \overline{A_2B_2}$ и дана какая-то ось *Ou*. Применяя к каждому из этих векторов формулу (2), получим

$$\text{Пр.}_u\overline{A_1B_1} = \text{Пр.}_u\overline{A_2B_2},$$

т. е. равные векторы имеют равные проекции на одну и ту же ось.

4. Проекции вектора на оси координат. Пусть в пространстве заданы некоторая прямоугольная система координат и произвольный вектор $\overline{AB}$. Пусть, далее, $X = \text{Пр.}_x\overline{AB}$, $Y = \text{Пр.}_y\overline{AB}$, $Z = \text{Пр.}_z\overline{AB}$. Проекции вектора $\overline{AB}$ на оси координат, будучи заданы, определяют его с точностью до положения в пространстве. Поэтому проекции X, Y, Z вектора $\overline{AB}$ называют его *координатами*. При этом пишут

$$\overline{AB} = \{X;\ Y;\ Z\}.$$

Теорема 9.2. *Каковы бы ни были две точки $A(x_1; y_1; z_1)$ и $B(x_2; y_2; z_2)$, координаты вектора $\overline{AB}$ определяются формулами*

$$X = x_2 - x_1,\quad Y = y_2 - y_1,\quad Z = z_2 - z_1. \tag{3}$$

Доказательство. Проведем через точки A и B плоскости, перпендикулярные к оси Oy, и обозначим

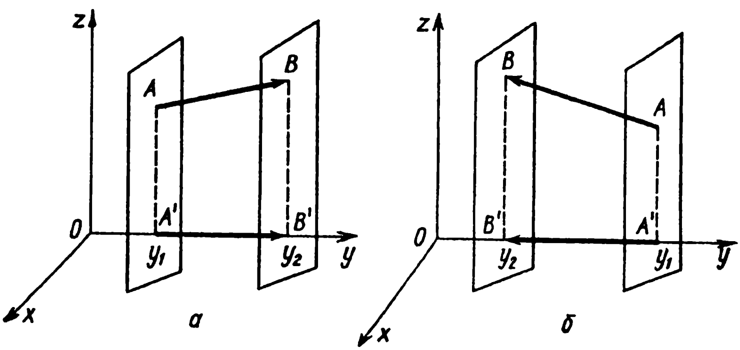


Рис. 6

точки их пересечения с осью Oy соответственно A' и B'. Точки A' и B' на оси Oy будут иметь координаты y_1 и y_2. Тогда, если $y_2 > y_1$ (рис. 6, *а*), то в силу (1)

$$\text{Пр.}_y\overline{AB} = |\overline{A'B'}| = y_2 - y_1 = Y.$$

Если $y_2 < y_1$ (рис. 6, *б*), то вектор $\overline{A'B'}$ имеет противоположное направление с осью Oy, и в силу (1)

$$\text{Пр.}_y\overline{AB} = -|\overline{A'B'}| = -(y_1 - y_2) = y_2 - y_1 = Y.$$

Аналогично устанавливаются и остальные формулы (3) ■

Замечание 2. Если вектор $\overline{AB}$ выходит из начала координат, т. е. $x_1=y_1=z_1=0$ и $x_2=x$, $y_2=y$, $z_2=z$, то координаты X, Y, Z вектора $\overline{AB}$ равны координатам его конца:

$$X=x, \quad Y=y, \; Z=z.$$

5. Направляющие косинусы вектора. Пусть дан произвольный вектор $\bar{a}=\{X;\,Y;\,Z\}$, для простоты будем считать, что $\bar{a}$ выходит из начала координат. Через A обозначим конец вектора и проведем через эту точку плоскости, перпендикулярные к осям. Вместе с координатными плоскостями они образуют прямоугольный параллелепипед, диагональю которого служит отрезок $\overline{OA}$ (рис. 7). Из элементарной геометрии известно, что квадрат длины диагонали прямоугольного параллелепипеда равен сумме квадратов длин трех его смежных сторон. Следовательно,

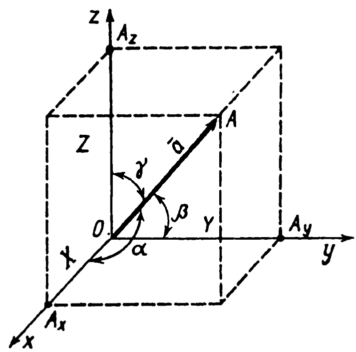


Рис. 7

$$|OA|^2=|OA_x|^2+|OA_y|^2+|OA_z|^2.$$

Но $|OA|=|\bar{a}|$, $|OA_x|=X$, $|OA_y|=Y$, $|OA_z|=Z$; таким образом, из последнего равенства получаем

$$|\bar{a}|^2=X^2+Y^2+Z^2,$$

или

$$|\bar{a}|=\sqrt{X^2+Y^2+Z^2}. \tag{4}$$

Формула (4) выражает длину произвольного вектора через его координаты.

Обозначим через α, β, γ углы наклона вектора $\bar{a}$ с осями координат. Из (2) и (4) получаем

$$\cos\alpha=\frac{X}{\sqrt{X^2+Y^2+Z^2}}, \quad \cos\beta=\frac{Y}{\sqrt{X^2+Y^2+Z^2}},$$

$$\cos\gamma=\frac{Z}{\sqrt{X^2+Y^2+Z^2}}; \tag{5}$$

$\cos\alpha$, $\cos\beta$, $\cos\gamma$ называются *направляющими косинусами* вектора $\bar{a}$.

Возводя в квадрат левую и правую части каждого из равенств (5) и суммируя полученные результаты, найдем:

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1, \qquad (6)$$

т. е. сумма квадратов направляющих косинусов любого вектора равна единице.

В заключение данного пункта рассмотрим задачу. Пусть даны две произвольные точки $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$. Нужно найти расстояние d между ними. Искомый результат получается сразу при помощи теоремы 9.2 и формулы (4). Имеем

$$\overline{M_1M_2}=\{x_2-x_1;\ y_2-y_1;\ z_2-z_1\},$$

а так как d есть длина вектора $\overline{M_1M_2}$, то

$$d=|\overline{M_1M_2}|=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}. \qquad (7)$$

Формула (7) и дает решение поставленной задачи.

§ 3. Линейные операции над векторами и их основные свойства

Линейными операциями над векторами называются операции сложения векторов и умножения векторов на числа.

Определение сложения двух векторов. *Пусть даны два вектора $\bar{a}$ и $\bar{b}$. Суммой $\bar{a}+\bar{b}$ называется вектор, который идет из начала вектора $\bar{a}$ в конец вектора $\bar{b}$ при условии, что вектор $\bar{b}$ приложен к концу вектора $\bar{a}$* (рис. 8, *а*).

Замечание 1. Действие вычитания векторов, как и в арифметике, обратно действию сложения, т. е. разностью двух векторов $\bar{a}$ и $\bar{b}$ называется вектор $\bar{b}-\bar{a}$, который в сумме с вектором $\bar{a}$ составляет вектор $\bar{b}$ (рис. 8, *б*).

Замечание 2. Определив сумму двух векторов, можно вычислить сумму любого числа векторов. Пусть, например, даны три вектора $\bar{a}$, $\bar{b}$ и $\bar{c}$. Сложив $\bar{a}$ и $\bar{b}$, получим вектор $\bar{a}+\bar{b}$. Прибавив теперь к нему вектор $\bar{c}$, получим вектор $\bar{a}+\bar{b}+\bar{c}$.

Определение произведения вектора на число. *Пусть даны вектор $\bar{a}$ и число λ. Произведением $\lambda\bar{a}$ (или $\bar{a}\lambda$) называется вектор, который коллинеарен вектору $\bar{a}$, имеет длину, равную $|\lambda||\bar{a}|$, и направление такое же, как у вектора $\bar{a}$, если $\lambda>0$, и противоположное, если $\lambda<0$* (рис. 9).

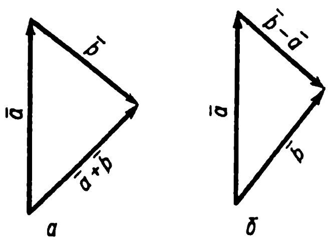


Рис. 8

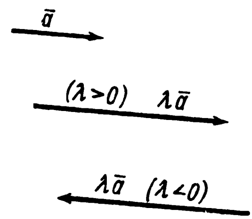


Рис. 9

Геометрический смысл операции умножения вектора на число можно выразить следующим образом: если $|\lambda|>1$, то при умножении вектора $\bar{a}$ на число λ вектор $\bar{a}$ «растягивается» в λ раз, а если $|\lambda|<1$ — «сжимается». При $\lambda<0$ происходит еще изменение направления вектора на противоположное. На рис. 9 $|\lambda|>1$.

Замечание 3. В случае, когда $\lambda=0$ или $\bar{a}=0$, произведение $\lambda\bar{a}$ имеет модуль, равный нулю, и, следовательно, представляет собой нулевой вектор.

Основные свойства линейных операций

1°. $\bar{a}+\bar{б}=\bar{б}+\bar{a}$ (*переместительное свойство суммы*).

Доказательство. Приложив вектор $\bar{a}$ и $\bar{б}$ к одной точке O, построим на них параллелограмм (рис. 10). Тогда

$$\bar{a}=\overline{OA},\ \bar{б}=\overline{AC},\ \bar{a}+\bar{б}=\overline{OC},$$
$$\bar{б}=\overline{OB},\ \bar{a}=\overline{BC},\ \bar{б}+\bar{a}=\overline{OC}.$$

Следовательно, $\overline{OC}=\bar{a}+\bar{б}=\bar{б}+\bar{a}$ ■

2°. $(\bar{a}+\bar{б})+\bar{c}=\bar{a}+(\bar{б}+\bar{c})$ (*сочетательное свойство суммы*).

Д о к а з а т е л ь с т в о. Расположим рассматриваемые векторы так, чтобы вектор $\bar{б}$ был приложен к концу вектора $\bar{a}$, а вектор $\bar{c}$ — к концу вектора $\bar{б}$. Обозначим буквой O начало вектора $\bar{a}$, буквой A — его конец, буквой B — конец вектора $\bar{б}$ и буквой C — конец вектора $\bar{c}$ (рис. 11). Тогда

$$(\bar{a}+\bar{б})+\bar{c}=(\overline{OA}+\overline{AB})+\overline{BC}=\overline{OB}+\overline{BC}=\overline{OC},$$

$$\bar{a}+(\bar{б}+\bar{c})=\overline{OA}+(\overline{AB}+\overline{BC})=\overline{OA}+\overline{AC}=\overline{OC}.$$

Следовательно, $\overline{OC}=(\bar{a}+\bar{б})+\bar{c}=\bar{a}+(\bar{б}+\bar{c})$ ■

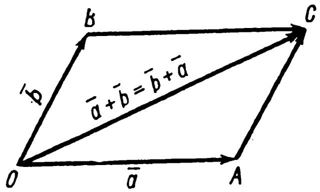


Рис. 10

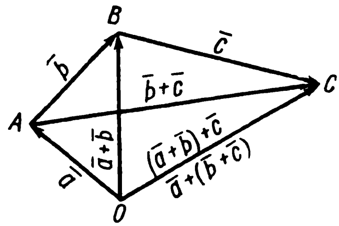


Рис. 11

3°. Рассмотрим теперь три свойства линейных операций, которые относятся одновременно к сложению векторов и умножению вектора на число. Пусть λ и μ — произвольные числа, $\bar{a}$ и $\bar{б}$ — любые векторы. Тогда

1) $\lambda(\mu\bar{a})=(\lambda\mu)\bar{a}$ (*сочетательное свойство умножения*);
2) $(\lambda+\mu)\bar{a}=\lambda\bar{a}+\mu\bar{a}$ (*распределительное свойство относительно суммы чисел*);
3) $\lambda(\bar{a}+\bar{б})=\lambda\bar{a}+\lambda\bar{б}$ (*распределительное свойство относительно суммы векторов*).

Д о к а з а т е л ь с т в о. *Докажем первое свойство.* Векторы $\lambda(\mu\bar{a})$ и $(\lambda\mu)\bar{a}$ коллинеарны (лежат на одной прямой), одинаково направлены (их направления либо совпадают с направлением вектора $\bar{a}$, если λ и μ одного знака, либо противоположны, если λ и μ разных знаков) и имеют одинаковые длины $(|\lambda(\mu\bar{a})|=|\lambda||\mu\bar{a}|=|\lambda||\mu||\bar{a}|=|\lambda\mu||\bar{a}|)$, следовательно, они равны. Если хотя бы одно из чисел λ, μ или вектор $\bar{a}$ равны нулю, то обе части равенства 1) обращаются в нуль. Свойство доказано.

Докажем второе свойство. Пусть сначала λ и μ одного знака. Тогда векторы $(\lambda+\mu)\bar{a}$ и $\lambda\bar{a}+\mu\bar{a}$ коллинеарны (лежат на одной прямой), одинаково направлены (при $\lambda>0$, $\mu>0$ их направления совпадают с направлением вектора $\bar{a}$, а при $\lambda<0$, $\mu<0$ противоположны $\bar{a}$). Так как векторы $\lambda\bar{a}$ и $\mu\bar{a}$ направлены одинаково, то длина вектора $\lambda\bar{a}+\mu\bar{a}$ равна $|\lambda\bar{a}+\mu\bar{a}|=|\lambda\bar{a}|+|\mu\bar{a}|=$ $=|\lambda||\bar{a}|+|\mu||\bar{a}|=(|\lambda|+|\mu|)|\bar{a}|=|\lambda+\mu||\bar{a}|$, т. е. равна длине вектора $(\lambda+\mu)\bar{a}$. Таким образом, векторы $(\lambda+\mu)\bar{a}$ и $\lambda\bar{a}+\mu\bar{a}$ коллинеарны, одинаково направлены и имеют равные длины, следовательно, они равны.

Пусть теперь λ и μ разных знаков и для определенности $|\lambda|>|\mu|$. В этом случае векторы $(\lambda+\mu)\bar{a}$ и $\lambda\bar{a}+\mu\bar{a}$ направлены так же, как вектор $\lambda\bar{a}$. Длина вектора $\lambda\bar{a}+\mu\bar{a}$ равна $|\lambda\bar{a}+\mu\bar{a}|=|\lambda\bar{a}|-|\mu\bar{a}|=$ $=|\lambda||\bar{a}|-|\mu||\bar{a}|=(|\lambda|-|\mu|)|\bar{a}|=|\lambda+\mu||\bar{a}|$, т. е. равна длине вектора $(\lambda+\mu)\bar{a}$. Следовательно, и в этом случае векторы $(\lambda+\mu)\bar{a}$ и $\lambda\bar{a}+\mu\bar{a}$ равны. Если же $|\lambda|=|\mu|$ и знаки λ и μ разные, то обе части равенства 2) равны нулю. То же самое будет, если вектор $\bar{a}=0$ или λ и μ равны нулю одновременно. Равенство 2) очевидно, если хотя бы одно из чисел λ, μ или вектор $\bar{a}$ равны нулю. Свойство доказано.

Докажем третье свойство. Пусть $\bar{a}$ и $\bar{б}$ — неколлинеарные векторы и $\lambda>0$. Построим векторы $\overline{OB}=\bar{a}+\bar{б}$, $\overline{OA'}=\bar{a}\cdot\lambda$ и $\overline{OB'}=(\bar{a}+\bar{б})\cdot\lambda$ (рис. 12). Из подобия треугольников OAB и $OA'B'$ и определения операции умножения вектора на число следует, что $\overline{A'B'}=\bar{б}\cdot\lambda$. Тогда из треугольника $OA'B'$ получим: $\overline{OA'}+\overline{A'B'}=\overline{OB'}$, или $\lambda\bar{a}+\lambda\bar{б}=\lambda(\bar{a}+\bar{б})$, т. е. равенство 3). Случай $\lambda<0$ рассматривается аналогично.

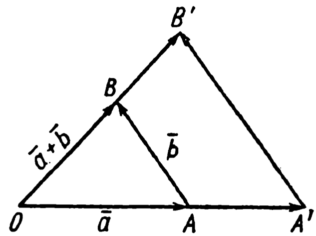


Рис. 12

Если $\bar{a}$ и $\bar{б}$ — коллинеарные векторы и λ — любое, то вектор $\bar{б}$ можно представить в виде $\bar{б}=\mu\bar{a}$ и равенство 3) следует из равенства 2). Действительно $(\lambda+\lambda\mu)\bar{a}=\lambda(1+$ $+\mu)\bar{a}=\lambda(\bar{a}+\mu\bar{a})=\lambda\bar{a}+\lambda\mu\bar{a}$. Отсюда получаем $\lambda(\bar{a}+\bar{б})=$

$=\lambda\bar{a}+\lambda\bar{b}$. Равенство 3) очевидно, если один из векторов $\bar{a}$, $\bar{b}$ или число λ равны нулю. Свойство доказано ■

Замечание 4. Доказанные свойства линейных операций имеют фундаментальное значение, так как они дают право проводить над векторами обычные алгебраические действия. Например, в силу второго и третьего свойств 3° можно проводить умножения скалярного многочлена на векторный многочлен «почленно».

§ 4. Теоремы о проекциях векторов

Рассмотрим две основные теоремы о проекциях.

Теорема 9.3. *Проекция суммы двух векторов на ось равна сумме их проекций на эту ось, т. е.*

$$\text{Пр.}_u(\bar{a}_1+\bar{a}_2)=\text{Пр.}_u\bar{a}_1+\text{Пр.}_u\bar{a}_2.$$

Доказательство. Обозначим концы векторов $\bar{a}_1$ и $\bar{a}_2$ соответственно через A_1, A_2, начало вектора $\bar{a}_1$ — через A, а через A', A_1' и A_2' — проекции на ось Ou точек A, A_1 и A_2. Для определенности рассмотрим случай, когда направления векторов $\overline{A'A_1'}$ и $\overline{A_1'A_2'}$ совпадают с направлением оси (рис. 13). Из построения следует:

$$\bar{a}_1+\bar{a}_2=\overline{AA_2}, \tag{1}$$

и, кроме того,

$$|\overline{A'A_1'}|=\text{Пр.}_u\bar{a}_1,\quad |\overline{A_1'A_2'}|=\text{Пр.}_u\bar{a}_2. \tag{2}$$

В силу равенства (1)

$$\text{Пр.}_u(\bar{a}_1+\bar{a}_2)=\text{Пр.}_u\overline{AA_2}=|\overline{A'A_2'}|=|\overline{A'A_1'}|+|\overline{A_1'A_2'}|. \tag{3}$$

Из равенства (3), принимая во внимание (2), находим, что

$$\text{Пр.}_u(\bar{a}_1+\bar{a}_2)=\text{Пр.}_u\bar{a}_1+\text{Пр.}_u\bar{a}_2.$$

Остальные случаи взаимного расположения векторов рассматриваются аналогично ■

Теорему можно обобщить на случай любого числа слагаемых.

Теорема 9.4. *При умножении вектора $\bar{a}$ на число λ*

его проекция на ось также умножается на это число, т. е.

$$\text{Пр.}_u\lambda\bar{a}=\lambda\text{Пр.}_u\bar{a}. \qquad (4)$$

Доказательство. Пусть φ — угол между вектором $\bar{a}$ и осью Ou, а φ' — угол между вектором $\lambda\bar{a}$ и осью Ou (рис. 14). Тогда, если $\lambda>0$, то векторы $\bar{a}$ и $\lambda\bar{a}$ направлены одинаково и $\varphi=\varphi'$. Если $\lambda<0$, то векторы

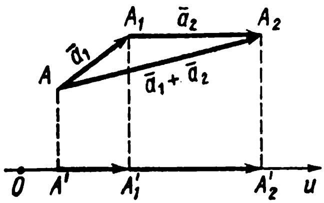


Рис. 13

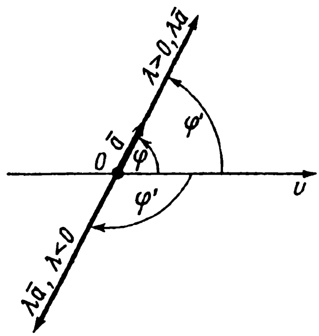


Рис. 14

$\bar{a}$ и $\lambda\bar{a}$ имеют противоположные направления и $\varphi'=\pi-\varphi$. По теореме 9.1 имеем:

при $\lambda>0$

$$\text{Пр.}_u\lambda\bar{a}=|\lambda\bar{a}|\cos\varphi'=\lambda|\bar{a}|\cos\varphi'=\lambda|\bar{a}|\cos\varphi=\lambda\text{Пр.}_u\bar{a},$$

при $\lambda<0$

$$\text{Пр.}_u\lambda\bar{a}=|\lambda\bar{a}|\cos\varphi'=|\lambda|\,|\bar{a}|\cos\varphi'=$$
$$=-\lambda|\bar{a}|\cos(\pi-\varphi)=-\lambda|\bar{a}|(-\cos\varphi)=$$
$$=\lambda|\bar{a}|\cos\varphi=\lambda\text{Пр.}_u\bar{a},$$

при $\lambda=0$ равенство (4) очевидно. Таким образом, при любом $\lambda\text{Пр.}_u\lambda\bar{a}=\lambda\text{Пр.}_u\bar{a}$ ■

Из доказанных теорем вытекают два важных следствия.

Следствие 1. Из теоремы 9.3 вытекает, что если $\bar{a}=\{X_1; Y_1; Z_1\}$ и $\bar{b}=\{X_2; Y_2; Z_2\}$, то $\bar{a}+\bar{b}=\{X_1+X_2; Y_1+Y_2; Z_1+Z_2\}$.

С л е д с т в и е 2. Из теоремы 9.4 вытекает, что если $\bar{a}=\{X; Y; Z\}$, то для любого числа λ будет $\lambda\cdot\bar{a}=\{\lambda X; \lambda Y; \lambda Z\}$.

В частности, из только что сказанного легко выводится условие коллинеарности двух векторов. В самом деле, равенство $\bar{b}=\lambda\bar{a}$ равносильно равенствам

$$X_2=\lambda X_1,\ Y_2=\lambda Y_1,\ Z_2=\lambda Z_1,$$

или
$$\frac{X_2}{X_1}=\frac{Y_2}{Y_1}=\frac{Z_2}{Z_1}, \qquad (5)$$

т. е. векторы $\bar{a}$ и $\bar{b}$ коллинеарны в том и только в том случае, когда их координаты пропорциональны.

§ 5. Разложение вектора по базису

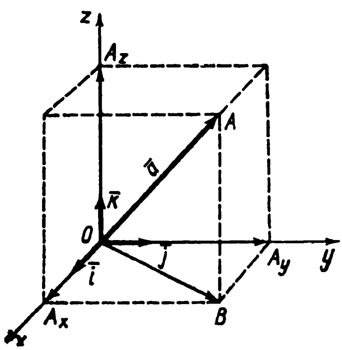


Рис. 15

Пусть векторы $\bar{i}$, $\bar{j}$ $\bar{k}$ — единичные, т. е. $|\bar{i}|=|\bar{j}|=|\bar{k}|=1$, приложены к началу координат и каждый направлен на своей оси в положительную сторону (рис. 15). Тогда имеет место

Т е о р е м а 9.5. *Любой вектор $\bar{a}$ может быть единственным образом разложен по базису $\bar{i}$, $\bar{j}$, $\bar{k}$, т. е. может быть представлен в виде*

$$\bar{a}=\lambda\bar{i}+\mu\bar{j}+\nu\bar{k}, \qquad (1)$$

где λ, μ, ν — некоторые числа.

Д о к а з а т е л ь с т в о. Приложив вектор $\bar{a}$ к началу координат, обозначим его конец через A. Проведем через точку A плоскости, перпендикулярные осям координат. Пусть A_x, A_y, A_z — точки пересечения этих плоскостей с осями координат. По определению сложения векторов имеем

$$\bar{a}=\overline{OB}+\overline{OA_z},\ \overline{OB}=\overline{OA_x}+\overline{OA_y}. \qquad (2)$$

Из равенства (2) получаем

$$\bar{a}=\overline{OA_x}+\overline{OA_y}+\overline{OA_z}. \qquad (3)$$

Кроме того, можем записать

$$\overline{OA_x}=\lambda\bar{i},\ \overline{OA_y}=\mu\bar{j},\ \overline{OA_z}=\nu\bar{k} \tag{4}$$

(λ, μ, ν — некоторые числа).

Из равенства (3) и соотношений (4) получаем

$$\bar{a}=\lambda\bar{i}+\mu\bar{j}+\nu\bar{k}.$$

Для доказательства единственности представления (1) установим, что

$$\lambda=X,\ \mu=Y,\ \nu=Z,$$

где X, Y, Z — координаты вектора $\bar{a}$.

Покажем, например, что $\lambda=X$. Так как $X=|\overline{OA_x}|$, если $\overline{OA_x}$ имеет то же направление, что и вектор $\bar{i}$, и $X=-|\overline{OA_x}|$, если вектор $\overline{OA_x}$ имеет направление, противоположное направлению вектора $\bar{i}$, то $\overline{OA_x}=X\bar{i}$. Сравнивая с равенством $\overline{OA_x}=\lambda\bar{i}$, получаем $\lambda=X$.

Аналогично показывается, что $\mu=Y$ и $\nu=Z$ ■

§ 6. Скалярное произведение и его основные свойства

Этот и следующие два параграфа посвящены рассмотрению некоторых нелинейных операций над векторами.

Определение. *Скалярным произведением двух векторов $\bar{a}$ и $\bar{b}$ называется число (скаляр), равное произведению длин этих векторов на косинус угла между ними. Если хоть один из векторов нулевой, то угол не определен, и скалярное произведение по определению полагают равным нулю.*

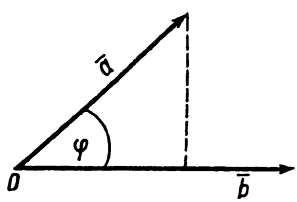


Рис. 16

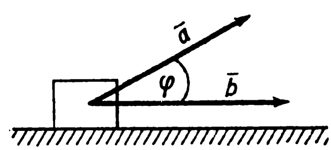


Рис. 17

Скалярное произведение обозначается $\bar{a}\cdot\bar{b}$. Итак,

$$\bar{a}\cdot\bar{b}=|\bar{a}|\,|\bar{b}|\cos\varphi.$$

Так как $|\bar{a}|\cos\varphi=\text{Пр.}_{\bar{b}}\,\bar{a}$, $|\bar{b}|\cos\varphi=\text{Пр.}_{\bar{a}}\,\bar{b}$ (рис. 16), то можем записать

$$\bar{a}\cdot\bar{b}=|\bar{b}|\,\text{Пр.}_{\bar{b}}\,\bar{a}=|\bar{a}|\,\text{Пр.}_{\bar{a}}\,\bar{b}. \qquad (1)$$

Понятие скалярного произведения имеет свой источник в физике. Именно, если вектор $\bar{a}$ изображает силу, точка приложения которой перемещается из начала в конец вектора $\bar{b}$ (рис. 17), то работа A этой силы определяется равенством

$$A=|\bar{a}|\,|\bar{b}|\cos\varphi.$$

Рассмотрим некоторые свойства скалярного произведения.

Алгебраические свойства скалярного произведения

1°. $\bar{a}\cdot\bar{b}=\bar{a}\cdot\bar{b}$ (*свойство перестановочности сомножителей*).

Доказательство. По определению скалярного произведения $\bar{a}\cdot\bar{b}=|\bar{a}|\,|\bar{b}|\cos\varphi$ и $\bar{b}\cdot\bar{a}=|\bar{b}|\,|\bar{a}|\cos\varphi$, но $|\bar{a}|\,|\bar{b}|=|\bar{b}|\,|\bar{a}|$, поскольку это есть произведение чисел. Следовательно, $\bar{a}\cdot\bar{b}=\bar{b}\cdot\bar{a}$ ■

2°. $(\lambda\bar{a})\cdot\bar{b}=\lambda(\bar{a}\cdot\bar{b})$ (*свойство сочетательности относительно умножения на число*).

Доказательство. По формуле (1)

$$(\lambda\bar{a})\cdot\bar{b}=|\bar{b}|\,\text{Пр.}_{\bar{b}}\,(\lambda\bar{a}),$$

но согласно теореме 9.4

$$\text{Пр.}_{\bar{b}}\,(\lambda\bar{a})=\lambda\,\text{Пр.}_{\bar{b}}\,\bar{a}.$$

Таким образом,

$$(\lambda\bar{a})\cdot\bar{b}=|\bar{b}|\,\text{Пр.}_{\bar{b}}\,(\lambda\bar{a})=|\bar{b}|\,\lambda\,\text{Пр.}_{\bar{b}}\,\bar{a}=\lambda\,(|\bar{b}|\,\text{Пр.}_{\bar{b}}\,\bar{a}).$$

С другой стороны, по той же формуле (1)

$$|\bar{b}|\,\text{Пр.}_{\bar{b}}\,\bar{a}=\bar{a}\cdot\bar{b}.$$

Следовательно

$$(\lambda\bar{a})\cdot\bar{b}=\lambda\,(|\bar{b}|\,\text{Пр.}_{\bar{b}}\,\bar{a})=\lambda\,(\bar{a}\cdot\bar{b})$$

Замечание 1. Из свойств 1° и 2° следует, что $(\lambda\bar{a})\cdot(\mu\bar{b})=(\lambda\cdot\mu)\,(\bar{a}\cdot\bar{b})$. Действительно, $(\lambda\bar{a})\cdot(\mu\bar{b})=$ $=\lambda[\bar{a}\cdot(\mu\bar{b})]=\lambda[(\mu\bar{b})\cdot\bar{a}]=\lambda[\mu(\bar{b}\cdot\bar{a})]=(\lambda\mu)\,(\bar{a}\cdot\bar{b})$.

3°. $\bar{a}\cdot(\bar{b}+\bar{c})=\bar{a}\cdot\bar{b}+\bar{a}\cdot\bar{c}$ (*свойство распределительности относительно сложения*).

Доказательство. По формуле (1)

$$\bar{a}\cdot(\bar{b}+\bar{c})=|\bar{a}|\,\text{Пр.}_{\bar{a}}\,(\bar{b}+\bar{c}),$$

но согласно теореме 9.3

$$\text{Пр}_{\bar{a}}\,(\bar{b}+\bar{c})=\text{Пр.}_{\bar{a}}\,\bar{b}+\text{Пр.}_{\bar{a}}\,\bar{c}.$$

Таким образом,

$$\bar{a}\cdot(\bar{b}+\bar{c})=|\bar{a}|\,\text{Пр.}_{\bar{a}}\,(\bar{b}+\bar{c})=|\bar{a}|(\text{Пр.}_{\bar{a}}\,\bar{b}+\text{Пр.}_{\bar{a}}\,\bar{c})=$$

$$=|\bar{a}|\,\text{Пр.}_{\bar{a}}\,\bar{b}+|\bar{a}|\,\text{Пр.}_{\bar{a}}\,\bar{c}.$$

С другой стороны, по той же формуле (1)

$$|\bar{a}|\,\text{Пр.}_{\bar{a}}\,\bar{b}=\bar{a}\cdot\bar{b}\ \text{и}\ |\bar{a}|\,\text{Пр.}_{\bar{a}}\,\bar{c}=\bar{a}\cdot\bar{c}.$$

Следовательно,

$$\bar{a}\cdot(\bar{b}+\bar{c})=|\bar{a}|\,\text{Пр}\cdot_{\bar{a}}\,\bar{b}+|\bar{a}|\,\text{Пр}\cdot_{\bar{a}}\,\bar{c}=\bar{a}\cdot\bar{b}+\bar{a}\cdot\bar{c}\ \blacksquare$$

Замечание 2. Доказанное свойство дает право при скалярном перемножении векторных многочленов выполнять действие почленно. В силу свойства 1° можно при этом не заботиться о порядке сомножителей, а свойство 2° позволяет (см. замечание 1) объединить числовые коэффициенты векторных сомножителей. Например:

$$(2\bar{a}+5\bar{b})\cdot(3\bar{c}+4\bar{d})=$$
$$=(2\bar{a}+5\bar{b})\cdot(3\bar{c})+(2\bar{a}+5\bar{b})\cdot(4\bar{d})=$$
$$=(2\bar{a})\cdot(3\bar{c})+(5\bar{b})\cdot(3\bar{c})+(2\bar{a})\cdot(4\bar{d})+(5\bar{b})\cdot(4\bar{d})=$$
$$=6\bar{a}\cdot\bar{c}+15\bar{b}\cdot\bar{c}+8\bar{a}\cdot\bar{d}+20\bar{b}\cdot\bar{d}.$$

4°. $\bar{a}\cdot\bar{a}=|\bar{a}|^2$, *если* $\bar{a}\neq 0$.

Доказательство. По определению скалярного произведения $\bar{a}\cdot\bar{a}=|\bar{a}|\,|\bar{a}|\cos 0=|\bar{a}|^2$, если $|\bar{a}|\neq 0$, т. е. если $\bar{a}\neq 0$ ■

Скалярное произведение $\bar{a}\cdot\bar{a}$ называется *скалярным квадратом* вектора $\bar{a}$ и обозначается $\bar{a}^2$. На основании только что доказанного мы имеем: $\bar{a}^2=|\bar{a}|^2$; отсюда, в частности, $\sqrt{\bar{a}^2}=|\bar{a}|$.

5° $\bar{a}\cdot\bar{b}=0$, *если* $\bar{a}\perp\bar{b}$.

Доказательство. По определению скалярного произведения $\bar{a}\cdot\bar{b}=|\bar{a}|\,|\bar{b}|\cos\varphi$ $(0\leqslant\varphi\leqslant\pi)$. Если $\varphi=\frac{\pi}{2}$, т. е. векторы $\bar{a}$ и $\bar{b}$ перпендикулярны друг другу, то $\cos\frac{\pi}{2}=0$; отсюда $\bar{a}\cdot\bar{b}=0$.

Верно и обратное, если $\bar{a}\cdot\bar{b}=0$ и $|\bar{a}|\,|\bar{b}|\neq 0$, то $\cos\varphi=0$ и $\varphi=\frac{\pi}{2}$, т. е. векторы $\bar{a}$ и $\bar{b}$ перпендикулярны ■

Пример.

$$(\bar{a}+\bar{b}+\bar{c})\cdot(2\bar{b}+3\bar{c})=\bar{a}\cdot 2\bar{b}+\bar{b}\cdot 2\bar{b}+\bar{c}\cdot 2\bar{b}+$$
$$+\bar{a}\cdot 3\bar{c}+\bar{b}\cdot 3\bar{c}+\bar{c}\cdot 3\bar{c}=2\bar{a}\cdot\bar{b}+2\bar{b}^2+2\bar{c}\cdot\bar{b}+3\bar{a}\cdot\bar{c}+$$
$$+3\bar{b}\cdot\bar{c}+3\bar{c}^2=2\bar{a}\cdot\bar{b}+2\bar{b}^2+5\bar{c}\cdot\bar{b}+3\bar{a}\cdot\bar{c}+3\bar{c}^2.$$

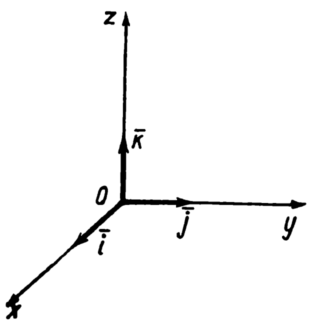


Рис. 18

Замечание 3. Из геометрических свойств 4° и 5° для базисных векторов $\bar{i}$, $\bar{j}$, $\bar{k}$ (рис. 18) непосредственно получаем следующие равенства:

$$\bar{i}^2=\bar{j}^2=\bar{k}^2=1 \qquad (2)$$

и

$$\bar{i}\cdot\bar{j}=\bar{i}\cdot\bar{k}=\bar{j}\cdot\bar{i}=\bar{j}\cdot\bar{k}=$$
$$\bar{k}\cdot\bar{i}=\bar{k}\cdot\bar{j}=0.$$

Выражение скалярного произведения через координаты векторов

Теорема 9.6. *Если векторы $\bar{a}$ и $\bar{b}$ заданы своими координатами: $\bar{a}=\{X_1; Y_1; Z_1\}$, $\bar{b}=\{X_2; Y_2; Z_2\}$, то их скалярное произведение выражается формулой*

$$\bar{a}\cdot\bar{b}=X_1X_2+Y_1Y_2+Z_1Z_2.$$

Доказательство. Согласно теореме 9.5 разложим векторы $\bar{a}$ и $\bar{b}$ по базису $\bar{i}$, $\bar{j}$, $\bar{k}$: $\bar{a}=X_1\bar{i}+Y_1\bar{j}+Z_1\bar{k}$, $\bar{b}=X_2\bar{i}+Y_2\bar{j}+Z_2\bar{k}$. На основании установленных алгебраических свойств 2° и 3° перемножим правые части равенств почленно:

$$\bar{a}\cdot\bar{b}=X_1X_2\bar{i}^2+X_1Y_2\bar{i}\cdot\bar{j}+X_1Z_2\bar{i}\cdot\bar{k}+Y_1X_2\bar{j}\cdot\bar{i}+$$
$$+Y_1Y_2\bar{j}^2+Y_1Z_2\bar{j}\cdot\bar{k}+Z_1X_2\bar{k}\cdot\bar{i}+Z_1Y_2\bar{k}\cdot\bar{j}+Z_1Z_2\bar{k}^2.$$

Откуда, пользуясь равенствами (2), находим: $\bar{a}\cdot\bar{b}=$ $=X_1X_2+Y_1Y_2+Z_1Z_2$ ■

Из теоремы 9.6 вытекают два важных следствия.

Следствие 1. Необходимым и достаточным условием перпендикулярности векторов $\bar{a}$ и $\bar{b}$ является равенство

$$X_1X_2+Y_1Y_2+Z_1Z_2=0. \quad (3)$$

В самом деле, в силу свойства 5° $\bar{a}\cdot\bar{b}=0$, если $\bar{a}$ и $\bar{b}$ взаимно перпендикулярны, т. е. векторы $\bar{a}$ и $\bar{b}$ перпендикулярны в том и только в том случае, если выполняется равенство (3).

Следствие 2. Угол между векторами $\bar{a}$ и $\bar{b}$ определяется равенством

$$\cos\varphi=\frac{X_1X_2+Y_1Y_2+Z_1Z_2}{\sqrt{X_1^2+Y_1^2+Z_1^2}\,\sqrt{X_2^2+Y_2^2+Z_2^2}}. \quad (4)$$

Действительно, по определению скалярного произведения $\bar{a}\cdot\bar{b}=|\bar{a}|\,|\bar{b}|\cos\varphi$, отсюда

$$\cos\varphi=\frac{\bar{a}\cdot\bar{b}}{|\bar{a}|\,|\bar{b}|}. \quad (5)$$

В силу теоремы 9.6 и формулы (4) § 2 из формулы (5) следует формула (4).

Пример. Даны три точки $A(1; 1; 1)$, $B(2; 2; 1)$ и $C(2; 1; 2)$. Найти угол $\varphi=\angle BAC$.

Решение. Применяя теорему 9.2, находим: $\overline{AB} = \{1; 1; 0\}$, $\overline{AC} = \{1; 0; 1\}$. Отсюда на основании формулы (4) получаем

$$\cos\varphi = \frac{1\cdot 1 + 1\cdot 0 + 0\cdot 1}{\sqrt{1^2+1^2+0^2}\,\sqrt{1^2+0^2+1^2}} = \frac{1}{\sqrt{2}\cdot\sqrt{2}} = \frac{1}{2}.$$

Следовательно, $\varphi = 60°$.

§ 7. Векторное произведение и его основные свойства

Векторы $\bar{a}$, $\bar{b}$ и $\bar{c}$ называются *компланарными*, если они лежат в одной плоскости или в параллельных плоскостях.

Тройка векторов, приведенных к общему началу, называется *упорядоченной*, если указано, какой из них считается первым, какой вторым и какой третьим. Например, если мы пишем: $(\bar{a}; \bar{b}; \bar{c})$, то $\bar{a}$ считается первым вектором, $\bar{b}$ — вторым, $\bar{c}$ — третьим; если мы пишем: $(\bar{b}; \bar{c}; \bar{a})$, то $\bar{b}$ считается первым вектором, $\bar{c}$ — вторым, $\bar{a}$ — третьим.

Упорядоченная тройка некомпланарных векторов называется *правой*, если из конца третьего вектора кратчайший поворот от первого ко второму виден против часовой стрелки. В противоположном случае тройка называется *левой*.

Перейдем теперь к рассмотрению второго вида умножения двух векторов.

Определение. *Векторным произведением вектора $\bar{a}$ на вектор $\bar{b}$ называется вектор, обозначаемый символом $\bar{a}\times\bar{b}$, который определяется следующими тремя условиями:*

1) *длина вектора $\bar{a}\times\bar{b}$ равна $|\bar{a}|\,|\bar{b}|\sin\varphi$, где φ — угол между векторами $\bar{a}$ и $\bar{b}$;*

2) *вектор $\bar{a}\times\bar{b}$ перпендикулярен к каждому из векторов $\bar{a}$ и $\bar{b}$;*

3) *векторы $\bar{a}$, $\bar{b}$, $\bar{a}\times\bar{b}$ образуют правую тройку векторов (рис. 19).*

Понятие векторного произведения имеет свой источник в механике.

Пусть к точке M твердого тела приложена сила $\bar{F} = \overline{MK}$ и пусть O — некоторая точка пространства. Как известно из механики, моментом силы $\bar{F}$ относительно точки O (центр момента) называется вектор $\bar{L}$, который: 1) имеет начало в точке O; 2) перпендикулярен к плоскости π, проходящей через точки O, M, K; 3) направлен так, что из конца его сила $\bar{F}$ представляется

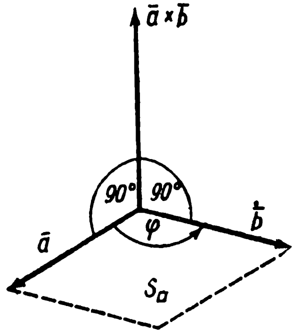


Рис. 19

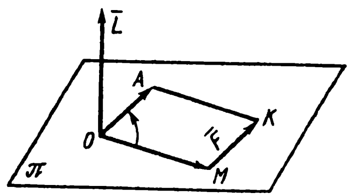


Рис. 20

вращающей плоскость π вокруг точки O против часовой стрелки (рис. 20). Из рисунка видно, что $\bar{L}$ представляет собой векторное произведение $\bar{L}=\overline{OM}\times\overline{OA}$, где $\overline{OA}=\bar{F}$.

Рассмотрим основные свойства векторного произведения.

Геометрические свойства векторного произведения

1°. $\bar{a}\times\bar{b}=0$, *если* $\bar{a}$ *и* $\bar{b}$ — *коллинеарные векторы.*

Доказательство. Если векторы $\bar{a}$ и $\bar{b}$ коллинеарны, то $\sin\varphi=0$. Следовательно, $|\bar{a}\times\bar{b}|=|\bar{a}|\,|\bar{b}|\sin\varphi=0$, т. е. длина вектора $\bar{a}\times\bar{b}$ равна нулю, а значит, и сам вектор $\bar{a}\times\bar{b}$ равен нулю ■

2°. *Длина векторного произведения* $|\bar{a}\times\bar{b}|$ *равна площади параллелограмма* S, *построенного на векторах* $\bar{a}$ *и* $\bar{b}$ *(см. рис. 19).*

Доказательство. Как известно из элементарной геометрии, площадь параллелограмма равна произведению его смежных сторон на синус угла между ними. Отсюда $|\bar{a}|\,|\bar{b}|\sin\varphi=S$, следовательно, $|\bar{a}\times\bar{b}|=S$ ■

3°. $\bar{a}\times\bar{b}=-\bar{b}\times\bar{a}$ (*свойство антиперестановочности сомножителей*).

Доказательство. Из определения векторного произведения следует, что векторы $\bar{a}\times\bar{b}$ и $\bar{b}\times\bar{a}$ имеют одинаковые длины (длина векторного произведения не зависит от порядка сомножителей), коллинеарны (они перпендикулярны одной и той же плоскости, в которой лежат векторы $\bar{a}$ и $\bar{b}$), но направлены противоположно

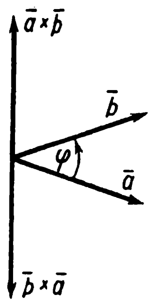


Рис. 21

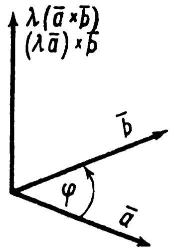


Рис. 22

(рис. 21), так как векторы $\bar{a}$, $\bar{b}$, $\bar{a}\times\bar{b}$ и $\bar{b}$, $\bar{a}$, $\bar{b}\times\bar{a}$ образуют правые тройки, следовательно, $\bar{a}\times\bar{b}=-\bar{b}\times\bar{a}$ ■

4°. $(\lambda\bar{a})\times\bar{b}=\lambda(\bar{a}\times\bar{b})$ * (*свойство сочетательности по отношению к скалярному множителю*).

Доказательство. Из определения векторного произведения следует, что $|\lambda(\bar{a}\times\bar{b})|=|\lambda||\bar{a}||\bar{b}|\sin\varphi$, но $|(\lambda\bar{a})\times\bar{b}|=|\lambda||\bar{a}||\bar{b}|\sin\varphi$, поэтому векторы $(\lambda\bar{a})\times\bar{b}$ и $\lambda(\bar{a}\times\bar{b})$ имеют одинаковую длину; кроме этого, они перпендикулярны каждому из векторов $\bar{a}$ и $\bar{b}$, а значит, коллинеарны друг другу и одинаково направлены (рис. 22) (при $\lambda>0$ это очевидно, при $\lambda<0$ векторы $\lambda\bar{a}$ и $\bar{a}$ имеют противоположные направления, отсюда вектор $(\lambda\bar{a})\times\bar{b}$ направлен противоположно вектору $\bar{a}\times\bar{b}$; но при $\lambda<0$ вектор $\lambda(\bar{a}\times\bar{b})$ также направлен противоположно вектору $\bar{a}\times\bar{b}$, т. е. и при $\lambda<0$ они имеют оди-

* К этому равенству сводится равенство $\bar{a}\times(\lambda\bar{b})=\lambda(\bar{a}\times\bar{b})$ при помощи перестановки сомножителей векторных произведений в его левой и правой частях с последующей заменой букв.

наковое направление), следовательно, они равны. При $\lambda=0$ свойство очевидно ■

5°. $(\bar{a}+\bar{b})\times\bar{c}=\bar{a}\times\bar{c}+\bar{b}\times\bar{c}$ * (*распределительное свойство относительно сложения*).

Доказательство. Введем для доказательства единичный вектор $\bar{c}_0$, направленный произвольным образом. Проведем через его начало O плоскость π, перпендикулярную $\bar{c}_0$, и рассмотрим треугольник OAB со сторонами $OA=\bar{a}$, $AB=\bar{b}$ и $OB=\bar{a}+\bar{b}$ (рис. 23). Спроектируем треугольник OAB на плоскость π, получим треугольник OA_1B_1 (если точка A_1 лежит на прямой OB_1, то треугольник OA_1B_1 вырождается в отрезок). Повернем треугольник OA_1B_1 вокруг $\bar{c}_0$ на 90° по часовой стрелке, если смотреть из конца $\bar{c}_0$, получим треугольник OA_2B_2. Обозначим через φ угол между векторами $\bar{c}_0$ и $\bar{a}$. Пусть для определенности $0<\varphi<\frac{\pi}{2}$ (как на рис. 23). Остальные случаи рассматриваются аналогично.

Рассмотрим вектор $\overline{OA_2}$. Длина этого вектора

$$|\overline{OA_2}|=|\overline{OA_1}|=|\bar{a}|\cos\left(\frac{\pi}{2}-\varphi\right)=|\bar{a}|\,|\bar{c}_0|\sin\varphi\ (|\bar{c}_0|=1).$$

Кроме этого, $\overline{OA_2}\perp\bar{c}_0$, $\overline{OA_2}\perp\bar{a}$, и векторы $\bar{a}$, $\bar{c}_0$ и $\overline{OA_2}$ образуют правую тройку. Следовательно, по определению векторного произведения

$$\overline{OA_2}=\bar{a}\times\bar{c}_0.$$

Проведя аналогичные рассуждения для каждого из векторов $\overline{OB_2}$ и $\overline{A_2B_2}$, получим

$$\overline{OB_2}=(\bar{a}+\bar{b})\times\bar{c}_0,\quad \overline{A_2B_2}=\bar{b}\times\bar{c}_0.$$

Но так как $\overline{OB_2}=\overline{OA_2}+\overline{A_2B_2}$, то

$$(\bar{a}+\bar{b})\times\bar{c}_0=\bar{a}\times\bar{c}_0+\bar{b}\times\bar{c}_0. \qquad (1)$$

Пусть теперь $\bar{c}$ — какой угодно вектор, направленный так же, как $\bar{c}_0$. Тогда имеем: $\bar{c}=|\bar{c}|\bar{c}_0$, где $|\bar{c}|$ — длина вектора $\bar{c}$ ($\bar{c}=\lambda\bar{c}_0$, $|\bar{c}|=\lambda|\bar{c}_0|=\lambda\cdot 1$, отсюда $\lambda=$

---

* К этому равенству сводится равенство $\bar{c}\times(\bar{a}+\bar{b})=\bar{c}\times\bar{a}+\bar{c}\times\bar{b}$ при помощи перестановки сомножителей в его левой и правой частях.

$=|\bar{c}|$). Умножим обе части равенства (1) на число $|\bar{c}|$, получим $|\bar{c}|\,[(\bar{a}+\bar{b})\times\bar{c}_0]=|\bar{c}|\,(\bar{a}\times\bar{c}_0+\bar{b}\times\bar{c}_0)$. Отсюда в силу свойства 4° $(\bar{a}+\bar{b})\times|\bar{c}|\bar{c}_0=\bar{a}\times|\bar{c}|\bar{c}_0+\bar{b}\times|\bar{c}|\bar{c}_0$. Заменяя $|\bar{c}|\bar{c}_0$ через $\bar{c}$, окончательно получаем: $(\bar{a}+\bar{b})\times\bar{c}=\bar{a}\times\bar{c}+\bar{b}\times\bar{c}$ ■

Замечание 1. Доказанное свойство дает право при перемножении векторных многочленов выполнять действие почленно, а свойство 4° позволяет объединить

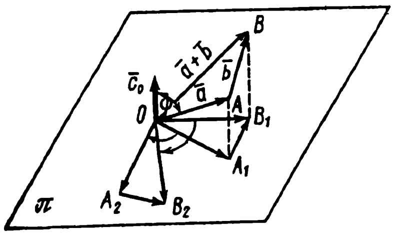


Рис. 23

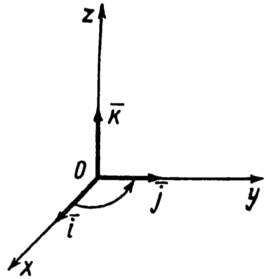


Рис. 24

числовые коэффициенты векторных сомножителей. Например:

$$(2\bar{a}+3\bar{b})\times(4\bar{c}+5\bar{d})=$$
$$=(2\bar{a}+3\bar{b})\times4\bar{c}+(2\bar{a}+3\bar{b})\times5\bar{d}=$$
$$=2\bar{a}\times4\bar{c}+3\bar{b}\times4\bar{c}+2\bar{a}\times5\bar{d}+3\bar{b}\times5\bar{d}=$$
$$=8(\bar{a}\times\bar{c})+12(\bar{b}\times\bar{c})+10(\bar{a}\times\bar{d})+15(\bar{b}\times\bar{d}).$$

Следует, однако, твердо понимать, что порядок сомножителей векторного произведения является существенным, так как по свойству 3° при перестановке сомножителей векторного произведения нужно ставить перед ним знак минус. Например:

$$(\bar{a}+\bar{b}+\bar{c})\times(2\bar{b}+3\bar{c})=2(\bar{a}\times\bar{b})+$$
$$+2(\bar{b}\times\bar{b})+2(\bar{c}\times\bar{b})+3(\bar{a}\times\bar{c})+3(\bar{b}\times\bar{c})+3(\bar{c}\times\bar{c})=$$
$$=2(\bar{a}\times\bar{b})-2(\bar{b}\times\bar{c})+3(\bar{a}\times\bar{c})+3(\bar{b}\times\bar{c})=$$
$$=2(\bar{a}\times\bar{b})+\bar{b}\times\bar{c}+3(\bar{a}\times\bar{c}).$$

Замечание 2. Из свойств 1°, 2° и 3° векторного произведения для базисных векторов $\bar{i}$, $\bar{j}$, $\bar{k}$ (рис. 24) непосредственно получаем следующие равенства:

$$\begin{aligned}&\bar{i}\times\bar{i}=0,\ \bar{i}\times\bar{j}=\bar{k},\ \bar{i}\times\bar{k}=-\bar{j},\\ &\bar{j}\times\bar{i}=-\bar{k},\ \bar{j}\times\bar{j}=0,\ \bar{j}\times\bar{k}=\bar{i},\\ &\bar{k}\times\bar{i}=\bar{j},\ \bar{k}\times\bar{j}=-\bar{i},\ \bar{k}\times\bar{k}=0.\end{aligned} \tag{2}$$

Выражение векторного произведения через координаты векторов

Теорема 9.7. *Если векторы $\bar{a}$ и $\bar{b}$ заданы своими координатами: $\bar{a}=\{X_1;\ Y_1;\ Z_1\}$, $\bar{b}=\{X_2;\ Y_2;\ Z_2\}$, то векторное произведение вектора $\bar{a}$ на вектор $\bar{b}$ определяется формулой*

$$\bar{a}\times\bar{b}=\{(Y_1Z_2-Y_2Z_1);\ \ (X_2Z_1-X_1Z_2);\ \ (X_1Y_2-X_2Y_1)\},$$

или, забегая немного вперед, эту формулу с помощью определителей 2-го порядка можно записать в виде *

$$\bar{a}\times\bar{b}=\left\{\begin{vmatrix}Y_1 & Z_1\\ Y_2 & Z_2\end{vmatrix};\ \ \begin{vmatrix}Z_1 & X_1\\ Z_2 & X_2\end{vmatrix};\ \ \begin{vmatrix}X_1 & Y_1\\ X_2 & Y_2\end{vmatrix}\right\}.$$

Доказательство. Разложим векторы $\bar{a}$ и $\bar{b}$ по базису $\bar{i}$, $\bar{j}$, $\bar{k}$:

$$\bar{a}=X_1\bar{i}+Y_1\bar{j}+Z_1\bar{k},\ \ \bar{b}=X_2\bar{i}+Y_2\bar{j}+Z_2\bar{k}. \tag{3}$$

На основании алгебраических свойств векторного произведения мы можем правые части равенств (3) перемножить почленно:

$$\begin{aligned}\bar{a}\times\bar{b}=X_1X_2(\bar{i}\times\bar{i})+X_1Y_2(\bar{i}\times\bar{j})+X_1Z_2(\bar{i}\times\bar{k})+\\ +Y_1X_2(\bar{j}\times\bar{i})+Y_1Y_2(\bar{j}\times\bar{j})+Y_1Z_2(\bar{j}\times\bar{k})+\\ +Z_1X_2(\bar{k}\times\bar{i})+Z_1Y_2(\bar{k}\times\bar{j})+Z_1Z_2(\bar{k}\times\bar{k}).\end{aligned}$$

Отсюда, пользуясь равенствами (2), находим

$$\bar{a}\times\bar{b}=(Y_1Z_2-Y_2Z_1)\bar{i}+(X_2Z_1-X_1Z_2)\bar{j}+(X_1Y_2-X_2Y_1)\bar{k},$$

или

$$\bar{a}\times\bar{b}=\begin{vmatrix}Y_1 & Z_1\\ Y_2 & Z_2\end{vmatrix}\bar{i}+\begin{vmatrix}Z_1 & X_1\\ Z_2 & X_2\end{vmatrix}\bar{j}+\begin{vmatrix}X_1 & Y_1\\ X_2 & Y_2\end{vmatrix}\bar{k}.$$

---

* См. сноску на с. 339.

Мы получили разложение вектора $\bar{a}\times\bar{b}$ по базису $\bar{i}, \bar{j}, \bar{k}$; коэффициенты этого разложения представляют собой координаты вектора $\bar{a}\times\bar{b}$. Таким образом,

$$\bar{a}\times\bar{b}=\{X; Y; Z\}, \tag{4}$$

где

$$X=\begin{vmatrix} Y_1 Z_1 \\ Y_2 Z_2 \end{vmatrix}; \ Y=\begin{vmatrix} Z_1 X_1 \\ Z_2 X_2 \end{vmatrix}; \quad Z=\begin{vmatrix} X_1 Y_1 \\ X_2 Y_2 \end{vmatrix} \blacksquare$$

Пример. Даны векторы $\bar{a}=\{2; 5; 7\}$ и $\bar{b}=\{1; 2; 4\}$. Найти координаты векторного произведения $\bar{a}\times\bar{b}$.

Решение. По формуле (4) находим

$$X=\begin{vmatrix} 5 & 7 \\ 2 & 4 \end{vmatrix}=6; \quad Y=\begin{vmatrix} 7 & 2 \\ 4 & 1 \end{vmatrix}=-1; \ Z=\begin{vmatrix} 2 & 5 \\ 1 & 2 \end{vmatrix}=-1;$$

искомые координаты

$$\bar{a}\times\bar{b}=\{6; -1; -1\}.$$

§ 8. Смешанное произведение трех векторов

Определение. *Смешанным произведением трех векторов $\bar{a}$, $\bar{b}$, $\bar{c}$ называется число, равное скалярному произведению вектора $\bar{a}$ на векторное произведение векторов $\bar{b}$ и $\bar{c}$, т. е.*

$$\bar{a}\cdot(\bar{b}\times\bar{c}).$$

Геометрический смысл смешанного произведения

Следующая теорема выражает геометрический смысл смешанного произведения.

Теорема 9.8. *Смешанное произведение $\bar{a}\cdot(\bar{b}\times\bar{c})$ равно объему V параллелепипеда, построенного на векторах $\bar{a}$, $\bar{b}$, $\bar{c}$, взятому со знаком $+$, если тройка $\bar{a}$, $\bar{b}$, $\bar{c}$ правая, и со знаком $-$, если тройка левая. Если же $\bar{a}$, $\bar{b}$, $\bar{c}$ компланарны, то $\bar{a}\cdot(\bar{b}\times\bar{c})=0$. Другими словами:*

$$\bar{a}\cdot(\bar{b}\times\bar{c})=\begin{cases} V, \textit{ если } \bar{a}, \bar{b}, \bar{c} \textit{ — правая}, \\ -V, \textit{ если } \bar{a}, \bar{b}, \bar{c} \textit{ — левая}, \\ 0, \textit{ если } \bar{a}, \bar{b}, \bar{c} \textit{ — компланарны}. \end{cases}$$

Доказательство. Пусть даны некомпланарные векторы $\bar{a}$, $\bar{b}$, $\bar{c}$, образующие правую тройку. Обозначим через V объем параллелепипеда, построенного на этих векторах, через S — площадь параллелограмма, построенного на векторах $\bar{b}$ и $\bar{c}$, а через h — высоту параллелепипеда (рис. 25). Тогда по определению скалярного произведения

$$\bar{a}\cdot(\bar{b}\times\bar{c}) = |\bar{a}|\,|\bar{b}\times\bar{c}|\cos\theta = |\bar{a}|\,|\bar{b}|\,|\bar{c}|\cos\theta\sin\varphi =$$
$$= |\bar{c}|\,|\bar{b}|\sin\varphi\,|\bar{a}|\cos\theta = S\cdot h = V,$$

где φ — угол между векторами $\bar{b}$ и $\bar{c}$, а θ — угол между векторами $\bar{a}$ и $\bar{b}\times\bar{c}$. Таким образом, $\bar{a}\cdot(\bar{b}\times\bar{c}) = V$.

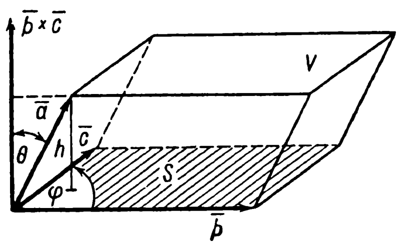


Рис. 25

Знак смешанного произведения совпадает со знаком $\cos\theta$. Он будет положителен

$$\left(\theta < \frac{\pi}{2},\ \cos\theta > 0 \text{ и } |\bar{a}|\cos\theta = +h\right),$$

когда вектор $\bar{b}\times\bar{c}$ направлен в ту же сторону от плоскости векторов $\bar{b}$ и $\bar{c}$, что и вектор $\bar{a}$, т. е. когда тройка $\bar{a}$, $\bar{b}$, $\bar{c}$ правая; и отрицателен

$$\left(\theta > \frac{\pi}{2},\ \cos\theta < 0 \text{ и } |\bar{a}|\cos\theta = -h\right),$$

когда тройка $\bar{a}$, $\bar{b}$, $\bar{c}$ левая, т. е. когда векторы $\bar{b}\times\bar{c}$ и $\bar{a}$ направлены в разные стороны от плоскости векторов $\bar{b}$ и $\bar{c}$. Первое утверждение теоремы доказано.

Докажем второе утверждение. Пусть векторы $\bar{a}$, $\bar{b}$, $\bar{c}$ компланарны. Тогда либо вектор $\bar{b}\times\bar{c}=0$ (векторы $\bar{b}$ и $\bar{c}$ коллинеарны), либо вектор $\bar{b}\times\bar{c}\perp\bar{a}$, и, следовательно, скалярное произведение $\bar{a}\cdot(\bar{b}\times\bar{c})=0$. Верно и обратное, если $\bar{a}\cdot(\bar{b}\times\bar{c})=0$, то векторы $\bar{a}$, $\bar{b}$, $\bar{c}$ компланарны. Действительно, если бы векторы $\bar{a}$, $\bar{b}$, $\bar{c}$ были некомпланарны, то по доказанному первому утверждению теоремы смешанное произведение $\bar{a}\cdot(\bar{b}\times\bar{c})=\pm V$, т. е. $\bar{a}\cdot(\bar{b}\times\bar{c})\neq 0$, что противоречит нашему допущению ■

Следствие. Из теоремы легко выводится следующее тождество:

$$\bar{a}\cdot(\bar{b}\times\bar{c})=(\bar{a}\times\bar{b})\cdot\bar{c}, \tag{1}$$

т. е. знаки $\cdot$ и $\times$ в смешанном произведении можно менять местами.

Действительно, в силу свойства 1° скалярного произведения

$$(\bar{a}\times\bar{b})\cdot\bar{c}=\bar{c}\cdot(\bar{a}\times\bar{b}). \tag{2}$$

Далее, по теореме 9.8 мы имеем

$$\bar{a}\cdot(\bar{b}\times\bar{c})=\pm V,\ \bar{c}\cdot(\bar{a}\times\bar{b})=\pm V. \tag{3}$$

Так как тройки $\bar{a}$, $\bar{b}$, $\bar{c}$ и $\bar{c}$, $\bar{a}$, $\bar{b}$ — одной ориентации, т. е. либо обе правые, либо обе левые, то на основании теоремы 9.8 в правых частях равенств (3) нужно брать один и тот же знак, откуда получаем

$$\bar{a}\cdot(\bar{b}\times\bar{c})=\bar{c}\cdot(\bar{a}\times\bar{b}),$$

и вследствие равенства (2)

$$\bar{a}\cdot(\bar{b}\times\bar{c})=(\bar{a}\times\bar{b})\cdot\bar{c},$$

т. е. получили тождество (1).

В силу тождества (1) смешанные произведения $\bar{a}\cdot(\bar{b}\times\bar{c})$ и $(\bar{a}\times\bar{b})\cdot\bar{c}$ обозначаются более простым символом: $\bar{a}\bar{b}\bar{c}$.

Выражение смешанного произведения через координаты векторов

Теорема 9.9. *Если векторы $\bar{a}$, $\bar{b}$, $\bar{c}$ заданы своими координатами:*

$$\bar{a}=\{X_1;\ Y_1;\ Z_1\},\ \bar{b}=\{X_2;\ Y_2;\ Z_2\},\ \bar{c}=\{X_3;\ Y_3;\ Z_3\},$$

то смешанное произведение $\bar{a}\bar{b}\bar{c}$ определяется формулой

$$\bar{a}\bar{b}\bar{c}=X_1\begin{vmatrix}Y_2 & Z_2\\ Y_3 & Z_3\end{vmatrix}+Y_1\begin{vmatrix}Z_2 & X_2\\ Z_3 & X_3\end{vmatrix}+Z_1\begin{vmatrix}X_2 & Y_2\\ X_3 & Y_3\end{vmatrix},$$

или, что одно и то же, но более компактно (см. § 2, гл. X)

$$\bar{a}\bar{b}\bar{c} = \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}.$$

Доказательство. Имеем $\bar{a}\bar{b}\bar{c}=\bar{a}\cdot(\bar{b}\times\bar{c})$. По теореме 9.7

$$\bar{b}\times\bar{c} = \left\{\begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix};\quad \begin{vmatrix} Z_2 & X_2 \\ Z_3 & X_3 \end{vmatrix};\quad \begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_2 \end{vmatrix}\right\}.$$

Умножая скалярно вектор $\bar{a}=\{X_1; Y_1; Z_1\}$ на вектор $\bar{b}\times\bar{c}$ и пользуясь теоремой 9.6, получаем

$$\bar{a}\bar{b}\bar{c} = X_1\begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix} + Y_1\begin{vmatrix} Z_2 & X_2 \\ Z_3 & X_3 \end{vmatrix} + Z_1\begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{vmatrix} =$$

$$= \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix} \blacksquare$$

Пример. В пространстве даны четыре точки: $A(1; 1; 1)$, $B(4; 4; 4)$, $C(3; 5; 5)$, $D(2; 4; 7)$. Найти объем тетраэдра $ABCD$.

Решение. Как известно из элементарной геометрии, объем $V_т$ тетраэдра $ABCD$ равен одной шестой объема параллелепипеда, построенного на вектора $\overline{AB}$, $\overline{AC}$ и $\overline{AD}$; отсюда и из теоремы 9.8 заключаем, что $V_т$ равняется одной шестой абсолютной величины смешанного произведения $\overline{AB}\cdot\overline{AC}\cdot\overline{AD}$. Остается подсчитать это смешанное произведение. Прежде всего найдем координаты векторов $\overline{AB}$, $\overline{AC}$ и $\overline{AD}$. По теореме 9.2 имеем: $\overline{AB}=\{3; 3; 3\}$, $\overline{AC}=\{2; 4; 4\}$, $\overline{AD}=\{1; 3; 6\}$. Применим теперь теорему 9.9:

$$\overline{AB}\cdot\overline{AC}\cdot\overline{AD} = \begin{vmatrix} 3 & 3 & 3 \\ 2 & 4 & 4 \\ 1 & 3 & 6 \end{vmatrix} = 3\begin{vmatrix} 4 & 4 \\ 3 & 6 \end{vmatrix} - 3\begin{vmatrix} 4 & 2 \\ 6 & 1 \end{vmatrix} + 3\begin{vmatrix} 2 & 4 \\ 1 & 3 \end{vmatrix} =$$

$$= 3\cdot 12 - 3\cdot 8 + 3\cdot 2 = 18.$$

Отсюда

$$V_T = \frac{1}{6}V_п = \frac{1}{6}\cdot 18 = 3.$$

§ 9. Уравнения поверхности и линии

Пусть заданы некоторая прямоугольная система координат $Oxyz$, произвольная поверхность S и равенство вида

$$F(x; y; z) = 0. \qquad (1)$$

(Равенство (1) называется *уравнением с тремя переменными*, если оно выполняется не для всех троек чисел x, y, z).

Будем говорить, что равенство (1) является уравнением поверхности S (рис. 26), если ему удовлетворяют координаты любой точки $M(x; y; z) \in S$ и не удовлетворяют координаты никакой точки, не лежащей на поверхности S. С точки зрения этого определения поверхность есть множество точек, координаты которых удовлетворяют уравнению (1).

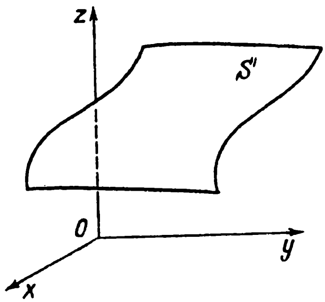


Рис. 26

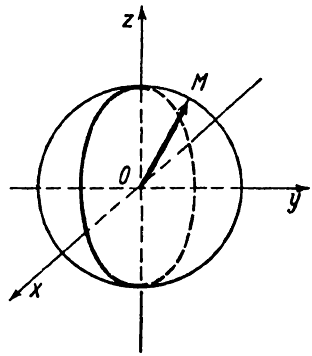


Рис. 27

В пространственной аналитической геометрии понятие уравнения поверхности является важнейшим понятием.

Пример. В прямоугольных координатах уравнение

$$x^2+y^2+z^2=R^2 \quad (F(x; y; z)=0:\ x^2+y^2+z^2-R^2=0)$$

определяет поверхность, являющуюся сферой с центром в точке $O(0; 0; 0)$ и радиусом, равным R (рис. 27).

В самом деле, если $M(x; y; z)$ — произвольная точка, то по формуле (7) (см. п. 5, § 2)

$$|\overline{OM}| = \sqrt{x^2 + y^2 + z^2}.$$

Отсюда ясно, что заданному уравнению удовлетворяют координаты тех и только тех точек, которые удалены от точки O на расстоянии R. Следовательно, множество точек, координаты которых удовлетворяют этому уравнению, есть сфера с центром в начале координат и радиусом R.

В пространственной аналитической геометрии каждая линия рассматривается как пересечение двух поверхностей, т. е. как множество точек, находящееся одновременно на двух поверхностях, и соответственно этому определяется заданием двух уравнений. Таким образом, два уравнения

$$\begin{cases} F_1(x;\ y;\ z) = 0, \\ F_2(x;\ y;\ z) = 0 \end{cases} \tag{2}$$

называются *уравнениями линии* L, если им удовлетворяют координаты любой точки, лежащей на L, и не удовлетворяют координаты ни одной точки, не лежащей на линии L.

Например, два уравнения

$$\begin{cases} (x-1)^2 + (y-2)^2 + (z-3)^2 = 14, \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

совместно определяют окружность (как пересечение двух сфер).

§ 10. Уравнение цилиндрической поверхности

Пусть в плоскости Oxy лежит некоторая линия L (рис. 28). Проведем через каждую точку L прямую, параллельную оси Oz. Множество этих прямых образует некоторую поверхность S, которая называется *цилиндрической*. Упомянутые прямые называ-

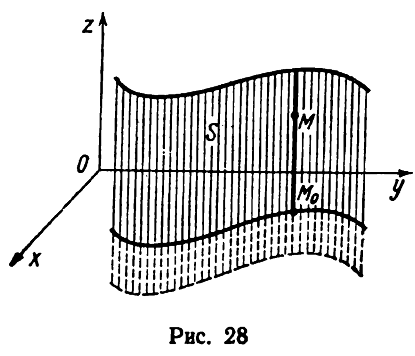


Рис. 28

ются *образующими* поверхности S, а линия L — ее *направляющей.*

Совершенно аналогично определяется цилиндрическая поверхность с образующими, параллельными осям Ox и Oy.

Для определенности будем рассматривать цилиндрическую поверхность S с образующими, параллельными оси Oz, и докажем, что она определяется уравнением вида

$$F(x;\ y)=0. \tag{1}$$

Действительно, пусть уравнение (1) есть уравнение направляющей L. Возьмем на S любую точку $M(x; y; z)$. Эта точка лежит на какой-то образующей. Если M_0 есть пересечение этой образующей с плоскостью Oxy, то точка $M_0 \in L$ и ее координаты x и y удовлетворяют уравнению (1). Но тогда числа x, y, z, где z — любое число, также удовлетворяют этому уравнению, поскольку $F(x; y)$ от z не зависит. Итак, координаты x, y, z произвольной точки $M \in S$ удовлетворяют уравнению (1). Очевидно, если $M(x; y; z) \notin S$, то $M_0(x; y) \notin L$, т. е. координаты x и y не удовлетворяют уравнению (1). Это доказывает, что (1) есть уравнение S.

Таким образом, мы видим, что уравнение цилиндрической поверхности не содержит координаты z и совпадает с уравнением направляющей. Например, если направляющей служит эллипс

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \tag{2}$$

то соответствующая цилиндрическая поверхность называется *эллиптическим цилиндром*, и (2) является ее уравнением.

Заметим, что на плоскости Oxy уравнение $F(x; y)=0$ определяет линию L, но эта же линия в пространственной системе координат $Oxyz$ задается двумя уравнениями:

$$\begin{cases} F(x;\ y)=0, \\ z=0. \end{cases}$$

Так, например, в пространственной системе координат $Oxyz$ уравнение $F(x; y)=0$: $x^2+y^2-R^2=0$ определяет цилиндрическую поверхность круговой цилиндр (рис. 29), а направляющая этого цилиндра L (окружность), лежащая в плоскости Oxy, определяется двумя уравнениями:

$$L\colon \begin{cases} F(x, y) = 0, \\ z = 0 \end{cases} \quad \text{или} \quad L\colon \begin{cases} x^2 + y^2 = R^2, \\ z = 0. \end{cases}$$

§ 11. Уравнения плоскости

Здесь мы покажем, что поверхности первого порядка суть плоскости и только плоскости, и познакомимся с двумя видами уравнений плоскости.

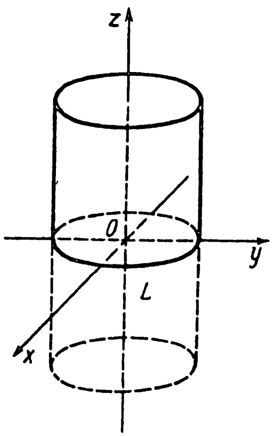


Рис. 29

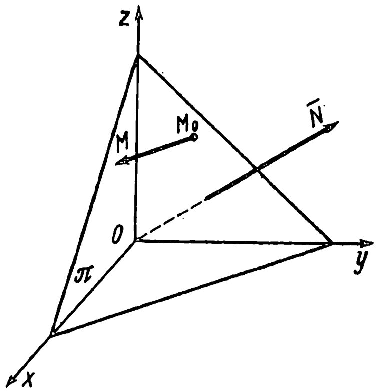


Рис. 30

1. Общее уравнение плоскости. Пусть заданы: некоторая прямоугольная система координат $Oxyz$, произвольная плоскость π, точка $M_0(x_0; y_0; z_0) \in \pi$, вектор $\overline{N} = \{A; B; C\}$, перпендикулярный к плоскости π, где A, B, C — его проекции на оси координат (рис. 30).

Возьмем произвольную точку $M(x; y; z)$. Ясно, что точка M будет лежать на плоскости π тогда и только тогда, когда векторы $\overline{M_0M}$ и $\overline{N}$ взаимно перпендикулярны. Так как проекции вектора $\overline{M_0M}$ на оси координат суть $x - x_0$, $y - y_0$, $z - z_0$, то в силу условия перпендикулярности двух векторов (см. формулу (3), § 6) получаем

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0. \qquad (1)$$

Это и есть искомое уравнение плоскости π, так как ему удовлетворяют координаты x, y, z точки M в том и только в том случае, когда M лежит на плоскости π.

Раскрывая скобки, представим уравнение (1) в виде

$$Ax+By+Cz+(-Ax_0-By_0-Cz_0)=0.$$

Далее, обозначая число $-Ax_0-By_0-Cz_0$ через D, получим

$$Ax+By+Cz+D=0. \qquad (2)$$

Уравнение вида (2) называется *общим уравнением плоскости*, которая действительно является поверхностью первого порядка, так как определяется уравнением первой степени.

Верно и обратное: всякое уравнение первой степени вида (2) определяет в некоторой прямоугольной системе координат плоскость.

Действительно, пусть заданы некоторая прямоугольная система координат $Oxyz$ и уравнение $Ax+By+Cz+$ $+D=0$ с произвольными коэффициентами A, B, C, D, причем из коэффициентов A, B, C хотя бы один отличен от нуля. Данное уравнение заведомо имеет хотя бы одно решение x_0, y_0, z_0 (если, например, $C\neq 0$, то, взяв произвольные x_0 и y_0, получим из уравнения значение $z_0=-\frac{A}{C}x_0-\frac{B}{C}y_0$). Таким образом, существует хотя бы одна точка $M_0(x_0; y_0; z_0)$, координаты которой удовлетворяют уравнению $Ax_0+By_0+Cz_0+D=0$. Вычитая его из уравнения $Ax+By+Cz+D=0$, получим уравнение

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0,$$

эквивалентное данному. Докажем, что последнее уравнение (а стало быть, и уравнение $Ax+By+Cz+D=0$) определяет произвольную плоскость π, проходящую через точку $M_0(x_0; y_0; z_0)$ и перпендикулярную вектору $\bar{N}=\{A; B; C\}$.

В самом деле, если точка $M(x; y; z)\in\pi$, то ее координаты удовлетворяют уравнению, так как векторы $\bar{N}=$

$=\{A; B; C\}$ и $\overline{M_0M}=\{x-x_0; y-y_0; z-z_0\}$ перпендикулярны и их скалярное произведение $A(x-x_0)+B(y-y_0)+$ $+C(z-z_0)$ равно нулю. Если же точка $M(x; y; z)\notin\pi$, то ее координаты не удовлетворяют уравнению, так как векторы $\bar{N}$ и $\overline{M_0M}$ не перпендикулярны и поэтому их скалярное произведение не равно нулю. Что и требовалось.

Вектор $\bar{N}=\{A; B; C\}$, перпендикулярный к некоторой плоскости, называется *нормальным вектором* этой плоскости.

Пример. Составить уравнение плоскости, проходящей через точку $M_0(1; 1; 1)$ перпендикулярно к вектору $\bar{N}=\{2; 2; 3\}$.

Решение. По формуле (1) искомое уравнение есть

$$2(x-1)+2(y-1)+3(z-1)=0,$$

или

$$2x+2y+3z-7=0.$$

В заключение этого пункта докажем следующую теорему.

Теорема 9.10. *Если два уравнения* A_1x+B_1y+ $+C_1z+D_1=0$ *и* $A_2x+B_2y+C_2z+D_2=0$ *определяют одну и ту же плоскость, то коэффициенты их пропорциональны.*

Доказательство. Действительно, векторы $\bar{N}_1=$ $=\{A_1; B_1; C_1\}$ и $\bar{N}_2=\{A_2; B_2; C_2\}$ перпендикулярны к одной плоскости, следовательно, коллинеарны друг другу. Но тогда числа A_1, B_1, C_1 пропорциональны числам A_2, B_2, C_2 (см. формулу (5), § 4), т. е.

$$\frac{A_2}{A_1}=\frac{B_2}{B_1}=\frac{C_2}{C_1}=\mu,$$

или $A_2=\mu A_1$, $B_2=\mu B_1$, $C_2=\mu C_1$ (μ — множитель пропорциональности). Умножая первое из заданных уравнений на μ и вычитая из второго, получим $D_2-\mu D_1=0$, т. е. $D_2=\mu D_1$, и, следовательно,

$$\frac{A_2}{A_1}=\frac{B_2}{B_1}=\frac{C_2}{C_1}=\frac{D_2}{D_1}=\mu \blacksquare$$

2. Угловые соотношения. 1°. *Угол между двумя плоскостями.* Рассмотрим две плоскости:

$$\pi_1:\ A_1x+B_1y+C_1z+D_1=0,$$
$$\pi_2:\ A_2x+B_2y+C_2z+D_2=0.$$

При любом расположении плоскостей π_1 и π_2 в пространстве один из углов φ между ними равен углу между их нормальными векторами $\bar{N}_1=\{A_1;\ B_1;\ C_1\}$ и $\bar{N}_2=\{A_2;\ B_2;\ C_2\}$ и вычисляется по следующей формуле:

$$\cos\varphi=\frac{\bar{N}_1\cdot\bar{N}_2}{|\bar{N}_1||\bar{N}_2|}=\frac{A_1A_2+B_1B_2+C_1C_2}{\sqrt{A_1^2+B_1^2+C_1^2}\sqrt{A_2^2+B_2^2+C_2^2}}. \tag{3}$$

Второй угол будет дополнять его до 180°.

2°. *Условие параллельности плоскостей.* Если плоскости π_1 и π_2 параллельны, то будут параллельны и их нормальные векторы $\bar{N}_1$ и $\bar{N}_2$, и наоборот. Но тогда, как известно, имеем

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}. \tag{4}$$

Условие (4) является условием параллельности плоскостей π_1 и π_2.

3°. *Условие перпендикулярности плоскостей.* Если плоскости π_1 и π_2 взаимно перпендикулярны, то их нормальные векторы $\bar{N}_1$ и $\bar{N}_2$ тоже перпендикулярны друг другу $\left(\varphi=\frac{\pi}{2}\right)$, и наоборот. Поэтому из формулы (3) непосредственно получаем условие перпендикулярности плоскостей π_1 и π_2:

$$A_1A_2+B_1B_2+C_1C_2=0.$$

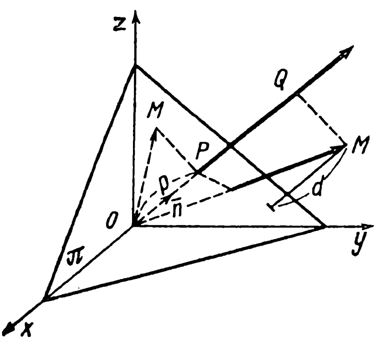


Рис. 31

3. Нормальное уравнение плоскости. *Расстояние от точки до плоскости.* Пусть заданы некоторая прямоугольная система координат $Oxyz$ и произвольная плоскость π (рис. 31). Проведем через начало координат прямую, перпендикулярную к плоскости π, и назовем ее нормалью. Через P обозначим точку, в которой нормаль пересечет плоскость π. На нормали введем положительное направление от точки O к точке P. Обозначим через α, β, γ углы, которые составляет направленная нормаль с осями координат, через p — длину отрезка OP.

Выведем уравнение данной плоскости π, считая известными числа $\cos\alpha$, $\cos\beta$, $\cos\gamma$ и p. С этой целью введем единичный вектор $\bar{n}$ на нормали, направление которого совпадает с положительным направлением нормали.

Так как $\bar{n}$ — единичный вектор, то

$$\bar{n}=\{\cos\alpha;\ \cos\beta;\ \cos\gamma\}. \qquad (5)$$

Пусть $M(x; y; z)$ — произвольная точка на плоскости π. Проекция вектора OM на нормаль равна p, т. е.

$$\text{Пр.}_{\bar{n}}\overline{OM} = p. \qquad (6)$$

Заметим теперь, что

$$\text{Пр.}_{\bar{n}}\overline{OM} = \bar{n}\cdot\overline{OM} \text{ и } \overline{OM} = \{x; y; z\}.$$

По теореме 9.6 с учетом равенства (5) имеем

$$\text{Пр.}_{\bar{n}}\overline{OM} = \bar{n}\cdot\overline{OM} = x\cos\alpha + y\cos\beta + z\cos\gamma. \qquad (7)$$

Из равенств (6) и (7) окончательно получаем

$$x\cos\alpha+y\cos\beta+z\cos\gamma-p=0. \qquad (8)$$

Это и есть искомое уравнение данной плоскости. Уравнение плоскости в виде (8) называется *нормальным*.

Теорема 9.11. *Если точка M^* имеет координаты x^*, y^*, z^*, а плоскость задана нормальным уравнением*

$$x\cos\alpha+y\cos\beta+z\cos\gamma-p=0,$$

то расстояние d от точки M^* до этой плоскости определяется по формуле

$$d=|x^*\cos\alpha+y^*\cos\beta+z^*\cos\gamma-p|.$$

Доказательство. Спроектируем точку M^* на нормаль, определяемую вектором $\bar{n}$. Пусть Q — ее проекция (см. рис. 31), тогда

$$d=|PQ|=|OQ-OP|.$$

Но $OQ = \text{Пр.}_{\bar{n}}\overline{OM^*}$, $OP = p$, следовательно,

$$d = |\text{Пр.}_{\bar{n}}\overline{OM^*} - p|. \qquad (9)$$

Вектор $\overline{OM^*} = \{x^*; y^*; z^*\}$, а $\text{Пр.}_{\bar{n}}\overline{OM^*} = \bar{n}\cdot\overline{OM^*}$. По теореме 9.5 с учетом равенства (5) находим

$$\text{Пр.}_{\bar{n}}\overline{OM^*} = \bar{n}\cdot\overline{OM^*} = x^*\cos\alpha + y^*\cos\beta + z^*\cos\gamma. \qquad (10)$$

Из равенств (9) и (10) окончательно получаем

$$d = |x^*\cos\alpha + y^*\cos\beta + z^*\cos\gamma - p| \ \blacksquare$$

Теперь покажем, как привести общее уравнение плоскости к нормальному виду. Пусть

$$Ax + By + Cz + D = 0 \qquad (11)$$

— общее уравнение некоторой плоскости, а

$$x\cos\alpha + y\cos\beta + z\cos\gamma - p = 0 \qquad (12)$$

— ее нормальное уравнение. Так как уравнения (11) и (12) определяют одну и ту же плоскость, то по теореме 9.10 коэффициенты этих уравнений пропорциональны. Это означает, что, умножив все члены (11) на некоторый множитель μ, получим уравнение

$$\mu Ax + \mu By + \mu Cz + \mu D = 0,$$

совпадающее с уравнением (12), т. е. будем иметь

$$\mu A = \cos\alpha, \ \mu B = \cos\beta, \ \mu C = \cos\gamma, \ \mu D = -p. \qquad (13)$$

Чтобы найти множитель μ, возведем первые три из равенств (13) в квадрат и сложим, тогда получим

$$\mu^2(A^2 + B^2 + C^2) = \cos^2\alpha + \cos^2\beta + \cos^2\gamma.$$

Но согласно (6) § 2 правая часть последнего равенства равна единице. Следовательно,

$$\mu = \pm\frac{1}{\sqrt{A^2 + B^2 + C^2}}.$$

Число μ, с помощью которого общее уравнение плоскости преобразуется в нормальное, называется *нормирующим множителем* этого уравнения. Знак μ определяется равенством $\mu D = -p$, т. е. имеет знак, противоположный свободному члену общего уравнения (11).

Замечание. Если в уравнении (11) $D=0$, то знак нормирующего множителя выбирается по желанию.

Пример. Даны плоскость $x+2y+2z-8=0$ и точка $M^*(1;\ 1;\ 1)$. Найти расстояние d от точки M^* до данной плоскости.

Решение. Чтобы применить теорему 9.10, надо прежде всего привести данное уравнение к нормальному виду. С этой целью находим нормирующий множитель:

$$\mu=\frac{1}{\sqrt{1^2+2^2+2^2}}=\frac{1}{3}.$$

Умножая данное уравнение на μ, получим искомое нормальное уравнение плоскости:

$$\frac{1}{3}x+\frac{2}{3}y+\frac{2}{3}z-\frac{8}{3}=0.$$

Подставляя в левую часть этого уравнения координаты точки M^*, окончательно получаем

$$d=\left|\frac{1}{3}+\frac{2}{3}+\frac{2}{3}-\frac{8}{3}\right|=\left|-\frac{3}{3}\right|=1.$$

§ 12. Уравнения прямой

Мы уже знаем, что в пространственной аналитической геометрии каждая линия рассматривается как пересечение двух поверхностей и определяется заданием двух уравнений. В частности, каждую прямую линию мы будем рассматривать как пересечение двух плоскостей и соответственно этому определять заданием двух уравнений первой степени.

Пусть задана некоторая прямоугольная система координат $Oxyz$ и произвольная прямая L. Обозначим через π_1 и π_2 какие-нибудь две различные плоскости, пересекающиеся по прямой L:

$$\left.\begin{aligned}\pi_1&: A_1x+B_1y+C_1z+D_1=0,\\ \pi_2&: A_2x+B_2y+C_2z+D_2=0.\end{aligned}\right\}\qquad(1)$$

Два уравнения вида (1) совместно определяют прямую L в том и только в том случае, когда плоскости π_1 и π_2 не параллельны и не совпадают друг с другом, т. е. нормальные векторы этих плоскостей $\bar{N}_1=$

$=\{A_1; B_1; C_1\}$ и $\bar{N}_2=\{A_2; B_2; C_2\}$ не коллинеарны (коэффициенты A_1, B_1, C_1 не пропорциональны коэффициентам A_2, B_2, C_2).

Уравнения (1) называются *общими уравнениями прямой*.

1. Канонические уравнения прямой. Для решения задач уравнения (1) не всегда удобны, поэтому используется специальный вид уравнений прямой, к выводу которых мы сейчас и перейдем.

Пусть дана какая-нибудь прямая L и ненулевой вектор $\bar{a}$, лежащий на данной прямой или параллельный ей (рис. 32). Вектор $\bar{a}$ называется *направляющим вектором* данной прямой, так как, будучи заданным, определяет направление прямой. Обозначим его координаты буквами l, m, n:

$$\bar{a}=\{l; m; n\}.$$

Выведем уравнения прямой, проходящей через данную точку $M_0(x_0; y_0; z_0)$ и имеющей данный направляющий вектор $\bar{a}=\{l; m; n\}$ (см. рис. 32).

Возьмем произвольную точку $M(x; y; z)\in L$. Тогда вектор

$$\overline{M_0M}=\{x-x_0;\ y-y_0;\ z-z_0\}$$

будет коллинеарным направляющему вектору $\bar{a}=\{l; m; n\}$. Следовательно, координаты вектора $\overline{M_0M}$ будут пропорциональны координатам вектора $\bar{a}$:

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}. \qquad (2)$$

Равенства (2) называются *каноническими уравнениями прямой*, проходящей через данную точку $M_0(x_0; y_0; z_0)$ в направлении вектора $\bar{a}=\{l; m; n\}$, при этом координаты l, m, n называются *направляющими параметрами* этой прямой.

Теперь покажем, как составить канонические уравнения (2), если прямая L задана уравнением (1). Для этого необходимо:

1) Найти какую-нибудь точку $M_0(x_0; y_0; z_0)\in L$; для этого следует задать численное значение одной из неизвестных координат x_0, y_0, z_0 и подставить его вместо соответствующей переменной в уравнения (1); после

этого две другие координаты определяются из уравнений (1) путем их совместного решения.

2) Найти направляющий вектор $\bar{a}=\{l; m; n\}$. Так как L определена пересечением плоскостей π_1 и π_2, то она перпендикулярна к каждому из нормальных векторов $\bar{N}_1$ и $\bar{N}_2$ (рис. 33). Поэтому в качестве вектора $\bar{a}$ можно взять любой вектор, перпендикулярный векто-

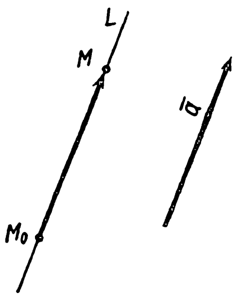


Рис. 32

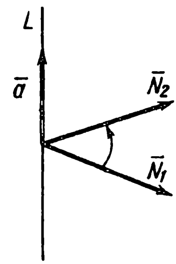


Рис. 33

рам $\bar{N}_1$ и $\bar{N}_2$, например их векторное произведение: $\bar{a}=\bar{N}_1\times\bar{N}_2$. Поскольку координаты векторов $\bar{N}_1$ и $\bar{N}_2$ известны: $\bar{N}_1=\{A_1; B_1; C_1\}$, $\bar{N}_2=\{A_2; B_2; C_2\}$, по теореме 9.7 найдем координаты вектора $\bar{a}$:

$$\bar{a}=\left\{\begin{vmatrix} B_1 C_1 \\ B_2 C_2 \end{vmatrix}; \quad \begin{vmatrix} C_1 A_1 \\ C_2 A_2 \end{vmatrix}; \begin{vmatrix} A_1 B_1 \\ A_2 B_2 \end{vmatrix}\right\}=\{l; m; n\}.$$

Пример. Найти канонические уравнения прямой

$$\begin{cases} 3x+2y+4z-11=0, \\ 2x+y-3z-1=0. \end{cases}$$

Решение. Полагая, например, $x_0=1$, находим из системы

$$\begin{cases} 2y+4z-8=0, \\ y-3z+1=0 \end{cases}$$

$y_0=2$ и $z_0=1$. Таким образом, точка прямой $M_0(1; 2; 1)$ найдена. Теперь найдем направляющий вектор $\bar{a}$. Имеем: $\bar{N}_1=\{3; 2; 4\}$, $\bar{N}_2=\{2; 1; -3\}$, отсюда $\bar{a}=\bar{N}_1\times\bar{N}_2=\{-10; 17; -1\}$, т. е. $l=-10$, $m=17$, $n=-1$. Подставляя найденные значения x_0, y_0, z_0 и l, m, n в равенства (2), получим канонические уравнения данной прямой:

$$\frac{x-1}{-10}=\frac{y-2}{17}=\frac{z-1}{-1}.$$

2. Параметрические уравнения прямой. Иногда прямую полезно задавать не в виде канонических уравнений (2), а несколько иначе. Пусть прямая L задана уравнениями (2). Обозначим через t каждое из равных отношений. Тогда

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}=t,$$

откуда

$$\left.\begin{aligned} x&=x_0+lt,\\ y&=y_0+mt,\\ z&=z_0+nt. \end{aligned}\right\} \qquad (3)$$

Равенства (3) называются *параметрическими уравнениями* прямой L, проходящей через точку $M_0(x_0;\, y_0;\, z_0)$ в направлении вектора $\bar{a}=\{l;\, m;\, n\}$. В (3) t рассматривается как произвольно изменяющийся параметр, x, y, z — как функции от t. При изменении t величины x, y, z меняются так, что точка $M(x;\, y;\, z)$ движется по данной прямой.

Параметрические уравнения удобны в тех случаях, когда требуется найти точку пересечения прямой с плоскостью. В самом деле, пусть даны непараллельные плоскость π и прямая L:

$$\pi:\ Ax+By+Cz+D=0,$$
$$L:\ x=x_0+lt,\ \ y=y_0+mt,\ \ z=z_0+nt.$$

Для определения точки пересечения прямой и плоскости подставим значения x, y, z из уравнения L в уравнение π. После преобразования получим

$$t=-\frac{Ax_0+By_0+Cz_0+D}{Al+Bm+Cn},$$

причем знаменатель не равен нулю, так как плоскость не параллельна прямой. Найдя t и подставляя его в уравнение прямой, получаем искомую точку $M(x;\, y;\, z)$ пересечения прямой L с плоскостью π.

3. Угловые соотношения. 1°. *Угол между прямыми.* Рассмотрим две прямые:

$$L_1: \frac{x-x_1}{l_1} = \frac{y-y_1}{m_1} = \frac{z-z_1}{n_1} \text{ и } L_2: \frac{x-x_2}{l_2} = \frac{y-y_2}{m_2} = \frac{z-z_2}{n_2}.$$

При любом расположении прямых L_1 и L_2 в пространстве один из двух углов между ними равен углу φ между направляющими векторами $\bar{a}_1=\{l_1; m_1; n_1\}$ и $\bar{a}_2=\{l_2; m_2; n_2\}$, а второй угол равен $180°-\varphi$, где φ вычисляется по следующей формуле:

$$\cos\varphi = \frac{l_1l_2+m_1m_2+n_1n_2}{\sqrt{l_1^2+m_1^2+n_1^2}\sqrt{l_2^2+m_2^2+n_2^2}}.$$

2°. *Условие параллельности прямых.* Прямые L_1 и L_2 параллельны в том и только в том случае, когда их направляющие векторы $\bar{a}_1=\{l_1; m_1; n_1\}$ и $\bar{a}_2=\{l_2; m_2; n_2\}$ коллинеарны. Отсюда получаем условие параллельности прямых L_1 и L_2:

$$\frac{l_2}{l_1} = \frac{m_2}{m_1} = \frac{n_2}{n_1}.$$

3°. *Условие перпендикулярности прямых.* Прямые L_1 и L_2 перпендикулярны в том и только в том случае, когда их направляющие векторы $\bar{a}_1=\{l_1; m_1; n_1\}$ и $\bar{a}_2=\{l_2; m_2; n_2\}$ перпендикулярны. Отсюда получаем условие перпендикулярности прямых L_1 и L_2:

$$l_1l_2+m_1m_2+n_1n_2=0.$$

4. Расстояние от точки до прямой. В заключение параграфа рассмотрим задачу: найти расстояние d между данной прямой и данной точкой в пространстве.

Пусть дана прямая $L: \frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$ и точка $M_1(x_1; y_1; z_1)$. Искомое расстояние d есть высота параллелограмма, построенного на векторах $\bar{a}=\{l; m; n\}$ и $\overline{M_0M_1}=\{x_1-x_0; y_1-y_0; z_1-z_0\}$ (рис. 34).

Пусть вектор $\bar{p}$ есть векторное произведение векторов $\overline{M_0M_1}$ и $\bar{a}$: $\bar{p}=\overline{M_0M_1}\times\bar{a}$. Так как $|\bar{p}|$ равен площади параллелограмма, построенного на векторах $\overline{M_0M_1}$ и $\bar{a}$, то

$$d=\frac{|\bar{p}|}{|\bar{a}|},\ |\bar{p}|=$$

$$=\sqrt{\begin{vmatrix} y_1-y_0 & z_1-z_0 \\ m & n \end{vmatrix}^2+\begin{vmatrix} z_1-z_0 & x_1-x_0 \\ n & l \end{vmatrix}^2+\begin{vmatrix} x_1-x_0 & y_1-y_0 \\ l & m \end{vmatrix}^2},$$

$$|\bar{a}|=\sqrt{l^2+m^2+n^2}.$$

Следовательно,

$$d=\frac{\sqrt{\begin{vmatrix} y_1-y_0 & z_1-z_0 \\ m & n \end{vmatrix}^2+\begin{vmatrix} z_1-z_0 & x_1-x_0 \\ n & l \end{vmatrix}^2+\begin{vmatrix} x_1-x_0 & y_1-y_0 \\ l & m \end{vmatrix}^2}}{\sqrt{l^2+m^2+n^2}}.$$

§ 13. Прямая и плоскость

1. Условия параллельности и перпендикулярности. Пусть даны прямая

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}$$

и плоскость

$$Ax+By+Cz+D=0.$$

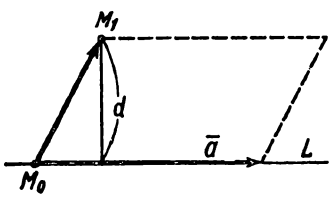


Рис. 34

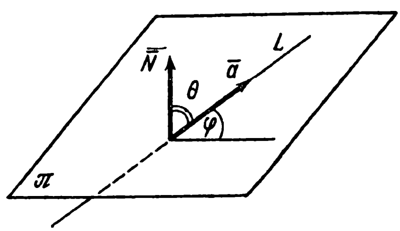


Рис. 35

Прямая параллельна плоскости в том и только в том случае, когда направляющий вектор $\bar{a}=\{l;\ m;\ n\}$ перпендикулярен к нормальному вектору $\bar{N}=\{A;\ B;\ C\}$. Отсюда получаем условие параллельности прямой и плоскости:

$$Al+Bm+Cn=0.$$

Прямая перпендикулярна к плоскости в том и только в том случае, когда направляющий вектор прямой коллинеарен нормальному вектору плоскости. Отсюда получаем условие перпендикулярности прямой и плоскости:

$$\frac{A}{l}=\frac{B}{m}=\frac{C}{n}.$$

2. Угол между прямой и плоскостью. Пусть даны плоскость π

$$Ax+By+Cz+D=0$$

и прямая L

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n},$$

не перпендикулярная плоскости. Под углом φ между прямой L и плоскостью π будем понимать острый угол между L и ее проекцией на π (рис. 35). Из построения видно, что угол φ является дополнительным углом углу θ, образованному векторами $\bar{N}=\{A; B; C\}$ и $\bar{a}=\{l; m; n\}$. Стало быть, $\varphi=90°-\theta$ и $\sin\varphi=\sin(90°-\theta)=\cos\theta$. Но для $\cos\theta$ формула нам известна (см. (4), § 6), следовательно,

$$\sin\varphi=\frac{Al+Bm+Cn}{\sqrt{A^2+B^2+C^2}\cdot\sqrt{l^2+m^2+n^2}}.$$

§ 14. Поверхности второго порядка

Поверхности второго порядка — это поверхности, которые в прямоугольной системе координат определяются алгебраическими уравнениями второй степени. Прежде всего мы рассмотрим эллипсоид и два типа гиперболоидов, являющихся пространственными аналогами эллипсов и гипербол на плоскости.

Геометрическое исследование поверхностей второго порядка будем проводить по заданным уравнениям с помощью так называемого *«метода параллельных сечений»*.

1. Эллипсоид. *Эллипсоидом называется поверхность, которая в некоторой прямоугольной системе координат определяется уравнением*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1. \qquad (1)$$

Уравнение (1) называется *каноническим уравнением* эллипсоида.

Выясним геометрический вид эллипсоида. Для этого рассмотрим сечения данного эллипсоида плоскостями, параллельными плоскости Oxy. Каждая из таких плоскостей определяется уравнением вида $z=h$, где h — любое число, а линия, которая получается в сечении, определяется двумя уравнениями:

$$\begin{cases} \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1-\dfrac{h^2}{c^2}, \\ z=h. \end{cases} \qquad (2)$$

Исследуем уравнения (2) при различных значениях h:

1) Если $|h|>c$ $(c>0)$, то $\frac{x^2}{a^2}+\frac{y^2}{b^2}<0$, и уравнения (2) определяют мнимый эллипс, т. е. точек пересечения плоскости $z=h$ с данным эллипсоидом не существует.

2) Если $h=\pm c$, то $\frac{x^2}{a^2}+\frac{y^2}{b^2}=0$, и линия сечения (2) вырождается в точки $(0; 0; +c)$ и $(0; 0; -c)$, т. е. плоскости $z=\pm c$ касаются эллипсоида.

3) Если $|h|<c$, то уравнения (2) можно представить в виде

$$\begin{cases} \dfrac{x^2}{a^{*2}}+\dfrac{y^2}{b^{*2}}=1, \\ z=h. \end{cases}$$

Отсюда видно, что плоскость $z=h$ пересекает эллипсоид по эллипсу с полуосями

$$a^*=a\sqrt{1-\frac{h^2}{c^2}},\ b^*=b\sqrt{1-\frac{h^2}{c^2}}.$$

С уменьшением $|h|$ величины a^* и b^* увеличиваются и достигают своих наибольших значений при $h=0$, т. е. в сечении эллипсоида координатной плоскостью Oxy получится самый большой эллипс с полуосями $a^*=a$ и $b^*=b$.

Аналогичная картина получается и при сечении данной поверхности плоскостями, параллельными координатным плоскостям Oxz и Oyz.

Таким образом, полученные линии сечений позволяют нам геометрически изобразить эллипсоид как замкнутую овальную поверхность (рис. 36).

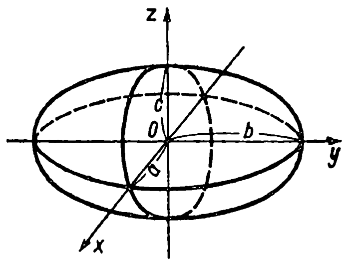


Рис. 36

Величины a, b, c называются *полуосями* эллипсоида. В случае $a=b=c$ эллипсоид является *сферой*.

2. Однополостный гиперболоид. *Однополостным гиперболоидом называется поверхность, которая в некоторой прямоугольной системе координат определяется уравнением*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1. \qquad (3)$$

Уравнение (3) называется *каноническим уравнением* однополостного гиперболоида.

Выясним геометрический вид поверхности (3). Для этого рассмотрим сечения ее координатными плоскостями Oxz ($y=0$) и Oyz ($x=0$). Получаем соответственно уравнения

$$\begin{cases}\dfrac{x^2}{a^2}-\dfrac{z^2}{c^2}=1,\\ y=0\end{cases} \quad \text{и} \quad \begin{cases}\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1,\\ x=0,\end{cases}$$

из которых видно, что линиями сечений являются гиперболы.

Теперь рассмотрим сечения данного гиперболоида плоскостями $z=h$, параллельными координатной плоскости Oxy. Линия сечения будет определяться уравнениями:

$$\begin{cases}\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1+\dfrac{h^2}{c^2},\\ z=h\end{cases} \quad \text{или} \quad \begin{cases}\dfrac{x^2}{a^{*2}}+\dfrac{y^2}{b^{*2}}=1,\\ z=h,\end{cases} \qquad (4)$$

из которых видно, что плоскость $z=h$ пересекает гиперболоид по эллипсу с полуосями

$$a^* = a\sqrt{1+\frac{h^2}{c^2}},\ b^* = b\sqrt{1+\frac{h^2}{c^2}},$$

достигающими своих наименьших значений при $h=0$, т. е. в сечении данного гиперболоида координатной плоскостью Oxy получается самый маленький эллипс с полуосями $a^*=a$ и $b^*=b$. При бесконечном возрастании $|h|$ величины a^* и b^* бесконечно возрастают.

Таким образом, полученные линии сечений позволяют нам геометрически изобразить однополостный гиперболоид в виде бесконечной трубки, бесконечно расширяющейся в обе стороны от плоскости Oxy (рис. 37).

Величины a, b, c называются *полуосями* однополостного гиперболоида, первые две из них изображены на рис. 37, а чтобы изобразить на чертеже полуось c, нужно было бы построить основной прямоугольник какой-нибудь из гипербол.

3. Двухполостный гиперболоид. *Двухполостным гиперболоидом называется поверхность, которая в некоторой прямоугольной системе координат определяется уравнением*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1. \tag{5}$$

Уравнение (5) называется *каноническим уравнением* двухполостного гиперболоида.

Выясним геометрический вид поверхности (5). Для этого рассмотрим ее сечения координатными плоскостями Oxz и Oyz. Получаем соответственно уравнения:

$$\begin{cases}\dfrac{x^2}{a^2}-\dfrac{z^2}{c^2}=-1,\\ y=0\end{cases} \quad \text{и} \quad \begin{cases}\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=-1,\\ x=0,\end{cases}$$

из которых видно, что линиями сечений являются гиперболы.

Теперь рассмотрим сечения данного гиперболоида плоскостями $z=h$, параллельными координатной плоскости Oxy. Линия сечения будет определяться уравнениями:

$$\begin{cases} \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = \dfrac{h^2}{c^2} - 1, \\ z = h \end{cases} \quad \text{или} \quad \begin{cases} \dfrac{x^2}{a^{*2}} + \dfrac{y^2}{b^{*2}} = 1, \\ z = h, \end{cases} \tag{6}$$

из которых видно, что при $|h|>c$ плоскость $z=h$ пересекает гиперболоид по эллипсу с полуосями

$$a^* = a\sqrt{\frac{h^2}{c^2} - 1}, \quad b^* = b\sqrt{\frac{h^2}{c^2} - 1}.$$

При возрастании $|h|$ величины a^* и b^* возрастают.

При $h=\pm c$ уравнения (6) удовлетворяются координатами лишь двух точек: $(0; 0; +c)$ и $(0; 0; -c)$. Это означает, что плоскости $z=\pm c$ касаются данной поверхности.

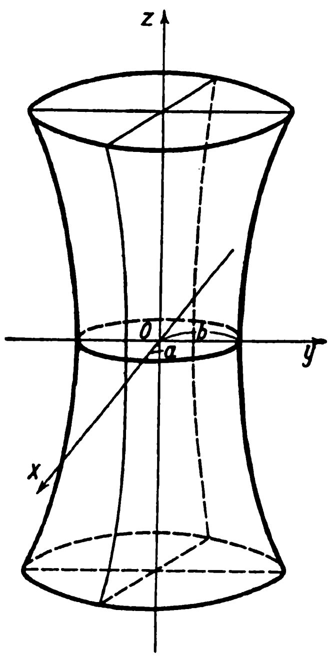


Рис. 37

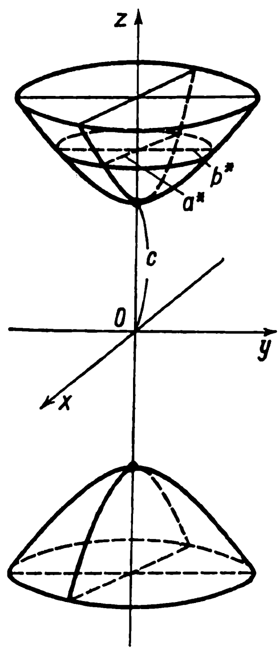


Рис. 38

При $|h|<c$ уравнения (6) определяют мнимый эл-

липс, т. е. точек пересечения плоскости $z=h$ с данным гиперболоидом не существует.

Таким образом, полученные линии сечений позволяют нам геометрически изобразить искомую поверхность как поверхность, состоящую из двух отдельных «полостей» (отсюда название — двухполостный), каждая из которых имеет вид бесконечно выпуклой чаши (рис. 38).

Величины a, b, c называются *полуосями* двухполостного гиперболоида. На рис. 38 изображена величина c. Чтобы изобразить на чертеже a и b, нужно построить основные прямоугольники гипербол в плоскостях Oxz и Oyz.

Теперь перейдем к рассмотрению двух поверхностей, которые являются пространственными аналогами парабол на плоскости.

4. Эллиптический параболоид. *Эллиптическим параболоидом называется поверхность, которая в некоторой прямоугольной системе координат определяется уравнением*

$$\frac{x^2}{p}+\frac{y^2}{q}=2z, \tag{7}$$

где $p>0$, $q>0$. Уравнение (7) называется *каноническим уравнением* эллиптического параболоида.

Исследуем геометрически эту поверхность. Рассмотрим сначала сечения данной поверхности координатными плоскостями Oxz и Oyz. При $y=0$ и $x=0$ из уравнения (7) имеем: $x^2=2pz$ и $y^2=2qz$. Таким образом, линии сечений определяются уравнениями

$$\begin{cases} x^2=2pz, \\ y=0 \end{cases} \text{ и } \begin{cases} y^2=2qz, \\ x=0, \end{cases}$$

т. е. являются параболами, симметричными относительно оси Oz, с вершинами в начале координат.

Теперь найдем линии сечения данного параболоида плоскостями $z=h$, параллельными координатной плоскости Oxy. Линия сечения будет определяться уравнениями:

$$\begin{cases} \dfrac{x^2}{p}+\dfrac{y^2}{q}=2h, \\ z=h \end{cases} \text{ или } \begin{cases} \dfrac{x^2}{a^{*2}}+\dfrac{y^2}{b^{*2}}=1, \\ z=h, \end{cases} \tag{8}$$

из которых видно, что при $h>0$ плоскость $z=h$ пересекает эллиптический параболоид по эллипсу с полуосями $a^*=\sqrt{2hp}$, $b^*=\sqrt{2hq}$. При возрастании h величины a^* и b^* возрастают; при $h=0$ эллипс вырождается в точку, т. е. плоскость $z=0$ касается данного параболоида.

При $h<0$ уравнения (8) определяют мнимый эллипс, т. е. точек пересечения плоскости $z=h$ с данным параболоидом не существует.

Таким образом, полученные линии сечений позволяют нам геометрически изобразить эллиптический параболоид в виде бесконечной выпуклой чаши (рис. 39).

Точка (0; 0; 0) называется *вершиной* эллиптического параболоида; числа p и q называются его *параметрами*.

В случае $p=q$ уравнения (8) определяют окружность с центром на оси Oz, т. е. эллиптический параболоид можно рассматривать как поверхность, образованную вращением параболы вокруг ее оси. Такая поверхность называется *параболоидом вращения*, уравнение которого $x^2+y^2=2pz$, а при $2p=1$ — $z=x^2+y^2$.

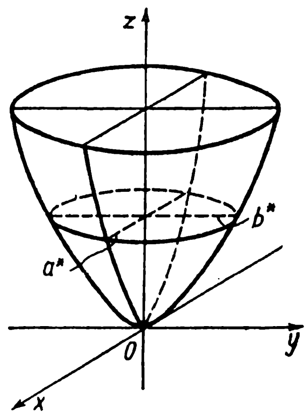


Рис. 39

5. Гиперболический параболоид. *Гиперболическим параболоидом называется поверхность, которая в некоторой прямоугольной системе координат определяется уравнением*

$$\frac{x^2}{p}-\frac{y^2}{q}=2z, \tag{9}$$

где $p>0$, $q>0$. Уравнение (9) называется *каноническим уравнением* гиперболического параболоида.

Выясним геометрический вид поверхности (9). Рассмотрим сечение параболоида координатной плоскостью Oxz ($y=0$). Получаем уравнения

$$\begin{cases} x^2=2pz, \\ y=0, \end{cases} \tag{10}$$

из которых видно, что линия сечения есть парабола, направленная вверх, симметричная относительно оси Oz, с вершиной в начале координат, а в сечениях поверхности с плоскостями, параллельными плоскости Oxz $(y=h)$, будут получаться параболы

$$\begin{cases} x^2 = 2p\left(z + \dfrac{h^2}{2q}\right), \\ y = h, \end{cases}$$

также направленные вверх.

Рассмотрим сечение данного параболоида плоскостью Oyz $(x=0)$. Получаем уравнения

$$\begin{cases} y^2 = -2qz, \\ x = 0, \end{cases}$$

из которых видно, что и в этом случае линией сечения является парабола, но теперь направленная вниз, симметричная относительно оси Oz, с вершиной в начале координат. Кроме этого, если рассмотрим сечения параболоида с плоскостями, параллельными плоскости Oyz $(x=h)$, то получим уравнения

$$\begin{cases} y^2 = -2q\left(z - \dfrac{h^2}{2p}\right), \\ x = h, \end{cases}$$

из которых видно, что при любом h линиями сечений будут параболы, направленные вниз, вершина каждой из которых лежит на параболе, определенной уравнениями (10).

Рассмотрим, наконец, сечения параболоида плоскостями $z=h$, параллельными плоскости Oxy. Линии сечений параболоида с плоскостями $z=h$ определяются уравнениями

$$\begin{cases} \dfrac{x^2}{p} - \dfrac{y^2}{q} = 2h, \\ z = h \end{cases} \quad \text{или} \quad \begin{cases} \dfrac{x^2}{2ph} - \dfrac{y^2}{2qh} = 1, \\ z = h, \end{cases}$$

из которых видно, что при $h>0$ в сечении будут получаться гиперболы, пересекающие плоскость Oxz; при $h<0$ — гиперболы, пересекающие плоскость Oyz; при

$h=0$ гипербола вырождается в пару пересекающихся прямых:

$$\begin{cases}\dfrac{x}{\sqrt{p}}-\dfrac{y}{\sqrt{q}}=0,\\ z=0\end{cases} \text{ и } \begin{cases}\dfrac{x}{\sqrt{p}}+\dfrac{y}{\sqrt{q}}=0,\\ z=0.\end{cases}$$

Таким образом, полученные линии сечений позволяют нам геометрически изобразить гиперболический параболоид в виде седлообразной поверхности (рис. 40). На чертеже изображено несколько сечений параболоида плоскостями $z=h$ для случаев $h>0$ и $h<0$.

Точка (0; 0; 0) называется *вершиной* гиперболического параболоида; числа p и q называется его *параметрами*.

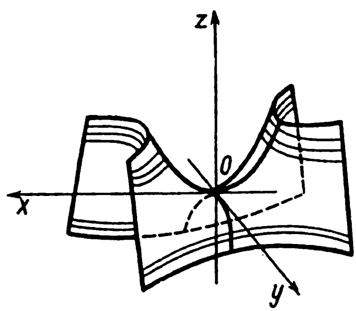


Рис. 40

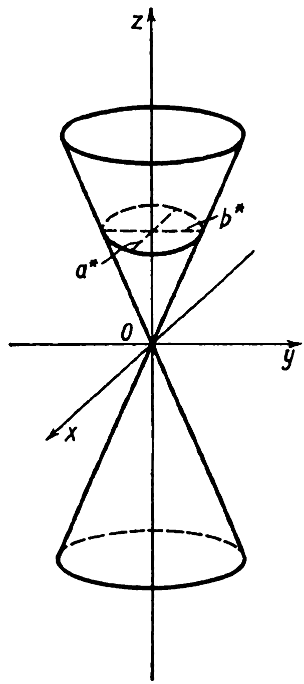


Рис. 41

В заключение параграфа мы рассмотрим поверхность, не имеющую аналога на плоскости.

6. Конус второго порядка. *Конусом второго порядка называется поверхность, которая в некоторой прямоугольной системе координат определяется уравнением*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0. \tag{11}$$

Уравнение (11) называется *каноническим уравнением* конуса второго порядка.

Исследуем геометрические свойства конуса. Рассмотрим сечение этой поверхности плоскостью Oxz ($y=0$). Получаем линию

$$\begin{cases} \dfrac{x^2}{a^2}-\dfrac{z^2}{c^2}=0, \\ y=0, \end{cases}$$

распадающуюся на пару пересекающихся прямых:

$$\begin{cases} \dfrac{x}{a}+\dfrac{z}{c}=0, \\ y=0 \end{cases} \text{ и } \begin{cases} \dfrac{x}{a}-\dfrac{z}{c}=0, \\ y=0. \end{cases}$$

Аналогично в сечении плоскостью Oyz ($x=0$) получается также пара пересекающихся прямых:

$$\begin{cases} \dfrac{y}{b}+\dfrac{z}{c}=0, \\ x=0 \end{cases} \text{ и } \begin{cases} \dfrac{y}{b}-\dfrac{z}{c}=0, \\ x=0. \end{cases}$$

Рассмотрим теперь сечения данной поверхности плоскостями $z=h$, параллельными плоскости Oxy. Линии сечений будут определяться уравнениями

$$\begin{cases} \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=\dfrac{h^2}{c^2}, \\ z=h \end{cases} \text{ или } \begin{cases} \dfrac{x^2}{a^{*2}}+\dfrac{y^2}{b^{*2}}=1, \\ z=h, \end{cases}$$

из которых видно, что при $h>0$ и $h<0$ в сечениях будут эллипсы с полуосями $a^*=\dfrac{a|h|}{c}$, $b^*=\dfrac{b|h|}{c}$. С ростом абсолютной величины h полуоси a^* и b^* будут увеличиваться.

При $h=0$ линия сечения поверхности с плоскостью $z=0$ вырождается в точку $(0; 0; 0)$.

Таким образом, полученные линии сечений позволяют нам представить конус в виде поверхности, изображенной на рис. 41.

Прямые, из которых составлен конус, называются его *образующими*; точка $(0; 0; 0)$, через которую они проходят, называется *вершиной* конуса.

7. Заключительные замечания. В нашем курсе мы ограничились изучением плоских линий и поверхностей первого и второго порядков. В гл. III рассмотрели линии первого и второго порядков, в данной главе — поверхности первого и второго порядков. Линии и поверхности выше второго порядка выходят за рамки учебных программ по аналитической геометрии.

Глава X

ЭЛЕМЕНТЫ ВЫСШЕЙ АЛГЕБРЫ

§ 1. Матрицы

1. Определение матрицы. Пусть даны прямоугольная система координат $Oxyz$, точка M с координатами x_1, x_2, x_3 и радиус-вектор* $\bar{r}$ точки M: $\bar{r}=\overline{OM}$. Обозначим через $\bar{e}_1$, $\bar{e}_2$, $\bar{e}_3$ единичные базисные векторы, тогда вектор $\bar{r}$ в данной системе координат запишется так:

$$\bar{r}=x_1\bar{e}_1+x_2\bar{e}_2+x_3\bar{e}_3.$$

Преобразуем данную систему координат путем ее поворота вокруг начала координат O и будем считать известными углы, которые образует каждая ось новой системы координат $Ox'y'z'$ с каждой осью старой. Обозначим через x'_1, x'_2, x'_3 координаты точки M в новой системе, а через $\bar{e}'_1$, $\bar{e}'_2$, $\bar{e}'_3$ единичные базисные векторы новых осей. Тогда вектор в новой системе координат запишется так:

$$\bar{r}=x'_1\bar{e}'_1+x'_2\bar{e}'_2+x'_3\bar{e}'_3.$$

Поскольку правые части вектора $\bar{r}$ представляют собой разложение одного и того же вектора, то имеет место векторное равенство

$$x_1\bar{e}_1+x_2\bar{e}_2+x_3\bar{e}_3=x'_1\bar{e}'_1+x'_2\bar{e}'_2+x'_3\bar{e}'_3. \qquad (1)$$

Рассмотрим преобразование координат x_1, x_2, x_3 точки M в координаты x'_1, x'_2, x'_3 той же точки.

Заметим, что данное преобразование обладает тем свойством, что вектор переходит в вектор той же дли-

* Радиусом-вектором точки M называется вектор, идущий из начала координат в эту точку.

ны. Такое преобразование называется *ортогональным.*

Разложим векторы $\bar{e}'_1$, $\bar{e}'_2$, $\bar{e}'_3$ по старому базису:

$$\begin{cases} \bar{e}'_1 = a_{11}\bar{e}_1 + a_{21}\bar{e}_2 + a_{31}\bar{e}_3, \\ \bar{e}'_2 = a_{12}\bar{e}_1 + a_{22}\bar{e}_2 + a_{32}\bar{e}_3, \\ \bar{e}'_3 = a_{13}\bar{e}_1 + a_{23}\bar{e}_2 + a_{33}\bar{e}_3. \end{cases} \tag{2}$$

Так как каждый из векторов $\bar{e}'_1$, $\bar{e}'_2$, $\bar{e}'_3$ является единичным, то для каждого из них коэффициентами разложения будут служить направляющие косинусы, т. е.

$$a_{11} = \cos(\bar{e}_1, \bar{e}'_1),\ a_{12} = \cos(\bar{e}_1, \bar{e}'_2),\ a_{13} = \cos(\bar{e}_1, \bar{e}'_3),$$

$$a_{21} = \cos(\bar{e}_2, \bar{e}'_1),\ a_{22} = \cos(\bar{e}_2, \bar{e}'_2),\ a_{23} = \cos(\bar{e}_2, \bar{e}'_3),$$

$$a_{31} = \cos(\bar{e}_3, \bar{e}'_1),\ a_{32} = \cos(\bar{e}_3, \bar{e}'_2),\ a_{33} = \cos(\bar{e}_3, \bar{e}'_3).$$

Заменяя в равенстве (1) векторы $\bar{e}'_1$, $\bar{e}'_2$, $\bar{e}'_3$ их разложениями (2) и группируя подобные члены, получим:

$$x_1\bar{e}_1 + x_2\bar{e}_2 + x_3\bar{e}_3 = (a_{11}x'_1 + a_{12}x'_2 + a_{13}x'_3)\cdot\bar{e}_1 +$$

$$+ (a_{21}x'_1 + a_{22}x'_2 + a_{23}x'_3)\cdot\bar{e}_2 + (a_{31}x'_1 + a_{32}x'_2 + a_{33}x'_3)\cdot\bar{e}_3.$$

Отсюда, сравнивая коэффициенты при одинаковых базисных векторах находим

$$\begin{cases} x_1 = a_{11}x'_1 + a_{12}x'_2 + a_{13}x'_3, \\ x_2 = a_{21}x'_1 + a_{22}x'_2 + a_{23}x'_3, \\ x_3 = a_{31}x'_1 + a_{32}x'_2 + a_{33}x'_3. \end{cases} \tag{3}$$

Таким образом, видим, что координаты x_1, x_2, x_3 представляют собой линейные комбинации координат x'_1, x'_2, x'_3, полностью определяемых совокупностью коэффициентов $a_{11}, \ldots, a_{33}$. Здесь мы подошли к понятию матрицы.

Определение. *Таблица, составленная из коэффициентов (3), записанная в виде*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \tag{4}$$

называется матрицей данного преобразования. Коротко матрицу обозначают так:

$$A=(a_{ij}) \quad (i=1, 2, 3;\ j=1, 2, 3),$$

a_{ij} называются *элементами* данной матрицы.

Элементы матрицы образуют *столбцы* и *строки.* Первый индекс (i) указывает номер строки, а второй (j) — номер столбца, на пересечении которых стоит элемент a_{ij}. Матрица (4) имеет три строки и три столбца.

В высшей алгебре рассматриваются матрицы с любым числом строк и с любым числом столбцов. Поэтому в общем виде матрица записывается следующим образом:

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{5}$$

Если в матрице число строк равно числу столбцов ($m=n$), то матрица называется *квадратной n-го порядка,* а в противном случае — *прямоугольной.* Так, матрица (4) является квадратной матрицей 3-го порядка. В матрице (5) m строк и n столбцов. Если $m=1$, $n>1$, то получается одностроечная матрица $(a_1\, a_2 \dots a_n)$, которая называется *вектор-строкой.* Если же $m>1$, а $n=1$, то получается одностолбцовая матрица

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix},$$

которая называется *вектор-столбцом.*

Две матрицы $A=(a_{ij})$ и $B=(b_{ij})$ считаются *равными,* если равны элементы, стоящие на одинаковых местах, т. е. $(a_{ij})=(b_{ij})$, если $a_{ij}=b_{ij}$ при всех i и j (при этом число строк и столбцов матриц A и B должно быть одинаковым).

2. Свойства матриц. Матрицы, подобно векторам, можно складывать, умножать на число и друг на друга. Рассмотрим эти свойства.

С у м м а д в у х м а т р и ц. *Суммой двух матриц* $A = (a_{ij})$ *и* $B = (b_{ij})$ *с одинаковым количеством* m *строк и* n *столбцов называется новая матрица* $C = (c_{ij})$, *элементы которой определяются равенством*

$$a_{ij} + b_{ij} = c_{ij} \quad (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n).$$

Сумма двух матриц обозначается $A + B = C$.

П р и м е р 1.

$$\underset{A}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}} + \underset{B}{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}} = \begin{pmatrix} 1+0 & 0+0 & 0+0 \\ 0+0 & 1+0 & 0+0 \\ 0+0 & 0+0 & 0+1 \end{pmatrix} = \underset{C}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}.$$

Аналогично определяется разность двух матриц.

У м н о ж е н и е м а т р и ц ы н а ч и с л о. *Чтобы умножить матрицу* $A = (a_{ij})$ *на число* λ, *нужно умножить на это число все элементы матрицы:*

$$\lambda A = \lambda(a_{ij}) = (\lambda a_{ij}) \quad (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n).$$

П р и м е р 2.

$$\underset{3(a_{ij})}{3\begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}} = \begin{pmatrix} 3\cdot 1 & 3\cdot 0 & 3\cdot 2 \\ 3\cdot 2 & 3\cdot 0 & 3\cdot 0 \\ 3\cdot 0 & 3\cdot 1 & 3\cdot 0 \end{pmatrix} = \underset{(3a_{ij})_{(i=1,2,3;\, j=1,2,3)}}{\begin{pmatrix} 3 & 0 & 6 \\ 6 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}}.$$

П р о и з в е д е н и е д в у х м а т р и ц. *Произведением матриц* $A = (a_{ij})$, *имеющей* m *строк и* k *столбцов, и* $B = (b_{ij})$, *имеющей* k *строк и* n *столбцов, называется новая матрица* $C = (c_{ij})$, *имеющая* m *строк и* n *столбцов, у которой элемент* c_{ij} *равен сумме произведений элементов i-й строки матрицы A на j-й столбец матрицы B, т. е.*

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj}$$
$$(i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n).$$

При этом число столбцов матрицы A должно быть равно числу строк матрицы B. В противном случае произведение не определено. Произведение двух матриц обозначается $A \cdot B = C$.

Пример 3.

$$\underset{A}{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}} \cdot \underset{B}{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}} =$$

$$= \begin{pmatrix} 0\cdot 1+1\cdot 0+0\cdot 1 & 0\cdot 0+1\cdot 1+0\cdot 0 & 0\cdot 1+1\cdot 0+0\cdot 1 \\ 1\cdot 1+0\cdot 0+1\cdot 1 & 1\cdot 0+0\cdot 1+1\cdot 0 & 1\cdot 1+0\cdot 0+1\cdot 1 \end{pmatrix} =$$

$$= \underset{C}{\begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix}}.$$

Пример 4. Пусть

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

тогда

$$A\cdot B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \text{ но } B\cdot A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Отсюда получаем

$$AB \neq BA,$$

т. е. умножение матриц не подчиняется перестановочному свойству.

Замечание. Правило произведения легко запомнить в таком виде: элемент c_{ij} матрицы C, стоящей на пересечении i-й строки и j-го столбца, есть скалярное произведение i-й вектор-строки матрицы A и j-го вектор-столбца матрицы B.

Путем непосредственной проверки можно убедиться, что для суммы и произведения матриц справедливы следующие соотношения:

$$(A+B)\cdot C = A\cdot C + B\cdot C,$$
$$C\cdot (A+B) = C\cdot A + C\cdot B,$$
$$A\cdot (B\cdot C) = (A\cdot B)\cdot C,$$
$$(A+B)+C = A+(B+C).$$

Умножение на единичную матрицу. Совокупность элементов $a_{11}, a_{22}, \ldots, a_{nn}$ квадратной матрицы $A=(a_{ij})$ называется *главной диагональю* матрицы. Матрица, у которой элементы, стоящие на главной диагонали равны единице, а все остальные элементы равны нулю, называется *единичной матрицей* и обозначается буквой E.

Так, единичной матрицей 3-го порядка будет

$$E=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}. \tag{6}$$

Единичная матрица обладает замечательным свойством, а именно: умножение квадратной матрицы любого порядка на соответствующую единичную матрицу не меняет матрицу. Это свойство и объясняет ее название.

Пример 5. Пусть

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} \text{ и } E=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

тогда согласно правилу умножения матриц имеем

$$A\cdot E=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

и

$$E\cdot A=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix},$$

отсюда видим, что

$$A\cdot E=A \text{ и } E\cdot A=A.$$

Заметим, что единичной матрице (6) соответствует специальное преобразование:

$$x_1=x_1',\ x_2=x_2',\ x_3=x_3'.$$

Такое преобразование называется *тождественным*.

С понятием матрицы тесно связано другое — понятие определителя, к рассмотрению которого мы и переходим.

§ 2. Определители

1. Определение определителя. Пусть дана квадратная матрица третьего порядка, элементы которой для удобства обозначим через a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}. \qquad (1)$$

Определение. *Определителем* третьего порядка, соответствующим матрице* (1), *называется число, обозначаемое символом*:

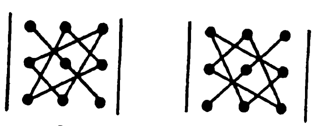
$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

и определяемое равенством

$$\Delta = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - \\ - c_1b_2a_3 - b_1a_2c_3 - a_1c_2b_3. \qquad (2)$$

Числа a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3 называются *элементами* определителя. Диагональ, образованная элементами a_1, b_2, c_3, называется *главной*, а диагональ, образованная элементами a_3, b_2, c_1, — *побочной*.

Чтобы запомнить, какие произведения в правой части равенства (2) берутся со знаком +, а какие со знаком —, полезно следующее правило треугольников:



+   −

Это правило позволяет легко написать формулу (2) и вычислить данный определитель. Например:

---

* Определители называются также детерминантами.

$$\begin{vmatrix} 3 & -2 & 1 \\ -2 & 1 & 3 \\ 2 & 0 & -2 \end{vmatrix} = 3\cdot 1\cdot(-2) + (-2)\cdot 3\cdot 2 +$$

$$+ (-2)\cdot 0\cdot 1 - 2\cdot 1\cdot 1 - 3\cdot 0\cdot 3 - (-2)\cdot(-2)\cdot(-2) = -12.$$

2. Свойства определителей. В этом пункте мы рассмотрим ряд свойств определителей. Эти свойства мы будем формулировать и доказывать для определителей третьего порядка, хотя сами свойства присущи более общим определителям любого порядка.

Свойство 1. *Величина определителя не изменится, если строки и столбцы этого определителя поменять ролями, т. е.*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Для доказательства этого свойства достаточно применить к определителям, стоящим в левой и правой частях равенства, формулу (2) и убедиться в равенстве полученных при этом членов.

Свойство 1 устанавливает полную равноправность строк и столбцов определителя. Поэтому все дальнейшие свойства определителя будем формулировать и для строк, и для столбцов, а доказывать только для строк или только для столбцов.

Свойство 2. *Перестановка двух столбцов или двух строк определителя равносильна умножению его на —1, т. е.*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix}.$$

Доказывается аналогично предыдущему.

Свойство 3. *Если определитель имеет два одинаковых столбца или две одинаковые строки, то он равен нулю.*

В самом деле, при перестановке двух одинаковых столбцов, с одной стороны, определитель Δ не изменится, а с другой стороны, в силу свойства 2, он изменит знак. Следовательно, $\Delta = -\Delta$, т. е. $2\Delta = 0$, или $\Delta = 0$.

Например:

$$\begin{vmatrix} 1 & 1 & 17 \\ 2 & 2 & 15 \\ 3 & 3 & 9 \end{vmatrix} = 0.$$

Свойство 4. *Умножение всех элементов одного столбца или одной строки определителя на любое число λ равносильно умножению определителя на это число λ, т. е.*

$$\begin{vmatrix} \lambda a_1 & b_1 & c_1 \\ \lambda a_2 & b_2 & c_2 \\ \lambda a_3 & b_3 & c_3 \end{vmatrix} = \lambda \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Для доказательства этого свойства достаточно заметить, что по формуле (2) определитель выражается в виде суммы, каждый член которой содержит множителем один элемент из каждой строки и из каждого столбца.

Свойство 5. *Если все элементы некоторого столбца или некоторой строки равны нулю, то сам определитель равен нулю.*

Это свойство вытекает из предыдущего (при $\lambda=0$).

Свойство 6. *Если элементы двух столбцов или двух строк определителя пропорциональны, то определитель равен нулю.*

В самом деле, если элементы двух столбцов определителя пропорциональны, то, в силу свойства 4, множитель пропорциональности можно вынести за знак определителя, после чего остается определитель с двумя одинаковыми столбцами, равный нулю согласно свойству 3.

Например:

$$\begin{vmatrix} 8 & 4 & 7 \\ 6 & 3 & 9 \\ 4 & 2 & 11 \end{vmatrix} = 2 \begin{vmatrix} 4 & 4 & 7 \\ 3 & 3 & 9 \\ 2 & 2 & 11 \end{vmatrix}.$$

Свойство 7. *Если каждый элемент n-го столбца (или n-й строки) определителя представляет собой сумму двух слагаемых, то определитель может быть представлен в виде суммы двух определителей, из которых один в n-м столбце (в n-й строке) имеет первые из упомянутых слагаемых, а другой — вторые; элементы,* *стоящие на остальных местах, у всех трех определителей одни и те же, т. е.*

$$\begin{vmatrix} a_1'+a_2'' & b_1 & c_1 \\ a_2'+a_2'' & b_2 & c_2 \\ a_3'+a_3'' & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1' & b_1 & c_1 \\ a_2' & b_2 & c_2 \\ a_3' & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1'' & b_1 & c_1 \\ a_2'' & b_2 & c_2 \\ a_3'' & b_3 & c_3 \end{vmatrix}.$$

Для доказательства этого свойства достаточно применить к определителям, стоящим в левой и правой частях равенства, формулу (2) и убедиться в равенстве полученных при этом членов.

Свойство 8. *Если к элементам некоторого столбца (или некоторой строки) прибавить соответствующие элементы другого столбца (или другой строки), умноженные на любой общий множитель* λ, *то величина определителя при этом не изменится.*

В самом деле, полученный в результате указанного прибавления определитель можно, в силу свойства 7, разбить на сумму двух определителей:

$$\begin{vmatrix} a_1+\lambda b_1 & b_1 & c_1 \\ a_2+\lambda b_2 & b_2 & c_2 \\ a_3+\lambda b_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \lambda b_1 & b_1 & c_1 \\ \lambda b_2 & b_2 & c_2 \\ \lambda b_3 & b_3 & c_3 \end{vmatrix},$$

первый из которых совпадает с исходным, а второй имеет два пропорциональных столбца и по свойству 6 равен нулю, т. е. получаем требуемое равенство.

Для формулировки следующего свойства определителя нам необходимо познакомиться с понятиями алгебраического дополнения и минора.

Минором некоторого элемента определителя называется определитель, получаемый из данного путем вычеркивания строки и столбца, на пересечении которых расположен этот элемент.

Например, минором элемента a_1 определителя Δ является определитель второго порядка* $\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$,

---

* Определителем второго порядка $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, соответствующим матрице $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$, называется число, равное $a_1b_2-a_2b_1$.

минором элемента b_1 является определитель второго порядка $\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$ и т. д.

Алгебраическим дополнением некоторого элемента определителя называется минор этого элемента, умноженный на $(-1)^p$, где p — сумма номеров строки и столбца, на пересечении которых расположен элемент. Алгебраическое дополнение элемента обозначается той же, но заглавной буквой, что и сам элемент. Так, алгебраическое дополнение элемента a_1 обозначается через A_1, элемента b_1 — через B_1 и т. д.

Например, так как элемент a_2 находится на пересечении первого столбца и второй строки, то для него $p=1+2=3$ и алгебраическим дополнением будет

$$A_2 = -\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} = b_3c_1 - b_1c_3.$$

Таким образом, соответствующие алгебраическое дополнение и минор могут отличаться только знаком.

Свойство 9. *Определитель равен сумме произведений элементов какого-нибудь столбца или строки на их алгебраические дополнения.*

Иначе говоря, имеют место следующие равенства:

$$\Delta=a_1A_1+a_2A_2+a_3A_3,\quad \Delta=a_1A_1+b_1B_1+c_1C_1, \tag{3}$$

$$\Delta=b_1B_1+b_2B_2+b_3B_3,\quad \Delta=a_2A_2+b_2B_2+c_2C_2, \tag{4}$$

$$\Delta=c_1C_1+c_2C_2+c_3C_3,\quad \Delta=a_3A_3+b_3B_3+c_3C_3. \tag{5}$$

Чтобы доказать, например, первое из этих равенств, достаточно записать правую часть формулы (2) в виде

$$\Delta=a_1(b_2c_3-b_3c_2)+a_2(b_3c_1-b_1c_3)+a_3(b_1c_2-b_2c_1).$$

Величины, стоящие в скобках, являются алгебраическими дополнениями элементов a_1, a_2, a_3, т. е.

$$b_2c_3-b_3c_2=A_1,\quad b_3c_1-b_1c_3=A_2,\quad b_1c_2-b_2c_1=A_3.$$

Отсюда и из предыдущего получаем

$$\Delta=a_1A_1+a_2A_2+a_3A_3,$$

что и требовалось. Остальные равенства (3)—(5) доказываются аналогично.

Запись определителя по какой-нибудь из формул (3)—(5) называется разложением его по элементам некоторого столбца или некоторой строки (первая формула дает разложение по элементам первого столбца и т. д.).

Пример. Вычислить определитель

$$\Delta=\begin{vmatrix} 2 & 4 & 6 \\ 5 & 12 & 19 \\ 3 & 9 & 17 \end{vmatrix},$$

разлагая его по элементам первой строки.

Решение.

$$\Delta=2\begin{vmatrix} 12 & 19 \\ 9 & 17 \end{vmatrix}-4\begin{vmatrix} 5 & 19 \\ 3 & 17 \end{vmatrix}+6\begin{vmatrix} 5 & 12 \\ 3 & 9 \end{vmatrix}=8.$$

Свойство 10. *Сумма произведений элементов какого-нибудь столбца или какой-нибудь строки определителя на алгебраические дополнения соответствующих элементов другого столбца или другой строки равна нулю.*

Докажем, например, что сумма произведений элементов второго и третьего столбцов на соответствующие алгебраические дополнения элементов первого столбца равна нулю. Для этого разложим определитель (1) по элементам первого столбца:

$$\Delta=a_1A_1+a_2A_2+a_3A_3. \tag{6}$$

Так как алгебраические дополнения A_1, A_2, A_3 не зависят от самих элементов a_1, a_2, a_3, то в равенстве (6) числа a_1, a_2, a_3 можно заменить произвольными числами h_1, h_2, h_3:

$$\begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix}=h_1A_1+h_2A_2+h_3A_3. \tag{7}$$

Беря теперь в равенстве (7) в качестве h_1, h_2, h_3 сначала элементы b_1, b_2, b_3 второго столбца, а затем элементы c_1, c_2, c_3 третьего столбца и учитывая, что по свойству 3 определитель с двумя одинаковыми столбцами равен нулю, получаем

$$b_1A_1+b_2A_2+b_3A_3=0,$$

$$c_1A_1+c_2A_2+c_3A_3=0,$$

что и требовалось.

Аналогично доказываются равенства

$$a_1B_1+a_2B_2+a_3B_3=0, \quad a_1C_1+a_2C_2+a_3C_3=0,$$
$$c_1B_1+c_2B_2+c_3B_3=0, \quad b_1C_1+b_2C_2+b_3C_3=0$$

и шесть подобных равенств, относящиеся не к столбцам, а к строкам:

$$a_1A_3+b_1B_3+c_1C_3=0, \quad a_2A_3+b_2B_3+c_2C_3=0,$$
$$a_1A_2+b_1B_2+c_1C_2=0, \quad a_3A_2+b_3B_2+c_3C_2=0,$$
$$a_2A_1+b_2B_1+c_2C_1=0, \quad a_3A_1+b_3B_1+c_3C_1=0.$$

§ 3. Матричная запись системы линейных уравнений. Понятие обратной матрицы

В начале главы на примере ортогонального преобразования мы ввели математическое понятие матрицы. В этом параграфе мы рассмотрим практическую сторону этого понятия. С этой целью возьмем любые соотношения вида

$$\left.\begin{aligned} y_1&=a_{11}x_1+a_{12}x_2+a_{13}x_3,\\ y_2&=a_{21}x_1+a_{22}x_2+a_{23}x_3,\\ y_3&=a_{31}x_1+a_{32}x_2+a_{33}x_3,\end{aligned}\right\} \tag{1}$$

считая, что матрица их коэффициентов задана:

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}.$$

Тогда, пользуясь правилом умножения матриц, систему (1) можно записать в матричной форме следующим образом:

$$\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\cdot\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix}, \tag{2}$$

здесь правая часть равенства понимается как произведение квадратной матрицы A на матрицу, состоящую из одного столбца, равное матрице (левая часть равенства), состоящей также из одного столбца. Две матрицы равны, если равны их элементы. Поэтому, приравняв соответствующие элементы, получим систему уравнений (1).

Обозначим теперь числа y_1, y_2, y_3 одной буквой Y, а числа x_1, x_2, x_3 одной буквой X; матричное равенство (2) принимает компактный вид

$$Y=A\cdot X. \tag{3}$$

Таким образом, в матричной записи систему (1) можно заменить одним эквивалентным ей матричным уравнением (3). Решение уравнения (3) заключается в отыскании такого вектор-столбца X, который при заданной матрице A и заданном вектор-столбце Y обращает уравнение в тождество.

Если определитель матрицы A равен нулю ($\Delta=0$), то матрица называется *вырожденной*, если же $\Delta\neq0$, то матрица называется *невырожденной*.

Пусть определитель матрицы A отличен от нуля ($\Delta\neq0$). Тогда существует единственное обратное преобразование, которое находится путем решения системы (1) относительно x_1, x_2, x_3*. Матрица обратного преобразования называется *обратной матрицей к* A и обозначается A^{-1}. В этом случае можем написать

$$X=A^{-1}\cdot Y. \tag{4}$$

Формулу (4) называют *матричной записью* решения системы (1).

Произведение матриц A и A^{-1} обладает важным свойством:

$$A^{-1}\cdot A=A\cdot A^{-1}=E. \tag{5}$$

Действительно, подставляя вместо Y в (4) его выражение (3), получаем тождественное преобразование: $X=A^{-1}\cdot(A\cdot X)=(A^{-1}\cdot A)\cdot X$, отсюда $A^{-1}\cdot A=E$. С другой стороны, если рассматривать данное преобразование и обратное преобразование в обратном порядке, то мы также получим тождественное преобразование $Y=(A\cdot A^{-1})\cdot Y$, откуда $A\cdot A^{-1}=E$.

---

* Это следует из результатов следующего § 4.

Равенства (5) показывают, что матрицы A и A^{-1} взаимно-обратные.

Покажем, что обратной матрицей A^{-1} будет матрица

$$A^{-1}=\begin{pmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \frac{A_{31}}{\Delta} \\ \frac{A_{12}}{\Delta} & \frac{A_{22}}{\Delta} & \frac{A_{32}}{\Delta} \\ \frac{A_{13}}{\Delta} & \frac{A_{23}}{\Delta} & \frac{A_{33}}{\Delta} \end{pmatrix}, \tag{6}$$

где A_{ij} есть алгебраическое дополнение элемента a_{ij} определителя Δ. С этой целью найдем матрицу C как произведение матриц $A \cdot A^{-1}$:

$$C = A\cdot A^{-1} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \frac{A_{31}}{\Delta} \\ \frac{A_{12}}{\Delta} & \frac{A_{22}}{\Delta} & \frac{A_{32}}{\Delta} \\ \frac{A_{13}}{\Delta} & \frac{A_{23}}{\Delta} & \frac{A_{33}}{\Delta} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E.$$

Полученный результат следует из того, что, согласно правилу умножения матриц и свойств 9 и 10 определителей, элементы матрицы C, стоящие на главной диагонали, суть суммы произведений элементов строки определителя Δ на соответствующие им алгебраические дополнения, деленные на определитель Δ, т. е. равны единице. А каждый недиагональный элемент есть сумма произведений элементов некоторой строки на алгебраические дополнения другой строки, деленная на определитель Δ, т. е. равен нулю. Например, элементы c_{11} и c_{23} определяются так:

$$c_{11} = a_{11}\frac{A_{11}}{\Delta} + a_{12}\frac{A_{12}}{\Delta} + a_{13}\frac{A_{13}}{\Delta} =$$

$$= \frac{a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}}{\Delta} = \frac{\Delta}{\Delta} = 1,$$

$$c_{23} = a_{21}\frac{A_{31}}{\Delta} + a_{22}\frac{A_{32}}{\Delta} + a_{23}\frac{A_{33}}{\Delta} =$$

$$= \frac{a_{21}A_{31} + a_{22}A_{32} + a_{23}A_{33}}{\Delta} = \frac{0}{\Delta} = 0.$$

Таким образом, формула (6) установлена.

Пример. Дана система

$$\begin{cases} y_1 = x_1 + 2x_2 + x_3, \\ y_2 = 2x_1 + x_2 + x_3, \\ y_3 = x_1 + 3x_2 + x_3. \end{cases}$$

Выразить неизвестные x_1, x_2, x_3 через y_1, y_2, y_3.

Решение. Имеем

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix},$$

$\Delta = 1 \neq 0$. Матрица A является невырожденной, следовательно, имеет обратную. Согласно (6) находим

$$A^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix}.$$

Отсюда

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

или

$$\begin{cases} x_1 = -2y_1 + y_2 + y_3, \\ x_2 = -y_1 \qquad\quad + y_3, \\ x_3 = 5y_1 - y_2 - 3y_3. \end{cases}$$

В заключение отметим, что матричная запись решения системы имеет практическое значение в тех задачах, где приходится многократно решать систему при одних и тех же коэффициентах в правых частях, но с различными значениями левых частей. В таких задачах достаточно один раз затратить труд на составление обратной матрицы, чтобы затем уже по готовым формулам находить неизвестные в зависимости от левых частей системы. Особое значение этот метод получает в задачах, сводящихся к линейным системам с большим числом неизвестных.

§ 4. Решение и исследование системы трех уравнений первой степени с тремя неизвестными

Теория матриц и определителей имеет широкое применение как в самой математике, так и в ее приложениях. Это очень удобный и часто употребляемый в самых разнообразных исследованиях математический аппарат.

Мы рассмотрим применение матриц и определителей в исследовании и решении системы трех уравнений первой степени с тремя неизвестными x, y, z:

$$\begin{cases} a_1x + b_1y + c_1z = h_1, \\ a_2x + b_2y + c_2z = h_2, \\ a_3x + b_3y + c_3z = h_3 \end{cases} \qquad (1)$$

(коэффициенты a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3 и свободные члены h_1, h_2, h_3 считаются заданными).

Тройка чисел x_0, y_0, z_0 называются *решением* системы (1), если подстановка этих чисел на место x, y, z обращает все три уравнения (1) в тождество.

В дальнейшем основную роль будут играть четыре определителя:

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \quad \Delta_x = \begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix}, \quad \Delta_y = \begin{vmatrix} a_1 & h_1 & c_1 \\ a_2 & h_2 & c_2 \\ a_3 & h_3 & c_3 \end{vmatrix},$$

$$\Delta_z = \begin{vmatrix} a_1 & b_1 & h_1 \\ a_2 & b_2 & h_2 \\ a_3 & b_3 & h_3 \end{vmatrix}.$$

Определитель Δ называется *определителем* системы (1). Определители Δ_x, Δ_y, Δ_z получаются из определителя системы Δ посредством замены свободными членами элементов соответственно первого, второго и третьего столбцов.

Мы рассмотрим отдельно два случая: 1) случай, когда· определитель Δ системы отличен от нуля, 2) случай, когда этот определитель равен нулю.

1. $\Delta \neq 0$. Здесь мы покажем, что решение системы (1) существует и оно единственно. С этой целью умножим обе части первого уравнения системы (1) на алгебраическое дополнение A_1, второго — на A_2, третьего — на A_3, а затем сложим эти уравнения. В результате получим

$$(a_1A_1+a_2A_2+a_3A_3)\cdot x+(b_1A_1+b_2A_2+b_3A_3)\cdot y+$$
$$+(c_1A_1+c_2A_2+c_3A_3)\cdot z=h_1A_1+h_2A_2+h_3A_3.$$

Отсюда и на основании свойств 9 и 10 определителей имеем

$$\Delta\cdot x=h_1A_1+h_2A_2+h_3A_3=\Delta_x.$$

Аналогично найдем:

$$\Delta\cdot y=h_1B_1+h_2B_2+h_3B_3=\Delta_y,$$
$$\Delta\cdot z=h_1C_1+h_2C_2+h_3C_3=\Delta_z.$$

Таким образом, из системы (1) мы получим систему уравнений

$$x=\frac{\Delta_x}{\Delta},\quad y=\frac{\Delta_y}{\Delta},\quad z=\frac{\Delta_z}{\Delta},\tag{2}$$

называемых *формулами Крамера* *.

Чтобы показать, что решение системы (1) *существует*, надо подставить на место x, y, z их значения, определяемые формулами Крамера, и убедиться, что все три уравнения (1) обращаются при этом в тождества. Проверим это относительно первого уравнения. Имеем

$$a_1x+b_1y+c_1z=a_1\frac{\Delta_x}{\Delta}+b_1\frac{\Delta_y}{\Delta}+c_1\frac{\Delta_z}{\Delta}=$$
$$=\frac{1}{\Delta}\{a_1(h_1A_1+h_2A_2+h_3A_3)+b_1(h_1B_1+h_2B_2+h_3B_3)+$$
$$+c_1(h_1C_1+h_2C_2+h_3C_3)\}=\frac{1}{\Delta}\{h_1(a_1A_1+b_1B_1+c_1C_1)+$$
$$+h_2(a_1A_2+b_1B_2+c_1C_2)+h_3(a_1A_3+b_1B_3+c_1C_3)\}=h_1,$$

* Крамер Габриель (1704—1752) — швейцарский математик.

так как согласно свойству 9 определителей

$$a_1A_1+b_1B_1+c_1C_1=\Delta,$$

а согласно свойству 10

$$a_1A_2+b_1B_2+c_1C_2=0, \quad a_1A_3+b_1B_3+c_1C_3=0.$$

Следовательно, обращение в тождество первого уравнения системы (1) установлено. Аналогично устанавливается обращение в тождество второго и третьего уравнений системы.

Таким образом, решение системы существует.

Кроме этого, формулы Крамера (2) доказывают *единственность* решения исходной системы (1), ибо система (2) является следствием системы (1), и всякое решение системы (1) обязано быть решением и системы (2), т. е. выражаться по формулам Крамера.

Все изложенное позволяет сделать следующий вывод: *если определитель* Δ *системы* (1) *отличен от нуля, то существует, и притом единственное, решение этой системы, определяемое формулами Крамера.*

Пример. Найти все решения системы

$$\begin{cases} x+2y+z=4, \\ 3x-5y+3z=1, \\ 2x+7y-z=8. \end{cases}$$

Решение. Так как $\Delta=33\neq0$, то данная система имеет единственное решение, определяемое формулами (2)

$$x=\frac{\Delta_x}{\Delta}=\frac{33}{33}=1, \quad y=\frac{\Delta_y}{\Delta}=\frac{33}{33}=1,$$

$$z=\frac{\Delta_z}{\Delta}=\frac{33}{33}=1.$$

Следовательно, $x=1$, $y=1$, $z=1$ — решение данной системы.

Для сравнения покажем, что факт существования и единственности решения системы (1) еще проще устанавливается матричным способом. Для этого заменим систему (1) эквивалентным ей матричным уравнением

$$AX=H, \qquad (3)$$

где A — матрица системы, а X и H — вектор-столбцы, первый из которых подлежит определению, а второй задан.

Так как A — невырожденная матрица ($\Delta\neq 0$), то существует обратная матрица A^{-1}. Умножая равенство (3) слева на A^{-1}, будем иметь

$$A^{-1}\cdot AX=A^{-1}\cdot H. \qquad (4)$$

Так как $A^{-1}AX=EX=X$, в силу свойства (5) § 3, из равенства (4) находим решение системы:

$$X=A^{-1}\cdot H. \qquad (5)$$

Покажем, что это решение единственное. Предположим обратное, т. е. пусть X_1 и X_2 — два различных решения системы. Это значит, что выполняются одновременно тождества: $AX_1=H$ и $AX_2=H$. Вычитая из первого второе, получим

$$A(X_1-X_2)=0.$$

Умножая это равенство слева на A^{-1}, будем иметь

$$A^{-1}\cdot A(X_1-X_2)=A^{-1}\cdot 0=0, \quad \text{или} \quad E\cdot(X_1-X_2)=0.$$

Отсюда $X_1-X_2=0$, $X_1=X_2$, т. е. решения X_1 и X_2 совпадают.

Легко проверить, что вектор-столбец X, определяемый равенством (5), в самом деле является решением матричного уравнения (3), т. е. при подстановке в это уравнение обращает его в тождество. Действительно, если вектор-столбец X определяется равенством (5), то

$$AX=A(A^{-1}H)=(AA^{-1})\cdot H=EH=H.$$

Далее, нетрудно заметить, что если развернуть равенство (5) с учетом формулы (6) § 3 для обратной матрицы, то получим для элементов вектор-столбца X формулы Крамера.

Теперь перейдем к рассмотрению случая, когда определитель Δ системы (1) равен нулю.

2. $\Delta=0$. Начнем с простого случая, когда хотя бы один из определителей Δ_x, Δ_y, Δ_z отличен от нуля. В этом случае хотя бы одно из равенств (2) является невозможным *, т. е. система (2) не имеет решений, а поэтому не имеет решений и система (1), так как система (2) является следствием системы (1).

Осталось рассмотреть случай, когда Δ_x, Δ_y и Δ_z равны нулю. Но сначала исследуем так называемые однородные системы.

Однородной системой трех уравнений первой степени с тремя неизвестными называется система вида

$$\begin{cases} a_1x + b_1y + c_1z=0, \\ a_2x + b_2y + c_2z=0, \\ a_3x + b_3y + c_3z=0. \end{cases} \tag{6}$$

Очевидно, что эта система всегда имеет так называемое нулевое решение $x=0$, $y=0$, $z=0$. Если $\Delta\neq 0$, то это решение является единственным (в силу п. 1).

Мы докажем, *что если определитель* $\Delta=0$, *то система (6) имеет бесконечно много ненулевых решений.* Разобьем доказательство на два этапа.

1°. Предположим, что хотя бы один из миноров определителя Δ отличен от нуля. Пусть, например,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0.$$

Тогда однородную систему, составленную из двух первых уравнений (6), мы можем представить в виде

$$\begin{cases} a_1x + b_1y = -c_1z, \\ a_2x + b_2y = -c_2z. \end{cases} \tag{7}$$

Для дальнейшего удобно ввести в рассмотрение алгебраические дополнения A_3, B_3, C_3 элементов третьей строки определителя Δ:

$$A_3 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, B_3 = -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}, C_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Так как по условию $C_3\neq 0$, то для каждого z существует единственное решение системы (7), определяемое формулами Крамера

$$x = \frac{A_3}{C_3} z, \ y = \frac{B_3}{C_3} z.$$

---

* Действительно, пусть, например, $\Delta_x\neq 0$. Тогда равенство $\Delta\cdot x=\Delta_x$ невозможно, так как его левая часть $\Delta\cdot x=0$ при любом x, а правая часть $\Delta_x\neq 0$.

Положим $z=C_3\cdot t$, где t может принимать любые значения. Тогда очевидно, что однородная система (7) будет иметь бесконечно много решений, определяемых формулами

$$x=A_3t, \quad y=B_3t, \quad z=C_3t \tag{8}$$

при любом значении t.

Осталось показать, что x, y и z, определяемые формулами (8), обращают в тождество и третье уравнение однородной системы (6). В самом деле, подставляя их вместо неизвестных в левую часть третьего уравнения, находим

$$a_3x+b_3y+c_3z=(a_3A_3+b_3B_3+c_3C_3)\cdot t=\Delta\cdot t=0\cdot t=0.$$

Таким образом, формулы (8) при любом t определяют решение однородной системы (6).

2°. Предположим теперь, что все миноры определителя Δ равны нулю. Это значит, что соответствующие коэффициенты всех трех уравнений (6) пропорциональны. Но тогда второе и третье уравнения (6) являются следствием первого и могут быть отброшены, а одно уравнение с тремя неизвестными $a_1x+b_1y+c_1z=0$, естественно, имеет бесконечно много решений (двум неизвестным можно предписывать произвольные значения, а третье неизвестное определять из уравнения).

Итак, мы доказали, что однородная система (6) с определителем Δ, равным нулю, имеет бесконечно много решений.

Теперь мы можем перейти к рассмотрению системы (1), когда все четыре определителя Δ, Δ_x, Δ_y, Δ_z равны нулю. Докажем, *что если* $\Delta=0$ *и система (1) имеет хотя бы одно решение, то она имеет бесконечно много решений.*

Пусть система (1) имеет решение x_0, y_0, z_0. Тогда справедливы тождества

$$\begin{cases} a_1x_0+b_1y_0+c_1z_0=h_1, \\ a_2x_0+b_2y_0+c_2z_0=h_2, \\ a_2x_0+b_3y_0+c_3z_0=h_3. \end{cases} \tag{9}$$

Вычитая почленно из уравнений (1) тождества (9), получим

$$\begin{cases} a_1(x-x_0)+b_1(y-y_0)+c_1(z-z_0)=0, \\ a_2(x-x_0)+b_2(y-y_0)+c_2(z-z_0)=0, \\ a_3(x-x_0)+b_3(y-y_0)+c_3(z-z_0)=0. \end{cases} \qquad (10)$$

Это есть однородная система трех уравнений первой степени с неизвестными $(x-x_0)$, $(y-y_0)$ и $(z-z_0)$ с определителем Δ, равным нулю. Но согласно только что доказанному эта система имеет бесконечно много решений, а следовательно, и данная система (1) имеет бесконечно много решений. Например, в случае, когда отличен от нуля минор C_3 (см. 1°), в силу формул (8) каждому решению системы (10) отвечает решение

$$x=x_0+A_3t, \quad y=y_0+B_3t, \quad z=z_0+C_3t$$

системы (1) (t принимает любые значения). Тем самым наше утверждение доказано, и мы можем сделать следующее заключение: *если* $\Delta=\Delta_x=\Delta_y=\Delta_z=0$, *то система (1) либо совсем не имеет решений, либо имеет бесконечно много решений (в последнем случае по крайней мере одно из уравнений системы будет следствием других)*.

В качестве примеров предлагаем читателю рассмотреть следующие три системы:

$$\begin{cases} x+y+z=2, \\ 3x+2y+2z=1, \\ 4x+3y+3z=4, \end{cases} \quad \begin{cases} x+y+z=1, \\ 2x+y+z=2, \\ 3x+2y+2z=3, \end{cases} \quad \begin{cases} x+y+z=1, \\ 2x+2y+2z=3, \\ 3x+3y+3z=4 \end{cases}$$

и убедиться в том, что первая система не имеет решений (для нее $\Delta=0$, $\Delta_y=1$), вторая система имеет бесконечно много решений (для нее $\Delta=\Delta_x=\Delta_y=\Delta_z=0$), определяемых при любом t формулами: $x=1$, $y=t$, $z=-t$, а третья система не имеет решений (для нее $\Delta=\Delta_x=\Delta_y=\Delta_z=0$), так как она не имеет хотя бы одного решения (в самом деле, даже первые два уравнения этой системы не имеют решения, так как, умножая первое из них на 2 и вычитая его почленно из второго, получаем невозможное равенство $0=1$).

В заключение данной главы заметим, что при рассмотрении общих задач исследования и решения систем уравнений первой степени со многими неизвестными и во многих других вычислительных задачах математики приходится иметь дело с матрицами и определителями n-го порядка ($n=2, 3, 4, 5, ...$). Теория матриц и определителей произвольного порядка строится в общих чертах аналогично изложенной нами теории матриц и определителей третьего порядка. Однако строгое ее построение со всеми деталями требует введения дополнительных понятий и доказательства ряда сложных теорем. Поэтому читатель, желающий расширить и углубить свои знания, может познакомиться с этой теорией, как и с теорией систем уравнений первой степени со многими неизвестными, в любом курсе высшей алгебры.

Глава XI

ПОНЯТИЕ, ПРЕДЕЛ И НЕПРЕРЫВНОСТЬ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 1. Понятие функции нескольких переменных

1. Вводные замечания. До сих пор мы занимались изучением функций одной переменной, т. е. изучением функции, значения которой зависят от значений одной независимой переменной.

При изучении многих вопросов геометрии, естествознания и т. д. часто приходится иметь дело с такими зависимостями между несколькими переменными величинами, когда численные значения одной из них полностью определяются значениями других. Так, например, температура нагретого тела в данный момент времени t меняется от точки к точке. Поскольку каждая точка тела определяется тремя координатами x, y, z, то температура тела определяется значениями трех переменных x, y, z, а если еще учитывать зависимость температуры от времени t, то значения ее будут определяться значениями уже четырех переменных x, y, z, t. Или, например, площадь прямоугольника со сторонами, длины которых равны x и y, определяется значениями двух переменных x и y, а объем прямоугольного параллелепипеда с ребрами, длины которых равны x, y, z, определяется значениями трех переменных x, y, z. Примеров таких зависимостей можно привести сколько угодно.

Эта часть курса и посвящается изучению такого рода зависимостей. С этой целью вводится понятие функции нескольких переменных и развивается аппарат для исследования таких функций.

2. Определение функции двух и более переменных. В полной аналогии с функцией одной переменной вводится понятие функции двух переменных.

Определение. *Пусть X, Y и Z — некоторые множества. Функцией двух переменных f называется множество троек чисел $(x; y; z)$ таких, что $x \in X$, $y \in Y$, $z \in Z$ и каждая пара чисел $(x; y)$ входит в одну и только одну тройку этого множества. В этом случае говорят, что паре чисел $(x; y)$ поставлено в соответствие число z и пишут $z = f(x, y)$. При этом z называют зависимой переменной, а x и y — независимыми переменными (или аргументами); множество $(x; y)$* — областью определения функции, а множество Z — множеством значений функции.*

Как и в случае одной переменной, функция двух переменных обозначается: $z = z(x, y)$, $z = \varphi(x, y)$, $z = h(x, y)$, $z = F(x, y)$ и т. д.

Пусть некоторая функция обозначена символом $f(x, y)$. Тогда значение этой функции, соответствующее некоторым численным значениям аргументов $x = x_0$ и $y = y_0$, обозначается $f(x_0, y_0)$. Например, если $f(x, y) = xy$, то $f(1, 3) = 1 \cdot 3 = 3$, $f(2, 3) = 2 \cdot 3 = 6$ и т. д.

Так как каждой паре чисел $(x; y)$ при фиксированной системе координат соответствует единственная точка M плоскости и, обратно, каждой точке M соответствует пара чисел $(x; y)$, то функцию двух переменных можно рассматривать как функцию точки M и вместо записи $z = f(x, y)$ писать $z = f(M)$, а областью определения функции в этом случае будет некоторое множество $\{M\}$ точек плоскости. В дальнейшем мы будем пользоваться этими двумя записями функции двух переменных.

Способы задания функции двух переменных, как и в случае одной переменной, могут быть различными. В нашем курсе наиболее важным является аналитический способ задания, когда функция задается с помо-

---

* В дальнейшем множество пар чисел $(x; y)$ будем обозначать символом $\{M\}$.

щью формулы. Областью определения функции в этом случае считается множество всех точек плоскости, для которых эта формула имеет смысл.

Рассмотрим примеры функций двух переменных:

1. $z=x^2+y^2$. Областью определения этой функции является множество $\{M\}$ всех пар чисел $(x;\ y)$, т. е. вся плоскость Oxy, а множеством значений — промежуток $Z=[0,\ +\infty)$.

2. $z=\sqrt{1-x^2-y^2}$. Областью определения этой функции является множество всех точек, для которых выражение $\sqrt{1-x^2-y^2}$ определено, т. е. множество точек, для которых $1-x^2-y^2\geqslant 0$, или $x^2+y^2\leqslant 1$. Множество всех таких точек образует круг с центром в начале координат и радиусом, равным единице. Множество значений функции представляет собой отрезок $[0,\ 1]$.

3. $z=\dfrac{1}{\sqrt{x^2+y^2-1}}$. Областью определения этой функции является множество точек, координаты которых удовлетворяют неравенству $x^2+y^2>1$, т. е. множество точек, лежащих вне круга радиуса 1 с центром в начале координат, а множество значений представляет собой промежуток $(0,\ +\infty)$.

Из рассмотренных примеров видно, что областью определения функции двух переменных может быть вся плоскость Oxy или ее часть.

Далее, из аналитической геометрии мы знаем, что множество всех троек чисел $(x;\ y;\ z)$ образует координатное пространство. При этом каждой тройке $(x;\ y;\ z)$ в пространстве соответствует точка $M(x;\ y;\ z)$. Если вместо множества $\{M\}$ точек плоскости возьмем множество $\{M\}$ точек пространства, то совершенно аналогично можно дать определение функции трех переменных: $u=f(M)$ или $u=f(x;\ y;\ z)$. Областью определения функции трех переменных будет уже все пространство $Oxyz$ или его часть. Так, например, функция $u=x^2+y^2+z^2$ определена во всем пространстве $Oxyz$, а функция $u=\ln xyz$ определена на множестве точек пространства, координаты которых удовлетворяют неравенству $xyz>0$. Причем в первом случае множеством значений функции будет промежуток $[0,\ +\infty)$, а во втором — $(-\infty,\ +\infty)$.

Аналогично можно дать определение функции четырех переменных: $u=f(x, y, z, t)$. В этом случае множество четверок чисел $(x; y; z; t)$ образует так называемое четырехмерное пространство. Точка $M(x; y; z; t)$ является точкой этого пространства. Однако область определения функции четырех переменных уже не имеет простого геометрического истолкования. Аналогично можно ввести понятия функций пяти и вообще n переменных: $u=f(x_1, x_2, ..., x_n)$.

В дальнейшем мы будем подробно рассматривать функции двух переменных, имея в виду, что обобщение определений и полученных результатов на функции трех и более переменных представляет собой, как правило, лишь технические трудности.

§ 2. Геометрическое изображение функции двух переменных

Как известно, функция одной переменной геометрически изображается на плоскости в виде линии, определенной уравнением $y=f(x)$. Подобным же образом функция двух переменных геометрически изображается в пространстве в виде поверхности, определенной уравнением $z=f(x, y)$, т. е. сама формула, задающая функцию, и есть уравнение этой поверхности.

В аналитической геометрии мы познакомились с различными поверхностями и их уравнениями. Так, например, уравнение $z-2x+3y+8=0$ есть уравнение плоскости. Данная плоскость есть график функции $z=2x-3y-8$.

Уравнение $x^2+y^2+z^2=R^2$ есть уравнение сферы радиуса R с центром в начале координат. С другой стороны, сфера есть объединение графиков двух однозначных функций

$$z=\sqrt{R^2-x^2-y^2} \text{ и } z=-\sqrt{R^2-x^2-y^2}.$$

Построение графиков функций двух переменных в большинстве случаев представляет значительные трудности. Поэтому существует еще один способ геометрического изображения функции двух переменных, основанный на сечении поверхности плоскостями $z=c$, где c — любое число, параллельными плоскости Oxy.

Назовем *линией уровня* функции $z=f(x, y)$ множество точек $(x; y)$ плоскости Oxy, в которых функция принимает одно и то же значение c Очевидно, давая c различные значения, мы будем получать различные линии уровня для данной функции

Если взять числа $c_1, c_2, ..., c_n, ...$ в арифметической прогрессии с разностью h, то получим ряд линий уров-

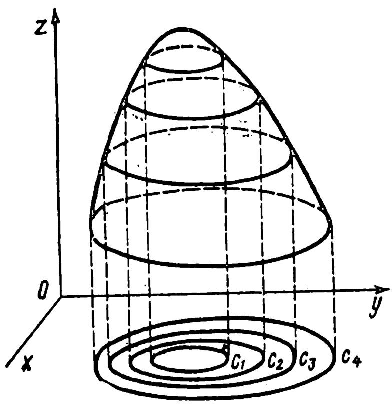


Рис. 42

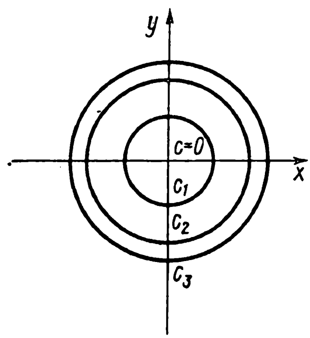


Рис. 43

ня, по взаимному расположению которых можно получить представление о графике функции, т. е. о форме поверхности. Там, где линии располагаются гуще, функция изменяется быстрее (поверхность идет круче), а в тех местах, где линии уровня располагаются реже, функция изменяется медленнее (поверхность будет более пологой) (рис. 42). Ясно, что чем меньше будет h, то тем точнее будет представление о графике функции.

Термин «линии уровня» заимствован из картографии. Там линии уровня — это линии, в которых высота точек земной поверхности над уровнем моря постоянна. По ним можно судить не только о высоте над уровнем моря интересующей нас точки местности, но и о характере рельефа.

Пример. Построить линии уровня функции $z=x^2+y^2$.

Решение. Для того чтобы найти линии уровня данной функции, пересечем поверхность $z=x^2+y^2$ плоскостью $z=c$. Получаем $x^2+y^2=c$ $(0 \leqslant c < +\infty)$. Задавая c различные значения, например $c=0, 1, 2, 3, \ldots$, получим семейство линий уровней, представляющих собой окружности. При $c=0$ окружность вырождается в точку $(0; 0)$ (рис. 43).

Из того, что линиями уровня оказались окружности с центрами в начале координат, следует, что графиком данной функции должна быть поверхность вращения вокруг оси Oz. Действительно, из аналитической геометрии известно, что уравнение $z=x^2+y^2$ определяет параболоид вращения.

§ 3. Предел функции двух переменных

Введем два вспомогательных понятия — понятие сходящейся последовательности точек и понятие δ-окрестности данной точки.

Определение 1. *Множество $\{M\}$ точек, координаты x и y которых удовлетворяют неравенству $\sqrt{(x-x_0)^2+(y-y_0)^2} < \delta$ или, короче, $\rho(M; M_0) < \delta$, называется δ-окрестностью точки $M_0(x_0; y_0)$.*

Другими словами, δ-окрестность точки M_0 — это все точки, лежащие внутри круга с центром M_0 радиуса δ.

Рассмотрим последовательность точек $M_1(x_1; y_1)$, $M_2(x_2; y_2)$, ..., $M_n(x_n; y_n)$, ...*. Сформулируем следующее определение.

Определение 2. *Последовательность $\{M_n\}$ точек называется сходящейся к точке M_0, если для любого $\varepsilon > 0$ существует номер N такой, что при $n > N$ выполняется неравенство $\rho(M_n, M_0) < \varepsilon$. При этом точка M_0 называется пределом последовательности $\{M_n\}$.*

Для обозначения предела M_0 последовательности $\{M_n\}$ используется следующая символика:

$$\lim_{n\to\infty} M_n = M_0 \text{ или } M_n \to M_0 \text{ при } n \to \infty.$$

Нетрудно заметить, что определение 2 является обобщением понятия сходящейся числовой последовательности. Действительно, задание последовательности

* Будем кратко обозначать эту последовательность символом $\{M_n\}$.

$\{M_n\}$ точек на прямой равносильно заданию числовой последовательности $\{x_n\}$ и неравенство $\rho(M_n, M_0)<\varepsilon$ переходит в этом случае в неравенство $|x_n-x_0|<\varepsilon$.

Теперь мы можем перейти к определению предела функции двух переменных. Оно будет совершенно аналогичным определению предела функции одной переменной.

Пусть функция $z=f(M)$ определена на некотором множестве $\{M\}$ и пусть точка $M_0\in\{M\}$ или $M_0\notin\{M\}$, но обладает тем свойством, что в любой δ-окрестности этой точки содержится хотя бы одна точка множества $\{M\}$, отличная от M_0.

Определение 3. *Число A называется пределом функции $z=f(M)$ в точке M_0, если для любой сходящейся к M_0 последовательности $M_1, M_2, \ldots, M_n, \ldots$ точек ($M_n\neq M_0$, $M_n\in\{M\}$) соответствующая последовательность $f(M_1), f(M_2), \ldots, f(M_n), \ldots$ значений функции сходится к A.*

Обозначается это следующим образом:

$$\lim_{M\to M_0} f(M) = A \text{ или } \lim_{\substack{x\to x_0\\ y\to y_0}} f(x, y) = A.$$

Так, например, функция $f(x, y)=x^2+y^2$ определена во всей плоскости. Найдем предел этой функции в точке $M_0(1;\ 2)$. Для любой последовательности точек $\{M_n\}$, сходящейся к точке M_0, будем иметь

$$\lim_{n\to\infty}(x_n^2+y_n^2)=\lim_{n\to\infty}x_n^2+\lim_{n\to\infty}y_n^2=1^2+2^2=5,$$

следовательно,

$$\lim_{\substack{x\to 1\\ y\to 2}}(x^2+y^2)=5.$$

Легко привести примеры функций, не имеющих предела в какой-либо точке. Например, функция $f(x, y)=\frac{x-y}{x+y}$ определена всюду, кроме точек прямой $x+y=0$. Покажем, что она не имеет предела в точке $(0,\ 0)$. Для этого выберем две сходящиеся к точке $(0,\ 0)$ последовательности точек: $M_n\left(\frac{1}{n}; 0\right)$ и $M_n'\left(0; \frac{1}{n}\right)$, тогда соответственно получим

$$\lim_{n\to\infty} \frac{\frac{1}{n}-0}{\frac{1}{n}+0} = 1 \text{ и } \lim_{n\to\infty} \frac{0-\frac{1}{n}}{0+\frac{1}{n}} = -1.$$

Таким образом, двум различным последовательностям точек, сходящимся к началу координат по разным направлениям, соответствуют две последовательности значений функции, имеющие разные пределы. Следовательно, по определению 3 данная функция не имеет предела в точке (0; 0).

Приведенное определение предела функции двух переменных дано с помощью понятия предела последовательности. Так же, как для функции одной переменной, можно дать эквивалентное ему определение, используя «ε—δ»-терминологию.

Определение 4. *Число A называется пределом функции $z=f(M)$ в точке M_0, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что для всех точек $M\in\{M\}$, удовлетворяющих условию $0<\rho(M, M_0)<\delta$, выполняется неравенство $|f(M)-A|<\varepsilon$.*

Доказательство эквивалентности определений 3 и 4 проводится точно так же, как и для функции одной переменной. Следует лишь в рассуждениях теоремы 4.1 заменить последовательность $\{x_n\}$ последовательностью точек $\{M_n\}$, точку x_0 — точкой M_0, разности $|x-x_0|$ и $|x_n-x_0|$ — расстояниями $\rho(M, M_0)$ и $\rho(M_n, M_0)$ соответственно, а числовую последовательность $\{f(x_n)\}$ заменить числовой последовательностью $\{f(M_n)\}$.

Пользуясь определением предела функции двух переменных, читатель может без особого труда перенести основные положения теории функции одной переменной на функции двух переменных. Например, имеет место следующая

Теорема. *Пусть функции $f(M)$ и $g(M)$ определены на одном и том же множестве $\{M\}$ и имеют в точке M_0 пределы B и C. Тогда функции $f(M)\pm g(M)$, $f(M)\cdot g(M)$ и $\frac{f(M)}{g(M)}$ $(C\neq 0)$ имеют в точке M_0 пределы, равные соответственно $B\pm C$, $B\cdot C$ и $\frac{B}{C}$.*

Доказательство этой теоремы совершенно аналогично доказательству теоремы 4.2 и может быть получено из него формальной заменой букв x и x_0 буквами M и M_0, только вместо определения 1 предела функции одной переменной следует использовать определение 3 предела функции двух переменных.

Функция $z=f(M)$ называется бесконечно малой в точке $M=M_0$ (или при $M \to M_0$), если $\lim\limits_{M \to M_0} f(M)=0$.

Если функция $z=f(M)$ имеет в точке M_0 предел, равный A, то функция $\alpha(M)=f(M)-A$ является бесконечно малой в точке M_0. Действительно,

$$\lim_{M \to M_0} \alpha(M) = \lim_{M \to M_0} [f(M)-A] = \lim_{M \to M_0} f(M) - \\ - \lim_{M \to M_0} A = A - A = 0.$$

Отсюда мы получаем специальное представление для функции, имеющей предел A в точке M_0: $f(M)=A+$ $+\alpha(M)$, где $\lim\limits_{M \to M_0} \alpha(M)=0$. При этом говорят, что функция $f(M)$ в окрестности точки M_0 отличается от A на бесконечно малую функцию.

Сравнение бесконечно малых функций двух переменных производится точно так же, как для бесконечно малых функций одной переменной, причем, как и в случае одной переменной, под символом $o(\beta)$ мы будем понимать любую бесконечно малую в данной точке M_0 функцию более высокого порядка малости, чем бесконечно малая в точке M_0 функция $\beta(M)$.

§ 4. Непрерывность функции двух переменных

Понятие непрерывности функции двух переменных вводится на основе понятия предела.

1. Определение непрерывности функции двух переменных. Пусть на некотором множестве $\{M\}$ определена функция $f(M)$, точка $M_0 \in \{M\}$ и любая δ-окрестность точки M_0 содержит точки множества $\{M\}$.

Определение 1. *Функция $z=f(M)$ называется непрерывной в точке M_0, если предел функции в этой точке существует и равен значению функции в этой точке, т. е.*

$$\lim_{M \to M_0} f(M) = f(M_0), \text{ или } \lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = f(x_0, y_0).$$

Точки, в которых функция не обладает свойством непрерывности, называются *точками разрыва* этой функции.

Например, функция

$$f(x, y) = \begin{cases} \frac{x-y}{x+y} \text{ при } x+y \neq 0, \\ 1 \quad \text{ при } x+y=0 \end{cases}$$

будет разрывна в точке (0; 0), так как не существует предела этой функции при $x \to 0$, $y \to 0$; функция

$$f(x, y) = \begin{cases} x^2 + y^2 \text{ всюду, кроме } x=1, y=2, \\ 0 \quad \text{ при } x=1, y=2 \end{cases}$$

в точке (1; 2) разрывна, так как

$$\lim_{\substack{x \to 1 \\ y \to 2}} f(x, y) = 5, \text{ а } f(1; 2) = 0.$$

Сформулируем определение непрерывности функции, используя определение предела функции в терминах «ε—δ».

Определение 2. *Функция $z=f(M)$ называется непрерывной в точке M_0, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех точек $M \in \{M\}$, удовлетворяющих условию $\rho(M, M_0) < \delta$, выполняется неравенство $|f(M) - f(M_0)| < \varepsilon$.*

Так же как для функции одной переменной, пользуясь данными определениями непрерывности и соответствующими теоремами о пределах, можно доказать, что арифметические операции над непрерывными функциями и построение сложных функций из непрерывных функций приводят также к непрерывным функциям.

В дальнейшем нам придется пользоваться определением 1 непрерывности функции, но записанным в ином виде.

Назовем *полным приращением* функции $z=f(M)$ в точке M_0 функцию Δz, определяемую формулой

$$\Delta z = f(M) - f(M_0),$$

где M — любая точка из области определения функции. Пусть точки M_0 и M имеют соответственно координаты $(x_0;\ y_0)$ и $(x;\ y)$. Обозначим $x-x_0=\Delta x$, $y-y_0=$ $=\Delta y$. Используя эти обозначения, получим для Δz следующее выражение:

$$\Delta z=f(x_0+\Delta x,\ y_0+\Delta y)-f(x_0,\ y_0).$$

Определение 3. *Функция $z=f(M)$ называется непрерывной в точке M_0, если ее полное приращение в этой точке есть бесконечно малая при $M\to M_0$ функция, т. е.*

$$\lim_{M\to M_0}\Delta z=\lim_{M\to M_0}[f(M)-f(M_0)]=0,\ \text{или}\ \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\Delta z=0.$$

Геометрически это означает, что при приближении точки M к точке M_0 аппликаты соответствующих точек поверхности функции $z=f(M)$ стремятся к аппликате поверхности в точке M_0.

Функция $z=f(M)$ называется непрерывной на некотором множестве $\{M\}$, если она непрерывна в каждой точке этого множества.

2. Основные свойства непрерывных функций двух переменных. В этом пункте мы лишь перечислим основные свойства непрерывных функций двух переменных, поскольку доказательства этих свойств в основном аналогичны доказательствам соответствующих свойств функций одной переменной. Для формулировки этих свойств введем ряд понятий для множеств $\{M\}$ точек плоскости.

Определение 4. *Множество $\{M\}$ точек плоскости называется линейно-связным, если любые две точки этого множества можно соединить непрерывной линией, состоящей из точек того же множества.*

Например, линейно-связным множеством будет любая фигура, не распадающаяся на отдельные, не соединенные между собой части.

Определение 5. *Точка M называется внутренней точкой некоторого множества $\{M\}$, если существует δ-окрестность этой точки, состоящая из точек данного множества.*

Определение 6. *Множество $\{M\}$, состоящее лишь из внутренних точек, называется открытым множеством.*

Определение 7. *Линейно-связное открытое мно-*

жество $\{M\}$ точек называется открытой областью, или, короче, областью.

Простейшими областями являются: внутренность треугольника, круга, эллипса и т. п.

Определение 8. *Точка M называется граничной точкой области, если в любой ее δ-окрестности есть точки, как принадлежащие, так и не принадлежащие этой области. Множество $\{M\}$ всех граничных точек области называется границей этой области.*

Например, для области, состоящей из точек, лежащих внутри круга, границей будет окружность.

Определение 9. *Множество $\{M\}$ точек, образованное областью и ее границей, называется замкнутой областью.*

Определение 10. *Область называется ограниченной, если существует круг, внутри которого она содержится.*

Отрезок и треугольник являются ограниченными областями. Прямая не является ограниченной областью.

Замкнутая ограниченная область, в которой определена функция двух переменных, является аналогом отрезка для функции одной переменной.

Теперь можем перейти к формулировкам основных свойств. К их числу мы отнесем следующие:

1) Если функция $z=f(M)$ непрерывна в замкнутой ограниченной области, то она ограничена в этой области, т. е. существует число k такое, что для всех точек области выполняется неравенство $|f(M)|<k$.

2) Если функция $z=f(M)$ непрерывна в замкнутой ограниченной области, то она достигает в этой области своих точных граней.

3) Если функция $z=f(M)$ непрерывна в замкнутой ограниченной области, то она принимает все промежуточные значения между любыми двумя своими значениями, т. е. если $A<C<B$, где A и B — какие-то значения функции $f(M)$ в данной области, то в этой области существует точка M_0, в которой $f(M_0)=C$.

Отсюда, в частности, следует, что если M_1 и M_2 — точки данной области и $f(M_1)<0$, а $f(M_2)>0$, то в области существует точка M_0, в которой $f(M_0)=0$.

4) Если функция $z=f(M)$ непрерывна в замкнутой ограниченной области, то она равномерно непрерывна в этой области, т. е. для любого $\varepsilon>0$ существует $\delta>0$

такое, что для любых двух точек M' и M'' области, удовлетворяющих условию $\rho(M', M'')<\delta$, выполняется неравенство $|f(M'')-f(M')|<\varepsilon$.

Доказательства этих свойств в данном курсе не рассматриваются.

Заканчивая данную главу, отметим, что понятия предела, непрерывности и перечисленные свойства функций двух переменных легко обобщаются на функции трех и более переменных.

Глава XII

ЧАСТНЫЕ ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИРУЕМОСТЬ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 1. Частные производные

Пусть на некотором множестве $\{M\}$ определена функция $z=f(M)$. Возьмем произвольную точку $M(x; y)$ этого множества и придадим переменной x произвольное приращение Δx, оставляя значение переменной y неизменным, т. е. перейдем на плоскости от точки $M(x; y)$ к точке $M_1(x+\Delta x; y)$. При этом Δx таково, что точка $M_1\in\{M\}$. Тогда соответствующее приращение функции

$$\Delta_x z=f(x+\Delta x, y)-f(x, y)$$

называется *частным приращением* функции по переменной x в точке $M(x; y)$.

Аналогично определяется частное приращение функции по переменной y:

$$\Delta_y z=f(x, y+\Delta y)-f(x, y).$$

Определение. *Если существует предел*

$$\lim_{\Delta x\to 0}\frac{\Delta_x z}{\Delta x}\ \left(\lim_{\Delta y\to 0}\frac{\Delta_y z}{\Delta y}\right),$$

то он называется частной производной функции $z=f(M)$ в точке M по переменной x (по переменной y) и обозначается одним из следующих символов:

$$z'_x,\ f'_x,\ \frac{\partial f}{\partial x},\ \frac{\partial z}{\partial x}\ \left(z'_y,\ f'_y,\ \frac{\partial f}{\partial y},\ \frac{\partial z}{\partial y}\right)$$

(иногда частные производные обозначаются без верхних индексов).

Из определения следует, что частная производная функции двух переменных по переменной x представляет собой обыкновенную производную функции одной переменной x при фиксированном значении переменной y. Поэтому вычисление частных производных производится по формулам и правилам вычисления производных функций одной переменной.

Примеры:

1. $z = x^2 - 2xy^2 + y^3,\ \frac{\partial z}{\partial x} = 2x - 2y^2,\ \frac{\partial z}{\partial y} = -4xy + 3y^2.$

2. $z = \operatorname{arc\,tg} \frac{x}{y},\ \frac{\partial z}{\partial x} = \frac{y}{x^2 + y^2},\ \frac{\partial z}{\partial y} = \frac{-x}{x^2 + y^2}.$

3. $z = x^2 \sin y,\ \frac{\partial z}{\partial x} = 2x \sin y,\ \frac{\partial z}{\partial y} = x^2 \cos y.$

4. $z = xye^{x+2y},\ \frac{\partial z}{\partial x} = e^{x+2y} \cdot y(1 + x),\ \frac{\partial z}{\partial y} = e^{x+2y} \cdot x(1 + 2y).$

§ 2. Понятие дифференцируемости функции

1. Определение дифференцируемости. Напомним, что полным приращением функции $z=f(M)$ в точке $M(x;\ y)$, соответствующим приращениям Δx и Δy переменных x и y, называется выражение

$$\Delta z = f(x+\Delta x,\ y+\Delta y) - f(x,\ y).$$

Функция $z=f(M)$ предполагается определенной в некоторой окрестности точки M.

Определение. *Функция $z=f(M)$ называется дифференцируемой в точке M, если ее полное приращение в этой точке может быть представлено в виде*

$$\Delta z = A\Delta x + B\Delta y + \alpha(\Delta x,\ \Delta y)\Delta x + \beta(\Delta x,\ \Delta y)\Delta y, \qquad (1)$$

где A и B — некоторые не зависящие от Δx и Δy числа, а $\alpha(\Delta x,\ \Delta y)$ и $\beta(\Delta x,\ \Delta y)$ — бесконечно малые при $\Delta x \to 0$, $\Delta y \to 0$ функции.

Известно, что если функция одной переменной дифференцируема в некоторой точке, то она непрерывна и имеет производную в этой точке. Из существования производной функции одной переменной в данной точке следует дифференцируемость функции в этой точке. Выясним, как переносятся эти свойства на функции двух переменных.

2. Необходимые условия дифференцируемости.

Теорема 12.1. *Если функция $z=f(M)$ дифференцируема в точке M, то она непрерывна в этой точке.*

Доказательство. Если функция $z=f(M)$ дифференцируема в точке M, то, как следует из соотношения (1), $\lim\limits_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\Delta z=0$, а это и означает, что функция непрерывна в точке M ■

Теорема 12.2. *Если функция $z=f(M)$ дифференцируема в точке M, то она имеет в этой точке частные производные $f_x'(x, y)$ и $f_y'(x, y)$, причем*

$$f_x'(x, y)=A, \quad f_y'(x, y)=B.$$

Доказательство. Так как функция $z=f(M)$ дифференцируема в точке M, то имеет место соотношение (1). Полагая в нем $\Delta y=0$, имеем $\Delta_x z=A\Delta x+$ $+\alpha(\Delta x, 0)\Delta x$, где $\alpha(\Delta x, 0)$ — бесконечно малая при $\Delta x\to 0$ функция. Деля на Δx и переходя к пределу при $\Delta x\to 0$, получаем

$$\lim_{\Delta x\to 0}\frac{\Delta_x z}{\Delta x}=\lim_{\Delta x\to 0}[A+\alpha(\Delta x, 0)]=A=f_x'(x, y).$$

Следовательно, в точке M существует частная производная $f_x'(x, y)$.

Аналогично доказывается, что в точке M существует частная производная $f_y'(x, y)=B$ ■

Обратные утверждения к теоремам 12.1 и 12.2 неверны, т. е. из непрерывности функции двух переменных в точке M, а также из существования ее частных производных в этой точке еще не следует дифференцируемость функции.

Например, функция $f(x, y)=\sqrt{x^2+y^2}$ непрерывна в точке $(0; 0)$, но не имеет в этой точке частных производных. В самом деле,

$$f'_x(0,\ 0)=\lim_{\Delta x\to 0}\frac{\sqrt{(0+\Delta x)^2+0}-0}{\Delta x}=\lim_{\Delta x\to 0}\frac{|\Delta x|}{\Delta x}.$$

Функция $\frac{|\Delta x|}{\Delta x}$ не имеет предела при $\Delta x\to 0$. Следовательно, $f_x'(0,\ 0)$ не существует. Аналогично доказывается несуществование $f_y'(0,\ 0)$.

Функция

$$f(x,\ y)=\begin{cases}1 \text{ при } x\neq 0,\ y\neq 0,\\ 0 \text{ при } x=0 \text{ или } y=0\end{cases}$$

имеет частные производные по x и y в точке $(0;\ 0)$. Это следует из того, что $f(x,\ 0)\equiv 0$ и $f(0,\ y)\equiv 0$, и поэтому $f'_x(0,\ 0)=0$ и $f'_y(0,\ 0)=0$, но не является непрерывной в этой точке.

И последний пример: функция $f(x,\ y)=\sqrt{|x|\cdot|y|}$ непрерывна в точке $(0;\ 0)$, так как $f(0,\ 0)=0$ и $\lim\limits_{\substack{x\to 0\\ y\to 0}}f(x,\ y)=0$, и имеет частные производные по x и y в этой точке

$$f'_x(0,\ 0)=\lim_{\Delta x\to 0}\frac{\sqrt{|\Delta x|\cdot 0}-0}{\Delta x}=0,$$

$$f'_y(0,\ 0)=\lim_{\Delta y\to 0}\frac{\sqrt{0\cdot|\Delta y|}-0}{\Delta y}=0,$$

но тем не менее данная функция не является дифференцируемой в точке $(0;\ 0)$. Действительно, возьмем x равным y, тогда $f(x,\ x)=\sqrt{|x|\cdot|x|}=|x|$, а функция $|x|$, как известно, не дифференцируема в точке $(0;\ 0)$; отсюда следует недифференцируемость данной функции, так как из ее дифференцируемости следовала бы дифференцируемость функции $f(x,\ x)=|x|$ в точке $(0;0)$.

Таким образом, функция $f(x,\ y)=\sqrt{|x||y|}$ непрерывна в точке $(0;\ 0)$, имеет в этой точке частные производные и тем не менее не является дифференцируемой.

Следующая теорема дает достаточные условия дифференцируемости функции двух переменных.

3. Достаточные условия дифференцируемости.

Теорема 12.3. *Если функция $z=f(M)$ имеет частные производные в некоторой окрестности точки M и эти производные непрерывны в самой точке M, то эта функция дифференцируема в точке M.*

Доказательство. Дадим переменным x и y столь малые приращения Δx и Δy, чтобы точка $M_1(x+\Delta x,\ y+\Delta y)$ не выходила за пределы указанной окрестности точки M. Полное приращение

$$\Delta z=f(x+\Delta x,\ y+\Delta y)-f(x,\ y)$$

можно записать в виде

$$\Delta z=[f(x+\Delta x,\ y+\Delta y)-f(x,\ y+\Delta y)]+ \\ +[f(x,\ y+\Delta y)-f(x,\ y)]. \qquad (2)$$

Выражение $[f(x+\Delta x,\ y+\Delta y)-f(x,\ y+\Delta y)]$ можно рассматривать как приращение функции $f(x,\ y+\Delta y)$ одной переменной x на отрезке $[x,\ x+\Delta x]$ (или $[x+\Delta x,\ x]$, если $\Delta x<0$). Так как по условию эта функция имеет производную, совпадающую с $f'_x(x,\ y+\Delta y)$, то по теореме Лагранжа получим

$$f(x+\Delta x,\ y+\Delta y)-f(x,\ y+\Delta y)= \\ =f'_x(x+\theta_1\Delta x,\ y+\Delta y)\,\Delta x,\quad 0<\theta_1<1.$$

Рассуждая совершенно аналогично, для выражения $[f(x,\ y+\Delta y)-f(x,\ y)]$ получим

$$f(x,\ y+\Delta y)-f(x,\ y)=f'_y(x,\ y+\theta_2\Delta y)\,\Delta y,\qquad 0<\theta_2<1.$$

Так как производные f'_x и f'_y непрерывны в точке $M(x;\ y)$, то

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f'_x(x+\theta_1\Delta x,\ y+\Delta y)=f'_x(x,\ y),$$

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f'_y(x,\ y+\theta_2\Delta y)=f'_y(x,\ y),$$

т. е.

$$f'_x(x+\theta_1\Delta x,\ y+\Delta y)=f'_x(x,\ y)+\alpha(\Delta x,\ \Delta y),$$

$$f'_y(x,\ y+\theta_2\Delta y)=f'_y(x,\ y)+\beta(\Delta x,\ \Delta y),$$

где $\alpha(\Delta x, \Delta y)$ и $\beta(\Delta x, \Delta y)$ — бесконечно малые при $\Delta x \to 0$, $\Delta y \to 0$ функции. Подставляя эти значения в выражения (2) для Δz, находим

$$\Delta z = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y + \alpha(\Delta x, \Delta y)\,\Delta x + \\ + \beta(\Delta x, \Delta y)\,\Delta y,$$

а это и означает, что функция $z=f(M)$ дифференцируема в точке M ■

Следствие. Из непрерывности частных производных следует непрерывность самой функции.

Теорема 12.3 имеет важное значение для установления дифференцируемости функций. Дело в том, что непосредственная проверка дифференцируемости функции с помощью определения часто бывает затруднительна, в то время как проверка непрерывности частных производных оказывается проще.

В заключение заметим, что понятие дифференцируемости для функций трех и более переменных вводится аналогично понятию, рассмотренному для случая функции двух переменных.

§ 3. Производные сложных функций

Пусть $z=f(x, y)$ — функция двух переменных x и y, каждая из которых, в свою очередь, является функцией независимой переменной t: $x=x(t)$, $y=y(t)$. Тогда функция $z=f[x(t), y(t)]$ будет сложной функцией независимой переменной t, а переменные x и y будут для нее промежуточными переменными. Имеет место

Теорема 12.4. *Если функции $x=x(t)$ и $y=y(t)$ дифференцируемы в точке t, а функция $z=f(x, y)$ дифференцируема в точке $M[x(t), y(t)]$, то сложная функция $z=f[x(t), y(t)]$ также дифференцируема в точке t. При этом производная этой сложной функции определяется формулой*

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}. \qquad (1)$$

Доказательство. Придадим переменной t произвольное приращение Δt, оно вызовет соответственно приращения Δx и Δy функций $x(t)$ и $y(t)$, а они, в свою очередь, вызовут приращение функции $z=f(x, y)$:

$$\Delta z=f(x+\Delta x, y+\Delta y)-f(x, y).$$

Так как эта функция дифференцируема в точке $M(x; y)$, где $x=x(t)$, $y=y(t)$, то полное приращение этой функции может быть записано в виде

$$\Delta z = f'_x(x, y)\Delta x + f'_y(x, y)\Delta y + \alpha(\Delta x, \Delta y)\Delta x + \\ + \beta(\Delta x, \Delta y)\Delta y,$$

где $\alpha(\Delta x, \Delta y)$ и $\beta(\Delta x, \Delta y)$ — бесконечно малые при $\Delta x\to 0$, $\Delta y\to 0$ функции.

Разделив обе части равенства на Δt, получим

$$\frac{\Delta z}{\Delta t} = f'_x(x, y)\frac{\Delta x}{\Delta t} + f'_y(x, y)\frac{\Delta y}{\Delta t} + \\ + \alpha(\Delta x, \Delta y)\frac{\Delta x}{\Delta t} + \beta(\Delta x, \Delta y)\frac{\Delta y}{\Delta t}. \qquad (2)$$

По условию $\lim\limits_{\Delta t\to 0}\frac{\Delta x}{\Delta t} = \frac{dx}{dt}$, $\lim\limits_{\Delta t\to 0}\frac{\Delta y}{\Delta t} = \frac{dy}{dt}$. Кроме того, так как функции $x(t)$ и $y(t)$ дифференцируемы в точке t, то они непрерывны в этой точке, т. е. при $\Delta t\to 0$ будет $\Delta x\to 0$, $\Delta y\to 0$ и, как следствие, $\alpha(\Delta x, \Delta y)\to 0$ и $\beta(\Delta x, \Delta y)\to 0$. Слагаемые

$$\alpha(\Delta x, \Delta y)\frac{\Delta x}{\Delta t} \text{ и } \beta(\Delta x, \Delta y)\frac{\Delta y}{\Delta t}$$

будут стремиться к нулю при $\Delta t\to 0$.

Таким образом, мы доказали, что при $\Delta t\to 0$ существует предел выражения, стоящего в правой части равенства (2), а следовательно, существует предел левой части при $\Delta t\to 0$:

$$\lim_{\Delta t\to 0}\frac{\Delta z}{\Delta t} = \frac{dz}{dt},$$

причем

$$\frac{dz}{dt} = f'_x(x, y)\frac{dx}{dt} + f'_y(x, y)\frac{dy}{dt}, \text{ или, короче, } \frac{dz}{dt} = \\ = \frac{\partial z}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial z}{\partial y}\cdot\frac{dy}{dt} \blacksquare$$

Замечание. Необходимо обратить внимание на то, когда пишется «∂» в обозначениях производных и когда «d».

Примеры.

1. Пусть $z=f(x, y)$, $x=t^3+2$, $y=3t^4=1$. По формуле (1) имеем

$$\frac{dz}{dt}=\frac{\partial z}{\partial x}\cdot 3t^2+\frac{\partial z}{\partial y}\cdot 12t^3.$$

2. Пусть $z=x\sin\frac{x}{y}$, $x=1+3t$, $y=\sqrt{1+t^2}$.

По формуле (1) имеем

$$\frac{dz}{dt}=\frac{\partial z}{\partial x}\cdot\frac{dx}{dt}+\frac{\partial z}{\partial y}\cdot\frac{dy}{dt}=\left(\sin\frac{x}{y}+\frac{x}{y}\cos\frac{x}{y}\right)\cdot 3-$$
$$-\frac{x^2}{y^2}\cos\frac{x}{y}\cdot\frac{t}{\sqrt{1+t^2}}.$$

3. Пусть $z=x^2y^3$, $x=t$, $y=t^2$. По формуле (1) имеем

$$\frac{dz}{dt}=2xy^3\cdot 1+3x^2y^2\cdot 2t.$$

В частности, если $z=f(x, y)$, где $y=\varphi(x)$, то $z=f[x, \varphi(x)]$ является сложной функцией x, зависящей от переменной x явно и через посредство переменной y. На основании формулы (1) получим

$$\frac{dz}{dx}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial x}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial x},$$

но так как $y=\varphi(x)$ — функция только одной переменной x, то частная производная $\frac{\partial y}{\partial x}$ обращается в обыкновенную. Кроме того, $\frac{\partial x}{\partial x}=1$, поэтому

$$\frac{dz}{dx}=\frac{\partial z}{\partial x}+\frac{\partial z}{\partial y}\cdot\frac{dy}{dx}.$$

Аналогичным образом решается вопрос, когда число промежуточных переменных больше двух. Например, если $u=f(x, y, z)$, где $x=x(t)$, $y=y(t)$, $z=z(t)$, то формула (1) принимает вид

$$\frac{du}{dt}=\frac{\partial u}{\partial x}\cdot\frac{dx}{dt}+\frac{\partial u}{\partial y}\cdot\frac{dy}{dt}+\frac{\partial u}{\partial z}\cdot\frac{dz}{dt}.$$

Рассмотрим теперь более общий случай. Пусть $z=f(x, y)$ — функция двух переменных x и y, которые, в свою очередь, зависят от двух или большего числа переменных. Например, пусть $x=x(u, v)$, $y=y(u, v)$. Тогда функция $z=f[x(u, v), y(u, v)]$ будет сложной функцией независимых переменных u и v.

Если функции $x(u, v)$ и $y(u, v)$ дифференцируемы в точке $M'(u;v)$, а функция $z=f(x, y)$ дифференцируема в точке $M(x; y)$, где $x=x(u, v)$, $y=y(u, v)$, то сложная функция $z=f[x(u, v), y(u, v)]$ дифференцируема в точке $M'(u; v)$, причем ее частные производные в этой точке находятся по формулам

$$\begin{cases}\dfrac{\partial z}{\partial u}=\dfrac{\partial z}{\partial x}\cdot\dfrac{\partial x}{\partial u}+\dfrac{\partial z}{\partial y}\cdot\dfrac{\partial y}{\partial u},\\[2mm] \dfrac{\partial z}{\partial v}=\dfrac{\partial z}{\partial x}\cdot\dfrac{\partial x}{\partial v}+\dfrac{\partial z}{\partial y}\cdot\dfrac{\partial y}{\partial v}.\end{cases} \tag{3}$$

Действительно, для вычислении частной производной фиксируем значение одной из переменных u или v. Тогда попадаем в условия только что доказанной теоремы, и из формулы (1) получаем формулы (3).

Примеры:

1. Пусть $z=f(x, y)$ $x=u^2+2v$, $y=\frac{u^2}{v}$. По формулам (3) имеем

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\cdot 2u+\frac{\partial z}{\partial y}\cdot\frac{2u}{v};\quad \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\cdot 2+$$
$$+\frac{\partial z}{\partial y}\cdot\left(-\frac{u^2}{v^2}\right).$$

2. Пусть $z=x^2y^2$, $x=u+v$, $y=\frac{u}{v}$. По формулам (3) имеем

$$\frac{\partial z}{\partial u}=2xy^2\cdot 1+2x^2y\cdot\frac{1}{v};\quad \frac{\partial z}{\partial v}=2xy^2\cdot 1+$$
$$+2x^2y\cdot\left(-\frac{u}{v^2}\right).$$

3. Пусть $z=x^2-y^2$, $x=u\cos v$, $y=u\sin v$. По формулам (3) имеем

$$\frac{\partial z}{\partial u}=2x\cos v-2y\sin v;\quad \frac{\partial z}{\partial v}=-2xu\sin v-2yu\cos v.$$

В частности, если $z=f(x)$, где $x=x(u, v)$, то $z=f[x(u, v)]$ является сложной функцией, зависящей через посредство переменной x от двух переменных u и v, и ее частные производные также находятся по формулам (3):

$$\begin{cases} \dfrac{\partial z}{\partial u}=\dfrac{dz}{dx}\cdot\dfrac{\partial x}{\partial u}, \\ \dfrac{\partial z}{\partial v}=\dfrac{dz}{dx}\cdot\dfrac{\partial x}{\partial v}. \end{cases}$$

Следует обратить внимание на обозначения производных.

Для случая большого числа переменных формулы (3) естественным образом обобщаются.

Например, если $w=f(x, y, z)$ есть функция трех переменных x, y, z, а каждая из них зависит от u и v, то формулы (3) принимают вид

$$\begin{cases} \dfrac{\partial w}{\partial u}=\dfrac{\partial w}{\partial x}\cdot\dfrac{\partial x}{\partial u}+\dfrac{\partial w}{\partial y}\cdot\dfrac{\partial y}{\partial u}+\dfrac{\partial w}{\partial z}\cdot\dfrac{\partial z}{\partial u}, \\ \dfrac{\partial w}{\partial v}=\dfrac{\partial w}{\partial x}\cdot\dfrac{\partial x}{\partial v}+\dfrac{\partial w}{\partial y}\cdot\dfrac{\partial y}{\partial v}+\dfrac{\partial w}{\partial z}\cdot\dfrac{\partial z}{\partial v}. \end{cases}$$

§ 4. Дифференциал

1. Определение. Напомним, что если функция $z=f(M)$ дифференцируема в точке M, то ее полное приращение в этой точке может быть представлено в виде

$$\Delta z=A\Delta x+B\Delta y+\alpha(\Delta x, \Delta y)\Delta x+\beta(\Delta x, \Delta y)\Delta y, \quad (1)$$

где $\alpha(\Delta x, \Delta y)$ и $\beta(\Delta x, \Delta y)$ — бесконечно малые при $\Delta x\to 0$, $\Delta y\to 0$ функции.

Определение. *Дифференциалом dz дифференцируемой в точке M функции $z=f(M)$ называется глав-*

ная линейная относительно приращений Δx и Δy часть полного приращения этой функции в точке M, т. е.

$$dz = A\Delta x + B\Delta y. \tag{2}$$

Если A и B равны нулю, то дифференциал dz функции в точке M считается равным нулю. Используя теорему 12.2, мы можем, очевидно, переписать выражения (2) следующим образом:

$$dz = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y.$$

Если положить $z = x$ (т. е. $f(x, y) \equiv x$), то $dz = dx = 1\cdot\Delta x + 0\cdot\Delta y$, т. е. $dx = \Delta x$. Аналогично, полагая $z = y$, получим, что $dy = \Delta y$. Отсюда заключаем, что дифференциалы независимых переменных совпадают с приращениями этих переменных, и мы можем записать дифференциал (2) в следующем виде:

$$dz = f'_x(x, y)\,dx + f'_y(x, y)\,dy.$$

Из соотношений (1) и (2) следует, что разность между полным приращением и дифференциалом функции в точке M

$$\Delta z - dz = \alpha(\Delta x, \Delta y)\Delta x + \beta(\Delta x, \Delta y)\Delta y$$

есть бесконечно малая более высокого порядка, чем $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$, при $\Delta x \to 0$, $\Delta y \to 0$ ($\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ есть расстояние между точками $M(x; y)$ и $M_1(x + \Delta x; y + \Delta y)$). Действительно,

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \frac{\Delta z - dz}{\rho} = \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \left(\alpha\frac{\Delta x}{\rho} + \beta\frac{\Delta y}{\rho}\right) = 0,$$

так как α и β — бесконечно малые, а $\frac{\Delta x}{\rho}$ и $\frac{\Delta y}{\rho}$ — ограниченные величины $\left(\left|\frac{\Delta x}{\rho}\right| \leqslant 1,\ \left|\frac{\Delta y}{\rho}\right| \leqslant 1\right)$, т. е. получаем

$$\Delta z = dz + 0\left(\sqrt{(\Delta x)^2 + (\Delta y)^2}\right). \tag{3}$$

При достаточно малых Δx и Δy допускаемая погрешность может быть сколь угодно малой, поэтому можно принять

$$\Delta z \approx dz.$$

Этим широко пользуются в приближенных вычислениях, так как дифференциал легче подсчитать, чем полное приращение.

2. Касательная плоскость и нормаль к поверхности. Геометрический смысл дифференциала. Аналогично тому, как дифференциал функции одной переменной геометрически представляет собой приращение ординаты касательной, дифференциал функции двух переменных есть приращение аппликаты касательной плоскости. Введем понятие касательной плоскости к поверхности в точке N_0.

Плоскость, проходящая через точку N_0 поверхности, называется касательной плоскостью в этой точке, если угол между секущей прямой, проходящей через точку N_0 и любую точку N поверхности, и ее проекцией на плоскость стремится к нулю, когда точка N стремится к точке N_0.

Покажем, что касательная плоскость к поверхности в точке $N_0(x_0;\ y_0;\ z_0)$, заданной уравнением $z=f(x,\ y)$, где $f(x,\ y)$ — функция, дифференцируемая в точке $M_0(x_0;\ y_0)$, определяется уравнением

$$z - z_0 = f'_x(x_0,\ y_0)(x - x_0) + f'_y(x_0,\ y_0)(y - y_0). \qquad (4)$$

Действительно, из аналитической геометрии известно, что уравнение вида (4) определяет в некоторой прямоугольной системе координат плоскость, проходящую через точку $N_0(x_0;\ y_0;\ z_0)$ и имеющую нормальный вектор $\bar{n}=\{f'_x;\ f'_y;\ -1\}$. Осталось установить, что плоскость является касательной плоскостью. Для этого, очевидно, надо показать, что угол φ между вектором $\bar{n}$ этой плоскости и любой секущей N_0N стремится к $\frac{\pi}{2}$, когда точка N стремится к N_0. Так как координаты вектора $\bar{n}$ равны f'_x, f'_y, -1, а координаты вектора $\overline{N_0N}$ секущей равны $x-x_0$, $y-y_0$, $z-z_0$, то

$$\cos\varphi = \frac{f'_x\cdot(x - x_0) + f'_y\cdot(y - y_0) - (z - z_0)}{\sqrt{f'^2_x + f'^2_y + 1}\,\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}}.$$

Но, как следует из соотношения (3),

$$f'_x\cdot(x-x_0)+f'_y\cdot(y-y_0)-(z-z_0)=o(\rho),\ \text{где}\ \rho=$$
$$=\sqrt{(x-x_0)^2+(y-y_0)^2}.$$

Поэтому

$$|\cos\varphi|\leqslant\frac{|o(\rho)|}{\sqrt{(x-x_0)^2+(y-y_0)^2}}=\frac{|o(\rho)|}{\rho}\to 0,\quad \text{когда}\quad \rho\to 0,$$

$$\text{т. е.}\ \lim_{\rho\to 0}\varphi=\frac{\pi}{2},$$

что и требовалось показать.

Нормальный вектор $\bar{n}=\{f'_x;\ f'_y;\ -1\}$ касательной плоскости принято называть нормалью к поверхности $z=f(x,\ y)$ в точке N_0.

Обозначив в (4) $x-x_0=\Delta x$, $y-y_0=\Delta y$, $z-z_0=\Delta z$, видим, что приращение Δz аппликаты касательной плоскости дается формулой

$$\Delta z=f'_x(x_0,\ y_0)\,\Delta x+f'_y(x_0,\ y_0)\,\Delta y,$$

т. е. действительно совпадает с дифференциалом dz функции $z=f(x,\ y)$.

§ 5. Производная по направлению. Градиент

Рассмотрим функцию $z=f(M)$, определенную на некотором множестве $\{M\}$, содержащем точку $M(x;\ y)$, и единичный вектор $\bar{n}=\{\cos\alpha;\ \cos\beta\}$ произвольного направления (рис. 44).

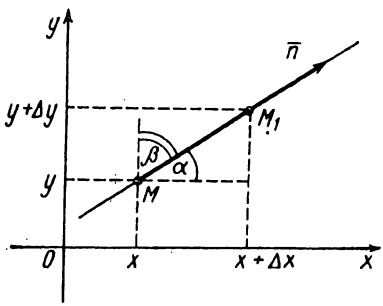


Рис. 44

Для характеристики скорости изменения функции в точке $M(x;\ y)$ в направлении вектора $\bar{n}$ введем понятие *производной по направлению*.

С этой целью проведем через точку M прямую параллельно вектору $\bar{n}$ и возьмем на ней точку $M_1(x+\Delta x;\ y+\Delta y)$. Обозначим MM_1 через l; $l=\sqrt{(\Delta x)^2+(\Delta y)^2}$, если точка M_1 расположена так, как показано на рис. 44,

$l=-\sqrt{(\Delta x)^2+(\Delta y)^2}$, если точка M_1 расположена по другую сторону от точки M. Функция $f(M)$ получит при этом полное приращение

$$\Delta z=f(x+\Delta x,\ y+\Delta y)-f(x,\ y).$$

Определение 1. *Предел отношения* $\frac{\Delta z}{l}$ *при* $l\to 0$ $(M_1\to M)$, *если он существует, называется производной функции* $z=f(M)$ *в точке* $M(x;\ y)$ *по направлению вектора* $\bar{n}$ *и обозначается* $\frac{\partial z}{\partial n}$, *т. е.*

$$\lim_{l\to 0}\frac{\Delta z}{l}=\frac{\partial z}{\partial n}.$$

Предположим теперь, что функция $f(M)$ дифференцируема в точке M. Тогда ее полное приращение в этой точке можем записать в виде

$$\Delta z=f'_x(x,\ y)\,\Delta x+f'_y(x,\ y)\,\Delta y+\alpha(\Delta x,\ \Delta y)\,\Delta x+$$
$$+\beta(\Delta x,\ \Delta y)\,\Delta y,$$

где α и β — бесконечно малые при $l\to 0$ функции. Деля обе части равенства на l и учитывая, что $\Delta x=l\cos\alpha$, $\Delta y=l\sin\alpha=l\cos\beta$, получим

$$\frac{\Delta z}{l}=f'_x(x,\ y)\cos\alpha+f'_y(x,\ y)\cos\beta+\alpha(\Delta x,\ \Delta y)\cos\alpha+$$
$$+\beta(\Delta x,\ \Delta y)\cos\beta.$$

Переходя в последнем равенстве к пределу при $l\to 0$, получаем формулу для производной по направлению:

$$\frac{\partial z}{\partial n}=\frac{\partial z}{\partial x}\cos\alpha+\frac{\partial z}{\partial y}\cos\beta. \qquad (1)$$

Из формулы (1) следует, что производная по направлению является линейной комбинацией частных производных, причем направляющие косинусы являются как бы весовыми множителями, показывающими вклад в производную по направлению соответствующей частной производной.

В частности, при

$$\alpha=0,\ \beta=\frac{\pi}{2}\ \frac{\partial z}{\partial n}=\frac{\partial z}{\partial x},\ \text{ а при } \alpha=\frac{\pi}{2},\ \beta=0\ \frac{\partial z}{\partial n}=\frac{\partial z}{\partial y}.$$

Отсюда видно, что частные производные по x и y являются частными случаями производной по направлению.

Пример. Вычислить производную функции $z=x^2+y^2x$ в точке $M(1;\ 2)$ по направлению вектора $\overline{MM_1}$, где M_1 — точка $M_1(3;\ 0)$.

Решение. Находим единичный вектор $\bar{n}$ данного направления:

$$\overline{MM_1}=\{2;\ -2\}=2\bar{i}-2\bar{j};\quad |\overline{MM_1}|=2\sqrt{2};$$

$$\bar{n}=\frac{\overline{MM_1}}{|\overline{MM_1}|}=\frac{2\bar{i}-2\bar{j}}{2\sqrt{2}}=\frac{1}{\sqrt{2}}\bar{i}-\frac{1}{\sqrt{2}}\bar{j}.$$

Отсюда

$$\cos\alpha=\frac{1}{\sqrt{2}},\quad \cos\beta=-\frac{1}{\sqrt{2}}.$$

Вычисляем частные производные функции в точке $M(1;\ 2)$:

$$f'_x(x,\ y)=2x+y^2,\ f'_y(x,\ y)=2xy,$$

т. е.

$$f'_x(1;\ 2)=6,\ f'_y(1;\ 2)=4.$$

Следовательно, по формуле (1) получаем

$$\frac{\partial z}{\partial n}=6\cdot\frac{1}{\sqrt{2}}-4\cdot\frac{1}{\sqrt{2}}=\sqrt{2}.$$

Введем понятие градиента дифференцируемой в точке $M(x;\ y)$ функции $z=f(M)$.

Определение 2. *Градиентом функции $z=f(M)$ в точке $M(x;\ y)$ называется вектор, проекции которого на координатные оси совпадают с соответствующими частными производными $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, взятыми в точке $M(x;\ y)$; он обозначается*

$$\operatorname{grad} z=\left\{\frac{\partial z}{\partial x};\ \frac{\partial z}{\partial y}\right\}.$$

Используя понятие градиента функции и учитывая, что вектор $\bar{n}$ имеет координаты $\cos\alpha$, $\cos\beta$, представим формулу (1) в виде скалярного произведения векторов $\operatorname{grad} z$ и $\bar{n}$:

$$\frac{\partial z}{\partial n} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta = \operatorname{grad} z\cdot\bar{n}. \qquad (2)$$

С другой стороны, по определению скалярного произведения имеем

$$\operatorname{grad} z\cdot\bar{n} = |\operatorname{grad} z|\cdot|\bar{n}|\cdot\cos\varphi, \qquad (3)$$

где $|\operatorname{grad} z|$ — длина вектора $\operatorname{grad} z$, φ — угол между векторами $\bar{n}$ и $\operatorname{grad} z$. Сопоставляя (2) и (3) и учитывая, что $|\bar{n}|=1$, получаем

$$\frac{\partial z}{\partial n} = |\operatorname{grad} z|\cdot\cos\varphi.$$

Из последнего равенства следует, что производная функции по направлению будет наибольшей по величине при $\cos\varphi=1$ ($\varphi=0$), т. е. когда направление вектора $\bar{n}$ совпадает с направлением $\operatorname{grad} z$, при этом

$$\frac{\partial z}{\partial n} = |\operatorname{grad} z|.$$

Таким образом, градиент функции $z=f(M)$ в точке $M(x;\ y)$ характеризует направление и величину максимального роста этой функции в точке $M(x;\ y)$.

Аналогичным образом определяется производная по направлению для функции трех переменных $u=f(x,\ y,\ z)$ и выводится формула

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma,$$

вводится понятие градиента

$$\operatorname{grad} u = \left\{\frac{\partial u}{\partial x};\ \frac{\partial u}{\partial y};\ \frac{\partial u}{\partial z}\right\}$$

и исследуются его свойства.

Понятия производной по направлению и градиента функции играют важную роль во многих приложениях.

§ 6. Частные производные и дифференциалы высших порядков

1. Частные производные высших порядков. Пусть частные производные $f'_x(x, y)$ и $f'_y(x, y)$ функции $z=f(M)$, определенной на некотором множестве $\{M\}$, существуют в каждой точке множества $\{M\}$. В этом случае частные производные представляют собой функции двух переменных x и y, определенные на $\{M\}$. Назовем их *частными производными 1-го порядка*.

В свою очередь, частные производные по переменным x и y от функций $f'_x(x, y)$ и $f'_y(x, y)$ в точке $M\in\{M\}$, если они существуют, называются *частными производными 2-го порядка* от функции $f(M)$ в этой точке и обозначаются следующими символами:

$$\frac{\partial^2 z}{\partial x^2}=f''_{xx}(x, y)=f^{(2)}_{x^2}(x, y);\quad \frac{\partial^2 z}{\partial y\,\partial x}=f''_{yx}(x, y)=f^{(2)}_{yx}(x, y);$$

$$\frac{\partial^2 z}{\partial y^2}=f''_{yy}(x, y)=f^{(2)}_{y^2}(x, y);\quad \frac{\partial^2 z}{\partial x\,\partial y}=f''_{xy}(x, y)=f^{(2)}_{xy}(x, y).$$

Частные производные 2-го порядка вида $f''_{yx}(x, y)$, $f''_{xy}(x, y)$ называются *смешанными частными производными*.

Примеры:

1. $z=x^4+4x^2y^3+7xy+1$. Имеем

$$\frac{\partial z}{\partial x}=4x^3+8xy^3+7y;\quad \frac{\partial z}{\partial y}=12x^2y^2+7x.$$

Следовательно,

$$\frac{\partial^2 z}{\partial x^2}=12x^2+8y^3,\quad \frac{\partial^2 z}{\partial x\,\partial y}=\frac{\partial^2 z}{\partial y\,\partial x}=24xy^2+7,$$

$$\frac{\partial^2 z}{\partial y^2}=24x^2y.$$

2. $z=\sin x\cos y$. Имеем

$$\frac{\partial z}{\partial x}=\cos x\cos y,\quad \frac{\partial z}{\partial y}=-\sin x\sin y.$$

Следовательно,

$$\frac{\partial^2 z}{\partial x^2} = -\sin x \cos y, \quad \frac{\partial^2 z}{\partial x\, \partial y} = \frac{\partial^2 z}{\partial y\, \partial x} = -\cos x \sin y,$$

$$\frac{\partial^2 z}{\partial y^2} = -\sin x \cos y.$$

В обоих примерах мы получили равенство смешанных частных производных $f''_{yx}(x, y)$ и $f''_{xy}(x, y)$. Но, вообще говоря, значения смешанных производных зависят от порядка, в котором производится дифференцирование. Так, например, функция

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{при } x^2 + y^2 \neq 0, \\ 0 & \text{при } x^2 + y^2 = 0 \end{cases}$$

в точке (0; 0) имеет смешанные частные производные $f''_{yx}(x, y)$ и $f''_{xy}(x, y)$, не равные друг другу. Действительно,

$$f'_x(x, y) = \begin{cases} \dfrac{y(x^4 - y^4 + 4x^2y^2)}{(x^2 + y^2)^2} & \text{при } x^2 + y^2 \neq 0, \\ 0 & \text{при } x^2 + y^2 = 0. \end{cases}$$

Следовательно,

$$f''_{yx}(0;\ 0) = \lim_{\Delta y \to 0} \frac{f'_x(0,\ 0 + \Delta y) - f'_x(0,\ 0)}{\Delta y} = -1.$$

Проводя аналогичные вычисления, получим $f''_{xy}(0, 0) = 1$. Таким образом, в точке (0; 0) $f''_{yx}(x, y) \neq f''_{xy}(x, y)$.

Спрашивается, при каких условиях значения смешанных производных какого-либо порядка не зависят от того, в каком порядке производится дифференцирование? Ответ дает следующая теорема.

Теорема 12.5. *Если в некоторой окрестности точки M производные $f''_{xy}(x, y)$ и $f''_{yx}(x, y)$ существуют и непрерывны в самой точке M, то они равны между собой в этой точке, т. е. имеет место следующее равенство:*

$$f''_{xy}(x, y) = f''_{yx}(x, y).$$

Доказательство. Рассмотрим выражение

$$A=[f(x+\Delta x,\ y+\Delta y)-f(x+\Delta x,\ y)]-$$
$$-[f(x,\ y+\Delta y)-f(x,\ y)],$$

где Δx и Δy — любые столь малые числа, что точка $M_1(x+\Delta x;\ y+\Delta y)$ находится в указанной окрестности точки M.

Введем вспомогательную функцию

$$\varphi(x)=f(x,\ y+\Delta y)-f(x,\ y),$$

тогда выражение A можно рассматривать как приращение

$$A=\Delta\varphi=\varphi(x+\Delta x)-\varphi(x)$$

дифференцируемой на отрезке $[x,\ x+\Delta x]$ функции $\varphi(x)=f(x,\ y+\Delta y)-f(x,\ y)$ одной переменной x. Поэтому, применяя к этой разности теорему Лагранжа, можем записать:

$$A=\Delta\varphi=\varphi'(x+\theta_1\Delta x)\,\Delta x=[f'_x(x+\theta_1\Delta x,\ y+\Delta y)-$$
$$-f'_x(x+\theta_1\Delta x,\ y)]\,\Delta x,\quad 0<\theta_1<1.$$

Выражение в квадратных скобках можно рассматривать как приращение дифференцируемой на отрезке $[y,\ y+\Delta y]$ функции $f'_x(x+\theta_1\Delta x,\ y)$ одной переменной y. Применяя еще раз теорему Лагранжа (по переменному y), получаем

$$A=f''_{yx}(x+\theta_1\Delta x,\ y+\theta_2\Delta y)\,\Delta x\,\Delta y,\quad 0<\theta_1,\ \theta_2<1. \qquad (1)$$

С другой стороны, если бы мы с самого начала положили

$$\psi(y)=f(x+\Delta x,y)-f(x,\ y),$$

то, поступая совершенно аналогично, получили бы

$$A=\Delta\psi=\psi(y+\Delta y)-\psi(y),$$

а затем

$$A=f''_{xy}(x+\theta_4\Delta x,\ y+\theta_3\Delta y)\,\Delta y\,\Delta x,\quad 0<\theta_3,\ \theta_4<1. \qquad (2)$$

Сравнивая (1) и (2), получаем

$$f''_{yx}(x+\theta_1\Delta x,\ y+\theta_2\Delta y)=f''_{xy}(x+\theta_4\Delta x,\ y+\theta_3\Delta y).$$

Переходя теперь в этом равенстве к пределу при $\Delta x \to 0$ и $\Delta y \to 0$

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f''_{yx}(x+\theta_1\Delta x,\ x+\theta_2\Delta y) = \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f''_{xy}(x+\theta_4\Delta x,\ y+\theta_3\Delta y)$$

и учитывая непрерывность частных производных $f''_{yx}(x, y)$, $f''_{xy}(x, y)$ в точке M, окончательно получим

$$f''_{xy}(x, y) = f''_{yx}(x, y) \blacksquare$$

Аналогично частным производным 2-го порядка вводятся *частные производные 3-го, 4-го, ..., n-го порядков* с соответствующей символикой обозначений и доказывается теорема типа 12.5 для смешанных производных любого порядка.

2. Дифференциалы высших порядков. В § 4 мы ввели понятие дифференциала дифференцируемой в точке M функции $z=f(M)$ и получили формулу

$$dz = f'_x(x, y)\,dx + f'_y(x, y)\,dy. \qquad (3)$$

Будем называть dz *дифференциалом 1-го порядка*. Для удобства будем обозначать дифференциалы не только символом d, но и символом δ (например, δx, δy).

Пусть функции $f'_x(x, y)$ и $f'_y(x, y)$ дифференцируемы в точке M. Будем рассматривать dx и dy в выражении для dz как постоянные множители. Тогда функция dz представляет собой функцию только переменных x и y, дифференцируемую в точке M, и ее дифференциал имеет вид

$$\delta(dz) = \delta[f'_x(x, y)\,dx + f'_y(x, y)\,dy].$$

Дифференциал $\delta(dz)$ от дифференциала dz в точке M, взятый при $\delta x=dx$, $\delta y=dy$, называется *дифференциалом 2-го порядка* функции $f(M)$ и обозначается d^2z. В свою очередь, дифференциал $\delta(d^2z)$ от d^2z, взятый при $\delta x=dx$, $\delta y=dy$, называется *дифференциалом 3-го порядка* функции $f(M)$ и обозначается d^3z и т. д. Дифференциал $\delta(d^{n-1}z)$ от дифференциала $d^{n-1}z$, взятый при $\delta x=dx$, $\delta y=dy$, называется *дифференциалом n-го порядка* (*или n-м дифференциалом*) функции $f(M)$ и обозначается d^nz.

Таким образом, для n-го дифференциала функции $f(M)$ справедлива формула

$$d^n z = \delta(d^{n-1}z)\Big|_{\substack{\delta x = dx\\ \delta y = dy}}$$

При конкретном вычислении второго и последующих дифференциалов обычно вычисление $\delta(dz)$ и приравнивание дифференциалов аргументов ($\delta x = dx$, $\delta y = dy$) при последовательных дифференцированиях совмещаются.

С помощью формулы (3) найдем выражение для дифференциала 2-го порядка:

$$d^2z = d(dz) = (f'_x dx + f'_y dy)'_x dx + (f'_x dx + f'_y dy)'_y dy =$$
$$= f''_{xx}(dx)^2 + f''_{xy}\,dx\,dy + f''_{yx}\,dy\,dx + f''_{yy}(dy)^2.$$

Предполагается, что f''_{xy} и f''_{yx} непрерывны, и поэтому, согласно теореме 12.5, слагаемые $f''_{xy}\,dx\,dy$ и $f''_{yx}\,dy\,dx$ равны, так что

$$d^2z = f''_{xx}(dx)^2 + 2f''_{xy}\,dx\,dy + f''_{yy}(dy)^2.$$

Аналогично

$$d^3z = f'''_{x^3}(dx)^3 + 3f'''_{x^2y}(dx)^2\,dy + 3f'''_{xy^2}\,dx\,(dy)^2 + f'''_{y^3}(dy)^3,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots,$$

$$d^nz = f^{(n)}_{x^n}(dx)^n + nf^{(n)}_{x^{n-1}y}(dx)^{n-1}\cdot dy + \dots + f^{(n)}_{y^n}(dy)^n.$$

Формула для любого d^nz напоминает формулу для разложения двучлена в n-ю степень по правилу бинома Ньютона. Поэтому выражение для d^nz символически можно записать в виде, более удобном для запоминания:

$$d^nz = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^n f(x,\ y).$$

Примеры:

1. Найти d^2z для функции $z = \operatorname{arctg}\frac{x}{y}$. Имеем

$$f'_x = \frac{y}{x^2+y^2},\quad f'_y = -\frac{x}{x^2+y^2},\quad f''_{xy} = \frac{x^2-y^2}{(x^2+y^2)^2} = f''_{yx},$$
$$f''_{x^2} = -\frac{2xy}{(x^2+y^2)^2},\quad f''_{y^2} = \frac{2xy}{(x^2+y^2)^2}.$$

Следовательно,

$$d^2z = \frac{-2xy(dx)^2 + 2(x^2 - y^2)\,dx\,dy + 2xy(dy)^2}{(x^2+y^2)^2}.$$

2. Найти d^3z для функции $z=\sin x\cos y$. Имеем

$$f'_x = \cos x\cos y,\ f'_y = -\sin x\sin y,\ f''_{x^2} = -\sin x\cos y,$$

$$f''_{y^2} = -\sin x\cos y,\ f''_{xy} = -\cos x\sin y,$$

$$f'''_{x^3} = -\cos x\cos y,\ f'''_{y^3} = \sin x\sin y,\ f'''_{x^2y} = \sin x\sin y,$$

$$f'''_{xy^2} = -\cos x\cos y.$$

Следовательно,

$$d^3z = -\cos x\cos y\,(dx)^3 + 3\sin x\sin y\,(dx)^2\,dy -$$

$$-3\cos x\cos y\,dx\,(dy)^2 + \sin x\sin y\,(dy)^3.$$

§ 7. Формула Тейлора для функции двух переменных

Аналогично функции одной переменной функцию двух переменных можно также представить в виде суммы многочлена n-го порядка по степеням Δx и Δy и некоторого остаточного члена. Докажем следующую теорему.

Теорема 12.6. *Пусть функция $z=f(x, y)$ непрерывна вместе со всеми частными производными до $(n+1)$-го порядка включительно в точке $M(x;\ y)$ и некоторой ее окрестности. Пусть точка $M_1(x+\Delta x;\ y+\Delta y)$ принадлежит этой окрестности. Тогда полное приращение $\Delta f=f(M_1)-f(M)$ этой функции в точке M может быть представлено в следующей форме:*

$$\Delta f = df(x, y) + \frac{d^2f(x,y)}{2!} + \ldots + \frac{d^nf(x,y)}{n!} +$$

$$+\frac{d^{n+1}f(x+\theta\Delta x, y+\theta\Delta y)}{(n+1)!},\ 0<\theta<1. \qquad (1)$$

Формула (1) называется *формулой Тейлора* для функции $z=f(x, y)$.

Доказательство. Для доказательства введем вспомогательную функцию

$$F(t)=f(x+t\Delta x,\ y+t\Delta y),$$

являющуюся сложной функцией новой независимой переменной t, изменяющейся в пределах от 0 до 1 ($t=0$ отвечает точка M, а $t=1$ — точка M_1), и имеющую $(n+1)$-ю производную по t на отрезке $[0, 1]$.

Дифференцируя функцию $F(t)$ по t, получаем

$$F'(t)=f'_x(x+t\Delta x, y+t\Delta y)\,\Delta x+f'_y(x+t\Delta x, y+t\Delta y)\,\Delta y=$$

$$=\left(\frac{\partial}{\partial x}\Delta x+\frac{\partial}{\partial y}\Delta y\right)f(x+t\Delta x, y+t\Delta y),$$

$$F''(t)=f''_{xx}(x+t\Delta x, y+t\Delta y)(\Delta x)^2+2f''_{xy}(x+t\Delta x, y+$$

$$+t\Delta y)\,\Delta x\Delta y+f''_{yy}(x+t\Delta x, y+t\Delta y)(\Delta y)^2=$$

$$=\left(\frac{\partial}{\partial x}\Delta x+\frac{\partial}{\partial y}\Delta y\right)^2 f(x+t\Delta x, y+t\Delta y).$$

По индукции найдем

$$F^{(n)}(t)=\left(\frac{\partial}{\partial x}\Delta x+\frac{\partial}{\partial y}\Delta y\right)^n f(x+t\Delta x, y+t\Delta y),$$

$$F^{(n+1)}(t)=\left(\frac{\partial}{\partial x}\Delta x+\frac{\partial}{\partial y}\Delta y\right)^{n+1} f(x+t\Delta x, y+t\Delta y).$$

С другой стороны, применяя к функции $F(t)$ как функции одной переменной t формулу Маклорена (см. п. 3, § 3, гл. VI) и полагая $t=1$, получим

$$F(1)=F(0)+\frac{F'(0)}{1!}+\frac{F''(0)}{2!}+$$

$$+\ldots+\frac{F^{(n)}(0)}{n!}+\frac{F^{(n+1)}(\theta)}{(n+1)!},\ 0<\theta<1.$$

Но

$$F(1)=f(x+\Delta x, y+\Delta y),$$

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

$$F'(0)=\left(\frac{\partial}{\partial x}\Delta x+\frac{\partial}{\partial y}\Delta y\right)f(x, y),$$

$$F''(0)=\left(\frac{\partial}{\partial x}\Delta x+\frac{\partial}{\partial y}\Delta y\right)^2 f(x, y),$$

. . . .

$$F^{(n)}(0) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^n f(x, y),$$

$$F^{(n+1)}(\theta) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^{n+1} f(x + \theta\Delta x, y + \theta\Delta y),$$

$$0 < \theta < 1.$$

Следовательно,

$$f(x + \Delta x, y + \Delta y) = f(x, y) + \frac{1}{1!}\left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right) \times$$

$$\times f(x, y) + \frac{1}{2!}\left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^2 f(x, y) + \ldots +$$

$$+ \ldots + \frac{1}{n!}\left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^n f(x, y) +$$

$$+ \frac{1}{(n+1)!}\left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^{n+1} f(x + \theta\Delta x, y + \theta\Delta y),$$

$$0 < \theta < 1,$$

отсюда получаем

$$\Delta f = df(x, y) + \frac{d^2 f(x, y)}{2!} + \ldots + \frac{d^n f(x, y)}{n!} +$$

$$+ \frac{d^{n+1} f(x + \theta\Delta x, y + \theta\Delta y)}{(n+1)!}, \quad 0 < \theta < 1,$$

т. е. формулу (1) ■

Формула Тейлора для функции двух переменных по внешнему виду напоминает формулу Тейлора для функции одной переменной. Но на самом деле, если раскрыть выражения дифференциалов функции $f(x, y)$ в формуле (1), то получим формулу более громоздкую и сложную, чем для функции одной переменной.

Формула Тейлора для функций большего числа переменных имеет аналогичный вид.

Замечание. При $n=0$ из (1) получается формула Лагранжа (или формула конечных приращений) для функций двух переменных:

$$\Delta f(x, y) = df(x + \theta\Delta x, y + \theta\Delta y) =$$

$$= f'_x(x + \theta\Delta x, y + \theta\Delta y)\Delta x + f'_y(x + \theta\Delta x, y + \theta\Delta y)\Delta y,$$

$$0 < \theta < 1,$$

из которой, в частности, следует, что если $f'_x = f'_y = 0$, то приращения функции равны нулю и функция $f(x, y)$ будет постоянной.

§ 8. Экстремумы функции двух переменных

1. Определение экстремума. Пусть функция $z = f(x, y)$ определена на некотором множестве $\{M\}$ и $M_0(x_0; y_0)$ — некоторая точка этого множества.

Определение. *Функция $z=f(x, y)$ имеет в точке M_0 локальный максимум (минимум), если для любой точки M из некоторой окрестности точки M_0 выполняется неравенство:* $f(x, y) \leqslant f(x_0, y_0)$ $(f(x, y) \geqslant f(x_0, y_0))$.

Такие точки называются *точками экстремума.*

Если функция $f(x, y)$ имеет экстремум в точке M_0, то полное приращение

$$\Delta z = f(M) - f(M_0)$$

этой функции в точке M_0 удовлетворяет одному из следующих условий:

$\Delta z \leqslant 0$ (в случае локального максимума),

$\Delta z \geqslant 0$ (в случае локального минимума).

Из сказанного следует, что полное приращение функции не меняет знака в некоторой окрестности точки экстремума.

2. Необходимые условия экстремума.

Теорема 12.7. *Если функция $f(x, y)$ имеет в точке $M_0(x_0; y_0)$ экстремум и имеет в точке M_0 частные производные 1-го порядка, то в этой точке частные производные 1-го порядка равны нулю, т. е.*

$$f'_x(x_0, y_0) = f'_y(x_0, y_0) = 0. \tag{1}$$

Доказательство. Докажем, например, равенство нулю частной производной $f'_x(x_0, y_0)$. Для этого рассмотрим в окрестности точки M_0 только те точки, для которых $y = y_0$. Мы получим функцию $f(x, y_0)$ одной переменной x. Эта функция имеет в точке $x = x_0$ экстремум и имеет в точке $x = x_0$ производную $f'_x(x_0, y_0)$. Это значит, что в этой точке выполняется необходимое условие экстремума функции одной переменной: $f'_x(x_0, y_0) = 0$, что и требовалось доказать.

Аналогично, рассматривая функцию $f(x_0, y)$ одной переменной y, находим: $f'_y(x_0, y_0) = 0$ ■

Условие (1) не является достаточным условием экстремума. Например, частные производные функции $z = x^2 - y^2$ равны нулю в точке (0; 0), но эта функция не имеет экстремума в этой точке, так как она равна нулю в точке (0; 0) и в окрестности точки (0; 0) не сохраняет знак, т. е. при $x=0$ $z<0$, а при $y=0$ $z>0$. Графиком данной функции является гиперболический параболоид (см. рис. 40).

Таким образом, условие (1) является только необходимым, а точки, в которых оно выполняется, мы будем по аналогии с функциями одной переменной называть *точками возможного экстремума.* Такие точки называются также *стационарными.*

3. Достаточные условия экстремума.

Теорема 12.8. *Пусть в точке M возможного экстремума и некоторой ее окрестности функция $f(x, y)$ имеет непрерывные частные производные 2-го порядка. Положим*

$$\Delta = \begin{vmatrix} f''_{xx}(x, y) & f''_{xy}(x, y) \\ f''_{xy}(x, y) & f''_{yy}(x, y) \end{vmatrix}.$$

Тогда

а) *если $\Delta > 0$, то в точке M функция имеет экстремум, причем при $f''_{xx}(x, y) < 0$ — локальный максимум, при $f''_{xx}(x, y) > 0$ — локальный минимум;*

б) *если $\Delta < 0$, то в точке M нет экстремума.*

Доказательство. а) Пусть $\Delta > 0$. Введем следующие обозначения: $f''_{xx}(x, y) = A$, $f''_{xy}(x, y) = B$, $f''_{yy}(x, y) = C$. По условию $f'_x(x, y) = f'_y(x, y) = 0$. Тогда, согласно формуле Тейлора (1), § 7, взятой для $n=1$, полное приращение функции $f(x, y)$ в точке M можно записать в виде

$$\Delta f = \frac{1}{2!}\left[A'(\Delta x)^2 + 2B'\Delta x\Delta y + C'(\Delta y)^2\right], \qquad (1)$$

где

$$A' = f''_{xx}(x + \theta\Delta x, y + \theta\Delta y),\ B' = f''_{xy}(x + \theta\Delta x, y + \theta\Delta y),$$

$$C' = f''_{yy}(x + \theta\Delta x, y + \theta\Delta y),\ 0 < \theta < 1.$$

Из непрерывности частных производных 2-го порядка в точке M следует

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} A' = f''_{xx}(x, y) = A \gtrless 0,$$

а также

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} (A'C' - B'^2) = f''_{xx}(x, y) f''_{yy}(x, y) - [f''_{xy}(x, y)]^2 = \Delta > 0.$$

Поэтому для достаточно малых Δy и Δx имеем

$$A' \lesseqgtr 0, A'C' - B'^2 = \Delta' > 0.$$

Поскольку $A' \gtrless 0$, то соотношение (1) можно переписать в виде

$$\Delta f = \frac{1}{2!}\cdot\frac{1}{A'}[A'^2(\Delta x)^2 + 2A'B'\Delta x\Delta y + A'C'(\Delta y)^2],$$

или, дополняя до полного квадрата,

$$\Delta f = \frac{1}{2!}\cdot\frac{1}{A'}[(A'\Delta x + B'\Delta y)^2 + (A'C' - B'^2)(\Delta y)^2].$$

Выражение в квадратных скобках неотрицательно, поэтому если $A' > 0\,(f''_{xx}(x, y) > 0)$, то $\Delta f \geqslant 0$, следовательно, в точке M — локальный минимум; если же $A' < 0\,(f''_{xx}(x, y) < 0)$, то $\Delta f \leqslant 0$, следовательно, в точке M — локальный максимум, что и требовалось доказать.

б) Пусть теперь $\Delta < 0$ и по-прежнему

$$A = f''_{xx}(x, y), B = f''_{xy}(x, y), C = f''_{yy}(x, y).$$

Рассмотрим многочлен

$$A + 2Bx + Cx^2.$$

Так как дискриминант $B^2 - AC > 0$, то существуют два разных корня и можно указать два числа x_1 и x_2 таких, что

$$A + 2Bx_1 + Cx_1^2 > 0,$$

$$A + 2Bx_2 + Cx_2^2 < 0.$$

Рассмотрим случай со значением x_1.

В силу непрерывности частных производных 2-го порядка

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} (A' + 2B'x_1 + C'x_1^2) = A + 2Bx_1 + Cx_1^2 > 0.$$

Следовательно, существует δ-окрестность точки M такая, что если точка $M_1(x+\Delta x;\ y+\Delta y)$ принадлежит этой окрестности, то

$$A' + 2B'x_1 + C'x_1^2 > 0. \tag{2}$$

Рассмотрим теперь произвольную δ'-окрестность точки M такую, что $\delta' < \delta$. Можно выбрать число $t > 0$ столь малым, что точка $M_2(x+t;\ y+tx_1)$ будет принадлежать δ'-окрестности точки M. Положив $\Delta x = t$, $\Delta y = tx_1$, из (1) и (2) получаем

$$\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y) =$$

$$= \frac{1}{2!} t^2 [A' + 2B'x_1 + C'x_1^2] > 0.$$

Поступая аналогично со значением x_2, получим, что в произвольной δ'-окрестности точки M существует точка $M_3(x+\Delta x;\ y+\Delta y)$, для которой

$$\Delta f = f(x+\Delta x,\ y+\Delta y) - f(x,\ y) < 0,$$

т. е. приращение функции $f(x, y)$ в сколь угодно малой окрестности точки M не сохраняет знак, следовательно, в точке M нет экстремума ■

Замечание. Если $\Delta = 0$, то функция $f(x, y)$ в точке M возможного экстремума может иметь экстремум, но может и не иметь его. В указанном случае для выяснения вопроса требуется специальное дальнейшее исследование, которое выходит за рамки данного курса.

Примеры:

1. Исследовать на экстремум функцию $z = x^2 + xy + y^2 - 2x - 3y$. Имеем

$$f'_x = 2x + y - 2,\ f'_y = x + 2y - 3,\ f''_{xx} = 2,\ f''_{xy} = 1,\ f''_{yy} = 2,$$

$$\Delta = 2\cdot 2 - 1 = 3 > 0.$$

Находим точки возможного экстремума. Для этого решаем систему уравнений

$$\begin{cases} 2x + y - 2 = 0,\\ x + 2y - 3 = 0.\end{cases}$$

Решением будет $x = \frac{1}{3}$, $y = \frac{4}{3}$. Следовательно, точка $M\left(\frac{1}{3}; \frac{4}{3}\right)$ — искомая точка возможного экстремума.

Так как $\Delta=3>0$ и $f''_{xx}=2>0$, то в точке $M\left(\frac{1}{3}; \frac{4}{3}\right)$ данная функция имеет минимум.

2. Исследовать на экстремум функцию $z=x^2-y^2$. Имеем

$$f'_x = 2x,\ f'_y = -2y,\ f''_{xx} = 2,\ f''_{xy} = 0,\ f''_{yy} = -2,$$
$$\Delta = 2\cdot(-2) - 0 = -4 < 0.$$

Решая систему уравнений $2x=0$ и $-2y=0$, находим точку $M(0;\ 0)$ — точку возможного экстремума. Поскольку $\Delta=-4<0$, то в точке $M(0;\ 0)$ нет экстремума.

3. Исследовать на экстремум функцию $z=x^4+y^4$. Имеем

$$f'_x = 4x^3,\ f'_y = 4y^3,\ f''_{xx} = 12x^2,\ f''_{xy} = 0,\ f''_{yy} = 12y^2.$$

Решая систему уравнений $4x^3=0$ и $4y^3=0$, находим точку $M(0;\ 0)$ — точку возможного экстремума. Тогда

$$f''_{xx}(0; 0) = 0,\ f''_{yy}(0; 0) = 0,\ \Delta = 0.$$

Согласно замечанию, в точке $M(0;\ 0)$ экстремум может быть, а может не быть. В данном случае он есть, так как $z>0$ во всех точках, кроме $M(0;\ 0)$, т. е. данная функция не меняет знак в окрестности точки $M(0;\ 0)$, а следовательно, в этой точке функция имеет минимум.

4. Исследовать на экстремум функцию $z=x^3+y^3$. Имеем

$$f'_x = 3x^2,\ f'_y = 3y^2,\ f''_{xx} = 6x,\ f''_{xy} = 0,\ f''_{yy} = 6y.$$

Решая систему уравнений $3x^2=0$ и $3y^2=0$, находим точку $M(0;\ 0)$ — точку возможного экстремума. Тогда

$$f''_{xx}(0; 0) = 0,\ f''_{yy}(0; 0) = 0,\ \Delta = 0.$$

В данном случае непосредственными рассуждениями убеждаемся, что точек экстремума нет. В самом деле,

$z(0;\ 0)=0$, $z(x,\ 0)=x^3$, отсюда при $x>0$ $z>0$, а при $x<0$ $z<0$, т. е. в любой окрестности точки $M(0;\ 0)$ данная функция меняет знак.

§ 9. Метод наименьших квадратов

В различных исследованиях приходится пользоваться формулами, составленными на основании эксперимента. Одним из лучших способов получения таких формул является *метод наименьших квадратов*.

Пусть на основании эксперимента необходимо установить функциональную зависимость между двумя переменными величинами x и y, например между температурой и удлинением прямолинейного металлического стержня. Производим n измерений и по результатам составляем таблицу

x	x_1	x_2	...	x_i	...	x_n
y	y_1	y_2	...	y_i	...	y_n

При этом вид функции $y=f(x)$ устанавливается или из теоретических исследований, или по характеру расположения на координатной плоскости экспериментальных точек. Пусть, например, точки, взятые из таблицы, расположены так, как показано на рис. 45. В данном случае естественно предположить, что между x и y существует линейная зависимость, выражающаяся формулой

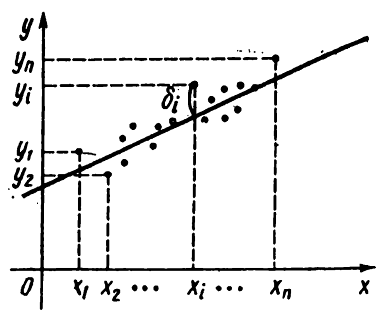


Рис. 45

$$y=ax+b. \tag{1}$$

Мы ограничимся рассмотрением только случая линейной зависимости.

Так как точки $(x_1; y_1)$, $(x_2; y_2)$, ..., $(x_n; y_n)$ приблизительно лежат на прямой, то формула (1) является приближенной. Поэтому, подставляя их координаты в формулу (1) вместо x и y, получим следующие равенства:

$$y_1-(ax_1+b)=\delta_1,$$
$$y_2-(ax_2+b)=\delta_2,$$
$$\cdots\cdots\cdots$$
$$y_n-(ax_n+b)=\delta_n,$$

где δ_1, δ_2, ..., δ_n — некоторые числа, которые назовем погрешностями.

Возникает задача — подобрать коэффициенты a и b таким образом, чтобы эти погрешности были возможно меньше по абсолютной величине. Методом решения этой задачи и является метод наименьших квадратов. Согласно этому методу рассмотрим сумму квадратов погрешностей:

$$S(a, b)=\sum_{i=1}^{n}[y_i-(ax_i+b)]^2=\sum_{i=1}^{n}\delta_i^2,$$

где x_i и y_i — заданные числа, а коэффициенты a и b — неизвестные величины, подлежащие определению, т. е. $S(a, b)$ можно рассматривать как функцию двух переменных a и b и исследовать ее на экстремум.

Таким образом, задача свелась к нахождению значений a и b, при которых функция $S(a, b)$ имеет минимум.

Имеем

$$\frac{\partial S}{\partial a}=-2\sum_{i=1}^{n}[y_i-(ax_i+b)]x_i,$$

$$\frac{\partial S}{\partial b}=-2\sum_{i=1}^{n}[y_i-(ax_i+b)].$$

Приравнивая эти частные производные нулю, получаем линейную систему двух уравнений с двумя неизвестными a и b:

$$\begin{cases} \sum_{i=1}^{n} y_i x_i = a \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i, \\ \sum_{i=1}^{n} y_i = a \sum_{i=1}^{n} x_i + bn. \end{cases} \qquad (2)$$

Система (2) называется *нормальной* системой метода наименьших квадратов.

Из этой системы мы находим числа a и b и затем, подставляя их в уравнение (1), получаем формулу искомой прямой.

То, что функция $S(a, b)$ в точке $M(a, b)$ имеет минимум, легко устанавливается с помощью частных производных 2-го порядка. Имеем

$$\frac{\partial^2 S}{\partial a^2} = 2\sum_{i=1}^{n} x_i^2, \quad \frac{\partial^2 S}{\partial a \partial b} = 2\sum_{i=1}^{n} x_i, \quad \frac{\partial^2 S}{\partial b^2} = 2n.$$

Следовательно,

$$\Delta = \frac{\partial^2 S}{\partial a^2} \cdot \frac{\partial^2 S}{\partial b^2} - \left(\frac{\partial^2 S}{\partial a \partial b}\right)^2 = 4n \sum_{i=1}^{n} x_i^2 -$$

$$-\left(2\sum_{i=1}^{n} x_i\right)^2 = 4(n-1)\sum_{i=1}^{n} x_i^2 > 0, \quad \frac{\partial^2 S}{\partial a^2} > 0,$$

что и требовалось. Рассмотрим

Пример. Пусть в результате эксперимента получены пять значений искомой функции y при пяти значениях аргумента $x (n=5)$, которые записаны в таблице. Будем искать функциональную зависимость между x и y в виде линейной функции $y = ax + b$.

x	y
-2	0,5
0	1
1	1,5
2	2
4	3

При составлении нормальной системы (2) для определения коэффициентов a и b предварительно вычисляем

$$\sum_{i=1}^{5} y_i x_i = 16{,}5; \sum_{i=1}^{5} x_i^2 = 25; \sum_{i=1}^{5} x_i = 5; \sum_{i=1}^{5} y_i = 8.$$

Система (2) принимает вид

$$\begin{cases} 25a + 5b = 16{,}5, \\ 5a + 5b = 8. \end{cases}$$

Решая эту систему, находим: $a=0{,}425$, $b=1{,}175$. Отсюда формула искомой прямой есть

$$y=0{,}425x+1{,}175.$$

Глава XIII
ИНТЕГРИРОВАНИЕ

В данной главе мы рассмотрим основные вопросы, связанные с интегрированием функций двух переменных. Определения и результаты, которые будут получены могут быть перенесены без особых трудностей на функции трех и более переменных.

§ 1. Двойные интегралы

Двойные интегралы представляют собой обобщение определенного интеграла на случай функций двух переменных.

1. Определение и существование двойного интеграла. Пусть G — некоторая замкнутая ограниченная область, а $z=f(x, y)$ — произвольная функция, определенная и ограниченная в области G.

Предполагается, что граница области G состоит из конечного числа кривых, заданных уравнениями вида $y=f(x)$ или $x=\varphi(y)$, где $f(x)$ и $\varphi(y)$ — непрерывные функции. Такой областью, например, является замкнутый многоугольник, граница которого состоит из конечного числа отрезков, представляющих собой графики непрерывных функций вида $y=kx+b$ или $x=a$.

Другой пример — область, ограниченная эллипсом, т. е. граница состоит из двух кривых:

$$y = \pm \frac{b}{a}\sqrt{a^2 - x^2}, \text{ и т. д.}$$

Разобьем область G произвольно на n частей G_i, не имеющих общих внутренних точек, с площадями Δs_i ($i=1, 2, ..., n$) (рис. 46). На каждой G_i выберем произвольную точку (ξ_i; η_i) и составим сумму

$$\sigma = \sum_{i=1}^{n} f(\xi_i; \eta_i)\,\Delta s_i, \qquad (1)$$

которую назовем *интегральной суммой* для функции $f(x, y)$ в области G. Назовем *диаметром* области наибольшее расстояние между граничными точками этой области. Обозначим через λ наибольший из диаметров частичных областей G_i ($\lambda = \max\limits_{1 \leqslant i \leqslant n} \{d(G_i)\}$) и дадим следующее

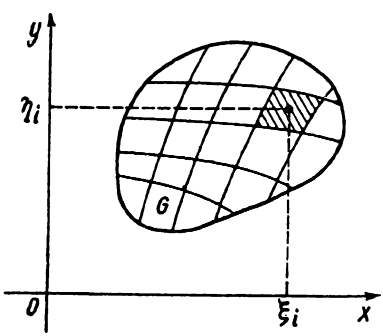


Рис. 46

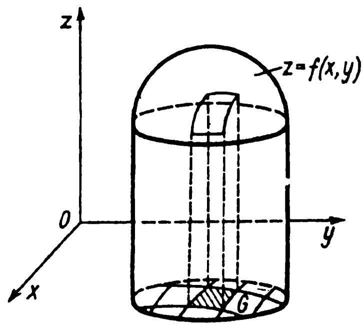


Рис. 47

Определение. *Если интегральная сумма (1) при $\lambda \to 0$ имеет конечный предел J*, то этот предел назы-*

* Число J называется пределом интегральной суммы (1) при $\lambda \to 0$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\lambda < \delta$ независимо от выбора точек (ξ_i; η_i) выполняется неравенство $|\sigma - J| < \varepsilon$.

вается двойным интегралом функции $f(x, y)$ по области G и обозначается одним из символов:

$$J = \iint\limits_G f(x, y)\, ds = \iint\limits_G f(x, y)\, dxdy.$$

В этом случае функция $f(x, y)$ называется *интегрируемой* в области G, G — *областью интегрирования*, x и y — *переменными интегрирования*, ds (или $dxdy$) — *элементом площади*.

Давая определение двойного интеграла, мы предполагаем, что функция $f(x, y)$ ограничена. Как и для функции одной переменной, это условие является необходимым условием интегрируемости. Однако это условие не является достаточным, так как существуют примеры ограниченных, но не интегрируемых функций. Примером таких функций может служить функция, определенная на квадрате $\{(x;\ y)\,|\,0\leqslant x\leqslant 1;\ 0\leqslant y\leqslant 1\}$ следующим образом:

$$f(x, y) = \begin{cases} 1, \text{ если } x \text{ и } y \text{ рациональны,} \\ 0, \text{ если } x \text{ или } y \text{ иррациональны.} \end{cases}$$

Доказательство неинтегрируемости такой функции непосредственно следует из определения и предоставляется читателю.

Для нахождения достаточных условий интегрируемости, как и в случае одной переменной, удобно воспользоваться теорией Дарбу, которая полностью переносится на случай двойного интеграла. Аналогично доказательству соответствующей теоремы для определенного интеграла доказывается следующая теорема.

Теорема 13.1. *Функция $f(x, y)$, непрерывная в области G, интегрируема в этой области.*

Однако не следует думать, что двойной интеграл существует только для непрерывных функций. Так же как для определенного интеграла, имеет место более общая теорема.

Теорема 13.2. *Функция $f(x, y)$, ограниченная в области G и непрерывная в ней всюду, кроме точек, лежащих на конечном числе кривых, являющихся графиками непрерывных функций вида $y=f(x)$ или $x=\varphi(y)$, интегрируема в этой области.*

2. Геометрический смысл двойного интеграла. Пусть в пространстве дано тело P (рис. 47), ограниченное сверху графиком непрерывной и неотрицательной функции $z=f(x, y)$, определенной в области G, с боков цилиндрической поверхностью, направляющей которой служит граница области G, снизу областью G, лежащей в плоскости Oxy. Тело такого вида называют *криволинейным цилиндром.*

Аналогично тому как задача о вычислении площади криволинейной трапеции привела нас к геометрическому понятию определенного интеграла, так задача о вычислении объема тела P приведет нас к геометрическому толкованию двойного интеграла.

Действительно, в данном случае интегральная сумма (1) представляет собой сумму объемов прямых цилиндров с площадями оснований Δs_i и высотами $f(\xi_i, \eta_i)$, которую можно принять за приближенное значение объема тела P:

$$v_P \approx \sum_{i=1}^{n} f(\xi_i, \eta_i)\,\Delta s_i.$$

Это приближенное равенство будет тем точнее, чем мельче разбиение области G на части. При переходе к пределу при $\lambda \to 0$ это приближенное равенство превращается в точное

$$v_P = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i, \eta_i)\,\Delta s_i.$$

Так как функция $f(x, y)$ интегрируема, то предел интегральной суммы существует и равен двойному интегралу от этой функции по области G. Следовательно,

$$v_P = \iint\limits_G f(x, y)\,dxdy.$$

Отсюда следует геометрический смысл двойного интеграла: *двойной интеграл от непрерывной, неотрицательной функции равен объему криволинейного цилиндра.*

З а м е ч а н и е. Если положить $f(x, y) \equiv 1$ всюду в области G, то непосредственно из определения двойного интеграла получаем выражение для площади s области G в виде двойного интеграла:

$$\iint_G 1 \cdot dxdy = \lim_{\lambda \to 0} \sum_{i=1}^{n} 1 \cdot \Delta s_i = \lim_{\lambda \to 0} s = s.$$

3. Свойства двойного интеграла. Основные свойства двойного интеграла вполне аналогичны соответствующим свойствам определенного интеграла. Поэтому мы ограничимся формулировкой этих свойств, не останавливаясь на доказательствах.

1) Если k — постоянное число и функция $f(x, y)$ интегрируема в области G, то функция $kf(x, y)$ тоже интегрируема в G и

$$\iint_G kf(x, y)\,dxdy = k \iint_G f(x, y)\,dxdy,$$

т. е. постоянный множитель можно выносить за знак интеграла.

2) Если функции $f(x, y)$ и $g(x, y)$ интегрируемы в области G, то их алгебраическая сумма также интегрируема в этой области и

$$\iint_G [f(x, y) \pm g(x, y)]\,dxdy = \iint_G f(x, y)\,dxdy \pm$$
$$\pm \iint_G g(x, y)\,dxdy.$$

3) Если область G является объединением областей G_1 и G_2, не имеющих общих внутренних точек, в каждой из которых функция $f(x, y)$ интегрируема, то в G эта функция также интегрируема и

$$\iint_G f(x, y)\,dxdy = \iint_{G_1} f(x, y)\,dxdy + \iint_{G_2} f(x, y)\,dxdy.$$

4) (Теорема о среднем.) Если функция $f(x, y)$ непрерывна в области G, то в этой области найдется такая точка $(\xi_i; \eta_i)$, что

$$\iint_G f(x, y)\,dxdy = f(\xi_i; \eta_i)\,s,$$

где s — площадь фигуры G.

Итак, мы познакомились с определением и основными свойствами двойного интеграла, условиями существования, выяснили его геометрический смысл. Теперь рассмотрим вопрос о способах вычисления двойных интегралов.

§ 2. Сведение двойного интеграла к повторному

Сведение двойного интеграла к повторному является одним из эффективных способов вычисления двойного интеграла.

1. Случай прямоугольной области. Сначала рассмотрим двойной интеграл по некоторому прямоугольнику D со сторонами, параллельными осям координат.

Теорема 13.3. *Пусть для функции $f(x, y)$ в прямоугольнике $D=\{(x;\ y)\,|\,a\leqslant x\leqslant b;\ c\leqslant y\leqslant d\}$ существует двойной интеграл*

$$\iint\limits_D f(x, y)\,dxdy. \tag{1}$$

Пусть, далее, для каждого x из отрезка $[a, b]$ существует определенный интеграл

$$J(x)=\int\limits_c^d f(x, y)\,dy. \tag{2}$$

Тогда существует интеграл

$$\int\limits_a^b J(x)\,dx=\int\limits_a^b dx\int\limits_c^d f(x, y)\,dy$$

(он называется повторным) и справедливо равенство

$$\iint\limits_D f(x, y)\,dxdy=\int\limits_a^b dx\int\limits_c^d f(x, y)\,dy. \tag{3}$$

Доказательство. Разобьем прямоугольник D с помощью точек $a=x_0<x_1<x_2<\ldots<x_n=b$ и $c=y_0<<y_1<y_2<\ldots<y_k=d$ на nk частичных прямоугольников $D_{ij}=\{(x;\ y)\,|\,x_{i-1}\leqslant x\leqslant x_i;\ y_{j-1}\leqslant y\leqslant y_j\}$. Положим $\Delta x_i=x_i-x_{i-1}$, $\Delta y_j=y_j-y_{j-1}$ и обозначим через m_{ij} и M_{ij}, соответственно, точную нижнюю и верхнюю грани

функции $f(x, y)$ на частичном прямоугольнике D_{ij} (рис. 48). Тогда всюду на этом прямоугольнике

$$m_{ij} \leqslant f(x, y) \leqslant M_{ij}. \tag{4}$$

Положим в этом неравенстве $x = \xi_i$, где ξ_i — произвольная точка отрезка $[x_{i-1}, x_i]$, и после этого проинтегрируем (4) по y в пределах от y_{j-1} до y_j. Получим

$$m_{ij}\Delta y_j \leqslant \int_{y_{j-1}}^{y_j} f(\xi_i, y)\,dy \leqslant M_{ij}\Delta y_j. \tag{5}$$

Суммируя (5) по всем j от 1 до k и используя обозначения (2), будем иметь

$$\sum_{j=1}^{k} m_{ij}\Delta y_j \leqslant J(\xi_i) \leqslant \sum_{j=1}^{k} M_{ij}\Delta y_j. \tag{6}$$

Далее, умножая (6) на Δx_i и суммируя по всем i от 1 до n, получим

$$\sum_{i=1}^{n}\sum_{j=1}^{k} m_{ij}\Delta x_i\Delta y_j \leqslant \sum_{i=1}^{n} J(\xi_i)\,\Delta x_i \leqslant \sum_{i=1}^{n}\sum_{j=1}^{k} M_{ij}\Delta x_i\Delta y_j. \tag{7}$$

Пусть наибольший диаметр частичных прямоугольников D_{ij} стремится к нулю ($\lambda \to 0$). Тогда и наибольшая из длин $\Delta x_i \to 0$. Крайние члены в (7), представляющие собой нижнюю и верхнюю суммы Дарбу, стремятся при этом к двойному интегралу (1). Стало быть, существует предел и среднего члена (7), равный тому же

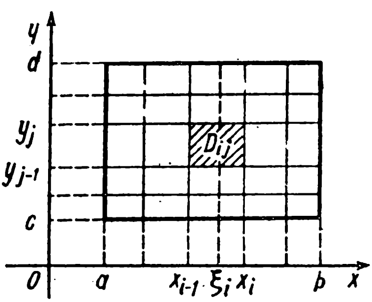


Рис. 48

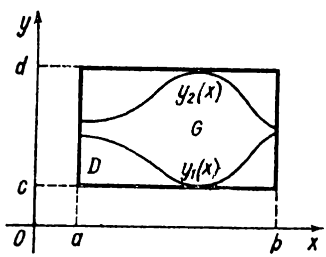


Рис. 49

самому двойному интегралу. Но этот предел по определению определенного интеграла равен

$$\int_a^b J(x)\,dx = \int_a^b dx \int_c^d f(x, y)\,dy.$$

Тем самым доказано существование повторного интеграла и равенство (3) ■

Замечание. Если в теореме 13.3 x и y поменять ролями, то теорема будет утверждать существование повторного интеграла

$$\int_c^d J(y)\,dy = \int_c^d dy \int_a^b f(x, y)\,dx$$

и равенство

$$\iint_D f(x, y)\,dxdy = \int_c^d dy \int_a^b f(x, y)\,dx. \qquad (8)$$

Формулы (3) и (8) приводят двойной интеграл к повторному, в котором, например, в (8) интегрирование сперва проводится по x при постоянном y, а затем полученный результат интегрируется по y, т. е. производится последовательное вычисление двух определенных интегралов.

Пример 1. Вычислить $\iint_D xydxdy$, где $D = \{(x;\ y)\,|\,1 \leqslant x \leqslant 2;\ 1 \leqslant y \leqslant 2\}$.

Решение. Имеем

$$\iint_D xydxdy = \int_1^2 dy \int_1^2 xydx = \int_1^2 \left[y\frac{x^2}{2}\right]_1^2 dy =$$

$$= \int_1^2 \left[2y - \frac{1}{2}y\right] dy = \frac{3}{4}y^2\Big|_1^2 = \frac{9}{4}.$$

2. Случай криволинейной области. Теперь рассмотрим двойной интеграл для случая криволинейной области G.

Теорема 13.4. *Пусть функция $z = f(x, y)$ определена в области $G = \{(x;\ y)\,|\,a \leqslant x \leqslant b;\ y_1(x) \leqslant y \leqslant y_2(x)\}$, где*

$y_1(x)$ и $y_2(x)$ — непрерывные функции, $y_1(x) \leqslant y_2(x)$ для $a \leqslant x \leqslant b$.

Пусть существует двойной интеграл

$$\iint\limits_G f(x, y)\, dxdy,$$

и для каждого x из отрезка $[a, b]$ существует определенный интеграл

$$J(x) = \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\, dy.$$

Тогда существует повторный интеграл

$$\int\limits_a^b J(x)\, dx = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\, dy$$

и справедливо равенство

$$\iint\limits_G f(x,y)\, dxdy = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\, dy. \tag{9}$$

Доказательство. Положим $c = \min y_1(x)$, $d = \max y_2(x)$ и заключим область G в прямоугольник $D = \{(x;\ y) \mid a \leqslant x \leqslant b;\ c \leqslant y \leqslant d\}$ (рис. 49). Рассмотрим в этом прямоугольнике вспомогательную функцию

$$F(x, y) = \begin{cases} f(x, y) & \text{в точках области } G, \\ 0 & \text{в остальных точках } D. \end{cases}$$

Функция $F(x, y)$ удовлетворяет условиям предыдущей теоремы. Действительно, она интегрируема в области G, так как совпадает в ней с $f(x, y)$, и интегрируема в остальной части прямоугольника $D (D-G)$, где она равна нулю. Следовательно, по свойству 3), § 1, она интегрируема и по всему прямоугольнику D. При этом

$$\iint\limits_G F(x, y)\, dxdy = \iint\limits_G f(x, y)\, dxdy \quad \text{и} \quad \iint\limits_{D-G} F(x, y)\, dxdy = 0,$$

откуда

$$\iint\limits_D F(x, y)\, dxdy = \iint\limits_G f(x, y)\, dxdy. \tag{10}$$

Далее, для каждого x из $[a, b]$ существует интеграл

$$\int_c^d F(x, y)\,dy = \int_c^{y_1(x)} F(x, y)\,dy + \int_{y_1(x)}^{y_2(x)} F(x, y)\,dy + \int_{y_2(x)}^{d} F(x, y)\,dy,$$

так как существует каждый из трех интегралов, стоящих справа. Действительно, отрезки $[c, y_1(x)]$ и $[y_2(x), d]$ лежат вне области G и на них $F(x, y)$ равна нулю, отсюда первый и третий интегралы равны нулю, а второй интеграл существует по условию, так как на $[y_1(x), y_2(x)]$ $F(x, y) = f(x, y)$. Поэтому окончательно получаем

$$\int_c^d F(x, y)\,dy = \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy. \tag{11}$$

Таким образом, мы видим, что для функции $F(x, y)$ выполнены все условия теоремы 13.3 и, следовательно, двойной интеграл от этой функции по прямоугольнику D может быть сведен к повторному:

$$\iint_D F(x, y)\,dxdy = \int_a^b dx \int_c^d F(x, y)\,dy.$$

Отсюда и из равенств (10) и (11) получаем

$$\iint_G f(x, y)\,dxdy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy,$$

т. е. формулу (9) ■

Замечание 1. Если в теореме 13.4 поменять ролями x и y, то теорема будет утверждать существование повторного интеграла

$$\int_c^d J(y)\,dy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y)\,dx$$

и равенства

$$\iint_G f(x, y)\,dxdy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y)\,dx. \tag{12}$$

Пример 2. Вычислить $\iint\limits_G (x^2+xy+2y^2)\,dxdy$ по области $G=\{(x;\ y)\,|\,0\leqslant x\leqslant 1;\ 0\leqslant y\leqslant 1-x\}$.

Решение. Легко видеть, что G представляет собой треугольник, ограниченный осями координат и прямой $y=-x+1$ (рис. 50). Следовательно, $y_1(x)=0$, $y_2(x)=1-x$. По формуле (9) имеем

$$\iint\limits_G (x^2+xy+2y^2)\,dxdy=\int\limits_0^1 dx\int\limits_0^{1-x}(x^2+xy+2y^2)\,dy=$$

$$=\int\limits_0^1\left[x^2(1-x)+\frac{x(1-x)^2}{2}+\frac{2(1-x)^3}{3}\right]dx=\frac{7}{24}.$$

Данный интеграл можно вычислить и по формуле (12), если в G x и y поменяем ролями. Тогда треугольник будет определен неравенствами $0\leqslant y\leqslant 1$, $0\leqslant x\leqslant 1-y$, отсюда

$$x_1(y)=0,\quad x_2(y)=1-y$$

и легко проверить, что интеграл

$$\iint\limits_G (x^2+xy+2y^2)\,dxdy=\int\limits_0^1 dy\int\limits_0^{1-y}(x^2+xy+2y^2)\,dx$$

имеет то же самое значение.

Замечание 2. Если область G не удовлетворяет условиям теоремы 13.4 (например, прямые (вертикальные или горизонтальные) пересекают ее границу более чем в двух точках), то необходимо область G разбить на части, каждая из которых удовлетворяла бы условиям теоремы 13.4, и сводить к повторному каждый из соответствующих двойных интегралов отдельно.

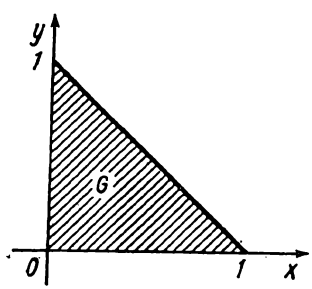


Рис. 50

§ 3. Замена переменных в двойном интеграле

Пусть в некоторой области G для функции $f(x, y)$ существует двойной интеграл

$$\iint_G f(x, y)\,dxdy. \tag{1}$$

Предположим, далее, что с помощью формул

$$\begin{aligned} x &= x(u, v), \\ y &= y(u, v) \end{aligned} \tag{2}$$

мы переходим к новым переменным u и v. Предполагается, что u и v определяются из (2) единственным образом:

$$\begin{aligned} u &= u(x, y), \\ v &= v(x, y). \end{aligned} \tag{3}$$

Обозначим через G^* множество точек $M^*(u; v)$. С помощью формул (3) каждой точке $M(x; y)$ из области G ставится в соответствие некоторая точка $M^*(u; v)$ из области G^*. Формулы (2) называют *формулами преобразования* координат, а формулы (3) — *формулами обратного преобразования*.

При сделанных предположениях доказывается, что если функции (2) имеют в области G^* непрерывные частные производные 1-го порядка и если выражение

$$\frac{D(x, y)}{D(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{4}$$

отлично в G^* от нуля, то для интеграла (1) справедлива формула замены переменных:

$$\iint_G f(x, y)\,dxdy = \iint_{G^*} f[x(u; v), y(u, v)] \left|\frac{D(x, y)}{D(u, v)}\right| dudv. \tag{5}$$

Выражение (4) называется *функциональным определителем* функций $x=x(u, v)$ и $y=y(u, v)$ или *якобианом* (по имени немецкого математика Якоби).

Коротко изложенное формулируется в виде следующей теоремы.

Теорема 13.5. *Если преобразование (2) переводит область G в G^* и является взаимно-однозначным и если функции (2) имеют в области G непрерывные частные производные 1-го порядка и отличный от нуля якобиан (4), то при условии существования интеграла (1) справедлива формула замены переменных (5).*

Доказательство теоремы опускается ввиду большой сложности .

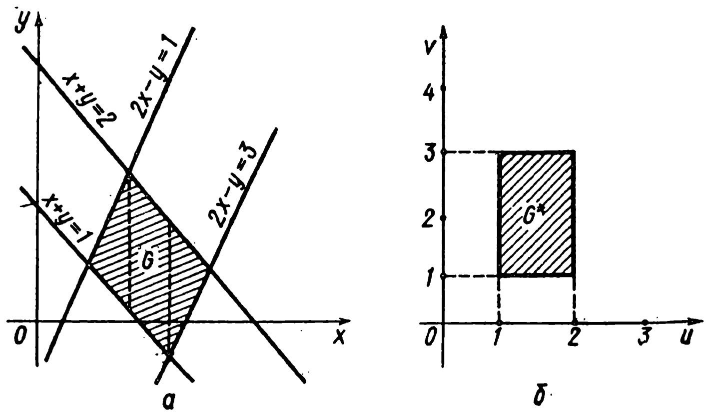


Рис. 51

В двойном интеграле, как и в определенном, замена переменных — важнейший способ приведения интеграла к виду, более удобному для его вычисления.

Пример 1. Вычислить интеграл $\iint\limits_G (2x-y)\,dxdy$, где G — параллелограмм, ограниченный прямыми: $x+y=1$, $x+y=2$, $2x-y=1$, $2x-y=3$ (рис. 51, *а*).

Решение. Непосредственное вычисление этого интеграла было бы затруднительным, так как для его вычисления необходимо область разбить на три области и вычислять соответственно три интеграла. Однако простая замена переменных

$$x+y=u, \quad 2x-y=v \tag{6}$$

позволяет избежать такого громоздкого способа решения. Прямые $x+y=1$ и $x+y=2$ в системе координат Oxy переходят в прямые $u=1$ и $u=2$ в системе координат Ouv (рис. 51, *б*), а прямые $2x-y=1$ и $2x-y=3$ — в прямые $v=1$ и $v=3$. Параллелограмм G однозначно преобразуется в прямоугольник G^*, который является более простой областью интегрирования. Осталось вычислить якобиан. Для этого выразим x и y через u и v из равенств (6):

$$x=\frac{u+v}{3},\quad y=\frac{2u-v}{3}.$$

Следовательно,

$$\frac{D(x,y)}{D(u,v)}=\begin{vmatrix}\frac{1}{3} & \frac{1}{3}\\ \frac{2}{3} & -\frac{1}{3}\end{vmatrix}=-\frac{1}{9}-\frac{2}{9}=-\frac{1}{3}.$$

По формуле (5) окончательно получаем

$$\iint\limits_G (2x-y)\,dxdy=\iint\limits_{G^*}\frac{1}{3}\,vdudv=\frac{1}{3}\int\limits_1^2 du\int\limits_1^3 vdv=$$

$$=\frac{1}{3}\int\limits_1^2\left[\frac{v^2}{2}\right]_1^3 du=\frac{1}{3}\left(\frac{9}{2}-\frac{1}{2}\right)\int\limits_1^2 du=\frac{1}{3}\cdot 4\cdot 1=\frac{4}{3}.$$

Замечание. Если подынтегральная функция или уравнение границы области интегрирования содержат сумму x^2+y^2, то в большинстве случаев упрощение интеграла достигается преобразованием его к полярным координатам, так как данная сумма в полярных координатах получает весьма простой вид:

$$(\rho\cos\varphi)^2+(\rho\sin\varphi)^2=\rho^2.$$

Пример 2. Вычислить интеграл $\iint\limits_G e^{x^2+y^2}\,dxdy$, где G — четверть круга $x^2+y^2=1$, расположенная в первом квадранте (рис. 52).

Решение. Преобразуем интеграл к полярным координатам по формулам: $x=\rho\cos\varphi$, $y=\rho\sin\varphi$. Тогда $x^2+y^2=\rho^2$ и

$$\frac{D(x,y)}{D(\rho,\varphi)}=\begin{vmatrix}\dfrac{\partial x}{\partial\rho} & \dfrac{\partial x}{\partial\varphi}\\ \dfrac{\partial y}{\partial\rho} & \dfrac{\partial y}{\partial\varphi}\end{vmatrix}=\begin{vmatrix}\cos\varphi & -\rho\sin\varphi\\ \sin\varphi & \rho\cos\varphi\end{vmatrix}=$$

$$=\rho\left[\cos^2\varphi+\sin^2\varphi\right]=\rho.$$

В области G ρ меняется в пределах от 0 до 1, а φ — от 0 до $\frac{\pi}{2}$. Таким образом, по формуле (5) получаем

$$\iint\limits_G e^{x^2+y^2}\,dxdy=\int\limits_0^{\frac{\pi}{2}}d\varphi\int\limits_0^1 e^{\rho^2}\rho d\rho=\frac{1}{2}(e-1)\int\limits_0^{\frac{\pi}{2}}d\varphi=$$

$$=\frac{1}{2}(e-1)\frac{\pi}{2}=\frac{\pi}{4}(e-1).$$

Заметим, что на практике при замене переменных нет необходимости детально строить область G^*. Выясняют пределы изменения новых координат, используя вид области G на плоскости Oxy и геометрический смысл полярных координат, что мы и сделали в нашем примере.

§ 4. Некоторые геометрические и физические приложения двойных интегралов

1. Вычисление объемов. Мы уже знаем, что объем v криволинейного цилиндра, ограниченного сверху поверхностью $z=f(x, y)>0$, снизу плоскостью $z=0$ и с боков цилиндрической поверхностью, у которой образующие параллельны оси Oz, а направляющей служит контур области G, равен

$$v=\iint\limits_G f(x,y)\,dxdy,$$

т. е. с помощью двойных интегралов можно вычислять объемы тел.

Пример 1. Вычислить объем тела, ограниченного поверхностями $x=0$, $y=0$, $z=0$ и $x+y+z=1$ (рис. 53).

Решение. Имеем

$$v = \iint_G (1 - x - y)\, dxdy,$$

где G — треугольная область интегрирования, ограниченная прямыми $x=0$, $y=0$, $x+y=1$. Расставляя пределы интегрирования в двойном интеграле, получим

$$v = \int_0^1 dx \int_0^{1-x} (1 - x - y)\, dy = \int_0^1 \left[(1-x)\,y - \frac{y^2}{2} \right]_0^{1-x} dx =$$

$$= \frac{1}{2} \int_0^1 (1-x)^2\, dx = \frac{1}{6}.$$

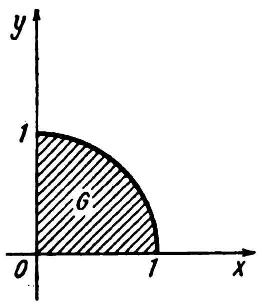


Рис. 52

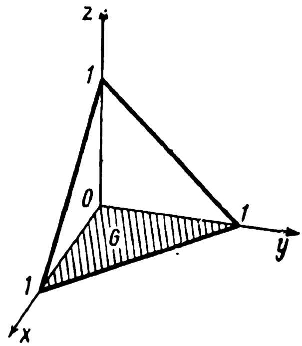


Рис. 53

2. Вычисление площадей. Как было установлено (см. замечание § 1), площадь s области G может быть вычислена с помощью двойного интеграла по формуле

$$s = \iint_G dxdy.$$

Отметим, что эта формула часто бывает удобнее, чем соответствующая формула, выражающая площадь криволинейной трапеции, с помощью определенного интеграла, так как данная формула применима не только к криволинейным трапециям, но и к фигурам, расположенным произвольным образом по отношению к координатным осям.

Пример 2. Вычислить площадь области G, ограниченной линиями $y^2=x+1$, $x+y=1$ (рис. 54).

Решение. Область G представляет собой фигуру, ограниченную слева параболой $y^2=x+1$, справа прямой $y=-x+1$. Решая совместно уравнения параболы и прямой, находим точки их пересечения: $M_1(3;\ -2)$, $M_2(0;\ 1)$. Следовательно, искомая площадь

$$s=\iint\limits_G dxdy=\int\limits_{-2}^{1}dy\int\limits_{y^2-x}^{1-y}dx=\int\limits_{-2}^{1}(2-y-y^2)\,dy=\frac{9}{2}.$$

При вычислении двойных интегралов одним из главных моментов является расстановка пределов интегрирования. Если бы мы выбрали в данном примере обрат-

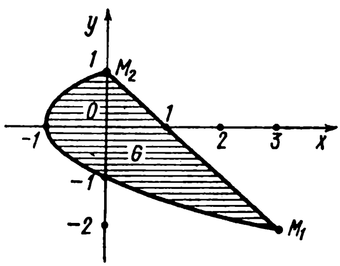


Рис. 54

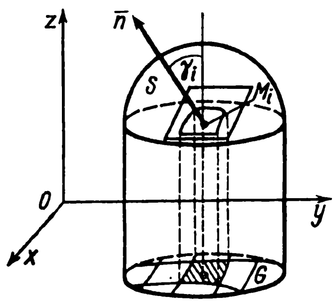


Рис. 55

ный порядок интегрирования, то область G предварительно пришлось бы разбить на две части, так как эта область ограничена сверху линией, заданной двумя различными уравнениями. Разумеется, мы получили бы тот же результат, но вычисления были бы более громоздкими.

Поэтому полезно запомнить следующее правило: если все прямые, параллельные оси Oy, входят в область интегрирования G на одной линии, а выходят на другой, то внутренний интеграл целесообразно брать по переменной y, а внешний — по x; если же все прямые, параллельные оси Ox, входят в область интегрирования G на одной линии (в нашем случае на параболе), а выходят на другой (в нашем случае на прямой), то внутренний интеграл целесообразно брать по переменной x, а внешний — по y: в этом случае мы получаем один повторный интеграл.

3. Вычисление площади поверхности. С помощью двойных интегралов можно вычислять площади не только плоских фигур, но и кривых поверхностей.

Пусть поверхность S задана уравнением $z=f(x, y)$, проекцией S на плоскость Oxy является область G (рис. 55) и в этой области функция $f(x, y)$ непрерывна и имеет непрерывные частные производные $f'_x(x, y)$ и $f'_y(x, y)$. Для определения площади поверхности S разобьем область G произвольно на n частей без общих внутренних точек G_i с площадями Δs_i ($i=1, 2, ..., n$) и обозначим через S_i часть поверхности S, проекцией которой на плоскость Oxy является частичная область G_i. Таким образом, поверхность S разобьется на n частей. В каждой части G_i выберем произвольную точку $(\xi_i; \eta_i)$, на поверхности S ей будет соответствовать точка $M_i[\xi_i; \eta_i; f(\xi_i; \eta_i)]$.

Проведем через точку M_i касательную плоскость к поверхности

$$f'_x(\xi_i, \eta_i)(x-\xi_i)+f'_y(\xi_i, \eta_i)(y-\eta_i)-(z-z_i)=0,$$

где x, y, z — текущие координаты, а $\xi_i, \eta_i, z_i=f(\xi_i, \eta_i)$ — координаты точки касания (см. п. 2, § 4, гл. XII). Напомним, что вектор (нормаль) $\bar{n}$, перпендикулярный к плоскости, имеет вид

$$\bar{\eta}=\{-f'_x(\xi_i, \eta_i);\ -f'_y(\xi_i, \eta_i);\ +1\}.$$

Рассмотрим на касательной плоскости ту ее часть, проекцией которой на плоскость Oxy является область G_i. Обозначим эту часть через σ_i, а ее площадь — через $\Delta\sigma_i$. При неограниченном измельчении разбиения области G естественно считать площадь части σ_i касательной плоскости, приближенно равной площади части S_i поверхности, а сумму всех таких площадей

$$\sum_{i=1}^{n}\Delta\sigma_i$$

приближенным значением площади всей поверхности S.

За точное значение площади поверхности S примем по определению предел такой суммы:

$$s = \lim_{\lambda \to 0} \sum_{i=1}^{n} \Delta\sigma_i \,* \tag{1}$$

где λ — наибольший из диаметров частичных областей G_i.

Докажем, что этот предел существует и равен двойному интегралу

$$s = \iint_G \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)}\, dx dy. \tag{2}$$

Обозначим через γ_i острый угол между вектором $\bar{n}$ и осью Oz. Он равен углу между касательной плоскостью в точке M_i и плоскостью Oxy. Так как область G_i есть проекция σ_i на плоскость Oxy, то площади этих областей связаны соотношением

$$\Delta\sigma_i = \frac{\Delta s_i}{\cos\gamma_i}.$$

Действительно, данная формула, как известно, справедлива для треугольников. Она, очевидно, остается справедливой и для плоских многоугольников, так как плоский многоугольник можно разбить на несколько треугольников. Она будет справедливой и для любой плоской фигуры площади $\Delta\sigma$, ограниченной некоторой кривой, так как ее площадь можно рассматривать как предел площадей вписанных в нее многоугольников.

С другой стороны, как известно из аналитической геометрии,

$$\cos\gamma_i = \frac{1}{\sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)}}.$$

---

* Определение предела суммы $\sum_{i=1}^{n} \Delta\sigma_i$ при $\lambda \to 0$ аналогично определению предела интегральной суммы для двойного интеграла (см. сноску на с. 398). В дальнейшем все аналогичные определения будем опускать.

Следовательно,

$$\Delta\sigma_i = \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)}\,\Delta s_i.$$

Подставляя значение $\Delta\sigma_i$ в сумму (1), получим

$$s = \lim_{\lambda \to 0} \sum_{i=1}^{n} \sqrt{1 + f_x'^2(\xi_i, \eta_i) + f_y'^2(\xi_i, \eta_i)}\,\Delta s_i.$$

Полученная сумма представляет собой интегральную сумму для функции $\sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)}$. Так как эта функция по условию непрерывна в области G, то предел этой суммы существует при $\lambda \to 0$ и равен двойному интегралу (2), что и требовалось доказать.

Формула (2) и есть формула, по которой вычисляются площади кривых поверхностей.

Пример 3. Вычислить площадь той части плоскости $6x+3y+2z=12$, которая заключена в первом октанте (рис. 56).

Решение. Так как функция $z = 6 - 3x - \frac{3}{2}y$ и область G, являющаяся проекцией данной части поверхности на плоскость Oxy, удовлетворяют сформулированным выше условиям, вычисляем искомую площадь по формуле (2). Имеем

$$f_x'(x, y) = -3, \quad f_y'(x, y) = -\frac{3}{2}$$

и

$$\sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} = \sqrt{1 + 9 + \frac{9}{4}} = \frac{7}{2}.$$

Областью G является треугольник, ограниченный осями Ox, Oy и прямой $6x+3y=12$, получаемой из уравнения данной плоскости при $z=0$. Расставляя пределы интегрирования в двойном интеграле, получим

$$s = \int_0^2 dx \int_0^{4-2x} \frac{7}{2}\,dy = \frac{7}{2}\int_0^2 [y]_0^{4-2x}\,dx =$$

$$= \frac{7}{2}\int_0^2 (4 - 2x)\,dx = \frac{7}{2}[4x - x^2]_0^2 = \frac{7}{2}\cdot 4 = 14.$$

4. Вычисление массы пластинки. Рассмотрим на плоскости Oxy материальную пластинку, т. е. некоторую область G, по которой распределена масса m с плотностью $\rho(x, y)$. Вычислим по заданной плотности $\rho(x, y)$ массу этой пластинки, считая, что $\rho(x, y)$ — непрерывная функция. Разобьем G произвольно на части G_i и обозначим через m_i массы этих частей. В каждой части G_i $(i=1, 2, ..., n)$ произвольно возьмем точку $(\xi_i; \eta_i)$. Массу m_i каждого такого элемента G_i можно считать равной приближенно $\rho(\xi_i, \eta_i)\Delta s_i$ (Δs_i — площадь G_i), а масса m всей пластинки приближенно равна сумме

$$m=\sum_{i=1}^{n} m_i \approx \sum_{i=1}^{n} \rho(\xi_i, \eta_i)\Delta s_i.$$

Но это — интегральная сумма для непрерывной функции $\rho(x, y)$ в области G. В пределе при $\lambda \to 0$, очевидно, мы получим точное значение массы пластинки, равное двойному интегралу от этой функции по области G, т. е.

$$m=\iint_G \rho(x, y)\,dxdy. \tag{3}$$

Этот пример показывает *физический смысл* двойного интеграла.

Пример 4. Определить массу квадратной пластинки со стороной $2a$, если плотность $\rho(x, y)$ в каждой точке пропорциональна квадрату расстояния от точки пересечения диагоналей.

Решение. Выберем систему координат так, как указано на рис. 57. После этого можно составить функцию плотности $\rho(x, y)$ по условию задачи. Пусть $M(x; y)$ — произвольная точка квадрата $\{(x; y) \mid |x| \leqslant a; |y| \leqslant a\}$. Тогда квадрат расстояния от точки пересечения диагоналей будет равен x^2+y^2. Следовательно, плотность в точке M представится в виде

$$\rho(M)=\rho(x, y)=k(x^2+y^2),$$

где k — коэффициент пропорциональности. По формуле (3) имеем

$$m=\iint_G k(x^2+y^2)\,dxdy.$$

Учитывая, что подынтегральная функция четна относительно x и y, а область интегрирования симметрична относительно осей координат, можно ограничиться вы-

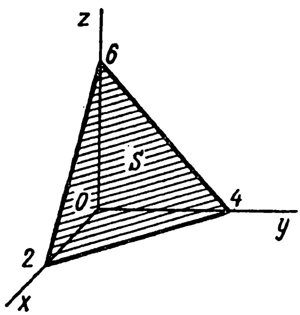


Рис. 56

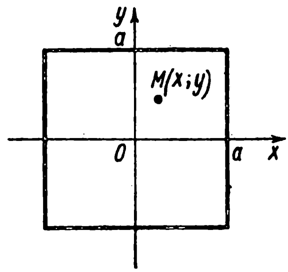


Рис. 57

числением интеграла по той части области G, которая расположена в первой четверти:

$$m = 4k\int_0^a dx\int_0^a (x^2 + y^2)\,dy = 4k\int_0^a \left[x^2y + \frac{y^3}{3}\right]_0^a dx =$$

$$= 4k\int_0^a \left(ax^2 + \frac{a^3}{3}\right) dx = 4k\left[\frac{ax^3}{3} + \frac{a^3x}{3}\right]_0^a =$$

$$= 4k\,\frac{2a^4}{3} = \frac{8}{3}ka^4.$$

5. Вычисление координат центра масс пластинки. Найдем координаты центра масс пластинки, занимающей в плоскости Oxy некоторую область G. Пусть $\rho(x, y)$ — плотность этой пластинки в точке $M(x; y)$. Разбив область G на части G_i $(i=1, 2, ..., n)$, выберем в каждой из этих частей некоторую точку $(\xi_i; \eta_i)$ и будем приближенно считать массу m_i каждой из частей пластинки равной $\rho(\xi_i, \eta_i)\,\Delta s_i$ (Δs_i — площадь G_i). Если считать, что каждая из этих масс сосредоточена в одной точке, а именно в точке $(\xi_i; \eta_i)$, то для координат x_c и y_c центра масс такой системы n материальных точек получаются выражения

$$x_c = \frac{\sum_{i=1}^{n} \xi_i \rho(\xi_i, \eta_i)\Delta s_i}{\sum_{i=1}^{n} \rho(\xi_i, \eta_i)\Delta s_i}, \quad y_c = \frac{\sum_{i=1}^{n} \eta_i \rho(\xi_i, \eta_i)\Delta s_i}{\sum_{i=1}^{n} \rho(\xi_i, \eta_i)\Delta s_i}, \tag{4}$$

которые представляют собой приближенные значения координат центра масс пластинки. Чтобы получить точные значения этих координат, необходимо в (4) перейти к пределу при $\lambda \to 0$. При этом стоящие в (4) интегральные суммы перейдут в соответствующие интегралы, и мы получим, что координаты центра масс пластинки определяются формулами

$$x_c = \frac{\iint\limits_G x\rho(x, y)\,dxdy}{m}, \quad y_c = \frac{\iint\limits_G y\rho(x, y)\,dxdy}{m}, \tag{5}$$

где $m = \iint\limits_G \rho(x, y)\,dxdy$ — масса пластинки.

Если пластинка однородна, т. е. $\rho = \text{const}$, то формулы координат центра масс упрощаются:

$$x_c = \frac{\iint\limits_G xdxdy}{\iint\limits_G dxdy}; \quad y_c = \frac{\iint\limits_G ydxdy}{\iint\limits_G dxdy}. \tag{6}$$

Выражения

$$M_y = \iint\limits_G x\rho(x, y)\,dxdy \text{ и } M_x = \iint\limits_G y\rho(x, y)\,dxdy,$$

стоящие в формулах (5), называются *статическими моментами* пластинки относительно осей Oy и Ox.

Таким образом, вычисление координат центра масс пластинки сводится к вычислению трех двойных интегралов.

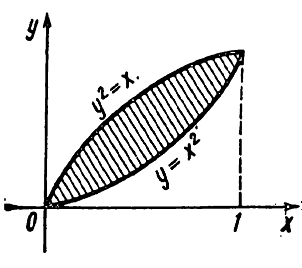


Рис. 58

Пример 5. Найти координаты центра масс однородной пластинки, ограниченной двумя параболами $y^2 = x$ и $x^2 = y$ (рис. 58).

Решение. Координаты центра масс данной пластинки будем находить по формулам (6). Сначала вычислим массу пластинки:

$$m = \iint\limits_G dxdy = \int\limits_0^1 dx \int\limits_{x^2}^{\sqrt{x}} dy = \frac{1}{3}.$$

Затем вычислим статические моменты ее относительно осей координат (числители в (6)):

$$M_y = \iint\limits_G xdxdy = \int\limits_0^1 xdx \int\limits_{x^2}^{\sqrt{x}} dy = \frac{3}{20};$$

$$M_x = \iint\limits_G ydxdy = \int\limits_0^1 dx \int\limits_{x^2}^{\sqrt{x}} ydy = \frac{3}{20}.$$

Теперь по формулам (6) находим

$$x_c = \frac{M_y}{m} = \frac{3}{20} : \frac{1}{3} = \frac{9}{20};\ y_c = \frac{M_x}{m} = \frac{3}{20} : \frac{1}{3} = \frac{9}{20}.$$

Следовательно, $x_c = y_c = \frac{9}{20}$.

6. Вычисление момента инерции пластинки. Как известно, момент инерции материальной точки относительно некоторой оси равен произведению массы точки на квадрат ее расстояния от этой оси, а момент инерции системы материальных точек равен сумме моментов инерций, составляющих эту систему масс.

Пусть область G плоскости Oxy занята пластинкой, имеющей плотность $\rho(x, y)$. Разбив область G на части G_i, площади которых равны Δs_i $(i=1, 2, ..., n)$, и выбрав в каждой из них некоторую точку $(\xi_i; \eta_i)$, заменим нашу пластинку системой масс $m_i=\rho(\xi_i, \eta_i)\,\Delta s_i$, лежащих в точках $(\xi_i; \eta_i)$. Момент инерции такой системы точечных масс относительно, например, оси Oy равен

$$\sum_{i=1}^{n} \xi_i^2 \rho(\xi_i, \eta_i)\,\Delta s_i.$$

Это выражение мы принимаем за приближенное значение момента инерции пластинки. Оно представляет собой интегральную сумму для непрерывной функции $x^2\rho(x, y)$. Переходя к пределу при $\lambda \to 0$, получим для момента инерции пластинки относительно оси Oy следующую формулу:

$$J_y = \iint_G x^2\rho(x, y)\,dxdy.$$

Аналогично момент инерции пластинки относительно оси Ox равен

$$J_x = \iint_G y^2\rho(x, y)\,dxdy.$$

Найдем еще момент инерции J_0 пластинки относительно начала координат. Приняв во внимание, что момент инерции материальной точки с массой m относительно начала координат равен

$$m(x^2+y^2),$$

и воспользовавшись рассуждениями, аналогичными проведенным выше, получим, что

$$J_0 = \iint_G (x^2 + y^2)\,\rho(x, y)\,dxdy, \tag{7}$$

т. е.

$$J_0 = J_x + J_y.$$

Пример 6. Найти момент инерции круга радиуса R с постоянной плотностью $\rho(x, y) = 1$ относительно начала координат.

Решение. По формуле (7) имеем

$$J_0 = \iint_G (x^2 + y^2)\,dxdy.$$

Для вычисления этого интеграла перейдем к полярным координатам. Уравнение окружности в полярных координатах есть $\rho = R$. Поэтому

$$J_0 = \int_0^{2\pi} d\varphi \int_0^R \rho^2\cdot\rho d\rho = \int_0^{2\pi} \left[\frac{\rho^4}{4}\right]_0^R d\varphi = \frac{1}{4}[R^4\varphi]_0^{2\pi} = \frac{\pi R^4}{2}.$$

§ 5. Криволинейные интегралы

В этом параграфе мы рассмотрим обобщение понятия определенного интеграла на случай, когда областью интегрирования является отрезок некоторой кривой, лежащей в плоскости.

Такого рода интегралы называются *криволинейными*. Они имеют широкое применение в различных разделах математики.

Различают два типа криволинейных интегралов, называемых обычно криволинейными интегралами первого и второго рода.

1. Определение криволинейного интеграла первого рода. Рассмотрим на плоскости Oxy некоторую кривую AB, гладкую или кусочно-гладкую *, и предположим, что функция $z=f(x, y)$ определена вдоль кривой AB.

Разобьем кривую AB произвольно на n частей точками

$$A=M_0,\ M_1,\ M_2,\ \ldots,\ M_{i-1},\ M_i,\ \ldots,\ M_n=B,$$

выберем на каждой из частичных дуг $M_{i-1}M_i$ произвольную точку M_i^* (рис. 59) и составим сумму

$$\sum_{i=1}^{n} f(M_i^*)\,\Delta l_i, \tag{1}$$

где Δl_i — длина дуги $M_{i-1}M_i$. Сумма (1) называется *интегральной суммой* для функции $z=f(x, y)=f(M)$. Обозначим через λ наибольшую из длин частичных дуг $M_{i-1}M_i$ $(\lambda=\max\limits_{1\leqslant i\leqslant n}\{\Delta l_i\})$ и дадим следующее

Определение. *Если интегральная сумма (1) при $\lambda\to 0$ имеет конечный предел J, то этот предел называется криволинейным интегралом первого рода от функции $f(x, y)$ по кривой AB и обозначается одним из символов*

$$J=\int\limits_{AB} f(M)\,dl=\int\limits_{AB} f(x, y)\,dl.$$

---

* Кривая, заданная уравнениями $x=\varphi(t)$, $y=\psi(t)$, $\alpha\leqslant t\leqslant\beta$, называется *гладкой*, если функции $\varphi(t)$ и $\psi(t)$ непрерывны и имеют непрерывные производные $\varphi'(t)$ и $\psi'(t)$, не обращающиеся в нуль одновременно (тем самым кривая в каждой точке имеет касательную). Непрерывная кривая, составленная из конечного числа гладких кусков, называется *кусочно-гладкой*.

В этом случае функция $f(x, y)$ называется *интегрируемой* вдоль кривой AB, кривая AB — контуром интегрирования, точка A — *начальной*, а точка B — *конечной* точками интегрирования.

Понятие криволинейного интеграла первого рода легко сводится к понятию определенного интеграла. Действительно, приняв на кривой AB за параметр длину дуги l, отсчитываемую от точки A, получим параметрическое представление кривой

$$x=x(l),\ y=y(l)\quad (0\leqslant l\leqslant L).$$

При этом функция $f(x, y)$ заданная вдоль AB, становится сложной непрерывной функцией параметра

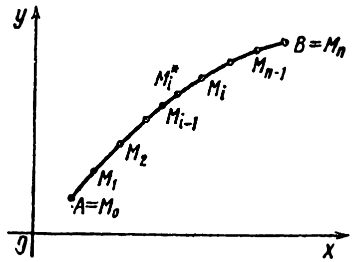


Рис. 59

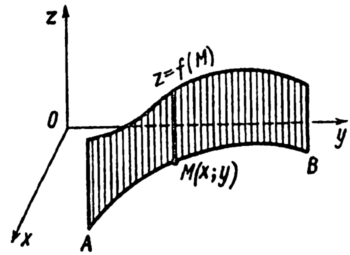


Рис. 60

$l: f[x(l), y(l)]$. Обозначив l_i^* значение параметра l, отвечающее точке M_i^*, перепишем интегральную сумму (1) в виде

$$\sum_{i=1}^{n} f[x(l_i^*),\ y(l_i^*)]\,\Delta l_i, \tag{2}$$

где $\Delta l_i=l_i-l_{i-1}$ и $l_{i-1}\leqslant l_i^*\leqslant l_i$. Сумма (2) является интегральной для определенного интеграла от непрерывной функции $f[x(l), y(l)]$ на отрезке $[0, L]$. А раз интегральные суммы (1) и (2) равны между собой, то равны и отвечающие им интегралы, т. е.

$$\int\limits_{AB} f(x, y)\,dl=\int\limits_{0}^{L} f[x(l),\ y(l)]\,dl. \tag{3}$$

Заметим, что (3) не только выражает криволинейный интеграл через определенный, но и доказывает существование криволинейного интеграла от функции $f(x, y)$, непрерывной на рассматриваемой кривой AB.*

Хотя, как это было показано, криволинейный интеграл первого рода непосредственно сводится к определенному, между этими понятиями имеется следующее различие. В интегральной сумме (1) величины Δl_i обязательно положительны, независимо от того, какую точку кривой AB мы считаем начальной, а какую — конечной, т. е.

$$\int\limits_{AB} f(x, y)\, dl = \int\limits_{BA} f(x, y)\, dl,$$

в то время как определенный интеграл $\int\limits_a^b f(x)\, dx$ при перестановке пределов интегрирования меняет знак. А в остальном криволинейный интеграл первого рода обладает теми же свойствами, что и определенный интеграл, что непосредственно вытекает из формулы (3).

Криволинейный интеграл первого рода, так же как и определенный, имеет геометрический смысл. Если определенный интеграл $\int\limits_a^b f(x)\, dx$ при $f(x) \geqslant 0$ представляет собой площадь криволинейной трапеции, то криволинейный интеграл $\int\limits_{AB} f(M)\, dl$ при $f(M) \geqslant 0$ численно равен площади куска цилиндрической поверхности, составленной из перпендикуляров к плоскости Oxy, восстановленных в точках $M(x; y)$ кривой AB и имеющих переменную длину $f(M)$ (рис. 60).

В частности, если AB — не кривая, а отрезок прямой $[a, b]$, расположенный на оси Ox, то $f(x, y) = f(x)$, $\Delta l_i = \Delta x_i$ и криволинейный интеграл будет обычным определенным интегралом.

Наконец, если положить $f(M) \equiv 1$, то получим криволинейный интеграл $\int\limits_{AB} dl$, значение которого есть длина дуги кривой AB.

Таким образом, с помощью криволинейного интеграла первого рода можно вычислить площади цилиндрических поверхностей и длины дуг. Кроме этого, криволинейный интеграл первого рода имеет широкое приме-

* Непрерывность функции $f(x, y) = f(M)$ вдоль кривой AB означает, что $\lim\limits_{M \to M_0} f(M) = f(M_0)$ в любой точке M_0 кривой AB, где M также точка этой кривой.

нение в физике. С его помощью можно, как мы делали для случая двойных интегралов, найти массу материальной кривой по ее плотности, найти ее моменты инерции относительно координатных осей, найти координаты центра масс такой кривой и т. д.

2. Вычисление криволинейных интегралов первого рода. Вычисление криволинейных интегралов первого рода сводится, как мы сейчас покажем, к вычислению определенных интегралов.

Пусть кривая AB задана параметрически уравнениями

$$x=\varphi(t), \ y=\psi(t) \quad (\alpha\leqslant t\leqslant\beta),$$

где $\varphi(t)$ и $\psi(t)$ — непрерывные вместе со своими производными $\varphi'(t)$ и $\psi'(t)$ функции, а $f(x, y)$ — функция, непрерывная вдоль этой кривой, причем для определенности будем считать, что точке A соответствует значение $t=\alpha$, точке B — значение $t=\beta$. Тогда для любой точки $M(x; y)$ кривой AB длину l дуги AM можно рассматривать как функцию параметра t, $l=l(t)$ и вычислить ее (см. ч. I, гл. VIII, § 10, п. 3) по формуле

$$l=l(t)=\int_{\alpha}^{t}\sqrt{[\varphi'(t)]^2+[\psi'(t)]^2}\,dt,$$

откуда, согласно правилу дифференцирования интеграла по верхнему пределу, находим

$$dl=\sqrt{[\varphi'(t)]^2+[\psi'(t)]^2}\,dt, \text{ или } dl=\sqrt{(dx)^2+(dy)^2}. \quad (4)$$

Производя замену переменной в определенном интеграле равенства (3), т. е. выражая переменную l через t, с учетом (4) получим

$$\int_{AB} f(x, y)\,dl=\int_{0}^{L} f[x(l), y(l)]\,dl=$$
$$=\int_{\alpha}^{\beta} f[\varphi(t), \psi(t)]\sqrt{[\varphi'(t)]^2+[\psi'(t)]^2}\,dt. \quad (5)$$

Пример 1. Вычислить криволинейный интеграл $\int_{AB} y^2dl$, где AB — часть окружности: $x=a\cos t$, $y=a\sin t$, $0\leqslant t\leqslant\frac{\pi}{2}$.

Решение. Так как

$$y^2 = a^2 \sin^2 t, \quad dl = \sqrt{a^2 \sin^2 t + a^2 \cos^2 t}\, dt = a dt,$$

то по формуле (5) получаем

$$\int\limits_{AB} y^2 dl = \int\limits_0^{\frac{\pi}{2}} a^2 \sin^2 t \cdot a dt = \frac{a^3}{2} \int\limits_0^{\frac{\pi}{2}} (1 - \cos 2t)\, dt =$$

$$= \frac{a^3}{2} \left[t - \frac{\sin 2t}{2} \right]_0^{\frac{\pi}{2}} = \frac{a^3 \pi}{4}.$$

В частности, если кривая AB задана уравнением вида $y = y(x)$, $a \leqslant x \leqslant b$, где $y(x)$ — непрерывно дифференцируемая функция*, то, принимая x за параметр ($t = x$), из формулы (5) получаем

$$\int\limits_{AB} f(x, y)\, dl = \int\limits_a^b f[x, y(x)] \sqrt{1 + y'^2(x)}\, dx. \tag{6}$$

Пример 2. Вычислить криволинейный интеграл $\int\limits_{AB} y dl$, где AB — дуга параболы $y^2 = 2x$ от точки (0; 0) до точки (2; 2).

Решение. Имеем

$$y = \sqrt{2x}, \; y' = \frac{2}{\sqrt{2x}}, \; dl = \sqrt{1 + y'^2}\, dx =$$

$$= \sqrt{1 + \frac{1}{2x}}\, dx.$$

По формуле (6) получаем

$$\int\limits_{AB} y dl = \int\limits_0^2 \sqrt{2x} \sqrt{1 + \frac{1}{2x}}\, dx = \int\limits_0^2 \sqrt{2x + 1}\, dx =$$

$$= \frac{1}{3} [(2x + 1)^{\frac{3}{2}}]_0^2 = \frac{1}{3} (5\sqrt{5} - 1).$$

Замечание. Формула (4) представляет самостоятельный интерес. Она дает простое геометрическое ис-

---

* То есть имеющая непрерывную производную.

толкование дифференциала дуги dl. Учитывая, что дифференциал функции равен приращению ординаты касательной (см. ч. I, гл. V, § 3, п. 1), из формулы (4) получаем, что дифференциал дуги dl (см. рис. 64) геометрически равен длине отрезка касательной к кривой AB от точки касания с абсциссой x до точки $(x+dx; y+dy)$, т. е. является гипотенузой прямоугольного треугольника с катетами $|dx|$ и $|dy|$.

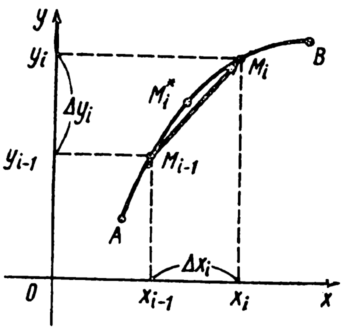


Рис. 61

3. Определение криволинейного интеграла второго рода. Пусть вдоль кривой AB определены две функции $P(x, y)$ и $Q(x, y)$. Обозначим через Δx_i и Δy_i проекции i-й частичной дуги $M_{i-1}M_i$* на оси координат (рис. 61) и составим для функции $P(x, y)\,[Q(x, y)]$ интегральную сумму следующего вида:

$$\sum_{i=1}^{n} P(M_i^*)\,\Delta x_i \left[\sum_{i=1}^{n} Q(M_i^*)\,\Delta y_i\right]. \tag{7}$$

Определение. *Если интегральная сумма (7) при* $\lambda \to 0$ $(\lambda = \max\limits_{1 \leqslant i \leqslant n}\{\Delta l_i\})$ *имеет конечный предел J, то этот предел называется криволинейным интегралом второго рода от функции $P(x, y)[Q(x, y)]$ по кривой AB и обозначается символом*

$$\int\limits_{AB} P(x, y)\,dx \left[\int\limits_{AB} Q(x, y)\,dy\right].$$

Сумму

$$\int\limits_{AB} P(x, y)\,dx + \int\limits_{AB} Q(x, y)\,dy$$

принято называть *общим криволинейным интегралом второго рода* и обозначать символом

* Под проекцией дуги $M_{i-1}M_i$ на ось координат понимается проекция вектора $\overline{M_{i-1}M_i}$.

$$\int_{AB} P(x, y)\,dx + Q(x, y)\,dy^*.$$

Криволинейные интегралы второго рода, так же как и первого рода, легко сводятся к определенным интегралам.

Действительно, пусть кривая AB задана параметрически уравнениями $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leqslant t \leqslant \beta$, где $\varphi(t)$ и $\psi(t)$ — непрерывные вместе со своими производными $\varphi'(t)$ и $\psi'(t)$ функции, причем точке A кривой соответствует значение $t = \alpha$, точке B — значение $t = \beta$, $\varphi'^2(t) + \psi'^2(t) \neq 0$. Пусть функции $P(x, y)$ и $Q(x, y)$ непрерывны вдоль кривой AB. Тогда справедливы следующие формулы:

$$\int_{AB} P(x, y)\,dx = \int_{\alpha}^{\beta} P[\varphi(t), \psi(t)]\,\varphi'(t)\,dt,$$

$$\int_{AB} Q(x, y)\,dy = \int_{\alpha}^{\beta} Q[\varphi(t), \psi(t)]\,\psi'(t)\,dt, \tag{8}$$

$$\int_{AB} P(x, y)\,dx + Q(x, y)\,dy =$$

$$= \int_{\alpha}^{\beta} \{P[\varphi(t), \psi(t)]\,\varphi'(t) + Q[\varphi(t), \psi(t)]\,\psi'(t)\}\,dt,$$

сводящие криволинейные интегралы к определенным интегралам.

Ограничимся доказательством первой из формул (8)

$$\int_{AB} P(x, y)\,dx = \int_{\alpha}^{\beta} P[\varphi(t), \psi(t)]\,\varphi'(t)\,dt, \tag{9}$$

остальные две формулы доказываются аналогично.

Пусть точкам M_i разбиения кривой AB соответствуют значения параметра t_i, M_i^* — значения t_i^*, т. е. M_i

---

* Вместо $P(x, y)$ и $Q(x, y)$ мы будем иногда писать просто P и Q, а криволинейный интеграл писать в виде $\int_{AB} P\,dx + Q\,dy$.

есть точка $[\varphi(t_i);\ \psi(t_i)]$, а M_i^* — точка $[\varphi(t_i^*);\ \psi(t_i^*)]$, $i=1, 2, ..., n$. Функция $P(x, y)$ на кривой является сложной функцией параметра t: $P[\varphi(t), \psi(t)]$. Так как функции $x=\varphi(t)$ и $y=\psi(t)$ непрерывны на отрезке $[\alpha, \beta]$, а функция $P(x, y)$ непрерывна вдоль кривой AB, то по теореме о непрерывности сложной функции функция $P[\varphi(t), \psi(t)]$ непрерывна на отрезке $[\alpha, \beta]$. Составим интегральную сумму (7) для функции $P(x, y)$:

$$\sigma=\sum_{i=1}^{n} P(M_i^*)\,\Delta x_i=\sum_{i=1}^{n} P\,[\varphi(t_i^*),\ \psi(t_i^*)]\,\Delta x_i.$$

Так как $\Delta x_i=x_i-x_{i-1}=\varphi(t_i)-\varphi(t_{i-1})$, то по формуле Ньютона—Лейбница

$$\Delta x_i=\varphi(t_i)-\varphi(t_{i-1})=\int\limits_{t_{i-1}}^{t_i}\varphi'(t)\,dt.$$

Поэтому

$$\sigma=\sum_{i=1}^{n} P\,[\varphi(t_i^*),\ \psi(t_i^*)]\int\limits_{t_{i-1}}^{t_i}\varphi'(t)\,dt=$$

$$=\sum_{i=1}^{n}\int\limits_{t_{i-1}}^{t_i} P\,[\varphi(t_i^*),\ \psi(t_i^*)]\,\varphi'(t)\,dt.$$

С другой стороны, функция $P[\varphi(t), \psi(t)]\,\varphi'(t)$ также является непрерывной функцией на $[\alpha, \beta]$ и для нее существует определенный интеграл, стоящий справа в формуле (9). Запишем его в виде суммы интегралов по частичным отрезкам $[t_{i-1}, t_i]$:

$$J=\sum_{i=1}^{n}\int\limits_{t_{i-1}}^{t_i} P\,[\varphi(t),\ \psi(t)]\,\varphi'(t)\,dt.$$

Рассмотрим и оценим разность

$$\sigma-J=\sum_{i=1}^{n}\int\limits_{t_{i-1}}^{t_i}\{P\,[\varphi(t_i^*),\ \psi(t_i^*)]-P\,[\varphi(t),\ \psi(t)]\}\,\varphi'(t)\,dt. \tag{10}$$

Из непрерывности функции $P[\varphi(t), \psi(t)]$ на $[\alpha, \beta]$ по теореме Кантора следует ее равномерная непрерывность. А это означает, что для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\mu = \max\limits_{1 \leqslant i \leqslant n} \{\Delta t_i\} < \delta$

$$|P[\varphi(\overset{*}{t}_i), \psi(\overset{*}{t}_i)] - P[\varphi(t), \psi(t)]| < \varepsilon. \tag{11}$$

Из непрерывности функции $\varphi'(t)$ на $[\alpha, \beta]$ следует ее ограниченность на $[\alpha, \beta]$, т. е. существует число k такое, что

$$|\varphi'(t)| \leqslant k. \tag{12}$$

Используя (11) и (12), получаем для разности (10) следующие оценки:

$$|\sigma - J| \leqslant \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |P[\varphi(\overset{*}{t}_i), \psi(\overset{*}{t}_i)] - P[\varphi(t), \psi(t)]| \times$$

$$\times |\varphi'(t)| \, dt < \varepsilon k \sum_{i=1}^{n} (t_i - t_{i-1}) = \varepsilon k (\beta - \alpha).$$

Отсюда, в силу произвольности ε, следует, что

$$\lim_{\mu \to 0} \sigma = J. \tag{13}$$

Но при $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta l_i\} \to 0$ также $\mu = \max\limits_{1 \leqslant i \leqslant n} \{\Delta t_i\} \to 0$ и наоборот*, а тогда из (13) следует, что и

$$\lim_{\lambda \to 0} \sigma = J,$$

т. е. существование криволинейного интеграла

---

* В самом деле, $\Delta l_i = \int\limits_{t_{i-1}}^{t_i} \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2} \, dt$. Из непрерывности функций $\varphi'(t)$ и $\psi'(t)$ на $[\alpha; \beta]$ следует непрерывность функции $\sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}$ на $[\alpha, \beta]$. Но тогда $m\Delta t_i \leqslant \Delta l_i \leqslant M\Delta t_i$, где m и M — минимальное и максимальное значения функции $\sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}$ на отрезке $[\alpha, \beta]$. Из левого неравенства следует, что $\mu \to 0$ при $\lambda \to 0$, а из правого, что $\lambda \to 0$ при $\mu \to 0$.

$$\int_{AB} P(x, y)\,dx$$

и справедливость формулы (9).

Криволинейный интеграл второго рода обладает свойствами, аналогичными свойствам определенного интеграла, что непосредственно вытекает из формул (8).

В отличие от криволинейного интеграла первого рода криволинейный интеграл второго рода зависит от того, в каком направлении (от A к B или от B к A) пробегается кривая AB, и меняет знак при изменении направления обхода кривой, т. е.

$$\int_{AB} P(x, y)\,dx = -\int_{BA} P(x, y)\,dx,$$

$$\int_{AB} Q(x, y)\,dy = -\int_{BA} Q(x, y)\,dy.$$

Действительно, изменив направление обхода кривой, мы соответственно изменим знаки проекций Δx_i и Δy_i в суммах (7), и, следовательно, сами суммы и их пределы изменят знак.

Из сказанного следует, что при вычислении криволинейных интегралов второго рода необходимо учитывать направление интегрирования.

В случае, когда L — замкнутая кривая, т. е. когда точка B совпадает с точкой A, из двух возможных направлений обхода замкнутого контура L условимся называть *положительным* то, при котором область, лежащая внутри этого контура, остается слева по отношению к точке, совершающей обход. Противоположное направление обхода контура L условимся называть *отрицательным*.

Криволинейный интеграл по замкнутому контуру L, пробегаемому в положительном направлении, часто обозначают символом

$$\oint_{L} P(x, y)\,dx + Q(x, y)\,dy.$$

4. Вычисление криволинейных интегралов второго рода. Криволинейные интегралы второго рода вычисляются путем сведения их к определенным интегралам по формулам (8).

В частности, если кривая AB задана уравнением вида $y=y(x)$, $a\leqslant x\leqslant b$, где $y(x)$ — непрерывно дифференцируемая функция, то, принимая x за параметр ($t=x$), из формул (8) получим

$$
\left.
\begin{aligned}
&\int_{AB} P(x, y)\,dx = \int_a^b P[x, y(x)]\,dx,\\
&\int_{AB} Q(x, y)\,dy = \int_a^b Q[x, y(x)]\,y'(x)\,dx,\\
&\int_{AB} P(x, y)\,dx + Q(x, y)\,dy =\\
&= \int_a^b \{P[x, y(x)] + Q[x, y(x)]\,y'(x)\}\,dx.
\end{aligned}
\right\} \qquad (14)
$$

Рассмотрим примеры.

Пример 3. Вычислить интеграл

$$\int_{AB} x^2\,dx + xy\,dy,$$

где AB — четверть окружности

$$x=\cos t,\ \ y=\sin t,\ \ 0\leqslant t\leqslant \frac{\pi}{2}.$$

A соответствует $t=0$, B соответствует $t=\frac{\pi}{2}$.

Решение. Имеем

$$x^2=\cos^2 t,\ dx=-\sin t\,dt,\ xy=\cos t\cdot\sin t,\ dy=\cos t\,dt.$$

По третьей формуле (8) получаем

$$\int_{AB} x^2\,dx + xy\,dy = \int_0^{\frac{\pi}{2}} (-\cos^2 t\sin t + \cos^2 t\sin t)\,dt = 0.$$

Пример 4. Вычислить интеграл $\oint_L (x+y)\,dy$, где L — контур прямоугольника, образованного прямыми $x=0$, $y=0$, $x=1$, $y=1$ (рис. 62).

Решение. На рис. 62 положительное направление обхода контура L обозначено стрелками. Разбивая весь промежуток интегрирования на части, можем записать

$$\oint_L = \int_{AB} + \int_{BC} + \int_{CD} + \int_{DA}.$$

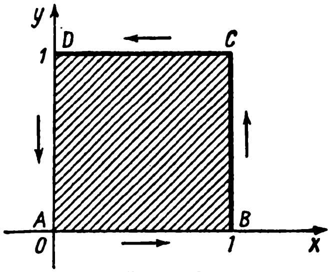


Рис. 62

Легко заметить, что интегралы вдоль участков AB и CD равны нулю, так как на них y постоянна и, следовательно, $dy=0$. Поэтому остается вычислить интегралы по участкам BC и DA. По второй формуле (14) получаем

$$\int_{BC} (x+y)\,dy = \int_0^1 (1+y)\,dy = \left[y + \frac{y^2}{2}\right]_0^1 = \frac{3}{2},$$

$$\int_{DA} (x+y)\,dy = \int_1^0 (0+y)\,dy = \left[\frac{y^2}{2}\right]_1^0 = -\frac{1}{2}.$$

Таким образом, окончательно имеем

$$\oint_L (x+y)\,dy = \frac{3}{2} - \frac{1}{2} = 1.$$

Пример 5. Вычислить интеграл

$$\int_{AB} 3x^2y\,dx + (x^3+1)\,dy,$$

где

а) AB — прямая $y=x$, соединяющая точки (0; 0) и (1; 1);

б) AB — парабола $y=x^2$, соединяющая те же точки;

в) AB — ломаная, проходящая через точки (0; 0); (1; 0), (1; 1) (рис. 63).

Решение. По третьей формуле (14) имеем

а) $\int_{AB} 3x^2y\,dx + (x^3+1)\,dy = \int_0^1 (4x^3+1)\,dx = 2;$

б) $\int\limits_{AB} 3x^2y\,dx + (x^3+1)\,dy = \int\limits_0^1 (5x^4+2x)\,dx = 2;$

в) $\int\limits_{AB} 3x^2y\,dx + (x^3+1)\,dy = \int\limits_0^1 3x^2y\,dx + \int\limits_0^1 (x^3+1)\,dy =$

$$= \int\limits_0^1 2dy = 2.$$

Заметим, что, взяв три различных пути, соединяющие одни и те же точки, мы получили три одинаковых результата. Это обстоятельство не является случайным. Причина его будет раскрыта в § 7.

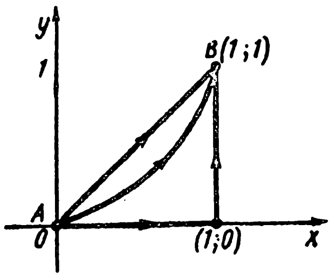


Рис. 63

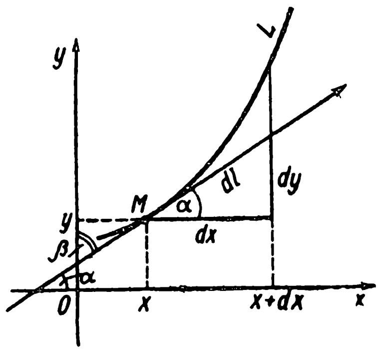


Рис. 64

5. Связь между криволинейными интегралами первого и второго рода. Для рассмотрения этого вопроса обозначим через α и β углы, составляемые с осями координат направленной* касательной к кривой AB в некоторой точке $M(x;\ y)$ (рис. 64), тогда получим соотношения

$$dx = \cos\alpha dl,\quad dy = \cos\beta dl^{**}. \qquad (15)$$

Заменяя в криволинейных интегралах второго рода dx и dy их выражениями (15), мы и производим преобра-

* За положительное направление касательной примем то, которое соответствует направлению движения точки по кривой.

** См. замечание в п. 2.

зование этих интегралов в криволинейные интегралы первого рода:

$$\int\limits_{AB} P(x, y)\,dx = \int\limits_{AB} P(x, y)\cos\alpha\,dl,$$

$$\int\limits_{AB} Q(x, y)\,dy = \int\limits_{AB} Q(x, y)\cos\beta\,dl, \tag{16}$$

$$\int\limits_{AB} P(x, y)\,dx + Q(x, y)\,dy =$$

$$= \int\limits_{AB} [P(x, y)\cos\alpha + Q(x, y)\cos\beta]\,dl.$$

Таким образом, формулы (16) выражают криволинейные интегралы второго рода через криволинейные интегралы первого рода и устанавливают связь между ними. При изменении направления движения точки по кривой на противоположное $\cos\alpha$, $\cos\beta$, dx и dy меняют знак, и формулы (16) остаются в силе.

Заканчивая данный параграф, заметим, что мы рассмотрели криволинейные интегралы для плоских кривых. Однако все сказанное о них может быть без особого труда перенесено и на пространственные кривые.

Пусть AB — пространственная кривая и вдоль этой кривой определены функции $f(x, y, z)$, $P(x, y, z)$, $Q(x, y, z)$ и $R(x, y, z)$. Тогда по аналогии со случаем плоской кривой можно определить криволинейные интегралы первого рода

$$\int\limits_{AB} f(x, y, z)\,dl$$

и криволинейные интегралы второго рода

$$\int\limits_{AB} P(x, y, z)\,dx, \quad \int\limits_{AB} Q(x, y, z)\,dy, \quad \int\limits_{AB} R(x, y, z)\,dz,$$

$$\int\limits_{AB} P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz.$$

Техника вычисления таких интегралов принципиально ничем не отличается от техники вычисления интегралов по плоской кривой.

§ 6. Формула Грина

Формула Грина* устанавливает связь между криволинейными интегралами и двойными интегралами. Она имеет широкое применение в математическом анализе и его приложениях.

Мы докажем эту формулу для замкнутой области, граница которой пересекается с прямыми, параллельными осям координат, не более чем в двух точках. Для краткости будем называть такие области *простыми*. Предполагается, что контур, ограничивающий область, гладкий или кусочно-гладкий.

Теорема 13.6. *Пусть G — некоторая простая замкнутая область, ограниченная контуром L, и пусть функции $P(x, y)$ и $Q(x, y)$ непрерывны вместе со своими частными производными $\frac{\partial P}{\partial y}$ и $\frac{\partial Q}{\partial x}$ в данной области. Тогда имеет место формула*

$$\iint\limits_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy = \oint\limits_L P\,dx + Q\,dy, \tag{1}$$

называемая формулой Грина.

Доказательство. Пусть контур L, ограничивающий область G, может быть задан как уравнениями $x=x_1(y)$, $x=x_2(y)$ $(c \leqslant y \leqslant d)$, так и уравнениями $y=y_1(x)$, $y=y_2(x)$ $(a \leqslant x \leqslant b)$ (рис. 65). Рассмотрим сначала область G, определенную неравенствами $a \leqslant x \leqslant b$, $y_1(x) \leqslant y \leqslant y_2(x)$, и преобразуем двойной интеграл

$$\iint\limits_G \frac{\partial P}{\partial y}\,dx\,dy$$

в криволинейный. Для этого сведем его к повторному интегралу и по формуле Ньютона — Лейбница выполним интегрирование по y. Получим

$$\iint\limits_G \frac{\partial P}{\partial y}\,dx\,dy =$$

$$= \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y}\,dy = \int\limits_a^b [P(x, y_2(x)) - P(x, y_1(x))]\,dx =$$

---

* Грин Джордж (1793—1841) — английский математик и физик.

$$= \int_a^b P(x, y_2(x))\,dx - \int_a^b P(x, y_1(x))\,dx.$$

Каждый из этих двух определенных интегралов представляет собой криволинейный интеграл, взятый по соответствующей кривой (см. (14), §5), а именно:

$$\int_a^b P(x, y_2(x))\,dx = -\int_{BDA} P(x, y)\,dx;$$

$$\int_a^b P(x, y_1(x))\,dx = \int_{ACB} P(x, y)\,dx.$$

Таким образом,

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy = -\left[\int_{BDA} P(x, y)\,dx + \int_{ACB} P(x, y)\,dx\right],$$

т. е.

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy = -\oint_L P(x, y)\,dx. \qquad (2)$$

Аналогично доказывается формула

$$\iint_G \frac{\partial Q}{\partial x}\,dx\,dy = \oint_L Q(x, y)\,dy. \qquad (3)$$

(При этом область G задается неравенствами $c \leqslant y \leqslant d$, $x_1(y) \leqslant x \leqslant x_2(y)$).

Вычитая из равенства (3) почленно равенство (2), получим искомую формулу (1) ■

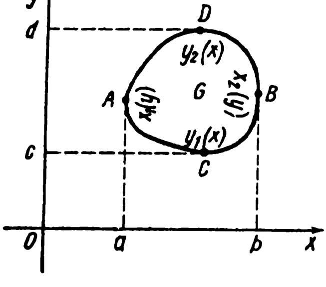


Рис. 65

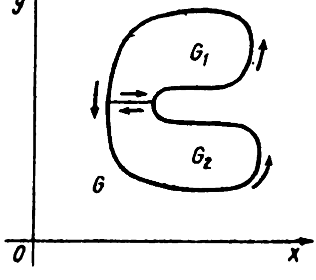


Рис. 66

Замечание 1. Формула Грина остается справедливой для всякой замкнутой области G, которую можно разбить проведением дополнительных линий на конечное число простых замкнутых областей. Действительно, пусть область G с границей L имеет вид, изображенный на рис. 66. Разобьем ее на две простые области G_1 и G_2, для каждой из которых справедлива формула (1). Напишем отдельно формулу Грина для G_1 и G_2 и сложим почленно полученные равенства. Будем иметь слева двойной интеграл по всей области G, а справа криволинейный интеграл по контуру L области G, так как криволинейный интеграл по вспомогательной кривой берется дважды в противоположных направлениях и при суммировании взаимно уничтожается.

Замечание 2. Можно доказать, что формула Грина справедлива для любой области G, ограниченной несколькими кусочно-гладкими контурами. Однако останавливаться на доказательстве этого факта мы не будем.

Пример. По формуле Грина вычислить криволинейный интеграл $\oint\limits_L (x-y)\,dx+(x+y)\,dy$, где L — окружность $x^2+y^2=R^2$.

Решение. Функции

$$P(x, y)=x-y,\ Q(x, y)=x+y \text{ и } \frac{\partial P}{\partial y}=-1,\ \frac{\partial Q}{\partial x}=1$$

непрерывны в замкнутом круге $x^2+y^2=R^2$. Следовательно, по теореме 13.6 формула Грина применима к данному интегралу. Получаем

$$\oint\limits_L (x-y)\,dx+(x+y)\,dy=\iint\limits_G [1-(-1)]\,dx\,dy=$$

$$=2\iint\limits_G dx\,dy=2s=2\pi R^2.$$

Заметим, что полученный результат легко проверить непосредственным вычислением данного интеграла.

§ 7. Условия независимости криволинейного интеграла от пути интегрирования

Как мы уже отмечали при решении примера 5 (см. п. 4, § 5), в некоторых случаях величина криволинейного интеграла $\int\limits_{AB} P\,dx + Q\,dy$ не зависит от формы пути интегрирования, а зависит только от начальной и конечной точек интегрирования. Выясним, при каких условиях такая независимость интеграла от выбора пути имеет место.

В исследовании этого вопроса существенную роль играет формула Грина.

Для дальнейшего нам необходимо уточнить, с какими областями мы будем иметь дело. Введем следующее

Определение. Плоская область G называется односвязной, если каков бы ни был замкнутый контур L, лежащий внутри этой области, ограниченная этим контуром часть плоскости целиком принадлежит области G.

Образно говоря, односвязность области означает, что область не имеет «дыр». Например, односвязными областями будут внутренности круга, эллипса, многоугольника и т. п. Простейшим примером неодносвязной области служит область, заключенная между окружностями $x^2+y^2=1$ и $x^2+y^2=3$, так как окружность $x^2+y^2=$ $=2$, лежащая в этой области, содержит внутри себя точки, не принадлежащие данной области, например начало координат $(0;\,0)$.

Перейдем теперь к доказательству основной теоремы нашего исследования.

Теорема 13.7. *Пусть функции $P(x,\,y)$ и $Q(x,\,y)$ определены и непрерывны вместе со своими частными производными $\frac{\partial P}{\partial y}$ и $\frac{\partial Q}{\partial x}$ в некоторой односвязной замкнутой области G. Тогда следующие четыре условия эквивалентны, т. е. выполнение любого из них влечет за собой выполнение остальных трех:*

1. *Для любой замкнутой кусочно-гладкой кривой L, расположенной в G,*

$$\oint\limits_L P\,dx + Q\,dy = 0.$$

2. Для любых двух точек A и B области G значение интеграла

$$\int_{AB} P\,dx + Q\,dy$$

не зависит от выбора пути интегрирования, целиком лежащего в G.

3. Выражение $Pdx+Qdy$ представляет собой полный дифференциал некоторой функции, определенной в области G. Иными словами, существует такая функция $F(x, y)$, определенная в G, что*

$$dF = Pdx + Qdy.$$

4. В области G всюду

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}. \qquad (1)$$

Доказательство. Доказательство теоремы проведем по схеме

$$1 \to 2 \to 3 \to 4 \to 1,$$

т. е. покажем, что из первого условия следует второе, из второго — третье, из третьего — четвертое, а из четвертого — снова первое. Тем самым будет доказана эквивалентность условий 1, 2, 3, 4.

Первый этап: $1 \to 2$. Рассмотрим в области G два произвольных пути, соединяющих точки A и B: ACB и ADB — любые две кусочно-гладкие кривые (рис. 67). В сумме они составляют замкнутую кривую $L = ACB + BDA$, расположенную в G. По условию 1

$$\oint_L P\,dx + Q\,dy = 0.$$

Но

$$\oint_L Pdx + Qdy = \int_{ACB} P\,dx + Q\,dy + \int_{BDA} P\,dx + Q\,dy =$$

$$= \int_{ACB} P\,dx + Q\,dy - \int_{ADB} P\,dx + Q\,dy.$$

Следовательно,

$$\int_{ACB} P\,dx + Q\,dy = \int_{ADB} P\,dx + Q\,dy,$$

т. е. условие 2 выполняется.

---

* В теории криволинейных интегралов второго рода дифференциал функции $F(x, y)$ обычно называют полным дифференциалом.

Второй этап: 2→3. Пусть интеграл $\int\limits_{AB} P\,dx + Q\,dy$ не зависит от пути интегрирования. Тогда если точку A зафиксировать: $A = A(x_0;\ y_0)$, то этот интеграл будет некоторой функцией координат x и y точки $B = B(x;\ y)$:

$$\int\limits_{AB} P\,dx + Q\,dy = F(x, y).$$

Покажем, что функция $F(x, y)$ дифференцируема и что

$$dF = P\,dx + Q\,dy. \tag{2}$$

Для этого достаточно доказать, что в каждой точке B области G существуют частные производные $\frac{\partial F}{\partial x}$ и $\frac{\partial F}{\partial y}$, причем

$$\frac{\partial F}{\partial x} = P(x, y), \quad \frac{\partial F}{\partial y} = Q(x, y). \tag{3}$$

Так как $P(x, y)$ и $Q(x, y)$ непрерывны в G, то из (3) следуют дифференцируемость функции $F(x, y)$ и равенство (2).

Для доказательства существования частных производных функции $F(x, y)$ и равенств (3) составим частное приращение по x функции $F(x, y)$ в некоторой точке $B(x, y)$:

$$\Delta_x F = F(x + \Delta x, y) - F(x, y) =$$
$$= \int\limits_{AC} P\,dx + Q\,dy - \int\limits_{AB} P\,dx + Q\,dy = \int\limits_{BC} P\,dx + Q\,dy,$$

взяв точку $C(x + \Delta x;\ y)$ так, чтобы кривая AC не выходила из области G (рис. 68). Так как по условию интеграл не зависит от вида кривой, то возьмем путь от A до B произвольным, а от $B(x, y)$ до $C(x + \Delta x;\ y)$ прямолинейным и параллельным оси Ox. Тогда

$$\Delta_x F = \int\limits_{BC} P\,dx + Q\,dy = \int\limits_{BC} P\,dx = \int\limits_{x}^{x+\Delta x} P\,dt.$$

Применяя к последнему интегралу теорему о среднем, получим

$$\Delta_x F = P(x + \theta\Delta x, y)\Delta x, \quad 0 < \theta < 1,$$

откуда

$$\frac{\Delta_x F}{\Delta x} = P(x + \theta\Delta x, y), \quad 0 < \theta < 1.$$

Следовательно,

$$\frac{\partial F}{\partial x} = \lim_{\Delta x \to 0} P(x + \theta\Delta x, y) = P(x, y),$$

поскольку по условию $P(x, y)$ непрерывна. Аналогично доказывается, что $\frac{\partial F}{\partial y} = Q(x,y)$. Таким образом, условие 3 установлено.

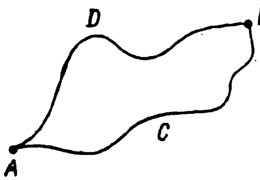


Рис. 67

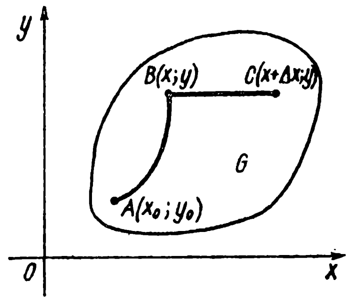


Рис. 68

Третий этап: 3→4. Пусть в области G определена функция $F(x, y)$ такая, что $dF = Pdx + Qdy$. Тогда

$$\frac{\partial F}{\partial x} = P, \quad \frac{\partial F}{\partial y} = Q$$

и по теореме о равенстве смешанных производных

$$\frac{\partial Q}{\partial x} = \frac{\partial^2 F}{\partial x\,\partial y} = \frac{\partial^2 F}{\partial y\,\partial x} = \frac{\partial P}{\partial y},$$

т. е. получили требуемое равенство.

Четвертый этап: 4→1. Пусть выполнено условие 4 и пусть L — кусочно-гладкая кривая, лежащая в области G и ограничивающая область G^*. Тогда, применяя формулу Грина к области G^* (здесь используется односвязность области G), получим

$$\oint_L P\,dx + Q\,dy = \iint_{G^*} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy.$$

В силу (1) интеграл справа равен нулю. Следовательно,

$$\oint_L P\,dx + Q\,dy = 0$$

для всякого замкнутого контура L, лежащего в области G ■

Замечание. Из эквивалентности условий 1, 2, 3, 4 теоремы 13.7, в частности, следует, что условие 3 представляет собой необходимое и достаточное условие, при котором криволинейный интеграл не зависит от выбора пути. Однако для приложений более удобным необходимым и достаточным условием является условие 4, т. е. равенство (1).

Теорема 13.7 позволяет легко решать вопрос о зависимости или независимости криволинейного интеграла от формы пути интегрирования. Так, например,

$$\int_{AB} e^y\,dx - y\,dy$$

в любой области зависит от формы пути, так как

$$\frac{\partial P}{\partial y} \neq \frac{\partial Q}{\partial x}.$$

Необходимо обратить внимание на то, что все условия теоремы существенны. Рассмотрим, например, интеграл

$$J = \oint_L \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy,$$

где L — окружность радиуса R с центром в начале координат. Имеем

$$P = -\frac{y}{x^2+y^2}, \quad \frac{\partial P}{\partial y} = \frac{-x^2-y^2+2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2},$$

$$Q = \frac{x}{x^2+y^2}, \quad \frac{\partial Q}{\partial x} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}.$$

Видим, что условие независимости интеграла от выбора пути выполнено, но интеграл нулю не равен. Действительно, задав окружность уравнениями $x=R\cos t$, $y=R\sin t$, получим

$$J=\int_0^{2\pi}\frac{-R\sin t\,(-R\sin t)+R\cos t\cdot R\cos t}{R^2\cos^2 t+R^2\sin^2 t}\,dt=\int_0^{2\pi}dt=2\pi.$$

Противоречия с теоремой здесь нет. Просто не выполнено одно из условий: ни функции P и Q, ни их частные производные $\frac{\partial P}{\partial y}$ и $\frac{\partial Q}{\partial x}$ не определены в точке (0; 0), в этом случае за область G можно взять всю плоскость Oxy с исключенной точкой (0; 0), область G, очевидно, не односвязна (начало координат играет роль «дырки»).

§ 8. Интегрирование полных дифференциалов

Из рассмотрения условий независимости криволинейного интеграла $\int\limits_{AB} P\,dx+Q\,dy$ от пути непосредственно вытекает рассмотрение вопроса об интегрировании полных дифференциалов и о нахождении функции по ее полному дифференциалу.

Мы доказали, что если функции $P(x, y)$ и $Q(x, y)$ и их частные производные $\frac{\partial P}{\partial y}$ и $\frac{\partial Q}{\partial x}$ непрерывны в некоторой замкнутой области G, то выражение

$$P\,dx+Q\,dy \tag{1}$$

служит полным дифференциалом некоторой функции в том и только в том случае, когда

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}.$$

Далее, мы показали, что если это равенство выполнено, то условию

$$dF=P\,dx+Q\,dy$$

удовлетворяет функция

$$F(x;\,y)=\int\limits_{(x_0;\,y_0)}^{(x;\,y)} P\,dx+Q\,dy=\int\limits_{AB} P\,dx+Q\,dy. \tag{2}$$

Пусть теперь выражение (1) является полным дифференциалом некоторой функции $\Phi(x, y)$. Тогда

$\frac{\partial \Phi}{\partial x}=P$, $\frac{\partial \Phi}{\partial y}=Q$, и разность $\Phi(x, y)-F(x, y)$ (см. замечание к теореме 12.6) будет постоянной величиной. Следовательно,

$$\Phi(x, y)=F(x, y)+c, \tag{3}$$

где c — произвольная постоянная. Полагая $x=x_0$, $y=y_0$, из (2) получаем $F(x_0, y_0)=0$, а из (3) — значение постоянной c: $c=\Phi(x_0, y_0)$. Теперь (3) можно представить в виде

$$F(x, y)=\Phi(x, y)-\Phi(x_0, y_0),$$

а равенство (2) в виде

$$\int\limits_{(x_0;y_0)}^{(x;y)} P\,dx+Q\,dy=\Phi(x, y)-\Phi(x_0, y_0).$$

Если, наконец, положить здесь $x=x_1$, $y=y_1$, то получим формулу

$$\int\limits_{(x_0;y_0)}^{(x_1;y_1)} P\,dx+Q\,dy=\Phi(x_1, y_1)-\Phi(x_0, y_0)=\Phi(x, y)\Big|_{(x_0;\,y_0)}^{(x_1;\,y_1)}. \tag{4}$$

Формула (4) аналогична формуле Ньютона—Лейбница, но она справедлива лишь при условии независимости криволинейного интеграла от пути интегрирования.

Пользуясь полученными результатами, мы можем теперь указать способ восстановления функции $F(x, y)$, полный дифференциал которой есть заданное выражение (1).

Формула

$$F(x, y)=\int\limits_{(x_0;y_0)}^{(x;y)} P\,dx+Q\,dy+c, \tag{5}$$

где $(x_0;\ y_0)$ — фиксированная точка, а c — произвольная постоянная, и дает возможность определить все функции, имеющие подынтегральное выражение своим полным дифференциалом.

Для отыскания $F(x, y)$ по формуле (5) достаточно, выбрав любую точку $(x_0;\ y_0)$ в области G, вычислить криволинейный интеграл по любой кривой, соединяющей точки $(x_0;\ y_0)$ и $(x;\ y)$. Так как в (5) интеграл не зависит от выбора пути, то удобно, например, за путь интегрирования взять ломаную, звенья которой параллельны осям координат (рис. 69). Тогда

$$\int\limits_{(x_0;y_0)}^{(x;y)} P\,dx + Q\,dy = \int\limits_{(x_0;\,y_0)}^{(x;y_0)} P\,dx + Q\,dy + \int\limits_{(x;y_0)}^{(x;y)} P\,dx + Q\,dy.$$

Так как на участке от $(x_0;\ y_0)$ до $(x;\ y_0)$ $y=y_0$ и $dy=0$, а на участке от $(x;\ y_0)$ до $(x;\ y)$ $dx=0$, то равенство (5) принимает вид

$$F(x,\ y) = \int\limits_{x_0}^{x} P(x,\ y_0)\,dx + \int\limits_{y_0}^{y} Q(x,\ y)\,dy + c,$$

где первый определенный интеграл справа вычисляется при постоянном y, а второй — при постоянном x.

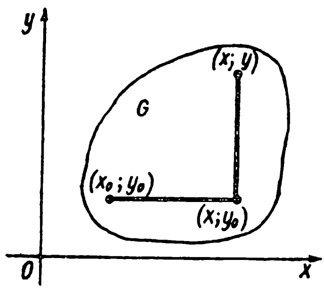


Рис. 69

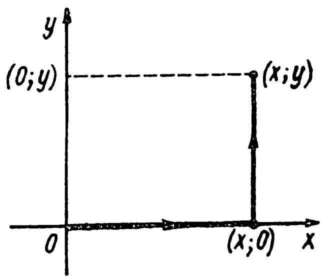


Рис. 70

Пример 1. Проверить, является ли выражение $(x^2+2xy-y^2)\,dx+(x^2-2xy-y^2)\,dy$ полным дифференциалом некоторой функции $F(x,\ y)$, и, если это так, найти $F(x,\ y)$.

Решение. В данном выражении

$$P(x,\ y)=x^2+2xy-y^2,\quad Q(x,\ y)=x^2-2xy-y^2 \qquad (6)$$

непрерывны вместе с частными производными

$$\frac{\partial P}{\partial y} = 2x-2y,\quad \frac{\partial Q}{\partial x} = 2x-2y,$$

которые равны между собой. Следовательно, выражение (6) является полным дифференциалом $dF(x, y)$. Для отыскания функции $F(x, y)$ воспользуемся формулой (2), где $A(x_0; y_0)$ — некоторая фиксированная точка, а $B(x; y)$ — переменная точка, взятые из области непрерывности функций P, Q,

$$\frac{\partial P}{\partial y} \text{ и } \frac{\partial Q}{\partial x}.$$

В данном случае за точку $A(x_0; y_0)$ удобно взять точку (0; 0). Учитывая, что криволинейный интеграл

$$\int\limits_{(0;0)}^{(x;y)} (x^2 + 2xy - y^2)\,dx + (x^2 - 2xy - y^2)\,dy$$

не зависит от пути интегрирования, выберем путь интегрирования от точки (0; 0) до точки $(x; y)$ в виде ломаной, звенья которой параллельны осям координат. Для этого достаточно взять точку $(x;\ 0)$ [или точку $(0;\ y)$] (рис. 70). Тогда одно звено ломаной будет совпадать с координатной осью. Получим

$$F(x, y) = \int\limits_{(0;0)}^{(x;y)} (x^2 + 2xy - y^2)\,dx + (x^2 - 2xy - y^2)\,dy =$$

$$= \int\limits_0^x x^2\,dx + \int\limits_0^y (x^2 - 2xy - y^2)\,dy = \frac{x^3}{3} + x^2y - xy^2 - \frac{y^3}{3}.$$

По формуле (5) окончательно получаем

$$F(x, y) = \frac{x^3}{3} + x^2y - xy^2 - \frac{y^3}{3} + c.$$

Практически при отыскании функции по ее полному дифференциалу удобно поступать следующим образом. Если

$$\frac{\partial F}{\partial x} = P, \quad \frac{\partial F}{\partial y} = Q,$$

то, интегрируя первое из этих равенств по x, получим

$$F(x, y) = \int P\,dx + f_1(y), \tag{7}$$

а интегрируя второе равенство по y, получим

$$F(x, y) = \int Q\,dy + f_2(x). \tag{8}$$

Если подобрать функции $f_1(y)$ и $f_2(x)$ так, чтобы правые части равенств (7) и (8) совпали, то полученная таким образом функция $F(x, y)$ и будет той функцией, полный дифференциал которой совпадает с выражением $Pdx + Qdy$.

Так, например, пусть $dF = (2xy + 1)\,dx + (x^2 + 3y^2)\,dy$. Интегрируя первое слагаемое, получим

$$\int (2xy + 1)\,dx = x^2y + x + f_1(y), \tag{9}$$

а интегрирование второго слагаемого дает

$$\int (x^2 + 3y^2)\,dy = yx^2 + y^3 + f_2(x). \tag{10}$$

Правые части равенств (9) и (10) совпадают, если мы положим

$$f_1(y) = y^3 + c,\; f_2(x) = x + c.$$

Таким образом, получаем, что

$$F(x, y) = yx^2 + y^3 + x + c.$$

Пример 2. Вычислить криволинейный интеграл

$$\int\limits_{(-1;2)}^{(2;3)} ydx + xdy.$$

Решение. В данном случае функции

$$P = y,\; Q = x,\; \frac{\partial P}{\partial y} = 1,\; \frac{\partial Q}{\partial x} = 1$$

непрерывны и частные производные равны между собой. Значит, выражение $ydx + xdy$ есть полный дифференциал $dF(x, y)$ и данный интеграл не зависит от пути интегрирования. По формулам (7) и (8) легко находим $F(x, y) = xy$, и по формуле (4) получаем

$$\int\limits_{(-1;2)}^{(2;3)} ydx + xdy = xy \Big|_{(-1;2)}^{(2;3)} = 2\cdot 3 - (-1)\cdot 2 = 6 + 2 = 8.$$

Заметим, что данный интеграл вычисляется непосредственно, если, например, взять в качестве пути интегрирования ломаную, соединяющую точки (—1, 2), (2, 2) и (2, 3), звенья которой будут параллельны координатным осям. Данное вычисление предлагается проделать самостоятельно.

§ 9. Некоторые приложения криволинейных интегралов второго рода

Криволинейные интегралы второго рода, так же как и первого, имеют широкое применение в геометрии, физике и технике. Мы ограничимся рассмотрением двух задач: вычисление площадей плоских фигур и определение работы силы.

1. Вычисление площади с помощью формулы Грина. Пусть G — некоторая область с границей L и s — площадь этой области. Известно, что двойной интеграл $\iint\limits_G f(x, y)\,dxdy$ при $f(x, y) \equiv 1$ выражает площадь области G. Поэтому если в формуле Грина подобрать функции $P(x, y)$ и $Q(x, y)$ таким образом, чтобы $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \equiv 1$, то площадь s области G представится в виде

$$s = \iint\limits_G dxdy = \oint\limits_L Pdx + Qdy.$$

Положим $Q(x, y) = x$ и $P(x, y) = 0$, тогда $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$, и получим

$$s = \oint\limits_L xdy.$$

Полагая $P(x, y) = -y$ и $Q(x, y) = 0$, аналогичным образом находим, что

$$s = -\oint\limits_L ydx,$$

а при $P(x, y) = -\frac{1}{2}y$, $Q(x, y) = \frac{1}{2}x$ получаем формулу

$$s = \frac{1}{2}\oint_L xdy - ydx. \qquad (1)$$

Таким образом, получили три формулы для вычисления площадей плоских фигур, ограниченных контуром L.

Пример 1. Вычислить площадь фигуры, ограниченной эллипсом.

Решение. Вычислим, например, площадь по формуле (1). Используя параметрические уравнения эллипса $x=a\cos t$, $y=b\sin t$, $0\leqslant t\leqslant 2\pi$, имеем $dx=-a\sin t dt$, $dy=b\cos t dt$, и по формуле (1) получаем

$$s = \frac{1}{2}\oint_L xdy - ydx = \frac{1}{2}\int_0^{2\pi} (a\cos t\, b\cos t +$$

$$+ b\sin t\, a\sin t)\, dt = \frac{ab}{2}\int_0^{2\pi} dt = \pi ab.$$

2. Работа силы. Известно, что работа, совершаемая переменной силой $\bar{F}$ по перемещению материальной точки вдоль оси Ox из точки $x=a$ в точку $x=b$ $(a<b)$, определяется с помощью определенного интеграла по формуле $A = \int_a^b \bar{F}(x)\, dx$ (см. ч. I, гл. VI, § 8, п. 6). Здесь мы рассмотрим более общую задачу.

Пусть материальная точка под действием силы $\bar{F}$ перемещается вдоль непрерывной плоской кривой BC в направлении от B к C. Сила $\bar{F}$ предполагается переменной, зависящей от положения точки на кривой BC. Вычислим работу силы $\bar{F}$, затраченную на перемещение точки из B в C. С этой целью разобьем произвольно кривую BC на n частей точками (рис. 71):

$$B=M_0,\ M_1,\ M_2,\ \dots,\ M_i,\ M_{i+1},\ \dots,\ M_n=C.$$

Заменим приближенно на участке M_iM_{i+1} силу $\bar{F}$ постоянным значением, равным ее значению в точке M_i, т. е. $\bar{F}_i=\bar{F}(M_i)$, и предположим также, что точка перемещается не по дуге $M_i M_{i+1}$, а по вектору $\overline{M_i M_{i+1}}$. Тогда работу постоянной силы $\bar{F}(M_i)$ вдоль вектора $\overline{M_i M_{i+1}}$ можно принять за приближенное значение работы A_i переменной силы вдоль дуги $M_i M_{i+1}$, т. е.

$$A_i \approx \bar{F}(M_i) \cdot \overline{M_i M_{i+1}}.$$

Правая часть этого приближенного равенства представляет собой скалярное произведение двух векторов $\bar{F}(M_i)$ и $\overline{M_i M_{i+1}}$. Оно равно сумме произведений соответствующих координат этих векторов, т. е. если $\bar{F}(M_i) = \{P_i;\ Q_i\}$, $\overline{M_i M_{i+1}} = \{\Delta x_i;\ \Delta y_i\}$, то

$$A_i \approx P_i \Delta x_i + Q_i \Delta y_i.$$

Суммируя по всем значениям i от 1 до n, получим приближенное значение работы A вдоль всей кривой BC:

$$A \approx \sum_{i=1}^{n} (P_i \Delta x_i + Q_i \Delta y_i) = \sum_{i=1}^{n} P_i \Delta x_i + \sum_{i=1}^{n} Q_i \Delta y_i. \quad (2)$$

За точное значение работы A естественно принимается предел, к которому стремится найденное приближенное значение при стремлении к нулю наибольшей из длин дуг $M_i M_{i+1}$. Но, с другой стороны, сумма (2) представляет собой сумму двух интегральных сумм для функций $P(x, y)$ и $Q(x, y)$, заданных вдоль кривой BC. По определению пределом этой суммы является криволинейный интеграл второго рода. Следовательно, работа силы определяется по формуле

$$A = \int_{BC} P dx + Q dy, \quad (3)$$

где P и Q — составляющие (или проекции) силы $\bar{F}$ по координатным осям.

Если рассмотреть данную задачу не на плоскости, а в пространстве, то решение ее сводится к вычислению криволинейного интеграла второго рода по пространственной кривой BC по формуле

$$A = \int_{BC} P dx + Q dy + R dz.$$

Пример 2. Вычислить работу силы $\bar{F}$ при перемещении материальной точки по эллипсу, если сила в каждой точке $(x;\ y)$ эллипса направлена к центру эллипса

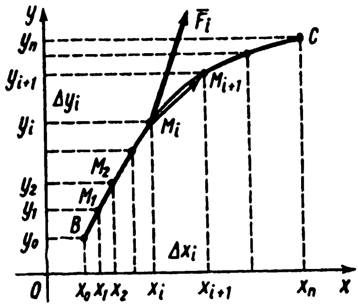


Рис. 71

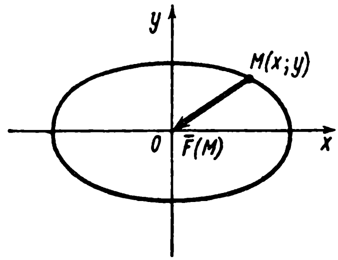


Рис. 72

и по величине равна расстоянию от точки $(x;\ y)$ до центра эллипса (рис. 72).

Решение. По условию $|\bar{F}|=\sqrt{x^2+y^2}$. Отсюда в каждой точке $(x;\ y)$ составляющие силы $\bar{F}$ по координатным осям будут: $P=-x$, $Q=-y$ (знак —, так как сила направлена к точке (0; 0)). По формуле (3): $A=\oint\limits_L Pdx+Qdy$, где L — эллипс: $x=a\cos t$, $y=b\sin t$, $0<t<2\pi$, имеем

$$A=-\oint\limits_L xdx+ydy=\frac{a^2-b^2}{2}\int\limits_0^{2\pi}\sin 2tdt=$$

$$=\frac{a^2-b^2}{4}(-\cos 2t)\Big|_0^{2\pi}=0.$$

Заметим, что из того, что интеграл оказался равным нулю, в частности, следует, что подынтегральное выражение является полным дифференциалом некоторой силовой функции $\bar{F}(M)$, которую нетрудно восстановить.

§ 10. Тройные интегралы

В начале главы мы ввели понятие двойного интеграла от функции двух переменных. В этом параграфе мы определим интеграл от функции трех переменных, так называемый *тройной интеграл*. Тройные интегралы, подобно двойным, также имеют широкое применение в различных физических и геометрических задачах.

1. Определение тройного интеграла. Тройной интеграл является полным аналогом двойного интеграла и строится для функции от трех переменных.

Пусть в некоторой замкнутой ограниченной области V трехмерного пространства задана произвольная функция $f(M)=f(x, y, z)$. Разобьем область V на n произвольных областей, не имеющих общих внутренних точек, с объемами $\Delta v_1, \Delta v_2, \ldots, \Delta v_n$. В каждой области возьмем произвольную точку $M_i(\xi_i; \eta_i; \zeta_i)$ и составим сумму

$$\sum_{i=1}^{n} f(\xi_i; \eta_i; \zeta_i)\, \Delta v_i, \tag{1}$$

которая называется *интегральной суммой* для функции $f(x, y, z)$ по области V. Обозначим через λ наибольший из диаметров частичных областей и дадим следующее

Определение. *Если интегральная сумма (1) при $\lambda \to 0$ имеет конечный предел J, то этот предел называется тройным интегралом функции $f(x, y, z)$ по области V и обозначается одним из следующих символов:*

$$J = \iiint\limits_V f(x, y, z)\, dv = \iiint\limits_V f(x, y, z)\, dx dy dz.$$

В этом случае функция $f(x, y, z)$ называется *интегрируемой в области* V, V — *областью* интегрирования, x, y и z — *переменными* интегрирования, dv (или $dxdydz$) — *элементом* объема.

В дальнейшем, поскольку результаты, полученные для двойных интегралов, вместе с их доказательствами могут быть перенесены на тройные интегралы, мы ограничимся лишь формулировками и краткими пояснениями.

Тройные интегралы являются непосредственным обобщением двойных на случай трехмерного пространства. Они обладают теми же необходимыми и достаточными условиями существования и свойствами. Если положить всюду в области V $f(x, y, z) \equiv 1$, то из определения тройного интеграла следует формула для вычисления объема тела V:

$$v = \iiint\limits_V dv = \iiint\limits_V dxdydz.$$

2. Вычисление тройных интегралов. Как и в случае двойных интегралов, вычисление тройных интегралов сводится к вычислению интегралов меньшей кратности.

Рассмотрим область V, ограниченную снизу и сверху поверхностями $z=z_1(x, y)$ и $z=z_2(x, y)$, а сбоку — некоторой цилиндрической поверхностью, и пусть область G — проекция области V на плоскость Oxy (рис. 73), в которой определены и непрерывны функции $z_1(x, y)$ и $z_2(x, y)$. Предположим, далее, что каждая прямая, параллельная оси Oz, пересекает границу области V не более чем в двух точках. Тогда для любой функции $f(x, y, z)$, непрерывной в области V, имеет место формула

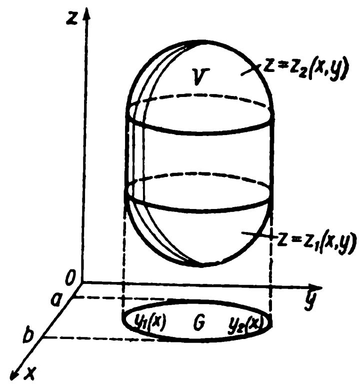


Рис. 73

$$\iiint\limits_V f(x, y, z)\,dxdy\,dz = \iint\limits_G dxdy \int\limits_{z_1(x,y)}^{z_2(x,y)} f(x, y, z)\,dz,$$

сводящая вычисление тройного интеграла к последовательному вычислению внутреннего определенного интеграла по переменной z (при постоянных x и y) и внешнего двойного интеграла по области G.

Выражение

$$J(x, y) = \int\limits_{z_1(x,y)}^{z_2(x,y)} f(x, y, z)\,dz$$

представляет собой функцию двух переменных. Если для этой функции и области G, по которой она интегрируется, выполнены условия теоремы 13.4, то, переходя от двойного интеграла $\iint\limits_G J(x, y)\,dxdy$ к повторному, получаем окончательную формулу:

$$\iiint\limits_V f(x, y, z)\,dxdydz = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x,y)}^{z_2(x,y)} f(x, y, z)\,dz, \qquad (1)$$

сводящую вычисление тройного интеграла к последовательному вычислению трех определенных интегралов. Разумеется, порядок интегрирования может быть другим, если область G будет проекцией области V на плоскость Oxz или Oyz, и тогда переменные x, y и z в формуле (1) поменяются соответственно ролями.

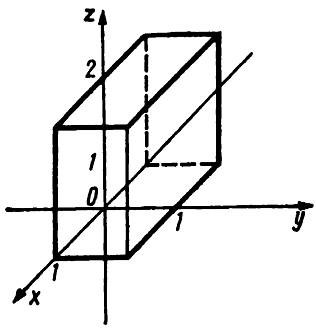


Рис. 74

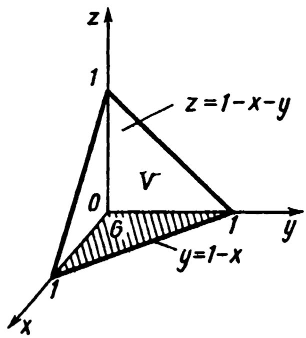


Рис. 75

В частности, если V — параллелепипед с гранями $x=a$, $x=b$ $(a<b)$, $y=c$, $y=d$ $(c<d)$, $z=k$, $z=l$ $(k<l)$, то формула (1) примет вид

$$\iiint\limits_V f(x, y, z)\,dxdydz = \int\limits_a^b dx \int\limits_c^d dy \int\limits_k^l f(x, y, z)\,dz. \qquad (2)$$

В этом случае интегрирование можно производить в любом порядке.

Пример 1. Вычислить интеграл

$$\iiint\limits_V (x + y - z)\,dxdydz,$$

где V — параллелепипед, ограниченный плоскостями $x=-1$, $x=+1$, $y=0$, $y=1$, $z=0$, $z=2$ (рис. 74)

Решение. По формуле (2) имеем

$$\iiint\limits_V (x+y-z)\,dxdydz = \int\limits_{-1}^{1} dx \int\limits_0^1 dy \int\limits_0^2 (x+y-z)\,dz =$$

$$= \int\limits_{-1}^{1} dx \int\limits_0^1 \left[xz + yz - \frac{z^2}{2}\right]_0^2 dy = \int\limits_{-1}^{1} dx \int\limits_0^1 (2x+2y-2)\,dy =$$

$$= \int\limits_{-1}^{1} [2xy + y^2 - 2y]_0^1\,dx = \int\limits_{-1}^{1} (2x-1)\,dx = [x^2 - x]_{-1}^{1} = -2.$$

Пример 2. Вычислить интеграл

$$\iiint\limits_V (x+y+z)\,dxdydz,$$

где V — треугольная пирамида, ограниченная плоскостью $x+y+z=1$ и координатными плоскостями $x=0$, $y=0$, $z=0$ (рис. 75).

Решение. Область V проектируется на плоскость Oxy в треугольник G, ограниченный прямыми $x=0$, $y=0$, $y=1-x$. По формуле (1) имеем

$$\iiint\limits_V (x+y+z)\,dxdydz = \int\limits_0^1 dx \int\limits_0^{1-x} dy \int\limits_0^{1-x-y} (x+y+z)\,dz =$$

$$= \int\limits_0^1 dx \int\limits_0^{1-x} \left[xz + yz + \frac{z^2}{2}\right]_0^{1-x-y} dy =$$

$$= \frac{1}{2}\int\limits_0^1 \left[y - yx^2 - xy^2 - \frac{y^3}{3}\right]_0^{1-x} dx =$$

$$= \frac{1}{6}\int\limits_0^1 (2 - 3x + x^3)\,dx =$$

$$= \frac{1}{6}\left[2x - \frac{3x^2}{2} + \frac{x^4}{4}\right]_0^1 = \frac{1}{6}\cdot\frac{3}{4} = \frac{1}{8}.$$

3. Замена переменных в тройном интеграле. Как для двойных интегралов, так и для тройных имеют место формулы перехода от прямоугольных координат к новым системам координат, наиболее употребительными из которых являются *цилиндрические* и *сферические*.

Замена переменных в тройном интеграле производится по следующему правилу.

Если ограниченная замкнутая область V пространства (x, y, z) взаимно-однозначно отображается на область V^* пространства (u, v, w) с помощью непрерывно дифференцируемых функций

$$x=x(u, v, w),\ y=y(u, v, w),\ z=z(u, v, w)$$

и якобиан J в области V^* не обращается в нуль:

$$J=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} \neq 0,$$

то справедлива формула

$$\iiint\limits_V f(x, y, z)\,dxdydz = \iiint\limits_{V^*} f[x(u, v, w),$$

$$y(u, v, w), z(u, v, w)]\,|J|\,dudvdw.$$

В частности, при переходе от прямоугольных координат x, y, z к цилиндрическим координатам ρ, φ, z (рис. 76), связанным с x, y, z формулами

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

$$(0\leqslant\rho<+\infty,\ 0\leqslant\varphi\leqslant 2\pi,\ -\infty<z<+\infty),$$

якобиан преобразования $J=\rho$, поэтому

$$\iiint\limits_V f(x, y, z)\,dxdydz = \iiint\limits_{V^*} f(\rho\cos\varphi, \rho\sin\varphi, z)\,\rho d\rho d\varphi dz. \quad (3)$$

Название «цилиндрические координаты» связано с тем, что координатная поверхность $\rho=\text{const}$ (т. е. поверхность, все точки которой имеют одну и ту же координату ρ) является цилиндром, прямолинейные образующие которого параллельны оси Oz.

При переходе от прямоугольных координат x, y, z к сферическим координатам ρ, φ, θ (рис. 77), связанным с x, y, z формулами

$$x=\rho\sin\theta\cos\varphi,\ y=\rho\sin\theta\sin\varphi,\ z=\rho\cos\theta$$

$$(0\leqslant\rho<+\infty,\ 0\leqslant\varphi\leqslant2\pi,\ 0\leqslant\theta\leqslant\pi),$$

якобиан преобразования $J=\rho^2\sin\theta$, поэтому

$$\iiint\limits_V f(x, y, z)\,dxdydz =$$

$$=\iiint\limits_{V^*} f[\rho\sin\theta\cos\varphi,\ \rho\sin\theta\sin\varphi,\ \rho\cos\theta]\,\rho^2\sin\theta\,d\rho\,d\varphi\,d\theta. \quad (4)$$

Наименование «сферические координаты» связано с тем, что координатная поверхность $\rho=\mathrm{const}$ (т. е. по-

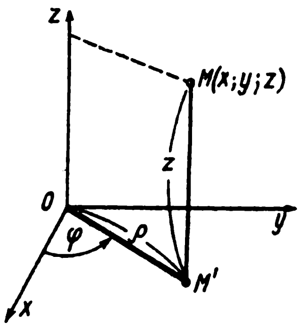


Рис. 76

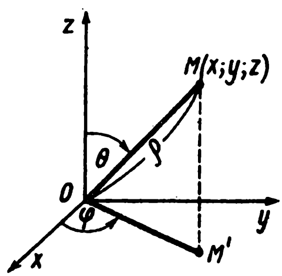


Рис. 77

верхность, все точки которой имеют одну и ту же координату ρ) является сферой. Сферические координаты иначе называются полярными координатами в пространстве.

При вычислении тройного интеграла в цилиндрических и сферических координатах область V^* обычно не изображают, а расстановку пределов интегрирования определяют непосредственно по области V, используя геометрический смысл координат.

Пример 3. Вычислить интеграл $\iiint\limits_V (x^2+y^2)\,dxdydz$ в цилиндрических координатах: $x=\rho\cos\varphi$, $y=\rho\sin\varphi$, $z=z$, где V есть область, ограниченная поверхностями $x^2+y^2=z$, $z=1$ (рис. 78).

Решение. Так как область V на плоскость Oxy проектируется в круг $x^2+y^2\leqslant 1$, то координата φ изменяется в пределах от 0 до 2π, координата ρ — в пределах от $\rho=0$ до $\rho=1$. Постоянному значению ρ, $0\leqslant\rho\leqslant 1$, в пространстве $Oxyz$ соответствует цилиндр $x^2+y^2=\rho^2$. В пересечении этого цилиндра с областью V получаем изменение координаты z от точек, лежащих на параболоиде, до точек, лежащих на плоскости $z=1$, т. е. от $z=\rho^2$ до $z=1$. Применяя формулу (3), получаем

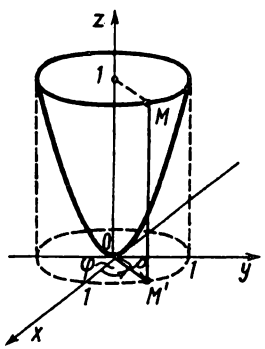


Рис. 78

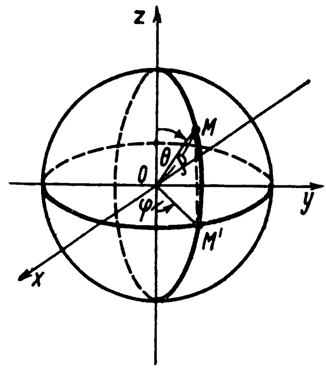


Рис. 79

$$\iiint\limits_V (x^2+y^2)dxdydz = \int\limits_0^{2\pi} d\varphi \int\limits_0^1 d\rho \int\limits_{\rho^2}^1 \rho^2\cdot\rho dz =$$

$$= \int\limits_0^{2\pi} d\varphi \int\limits_0^1 [\rho^3 z]_{\rho^2}^1 d\rho = \int\limits_0^{2\pi} \left[\frac{\rho^4}{4}-\frac{\rho^6}{6}\right]_0^1 d\varphi =$$

$$= \frac{1}{12}\int\limits_0^{2\pi} d\varphi = \frac{1}{12}\,\varphi\Big|_0^{2\pi} = \frac{\pi}{6}.$$

Трудно дать общую рекомендацию, когда следует применять ту или иную систему координат. Это зависит и от области интегрирования, и от вида подынтегральной функции. Однако, например, формулой (4) удобнее пользоваться, когда $f(x, y, z)$ имеет вид $f(x^2+y^2+z^2)$, а также когда областью V служит шар $x^2+y^2+z^2\leqslant R^2$ или часть шара.

Пример 4. Вычислить интеграл

$$\iiint\limits_V (x^2+y^2+z^2)\,dxdydz,$$

где V — шар $x^2+y^2+z^2\leqslant R^2$ (рис. 79).

Решение. В данном случае удобнее перейти к сферическим координатам $x=\rho\sin\theta\cos\varphi$, $y=\rho\sin\theta\sin\varphi$, $z=\rho\cos\theta$. Из чертежа области V следует, что координаты ρ, φ и θ меняются в следующих пределах: ρ от 0 до R, φ от 0 до 2π, θ от 0 до π. Так как подынтегральная функция $x^2+y^2+z^2=\rho^2\sin^2\theta\cos^2\varphi+\rho^2\sin^2\theta\sin^2\varphi+$ $+\rho^2\cos^2\theta=\rho^2\sin^2\theta+\rho^2\cos^2\theta=\rho^2$, то по формуле (4) получаем

$$\iiint\limits_V (x^2+y^2+z^2)\,dxdydz=\int\limits_0^R d\rho\int\limits_0^\pi d\theta\int\limits_0^{2\pi}\rho^2\cdot\rho^2\sin\theta\,d\varphi=$$

$$=\int\limits_0^R\rho^4d\rho\int\limits_0^\pi\sin\theta\,d\theta\int\limits_0^{2\pi}d\varphi=2\pi\int\limits_0^R\rho^4d\rho\int\limits_0^\pi\sin\theta\,d\theta=$$

$$=4\pi\int\limits_0^R\rho^4d\rho=4\pi\frac{\rho^5}{5}\Big|_0^R=\frac{4\pi R^5}{5}.$$

4. Некоторые приложения тройных интегралов. В этом пункте мы коротко рассмотрим типичные задачи, связанные с вычислением тройных интегралов, ограничившись лишь написанием формул, с помощью которых решается та или иная задача, так как их вывод аналогичен выводу соответствующих формул в случае двойных интегралов.

Если дано некоторое тело V с объемной плоскостью $\rho(M)=\rho(x, y, z)$, представляющей собой непрерывную функцию, то тройной интеграл

$$\iiint\limits_V \rho(x, y, z)\, dxdydz,$$

взятый по объему, занимаемому этим телом, представляет собой массу m данного тела. В частности, в этом заключается физический смысл тройного интеграла от неотрицательной функции. Установить геометрический смысл тройного интеграла для произвольной функции $f(x, y, z)$ не представляется возможным.

Моменты инерции пространственного тела с объемной плотностью $\rho(x, y, z)$ относительно координатных осей определяются по следующим формулам:

$$J_z = \iiint\limits_V (x^2 + y^2)\rho(M)\, dv; \quad J_y = \iiint\limits_V (x^2 + z^2)\rho(M)\, dv;$$

$$J_x = \iiint\limits_V (y^2 + z^2)\rho(M)\, dv.$$

Для момента инерции относительно начала координат формула имеет вид

$$J_0 = \iiint\limits_V (x^2 + y^2 + z^2)\rho(M)\, dv.$$

Координаты центра масс некоторого пространственного тела, имеющего объемную плотность $\rho(x, y, z)$, выражаются формулами

$$x_c = \frac{\iiint\limits_V x\rho(M)\, dv}{m}; \quad y_c = \frac{\iiint\limits_V y\rho(M)\, dv}{m}; \quad z_c = \frac{\iiint\limits_V z\rho(M)\, dv}{m},$$

где x_c, y_c, z_c — координаты центра масс, а m — масса данного тела. В частности, если рассматриваемое тело однородно, т. е. $\rho(x, y, z) \equiv \text{const}$, то выражения для координат центра масс соответственно упрощаются и принимают более простой вид:

$$x_c = \frac{\iiint\limits_V x\, dv}{v}; \quad y_c = \frac{\iiint\limits_V y\, dv}{v}; \quad z_c = \frac{\iiint\limits_V z\, dv}{v},$$

где v — объем данного тела.

Как мы уже отмечали, тройной интеграл

$$\iiint\limits_V dxdydz$$

равен объему тела V. Тройные интегралы в некоторых случаях бывают удобнее для вычисления объемов, чем двойные, так как с их помощью можно записать сразу объем не только криволинейного цилиндра, но и других тел.

Пример 5. Определить координаты центра масс верхней половины однородного шара V радиуса R с центром в начале координат.

Решение. Полушар ограничен поверхностями: $z=\sqrt{R^2-x^2-y^2}$, $z=0$. В силу симметрии полушара $x_c=y_c=0$. Координата z_c определяется по формуле

$$z_c=\frac{\iiint\limits_V zdxdydz}{\iiint\limits_V dxdydz}.$$

Переходя к сферическим координатам, получаем

$$z_c=\frac{\int\limits_0^{2\pi} d\varphi \int\limits_0^{\pi/2} \sin\theta\cos\theta d\theta \int\limits_0^R \rho^3 d\rho}{\int\limits_0^{2\pi} d\varphi \int\limits_0^{\pi/2} \sin\theta d\theta \int\limits_0^R \rho^2 d\rho}=\frac{2\pi\frac{R^4}{4}\cdot\frac{1}{2}}{\frac{2}{3}\pi R^3}=\frac{3}{8}R.$$

§ 11. Поверхностные интегралы

Этот параграф посвящен изучению интегралов от функций, заданных на поверхности, так называемых *поверхностных интегралов*.

Теория поверхностных интегралов во многом аналогична теории криволинейных интегралов. Различают поверхностные интегралы первого и второго рода.

1. Определение поверхностного интеграла первого рода. Пусть в точках некоторой поверхности S, гладкой или кусочно-гладкой *, определена функция $f(M)=f(x,$

---

* Поверхность называется *гладкой*, если в каждой ее точке существует касательная плоскость и при переходе от точки к точке положение этой касательной плоскости меняется непрерывно. Поверхность, состоящая из конечного числа гладких кусков, которые соединены непрерывно, называется кусочно-гладкой.

y, z). Для простоты ограничимся гладким случаем. Разобьем поверхность S произвольно на n частей, не имеющих общих внутренних точек (рис. 80), с площадями ΔS_1, ΔS_2, ..., ΔS_n. Выбрав на каждой частичной поверхности произвольную точку $M_i(\xi_i;\ \eta_i;\ \zeta_i)$, составим сумму

$$\sum_{i=1}^{n} f(M_i)\,\Delta S_i. \qquad (1)$$

Сумма (1) называется *интегральной суммой* для функции $f(M)=f(x,\ y,\ z)$. Обозначим через λ наибольший из диаметров частей поверхности и дадим следующее

Определение. *Если интегральная сумма (1) при* $\lambda\to 0$ *имеет конечный предел* J, *то этот предел называется поверхностным интегралом первого рода от функции* $f(x,\ y,\ z)$ *по поверхности* S *и обозначается одним из символов*

$$J=\iint\limits_S f(M)\,dS=\iint\limits_S f(x,\ y,\ z)\,dS.$$

В этом случае функция $f(x,\ y,\ z)$ называется *интегрируемой* по поверхности S; S — *областью* интегрирования.

Данное определение, по сути дела, аналогично определению двойного интеграла. Поэтому свойства двойных интегралов и условия их существования без особых изменений переносятся на поверхностные интегралы.

В частности, если на поверхности S $f(x,\ y,\ z)\equiv 1$, то

$$\iint\limits_S dS=\lim_{\lambda\to 0}\sum_{i=1}^{n}\Delta S_i=s,$$

т. е. с помощью поверхностного интеграла первого рода можно вычислять площади поверхностей.

Кроме этого, можно определять массы, статические моменты, моменты инерции, координаты центра масс и тому подобные величины для материальных поверхностей с известной плотностью распределения массы. Эти задачи решаются аналогично соответствующим задачам для случая материальной кривой, материальной плоской и пространственной области.

2. Вычисление поверхностных интегралов первого рода. Вычисление поверхностного интеграла первого рода решается путем сведения поверхностного интеграла к двойному. Покажем это.

Пусть поверхность S задана уравнением $z=z(x, y)$, где функция $z(x, y)$ вместе с производными $z'_x(x, y)$ и

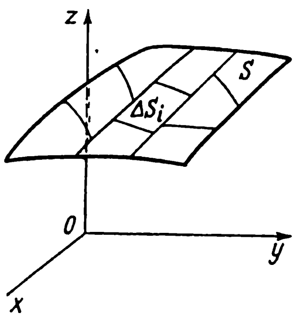


Рис. 80

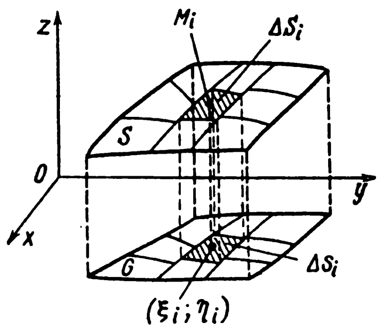


Рис. 81

$z'_y(x, y)$ непрерывна в замкнутой области G — проекции S на плоскость Oxy (рис. 81), и пусть функция $f(x, y, z)$ непрерывна на поверхности S и, следовательно, интегрируема по поверхности S. Разобьем поверхность S произвольно на n частей, не имеющих общих внутренних точек, и спроектируем это разбиение на плоскость Oxy. Мы получим соответственно разбиение области G на части $G_1, G_2, ..., G_n$.

Площадь ΔS_i каждой части поверхности может быть представлена в виде (см. (2), п. 3, § 4):

$$\Delta S_i = \iint\limits_{G_i} \sqrt{1 + z_x'^2(x, y) + z_y'^2(x, y)}\, dx\, dy.$$

Применяя к двойному интегралу теорему о среднем, получаем

$$\Delta S_i = \sqrt{1 + z_x'^2(\xi_i, \eta_i) + z_y'^2(\xi_i, \eta_i)}\, \Delta s_i, \qquad (2)$$

где $(\xi_i; \eta_i)$ — некоторая точка области G_i, Δs_i — площадь G_i. Обозначим через M_i точку на частичной поверхности с координатами $(\xi_i;\ \eta_i;\ \zeta_i)$, где $\zeta_i = z(\xi_i, \eta_i)$, а $(\xi_i;\ \eta_i)$ — та точка, которая фигурирует в формуле (2). Составим интегральную сумму для функции $f(x, y, z)$ по поверхности S, выбирая точки M_i в качестве промежуточных точек:

$$\sum_{i=1}^{n} f(\xi_i, \eta_i\ \xi_i)\,\Delta S_i = \sum_{i=1}^{n} f[\xi_i, \eta_i, z(\xi_i, \eta_i)] \times$$

$$\times \sqrt{1 + z_x'^2(\xi_i, \eta_i) + z_y'^2(\xi_i, \eta_i)}\,\Delta s_i. \tag{3}$$

В правой части равенства стоит интегральная сумма для двойного интеграла от непрерывной в области G функции

$$f[x, y, z(x, y)]\sqrt{1 + z_x'^2(x, y) + z_y'^2(x, y)}.$$

Поэтому предел правой части (3) при $\lambda \to 0$ равен двойному интегралу

$$\iint\limits_G f[x, y, z(x, y)]\sqrt{1 + z_x'^2(x, y) + z_y'^2(x, y)}\,dx\,dy.$$

Так как функция $f(x, y, z)$ интегрируема по поверхности S, то предел левой части (3) при $\lambda \to 0$ равен поверхностному интегралу

$$\iint\limits_S f(x, y, z)\,dS.$$

Следовательно, переходя к пределу в (3) при $\lambda \to 0$, получаем искомую формулу

$$\iint\limits_S f(x, y, z)\,dS = \iint\limits_G f[x, y, z(x, y)] \times$$

$$\times \sqrt{1 + z_x'^2(x, y) + z_y'^2(x, y)}\,dx\,dy, \tag{4}$$

выражающую поверхностный интеграл первого рода через двойной по проекции поверхности S на плоскость Oxy.

Аналогично получаются формулы, выражающие интеграл по поверхности S через двойные по ее проекциям на плоскости Oyz и Oxz.

Пример 1. Вычислить интеграл

$$\iint\limits_S \sqrt{1+4x^2+4y^2}\,dS,$$

где S — часть поверхности параболоида вращения $z=1-x^2-y^2$, отсеченная плоскостью $z=0$ (рис. 82).

Решение. Поверхность S, заданная уравнением $z=1-x^2-y^2$, проектируется на плоскость Oxy в область G, ограниченную окружностью $x^2+y^2=1$ (уравнение окружности получается из уравнения параболоида при $z=0$). Следовательно, областью G является круг $x^2+y^2\leqslant 1$. В этом круге функции $z=1-x^2-y^2$, $z'_x(x, y)=-2x$, $z'_y(x, y)=-2y$ — непрерывны. По формуле (4) получаем

$$\iint\limits_S f(x, y, z)\,dS = \iint\limits_S \sqrt{1+4x^2+4y^2}\,dS =$$

$$= \iint\limits_G \sqrt{1+4x^2+4y^2}\cdot\sqrt{1+4x^2+4y^2}\,dxdy =$$

$$= \iint\limits_G (1+4x^2+4y^2)\,dxdy.$$

Переходя в полученном двойном интеграле к полярным координатам $x=\rho\cos\varphi$, $y=\rho\sin\varphi$, находим

$$\iint\limits_G (1+4x^2+4y^2)\,dxdy = \int\limits_0^{2\pi} d\varphi \int\limits_0^1 (1+4\rho^2)\,\rho d\rho =$$

$$= \int\limits_0^{2\pi}\left[\frac{\rho^2}{2}+\rho^4\right]_0^1 d\varphi = \frac{3}{2}\int\limits_0^{2\pi} d\varphi = \frac{3}{2}\,\varphi\Big|_0^{2\pi} = 3\pi.$$

3. Определение поверхностного интеграла второго рода. Для того чтобы определить поверхностный интеграл второго рода, нам необходимо сначала ввести понятие стороны поверхности.

Возьмем на поверхности S произвольную точку M, проведем через нее нормаль к поверхности и выберем на этой нормали одно из двух возможных направлений. Проведем теперь на поверхности S через точку M какой-либо замкнутый контур, не имеющий общих точек с границей поверхности S, и начнем перемещать точку M по замкнутому контуру так, чтобы вектор все время оставался нормальным к S и чтобы его направление менялось при этом перемещении непрерывно (рис. 83). В прежнее положение точка M вернется либо с тем же направлением нормали, либо с прямо противоположным.

Если обход по любому замкнутому контуру, лежащему на поверхности S и не пересекающему ее границы, при возвращении в исходную точку не меняет направления нормали к поверхности, то поверхность называется *двусторонней*.

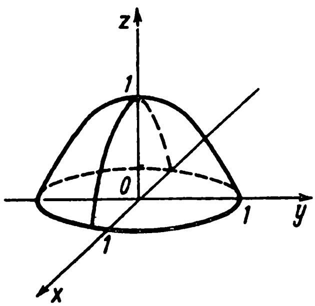


Рис. 82

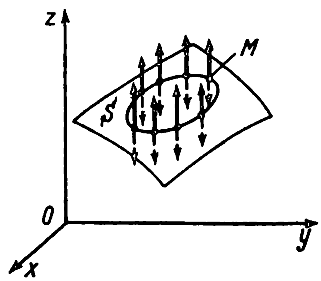


Рис. 83

Примерами двусторонних поверхностей могут служить плоскость, сфера, любая поверхность, задаваемая уравнением $z=f(x, y)$, где $f(x, y)$, $f'_x(x, y)$ и $f'_y(x, y)$ — функции, непрерывные в некоторой области G плоскости Oxy и т. д.

Если же на поверхности S существует замкнутый контур, при обходе которого направление нормали меняется после возвращения в исходную точку на противоположное, то поверхность называется *односторонней*.

Простейшим примером односторонней поверхности может служить так называемый лист Мёбиуса*, изображенный на рис. 84. Его можно получить, взяв полос-

* Мёбиус Август Фердинанд (1790—1868) — немецкий математик.

ку бумаги $ABCD$ и склеив ее так, чтобы точка A совпала с точкой C, а точка B с точкой D, т. е. повернув перед склеиванием один из ее краев на 180°. Легко видеть, что при обходе листа Мёбиуса по его средней линии направление нормали к нему меняется на противоположное.

В дальнейшем мы будем рассматривать лишь двусторонние поверхности.

Для двусторонней поверхности совокупность всех точек с выбранными в них направлениями нормали называется *стороной поверхности*, а выбор определенной ее стороны — *ориентацией поверхности*. Двустороннюю поверхность называют также *ориентируемой*, а одностороннюю — *неориентируемой*.

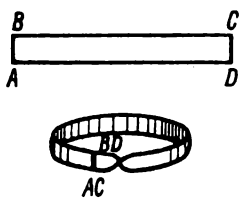


Рис. 84

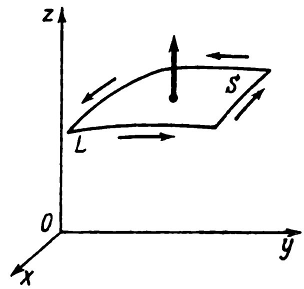


Рис. 85

С понятием стороны поверхности тесно связано понятие ориентации ее границы.

Пусть S — ориентированная (сторона уже выбрана) поверхность, ограниченная контуром L, не имеющим точек самопересечения. Будем считать *положительным* направлением обхода контура L то, при движении по которому сама поверхность остается слева по отношению к точке, совершающей обход (рис. 85). Противоположное направление будем считать *отрицательным*. Если изменить ориентацию поверхности, то положительное и отрицательное направления обхода контура L поменяются ролями.

Теперь мы можем перейти к определению поверхностного интеграла второго рода.

Пусть S — некоторая ориентирумая поверхность, заданная уравнением $z=f(x, y)$, и пусть $R(x, y, z)$ — функция, определенная в точках поверхности S. Выберем одну из двух сторон поверхности, т. е. выберем одно из двух возможных направлений векторов нормали в точках поверхности (тем самым мы ориентировали поверхность). Если векторы нормалей составляют острые углы с осью Oz, то будем говорить, что выбрана *верхняя сторона* поверхности $z=f(x, y)$, если тупые углы, то — *нижняя сторона* поверхности. Разобьем поверхность S произвольно на n частей, не имеющих общих внутренних точек. Обозначим через G_i проекцию i-й части поверхности на плоскость Oxy. Выбрав на каждой частичной поверхности произвольную точку $M_i(\xi_i; \eta_i; \zeta_i)$, составим сумму

$$\sum_{i=1}^{n} R(\xi_i, \eta_i, \zeta_i)\,\Delta s_i, \tag{5}$$

где Δs_i — площадь G_i, взятая со знаком плюс, если выбрана верхняя сторона поверхности S, и со знаком минус, если выбрана нижняя сторона поверхности S. Сумма (5) называется *интегральной суммой* для функции $R(M)=R(x, y, z)$. Обозначим через λ наибольший из диаметров частей поверхности S и дадим следующее

Определение. *Если интегральная сумма (5) при* $\lambda\to 0$ *имеет конечный предел I, то этот предел называется поверхностным интегралом второго рода от функции $R(x, y, z)$ по выбранной стороне поверхности S и обозначается одним из символов:*

$$J = \iint\limits_S R(M)\,dxdy = \iint\limits_S R(x, y, z)\,dxdy.$$

В этом случае функция $R(x, y, z)$ называется *интегрируемой* по поверхности S по переменным x и y.

Аналогичным образом определяется поверхностный интеграл второго рода по ориентируемой поверхности S по переменным y и z [z и x] от функции $P(x, y, z)$[$Q(x, y, z)$], определенной на поверхности S:

$$\iint\limits_S P(x, y, z)\,dydz \left[\iint\limits_S Q(x, y, z)\,dzdx\right].$$

Сумму

$$\iint\limits_S P(x, y, z)\,dydz + \iint\limits_S Q(x, y, z)\,dzdx + \iint\limits_S R(x, y, z)\,dxdy$$

принято называть общим поверхностным интегралом второго рода и обозначать символом

$$\iint\limits_S P(x, y, z)\,dydz + Q(x, y, z)\,dzdx + R(x, y, z)\,dxdy. \quad (6)$$

Поверхностный интеграл второго рода обладает всеми свойствами поверхностного интеграла первого рода, за исключением одного: при изменении стороны поверхности (переориентации) интеграл меняет знак.

К понятию поверхностного интеграла второго рода приводит, например, решение так называемой задачи *о потоке векторного поля*, которая будет рассмотрена в § 14.

Для односторонней поверхности понятие поверхностного интеграла второго рода не вводится.

4. Вычисление поверхностных интегралов второго рода. Вычисление и условия существования поверхностного интеграла второго рода вытекают из сведения поверхностного интеграла к двойному.

Пусть ориентированная (выберем верхнюю сторону) поверхность S задана уравнением $z=f(x, y)$, где $f(x, y)$ — непрерывная функция в замкнутой области G — проекции поверхности S на плоскость Oxy, а $R(x, y, z)$ — непрерывная функция на поверхности S.

Разобьем произвольно поверхность S на n частей, не имеющих общих внутренних точек, и спроектируем это разбиение на плоскость Oxy (рис. 86). Область G разобьется соответственно на части $G_1, G_2, \ldots, G_n$. Выберем на каждой части поверхности произвольную точку $M_i(\xi_i; \eta_i; \zeta_i)$ и составим интегральную сумму

$$\sum_{i=1}^{n} R(\xi_i, \eta_i, \zeta_i)\,\Delta s_i,$$

где Δs_i — площадь G_i. Так как $\zeta_i=f(\xi_i, \eta_i)$, то

$$\sum_{i=1}^{n} R(\xi_i, \eta_i, \zeta_i)\,\Delta s_i = \sum_{i=1}^{n} R[\xi_i, \eta_i, f(\xi_i, \eta_i)]\,\Delta s_i. \quad (7)$$

В правой части равенства стоит интегральная сумма для двойного интеграла от непрерывной в области G функции $R[x, y, f(x, y)]$. Переходя к пределу в (7) при $\lambda \to 0$, получаем требуемую формулу

$$\iint_S R(x, y, z)\,dxdy = \iint_G R[x, y, f(x, y)]\,dxdy, \tag{8}$$

выражающую поверхностный интеграл второго рода по переменным x и y через двойной. Кроме этого, формула (8) доказывает существование поверхностного ин-

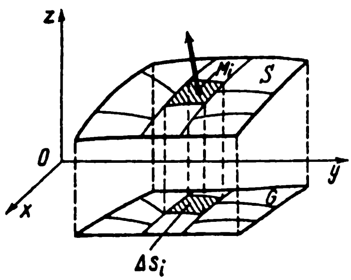


Рис. 86

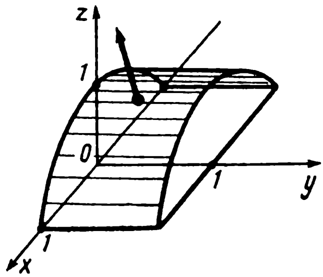


Рис. 87

теграла от функции $R(x, y, z)$, непрерывной на рассматриваемой поверхности S. Если выбрать нижнюю сторону поверхности, то перед интегралом в правой части (8) появится знак минус.

Аналогично устанавливается справедливость формул

$$\iint_S P(x, y, z)\,dydz = \iint_{G_1} P[f(y, z), y, z]\,dydz, \tag{9}$$

$$\iint_S Q(x, y, z)\,dzdx = \iint_{G_2} Q[x, f(x, z), z]\,dzdx, \tag{10}$$

где G_1 и G_2 — проекции поверхности S соответственно на плоскости Oyz и Oxz.

Для вычисления интеграла общего вида (6) используются те же формулы (8)—(10), если поверхность S однозначно проектируется на все три плоскости. В более сложных случаях поверхность S разбивается на части, обладающие указанными свойствами, а интеграл (6) — на сумму интегралов по этим частям.

П р и м е р 2. Вычислить интеграл $\iint\limits_S (y^2+z^2)\,dxdy$, где S — верхняя сторона поверхности $z=\sqrt{1-x^2}$, отсеченной плоскостями $y=0$, $y=1$ (рис. 87).

Р е ш е н и е. Проекцией G данной поверхности на плоскость Oxy является прямоугольник, определяемый неравенствами $-1\leqslant x\leqslant 1$, $0\leqslant y\leqslant 1$. По формуле (8) находим

$$\iint\limits_S (y^2+z^2)\,dxdy=\iint\limits_G [y^2+(\sqrt{1-x^2})^2]\,dxdy=$$

$$=\int\limits_{-1}^{1}dx\int\limits_0^1 (y^2+1-x^2)\,dy=\int\limits_{-1}^{1}\left[\frac{y^3}{3}+y-x^2y\right]_0^1 dx=$$

$$=\int\limits_{-1}^{1}\left(\frac{4}{3}-x^2\right)dx=\left[\frac{4}{3}x-\frac{x^3}{3}\right]_{-1}^{1}=2.$$

П р и м е р 3. Вычислить интеграл

$$\iint\limits_S xdydz+ydzdx+zdxdy,$$

где S — верхняя сторона части плоскости $x+z-1=0$, отсеченной плоскостями $y=0$, $y=4$ и лежащей в первом октанте (рис. 88).

Р е ш е н и е. По определению

$$\iint\limits_S xdydz+ydzdx+zdxdy=\iint\limits_{G_1} x(y,z)\,dydz+$$

$$+\iint\limits_S ydzdx+\iint\limits_{G_2} z(x,y)\,dxdy.$$

Здесь G_1 и G_2 — проекции поверхности S на плоскости Oyz и Oxy, а $\iint\limits_S ydzdx=0$, так как плоскость S параллельна оси Oy. По формуле (8) находим

$$\iint_S z\,dx\,dy = \iint_{G_2} (1-x)\,dx\,dy = \int_0^4 dy \int_0^1 (1-x)\,dx = 2.$$

По формуле (9) находим

$$\iint_S x\,dy\,dz = \iint_{G_1} (1-z)\,dy\,dz = \int_0^4 dy \int_0^1 (1-z)\,dz = 2.$$

Следовательно,

$$\iint_S x\,dy\,dz + y\,dz\,dx + z\,dx\,dy = 2 + 0 + 2 = 4.$$

5. Связь между поверхностными интегралами первого и второго рода. Поверхностные интегралы второго рода можно ввести и другим способом, а именно как поверхностные интегралы первого рода, в которых под

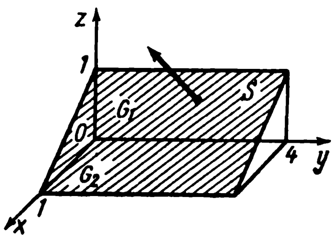


Рис. 88

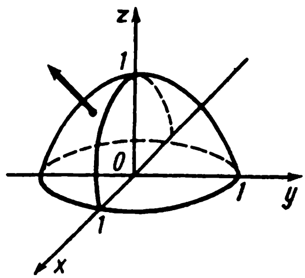


Рис. 89

знаком интеграла стоят некоторые специальные выражения. Обозначим через $\cos\alpha$, $\cos\beta$, $\cos\gamma$ направляющие косинусы нормали ориентированной поверхности S в произвольной ее точке. Поверхностные интегралы второго рода различаются своими отношениями к координатным плоскостям:

1) Поверхностный интеграл второго рода для плоскости Oxy от функции $R(x, y, z)$ выражается через поверхностный интеграл первого рода следующей формулой:

$$\iint_S R(x, y, z)\,dx\,dy = \iint_S R(x, y, z)\cos\gamma\,dS. \qquad (11)$$

2) Поверхностный интеграл второго рода для плоскости Oxz от функции $Q(x, y, z)$ выражается через поверхностный интеграл первого рода следующей формулой:

$$\iint\limits_S Q(x, y, z)\,dzdx = \iint\limits_S Q(x, y, z)\cos\beta dS. \qquad (12)$$

3) Поверхностный интеграл второго рода для плоскости Oyz от функции $P(x, y, z)$ выражается через поверхностный интеграл первого рода следующей формулой:

$$\iint\limits_S P(x, y, z)\,dydz = \iint\limits_S P(x, y, z)\cos\alpha dS. \qquad (13)$$

Суммируя (11), (12) и (13), получаем формулу, выражающую поверхностный интеграл второго рода общего вида по выбранной стороне поверхности через поверхностный интеграл первого рода:

$$\iint\limits_S Pdydz + Qdzdx + Rdxdy =$$

$$= \iint\limits_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,dS. \qquad (14)$$

Если выбрать другую сторону поверхности, то направляющие косинусы нормали $\cos\alpha$, $\cos\beta$, $\cos\gamma$ изменят знак, и, следовательно, поверхностный интеграл второго рода изменит знак.

Пример 4. Вычислить интеграл $\iint\limits_S z\cos\gamma dS$, где S — внешняя сторона сферы $x^2+y^2+z^2=1$, расположенной над плоскостью Oxy, а γ — острый угол между нормалью к поверхности S с осью Oz (рис. 89).

Решение. Воспользуемся формулой (11), связывающей поверхностные интегралы обоих типов. По этой формуле

$$\iint\limits_S z\cos\gamma dS = \iint\limits_S z\,dxdy.$$

Проекцией G данной поверхности S на плоскость Oxy является круг $x^2+y^2\leqslant 1$. По формуле (8) получаем

$$\iint\limits_S z\,dxdy = \iint\limits_G \sqrt{1-x^2-y^2}\,dxdy.$$

Переходя в двойном интеграле к полярным координатам, находим

$$\iint\limits_G \sqrt{1-x^2-y^2}\,dxdy = \int\limits_0^{2\pi} d\varphi \int\limits_0^1 \sqrt{1-\rho^2}\,\rho d\rho =$$

$$= \int\limits_0^{2\pi} \left[-\frac{(1-\rho^2)^{3/2}}{3}\right]_0^1 d\varphi = \frac{1}{3}\int\limits_0^{2\pi} d\varphi = \frac{2}{3}\pi.$$

§ 12. Формула Остроградского

Формула Остроградского* устанавливает связь между поверхностным интегралом по замкнутой поверхности и тройным интегралом по пространственной области, ограниченной этой поверхностью. Эта формула является аналогом формулы Грина, которая, как известно, связывает криволинейный интеграл по замкнутой кривой с двойным интегралом по плоской области, ограниченной этой кривой. Формула Остроградского имеет широкое применение как в самом анализе, так и в его приложениях.

Мы выведем эту формулу для замкнутой пространственной области, граница которой пересекается с любой прямой, параллельной осям координат, не более чем в двух точках. При этом мы будем рассматривать внешнюю сторону поверхности, ограничивающей эту область. Назовем, для краткости, такие области *простыми*. Предполагается, что поверхность гладкая или кусочно-гладкая.

Теорема 13.8. *Пусть V — простая замкнутая область, ограниченная поверхностью S, и пусть функции $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ непрерывны вместе со своими частными производными первого порядка в данной области. Тогда имеет место следующая формула:*

---

* Остроградский Михаил Васильевич (1801—1861) — выдающийся русский математик.

$$\iiint\limits_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dxdydz =$$

$$= \iint\limits_S Pdydz + Qdzdx + Rdxdy, \qquad (1)$$

называемая формулой Остроградского.

Доказательство. Пусть область G есть проекция поверхности S (и области V) на плоскость Oxy (рис. 90), а $z=z_1(x, y)$ и $z=z_2(x, y)$ суть уравнения соответствующих частей S— нижней части S_1 и верхней S_2, где $z_1(x, y)$ и $z_2(x, y)$ — непрерывные функции в области G. Преобразуем тройной интеграл

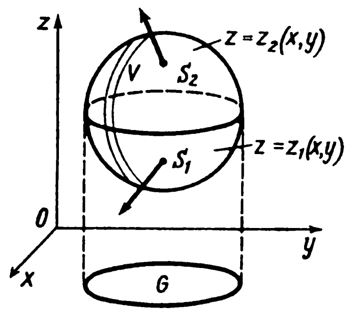


Рис. 90

$$\iiint\limits_V \frac{\partial R}{\partial z} dxdydz$$

в поверхностный. Для этого сведем его к повторному интегралу и по формуле Ньютона—Лейбница выполним интегрирование по z. Получим

$$\iiint\limits_V \frac{\partial R}{\partial z} dxdydz = \iint\limits_G dxdy \int\limits_{z_1(x,y)}^{z_2(x,y)} \frac{\partial R}{\partial z} dz =$$

$$= \iint\limits_G R[x, y, z_2(x, y)]\, dxdy - \iint\limits_G R[x, y, z_1(x, y)]\, dxdy.$$

Так как область G является проекцией на плоскость Oxy и поверхности S_2, и поверхности S_1, то двойные интегралы могут быть заменены равными им поверхностными интегралами (см. (8), п. 4, § 11), взятыми соответственно по верхней стороне поверхности $z=z_2(x, y)$ и верхней стороне поверхности $z=z_1(x, y)$, т. е.

$$\iiint\limits_V \frac{\partial R}{\partial z} dxdydz = \iint\limits_{S_2} R(x, y, z)\, dxdy - \iint\limits_{S_1} R(x, y, z)\, dxdy.$$

Меняя в интеграле по S_1 сторону поверхности, получаем

$$\iiint\limits_V \frac{\partial R}{\partial z}\,dxdydz = \iint\limits_{S_2} R(x, y, z)\,dxdy +$$
$$+ \iint\limits_{S_1} R(x, y, z)\,dxdy = \iint\limits_S Rdxdy, \tag{2}$$

где S — внешняя сторона поверхности, ограничивающей область V.

Аналогичным образом доказываются формулы

$$\iiint\limits_V \frac{\partial P}{\partial x}\,dxdydz = \iint\limits_S P\,dydz, \tag{3}$$

$$\iiint\limits_V \frac{\partial Q}{\partial y}\,dxdydz = \iint\limits_S Qdzdx. \tag{4}$$

Складывая почленно равенства (2), (3), (4), мы и приходим к формуле (1) ■

Замечание 1. Формула Остроградского верна для любой замкнутой пространственной области V, которую можно разбить на конечное число простых областей. В самом деле, применяя формулу (1) к каждой из областей разбиения и складывая результаты, получим слева тройной интеграл по всей области V, а справа — поверхностный интеграл по поверхности S, ограничивающей область V, так как поверхностные интегралы по вспомогательным поверхностям берутся дважды по противоположным сторонам и при суммировании взаимно уничтожаются.

Замечание 2. Можно показать, что формула Остроградского справедлива для любой области V, ограниченной несколькими кусочно-гладкими поверхностями. Но мы не будем на этом останавливаться.

С помощью формулы Остроградского удобно вычислять поверхностные интегралы по замкнутым поверхностям. Рассмотрим

Пример. Вычислить интеграл

$$\iint\limits_S x^3dydz + y^3dzdx + z^3dxdy,$$

где S — внешняя сторона сферы $x^2+y^2+z^2=R^2$.

Решение. Применяя формулу Остроградского, будем иметь

$$\iint_S x^3dydz + y^3dzdx + z^3dxdy =$$

$$= 3\iiint_V (x^2 + y^2 + z^2)\,dxdydz,$$

откуда, введя сферические координаты, получаем

$$3\iiint_V (x^2 + y^2 + z^2)\,dxdydz =$$

$$= 3\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta \int_0^R \rho^4 d\rho = \frac{12}{5}\pi R^5.$$

Мы отмечали (см. п. 1, § 9), что формула Грина дает выражение для площади области через криволинейный интеграл по ее границе. Точно так же из формулы Остроградского легко получить выражение для объема области в виде поверхностного интеграла по замкнутой поверхности S — границе этой области. Действительно, подберем функции P, Q и R так, чтобы

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \equiv 1.$$

Тогда получим

$$v = \iiint_V dxdydz = \iint_S Pdydz + Qdzdx + Rdxdy,$$

где v — объем, ограниченный поверхностью S. В частности, положив

$$P = \frac{1}{3}x, Q = \frac{1}{3}y, R = \frac{1}{3}z,$$

мы получим для вычисления объема формулу

$$v = \frac{1}{3}\iint_S xdydz + ydzdx + zdxdy.$$

§ 13. Формула Стокса

Формула Стокса* устанавливает связь между поверхностным и криволинейным интегралами. Подобно формулам Грина и Остроградского, формула Стокса широко применяется как в самом анализе, так и в его приложениях.

Пусть S — поверхность, заданная уравнением $z=z(x, y)$, где функции $z(x, y)$, $z'_x(x, y)$, $z'_y(x, y)$ непрерывны в замкнутой области G — проекции S на плоскость Oxy, L — контур, ограничивающий S, а l — его проекция на плоскость Oxy, являющаяся контуром, ограничивающим область G. Выберем для ориентации верхнюю сторону поверхности S (рис. 91). Тогда при выполнении сделанных предположений справедлива следующая

Теорема 13.9. *Если функции $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ непрерывны вместе со своими частными производными первого порядка на поверхности S, то имеет место следующая формула:*

$$\oint_L Pdx + Qdy + Rdz = \iint_S \left[\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \cos\gamma + \right.$$

$$\left. + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \cos\beta \right] dS, \qquad (1)$$

называемая формулой Стокса, где $\cos\alpha$, $\cos\beta$, $\cos\gamma$ — *направляющие косинусы нормали к поверхности S, а контур L пробегается в положительном направлении.*

Доказательство. Преобразуем криволинейный интеграл

$$\oint_L Pdx + Qdy + Rdz,$$

взятый по контуру L, в интеграл по поверхности S. Это преобразование проведем по следующей схеме:

$$\oint_L \to \oint_l \to \iint_G \to \iint_S.$$

---

* Стокс Джордж Габриель (1819—1903) — английский физик и математик.

т. е. криволинейный интеграл по пространственному контуру L преобразуем сперва в криволинейный интеграл по плоскому контуру l, затем переведем его в двойной интеграл по области G и, наконец, этот последний преобразуем в интеграл по поверхности S.

Рассмотрим сначала криволинейный интеграл вида

$$\oint_L P(x, y, z)\,dx.$$

Так как контур L лежит на поверхности S, то координаты его точек удовлетворяют уравнению $z=z(x, y)$, и поэтому значения функции $P(x, y, z)$ в точках контура L равны значениям функции $P[x, y, z(x, y)]$ в соответствующих точках контура l, являющегося проекцией L. Проекции же соответствующих участков разбиения контуров L и l на ось Ox совпадают. Значит, совпадают интегральные суммы для криволинейных интегралов второго рода от функции P по контурам L и l, а значит, равны и интегралы

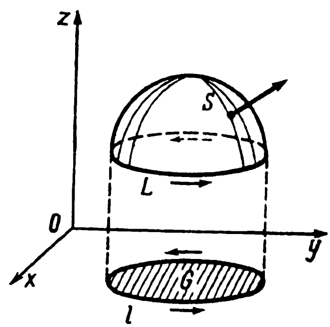


Рис. 91

$$\oint_L P(x, y, z)\,dx = \oint_l P\,[x, y, z\,(x, y)]\,dx.$$

Далее, применяя формулу Грина, перейдем к двойному интегралу по области G, получаем

$$\oint_l P\,[x, y, z\,(x, y)]\,dx = -\iint_G \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\,z'_y\right) dxdy.$$

Здесь подынтегральная функция равна частной производной по y от сложной функции, получающейся из $P(x, y, z)$ после подстановки $z(x, y)$ вместо z.

Так как S — верхняя сторона поверхности, т. е. $\cos\gamma>0$ (γ — острый угол между нормалью и осью Oz), то нормальный вектор имеет проекции $-z'_x$, $-z'_y$, 1. А так как направляющие косинусы нормали пропорциональны соответствующим проекциям, то

$$\frac{\cos\beta}{\cos\gamma} = \frac{-z'_y}{1} = -z'_y.$$

Поэтому

$$-\iint_G \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} z'_y\right) dxdy =$$

$$= -\iint_G \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\cdot\frac{\cos\beta}{\cos\gamma}\right) dxdy.$$

Теперь, воспользовавшись формулами (8) и (11) § 11, можно этот двойной интеграл преобразовать в поверхностный. Получаем

$$-\iint_G \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\cdot\frac{\cos\beta}{\cos\gamma}\right) dxdy =$$

$$= -\iint_S \left(\frac{\partial P}{\partial y}\cos\gamma - \frac{\partial P}{\partial z}\cos\beta\right) dS.$$

Итак,

$$\oint_L P(x, y, z)\, dx = \iint_S \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right) dS. \quad (2)$$

Аналогично доказывается при соответствующих условиях справедливость еще двух формул:

$$\oint_L Q(x, y, z)\, dy = \iint_S \left(\frac{\partial Q}{\partial x}\cos\gamma - \frac{\partial Q}{\partial z}\cos\alpha\right) dS, \quad (3)$$

$$\oint_L R(x, y, z)\, dz = \iint_S \left(\frac{\partial R}{\partial y}\cos\alpha - \frac{\partial R}{\partial x}\cos\beta\right) dS. \quad (4)$$

Складывая почленно равенства (2), (3), (4), получаем доказываемую формулу (1) ■

Формулу Стокса с помощью формулы связи поверхностных интегралов (14) § 11 можно переписать в следующем виде:

$$\oint_L Pdx + Qdy + Rdz = \iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy +$$

$$+\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)dydz+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)dzdx.$$

Формулу Стокса легко запомнить, заметив, что первое слагаемое в правой ее части — это то же выражение, которое стоит под знаком двойного интеграла в формуле Грина, а второе и третье получаются из него циклической перестановкой координат x, y, z и функций P, Q, R.

В частности, если поверхность S — область плоскости Oxy, ограниченная контуром L, то интегралы по $dzdx$ и $dydz$ обращаются в нуль, и формула Стокса переходит в формулу Грина.

Формула Стокса позволяет вычислять криволинейные интегралы по замкнутым контурам с помощью поверхностных интегралов. Рассмотрим

Пример. Вычислить с помощью формулы Стокса интеграл $\oint\limits_L x^2y^3dx+dy+zdz$, где L — окружность, заданная уравнениями $x^2+y^2=1$, $z=0$, а поверхностью S служит верхняя сторона полусферы $x^2+y^2+z^2=1$ ($z>0$), контур L пробегается в положительном направлении.

Решение. Применяя формулу Стокса, имеем

$$\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}=-3x^2y^2;\ \frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}=0;\ \frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}=0.$$

Тогда

$$\oint\limits_L x^2y^3dx+dy+zdz=-3\iint\limits_S x^2y^2dxdy=-\frac{\pi}{8}.$$

Из формулы Стокса следует, что если

$$\frac{\partial Q}{\partial x}=\frac{\partial P}{\partial y};\ \frac{\partial R}{\partial y}=\frac{\partial Q}{\partial z};\ \frac{\partial P}{\partial z}=\frac{\partial R}{\partial x}, \tag{5}$$

то криволинейный интеграл по любой пространственной замкнутой кривой L равен нулю:

$$\oint\limits_L Pdx+Qdy+Rdz=0. \tag{6}$$

А это значит, что в данном случае криволинейный интеграл не зависит от пути интегрирования.

Как и в случае плоской кривой, условия (5) являются необходимыми и достаточными для вычисления равенства (6).

При выполнении этих условий подынтегральные выражения $Pdx+Qdy+Rdz$ представляют собой полный дифференциал некоторой функции $U(x, y, z)$:

$$dU=Pdx+Qdy+Rdz$$

и, следовательно,

$$\int\limits_{(x_0;y_0;z_0)}^{(x_1;y_1;z_1)} Pdx+Qdy+Rdz=U(x_1;y_1;z_1)-U(x_0;y_0;z_0).$$

Это устанавливается так же, как соответствующая формула (4), § 8, для функции двух переменных.

§ 14. Скалярное и векторное поля. Понятие потенциального поля

Понятие поля лежит в основе многих представлений современной физики. Изучение теории поля выходит за рамки данного курса. Поэтому мы ограничимся краткими сведениями.

В общем случае говорят, что в пространстве задано *поле* некоторой величины u, если в каждой точке пространства (или некоторой его части) определено значение этой величины. Так, например, при изучении потока газа приходится исследовать несколько полей: температурное поле (в каждой точке температура имеет определенное значение), поле давлений, поле скоростей и т. д.

Поле величины u называется *стационарным* (или *установившимся*), если u не зависит от времени t. В противоположном случае называется *нестационарным* (или *неустановившимся*). Таким образом, величина u есть функция точки M и времени t.

В физических задачах чаще всего приходится иметь дело с скалярными и векторными величинами. В соответствии с этим различают два вида полей: *скалярные* и *векторные*. Для простоты будем считать их стационарными.

1. Скалярное поле. Пусть G — некоторая область на плоскости или в пространстве. *Если в каждой точке M из G определена скалярная величина u, то говорят, что в области G задано скалярное поле.* Понятия скалярного поля и функции, определенной в области G, совпадают. Обычно используется следующая терминология: скалярное поле задается с помощью функции $u=F(M)$, которая называется *скалярной функцией*. Если в пространстве ввести систему координат $Oxyz$, то каждая точка M в этой системе будет иметь определенные координаты x, y, z и скалярная величина u станет функцией этих координат: $u=F(M)=F(x, y, z)$.

Примером скалярного поля может служить поле температур воздуха в некотором помещении, если температуру рассматривать как функцию точки. В точках, расположенных ближе к источнику тепла, температура будет выше, дальше от источника тепла — ниже. Если окажется, что температура везде одинаковая, то в этом случае скалярное поле постоянно.

2. Векторное поле. В полной аналогии с понятием скалярного поля вводится понятие векторного поля: *если в каждой точке M из G определен вектор $\bar{F}(M)$, то говорят, что в области G задано векторное поле.* Функция $\bar{F}(M)$, с помощью которой задается векторное поле, называется *векторной функцией.*

Примером векторного поля может служить поле сил любой природы. Каждой точке области соответствует определенный вектор, имеющий численную величину и направление силы в этой точке.

3. Потенциальное поле. Перейдем теперь к определению понятия потенциального поля. Рассмотрим некоторое скалярное поле $F(M)$. *Если в каждой точке M из G определен вектор* $\operatorname{grad} F$, *то поле этого вектора называется потенциальным полем.* Само скалярное поле называется при этом *потенциалом* векторного поля, а вектор, определяющий потенциальное поле, часто также называют *потенциальным* вектором, т. е. вектор $\bar{a}(M)$ потенциальный, если найдется такая скалярная функция $F(M)$, что

$$\bar{a} = \operatorname{grad} F = \frac{\partial F}{\partial x}\bar{i} + \frac{\partial F}{\partial y}\bar{j} + \frac{\partial F}{\partial z}\bar{k}. \qquad (1)$$

Естественно возникает вопрос, при каких условиях данное векторное поле $\bar{a}(M)$ потенциально. Фактически этот вопрос мы уже рассмотрели в § 7. Пусть P, Q и R — проекции вектора $\bar{a}$ на координатные оси Ox, Oy, Oz соответственно, т. е.

$$\bar{a}=\bar{a}(M)=P\bar{i}+Q\bar{j}+R\bar{k}.$$

В силу соотношения (1) векторное поле $\bar{a}(M)$ будет потенциальным, если найдется функция $F(M)$ такая, что

$$\partial F/\partial x = P;\ \partial F/\partial y = Q;\ \partial F/\partial z = R. \tag{2}$$

В теореме 13.7 было показано, что выражение $Pdx+$ $+Qdy+Rdz$ (где P, Q, R — непрерывные функции, имеющие непрерывные частные производные 1-го порядка) служит полным дифференциалом некоторой функции $F(x, y, z)$ в том и только том случае, если P, Q, R удовлетворяют условиям

$$\partial P/\partial y = \partial Q/\partial x;\ \partial Q/\partial z = \partial R/\partial y;\ \partial R/\partial x = \partial P/\partial z. \tag{3}$$

Но если $Pdx+Qdy+Rdz=dF$, то справедливы и равенства (2), т. е. условие (3) как раз и означает, что данное векторное поле потенциально. Функция $F(x, y, z)$ в этом случае называется *потенциальной* функцией поля.

Примером потенциального поля может служить поле сил тяготения. Если в начале координат помещена масса m, то эта масса создает поле сил тяготения, так как в каждой точке M пространства на помещенную единичную массу, по закону Ньютона, действует сила $\bar{F}(M)$, равная по величине km/r^2 и направленная к началу координат. Здесь $r=|OM|=\sqrt{x^2+y^2+z^2}$ — расстояние от начала координат до точки M, а k — коэффициент пропорциональности.

Пусть x, y, z — координаты точки M. Тогда проекции P, Q, R силы $\bar{F}(M)$ определяются следующим образом:

$$P=|\bar{F}|\cos\alpha = km/r^2\,(-x/r) = -kmx/r^3,$$

$$Q=|\bar{F}|\cos\beta = km/r^2\,(-y/r) = -kmy/r^3,$$

$$R=|\bar{F}|\cos\gamma = km/r^2\,(-z/r) = -kmz/r^3,$$

где $\cos\alpha$, $\cos\beta$, $\cos\gamma$ — направляющие косинусы вектора $\bar{F}(M)$. Следовательно,

$$\bar{F}(M) = -\frac{kmx}{r^3}\bar{i} - \frac{kmy}{r^3}\bar{j} - \frac{kmz}{r^3}\bar{k}.$$

Легко проверить, что данное векторное поле потенциально и его потенциальной функцией будет функция $u(r) = km/r$.

И в заключение данного пункта найдем работу силы $\bar{F}(M)$ при перемещении единичной массы из точки $B(x_1;\ y_1;\ z_1)$ в точку $C(x_2;\ y_2;\ z_2)$.

Как мы знаем, работа A выражается криволинейным интегралом

$$A = \int_{BC} P dx + Q dy + R dz,$$

где P, Q и R — проекции силы $\bar{F}(M)$ на координатные оси. Так как данное силовое поле является потенциальным, то подынтегральное выражение представляет собой полный дифференциал, и в этом случае интеграл не зависит от пути интегрирования и может быть вычислен по формуле

$$A = \int_{B}^{C} P dx + Q dy + R dz = u(C) - u(B),$$

т. е. величина работы силы $\bar{F}(M)$ равна разности значений потенциальной функции в точках C и B. В данном случае

$$A = u(r_2) - u(r_1) = (km/r_2) - (km/r_1) = km[(1/r_2) - (1/r_1)],$$

где r_1 и r_2 — расстояния точек B и C от начала координат.

Заметим, что область, в которой определено поле сил тяготения, есть все пространство, за исключением начала координат.

Закончим этот параграф рассмотрением задачи, о которой говорилось в п. 3 § 11.

4. Задача о потоке векторного поля. Пусть в пространстве задано векторное поле $\bar{v}(M)$ скоростей жидкости, т. е. пространство заполнено движущейся жидкостью, скорость которой в каждой точке $(x;\ y;\ z)$ задается вектором

$$\bar{v}(M) = P(x, y, z)\bar{i} + Q(x, y, z)\bar{j} + R(x, y, z)\bar{k},$$

где P, Q, R — проекции скорости на координатные оси.

Пусть P, Q, R — непрерывные функции координат. Вычислим количество Π жидкости, протекающей за единицу времени через некоторую ориентированную поверхность S, ограниченную пространственной кривой L, считая плотность жидкости $\rho=1$.

Пусть $\bar{n}=\cos\alpha\bar{i}+\cos\beta\bar{j}+\cos\gamma\bar{k}$ — единичный вектор нормали к поверхности S и пусть его направляющие косинусы являются непрерывными функциями координат x, y, z точек данной поверхности.

Разобьем поверхность S произвольно на n частей, не имеющих общих внутренних точек, с площадями ΔS_1, ΔS_2, ..., ΔS_n, и в каждой из них выберем точку $M_i(x_i;\ y_i;\ z_i)$. Подсчитаем количество Π жидкости, протекающей за единицу времени через i-ю часть поверхности (рис. 92). Обозначим через φ угол между векторами $\bar{n}_i$ и $\bar{v}_i=\bar{v}(M_i)$. Приближенно можно считать, что при достаточно мелком разбиении поверхности S скорость $\bar{v}$ во всех точках i-й части постоянна и равна $\bar{v}(M_i)$, а частичные поверхности — плоские. Тогда количество Π жидкости, протекающей через i-ю часть за единицу времени в направлении нормали $\bar{n}_i$, приближенно равно объему цилиндра с основанием ΔS_i и высотой h, равной проекции вектора $\bar{v}_i$ на нормаль $\bar{n}_i$, т. е. $\Pi_i\approx\Delta S_i\cdot h$. А так как $h=|\bar{v}_i|\cos\varphi=|\bar{v}_i|\,|\bar{n}_i|\cos\varphi=(\bar{v}_i\cdot\bar{n}_i)$, то

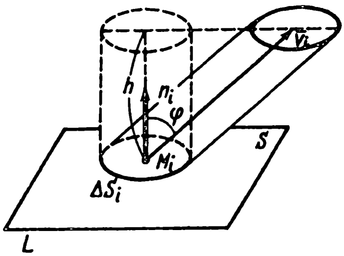


Рис. 92

$$\Pi_i\approx(\bar{v}_i\cdot\bar{n}_i)\Delta S_i.$$

Проведя аналогичные рассуждения для каждой i-й части разбиения, получим приближенное значение количества Π жидкости, протекающей через поверхность S за единицу времени:

$$\Pi\approx\sum_{i=1}^{n}(\bar{v}_i\cdot\bar{n}_i)\,\Delta S_i,$$

т.е. получили интегральную сумму для функции $(\bar{v}\cdot\bar{n})$. Так как проекции P, Q, R вектора $\bar{v}$ и направляющие косинусы вектора $\bar{n}$ являются непрерывными функциями координат x, y, z точек поверхности S, то скалярное произведение $\bar{v}\cdot\bar{n}=P\cos\alpha+Q\cos\beta+R\cos\gamma$ есть непрерывная функция. Следовательно, предел этой суммы при стремлении к нулю наибольшего из диаметров частей поверхности существует и численно равен поверхностному интегралу первого рода по поверхности S от функции $(\bar{v}\cdot\bar{n})$:

$$\Pi=\lim_{\lambda\to 0}\sum_{i=1}^{n}(\bar{v}_i\cdot\bar{n}_i)\,\Delta S_i=\iint_S(\bar{v}\cdot\bar{n})\,dS,$$

или, выражая скалярное произведение через координаты векторов,

$$\Pi=\lim_{\lambda\to 0}\sum_{i=1}^{n}[P(x_i;y_i;z_i)\cos\alpha_i+Q(x_i;y_i;z_i)\cos\beta_i+$$

$$+R(x_i;y_i;z_i)\cos\gamma_i]\,\Delta S_i=\iint_S(P\cos\alpha+Q\cos\beta+R\cos\gamma)\,dS.$$

Воспользовавшись формулой (14), § 11, связывающей поверхностные интегралы, окончательно получим

$$\Pi=\iint_S Pdydz+Qdzdx+Rdxdy. \tag{4}$$

Таким образом, *количество* Π *жидкости, протекающей за единицу времени через некоторую ориентированную поверхность* S, *представляет собой поверхностный интеграл второго рода.*

Для произвольного векторного поля $\bar{v}(M)$ поверхностный интеграл второго рода (4) называется *потоком вектора* $\bar{v}(M)$ (или *потоком векторного поля* $\bar{v}(M)$) через поверхность S. Для векторного поля иной природы, чем в примере, поток, разумеется, имеет другой физический смысл.

Часть III

РЯДЫ. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Глава XIV

РЯДЫ

В настоящей главе мы займемся изучением рядов, являющихся очень важным аппаратом, применяемым для вычислений и исследований как в различных разделах самой математики, так и во многих ее приложениях.

§ 1. Понятие числового ряда

1. Основные определения. Пусть дана числовая последовательность $a_1, a_2, a_3, \ldots, a_n, \ldots$ *Выражение вида*

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n \qquad (1)$$

называется числовым рядом или просто рядом.

Числа $a_1, a_2, \ldots, a_n, \ldots$ называются *членами* ряда, член a_n с произвольным номером — *общим членом* ряда.

Суммы конечного числа членов ряда:

$$S_1 = a_1,\ S_2 = a_1 + a_2,\ S_3 = a_1 + a_2 + a_3,\ \ldots,\ S_n = a_1 + a_2 + a_3 + \ldots + a_n,\ \ldots$$

называются *частичными* суммами S_n ряда (1). Так как число членов ряда бесконечно, то частичные суммы ряда образуют бесконечную последовательность частичных сумм:

$$S_1,\ S_2,\ S_3,\ \ldots,\ S_n,\ \ldots \qquad (2)$$

Ряд (1) называется *сходящимся*, если последовательность его частичных сумм (2) сходится к какому-нибудь числу S, которое в этом случае называется *суммой* ряда (1). Символически это записывается так:

$$S = a_1 + a_2 + a_3 + \ldots + a_n + \ldots \text{ или } S = \sum_{n=1}^{\infty} a_n.$$

Если же последовательность частичных сумм (2) расходится, т. е. или не имеет предела, или ее предел равен бесконечности, то ряд (1) называется *расходящимся*.

Рассмотрим несколько примеров.

Пример 1. Покажем, что ряд

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots + \frac{1}{n(n+1)} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n(n+1)}$$

сходится. Возьмем сумму S_n первых n членов ряда:

$$S_n = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots + \frac{1}{n(n+1)}.$$

Легко видеть, что слагаемые могут быть представлены в виде

$$\frac{1}{1\cdot 2} = 1 - \frac{1}{2};\ \frac{1}{2\cdot 3} = \frac{1}{2} - \frac{1}{3};\ \frac{1}{3\cdot 4} =$$
$$= \frac{1}{3} - \frac{1}{4};\ \ldots;\ \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}.$$

Поэтому

$$S_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) +$$
$$+ \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}.$$

Отсюда следует, что предел последовательности частичных сумм этого ряда равен единице:

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty}\left(1 - \frac{1}{n+1}\right) = 1 - \lim_{n\to\infty}\frac{1}{n+1} = 1.$$

Таким образом, данный ряд сходится, и его сумма S равна 1.

Пример 2. Изучим вопрос о сходимости ряда

$$1-1+1-1+\ldots+(-1)^{n-1}+\ldots=\sum_{n=1}^{\infty}(-1)^{n-1}.$$

Поскольку последовательность его частичных сумм имеет вид $S_1=1$, $S_2=0$, $S_3=1$, $S_4=0$, ..., т. е. не сходится ни к какому пределу, то данный ряд расходится.

Пример 3. Рассмотрим ряд, составленный из элементов геометрической прогрессии:

$$a+aq+aq^2+\ldots+aq^{n-1}+\ldots=\sum_{n=1}^{\infty}aq^{n-1}, \qquad (3)$$

n-я частичная сумма S_n этого ряда при $q\neq 1$ имеет вид

$$S_n=a+aq+aq^2+\ldots+aq^{n-1}=\frac{a-aq^n}{1-q}=$$

$$=\frac{a}{1-q}-\frac{aq^n}{1-q}.$$

Отсюда

1) При $|q|<1$ $\lim\limits_{n\to\infty}S_n=\lim\limits_{n\to\infty}\frac{a}{1-q}-\lim\limits_{n\to\infty}\frac{aq^n}{1-q}=\frac{a}{1-q}$, т. е. ряд сходится и его сумма $S=\frac{a}{1-q}$. Например, при $a=1$, $q=\frac{1}{2}$ имеем

$$S=2=1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^{n-1}}+\ldots$$

2) При $|q|>1$ $\lim\limits_{n\to\infty}S_n=\lim\limits_{n\to\infty}\frac{a-aq^n}{1-q}=\infty$, т. е. ряд расходится.

3) При $q=1$ ряд (3) принимает вид $a+a+a+\ldots+a+\ldots$ В этом случае $\lim\limits_{n\to\infty}S_n=\lim\limits_{n\to\infty}(\underbrace{a+a+a+\ldots+a}_{n\text{ раз}})=$ $=\lim\limits_{n\to\infty}na=\infty$, т. е. ряд расходится.

4) При $q=-1$ ряд (3) принимает вид $a-a+a-a+\ldots$ Для него $S_n=\frac{a}{2}-\frac{a(-1)^n}{2}$. В этом случае $S_n=0$ при

n четном и $S_n=a$ при n нечетном. Следовательно, $\lim\limits_{n\to\infty} S_n$ не существует и ряд расходится.

Суммируя все изложенное, можем сказать, что ряд (3) является сходящимся рядом при $|q|<1$ и расходящимся при всех остальных q.

2. Свойства сходящихся рядов.

Теорема 14.1. *Если сходится ряд*

$$a_1+a_2+a_3+\ldots+a_{k-1}+a_k+a_{k+1}+\ldots$$

$$\ldots+a_{n-1}+a_n+\ldots=\sum_{n=1}^{\infty}a_n, \qquad (4)$$

то сходится и ряд

$$a_{k+1}+\ldots+a_{n-1}+a_n+\ldots=\sum_{n=k+1}^{\infty}a_n \qquad (5)$$

и, обратно, если сходится ряд (5), то сходится и ряд (4).

Другими словами, на сходимость ряда не влияет отбрасывание любого конечного числа его первых членов.

Доказательство. Пусть ряд (4) сходится и имеет сумму S, т. е. $\lim\limits_{n\to\infty} S_n=S$. Обозначим через S_k сумму отброшенных членов ряда (4), а через σ_{n-k} — сумму $n-k$ первых членов ряда (5). Тогда

$$S_n=S_k+\sigma_{n-k}, \qquad (6)$$

где S_k — некоторое число, не зависящее от n. Из равенства (6) следует

$$\lim_{n\to\infty}\sigma_{n-k}=\lim_{n\to\infty}(S_n-S_k)=\lim_{n\to\infty}S_n-\lim_{n\to\infty}S_k=S-S_k,$$

т. е. последовательность частичных сумм $\{\sigma_{n-k}\}$ ряда (5) имеет предел, что означает сходимость ряда (5).

Пусть теперь ряд (5) сходится и имеет сумму σ, т. е. $\lim\limits_{n\to\infty}\sigma_{n-k}=\sigma$. Тогда из (6) следует

$$\lim_{n\to\infty}S_n=\lim_{n\to\infty}(S_k+\sigma_{n-k})=\lim_{n\to\infty}S_k+\lim_{n\to\infty}\sigma_{n-k}=S_k+\sigma,$$

т. е. сходимость ряда (4) ■

Над сходящимися рядами можно выполнять обычные арифметические действия.

Теорема 14.2. *Если ряд* $\sum_{n=1}^{\infty} a_n$ *сходится и его сумма равна* S, *то и ряд* $\sum_{n=1}^{\infty} ca_n$, *где* c — *некоторое число, также сходится и его сумма равна* cS.

Доказательство. Пусть S_n — частичная сумма ряда $\sum_{n=1}^{\infty} a_n$, а σ_n — частичная сумма ряда $\sum_{n=1}^{\infty} ca_n$. Тогда

$$\sigma_n = ca_1 + ca_2 + ca_3 + \dots + ca_n = c(a_1 + a_2 + a_3 + \dots + a_n) = cS_n.$$

Отсюда, переходя к пределу при $n \to \infty$, получаем

$$\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} cS_n = c \lim_{n\to\infty} S_n = cS,$$

т. е. последовательность частичных сумм $\{\sigma_n\}$ ряда $\sum_{n=1}^{\infty} ca_n$ сходится к cS. Следовательно, $\sum_{n=1}^{\infty} ca_n = cS$ ■

Теорема 14.3. *Если ряды* $\sum_{n=1}^{\infty} a_n$ *и* $\sum_{n=1}^{\infty} b_n$ *сходятся и их суммы соответственно равны* S *и* σ, *то и ряд* $\sum_{n=1}^{\infty} (a_n \pm b_n)$ *сходится и его сумма равна* $S \pm \sigma$.

Доказательство. Пусть S_n и σ_n — частичные суммы рядов $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} b_n$, а τ_n — частичная сумма ряда $\sum_{n=1}^{\infty} (a_n \pm b_n)$. Тогда

$$\tau_n = (a_1 \pm b_1) + (a_2 \pm b_2) + \dots + (a_n \pm b_n) = (a_1 + a_2 + \dots + a_n) \pm (b_1 + b_2 + \dots + b_n) = S_n \pm \sigma_n.$$

Отсюда, переходя к пределу при $n \to \infty$, получаем

$$\lim_{n\to\infty} \tau_n = \lim_{n\to\infty} (S_n \pm \sigma_n) = \lim_{n\to\infty} S_n \pm \lim_{n\to\infty} \sigma_n = S \pm \sigma,$$

т. е. последовательность частичных сумм $\{\tau_n\}$ ряда $\sum_{n=1}^{\infty} (a_n \pm b_n)$ сходится к $S \pm \sigma$. Следовательно, $\sum_{n=1}^{\infty} (a_n \pm b_n) = S + \sigma$ ■

Таким образом, установлено, что сходящиеся ряды можно умножать на числа, почленно складывать и вычитать так же, как это обычно делается с конечными суммами.

3. Необходимое условие сходимости ряда. При изучении рядов возникают две задачи: 1) исследовать ряд на сходимость и 2) зная, что ряд сходится, найти его сумму. Мы будем решать в основном первую задачу, имеющую теоретический характер. Начнем с необходимого условия сходимости рядов.

Теорема 14.4. *Если ряд* $\sum_{n=1}^{\infty} a_n$ *сходится, то его общий член стремится к нулю, т. е.* $\lim_{n\to\infty} a_n = 0$.

Доказательство. По условию ряд $\sum_{n=1}^{\infty} a_n$ сходится. Обозначим через S его сумму. Рассмотрим частичные суммы ряда $S_n = a_1 + a_2 + \ldots + a_{n-1} + a_n$ и $S_{n-1} = a_1 + a_2 + \ldots + a_{n-1}$. Отсюда $a_n = S_n - S_{n-1}$. Так как $S_n \to S$ и $S_{n-1} \to S$ при $n \to \infty$, то

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0 ■$$

Заметим, однако, что условие $\lim_{n\to\infty} a_n = 0$ является лишь необходимым, но не достаточным условием сходимости ряда.

В качестве примера рассмотрим ряд

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

который обычно называют *гармоническим* рядом. Очевидно, что для гармонического ряда выполнено необходимое условие сходимости, ибо $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \frac{1}{n} = 0$. Докажем, однако, что этот ряд расходится. Действительно, если бы этот ряд сходился, то, обозначая его сумму через S, мы бы имели

$$\lim_{n\to\infty}(S_{2n} - S_n) = \lim_{n\to\infty} S_{2n} - \lim_{n\to\infty} S_n = S - S = 0.$$

Но

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} >$$

$$> \frac{1}{2n} + \frac{1}{2n} + \ldots + \frac{1}{2n} = n \cdot \frac{1}{2n} = \frac{1}{2},$$

т. е. $S_{2n} - S_n > \frac{1}{2}$.

Отсюда следует, что равенство $\lim\limits_{n\to\infty}(S_{2n} - S_n) = 0$ невозможно. Следовательно, гармонический ряд расходится.

Таким образом, если общий член ряда стремится к нулю, то о сходимости ряда ничего еще сказать нельзя. Необходимо дополнительное исследование, которое может быть проведено с помощью достаточных условий (или признаков) сходимости ряда.

Если же для некоторого ряда его общий член не стремится к нулю, то теорема 14.4 позволяет сразу сказать, что такой ряд расходится.

§ 2. Ряды с неотрицательными членами

Теперь перейдем к рассмотрению некоторых достаточных условий сходимости рядов с неотрицательными членами. Но предварительно докажем теорему, которая будет использована в последующих рассуждениях.

Теорема 14.5. *Для того чтобы ряд* $\sum\limits_{n=1}^{\infty} a_n$ *с неотрицательными членами сходился, необходимо и достаточно, чтобы последовательность частичных сумм этого ряда была ограничена.*

Доказательство. *Необходимость.* Пусть ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится. Это значит, что последовательность его частичных сумм имеет конечный предел. В силу теоремы 2.6 всякая сходящаяся последовательность является ограниченной.

Достаточность. Пусть последовательность частичных сумм $\{S_n\}$ ряда $\sum\limits_{n=1}^{\infty} a_n$ ограничена. Так как ряд $\sum\limits_{n=1}^{\infty} a_n$ — с неотрицательными членами, то его частичные суммы образуют неубывающую монотонную последовательность: $0 \leqslant S_1 \leqslant S_2 \leqslant \ldots \leqslant S_n \leqslant S_{n+1} \leqslant \ldots$ В силу теоремы 2.12 о монотонных ограниченных последовательностях она сходится, т. е. сходится ряд $\sum\limits_{n=1}^{\infty} a_n$ ■

Достаточные условия сходимости ряда. В этом пункте мы установим ряд признаков, позволяющих сделать заключение о сходимости (или расходимости) рассматриваемого ряда.

1°. *Признак сравнения.*

Теорема 14.6. *Пусть даны два ряда с неотрицательными членами* $\sum\limits_{n=1}^{\infty} a_n$ *и* $\sum\limits_{n=1}^{\infty} b_n$ *и для всех n выполняется неравенство* $a_n \leqslant b_n$. *Тогда из сходимости ряда* $\sum\limits_{n=1}^{\infty} b_n$ *следует сходимость ряда* $\sum\limits_{n=1}^{\infty} a_n$, *а из расходимости ряда* $\sum\limits_{n=1}^{\infty} a_n$ *следует расходимость ряда* $\sum\limits_{n=1}^{\infty} b_n$.

Доказательство. Обозначим через S_n и σ_n соответственно частичные суммы рядов $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$. Из неравенства $a_n \leqslant b_n$ заключаем, что

$$S_n \leqslant \sigma_n. \tag{7}$$

Если ряд $\sum\limits_{n=1}^{\infty} b_n$ сходится, то по теореме 14.5 (необходимость) последовательность его частичных сумм ограничена, т. е. для любого n $\sigma_n \leqslant M$, где M — некоторое число. Но в силу (7) и $S_n \leqslant M$, откуда по той же теореме 14.5 (достаточность) ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится.

Если же ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится, то ряд $\sum\limits_{n=1}^{\infty} b_n$ также расходится, так как, допустив сходимость ряда $\sum\limits_{n=1}^{\infty} b_n$, получим по только что доказанному сходимость ряда $\sum\limits_{n=1}^{\infty} a_n$, а это противоречит условию теоремы ■

Рассмотрим примеры.

Пример 1. Ряд $\sum\limits_{n=1}^{\infty} \frac{1}{(n+1)^{n+1}}$ сходится, так как сходится ряд из членов геометрической прогрессии $\sum\limits_{n=1}^{\infty} \frac{1}{2^{n+1}} \left(q = \frac{1}{2} < 1\right)$, а члены данного ряда не больше соответствующих членов ряда сходящейся геометрической прогрессии $\frac{1}{(n+1)^{n+1}} \leqslant \frac{1}{2^{n+1}}$.

Пример 2. Ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n^s}$ $(s \leqslant 1)$ расходится, так как его члены не меньше членов гармонического ряда $\sum\limits_{n=1}^{\infty} \frac{1}{n}$: $\frac{1}{n} \leqslant \frac{1}{n^s}$, а гармонический ряд расходится.

Существуют признаки сходимости рядов, позволяющие непосредственно заключать о сходимости (или расходимости) данного ряда, не сравнивая его с другим рядом, относительно которого известно, сходится он или расходится. На практике такие признаки чаще применяются. Рассмотрим два из них.

2°. *Признак Даламбера**.

Теорема 14.7. *Пусть дан ряд* $\sum_{n=1}^{\infty} a_n$ *с положительными членами, и существует предел* $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \rho$. *Тогда а) при* $\rho<1$ *ряд сходится, б) при* $\rho>1$ *ряд расходится.*

Доказательство. а) Пусть $\rho<1$ и $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \rho$. Покажем, что ряд $\sum_{n=1}^{\infty} a_n$ сходится. На основании определения предела числовой последовательности для любого $\varepsilon>0$ существует номер N такой, что при $n \geqslant N$ выполняется неравенство $\left|\frac{a_{n+1}}{a_n} - \rho\right| < \varepsilon$. Отсюда следует, что

$$\rho - \varepsilon < \frac{a_{n+1}}{a_n} < \rho + \varepsilon. \tag{8}$$

Так как $\rho<1$, то ε можем взять настолько малым, что будет выполнено неравенство $\rho+\varepsilon<1$. Полагая $\rho+\varepsilon=q$, на основании правого неравенства (8) имеем

$$a_{n+1}/a_n<q, \text{ или } a_{n+1}<a_n q$$

для $n=N$, $N+1$, $N+2$, ... Давая n эти значения, из последнего неравенства получаем

$$\begin{aligned}
&a_{N+1}<a_N q,\\
&a_{N+2}<a_{N+1} q<a_N q^2,\\
&a_{N+3}<a_{N+2} q<a_N q^3,\\
&\ldots\ldots\ldots
\end{aligned}$$

т. е. члены ряда

$$a_{N+1}+a_{N+2}+a_{N+3}+\ldots \tag{9}$$

меньше соответствующих членов геометрической прогрессии

* Даламбер Жан Лерон (1717—1783) — французский математик, механик и философ-просветитель.

$$a_N q + a_N q^2 + a_N q^3 + \ldots \tag{10}$$

Так как $q<1$, то ряд (10) сходится (см. пример 3, § 1). Тогда по признаку сравнения ряд (9) также сходится. Но ряд (9) получен из данного ряда $\sum\limits_{n=1}^{\infty} a_n$ путем отбрасывания конечного числа первых членов, следовательно, по теореме 14.1 ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится.

б) Пусть теперь $\rho>1$. Покажем, что ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится. Возьмем ε настолько малым, чтобы $\rho-\varepsilon>1$. Тогда при $n\geqslant N$ в силу левого неравенства (8) будет выполняться неравенство $a_{n+1}/a_n>1$, или $a_{n+1}>a_n$. Таким образом, члены ряда, начиная с некоторого номера N, возрастают с увеличением их номеров, т. е. общий член ряда a_n не стремится к нулю при $n\to\infty$. Следовательно, согласно теореме 14.4 ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится ■

Замечание. При $\rho=1$, как показывают примеры, ряд $\sum\limits_{n=1}^{\infty} a_n$ может как сходиться, так и расходиться. В этом случае необходимо дополнительное исследование ряда с помощью признака сравнения или других признаков.

Пример 3. Ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n!}$ сходится, так как

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0 < 1.$$

Пример 4. Ряд $\sum\limits_{n=1}^{\infty} \frac{n^n}{n!}$ расходится, так как

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{(n+1)^{n+1}\cdot n!}{(n+1)!\, n^n} = \lim_{n\to\infty} \left(\frac{n+1}{n}\right)^n =$$

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

Пример 5. Рассмотрим ряд $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$. Находим

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \lim_{n \to \infty} \frac{\sqrt{n}}{\sqrt{n+1}} = 1.$$

По признаку Даламбера сделать заключение о сходимости или расходимости ряда мы не можем. Однако, как было показано ранее (см. пример 2), этот ряд расходится.

3°. *Интегральный признак.*

Теорема 14.8. *Пусть дан ряд*

$$f(1) + f(2) + f(3) + \ldots + f(n) + \ldots = \sum_{n=1}^{\infty} f(n),$$

члены которого являются значениями некоторой функции $f(x)$, *положительной, непрерывной, убывающей на полуинтервале* $[1, +\infty)$. *Тогда, если* $\int_1^{+\infty} f(x)\,dx$ *сходится, то сходится и ряд* $\sum_{n=1}^{\infty} f(n)$; *если же* $\int_1^{+\infty} f(x)\,dx$ *расходится, то ряд* $\sum_{n=1}^{\infty} f(n)$ *также расходится.*

Доказательство. Рассмотрим криволинейную трапецию, ограниченную сверху графиком функции $y=f(x)$, с боков прямыми $x=1$, $x=n$, снизу осью Ox. Впишем в эту трапецию и опишем около нее две ступенчатые фигуры, состоящие из прямоугольников с основаниями $[1, 2]$, $[2, 3]$, ..., $[n-1, n]$ и высотами $f(1)$,

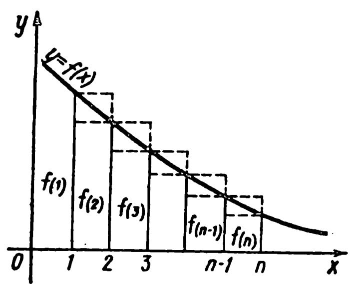


Рис. 93

$f(2)$, $f(3)$, ..., $f(n-1)$, $f(n)$ (рис. 93). Тогда, учитывая геометрический смысл определенного интеграла, имеем

$$f(2)+f(3)+\ldots+f(n)<$$
$$<\int_1^n f(x)\,dx<f(1)+f(2)+\ldots+f(n-1),$$

или, короче,

$$S_n-f(1)<\int_1^n f(x)\,dx<S_n-f(n).$$

Отсюда получаем

$$S_n<f(1)+\int_1^n f(x)\,dx, \tag{11}$$

$$S_n>f(n)+\int_1^n f(x)\,dx, \tag{12}$$

где S_n — частичные суммы рассматриваемого ряда.

Пусть интеграл $\int_1^\infty f(x)\,dx$ сходится. Это значит, что существует $\lim\limits_{n\to\infty}\int_1^n f(x)\,dx=J$. Так как $f(x)>0$, то последовательность $\int_1^n f(x)\,dx$ возрастает с увеличением n и ограничена сверху своим пределом: $\int_1^n f(x)\,dx<J$. Из неравенства (11) следует, что $S_n<f(1)+J$, т. е. последовательность частичных сумм $\{S_n\}$ ряда $\sum\limits_{n=1}^{\infty} f(n)$ ограничена. По теореме 14.5 ряд $\sum\limits_{n=1}^{\infty} f(n)$ сходится.

Пусть теперь интеграл $\int\limits_1^\infty f(x)\,dx$ расходится. В этом случае $\int\limits_1^n f(x)\,dx \to +\infty$ при $n\to\infty$ (как монотонно возрастающая неограниченная последовательность). Из неравенства (12) следует, что $S_n \to +\infty$ при $n\to\infty$, т. е. последовательность частичных сумм $\{S_n\}$ ряда $\sum\limits_{n=1}^\infty f(n)$ расходится и, следовательно, ряд расходится ■

Пример 6. Рассмотрим ряд

$$1+\frac{1}{2^\alpha}+\frac{1}{3^\alpha}+\ldots+\frac{1}{n^\alpha}+\ldots=\sum_{n=1}^\infty \frac{1}{n^\alpha} \quad (\alpha>0).$$

С помощью интегрального признака выясним вопрос о поведении данного ряда при $\alpha>0$. Возьмем в качестве функции $f(x)$ функцию $1/x^\alpha$ $(x\geqslant 1)$; она удовлетворяет условиям теоремы 14.8. Члены ряда равны значениям этой функции при $x=1, 2, 3, \ldots, n, \ldots$. Как мы знаем (см. гл. VIII, § 11, п. 1, пример 4), несобственный интеграл $\int\limits_1^{+\infty} \frac{dx}{x^\alpha}$ при $\alpha>1$ сходится, а при $\alpha\leqslant 1$ расходится. Следовательно, данный ряд сходится при $\alpha>1$ и расходится при $\alpha\leqslant 1$.

Заметим, что при $\alpha\leqslant 0$ такие ряды также расходятся, так как их общий член не стремится к нулю при $n\to\infty$, т. е. нарушается необходимое условие сходимости ряда (см. теорему 14.4).

В частности, при $\alpha=2$ получим сходящийся ряд $\sum\limits_{n=1}^\infty \frac{1}{n^2}$; при $\alpha=1$ — расходящийся гармонический ряд $\sum\limits_{n=1}^\infty \frac{1}{n}$; при $\alpha=1/2$ — расходящийся ряд $\sum\limits_{n=1}^\infty \frac{1}{\sqrt{n}}$ и т. д.

§ 3. Знакочередующиеся ряды

До сих пор мы изучали ряды с неотрицательными членами. Что касается рядов с неположительными членами, то они отличаются от соответствующих рядов с неотрицательными членами только множителем -1, вследствие чего ведут себя относительно сходимости аналогичным образом.

Теперь перейдем к рассмотрению так называемых *знакочередующихся* рядов, члены которых имеют чередующиеся знаки. Для удобства будем считать, что первый член такого ряда положителен. Тогда знакочередующийся ряд можно записать в виде

$$a_1-a_2+a_3-a_4+\ldots+(-1)^{n+1}a_n+\ldots, \tag{1}$$

где $a_n>0$.

Для знакочередующихся рядов имеет место очень простой достаточный признак сходимости.

Признак Лейбница.

Теорема 14.9. *Если абсолютные величины членов знакочередующегося ряда (1) монотонно убывают: $a_1>a_2>a_3>\ldots$ и общий член ряда стремится к нулю: $\lim\limits_{n\to\infty} a_n=0$, то ряд сходится.*

Доказательство. Пусть дан ряд (1) и пусть $a_n>a_{n+1}$ и $a_n\to 0$ при $n\to\infty$.

Рассмотрим частичную сумму ряда с четным числом членов:

$$S_{2n}=a_1-a_2+a_3-a_4+\ldots+a_{2n-1}-a_{2n}=$$
$$=(a_1-a_2)+(a_3-a_4)+\ldots+(a_{2n-1}-a_{2n}).$$

Так как все разности в скобках в силу первого условия положительны, то последовательность частичных сумм $\{S_{2n}\}$ — возрастающая при увеличении n. Докажем, что она ограничена. Для этого представим S_{2n} следующим образом:

$$S_{2n}=a_1-[(a_2-a_3)+(a_4-a_5)+\ldots+(a_{2n-2}-a_{2n-1})+a_{2n}].$$

Отсюда следует, что $S_{2n}<a_1$ для любого n, т. е. $\{S_{2n}\}$ ограничена.

Итак, последовательность $\{S_{2n}\}$ монотонно возрастающая и ограниченная, следовательно, имеет предел:

$$\lim_{n\to\infty} S_{2n}=S.$$

Покажем теперь, что и последовательность частичных сумм нечетного числа членов сходится к тому же пределу S. Действительно, $S_{2n+1}=S_{2n}+a_{2n+1}$. Переходя в этом равенстве к пределу при $n\to\infty$ и используя второе условие ($a_n\to 0$ при $n\to\infty$), получаем

$$\lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} (S_{2n} + a_{2n+1}) =$$
$$= \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} a_{2n+1} = S + 0 = S.$$

Таким образом, последовательность частичных сумм $\{S_n\}$ ряда (1) сходится к пределу S. Это означает, что ряд (1) сходится ■

Рассмотрим

Пример. Ряд

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + (-1)^{n+1}\frac{1}{n} + \ldots$$
$$\ldots = \sum_{n=1}^{\infty} (-1)^{n+1}\frac{1}{n}$$

сходится, так как удовлетворяет условиям признака Лейбница: 1) $1 > \frac{1}{2} > \frac{1}{3} > \ldots$; 2) $\lim\limits_{n\to\infty} \frac{1}{n} = 0$. Заметим, что этот ряд отличается от гармонического ряда только знаками четных членов.

§ 4. Абсолютная и условная сходимость рядов

Нам осталось рассмотреть ряды с членами произвольных знаков. Такие ряды называются *знакопеременными* рядами.

Возьмем какой-нибудь знакопеременный ряд

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n, \qquad (1)$$

где числа $a_1, a_2, a_3, \ldots, a_n, \ldots$ могут быть как положительными, так и отрицательными, причем расположение положительных и отрицательных членов в ряде совершенно произвольно. Одновременно рассмотрим ряд, составленный из абсолютных величин членов ряда (1):

$$|a_1|+|a_2|+|a_3|+\ldots+|a_n|+\ldots=\sum_{n=1}^{\infty}|a_n|. \qquad (2)$$

Для знакопеременных рядов имеет место следующий признак сходимости.

Теорема 14.10. *Если ряд (2) сходится, то сходится и ряд (1).*

Доказательство. Пусть ряд (2) сходится. Обозначим через S_n частичную сумму ряда (1), а через σ_n — частичную сумму ряда (2):

$$S_n=a_1+a_2+a_3+\ldots+a_n;\ \sigma_n=|a_1|+|a_2|+|a_3|+ +\ldots+|a_n|.$$

Так как ряд (2) сходится, то последовательность его частичных сумм $\{\sigma_n\}$ имеет предел $\lim\limits_{n\to\infty}\sigma_n=\sigma$, при этом для любого n имеет место неравенство

$$\sigma_n\leqslant\sigma, \qquad (3)$$

поскольку члены ряда (2) неотрицательны.

Обозначим через S'_n сумму положительных членов, а через S''_n — сумму модулей отрицательных членов, содержащихся в сумме S_n. Тогда

$$S_n=S'_n-S''_n \qquad (4)$$

и

$$\sigma_n=S'_n+S''_n. \qquad (5)$$

Из равенства (5) следует, что при возрастании n последовательности $\{S'_n\}$ и $\{S''_n\}$ монотонно возрастают, а из (3) — что они являются ограниченными: $S'_n\leqslant\sigma_n\leqslant\sigma$ и $S''_n\leqslant\sigma_n\leqslant\sigma$. Следовательно, существуют

$$\lim_{n\to\infty}S'_n=S' \text{ и } \lim_{n\to\infty}S''=S''.$$

Но в таком случае, в силу равенства (4), последовательность частичных сумм ряда (1) имеет предел

$$\lim_{n\to\infty}S_n=\lim_{n\to\infty}(S'_n-S''_n)=\lim_{n\to\infty}S'_n-\lim_{n\to\infty}S''_n=S'-S''.$$

Это значит, что ряд (1) сходится ■

Рассмотрим примеры.

Пример 1. Ряд

$$1-\frac{1}{2^2}-\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}-\frac{1}{6^2}-\frac{1}{7^2}+\dots$$

на основании доказанного признака сходится, так как сходится ряд, составленный из абсолютных величин членов ряда:

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\dots$$

(см. пример 6 § 2).

Рассмотренный признак сходимости знакопеременного ряда является достаточным, но не необходимым, так как существуют знакопеременные ряды, которые сходятся, а ряды, составленные из абсолютных величин их членов, расходятся. Так, например, ряд $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ по признаку Лейбница сходится (см. пример из § 3), а ряд $\sum_{n=1}^{\infty}\frac{1}{n}$, составленный из абсолютных величин его членов, расходится (гармонический ряд).

Поэтому все сходящиеся ряды делятся на *абсолютно* сходящиеся и *условно* сходящиеся.

К абсолютно сходящимся рядам относятся сходящиеся ряды, для которых ряды, составленные из абсолютных величин их членов, также сходятся.

К условно сходящимся рядам относятся сходящиеся ряды, для которых ряды, составленные из абсолютных величин их членов, расходятся.

Пример 2. Ряд

$$1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{1}{16}-\frac{1}{32}+\dots$$

есть ряд абсолютно сходящийся, так как ряд, составленный из абсолютных величин:

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+\dots,$$

тоже сходится (оба ряда — геометрические прогрессии со знаменателями, равными $-{}^1/_2$ и ${}^1/_2$).

Пример 3. Ряд

$$1-\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}-\frac{1}{\sqrt{4}}+\ldots$$

условно сходящийся, так как он сходится по признаку Лейбница, а ряд, составленный из абсолютных величин:

$$1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{4}}+\ldots,$$

расходится (см. пример 6, § 2).

Следует заметить, что деление сходящихся рядов на абсолютно и условно сходящиеся весьма существенно. Дело в том, что абсолютно сходящиеся ряды обладают рядом важных свойств, тогда как условно сходящиеся ряды некоторыми из этих свойств не обладают. Например, для условно сходящихся рядов сумма ряда не равна сумме положительных и сумме отрицательных членов ряда, как это имеет место для абсолютно сходящихся рядов, что было показано по ходу доказательства теоремы 14.10.

§ 5. Степенные ряды

1. Определение и общие замечания. *Ряд вида*

$$a_0+a_1x+a_2x^2+a_3x^3+\cdots+a_nx^n+\ldots=\sum_{n=0}^{\infty}a_nx^n \qquad (1)$$

называется степенным рядом.

Постоянные числа a_0, a_1, a_2, a_3, ..., a_n... называются *коэффициентами* ряда.

Давая x определенные числовые значения, мы получаем различные числовые ряды, которые могут оказаться сходящимися или расходящимися. Совокупность тех значений x, при которых ряд (1) сходится, называется *областью* его сходимости.

Очевидно, что частичная сумма степенного ряда $S_n(x)=a_0+a_1x+\ldots+a_nx^n$ является функцией от переменной x. Кроме этого, в области сходимости ряда его сумма S также является некоторой функцией от переменной x, определенной в области сходимости ряда. Поэтому пишут

$$S=S(x)=\sum_{n=0}^{\infty} a_n x^n \quad \left(\text{или } f(x)=\sum_{n=0}^{\infty} a_n x^n\right).$$

2. Интервал сходимости степенного ряда. Докажем теорему, имеющую важное значение в теории степенных рядов и касающуюся области сходимости степенного ряда.

Теорема 14.11 (Теорема Абеля)*. *1) Если степенной ряд (1) сходится при $x=x_0$ ($x_0 \neq 0$), то он сходится, и притом абсолютно, для всех x, удовлетворяющих условию $|x|<|x_0|$; 2) если ряд расходится при $x=x_1$, то он расходится для всех x, удовлетворяющих условию $|x|>|x_1|$.*

Доказательство. 1) Так как, по условию, числовой ряд $\sum_{n=0}^{\infty} a_n x_0^n$ сходится, то его общий член $a_n x_0^n \to 0$ при $n \to \infty$, а это значит, что последовательность $\{a_n x_0^n\}$ ограничена, т. е. существует число $M>0$ такое, что

$$|a_n x_0^n|<M, \quad n=0, 1, 2, \ldots \tag{2}$$

Перепишем ряд (1) в виде

$$a_0+a_1x_0\left(\frac{x}{x_0}\right)+a_2x_0^2\left(\frac{x}{x_0}\right)^2+\ldots+ a_n x_0^n\left(\frac{x}{x_0}\right)^n+\ldots \tag{3}$$

и рассмотрим ряд из абсолютных величин его членов:

$$|a_0|+|a_1x_0|\left|\frac{x}{x_0}\right|+|a_2x_0^2|\left|\frac{x}{x_0}\right|^2+\ldots \ldots+|a_n x_0^n|\left|\frac{x}{x_0}\right|^n+\ldots \tag{4}$$

Члены ряда (4), в силу неравенства (2), меньше соответствующих членов ряда

$$M+M\left|\frac{x}{x_0}\right|+M\left|\frac{x}{x_0}\right|^2+\ldots+M\left|\frac{x}{x_0}\right|^n+\ldots \tag{5}$$

---

* Абель Нильс Хенрик (1802—1829) — норвежский математик.

При $|x|<|x_0|$ ряд (5) представляет геометрическую прогрессию со знаменателем $q=\left|\frac{x}{x_0}\right|<1$ и, следовательно, сходится. Так как члены ряда (4) меньше соответствующих членов ряда (5), то, по признаку сравнения, ряд (4) тоже сходится, а это значит, что ряд (1) при $|x|<|x_0|$ сходится абсолютно.

2) Теперь докажем вторую часть теоремы. По условию в некоторой точке x_1 ряд (1) расходится. Требуется показать, что он расходится для всех x, удовлетворяющих условию $|x|>|x_1|$. Предположим обратное, т. е. допустим, что при некотором значении x таком, что $|x|>|x_1|$, ряд (1) сходится. Тогда по только что доказанной первой части теоремы ряд (1) должен сходиться и в точке x_1, так как $|x_1|<|x|$. Но это противоречит условию, что в точке x_1 ряд расходится ■

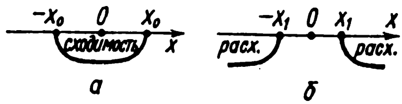


Рис. 94

Теорема Абеля геометрически утверждает, что если x_0 — точка сходимости, то во всех точках, расположенных на интервале $(-|x_0|, |x_0|)$ (рис. 94, *а*), ряд сходится абсолютно, а если x_1 — точка расходимости, то во всех точках, расположенных вне интервала $(-|x_1|, |x_1|)$ (рис. 94, *б*), ряд расходится.

Отсюда вытекает следующая

Теорема 14.12. *Если ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится не при всех значениях x, то существует число $R>0$ такое, что ряд абсолютно сходится при $|x|<R$ и расходится при $|x|>R$.*

Доказательство. Обозначим через X_0 множество точек x_0, в которых ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится. Покажем, что множество X_0 ограничено. Действительно, если возьмем точку x_1, в которой ряд расходится (по условию такие точки существуют), то по теореме Абеля для любого x_0 из X_0 выполняется неравенство $|x_0|<|x_1|$. Известно, что у ограниченного сверху множества существует точная верхняя грань. Обозначим ее через $R(0<<R<+\infty):R=\sup X_0$.

Возьмем теперь любое x, для которого $|x|<R$. По свойству точной верхней грани найдется $x_0\in X_0$ такое, что $|x|<|x_0|<R$, а это по теореме Абеля означает абсолютную сходимость ряда при данном x.

Возьмем теперь любое x, для которого $|x|>R$. Такое $x\notin X_0$. Следовательно, при этом x ряд расходится ■

Таким образом, разрешен вопрос об области сходимости степенного ряда: *интервал* $(-R, R)$ *называется интервалом сходимости степенного ряда.* Число R называется *радиусом* сходимости степенного ряда.

Отметим, что у некоторых рядов интервал сходимости охватывает всю числовую прямую (в этом случае пишут $R=+\infty$), у других вырождается в одну точку $(R=0)$.

Таким образом, всякий степенной ряд имеет свой радиус сходимости R.

При $x=\pm R$ ряд может либо сходиться, либо расходиться. Этот вопрос решается индивидуально для каждого конкретного ряда.

Укажем способ определения радиуса сходимости степенного ряда. Этому посвящается следующая

Теорема 14.13. *Пусть дан ряд* $\sum\limits_{n=0}^{\infty} a_n x^n$. *Тогда, если существует конечный предел* $\lim\limits_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|\neq 0$, *то радиусом сходимости ряда будет число*

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|.$$

Доказательство. Рассмотрим ряд $\sum\limits_{n=0}^{\infty}|a_n x^n|$.

По условию существует $\lim\limits_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|\neq 0$. Обозначим его через $\frac{1}{R}$. Тогда

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right|=|x|\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=\frac{|x|}{R}.$$

При каждом конкретном значении x степенной ряд становится числовым. Поэтому по признаку Даламбера ряд $\sum_{n=0}^{\infty} |a_n x^n|$ сходится, если $\frac{|x|}{R} < 1$, т. е. $|x| < R$.

Следовательно, по теореме 14.10 о сходимости знакопеременных рядов, ряд $\sum_{n=0}^{\infty} a_n x^n$ также сходится при $|x| < R$, причем абсолютно. При $|x| > R$ ряд $\sum_{n=0}^{\infty} a_n x^n$ расходится, так как $\lim_{n\to\infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = \frac{|x|}{R} > 1$ и, следовательно, общий член ряда $a_n x^n$ не стремится к нулю при $n \to \infty$.

Таким образом, данный ряд сходится внутри интервала $(-R, R)$ и расходится вне его, т. е. радиус сходимости равен $R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|$ ■

Пример 1. Рассмотрим ряд $\sum_{n=1}^{\infty} \frac{x^n}{n}$. Здесь

$$a_n = \frac{1}{n} \text{ и } a_{n+1} = \frac{1}{n+1}.$$

Поэтому

$$R = \lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = 1.$$

Следовательно, данный ряд сходится на интервале $(-1, 1)$. Исследуем поведение ряда на концах интервала сходимости, т. е. в точках $x = \pm 1$. При $x = 1$ получаем гармонический ряд

$$\sum_{n=1}^{\infty} \frac{1}{n}, \text{ а при } x = -1 \text{ ряд } \sum_{n=1}^{\infty} (-1)^n \frac{1}{n},$$

который сходится в силу признака Лейбница. Таким образом, данный ряд сходится в любой точке полуинтервала $[-1, 1)$ и расходится вне его.

Пример 2. Ряд $\sum_{n=1}^{\infty} n!\,x^n$ расходится на всей числовой прямой, кроме точки $x=0$, так как его радиус сходимости

$$R=\lim_{n\to\infty}\frac{a_n}{a_{n+1}}=\lim_{n\to\infty}\frac{n!}{(n+1)!}=\lim_{n\to\infty}\frac{1}{1+n}=0.$$

Пример 3. Ряд $\sum_{n=1}^{\infty}\frac{x^n}{n!}$ абсолютно сходится на всей числовой прямой, так как его радиус сходимости

$$R=\lim_{n\to\infty}\frac{a_n}{a_{n+1}}=\lim_{n\to\infty}\frac{(n+1)!}{n!}=\lim_{n\to\infty}(n+1)=+\infty.$$

3. Свойства степенных рядов. Пусть функция $f(x)$ является суммой степенного ряда

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n+\ldots, \tag{6}$$

интервал сходимости которого $(-R, R)$.

В этом случае говорят, что функция $f(x)$ разлагается в ряд по степеням x.

Имеют место две теоремы о свойствах степенных рядов, которые приведем без доказательства.

Теорема 14.14. *Если функция $f(x)$ на интервале $(-R, R)$ является суммой ряда (6), то она дифференцируема на интервале $(-R, R)$ и ее производная $f'(x)$ находится почленным дифференцированием ряда (6), т. е.*

$$f'(x)=(a_0+a_1x+a_2x^2+\ldots+a_nx^n+\ldots)'=a_1+2a_2x+$$
$$+3a_3x^2+\ldots+na_nx^{n-1}+\ldots$$

Аналогично могут быть вычислены производные любого порядка. При этом все они имеют тот же интервал сходимости, что и ряд (6).

Теорема 14.15. *Если функция $f(x)$ на интервале $(-R, R)$ является суммой ряда (6), то она интегрируема в интервале $(-R, R)$ и интеграл от нее находится почленным интегрированием ряда (6), т. е. если $x_1, x_2\in$ $\in(-R, R)$, то*

$$\int_{x_1}^{x_2} f(x)\,dx = \int_{x_1}^{x_2} (a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots)\,dx =$$

$$= \int_{x_1}^{x_2} a_0\,dx + \int_{x_1}^{x_2} a_1x\,dx + \ldots + \int_{x_1}^{x_2} a_nx^n\,dx + \ldots$$

При этом ряд, полученный после интегрирования ряда (6) по $[0, x]$, $|x|<R$, *имеет тот же интервал сходимости, что и ряд (6).*

Сформулированные теоремы о дифференцировании и интегрировании степенных рядов имеют важное значение. Ими будем неоднократно пользоваться.

Представляет интерес интегрирование степенного ряда по отрезку $[0, x]$, где $|x|<R$, так как результатом интегрирования каждого члена ряда будет не число, а функция от переменного верхнего предела x:

$$\int_0^x f(x)\,dx = a_0x + \frac{a_1x^2}{2} + \frac{a_2x^3}{3} + \ldots + \frac{a_nx^{n+1}}{n+1} + \ldots$$

В этом случае получаем снова степенной ряд, сумма которого, по теореме 14.15, равна интегралу по $[0, x]$ от суммы $f(x)$ ряда (6).

Отметим, что в ряде случаев рассматриваются степенные ряды более общего вида

$$a_0 + a_1(x-a) + a_2(x-a)^2 + \ldots + a_n(x-a)^n + \ldots =$$

$$= \sum_{n=0}^{\infty} a_n(x-a)^n. \qquad (7)$$

Ряд вида (7) приводится к виду (1) заменой переменной $x-a=t$.

Если функция $f(x)$ является суммой ряда (7), то в этом случае говорят, что функция $f(x)$ разлагается в ряд по степеням $(x-a)$.

Все изложенное полностью переносится и на ряды вида (7). Поэтому для простоты последующие рассуждения будут проводиться для рядов вида (1).

4. Разложение функций в степенные ряды. Пусть функция $f(x)$ является суммой степенного ряда (1), интервал сходимости которого $(-R, R)$. Тогда функция $f(x)$ может быть разложена в степенной ряд, причем, как показывает следующая теорема, это разложение единственно.

Теорема 14.16. *Если функция $f(x)$ на интервале $(-R, R)$ разлагается в степенной ряд*

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n+\ldots, \tag{8}$$

то это разложение единственно.

Доказательство. По условию ряд (8) сходится и функция $f(x)$ — его сумма. Следовательно, по теореме 14.14 его можно почленно дифференцировать любое число раз.

Дифференцируя, получим

$$f'(x)=1\cdot a_1+2a_2x+3a_3x^2+4a_4x^3+\ldots+na_nx^{n-1}+\ldots,$$

$$f''(x)=1\cdot 2\cdot a_2+2\cdot 3a_3x+3\cdot 4a_4x^2+4\cdot 5a_5x^3+\ldots+$$
$$+n(n-1)a_nx^{n-2}+\ldots,$$

$$f'''(x)=1\cdot 2\cdot 3a_3+2\cdot 3\cdot 4a_4x+\ldots+n(n-1)(n-2)a_nx^{n-3}+\ldots,$$

. . . .

$$f^{(n)}(x)=n!a_n+2\cdot 3\cdot 4\ldots(n+1)a_{n+1}x+\ldots,$$

. . . .

Полагая в полученных равенствах и в равенстве (8) $x=0$, получим

$$f(0)=a_0,\ f'(0)=1\cdot a_1,\ f''(0)=2!\,a_2,$$
$$f'''(0)=3!\,a_3,\ \ldots,\ f^{(n)}(0)=n!\,a_n,\ \ldots,$$

откуда находим

$$a_0=f(0),\quad a_1=\frac{f'(0)}{1!},\ a_2=\frac{f''(0)}{2!},$$

$$a_3=\frac{f'''(0)}{3!},\ \ldots,\ a_n=\frac{f^{(n)}(0)}{n!},\ \ldots \tag{9}$$

Из (9) следует, что все коэффициенты ряда (8) определяются единственным образом формулами (9), что и доказывает теорему ■

Подставляя полученные выражения коэффициентов в равенство (8), получим

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

Таким образом, если функция $f(x)$ разлагается в степенной ряд, то этот ряд имеет вид

$$f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+\ldots \tag{10}$$

Ряд (10) называется *рядом Маклорена* для функции $f(x)$.

Пусть теперь $f(x)$ — произвольная бесконечно дифференцируемая функция. Для нее можно составить ряд (10). Как всякий степенной ряд, он будет иметь интервал сходимости и сумму. Возникает вопрос, при каких условиях сумма ряда (10) будет совпадать с функцией $f(x)$. Ответ на поставленный вопрос может быть получен с помощью формулы Маклорена. В § 3, гл. VI, ч. I мы показали, что для любых бесконечно дифференцируемых функций справедлива формула Маклорена

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

где

$$R_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1},\quad \xi=\theta x,\ 0<\theta<1. \tag{11}$$

Если обозначим через $S_n(x)$ частичную сумму ряда Маклорена, то формулу Маклорена можно записать так:

$$f(x)=S_n(x)+R_n(x). \tag{12}$$

Теорема 14.17. *Для того чтобы ряд Маклорена (10) сходился на $(-R, R)$ и имел своей суммой функцию $f(x)$, необходимо и достаточно, чтобы остаточный член формулы Маклорена (11) стремился к нулю при $n\to\infty$, т. е. $\lim\limits_{n\to\infty} R_n(x)=0$ для любого $x\in(-R, R)$.*

Доказательство. *Необходимость.* Пусть функция $f(x)$ есть сумма ряда Маклорена на $(-R, R)$, т. е. $\lim\limits_{n\to\infty} S_n(x)=f(x)$. Тогда из равенства (12) следует, что $\lim\limits_{n\to\infty} R_n(x)=0$ для любого $x\in(-R, R)$.

Достаточность. Пусть $\lim\limits_{n\to\infty} R_n(x)=0$ для любого $x\in(-R, R)$. Тогда из того же равенства (12) следует, что $\lim\limits_{n\to\infty}[f(x)-S_n(x)]=0$, т. е. $\lim\limits_{n\to\infty} S_n(x)=f(x)$. Это и означает, что ряд Маклорена (10) сходится на $(-R, R)$ и сумма его равна $f(x)$ ■

Из теоремы вытекает, что вопрос о разложении функции в ряд Маклорена сводится к исследованию поведения остаточного члена $R_n(x)$ при $n\to\infty$.

Рассмотрим разложения в ряд Маклорена некоторых элементарных функций.

Разложение функции $f(x)=e^x$. Имеем $f'(x)=f''(x)=\ldots=f^{(n)}(x)=e^x$, откуда при $x=0$ получим $f(0)=f'(0)=f''(0)=\ldots=f^{(n)}(0)=1$. По формуле (10) для функции $f(x)=e^x$ составим ряд Маклорена

$$1+x/1!+x^2/2!+\ldots+x^n/n!+\ldots \tag{13}$$

Найдем интервал сходимости ряда (13):

$$R=\lim_{n\to\infty}\frac{a_n}{a_{n+1}}=\lim_{n\to\infty}\frac{n!\,(n+1)}{n!}=+\infty.$$

Следовательно, ряд абсолютно сходится на всей числовой прямой.

Теперь покажем, что ряд (13) имеет своей суммой функцию e^x.

Отметим, что в силу необходимого условия сходимости ряда для любого x справедливо равенство

$$\lim_{n\to\infty}\frac{x^n}{n!}=0. \tag{14}$$

Так как $f^{(n+1)}(\xi)=e^{\xi}$, то

$$R_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}=\frac{e^{\xi}}{(n+1)!}x^{n+1},$$

где $\xi=\theta x$, $0<\theta<1$. Отсюда, учитывая, что $e^{\xi}<e^{|x|}$, получим

$$R_n(x)=\frac{e^{\xi}}{(n+1)!}x^{n+1}<\frac{e^{|x|}}{(n+1)!}x^{n+1}.$$

Так как в силу (14) $\lim\limits_{n\to\infty}\frac{x^n}{n!}=0$, то и $\lim\limits_{n\to\infty}\frac{x^{n+1}}{(n+1)!}=0$.

Поэтому, переходя к пределу в последнем неравенстве при $n\to\infty$, получим, что $\lim\limits_{n\to\infty} R_n(x)=0$ при любом x, и, следовательно, функция e^x является суммой ряда (13).

Таким образом, при любом x имеет место разложение

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots$$

Разложение функции $f(x)=\sin x$. Имеем

$$f'(x) = \cos x = \sin\left(x + \frac{\pi}{2}\right),$$

$$f''(x) = -\sin x = \sin\left(x + 2\frac{\pi}{2}\right), \ldots$$

$$\ldots, f^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right)$$

(см. п. 2, § 10, гл. V, ч. I), откуда, полагая $x=0$, получим $f(0)=0$, $f'(0)=1$, $f''(0)=0$, $f'''(0)=-1$, $f^{(4)}(0)=0$, ... По формуле (10) для функции $\sin x$ составляем ряд Маклорена

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{(-1)^{n-1}x^{2n-1}}{(2n-1)!} + \ldots$$

Легко проверить, что полученный ряд абсолютно сходится на всей числовой прямой. Исследуем остаточный член

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1} = \frac{\sin\left[\xi + (n+1)\frac{\pi}{2}\right]}{(n+1)!}x^{n+1},$$

где $\xi=\theta x$, $0<\theta<1$. Так как

$$\left|\sin\left[\xi + (n+1)\frac{\pi}{2}\right]\right| \leqslant 1, \text{ то } |R_n(x)| \leqslant \frac{|x|^{n+1}}{(n+1)!}.$$

В силу (14) $\lim\limits_{n\to\infty}\frac{|x|^{n+1}}{(n+1)!}=0$. Следовательно, $\lim\limits_{n\to\infty} R_n(x)=0$

при любом x. А это означает, что функция $\sin x$ является суммой ряда и имеет место разложение

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{(-1)^{n-1} x^{2n-1}}{(2n-1)!} + \ldots$$

Разложение функции $f(x) = \cos x$. Аналогично предыдущему можно получить разложение функции $\cos x$ в ряд Маклорена, справедливое при любом x. Однако еще проще разложение $\cos x$ получается путем почленного дифференцирования ряда для $\sin x$:

$$(\sin x)' = (x)' - \left(\frac{x^3}{3!}\right)' + \left(\frac{x^5}{5!}\right)' - \left(\frac{x^7}{7!}\right)' + \ldots$$

$$\ldots + \left[\frac{(-1)^{n-1} x^{2n-1}}{(2n-1)!}\right]' + \ldots,$$

откуда получаем разложение, справедливое при любом x:

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots$$

Кроме рассмотренных функций e^x, $\sin x$, $\cos x$ в ряд Маклорена могут быть разложены и многие другие функции. Вместо ряда Маклорена можно было бы рассмотреть более общий ряд Тейлора по степеням $(x-a)$, где $a \neq 0$, т. е. ряд вида

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots$$

$$\ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots$$

Но мы делать этого не будем, так как все изложенное полностью переносится и на эти ряды.

При разложении функции $\cos x$ в ряд Маклорена было использовано свойство почленной дифференцируемости степенных рядов. Аналогичным образом можно использовать и другое свойство степенных рядов — их почленную интегрируемость. В качестве примеров получим разложение в степенные ряды функций $\ln(1+x)$ и $\operatorname{arc\,tg} x$ с помощью почленного интегрирования.

Рассмотрим ряд $1+x+x^2+x^3+\ldots+x^n+\ldots$ Очевидно, данный ряд является геометрической прогрессией с первым членом, равным единице, и со знаменателем $q = x$. Как известно, при $|x|<1$ данный ряд сходится и его сумма равна $\frac{1}{1-x}$. Следовательно,

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots + x^n + \ldots \qquad (15)$$

Равенство (15) является разложением функции

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

в степенной ряд.

Подставляя в равенство (15) $-t$ вместо x, получим равенство

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

справедливое при $|t| < 1$. Проинтегрируем этот степенной ряд почленно в пределах от 0 до x, $|x| < 1$:

$$\int_0^x \frac{dt}{1+t} = \ln(1+t)\Big|_0^x = \ln(1+x) =$$

$$= \int_0^x (1 - t + t^2 - t^3 + \ldots + (-1)^n t^n + \ldots)\, dt =$$

$$= \int_0^x dt - \int_0^x t\,dt + \int_0^x t^2\,dt - \int_0^x t^3\,dt + \ldots + (-1)^n \int_0^x t^n dt + \ldots =$$

$$= t\Big|_0^x - \frac{t^2}{2}\Big|_0^x + \frac{t^3}{3}\Big|_0^x - \frac{t^4}{4}\Big|_0^x + \ldots$$

$$\ldots + (-1)^n \frac{t^{n+1}}{n+1}\Big|_0^x + \ldots =$$

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

Отсюда

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

$$\ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots \qquad (16)$$

Равенство (16) есть разложение функции $\ln(1+x)$ в степенной ряд. Оно справедливо при $|x|<1$. Впрочем, можно доказать, что (16) верно и для $x=1$.

Действительно, при $x=-1$ равенство (16) теряет смысл, а при $x=1$ сохраняет его, так как функция $\ln(1+x)$ обращается в число $\ln 2$, а ряд — в сходящийся по признаку Лейбница числовой ряд

$$1-\frac{1}{2}+\frac{1}{3}-\ldots+(-1)^{n-1}\frac{1}{n}+\ldots \tag{17}$$

Осталось проверить справедливость равенства

$$\ln 2=1-\frac{1}{2}+\frac{1}{3}-\ldots+(-1)^{n-1}\frac{1}{n}+\ldots \tag{18}$$

Для этого проинтегрируем от 0 до 1 следующее выражение:

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

полученное в результате деления единицы на $1+t$. Имеем

$$\int_0^1\frac{dt}{1+t}=\ln(1+t)\Big|_0^1=\ln 2=$$

$$=\int_0^1\left(1-t+t^2-t^3+\ldots+(-1)^{n-1}t^{n-1}+(-1)^n\frac{t^n}{1+t}\right)dt=$$

$$=\int_0^1 dt-\int_0^1 t\,dt+\int_0^1 t^2\,dt-\int_0^1 t^3\,dt+\ldots$$

$$\ldots+(-1)^{n-1}\int_0^1 t^{n-1}\,dt+(-1)^n\int_0^1\frac{t^n}{1+t}\,dt=$$

$$=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

$$\ldots + (-1)^{n-1}\frac{1}{n} + (-1)^n \int_0^1 \frac{t^n}{1+t}\, dt.$$

Откуда

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

$$\ldots + (-1)^{n-1}\frac{1}{n} + (-1)^n \int_0^1 \frac{t^n\, dt}{1+t}. \qquad (19)$$

В равенстве (19) справа сумма первых n слагаемых является частичной суммой S_n ряда (17). Запишем (19) в виде

$$\ln 2 - S_n = (-1)^n \int_0^1 \frac{t^n\, dt}{1+t}. \qquad (20)$$

Так как при $0 \leqslant t \leqslant 1$ $\frac{t^n}{1+t} \leqslant t^n$, то

$$\left| (-1)^n \int_0^1 \frac{t^n\, dt}{1+t} \right| = \int_0^1 \frac{t^n\, dt}{1+t} \leqslant$$

$$\leqslant \int_0^1 t^n\, dt = \frac{t^{n+1}}{n+1}\bigg|_0^1 = \frac{1}{n+1} \to 0 \quad \text{при } n \to \infty.$$

Отсюда заключаем, что интеграл в правой части (20) стремится к нулю при $n \to \infty$, а следовательно, $\lim\limits_{n\to\infty} S_n = \ln 2$, что и означает справедливость равенства (18).

Получим теперь разложение функции $\operatorname{arctg} x$. Подставляя в (15) $-t^2$ вместо x и интегрируя по t от 0 до x, будем иметь

$$\operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots \qquad (21)$$

Равенство (21) справедливо при $|x| < 1$. Однако можно показать аналогично предыдущему, что оно верно и для $x = \pm 1$.

В заключение заметим, что степенные ряды имеют самые разнообразные приложения. С их помощью, например, с любой заданной точностью вычисляют значение функций, в частности вычисляют значения таких чисел, как π и e. С помощью степенных рядов вычисляют приближенные значения определенных интегралов, которые или не выражаются через элементарные функции, или слишком сложны для вычислений. Так, например, интеграл $\int\limits_0^a \frac{\sin x}{x}\,dx$ не берется в элементарных функциях, так как первообразная функция $\frac{\sin x}{x}$ не является элементарной функцией. В то же время эта первообразная легко выражается в виде суммы степенного ряда. Действительно, так как

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots,$$

то, умножая этот ряд на $\frac{1}{x}$, получим

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \ldots,$$

причем последний ряд сходится при любом x. Интегрируя его почленно от 0 до a, получаем

$$\int\limits_0^a \frac{\sin x}{x}\,dx = a - \frac{a^3}{3!\cdot 3} + \frac{a^5}{5!\cdot 5} - \frac{a^7}{7!\cdot 7} + \ldots$$

С помощью этого равенства можно при любом a с любой степенью точности вычислить данный интеграл.

Наконец, значительное место отводится степенным рядам в приближенных методах решений дифференциальных уравнений.

§ 6. Комплексные ряды

1. Краткие сведения о комплексных числах. Определение комплексного числа и его геометрический смысл. *Комплексным числом z называется упорядоченная пара* *вещественных чисел $(x;\ y)$, т. е. $z=(x;\ y)$. При этом x называется вещественной частью, а y — мнимой частью комплексного числа.*

Геометрически комплексное число $z=(x;\ y)$ изображается на плоскости Oxy или *точкой* M с координатами $(x;\ y)$, или *вектором* $\overline{OM}$, проекции которого на оси Ox и Oy соответственно равны x и y (рис. 95). Плоскость Oxy в этом случае называется *условно комплексной* плоскостью.

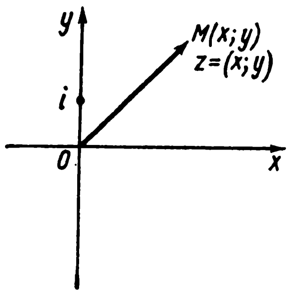


Рис. 95

Комплексное число вида $(x;\ 0)$ отождествляется с вещественным числом x, т. е. $(x;\ 0)=x$. Это позволяет рассматривать множество всех вещественных чисел как подмножество множества комплексных чисел. На комплексной плоскости вещественные числа изображаются точками на оси Ox, которая называется вещественной осью.

Комплексное число $(x;\ y)$ при $y\neq 0$ называется мнимым. Мнимое число $(0;\ y)$ называется *чисто мнимым* и символически обозначается $z=iy$. Чисто мнимое число $(0;\ 1)$ называется *мнимой единицей* и обозначается буквой i, т. е. $i=(0;\ 1)$. Чисто мнимые числа на комплексной плоскости изображаются точками на оси Oy, которая называется мнимой осью.

Два комплексных числа $z_1=(x_1;\ y_1)$ и $z_2=(x_2;\ y_2)$ называются *равными*, если $x_1=x_2$, $y_1=y_2$. Отсюда векторы, изображающие равные комплексные числа, также равны между собой. Комплексное число $0=(0;\ 0)$ называется *нулем*.

Действия над комплексными числами. Определим алгебраические действия над комплексными числами.

Суммой комплексных чисел $z_1=(x_1;\ y_1)$ и $z_2=(x_2;\ y_2)$ называется комплексное число z вида

$$z=(x_1+x_2;\ y_1+y_2).$$

Произведением комплексных чисел $z_1=(x_1;\ y_1)$ и $z_2=(x_2;\ y_2)$ называется комплексное число z вида

$$z=(x_1x_2-y_1y_2;\ x_1y_2+x_2y_1).$$

Действие *вычитания* комплексных чисел определяется как действие, обратное сложению, т. е. число z называется разностью чисел z_1 и z_2, если $z+z_2=z_1$. Из этого определения вытекает, что комплексное число z вида

$$z=(x_1-x_2;\ y_1-y_2)$$

является разностью комплексных чисел $z_1=(x_1;\ y_1)$ и $z_2=(x_2;\ y_2)$.

Действие *деления* комплексных чисел определяется как действие, обратное умножению, т. е. число z называется частным чисел z_1 и $z_2\neq 0$, если $z\cdot z_2=z_1$. Из этого определения вытекает, что комплексное число z вида

$$z=\left(\frac{x_1x_2+y_1y_2}{x_2^2+y_2^2};\ \frac{x_2z_1-x_1y_2}{x_2^2+y_2^2}\right)$$

является частным комплексных чисел $z_1=(x_1;\ y_1)$ и $z_2=(x_2;\ y_2)$.

В действиях с комплексными числами особую роль играет мнимая единица $i=(0;\ 1)$. Умножая ее саму на себя (т. е. возводя ее в квадрат), получим в силу определения произведения комплексных чисел: $(0;\ 1)\cdot(0;\ 1)=(-1;\ 0)=-1$, т. е. $i^2=-1$. Заметив это, мы можем любое комплексное число $z=(x;\ y)$ представить в виде

$$z=(x;\ y)=(x;\ 0)+(0;\ y)=(x;\ 0)+(y;\ 0)\cdot(0;\ 1)=$$
$$=x+iy \qquad (1)$$

и производить над комплексными числами действия по обычным правилам алгебры многочленов. Запись (1) называется *алгебраической формой* комплексного числа.

Комплексное число $\bar{z}=(x;\ -y)=x-iy$ называется *комплексно сопряженным* числу $z=(x;\ y)=x+iy$ и изображается на комплексной плоскости точкой, симметричной точке z относительно оси Ox.

Пример. Найти произведение сопряженных чисел $z=x+iy$ и $\bar{z}=x-iy$.

Решение. Имеем:

$$z\cdot\bar{z}=(x+iy)\cdot(x-iy)=xx+ixy-iyx+yy=x^2+y^2.$$

Тригонометрическая форма комплексного числа. Введем на комплексной плоскости Oxy полярную систему координат так, чтобы полюс находился в начале O прямоугольной системы, а полярная ось была направлена вдоль положительного направления оси Ox (рис. 96). Рассмотрим комплексное число $z=x+iy$. Воспользовавшись формулами $x=\rho\cos\varphi$, $y=\rho\sin\varphi$, связывающими полярные и прямоугольные координаты, получим так называемую *тригонометрическую форму* записи комплексного числа $z=x+iy$:

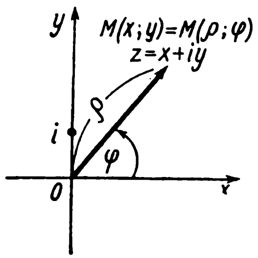


Рис. 96

$$z=\rho(\cos\varphi+i\sin\varphi).$$

Число ρ называется *модулем*, а число φ — *аргументом* комплексного числа z и обозначаются $\rho=|z|$, $\varphi=\operatorname{Arg} z$. Отметим, что аргумент φ определен неоднозначно, а с точностью до слагаемого $2\pi n$, где $n=0, \pm1, \pm2, \ldots$, а модуль имеет определенное значение:

$$\rho=|z|=|\overline{OM}|=\sqrt{x^2+y^2}.$$

Значение аргумента, удовлетворяющее неравенствам $0\leqslant\varphi<2\pi$, называется *главным* и обозначается $\varphi=\arg z$. Следует заметить, что аргумент комплексного числа $z=0$ не определен, а его модуль равен нулю.

Пример. Представить в тригонометрической форме следующие числа: 1) $z=2+i2\sqrt{3}$; 2) $z=5$; 3) $z=i$.

Решение. 1) Для числа $z=2+i2\sqrt{3}$ имеем $x=2$, $y=2\sqrt{3}$,

$$\rho=\sqrt{(2)^2+(2\sqrt{3})^2}=4,\ \cos\varphi=\frac{x}{\rho}=\frac{1}{2},\ \sin\varphi=\frac{y}{\rho}=\frac{\sqrt{3}}{2}.$$

Этим значениям косинуса и синуса соответствует значение аргумента $\varphi=\pi/3$. Следовательно,

$$z=2+i2\sqrt{3}=4\left(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}\right).$$

2) Для числа $z=5$ имеем $x=5$, $y=0$, $\rho=5$, $\cos\varphi=1$, $\sin\varphi=0$. Таким образом,

$$z=5=5(\cos 0+i\sin 0).$$

3) Для числа $z=i$ имеем $x=0$, $y=1$, $\rho=1$, $\cos\varphi=0$, $\sin\varphi=1$. Следовательно,

$$z=i=1\cdot(\cos\pi/2+i\sin\pi/2).$$

В тригонометрической форме удобно производить действия умножения и деления комплексных чисел. Пусть $z_1=\rho_1(\cos\varphi_1+i\sin\varphi_1)$ и $z_2=\rho_2(\cos\varphi_2+i\sin\varphi_2)$. Тогда

$$z=\rho(\cos\varphi+i\sin\varphi)=z_1\cdot z_2=$$
$$=\rho_1(\cos\varphi_1+i\sin\varphi_1)\cdot\rho_2(\cos\varphi_2+i\sin\varphi_2)=$$
$$=\rho_1\cdot\rho_2[\cos\varphi_1\cos\varphi_2+i\cos\varphi_1\sin\varphi_2+$$
$$+i\sin\varphi_1\cos\varphi_2-\sin\varphi_1\sin\varphi_2]=$$
$$=\rho_1\rho_2[\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)].$$

Таким образом,

$$\rho=\rho_1\cdot\rho_2,\ \varphi=\varphi_1+\varphi_2,$$

т. е. при умножении комплексных чисел их модули перемножаются, а аргументы складываются. В случае деления комплексных чисел при $\rho_2\neq0$ имеют место соотношения

$$\rho=\rho_1/\rho_2,\ \varphi=\varphi_1-\varphi_2.$$

В частности, если перемножаются n равных комплексных чисел, то $z^n=\rho^n(\cos n\varphi+i\sin n\varphi)$. Если положить $\rho=1$, то получим формулу

$$(\cos\varphi+i\sin\varphi)^n=\cos n\varphi+i\sin n\varphi,$$

называемую *формулой Муавра**.

Предел последовательности комплексных чисел. Теория пределов, рассмотренная нами ранее для последовательностей вещественных чисел, обобщается и на случай последовательностей комплексных чисел, при этом многие определения, связанные с предельным переходом, полностью повторяются.

Рассмотрим последовательность комплексных чисел $z_1, z_2, \ldots, z_n, \ldots$, где $z_n=(x_n+iy_n)$ $(n=1, 2, \ldots)$.

---

* Муавр Абрахам де (1667—1754) — английский математик, француз по происхождению.

Комплексное число $a=x_0+iy_0$ называется *пределом* последовательности $\{z_n\}$, если для любого $\varepsilon>0$ существует номер N такой, что при $n>N$ выполняется неравенство $|z_n-a|<\varepsilon$. Последовательность $\{z_n\}$ в этом случае называется сходящейся к числу a, что записывается в виде $\lim\limits_{n\to\infty} z_n=a$ или $z_n\to a$ при $n\to\infty$.

Геометрически это означает, что для любого $\varepsilon>0$, начиная с некоторого номера N, зависящего от ε, все элементы последовательности $\{z_n\}$ попадут в круг радиуса ε с центром в точке a, и точки z_n с возрастанием n будут неограниченно приближаться к точке a, так как

$$|z_n-a|=|(x_n-x_0)+i(y_n-y_0)|=$$
$$=\sqrt{(x_n-x_0)^2+(y_n-y_0)^2}=\rho(M_n;\ M_0)<\varepsilon,$$

где $\rho(M_n, M_0)$ есть расстояние между точками $M_n(x_n;\ y_n)$ и $M_0(x_0;\ y_0)$, т. е. между точками, изображающими числа z_n и a.

Поскольку каждое комплексное число z_n есть упорядоченная пара вещественных чисел $(x_n;\ y_n)$, то последовательности $\{z_n\}$ соответствуют две последовательности вещественных чисел $\{x_n\}$ и $\{y_n\}$, составленные соответственно из вещественных и мнимых частей элементов z_n последовательности $\{z_n\}$.

Имеет место следующая

Теорема 14.18. *Необходимым и достаточным условием сходимости последовательности* $\{z_n\}=\{x_n+iy_n\}$ *является сходимость последовательностей вещественных чисел* $\{x_n\}$ *и* $\{y_n\}$.

Доказательство. *Необходимость.* Пусть $z_n\to a=x_0+iy_0$. Тогда для любого $\varepsilon>0$ $|x_n-x_0|\leqslant|z_n-a|<\varepsilon$, $|y_n-y_0|\leqslant|z_n-a|<\varepsilon$ при $n>N$ и, следовательно, $x_n\to x_0$ и $y_n\to y_0$.

Достаточность. Пусть $x_n\to x_0$, $y_n\to y_0$ и пусть $a=x_0+iy_0$. Тогда из соотношения

$$|z_n-a|=\sqrt{(x_n-x_0)^2+(y_n-y_0)^2}$$

следует, что $z_n\to a$ при $n\to\infty$ ■

Доказанная теорема позволяет перенести основные результаты из теории пределов для последовательностей вещественных чисел на последовательности комплексных чисел.

2. Числовые ряды с комплексными членами. Составим ряд из комплексных чисел

$$z_1 + z_2 + \ldots + z_n + \ldots = \sum_{n=1}^{\infty} z_n, \qquad (2)$$

где $z_n = x_n + iy_n$. Ряд (2) называется сходящимся, если сходится последовательность $\{S_n\}$ его частичных сумм $S_n = z_1 + z_2 + \ldots + z_n = \left(\sum_{n=1}^{\infty} x_n\right) + i\left(\sum_{n=1}^{\infty} y_n\right)$ при $n \to \infty$, а предел $S = \lim\limits_{n\to\infty} S_n$ называется *суммой ряда.*

Ряду (2) с комплексными членами соответствует два вещественных ряда $\sum\limits_{n=1}^{\infty} x_n$ и $\sum\limits_{n=1}^{\infty} y_n$. Так же, как для последовательности комплекеных чисел (см. теорему 14.18), легко показать, что для сходимости ряда (2) необходимо и достаточно, чтобы сходились ряды $\sum\limits_{n=1}^{\infty} x_n$ и $\sum\limits_{n=1}^{\infty} y_n$. При этом если

$$\sum_{n=1}^{\infty} x_n = S' \text{ и } \sum_{n=1}^{\infty} y_n = S'', \text{ то } S = \sum_{n=1}^{\infty} z_n = S' + iS''.$$

Для изучения сходимости ряда (2) полезна следующая

Теорема 14.19. *Пусть дан ряд с комплексными членами*

$$z_1 + z_2 + \ldots + z_n + \ldots = \sum_{n=1}^{\infty} z_n. \qquad (3)$$

Тогда если сходится ряд, составленный из модулей членов ряда (3)

$$|z_1| + |z_2| + \ldots + |z_n| + \ldots = \sum_{n=1}^{\infty} |z_n|, \qquad (4)$$

то сходится ряд (3).

Доказательство. Пусть $z_n=x_n+iy_n$, тогда $|z_n|=\sqrt{x_n^2+y_n^2}$ и, очевидно, что $|x_n|\leqslant\sqrt{x_n^2+y_n^2}$, $|y_n|\leqslant\sqrt{x_n^2+y_n^2}$. Отсюда в силу признака сравнения рядов с вещественными членами (см. теорему 14.6) из сходимости ряда (4) вытекает, что неотрицательные ряды $\sum\limits_{n=1}^{\infty}|x_n|$ и $\sum\limits_{n=1}^{\infty}|y_n|$ сходятся. Следовательно, сходятся, и притом абсолютно, ряды $\sum\limits_{n=1}^{\infty}x_n$ и $\sum\limits_{n=1}^{\infty}y_n$, а этого достаточно для сходимости ряда (3) ■

Если ряд (4) сходится, то говорят, что ряд (3) сходится абсолютно.

Доказанная теорема позволяет применять при исследовании сходимости рядов с комплексными членами все достаточные признаки сходимости рядов с неотрицательными вещественными членами.

Пример. Исследовать на сходимость ряд

$$\sum_{n=1}^{\infty}\frac{(1+i)^n}{n!}.$$

Решение. Данный ряд сходится абсолютно, так как, применяя к ряду, составленному из модулей его членов, признак Даламбера, получим

$$\lim_{n\to\infty}\frac{|z_{n+1}|}{|z_n|}=\lim_{n\to\infty}\frac{|1+i|^{n+1}\cdot n!}{(n+1)!\,|1+i|^n}=\lim_{n\to\infty}\frac{|1+i|}{n+1}=$$

$$=\lim_{n\to\infty}\frac{\sqrt{1^2+1^2}}{n+1}=\lim_{n\to\infty}\frac{\sqrt{2}}{n+1}=0<1.$$

3. Степенные ряды с комплексными членами. *Ряд вида*

$$a_0+a_1(z-a)+a_2(z-a)^2+\ldots+a_n(z-a)^n+$$
$$+\ldots=\sum_{n=0}^{\infty}a_n(z-a)^n, \tag{5}$$

где z — комплексная переменная, a_n — комплексные *числа, a — некоторое комплексное число, называется степенным рядом.*

Как и для вещественных чисел, мы ограничимся рассмотрением степенных рядов вида

$$a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots = \sum_{n=0}^{\infty} a_n z^n. \qquad (6)$$

Ряд (5) приводится к виду (6) с помощью подстановки $z-a=t$.

Для определения области сходимости степенного ряда (6) используется теорема Абеля (см. теорему 14.11), которая формулируется и доказывается аналогично вещественному случаю. Приведем формулировку теоремы.

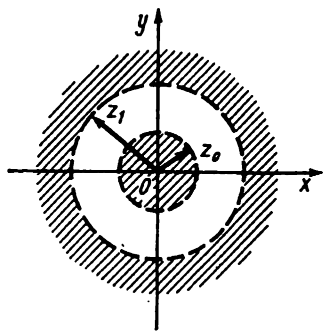


Рис. 97

Теорема 14.20. 1) *Если степенной ряд (6) сходится при $z=z_0$ ($z_0 \neq 0$), то он сходится, и притом абсолютно, для всех z, удовлетворяющих условию $|z|<|z_0|$;* 2) *если ряд (6) расходится при $z=z_1$, то он расходится для всех z, удовлетворяющих условию $|z|>|z_1|$.*

Рассмотрим геометрическое истолкование теоремы Абеля. Если z_0 — точка сходимости ряда (6), то во всех точках, расположенных внутри круга радиуса $|z_0|$ с центром в начале координат, ряд сходится абсолютно, а если z_1 — точка расходимости, то во всех точках, расположенных вне круга радиуса $|z_1|$ с центром в начале координат, ряд расходится (рис. 97).

Аналогично тому, как это было сделано для степенных рядов в вещественном случае, можно установить вид области сходимости ряда (6).

Теорема 14.21. *Если ряд (6) сходится не при всех значениях z, то существует число $R>0$ такое, что ряд абсолютно сходится при $|z|<R$ и расходится при $|z|>R$.*

Доказательство теоремы аналогично доказательству теоремы 14.12. Круг на плоскости комплексного переменного z с центром в начале координат и радиусом R

называют *кругом сходимости* степенного ряда (6), а число R — *радиусом сходимости*. На границе круга сходимости ($|z|=R$) вопрос о сходимости ряда решается, как и в вещественном случае, дополнительным исследованием.

Если степенной ряд (6) сходится только в точке $z=0$, то полагают $R=0$, если же ряд сходится при любом значении z, т. е. во всей комплексной области, то полагают $R=\infty$.

Радиус сходимости можно находить так же, как и в вещественном случае.

Пример. Найти радиус сходимости степенного ряда $\sum\limits_{n=1}^{\infty}\frac{z^n}{n!}$.

Решение. Находим

$$R=\lim_{n\to\infty}\frac{|a_n|}{|a_{n+1}|}=\lim_{n\to\infty}\frac{(n+1)!}{n!}=\lim_{n\to\infty}(n+1)=\infty,$$

следовательно, ряд сходится при любом значении z.

Рассмотрим ряд

$$f(z)=a_0+a_1z+a_2z^2+\ldots+a_nz^n+\ldots=\sum_{n=0}^{\infty}a_nz^n. \quad (7)$$

Очевидно, внутри круга сходимости ряда (7) его сумма $f(z)$ есть функция комплексной переменной. Такие функции, т. е. функции, представимые в виде степенных рядов, называются *аналитическими*. Таким образом, мы пришли к понятию функции комплексной переменной. Изучение таких функций выходит за рамки данного курса.

4. Формулы Эйлера*. В заключение данного параграфа с помощью степенных рядов установим две формулы, которые будут нам полезны в дальнейшем.

В § 5 было получено разложение функции e^x в степенной ряд, сходящийся при любом значении x:

---

* Эйлер Леонард (1707—1783) — известный математик, механик, физик и астроном, член Петербургской Академии наук, большую часть жизни провел в России, по происхождению швейцарец.

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots \qquad (8)$$

Если в нем вещественную переменную x заменить комплексной переменной z, то получим ряд по степеням z:

$$1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots,$$

сходящийся, как это следует из только что рассмотренного примера, при всех значениях z. Обозначим его сумму через e^z. Таким образом, по определению для любого комплексного числа z

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots \qquad (9)$$

Естественно сумму ряда (9) назвать показательной функцией комплексной переменной z, так как при вещественных значениях z $(z=x)$ сумма ряда равна e^x.

Аналогично определяются тригонометрические функции $\sin z$ и $\cos z$ комплексной переменной z:

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots + (-1)^{n-1} \frac{z^{2n-1}}{(2n-1)!} + \ldots, \qquad (10)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots + (-1)^n \frac{z^{2n}}{(2n)!} + \ldots \qquad (11)$$

Между показательной функцией e^z и тригонометрическими функциями $\sin z$ и $\cos z$ существует простая связь. Подставим в (9) iz вместо z и сгруппируем отдельно в ряде все слагаемые, содержащие множитель i, и отдельно — не содержащие множителя i:

$$e^{iz} = 1 + iz + \frac{i^2z^2}{2!} + \frac{i^3z^3}{3!} + \frac{i^4z^4}{4!} + \frac{i^5z^5}{5!} + \ldots =$$

$$= 1 + iz - \frac{z^2}{2!} - i\frac{z^3}{3!} + \frac{z^4}{4!} + i\frac{z^5}{5!} - \ldots =$$

$$= \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots\right) + i\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots\right).$$

Сравнивая полученный результат с формулами (10) и (11), получаем

$$e^{iz} = \cos z + i \sin z. \qquad (12)$$

Далее, подставляя в (9) $-iz$ вместо z, получим

$$e^{-iz} = \cos z - i \sin z. \tag{13}$$

Формулы (12) и (13) называются *формулами Эйлера.* Они устанавливают связь между показательной и тригонометрическими функциями комплексной переменной z. Если почленно сложим и вычтем равенства (12) и (13), то получим другую запись тех же формул Эйлера:

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}; \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

Таким образом, функции e^z, $\sin z$, $\cos z$ тесно связаны.

Заметим, что при $z = x$ (x — вещественная переменная) определенные выше функции комплексной переменной совпадают соответственно с функциями e^x, $\sin x$ и $\cos x$ вещественной переменной.

§ 7. Ряды Фурье

1. Тригонометрический ряд и его основные свойства. *Ряд вида*

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots +$$

$$+ a_n \cos nx + b_n \sin nx + \ldots =$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1}$$

называется тригонометрическим рядом. Числа a_0, a_1, b_1, a_2, b_2, ..., a_n, b_n, ... называются *коэффициентами* тригонометрического ряда.

В отличие от степенного ряда, рассмотренного ранее, в тригонометрическом ряду вместо простейших функций — степеней x: 1, x, x^2, ..., x^n, ... выбраны другие функции — тригонометрические:

$$1/2, \ \cos x, \ \sin x, \ \cos 2x, \ \sin 2x, \ \ldots, \ \cos nx, \ \sin nx, \ \ldots, \tag{2}$$

которые также являются достаточно простыми и хорошо изученными.

Прежде всего отметим, что все функции системы (2) являются периодическими с периодами 2π. В самом деле, $1/2$ имеет любой период, а период функций $\sin nx$ и $\cos nx$ $(n=1, 2, \ldots)$ равен $2\pi/n$* и, следовательно, число $2\pi=n(2\pi/n)$ также является их периодом. Очевидно, что каждый член тригонометрического ряда (1) является периодической функцией периода 2π. Но тогда и частичная сумма ряда (1) должна быть 2π-периодична (если все члены ряда не меняются от замены x на $x+2\pi$, то и сумма его не меняется от этой замены). Отсюда следует, что если ряд (1) сходится на отрезке $[-\pi, \pi]$, то его сумма, будучи пределом периодической частичной суммы, является периодической функцией $f(x)$ с периодом 2π. Поэтому тригонометрические ряды особенно удобны при изучении периодических функций, математически описывающих различные периодические процессы, происходящие в природе и технике. Примерами периодических процессов могут служить колебательные и вращательные движения различных деталей машин и приборов, периодическое движение небесных тел и элементарных частиц, акустические и электромагнитные колебания и многие другие.

Другим важным свойством функций системы (2) является их ортогональность на отрезке $[-\pi, \pi]$ в следующем смысле: интеграл по отрезку $[-\pi, \pi]$ от произведения любых двух различных функций этой системы равен нулю, а интеграл по отрезку $[-\pi, \pi]$ от квадрата любой функции этой системы отличен от нуля.

Действительно,

$$\int_{-\pi}^{\pi} \frac{1}{2}\cos kx\,dx = \frac{1}{2k}\sin kx\Big|_{-\pi}^{\pi} = 0; \quad \int_{-\pi}^{\pi} \frac{1}{2}\sin kx\,dx = -\frac{1}{2k}\cos kx\Big|_{-\pi}^{\pi} = 0. \tag{3}$$

Далее,

$$\int_{-\pi}^{\pi} \cos kx\cdot\cos nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} [\cos(k+n)x + \cos(k-n)x]\,dx =$$

* Действительно, $\sin\left[n\left(x \pm \frac{2\pi}{n}\right)\right] = \sin(nx \pm 2\pi) = \sin nx.$

$$= \frac{1}{2}\left[\frac{\sin(k+n)x}{k+n} + \frac{\sin(k-n)x}{k-n}\right]_{-\pi}^{\pi} = 0 \text{ при } k \neq n. \qquad (4)$$

Аналогично находим

$$\int_{-\pi}^{\pi} \sin kx \cdot \sin nx\,dx = 0, \quad \int_{-\pi}^{\pi} \sin kx \cdot \cos nx\,dx = 0 \text{ при } k \gtreqless n. \qquad (5)$$

Наконец,

$$\left.\begin{aligned} &\int_{-\pi}^{\pi} \cos^2 kx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} (1 + \cos 2kx)\,dx = \\ &\qquad = \frac{1}{2}\left[x + \frac{1}{2k}\sin 2kx\right]_{-\pi}^{\pi} = \pi, \\ &\int_{-\pi}^{\pi} \sin^2 kx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} (1 - \cos 2kx)\,dx = \\ &\qquad = \frac{1}{2}\left[x - \frac{1}{2k}\sin 2kx\right]_{-\pi}^{\pi} = \pi, \\ &\qquad \int_{-\pi}^{\pi} \left(\frac{1}{2}\right)^2 dx = \frac{1}{4}x\Big|_{-\pi}^{\pi} = \frac{\pi}{2}, \end{aligned}\right\} \qquad (6)$$

что и требовалось.

2. Ряд Фурье. Аналогично степенному ряду для тригонометрического ряда имеет место следующая

Теорема 14.22. *Если функция $f(x)$, определенная и непрерывная на отрезке $[-\pi, \pi]$, разлагается в тригонометрический ряд*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \qquad (7)$$

который можно интегрировать почленно, то это разложение единственно.

Доказательство. Интегрируя (7), получим

$$\int_{-\pi}^{\pi} f(x)\,dx = \frac{a_0}{2}\int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty}\left[a_n\int_{-\pi}^{\pi} \cos nx\,dx + b_n\int_{-\pi}^{\pi} \sin nx dx\right],$$

откуда, с учетом (3), находим

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\,dx. \tag{8}$$

Для нахождения коэффициента a_k при $\cos kx$ (k — натуральное число) умножим равенство (7) на $\cos kx$ и проинтегрируем по x от $-\pi$ до π. Тогда в силу (3) — (6) получим

$$\int_{-\pi}^{\pi} f(x) \cos kx\,dx = \frac{a_0}{2} \int_{-\pi}^{\pi} \cos kx\,dx +$$

$$+ \sum_{n=1}^{\infty} \left[a_n \int_{-\pi}^{\pi} \cos kx \cdot \cos nx\,dx + b_n \int_{-\pi}^{\pi} \cos kx \cdot \sin nx\,dx \right] =$$

$$= a_k \int_{-\pi}^{\pi} \cos^2 kx dx = a_k \pi,$$

откуда

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx\,dx. \tag{9}$$

Аналогично, умножая равенство (7) на $\sin kx$ и интегрируя в пределах от $-\pi$ до π, на основании (3)—(6) получим

$$\int_{-\pi}^{\pi} f(x) \sin kx\,dx = b_k \pi,$$

откуда находим выражения для коэффициентов b_k:

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx\,dx. \tag{10}$$

Таким образом, свободный член a_0 ряда (7) и его коэффициенты a_k и b_k определяются единственным образом формулами (8)—(10), что и доказывает теорему ■

Эта теорема дает основание ввести следующее

Определение. *Пусть $f(x)$ — функция, определенная и интегрируемая на отрезке* $[-\pi, \pi]$. *Тогда числа* a_0, a_n, b_n, *найденные по формулам* (8)—(10), *называются коэффициентами Фурье, а ряд*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

с этими коэффициентами называется рядом Фурье функции $f(x)$.*

3. Сходимость ряда Фурье. Естественно возникает вопрос, является ли ряд Фурье сходящимся, и если он сходится, то каким условиям должна удовлетворять функция $f(x)$, породившая этот ряд, чтобы он сходился к самой этой функции?

Аналогичный вопрос ставился при изучении степенных рядов.

Ответ на поставленный вопрос дает следующая

Теорема 14.23. *Пусть функция $f(x)$ и ее производная $f'(x)$ — непрерывные функции на отрезке $[-\pi, \pi]$ или же имеют на нем конечное число точек разрыва 1-го рода. Тогда ряд Фурье функции $f(x)$ сходится на всей числовой прямой, причем в каждой точке $x \in (-\pi, \pi)$, в которой $f(x)$ непрерывна, сумма ряда равна $f(x)$, а в каждой точке x_0 разрыва функции сумма ряда равна*

$$\frac{f(x_0-)+f(x_0+)}{2},$$

где $f(x_0-)=\lim\limits_{x\to x_0-} f(x)$ *и* $f(x_0+)=\lim\limits_{x\to x_0+} f(x)$.

Наконец, на концах отрезка $[-\pi, \pi]$ сумма ряда равна

$$\frac{f(-\pi)+f(\pi)}{2}.$$

Доказательство теоремы опускается.

4. Ряды Фурье для четных и нечетных функций. Пусть функция $f(x)$ определена на отрезке $[-\pi, \pi]$ и является четной, т. е. $f(-x)=f(x)$. Тогда ее коэффициенты Фурье b_n равны нулю. Действительно,

$$b_n=\frac{1}{\pi}\int\limits_{-\pi}^{\pi} f(x)\sin nx\,dx=\frac{1}{\pi}\left[\int\limits_{-\pi}^{0} f(x)\sin nx\,dx+\right.$$

$$\left.+\int\limits_{0}^{\pi} f(x)\sin nx\,dx\right].$$

* Фурье Жан Батист Жозеф (1768—1830) — французский математик и физик.

В первом интеграле в квадратных скобках сделаем замену переменной. Положим $x=-t$. Тогда $dx=-dt$; если $x=-\pi$, то $t=\pi$; если $x=0$, то $t=0$. Пользуясь четностью $f(x)$ и нечетностью синуса, получим

$$\int_{-\pi}^{0} f(x)\sin nx\,dx = -\int_{\pi}^{0} f(-t)\sin n(-t)\,dt = -\int_{0}^{\pi} f(t)\sin nt\,dt.$$

Следовательно,

$$b_n = \frac{1}{\pi}\left[-\int_{0}^{\pi} f(t)\sin nt\,dt + \int_{0}^{\pi} f(x)\sin nx\,dx\right] = 0,$$

так как определенный интеграл не зависит от обозначения переменной интегрирования.

Аналогично, пользуясь четностью $f(x)$ и четностью косинуса, получим следующие выражения для коэффициентов a_n:

$$a_0 = \frac{2}{\pi}\int_{0}^{\pi} f(x)\,dx, \quad a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos nx\,dx. \qquad (11)$$

Пусть теперь функция $f(x)$, определенная на отрезке $[-\pi, \pi]$, нечетна, т. е. $f(x)=-f(-x)$. Тогда, используя рассуждения, аналогичные приведенным выше, можно показать, что в этом случае коэффициенты Фурье a_n равны нулю, а коэффициенты b_n определяются выражениями

$$b_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\sin nx\,dx. \qquad (12)$$

Таким образом, если функция $f(x)$ четная, то ряд Фурье содержит только косинусы, а если функция $f(x)$ нечетная — только синусы. Формулы (11) и (12) позволяют упрощать вычисления коэффициентов Фурье, когда заданная функция является четной или нечетной.

Пример 1. Рассмотрим функцию $f(x)=x$. Эта функция удовлетворяет условиям теоремы 14.23 и, следовательно, может быть разложена в ряд Фурье. Так как она нечетная, то ее коэффициенты Фурье $a_n=0$, а b_n находятся по формуле (12). Имеем

$$b_n = \frac{2}{\pi}\int_0^{\pi} x\sin nx\,dx = \frac{2}{\pi}\left[-\frac{1}{n}x\cos nx\Big|_0^{\pi} + \right.$$

$$\left. + \frac{1}{\pi}\int_0^{\pi}\cos nx\,dx\right] = (-1)^{n+1}\frac{2}{n}.$$

Таким образом, получаем ряд Фурье данной функции

$$x = 2\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \right.$$

$$\left. + \ldots + (-1)^{n+1}\frac{\sin nx}{n} + \ldots\right).$$

Это равенство справедливо для любого $x\in(-\pi, \pi)$. В точках $x=\pm\pi$ сумма ряда Фурье по теореме 14.23 не совпадает со значениями функции $f(x)=x$, а равна

$$\frac{f(-\pi)+f(\pi)}{2} = \frac{-\pi+\pi}{2} = 0.$$

Вне отрезка $[-\pi, \pi]$ сумма ряда имеет периодическое продолжение, график которого изображен на рис. 98, *а*.

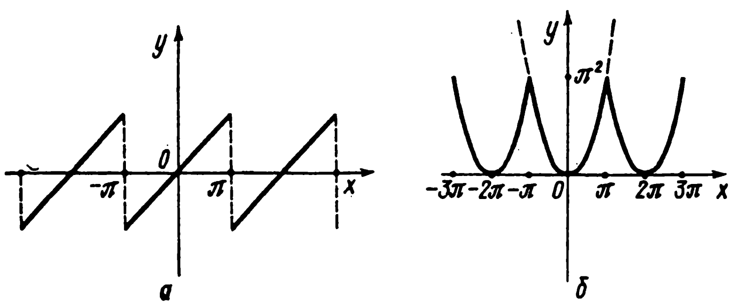


Рис. 98

Пример 2. Рассмотрим функцию $f(x)=x^2$. Эта функция удовлетворяет условиям теоремы 14.23 и, следовательно, может быть разложена в ряд Фурье. Так как она четная, то ее коэффициенты Фурье $b_n=0$, а a_n находятся по формулам (11). Имеем

$$a_0=\frac{2}{\pi}\int_0^\pi x^2\,dx=\frac{2\pi^2}{3};\quad a_n=\frac{2}{\pi}\int_0^\pi x^2\cos nx\,dx=$$

$$=\frac{2}{\pi}\left[\frac{x^2\sin nx}{n}\Big|_0^\pi-\frac{2}{\pi}\int_0^\pi x\sin nx\,dx\right]=(-1)^n\frac{4}{n^2}.$$

Значит, ряд Фурье данной функции имеет вид

$$x^2=\frac{\pi^2}{3}-4\left(\frac{\cos x}{1}-\frac{\cos 2x}{2^2}+\frac{\cos 3x}{3^2}-\ldots\right).$$

Это равенство справедливо для любого $x\in[-\pi,\pi]$, так как в точках $x=\pm\pi$ сумма ряда по теореме 14.23 совпадает со значениями функции $f(x)=x^2$, ибо $(f(-\pi)+f(\pi))/2=(\pi^2+\pi^2)/2=\pi^2=f(\pi)=f(-\pi)$. Графики функции $f(x)=x^2$ и суммы данного ряда Фурье изображены на рис. 98, *б*.

5. **Ряд Фурье для функции с периодом $2l$.** В ряд Фурье можно разлагать функции, период которых отличен от 2π.

Пусть $f(x)$ есть периодическая функция с периодом $2l$ (l — произвольное положительное число) и на отрезке $[-l, l]$ удовлетворяет условиям теоремы 14.23. Разложим ее в ряд Фурье.

Введем новую переменную ξ по формуле

$$x=l\xi/\pi$$

и рассмотрим функцию $\varphi(\xi)=f(l\xi/\pi)=f(x)$. Покажем, что $\varphi(\xi)$ есть функция с периодом 2π. Из последнего равенства следует, что

$$\varphi(\xi+2\pi)=f\left[\frac{l}{\pi}(\xi+2\pi)\right]=f\left[\frac{l}{\pi}\xi+2l\right]=f(x+2l).$$

Но так как $f(x)$ имеет период $2l$, то $f(x+2l)=f(x)=\varphi(\xi)$. Следовательно, $\varphi(\xi+2\pi)=\varphi(\xi)$, что и требовалось показать.

Таким образом, функция $\varphi(\xi)$ определена на отрезке $[-\pi,\pi]$ и удовлетворяет на нем условиям теоремы 14.23, поскольку замена x новой переменной ξ линейна и не может привести к нарушению требований теоремы.

Разложим функцию $\varphi(\xi)$ на отрезке $[-\pi, \pi]$ в ряд Фурье:

$$\varphi(\xi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\xi + b_n \sin n\xi), \qquad (13)$$

где

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(\xi)\, d\xi; \quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(\xi) \cos n\xi\, d\xi;$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(\xi) \sin n\xi\, d\xi.$$

Вернемся теперь от ξ обратно к старой переменной x: $x = \frac{l}{\pi}\xi$, $\xi = x\frac{\pi}{l}$, $d\xi = \frac{\pi}{l}dx$. Тогда формула (13) примет вид

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi x}{l} + b_n \sin\frac{n\pi x}{l}\right), \qquad (14)$$

где

$$a_0 = \frac{1}{l}\int_{-l}^{l} f(x)\, dx; \quad a_n = \frac{1}{l}\int_{-l}^{l} f(x) \cos\frac{n\pi x}{l}\, dx;$$

$$b_n = \frac{1}{l}\int_{-l}^{l} f(x) \sin\frac{n\pi x}{l}\, dx.$$

Формула (14) и есть ряд Фурье для периодической функции с периодом $2l$.

Пример 3. Разложить в ряд Фурье функцию $f(x)$ с периодом $2l$, которая на отрезке $[-l, l]$ задается формулой $f(x) = |x|$.

Решение. Так как функция $f(x) = |x|$ — четная, то

$$b_n = 0; \quad a_0 = \frac{2}{l}\int_{0}^{l} x dx = \frac{2}{l}\cdot\frac{x^2}{2}\Big|_0^l = l;$$

$$a_n = \frac{2}{l}\int_0^l x\cos\frac{n\pi x}{l}\,dx = \frac{2}{l}\left[\frac{lx\sin\frac{n\pi x}{l}}{n\pi}\Bigg|_0^l - \right.$$

$$\left. - \frac{l}{n\pi}\int_0^l \sin\frac{n\pi x}{l}\,dx\right] = \frac{2l}{n^2\pi^2}\cos\frac{n\pi x}{l}\Bigg|_0^l =$$

$$= \frac{2l}{n^2\pi^2}[(-1)^n - 1] = \begin{cases} 0 & \text{при } n \text{ четном} \\ -\dfrac{4l}{n^2\pi^2} & \text{при } n \text{ нечетном.} \end{cases}$$

Следовательно, ряд Фурье функции $f(x)$ имеет вид

$$|x| = \frac{l}{2} - \frac{4l}{\pi^2}\left[\cos\frac{\pi x}{l} + \frac{1}{3^2}\cos\frac{3\pi x}{l} + \frac{1}{5^2}\cos\frac{5\pi x}{l} + \ldots\right].$$

Функция $|x|$ удовлетворяет условиям теоремы 14.23, и полученное равенство справедливо для любого $x \in [-l, l]$, а это значит, что ряд сходится на всей числовой прямой и имеет своей суммой функцию, график которой изображен на рис. 99.

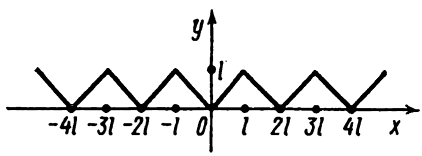


Рис. 99

В заключение заметим, что ряды Фурье имеют очень широкое применение как в теоретических исследованиях, так и в практических задачах.

Глава XV

ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

В математике дифференциальные уравнения занимают особое место. Дело в том, что математическое исследование самых разнообразных явлений, происходящих в природе, часто приводит к решению таких уравнений, поскольку сами законы, которым подчиняется то или иное явление, записываются в виде дифференциальных уравнений.

Дифференциальные уравнения — это уравнения, в которые неизвестная функция входит под знаком производной. Основной задачей теории дифференциальных уравнений является изучение функций, являющихся решениями таких уравнений.

Дифференциальные уравнения делятся на обыкновенные дифференциальные уравнения, в которых неизвестные функции являются функциями одной переменной, и на дифференциальные уравнения в частных производных, в которых неизвестные функции являются функциями двух и большего числа переменных.

Теория дифференциальных уравнений в частных производных более сложна и изучается в полных или специальных математических курсах. С элементами теории обыкновенных дифференциальных уравнений мы познакомимся в настоящей главе. В дальнейшем, говоря о дифференциальных уравнениях, будем иметь в виду только обыкновенные дифференциальные уравнения.

§ 1. Дифференциальные уравнения первого порядка

Изучение теории дифференциальных уравнений начнем с наиболее простого уравнения — с уравнения первого порядка.

1. Определение дифференциального уравнения первого порядка

Определение 1. *Уравнение вида*

$$F(x, y, y')=0, \qquad (1)$$

где x — независимая переменная, y — искомая функция, y' — ее производная, называется дифференциальным уравнением первого порядка.

Если уравнение (1) можно разрешить относительно y', то его можно записать в виде

$$y'=f(x, y). \qquad (2)$$

Уравнение (2) называется *уравнением первого поряд-*

ка, разрешенным относительно производной. Мы будем рассматривать именно такие уравнения.

В некоторых случаях уравнение (2) удобно записать в виде $dy/dx=f(x, y)$ или в виде $f(x, y)dx-dy=0$, являющемся частным случаем более общего уравнения

$$P(x, y)dx+Q(x, y)dy=0, \tag{3}$$

где $P(x, y)$ и $Q(x, y)$ — известные функции. Уравнение в симметричной форме (3) удобно тем, что переменные x и y в нем равноправны, т. е. каждую из них можно рассматривать как функцию другой.

Приведем примеры дифференциальных уравнений вида (2) и (3):

$$y'=xe^y,\ y'=(y\cdot\ln x)/x,\ y'=x+y,\ xdx+ydy=0 \quad \text{и т. д.}$$

2. Решение уравнения. Задача Коши

Определение 2. *Решением дифференциального уравнения первого порядка называется всякая функция $y=\varphi(x)$, которая при подстановке в уравнение обращает его в тождество.*

Так, например, функция $y=x^3$ является решением уравнения $3y-xy'=0$, т. е. при подстановке в уравнение обращает его в тождество: $3x^3-x3x^2\equiv0$.

График решения дифференциального уравнения называется *интегральной кривой.*

Возникает вопрос, при каких условиях уравнение (2) имеет решение. Ответ дает теорема Коши, называемая теоремой о существовании и единственности решения дифференциального уравнения (2) и являющаяся основной в теории дифференциальных уравнений.

Теорема 15.1. (Теорема Коши). *Если в уравнении $y'=f(x, y)$ функция $f(x, y)$ и ее частная производная $f'_y(x, y)$ определены и непрерывны в некоторой области G плоскости Oxy, то какова бы ни была внутренняя точка $(x_0;\ y_0)$ области G, существует единственное решение $y=\varphi(x)$ данного уравнения, удовлетворяющее условию*

$$y=y_0 \text{ при } x=x_0. \tag{4}$$

Теорема принимается без доказательства.

Геометрически теорема утверждает, что через каждую внутреннюю точку $(x_0;\ y_0)$ области G проходит единственная интегральная кривая. Очевидно, что в области G уравнение (2) имеет бесконечное число различных решений.

Условия (4), в силу которых функция $y=\varphi(x)$ принимает заданное значение y_0 в заданной точке x_0, называются *начальными условиями* решения и записываются обычно так:

$$y|_{x=x_0}=y_0. \tag{5}$$

Отыскание решения уравнения (2), удовлетворяющего начальным условиям (5), является одной из важнейших задач теории дифференциальных уравнений. Эта задача называется *задачей Коши*. С геометрической точки зрения решить задачу Коши — значит из множества интегральных кривых выделить ту, которая проходит через заданную точку $(x_0; y_0)$ плоскости Oxy.

Точки плоскости, в которых не выполняются условия теоремы Коши, называются *особыми* точками. Через каждую из них может проходить либо несколько интегральных кривых, либо ни одной.

3. Общее и частное решения уравнения. Дадим теперь два основных определения.

Определение 3. *Общим решением уравнения* (2) *в некоторой области G называется функция $y=\varphi(x, C)$, зависящая от x и одной произвольной постоянной C, если она является решением уравнения* (2) *при любом значении постоянной C и если при любых начальных условиях* (5) *таких, что $(x_0; y_0)\in G$, существует единственное значение постоянной $C=C_0$ такое, что функция $y=\varphi(x, C_0)$ удовлетворяет данным начальным условиям: $\varphi(x_0, C_0)=y_0$.*

Определение 4. *Частным решением уравнения* (2) *в некоторой области G называется функция $y=\varphi(x, C_0)$, которая получается из общего решения $y=\varphi(x, C)$ при определенном значении постоянной $C=C_0$.*

Геометрически общее решение $y=\varphi(x, C)$ представляет собой семейство интегральных кривых на плоскости Oxy, зависящее от одной произвольной постоянной C, а частное решение $y=\varphi(x, C_0)$ — одну интегральную кривую этого семейства, проходящую через заданную точку $(x_0; y_0)$.

Иногда начальные условия (5) называют *условиями*

Коши, а частным решением — решение какой-нибудь задачи Коши.

Обратимся к примерам.

Пример 1. Рассмотрим уравнение $y'=3x^2$.

Данное уравнение является дифференциальным уравнением первого порядка. Оно удовлетворяет всем условиям теоремы Коши, так как функции $f(x, y)=3x^2$ и $f'_y(x, y)=0$ определены и непрерывны на всей плоскости Oxy.

Легко проверить, что функция $y=x^3+C$, где C — произвольная постоянная, является общим решением данного уравнения во всей плоскости Oxy.

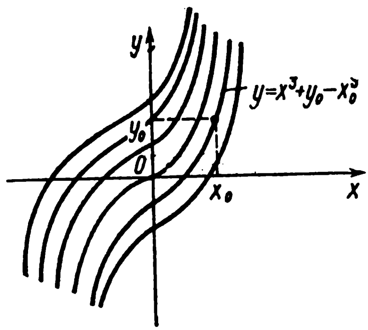


Рис. 100

Геометрически общее решение представляет собой семейство кубических парабол $y=x^3+C$. При различных значениях постоянной C получаются различные решения данного уравнения. Например, если $C=0$, то $y=x^3$, если $C=-1$, то $y=x^3-1$, если $C=2$, то $y=x^3+2$ и т. д.

Для решения какой-нибудь задачи Коши, т. е. отыскания частного решения, зададим любые начальные условия: $x=x_0$, $y=y_0$. Подставляя эти значения в общее решение $y=x^3+C$ вместо x и y, получим $y_0=x^3_0+C$, откуда $C=y_0-x^3_0$, и соответствующее частное решение $y=x^3+y_0-x^3_0$. Геометрически это означает, что из семейства кубических парабол $y=x^3+C$ выбрана одна парабола, проходящая через заданную точку $(x_0; y_0)$ (рис. 100).

Пример 2. Рассмотрим уравнение $y'=-y/x$.

Данное уравнение есть дифференциальное уравнение первого порядка. Функции $f(x, y) = -y/x$ и $f'_y(x, y) = -1/x$ непрерывны при $x \neq 0$. Следовательно, во всей плоскости Oxy, кроме оси Oy, данное уравнение удовлетворяет условиям теоремы Коши. Точки, лежащие на оси Oy, являются особыми.

Нетрудно проверить, что общим решением данного уравнения в областях $y>0$ и $y<0$ является функция $y=C/x$, где C — произвольная постоянная. При различных значениях постоянной C: $C=1/2$, $C=-1$, $C=1$, $C=2$ и т. д. получаются различные решения.

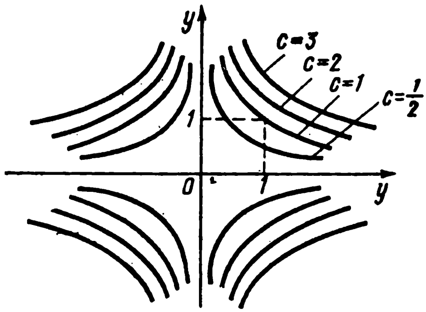


Рис. 101

Найдем частное решение, удовлетворяющее, например, начальным условиям $x_0=1$, $y_0=1$. Имеем $1=C/1$. Отсюда $C=1$ и искомое частное решение $y=1/x$.

Геометрически общее решение данного уравнения — семейство гипербол $y=C/x$, каждая из которых изображает частное решение данного уравнения. Задавая начальные условия $x_0=1$, $y_0=1$, мы выделим из всего семейства ту гиперболу, которая проходит через точку $(1; 1)$ плоскости Oxy (рис. 101).

Заметим, что через особые точки, лежащие на оси Oy, не проходит ни одна интегральная кривая.

4. Геометрический смысл уравнения. Пусть дано дифференциальное уравнение первого порядка $y'=f(x, y)$ и пусть функция $y=\varphi(x)$ есть решение данного уравнения. График решения есть непрерывная интегральная кривая, через каждую точку которой можно провести касательную. Из уравнения следует, что угловой коэффициент y' касательной к интегральной кривой в каждой ее точке (x; y) равен значению в этой точке правой части уравнения $f(x, y)$. Таким образом, уравнение $y'=f(x, y)$ устанавливает зависимость между координатами точки (x; y) и угловым коэффициентом y' касательной к графику интегральной кривой в той же точке. Зная x и y, мы можем указать направление касательной к этой интегральной кривой в точке (x; y).

Сопоставим каждой точке (x; y) интегральной кривой направленный отрезок, угловой коэффициент которого равен $f(x, y)$. Получим так называемое *поле направлений* данного уравнения, раскрывающее геометрический смысл дифференциального уравнения первого порядка.

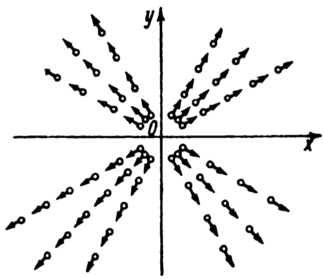


Рис. 102

Итак, с геометрической точки зрения уравнение $y'=f(x, y)$ определяет на плоплоскости Oxy поле направлений, а решение этого уравнения есть интегральная кривая, направление касательной к которой в каждой точке совпадает с направлением поля в этой точке.

Построив на плоскости поле направлений данного дифференциального уравнения, можно приближенно построить интегральные кривые и по их форме определить решения уравнения.

Пример 3. Рассмотрим уравнение $y'=y/x$.

Функция $f(x, y)=y/x$ не определена при $x=0$, следовательно, поле направлений для данного уравнения можно построить на всей плоскости, кроме оси Oy.

В каждой точке (x; y) ($x\neq0$) угловой коэффициент y' касательной к интегральной кривой равен y/x и совпадает с угловым коэффициентом прямой, направленной из начала координат в точку (x; y). На рис. 102 изображено поле направлений данного уравнения. Очевидно, что интегральными кривыми будут прямые $y=Cx$ (C — произвольная постоянная), т. е. направление этих прямых всюду совпадает с направлением поля.

Теперь перейдем к изучению методов нахождения решений дифференциальных уравнений первого порядка Сразу отметим, что общего метода нахождения решений не существует. Обычно рассматриваются отдельные типы таких уравнений, для каждого из которых дается свой способ нахождения решения.

5. Уравнения с разделяющимися переменными.

Определение 5. *Уравнение вида*

$$y'=f_1(x)\,f_2(y), \qquad (6)$$

где $f_1(x)$ и $f_2(y)$ — непрерывные функции, зависящие только от одного аргумента, называется дифференциальным уравнением с разделяющимися переменными.

Для отыскания решения уравнения (6) нужно, как говорят, разделить в нем переменные. Для этого заменим в (6) y' на dy/dx, разделим на $f_2(y)$ (предполагается $f_2(y)\neq0$) и умножим на dx. Тогда уравнение (6) преобразуется в уравнение

$$dy/f_2(y)=f_1(x)dx, \qquad (7)$$

в котором переменная x входит только в правую часть, а переменная y — только в левую (в таком случае говорят, что переменные разделены).

Предполагая, что функция $y=\varphi(x)$ является решением уравнения, и подставляя ее в (7), получим тождество. Интегрирование этого тождества дает соотношение

$$\int dy/f_2(y)=\int f_1(x)dx+C, \qquad (8)$$

где C — произвольная постоянная.

Соотношение (8) определяет неявным образом общее решение уравнения (6).

Пример 4. Решить уравнение $y'=y/x$ (сравните с примером 3).

Решение. Данное уравнение есть уравнение вида (6), где $f_1(x)=1/x$ и $f_2(y)=y$. Разделяя переменные, получим $dy/y=dx/x$. Интегрируя, будем иметь $\int dy/y=\int dx/x+\ln|C_1|$, $C_1\neq0$*, или $\ln|y|=\ln|x|+\ln|C_1|$. Потенцируя, находим $|y|=|C_1||x|$, что эквивалентно уравнению $y=\pm C_1x$. Полагая $\pm C_1=C$, окончательно получим

* Для упрощения полученного решения мы обозначили произвольную постоянную через $\ln|C_1|$, что законно, так как $\ln|C_1|$ может принимать любое значение от $-\infty$ до $+\infty$.

$$y = Cx \tag{9}$$

— общее решение данного уравнения, где C — произвольная постоянная, которая может принимать как положительные, так и отрицательные значения, но $C \neq 0$. Заметим, однако, что $y = 0$ является решением уравнения, оно было потеряно при делении на y. Это решение можно включить в (9), если считать, что постоянная C принимает и значение $C = 0$. Геометрически общее решение (9) представляет собой семейство прямых, проходящих через начало координат.

Пусть требуется выделить из общего решения (9) частное решение, удовлетворяющее начальным условиям: $x_0 = 1$, $y_0 = 2$. Подставляя эти значения в общее решение (9) вместо x и y, получим $2 = C \cdot 1$, отсюда $C = 2$. Таким образом, искомое частное решение $y = 2x$.

6. Линейные уравнения

Определение 6. *Уравнение вида*

$$y' + p(x)y = f(x), \tag{10}$$

где $p(x)$ и $f(x)$ — непрерывные функции, называется линейным дифференциальным уравнением первого порядка.

Название уравнения объясняется тем, что неизвестная функция y и ее производная y' входят в уравнение линейно, т. е. в первой степени.

Если $f(x) \equiv 0$, то уравнение (10) называется *линейным однородным* уравнением. Если $f(x) \not\equiv 0$, то уравнение (10) называется *линейным неоднородным* уравнением.

Для нахождения общего решения уравнения (10) может быть применен так называемый *метод вариации постоянной.*

Сначала находится общее решение линейного однородного уравнения

$$y' + p(x)y = 0, \tag{11}$$

соответствующего данному неоднородному (10). Уравнение (11) — уравнение с разделяющимися переменными. Разделяя переменные и интегрируя, будем иметь

$$dy/y = -p(x)dx, \quad \ln|y| = -\int p(x)dx + \ln|C_1|.$$

Отсюда, потенцируя, находим общее решение уравнения (11):

$$y = \pm C_1 e^{-\int p(x)\,dx},$$

или

$$y = Ce^{-\int p(x)\,dx}, \tag{12}$$

где $C = \pm C_1$ — произвольная постоянная.

Теперь с помощью (12) найдем общее решение линейного неоднородного уравнения (10). Будем искать общее решение уравнения (10) в виде (12), где C будем считать не постоянной, а функцией от x (в этом смысл метода!), т. е. в виде

$$y = C(x)\,e^{-\int p(x)\,dx}, \tag{13}$$

где $C(x)$ — новая неизвестная функция от x. Чтобы найти функцию $C(x)$ и тем самым решение в виде (13), подставим функцию (13) в уравнение (10). Получим.

$$C'(x)\,e^{-\int p(x)\,dx} - C(x)\,p(x)\,e^{-\int p(x)\,dx} +$$
$$+ p(x)\,C(x)\,e^{-\int p(x)\,dx} = f(x),$$

или

$$C'(x) = f(x)\,e^{\int p(x)\,dx}. \tag{14}$$

Чтобы соотношение (13) было решением уравнения (10), очевидно, функция $C(x)$ должна удовлетворять уравнению (14). Интегрируя его, находим

$$C(x) = \int f(x)\,e^{\int p(x)\,dx}\,dx + C_1,$$

где C_1 — произвольная постоянная. Подставляя найденное выражение в соотношение (13), получаем общее решение линейного уравнения (10):

$$y(x) = C_1 e^{-\int p(x)\,dx} + e^{-\int p(x)\,dx}\int f(x)\,e^{\int p(x)\,dx}\,dx. \tag{15}$$

Заметим, что при решении конкретных примеров проще повторять каждый раз все приведенные выше выкладки, чем пользоваться громоздкой формулой (15).

Пример 5. Найти общее решение уравнения $y' + 3y = e^{2x}$.

Решение. Данное уравнение является линейным. Здесь $p(x) = 3$, $f(x) = e^{2x}$. Решаем сначала соответствующее однородное уравнение $y'+3y=0$. Разделяя переменные $dy/y=-3dx$ и интегрируя, находим

$$\ln|y|=-3x+\ln|C_1|, \text{ или } y=\pm C_1e^{-3x}=Ce^{-3x}.$$

Ищем общее решение данного уравнения в виде $y=C(x)e^{-3x}$. Дифференцируя, находим $y'=C'(x)e^{-3x}-3C(x)e^{-3x}$. Подставляя в данное уравнение выражения для y и y', получаем

$$C'(x)e^{-3x}=e^{2x},\ c'(x)=e^{5x}, \text{ или } dC=e^{5x}dx,$$

откуда $C(x)=1/5e^{5x}+C_2$, где C_2 — произвольная постоянная. Следовательно, общее решение данного уравнения имеет вид

$$y=C(x)e^{-3x}=(1/5e^{5x}+C_2)e^{-3x} \text{ или } y=1/5e^{2x}+C_2e^{-3x}.$$

7. Уравнение в полных дифференциалах

Определение 7. *Уравнение вида*

$$P(x, y)dx+Q(x, y)dy=0, \tag{16}$$

где левая часть представляет собой полный дифференциал некоторой функции $F(x, y)$ в некоторой области G, называется уравнением в полных дифференциалах.

Если уравнение (16) является уравнением в полных дифференциалах, то его можно записать следующим образом:

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

где $F(x, y)$ — такая функция, что $dF(x, y)=P(x, y)dx+Q(x, y)dy$. Отсюда следует, что общее решение уравнения (16) в неявном виде определяется уравнением

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

где C — произвольная постоянная. Действительно, если $y=\varphi(x)$ является решением уравнения (16), то $dF(x, \varphi(x))\equiv 0$, т. е. $F(x, \varphi(x))=C$, и, наоборот, для любой функции $y=\varphi(x)$, обращающей в тождество уравнение $F(x, y)=C$, получим $dF(x, \varphi(x))=0$, т. е. $y=\varphi(x)$ — решение уравнения (16).

Таким образом, нахождение решения уравнения (16) сводится к отысканию такой функции $F(x, y)$, дифференциал которой равен $P(x, y)dx+Q(x, y)dy$.

Как известно (см. теорему 13.7), для того чтобы выражение $P(x, y)dx+Q(x, y)dy$ было полным дифференциалом некоторой функции $F(x, y)$, необходимо и достаточно, чтобы

$$\partial P/\partial y=\partial Q/\partial x. \tag{17}$$

Допустим, что условие (17) выполнено. Тогда существует функция $F(x, y)$ такая, что

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = P(x, y)\,dx + Q(x, y)\,dy.$$

Отсюда

$$\partial F/\partial x = P(x, y), \quad \partial F/\partial y = Q(x, y). \tag{18}$$

Из соотношения $\partial F/\partial x = P(x, y)$ находим

$$F(x, y) = \int P(x, y)\,dx + C(y), \tag{19}$$

где $C(y)$ — произвольная функция от y. Теперь подберем функцию $C(y)$ так, чтобы выполнялось второе из соотношений (18). Для этого продифференцируем найденную функцию $F(x, y)$ по y и результат приравняем $Q(x, y)$:

$$\frac{\partial}{\partial y}\left(\int P(x, y)\,dx\right) + C'(y) = Q(x, y). \tag{20}$$

Из полученного уравнения (20) определяем $C'(y)$ и, интегрируя, находим $C(y)$. Подставляя найденную функцию $C(y)$ в соотношение (19), получаем в неявном виде общее решение уравнения (16).

Чтобы выделить из общего решения частное, удовлетворяющее начальным условиям $x = x_0$, $y = y_0$, надо в общем решении $F(x, y) = C$ x и y заменить начальными значениями. Тогда $C = F(x_0, y_0)$, и $F(x, y) = F(x_0, y_0)$ будет искомым частным решением.

Пример 6. Найти общее решение уравнения $(x+y+1)dx+(x-y^2+3)dy=0$ и выделить из него частное решение, удовлетворяющее начальным условиям $x_0 = 1$, $y_0 = 0$.

Решение. Здесь $P(x, y) = x+y+1$, $Q(x, y) = x-y^2+3$. Так как

$$\partial(x+y+1)/\partial y = 1 = \partial(x-y^2+3)/\partial x,$$

то выражение $(x+y+1)dx+(x-y^2+3)dy$ является полным дифференциалом некоторой функции $F(x, y)$. Имеем

$$F(x, y) = \int P(x, y)\,dx + C(y) = \int (x+y+1)\,dx + C(y) =$$
$$= x^2/2 + xy + x + C(y). \tag{21}$$

Ищем функцию $C(y)$ по формуле (20):

$$\frac{\partial}{\partial y}\left(\frac{x^2}{2}+xy+x+C(y)\right)=x-y^2+3,$$

$$C'(y)=-y^2+3,\ dC/dy=-y^2+3,\ dC=(-y^2+3)\,dy,$$

$$C(y)=\int(-y^2+3)\,dy+C_1=-y^3/3+3y+C_1.$$

Подставляя найденное $C(y)$ в (21), получаем

$$F(x,\ y)=x^2/2+xy+x-y^3/3+3y+C_1.$$

Данное уравнение принимает вид $dF(x,\ y)=0$, а его общее решение определяется уравнением

$$F(x,\ y)=C_2,\ \text{или}\ x^2/2+xy+x-y^3/3+3y+C_1=C_2.$$

Объединяя произвольные постоянные C_1 и C_2 в одну, получаем окончательное уравнение, определяющее неявным образом общее решение исходного дифференциального уравнения:

$$3x^2+6xy+6x-2y^3+18y=C_3.$$

Найдем теперь значение постоянной C_3, при котором частное решение удовлетворяет заданным начальным условиям: $3\cdot1+6\cdot1\cdot0+6\cdot1-2\cdot0+18\cdot0=C_3$, отсюда $C_3=9$, и искомое частное решение определяется уравнением

$$3x^2+6xy+6x-2y^3+18y=9.$$

8. Приближенное решение дифференциальных уравнений первого порядка методом Эйлера. Мы познакомились с несколькими из способов нахождения точных решений дифференциальных уравнений первого порядка. А как быть, если ни один из них не приводит к цели или требует сложных вычислений? В таких случаях прибегают к приближенным методам решений уравнений. Здесь мы познакомимся с простейшим из них, с так называемым *методом Эйлера.*

Суть этого метода состоит в том, что искомая интегральная кривая, являющаяся графиком частного решения, приближенно заменяется ломаной.

Пусть даны дифференциальное уравнение

$$y'=f(x,\ y)$$

и начальные условия $y|_{x=x_0}=y_0$. Найдем приближенно решение уравнения на отрезке $[x_0, b]$, удовлетворяющее заданным начальным условиям.

Разобьем отрезок $[x_0, b]$ точками $x_0<x_1<x_2<\ldots<x_n=b$ на n равных частей. Обозначим $x_1-x_0=x_2-x_1=\ldots=x_n-x_{n-1}=\Delta x$. Через y_i обозначим приближенные значения искомого решения в точках x_i $(i=1, 2, \ldots, n)$. Проведем через точки разбиения x_i прямые (рис. 103), параллельные оси Oy, и последовательно проделаем следующие однотипные операции.

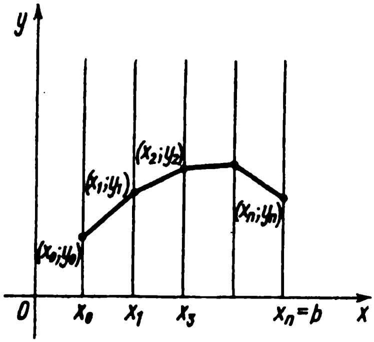


Рис. 103

Подставляем значения x_0 и y_0 в правую часть уравнения $y'=f(x, y)$ и вычисляем угловой коэффициент $y'=f(x_0, y_0)$ касательной к интегральной кривой в точке $(x_0; y_0)$. Для вычисления приближенного значения y_1 искомого решения заменяем на отрезке $[x_0, x_1]$ интегральную кривую отрезком ее касательной в точке $(x_0; y_0)$. При этом получаем

$$y_1-y_0=f(x_0, y_0)(x_1-x_0),$$

откуда, так как величины x_0, x_1, y_0 известны, находим

$$y_1=y_0+f(x_0, y_0)(x_1-x_0), \text{ или } y_1=y_0+f(x_0, y_0)\Delta x.$$

Подставляя значения x_1 и y_1 в правую часть уравнения $y'=f(x, y)$, вычисляем угловой коэффициент $y'=f(x_1, y_1)$ касательной к интегральной кривой в точке $(x_1; y_1)$. Далее, заменяя на отрезке $[x_1, x_2]$ интегральную кривую отрезком ее касательной, находим приближенное значение решения y_2 в точке x_2:

$$y_2=y_1+f(x_1, y_1)(x_2-x_1), \text{ или } y_2=y_1+f(x_1, y_1)\Delta x.$$

В этом равенстве известными величинами являются x_1, y_1, x_2, а y_2 определяется через них.

Аналогично вычисляем

$$y_3=y_2+f(x_2, y_2)\Delta x,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$y_n=y_{n-1}+f(x_{n-1}, y_{n-1})\Delta x.$$

Таким образом, мы приближенно построили искомую интегральную кривую в виде ломаной и получили приближенные значения y_i искомого решения в точках x_i. При этом значения y_i вычисляются по формуле

$$y_i = y_{i-1} + f(x_{i-1}, y_{i-1})\Delta x \quad (i=1, 2, \ldots, n). \qquad (22)$$

Формула (22) есть *основная* расчетная формула метода Эйлера. Ее точность тем выше, чем меньше разность Δx.

Следует отметить, что степень точности метода Эйлера, вообще говоря, невелика. Существуют гораздо более точные методы приближенного решения дифференциальных уравнений. С ними можно познакомиться в специальных курсах.

Пример 7. Найти приближенное решение уравнения $y' = y + x$ на отрезке $[0, 1]$, удовлетворяющее начальным условиям $x_0 = 0$, $y_0 = 1$, и вычислить y при $x = 1$.

Решение. Разделим отрезок $[0, 1]$ на 10 равных частей точками $x_0 = 0$, $x_1 = 0,1$, $x_2 = 0,2, \ldots$, $x_{10} = 1$. Обозначим через $y_1, y_2, \ldots, y_{10}$ приближенные значения решения, которые будем искать по формуле (22). Имеем

$$y_1 = 1 + (1+0)\cdot 0,1 = 1,1,$$
$$y_2 = 1,1 + (1,1+0,1)\cdot 0,1 = 1,22.$$

Аналогично находятся остальные значения y, причем результаты вычисления удобно расположить в виде таблицы, заполняя последовательно одну строку за другой.

x	y	$f(x, y)\,\Delta x$
$x_0 = 0$	$y_0 = 1$	0,1
$x_1 = 0,1$	$y_1 = 1,1$	0,12
$x_2 = 0,2$	$y_2 = 1,22$	0,142
$x_3 = 0,3$	$y_3 = 1,36$	0,1662
$x_4 = 0,4$	$y_4 = 1,5282$	0,1928
$x_5 = 0,5$	$y_5 = 1,7210$	0,2221
$x_6 = 0,6$	$y_6 = 1,9431$	0,2543
$x_7 = 0,7$	$y_7 = 2,1974$	0,2897
$x_8 = 0,8$	$y_8 = 2,4871$	0,3287
$x_9 = 0,9$	$y_9 = 2,8158$	0,3715
$x_{10} = 1$	$y_{10} = 3,1873$	

Второй столбец таблицы содержит приближенные значения y_i искомого решения данного уравнения на $[0, 1]$, удовлетворяющего заданным начальным условиям.

Приближенное значение функции y при $x=1$: $y_{10}=$ $=3{,}1873$.

Чтобы сравнить приближенный результат с точным, найдем точное решение данного уравнения при тех же начальных условиях. Так как уравнение линейное, то применяем метод вариации постоянной. Вначале находим общее решение однородного уравнения:

$$dy/dx=y,\ dy/y=dx,\ \ln|y|=x+\ln|C|,\ y=Ce^x.$$

Варьируем постоянную: $y=C(x)e^x$, $y'=C'(x)e^x+C(x)e^x$ и, подставляя в данное уравнение, получим

$$C'(x)e^x+C(x)e^x=C(x)e^x+x,$$
$$C'(x)=xe^{-x},\ C(x)=-xe^{-x}-e^{-x}+C_1,$$
$$y=C_1e^x-x-1$$

— общее решение данного уравнения. Подставляя вместо x и y начальные условия $x_0=0$, $y_0=1$, находим $C_1=2$, и точное решение данного уравнения, удовлетворяющее заданным начальным условиям, $y=2e^x-x-1$. Значение точного решения при $x=1$ равно: $y(1)=$ $=2(e-1)\approx3{,}4366$. Сравнивая с приближенным значением, видим, что абсолютная ошибка не больше 0,2293.

9. Некоторые применения дифференциальных уравнений первого порядка. К дифференциальным уравнениям первого порядка приводят различные задачи из физики. Основную трудность при решении таких задач представляет составление самих дифференциальных уравнений. Здесь нет универсального метода. Каждая задача требует индивидуального подхода, основанного на глубоком понимании соответствующего закона физики и умения переводить физические задачи на математический язык.

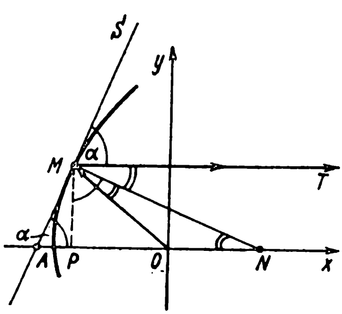


Рис. 104

Рассмотрим несколько таких задач.

1°. *Задача о прожекторе.* Определить форму зеркала, обладающего тем свойством, чтобы все лучи, выходящие из источника света, помещенного в точке O на оси вращения, отражались бы зеркалом параллельно этой оси.

Для решения задачи будем рассматривать плоское сечение зеркала, проходящее через ось вращения. Поместим источник света в начале координат, и пусть ось вращения совпадает с осью Ox (рис. 104). Обозначим через α угол, образованный осью Ox и касательной AS в произвольной точке сечения $M(x;\ y)$. Наша цель — найти форму сечения, т. е. зависимость координаты y от координаты x: $y=y(x)$. Ломаная OMT изображает путь луча, исходящего из источника света в точке O и отражающегося в точке M от поверхности зеркала параллельно оси Ox. Проведем нормаль MN и опустим из точки M на ось Ox перпендикуляр MP. Так как $\angle AMO=\angle TMS$ (угол падения равен углу отражения), имеем $\angle OMN=\angle NMT=\angle ONM$. Следовательно, ΔNOM — равнобедренный и поэтому $OM=ON$. Кроме этого, по построению $\angle PMN=\alpha$. Можно переходить к составлению дифференциального уравнения:

$$ON=PN-PO,\ PN=y\operatorname{tg}\alpha,\ PO=-x,$$

$$ON=OM=\sqrt{x^2+y^2},$$

$$ON=y\operatorname{tg}\alpha+x=\sqrt{x^2+y^2}.$$

Учитывая геометрический смысл производной: $dy/dx=\operatorname{tg}\alpha$, получаем для определения зависимости y от x дифференциальное уравнение первого порядка

$$x+dy/dx\,y=\sqrt{x^2+y^2}.$$

Для нахождения решения уравнения преобразуем его следующим образом. Умножаем обе части равенства на $2dx$:

$$2xdx+2ydy=2\sqrt{x^2+y^2}\,dx \text{ или } d(x^2+y^2)=2\sqrt{x^2+y^2}\,dx.$$

Подстановкой $z=x^2+y^2$ приводим уравнение к уравнению с разделяющимися переменными:

$$dz=2\sqrt{z}\,dx,$$

которое преобразуется к виду

$$z^{-1/2}\,dz = 2dx.$$

Отсюда находим

$$\sqrt{z} = x + C.$$

Заменяя переменную z ее выражением через x и y, получаем

$$\sqrt{x^2+y^2} = x + C.$$

Упрощая полученное уравнение возведением в квадрат обеих его частей, получаем

$$y^2 = 2C(x + C/2).$$

Таким образом, искомая кривая — парабола с параметром $p = C$ и вершиной, лежащей на расстоянии $C/2$ влево от начала координат (рис. 105). Следовательно, искомая отражательная поверхность — *параболоид вращения.*

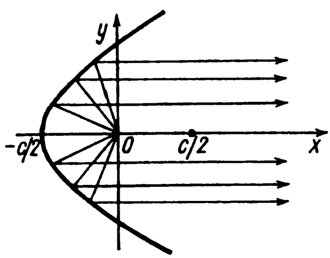


Рис. 105

2°. *Задача о радиоактивном распаде.* Экспериментальным путем установлено, что скорость распада радиоактивного вещества прямо пропорциональна его количеству. Найдем закон изменения массы m радиоактивного вещества в зависимости от времени t, считая, что начальное количество вещества при $t=0$ было m_0.

Сначала определим скорость радиоактивного распада как скорость изменения массы m в зависимости от времени t. Пусть в момент времени t была масса m, в момент времени $t+\Delta t$ — масса $m+\Delta m$. За время Δt распалась масса Δm. Отношение $\Delta m/\Delta t$ есть средняя скорость распада, а $\lim\limits_{\Delta t \to 0} \dfrac{\Delta m}{\Delta t} = \dfrac{dm}{dt}$ есть мгновенная скорость распада в момент времени t. Но по условию задачи

$$dm/dt = -km, \tag{23}$$

где k — коэффициент пропорциональности (знак «минус» поставлен потому, что масса вещества убывает, а производная убывающей функции отрицательна). Получили дифференциальное уравнение первого порядка, из которого надо найти значение массы m в зависимости от времени t.

Решая уравнение, получим

$$dm/m = -kdt,\ \ln m = -kt + \ln C,$$

откуда

$$m = Ce^{-kt}. \tag{24}$$

Формула (24) и дает значение массы вещества как функцию времени. В условиях нашей задачи произвольная постоянная C имеет вполне определенное значение, а именно, при $t=0$ $m_0 = Ce^0 = C$. Подставляя это значение C в формулу (24), получим искомую зависимость массы радиоактивного вещества от времени:

$$m = m_0 e^{-kt}. \tag{25}$$

Заметим, что равенство (24) представляет собой общее решение дифференциального уравнения, а равенство (25) — частное решение, отвечающее начальным условиям данной задачи.

Коэффициент k определяется экспериментально. Например, для радия $k \approx 0{,}000447$. Промежуток времени T, за который распадается половина первоначальной массы радиоактивного вещества, называют «периодом полураспада» этого вещества. Подставляя в формулу (25) вместо m значение $m_0/2$, вместо k — значение 0,000447, получим уравнение для определения периода полураспада T для радия:

$$m_0/2 = m_0 e^{-0{,}000447T},\ -0{,}000447T = -\ln 2,$$

откуда

$$T = \ln 2/0{,}000447 \approx 1550 \text{ лет}.$$

3°. *Задача о законе «естественного роста».* Закон «естественного роста» — это закон, при котором скорость роста вещества прямо пропорциональна его количеству. Найдем формулу для определения изменения количества вещества y в зависимости от времени t, считая, что в начальный момент при $t=0$ количество вещества было y_0.

Здесь независимой переменной является время t, а искомой функцией — количество вещества в любой момент времени. Скорость роста вещества есть скорость изменения величины y в зависимости от переменной t.

Используя, как и в предыдущей задаче, физический смысл производной, можно записать закон «естественного роста» следующим образом:

$$dy/dt = ky, \tag{26}$$

где $k>0$ — коэффициент пропорциональности. Уравнение (26) (отличается от уравнения (23) лишь знаком правой части) описывает многие процессы «размножения».

Решение уравнения (26), удовлетворяющее заданным начальным условиям $t_0=0$, $y=y_0$, имеет вид

$$y = y_0 e^{kt}. \tag{27}$$

Формула (27) и является формулой, выражающей закон «естественного роста». По этому закону, например, происходит «размножение» числа нейтронов в ядерных реакциях, размножение числа бактерий, рост кристаллов, населения и т. п.

На этом закончим изучение дифференциальных уравнений первого рода.

§ 2. Дифференциальные уравнения второго порядка

1. Основные понятия. Определение. *Уравнение вида*

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

где x — независимая переменная, y — искомая функция, y' и y'' — ее производные, называется дифференциальным уравнением второго порядка.

Обычно изучают уравнения, которые могут быть записаны в виде, разрешенном относительно второй производной:

$$y'' = f(x, y, y'). \tag{1}$$

Так же, как и для дифференциального уравнения первого порядка, решением уравнения (1) называется всякая функция $y=\varphi(x)$, которая при подстановке в уравнение обращает его в тождество. График решения также называется интегральной кривой.

Для уравнения второго порядка имеет место теорема существования и единственности решения (теорема Коши), аналогичная соответствующей теореме для уравнения первого порядка, которую мы приведем без доказательства.

Теорема 15.2 (Теорема Коши). *Если в уравнении* $y''=f(x, y, y')$ *функция* $f(x, y, y')$ *и ее частные производные* $f'_y(x, y, y')$ *и* $f'_{y'}(x, y, y')$ *определены и непрерывны в некоторой области* G *пространства переменных:* $(x; y; y')$, *то, какова бы ни была внутренняя точка* $(x_0; y_0; y'_0)$ *области* G, *существует единственное решение* $y=\varphi(x)$ *данного уравнения, удовлетворяющее условиям*

$$y=y_0, \ y'=y'_0 \ \text{при} \ x=x_0. \tag{2}$$

Геометрически это означает, что через заданную точку $(x_0; y_0)$ плоскости с заданным угловым коэффициентом y'_0 касательной проходит единственная интегральная кривая.

Условия (2) называются *начальными условиями* решения и часто записываются в виде

$$y|_{x=x_0} = y_0, \ y'|_{x=x_0} = y'_0. \tag{3}$$

Как и для уравнения первого порядка, задачу отыскания решения по заданным начальным условиям называют *задачей Коши.*

Дадим теперь определения общего и частного решений уравнения (1), удовлетворяющего условиям теоремы Коши.

Функция $y=\varphi(x, C_1, C_2)$, зависящая от x и двух произвольных постоянных C_1 и C_2, называется *общим решением* уравнения (1) в некоторой области G, если она является решением уравнения (1) при любых значениях постоянных C_1 и C_2 и если при любых начальных условиях (3) существуют единственные значения постоянных $C_1=C_1^\circ$, $C_2=C_2^\circ$ такие, что функция $y=\varphi(x, C_1^\circ, C_2^\circ)$ удовлетворяет данным начальным условиям.

Всякая функция $y=\varphi(x, C_1^\circ, C_2^\circ)$, получающаяся из общего решения $y=\varphi(x, C_1, C_2)$ уравнения (1) при определенных значениях постоянных $C_1=C_1^\circ$, $C_2=C_2^\circ$, называется *частным решением.*

Обратимся к *примеру*. Рассмотрим уравнение $y''=2$. Данное уравнение есть уравнение второго порядка. Так как функции

$$f(x,\ y,\ y')=2,\ f'_y(x,\ y,\ y')=0 \text{ и } f'_{y'}(x,\ y,\ y')=0$$

определены и непрерывны во всем пространстве $(x;\ y;\ y')$, оно удовлетворяет всем требованиям теоремы Коши.

Общее решение данного уравнения найдем путем его двукратного последовательного интегрирования. Последовательно интегрируя, находим сначала первую производную $y'=2x+C_1$, а затем и общее решение

$$y=x^2+C_1x+C_2,$$

где C_1 и C_2 — произвольные постоянные. Геометрически общее решение представляет собой семейство парабол, причем так как оно зависит от двух произвольных постоянных, то через каждую точку плоскости проходит бесконечное множество парабол, имеющих различные касательные. Поэтому для выделения одной параболы из их множества помимо точки $(x_0;\ y_0)$, через которую проходят параболы, задают еще угловой коэффициент y'_0 касательной к искомой кривой.

Найдем, например, частное решение данного уравнения при начальных условиях $y|_{x=1}=1$, $y'|_{x=1}=1$. Подставляя эти значения в выражения для общего решения $y=x^2+C_1x+C_2$ и его производной $y'=2x+C_1$, получим для определения C_1 и C_2 систему уравнений

$$\begin{cases} 1=1+C_1+C_2, \\ 1=2+C_1, \end{cases}$$

откуда находим $C_1=-1$ и $C_2=1$. Следовательно, искомым частным решением будет

$$y=x^2-x+1$$

— парабола, проходящая через точку $(1;1)$ с угловым коэффициентом, равным единице.

2. Уравнения второго порядка, допускающие понижение порядка. Отметим три частных вида уравнения (1), когда решение его с помощью замены переменной сводится к решению уравнения первого порядка. Такое преобразование уравнения (1) называется *понижением порядка*. Рассмотрим эти частные случаи.

1°. *Уравнение вида* $y''=f(x)$. Уравнение не содержит y и y'. Введем новую функцию $z(x)$, положив $z(x)=y'$. Тогда $z'(x)=y''$, и уравнение превращается в уравнение первого порядка: $z'(x)=f(x)$ с искомой функцией $z(x)$. Решая его, находим $z(x)=\int f(x)dx+C_1$ или, так как $z(x)=y'$, $y'=\int f(x)dx+C_1$. Отсюда, интегрируя еще раз, получим искомое решение:

$$y=\int\left[\int f(x)dx\right]dx+C_1x+C_2,$$

где C_1 и C_2 — произвольные постоянные.

Пример 1. Найти общее решение уравнения $y''=x$.

Решение. Полагая $z(x)=y'$, получаем уравнение первого порядка $z'(x)=x$. Интегрируя его, находим: $z(x)=x^2/2+C_1$. Заменяя $z(x)$ на y' и интегрируя еще раз, находим искомое общее решение:

$$y=\int[x^2/2+C_1]dx+C_2=x^3/6+C_1x+C_2.$$

2°. *Уравнение вида* $y''=f(x, y')$. Уравнение не содержит y. Положим, как и в предыдущем случае, $z(x)=y'$, тогда $z'(x)=y''$, и уравнение преобразуется в уравнение первого порядка относительно $z(x)$: $z'=f(x, z)$. Решив его, находим $z(x)=\varphi(x, C_1)$ или, так как $z(x)=y'$, $y'=\varphi(x, C_1)$. Отсюда, интегрируя еще раз, получим искомое решение:

$$y=\int\varphi(x, C_1)dx+C_2,$$

где C_1 и C_2 — произвольные постоянные.

Пример 2. Найти общее решение уравнения $y''-3y'/x=x$.

Решение. Положив $z(x)=y'$ и $z'(x)=y''$, получим линейное уравнение первого порядка $z'-3z/x=x$. Решив его, найдем $z(x)=C_1x^3-x^2$. Тогда $y'=C_1x^3-x^2$ и $y=C_1x^4/4-x^3/3+C_2$.

3°. *Уравнение вида* $y''=f(y, y')$. Уравнение не содержит x. В этом случае вводим новую функцию $z(y)$, полагая $y'=z$. Тогда

$$y''=\frac{d(y')}{dx}=\frac{dy'}{dy}\cdot\frac{dy}{dx}=\frac{dz}{dy}\cdot\frac{dy}{dx}=\frac{dz}{dy}z(y).$$

Подставим в уравнение выражения y' и y'', получим уравнение первого порядка относительно z как функции от y:

$$z\frac{dz}{dy}=f(y,z).$$

Решив его, находим $z=\varphi(y, C_1)$ или, так как $z=dy/dx$, $dy/dx=\varphi(y, C_1)$. Отсюда $dy/\varphi(y, C_1)=dx$. Получили уравнение с разделяющимися переменными, из которого находим общее решение данного уравнения:

$$\int\frac{dy}{\varphi(y,C_1)}=x+C_2,$$

где C_1 и C_2 — произвольные постоянные.

Пример 3. Найти общее решение уравнения

$$yy''-2y'^2=0.$$

Решение. Полагая $y'=z(y)$ и $y''=z\frac{dz}{dy}$, получим $zy\frac{dz}{dy}-2z^2=0$. Это уравнение первого порядка с разделяющимися переменными. Приводя его к виду $dz/z=2dy/y$ и интегрируя, получим $\ln|z|=2\ln|y|+\ln|C_1|$, т. е. $z=C_1y^2$. Вспоминая, что $z=dy/dx$, находим $dy/y^2=C_1dx$, откуда приходим к искомому решению:

$$-1/y=C_1x+C_2, \text{ или } y=-1/(C_1x+C_2).$$

Следует заметить, что при сокращении на z было потеряно решение уравнения $z=y'=0$, т. е. $y=C=\text{const}$. В данном случае оно содержится в общем решении, так как получается из него при $C_1=0$ (за исключением решения $y=0$).

3. Дифференциальные уравнения высших порядков. Здесь мы ограничимся только основными определениями и общими замечаниями, относящимися к дифференциальным уравнениям n-го порядка.

Дифференциальные уравнения n-го порядка имеют вид

$$F(x, y, y', \ldots, y^{(n)})=0,$$

или, если они разрешены относительно старшей производной,

$$y^{(n)}=f(x, y, y', \ldots, y^{(n-1)}). \tag{4}$$

Решением уравнения (4), как и для уравнений 1-го и 2-го порядков, называется *всякая функция* $y=\varphi(x)$, которая при подстановке в уравнение обращает его в тождество.

Теорема существования и единственности решения уравнения (4) аналогична соответствующим теоремам, приведенным ранее для случаев $n=1$ и $n=2$.

Общее решение уравнения (4) зависит от x и n произвольных постоянных и может быть записано в виде

$$y=\varphi(x,\ C_1,\ C_2, \ldots, C_n).$$

Решения, получающиеся из общего при определенных значениях постоянных $C_1,\ C_2, \ldots, C_n$, называются *частными решениями* уравнения (4). Чтобы выделить частное решение из общего решения уравнения (4), необходимо задать начальные условия:

$$y\,|_{x=x_0}=y_0,\ y'\,|_{x=x_0}=y_0',\ y''\,|_{x=x_0}=y_0'', \ldots,\ y^{(n-1)}\,|_{x=x_0}=y_0^{(n-1)}. \tag{5}$$

Отыскание решения уравнения (4), удовлетворяющего заданным начальным условиям (5), называется *решением задачи Коши* для этого уравнения.

Простейшим уравнением (4) является уравнение, в котором правая часть зависит только от x, т. е. уравнение вида

$$y^{(n)}=f(x). \tag{6}$$

Это уравнение легко решается. Действительно, производя последовательно n интегрирований, получим

$$y^{(n-1)}=\int f(x)\,dx+C_1,$$

$$y^{(n-2)}=\int\left[\int f(x)\,dx+C_1\right]dx+C_2=\int dx\int f(x)\,dx+C_1x+C_2,$$

. . . .

$$y=\int dx\int dx \ldots \int f(x)\,dx+C_1\frac{x^{n-1}}{(n-1)!}+C_2\frac{x^{n-2}}{(n-2)!}+\ldots+C_n, \tag{7}$$

где $C_1,\ C_2, \ldots, C_n$ — произвольные постоянные. Функция (7) и является общим решением уравнения (6).

Пример 4. Найти общее решение уравнения третьего порядка $y'''=e^x$ и выделить из него частное решение, удовлетворяющее начальным условиям

$$y\,|_{x=0}=0,\ y'\,|_{x=0}=0,\ y''\,|_{x=0}=1.$$

Решение. Последовательно интегрируя, находим $y''=e^x+C_1$, $y'=e^x+C_1x+C_2$. Интегрируя еще раз, получим общее решение данного уравнения:

$$y=e^x+C_1\frac{x^2}{2}+C_2x+C_3,$$

где C_1, C_2, C_3 — произвольные постоянные.

Подставляя в выражение для y, y', y'' начальные условия, получим $0=1+C_3$, $0=1+C_2$, $1=1+C_1$, откуда находим $C_1=0$, $C_2=-1$, $C_3=-1$ и искомое частное решение $y=e^x-x-1$.

§ 3. Линейные дифференциальные уравнения второго порядка

Линейные дифференциальные уравнения второго порядка в теории дифференциальных уравнений занимают важное место не только потому, что они представляют простой и хорошо изученный тип уравнений, но и в связи с тем, что значительное количество практических задач физики, механики, техники и особенно электротехники приводят к решению этих уравнений.

1. Основные понятия. Определение. *Уравнение вида*

$$y''+p(x)y'+q(x)y=f(x), \tag{1}$$

где y — искомая функция, а $p(x)$, $q(x)$ и $f(x)$ — непрерывные функции на некотором интервале (a, b), называется линейным дифференциальным уравнением второго порядка.

Если $f(x)\equiv 0$, то уравнение (1) называется *линейным однородным* уравнением. Если же $f(x)\not\equiv 0$, то уравнение (1) называется *линейным неоднородным* уравнением.

Разрешая уравнение (1) относительно y'': $y''=-p(x)y'-q(x)y+f(x)$, видим, что оно является частным случаем уравнения $y''=f(x, y, y')$. Следовательно, уравнение (1) удовлетворяет условиям теоремы о существовании и единственности решения. Действительно, функция $f(x, y, y')=-p(x)y'-q(x)y+f(x)$ непрерывна как функция трех переменных x, y и y' (она зависит от y и y' линейно, а функции $p(x)$, $q(x)$ и $f(x)$ непрерывны по условию), частные производные $f'_y(x, y, y')=-q(x)$ и $f'_{y'}(x, y, y')=-p(x)$ также являются непрерывными функциями трех переменных x, y и y' (от y и y' $p(x)$ и $q(x)$ не зависят, а как функции x непрерывны по условию). Поэтому при любых начальных условиях $y|_{x=x_0}=y_0$, $y'|_{x=x_0}=y'_0$, где $x_0 \in (a, b)$, уравнение (1) имеет единственное решение задачи Коши, удовлетворяющее этим условиям.

Изучение линейных дифференциальных уравнений мы начнем с однородных уравнений.

2. Линейные однородные дифференциальные уравнения второго порядка. Рассмотрим некоторые свойства решений линейных однородных уравнений.

Теорема 15.3. *Если функции $y_1(x)$ и $y_2(x)$ — решения уравнения*

$$y''+p(x)y'+q(x)y=0, \qquad (2)$$

то функция $y=C_1y_1(x)+C_2y_2(x)$ при любых значениях постоянных C_1 и C_2 также является решением уравнения (2).

Доказательство. Продифференцировав дважды функцию $y=C_1y_1(x)+C_2y_2(x)$ и подставив выражения для y, y' и y'' в левую часть уравнения (2), получим

$$C_1y_1''(x)+C_2y_2''(x)+p(x)(C_1y_1'(x)+C_2y_2'(x))+$$
$$+q(x)(C_1y_1(x)+C_2y_2(x))=C_1[y_1''(x)+p(x)y_1'(x)+$$
$$+q(x)y_1(x)]+C_2[y_2''(x)+p(x)y_2'(x)+q(x)y_2(x)].$$

Так как функции $y_1(x)$ и $y_2(x)$ по условию являются решениями уравнения (2), то выражения в квадратных скобках тождественно равны нулю, а это значит, что функция $y=C_1y_1(x)+C_2y_2(x)$ является решением уравнения (2) ■

Итак, доказано, что функция вида $y=C_1y_1(x)+$ $+C_2y_2(x)$ с произвольными постоянными C_1 и C_2 является решением уравнения (2). Естественно возникает вопрос, не является ли это решение общим решением уравнения (2). Будет доказано, что при некоторых условиях функция $y=C_1y_1(x)+C_2y_2(x)$ является общим решением уравнения (2). Для выяснения этих условий введем понятия линейной зависимости и линейной независимости функций $y_1(x)$ и $y_2(x)$.

Функции $y_1(x)$ и $y_2(x)$ называются *линейно-зависимыми* на (a, b), если существуют такие числа α_1 и α_2, из которых хотя бы одно отлично от нуля, что для любого $x \in (a, b)$ имеет место равенство

$$\alpha_1 y_1(x) + \alpha_2 y_2(x) = 0. \tag{3}$$

Очевидно, что если функции $y_1(x)$ и $y_2(x)$ линейно-зависимы, то они *пропорциональны*. Действительно, если $\alpha_1 y_1(x) + \alpha_2 y_2(x) = 0$, причем $\alpha_1 \neq 0$ и $y_2(x) \neq 0$, то $y_1(x)/y_2(x) = -\alpha_2/\alpha_1 = \text{const}$. Верно и обратное.

Функции $y_1(x)$ и $y_2(x)$ называются *линейно-независимыми* на (a, b), если не существует таких чисел α_1 и α_2, из которых хотя бы одно отлично от нуля, что для любого $x \in (a, b)$ имеет место равенство (3).

Другими словами, равенство (3) выполняется сразу для всех $x \in (a, b)$, только если $\alpha_1 = \alpha_2 = 0$.

Очевидно, что если функции $y_1(x)$ и $y_2(x)$ линейно-независимы, то их отношение $y_1(x)/y_2(x) \neq \text{const}$, т. е. они *не пропорциональны*. Так, например, функции $y_1(x) = x^2$ и $y_2(x) = x^3$ линейно-независимы на любом интервале (a, b), так как $y_1(x)/y_2(x) = 1/x \neq \text{const}$, а функции $y_1(x) = 4x^2$ и $y_2(x) = x^2$ линейно-зависимы в любом промежутке, так как $y_1(x)/y_2(x) = 4 = \text{const}$.

Предположим теперь, что функции $y_1(x)$ и $y_2(x)$ являются решениями уравнения (2). Как узнать, являются ли они линейно-зависимыми или линейно-независимыми. Удобным аппаратом для исследования этого вопроса является так называемый *определитель Вронского**, составленный из этих решений:

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'. \tag{4}$$

Определитель Вронского (или вронскиан) является функцией, определенной на (a, b), и обозначается $W(y_1, y_2)$ или просто $W(x)$.

Теорема 15.4. *Если функции $y_1(x)$ и $y_2(x)$ линейно-зависимы на (a, b), то определитель Вронского, составленный из них, равен нулю на этом интервале.*

Доказательство. Так как по условию функции $y_1(x)$ и $y_2(x)$ линейно-зависимы, то по определению существуют числа α_1 и α_2, из которых одно обязательно отлично от нуля, такие, что имеет место равенство (3):

---

* Вронский Юзеф (1775—1853) — польский математик.

$\alpha_1 y_1(x)+\alpha_2 y_2(x)=0$. Пусть, например, $\alpha_1 \neq 0$. Тогда из равенства (3) следует, что

$$y_1(x) = -\frac{\alpha_2}{\alpha_1} y_2(x),\ y_1'(x) = -\frac{\alpha_2}{\alpha_1} y_2'(x).$$

Подставляя выражения для $y_1(x)$ и $y'_1(x)$ в определитель Вронского, получаем

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} -\frac{\alpha_2}{\alpha_1} y_2 & y_2 \\ -\frac{\alpha_2}{\alpha_1} y_2' & y_2' \end{vmatrix} = 0 \ \blacksquare$$

Теорема 15.5. *Если решения $y_1(x)$ и $y_2(x)$ уравнения (2) линейно-независимы на (a, b), то определитель Вронского, составленный из них, отличен от нуля на этом интервале.*

Доказательство. Допустим обратное, т. е. предположим, что существует точка $x_0 \in (a, b)$, в которой определитель Вронского $W(x_0)=0$. Составим систему уравнений

$$\begin{cases} \alpha_1 y_1(x_0) + \alpha_2 y_2(x_0) = 0, \\ \alpha_1 y_1'(x_0) + \alpha_2 y_2'(x_0) = 0, \end{cases} \tag{5}$$

в которой α_1 и α_2 — неизвестные числа. Так как определитель этой системы $W(x_0)=0$, то система (5) имеет (см. гл. X, § 4, п. 2) ненулевое решение α_1 и α_2. Рассмотрим функцию

$$y(x)=\alpha_1 y_1(x)+\alpha_2 y_2(x).$$

По теореме 15.3 эта функция является решением уравнения (2). Кроме того, поскольку α_1 и α_2 — решения системы (5), функция $y(x)$ удовлетворяет нулевым начальным условиям

$$y\,|_{x=x_0} = 0,\ y'\,|_{x=x_0} = 0. \tag{6}$$

Таким начальным условиям, очевидно, удовлетворяет решение $y(x)=0$. По теореме о существовании и единственности решения, решение $y(x)=0$ является единственным решением уравнения (2) с начальными условиями (6). Следовательно, $\alpha_1 y_1(x)+\alpha_2 y_2(x)=0$, а это означает, что функции $y_1(x)$ и $y_2(x)$ линейно-зависимы на (a, b), что противоречит условию теоремы. Таким образом, $W(x) \neq 0$ для всех $x \in (a, b)$ ■

Итак, установлено, что если функции $y_1(x)$ *и* $y_2(x)$ являются на (a, b) решениями линейного однородного уравнения (2), то составленный из них определитель Вронского либо равен нулю на (a, b) ($y_1(x)$ и $y_2(x)$ линейно-зависимы), либо отличен от нуля на (a, b) ($y_1(x)$ и $y_2(x)$ линейно-независимы).

Теперь мы, наконец, можем доказать, при каких условиях функция $y = C_1y_1(x) + C_2y_2(x)$ будет общим решением линейного однородного уравнения (2).

Теорема 15.6. *Если функции* $y_1(x)$ и $y_2(x)$ — *линейно-независимые на* (a, b) *решения уравнения* (2), *то функция*

$$y = C_1y_1(x) + C_2y_2(x), \tag{7}$$

где C_1 *и* C_2 — *произвольные постоянные, является общим решением уравнения* (2) *на* (a, b).

Доказательство. Прежде всего напомним, что в силу теоремы 15.3 функция $y = C_1y_1(x) + C_2y_2(x)$ при любых значениях постоянных C_1 и C_2 является решением уравнения (2). Для того чтобы доказать, что эта функция является общим решением уравнения (2), достаточно установить, что из него можно выделить частное решение, удовлетворяющее заданным начальным условиям.

Пусть $x_0 \in (a, b)$ и

$$y|_{x=x_0} = y_0 \text{ и } y'|_{x=x_0} = y_0' \tag{8}$$

— произвольные начальные условия. Покажем, что постоянные C_1 и C_2 можно подобрать так, что решение (7) при этих значениях постоянных будет частным решением, удовлетворяющим заданным начальным условиям (8).

Составим систему уравнений:

$$\begin{cases} y_0 = C_1y_1(x_0) + C_2y_2(x_0), \\ y_0' = C_1y_1'(x_0) + C_2y_2'(x_0), \end{cases} \tag{9}$$

в которой C_1 и C_2 — неизвестные числа. Определитель этой системы есть определитель Вронского $W(x_0)$. Так как по условию функции $y_1(x)$ *и* $y_2(x)$ — линейно-независимы на (a, b), то, в силу теоремы 15.5, $W(x_0) \neq 0$.

Поэтому система (9) имеет единственное решение, которое обозначим $C_1=C^0{}_1$, $C_2=C^0{}_2$. Подставляя $C^0{}_1$ и $C^0{}_2$ в равенство (7), получим искомое частное решение уравнения (2) $y=C^0{}_1y_1(x)+C^0{}_2y_2(x)$, удовлетворяющее условиям (8). Это и означает, что решение (7) является общим решением уравнения (2) ■

Из доказанной теоремы следует, что для отыскания общего решения уравнений (2) достаточно найти два линейно-независимых частных решения и составить выражение (7) с произвольными постоянными.

Пример 1. Рассмотрим уравнение $y''-y=0$.

Имеем линейное однородное уравнение. Легко заметить, что частными решениями будут $y_1(x)=e^x$ и $y_2(x)=e^{-x}$. Так как определитель Вронского

$$W(x)=\begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix}=-2$$

отличен от нуля, то эти решения линейно-независимы на всей числовой прямой. Следовательно, общее решение данного уравнения можно записать в виде $y=C_1e^x+$ $+C_2e^{-x}$, где C_1 и C_2 — произвольные постоянные.

В заключение покажем, как можно найти общее решение уравнения (2), если известно только одно частное решение этого уравнения.

Пусть $y_1(x)$ — частное решение уравнения (2). Введем новую функцию z, полагая $y=y_1z$. Тогда $y'=y'_1z+$ $+y_1z'$, $y''=y''_1z+2y'_1z'+y_1z''$. Подставляя выражения для y, y' и y'' в уравнение (2) и группируя слагаемые, получаем

$$[y_1''+p(x)\,y_1'+q(x)\,y_1]\,z+z'\,[2y_1'+p(x)\,y_1]+y_1z''=0.$$

Так как $y_1(x)$ — решение уравнения (2), то выражение в первых квадратных скобках равно нулю, и уравнение примет вид

$$z'\,[2y_1'+p(x)\,y_1]+y_1z''=0.$$

Порядок этого уравнения можно понизить, положив $z'=u$, где u — новая искомая функция:

$$u\,[2y_1'+p(x)\,y_1]+u'y_1=0.$$

Получили уравнение первого порядка с разделяющимися переменными относительно функции u. Решая его, находим

$$\int \frac{du}{u} = -2\int \frac{dy_1}{y_1} - \int p(x)\,dx,$$

$$\ln|u| = -2\ln|y_1| - \int p(x)\,dx + \ln|C|,$$

$$u = \pm \frac{C}{y_1^2} e^{-\int p(x)dx}, \text{ или } u = \frac{C_1}{y_1^2} e^{-\int p(x)dx},$$

где C_1 — произвольная постоянная. Возвращаясь к переменной z и умножая выражение для z на y_1, получаем общее решение уравнения (2):

$$y = C_1 y_1 \int \frac{1}{y_1^2} e^{-\int p(x)dx}\,dx + C_2 y_1,$$

где C_1 и C_2 — произвольные постоянные.

В качестве примера предлагается самостоятельно найти общее решение только что рассмотренного уравнения $y''-y=0$ (см. пример 1), беря одно из его частных решений за известное.

Заметим, что линейное однородное уравнение вида (2) в общем виде не решается, поэтому рассмотренный выше метод отыскания общего решения представляет определенный интерес.

3. Линейные неоднородные дифференциальные уравнения второго порядка. Рассмотрим теперь основные свойства решений линейного неоднородного дифференциального уравнения второго порядка (1):

$$y''+p(x)y'+q(x)y=f(x).$$

Имеет место

Теорема 15.7. *Общее решение уравнения* (1) *есть сумма любого его частного решения и общего решения соответствующего однородного уравнения.*

Доказательство. *Пусть* $\tilde{y}(x)$ — частное решение уравнения (1), а $Y(x)=C_1y_1(x)+C_2y_2(x)$ — общее решение соответствующего однородного уравнения (2), где C_1 и C_2 — произвольные постоянные. Покажем, что функция

$$y=\tilde{y}(x)+Y(x) \tag{10}$$

является решением уравнения (1). Для этого найдем $y'=\tilde{y}'(x)+Y'(x)$, $y''=\tilde{y}''(x)+Y''(x)$ и подставим их в уравнение (1):

$$\tilde{y}''(x)+Y''(x)+p(x)[\tilde{y}'(x)+Y'(x)]+q(x)[\tilde{y}(x)+ \\ +Y(x)]=[Y''(x)+p(x)Y'(x)+q(x)Y(x)]+ \\ +[\tilde{y}''(x)+p(x)\tilde{y}'(x)+q(x)\tilde{y}(x)]=0+f(x)=f(x).$$

Отсюда следует, что функция $y=\tilde{y}(x)+Y(x)$ действительно является решением уравнения (1).

Покажем теперь, что функция (10) является общим решением уравнения (1). Для этого возьмем решение y уравнения (1) и рассмотрим разность $y-\tilde{y}(x)$. Оказывается, эта разность является решением однородного уравнения (2). Действительно,

$$[y-\tilde{y}(x)]''+p(x)[y-\tilde{y}(x)]'+q(x)[y-\tilde{y}(x)]= \\ =[y''+p(x)y'+q(x)y]-[\tilde{y}''(x)+p(x)\tilde{y}'(x)+ \\ +q(x)\tilde{y}(x)]=f(x)-f(x)=0.$$

Это значит, что разность $y-\tilde{y}(x)$ может быть записана в виде

$$y-\tilde{y}(x)=C_1^0y_1(x)+C_2^0y_2(x),$$

или

$$y=\tilde{y}(x)+C_1^0y_1(x)+C_2^0y_2(x),$$

где $C^0{}_1$ и $C^0{}_2$ — определенные значения постоянных C_1 и C_2. Отсюда заключаем, что любое решение y уравнения (1) получается из решения (10) при соответствующем подборе произвольных постоянных C_1 и C_2 в решении $Y(x)$ т. е. решение (10) является общим решением уравнения (1) $\blacksquare$

Таким образом, чтобы найти общее решение линейного неоднородного уравнения, необходимо знать общее решение соответствующего однородного уравнения и хотя бы одно его частное решение. В общем случае задача отыскания частного решения является сложной. Поэтому мы покажем, как можно найти частное решение неоднородного уравнения методом вариации произвольных постоянных, если известно общее решение соответствующего однородного уравнения.

Пусть $y=C_1y_1(x)+C_2y_2(x)$ — общее решение однородного уравнения. Будем искать частное решение неоднородного уравнения (1) в виде

$$y = C_1(x)y_1(x) + C_2(x)y_2(x), \tag{11}$$

рассматривая C_1 и C_2 как некоторые искомые функции от x. Продифференцируем равенство (11):

$$y' = C_1'(x)y_1(x) + C_1(x)y_1'(x) + C_2'(x)y_2(x) + C_2(x)y_2'(x). \tag{12}$$

Будем подбирать функции $C_1(x)$ и $C_2(x)$ так, чтобы выполнялось равенство

$$C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0. \tag{13}$$

Тогда равенство (12) примет вид

$$y' = C_1(x)y_1'(x) + C_2(x)y_2'(x).$$

Дифференцируя теперь это равенство, найдем y'':

$$y'' = C_1'(x)y_1'(x) + C_1(x)y_1''(x) + C_2'(x)y_2'(x) + C_2(x)y_2''(x).$$

Подставляя выражения для y, y' и y'' в уравнение (1) и группируя слагаемые, получим

$$C_1(x)[y_1''(x) + p(x)y_1'(x) + q(x)y_1(x)] +$$
$$+ C_2(x)[y_2''(x) + p(x)y_2'(x) + q(x)y_2(x)] +$$
$$+ C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = f(x).$$

Выражения, стоящие в квадратных скобках, равны нулю, так как $y_1(x)$ и $y_2(x)$ — решения однородного уравнения. Поэтому последнее равенство принимает вид

$$C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = f(x). \tag{14}$$

Таким образом, функция (11) будет решением уравнения (1), если функции $C_1(x)$ и $C_2(x)$ удовлетворяют уравнениям (13) и (14). Объединяя их, получаем систему уравнений

$$\begin{cases} C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0, \\ C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = f(x), \end{cases} \tag{15}$$

в которой $C'_1(x)$ и $C'_2(x)$ неизвестны, а $y_1(x)$, $y_2(x)$, $y'_1(x)$, $y'_2(x)$ и $f(x)$ известны. Так как определителем этой системы является определитель Вронского

$$W(x)=\begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix},$$

составленный из линейно-независимых решений $y_1(x)$ и $y_2(x)$ однородного уравнения (2), то он по теореме 15.5 не равен нулю, а значит, система (15) имеет единственное решение относительно $C'_1(x)$ и $C'_2(x)$. Решив эту систему, получим

$$C_1'(x)=\varphi_1(x),\ C_2'(x)=\varphi_2(x),$$

где $\varphi_1(x)$ и $\varphi_2(x)$ — известные функции, откуда, интегрируя, найдем $C_1(x)$ и $C_2(x)$. Подставляя полученные выражения $C_1(x)$ и $C_2(x)$ в равенство (11), получим искомое частное решение уравнения (1).

Пример 2. Найти частное решение уравнения $y''-y=x$.

Решение. В примере 1 п. 3 было найдено общее решение $Y(x)=C_1e^x+C_2e^{-x}$ соответствующего однородного уравнения $y''-y=0$. Поэтому частное решение неоднородного уравнения будем искать в виде

$$\tilde{y}(x)=C_1(x)e^x+C_2(x)e^{-x}. \qquad (16)$$

Система (15) для нахождения $C'_1(x)$ и $C'_2(x)$ в данном случае имеет вид

$$\begin{cases} C_1'(x)e^x+C_2'e^{-x}=0, \\ C_1'(x)e^x-C_2'(x)e^{-x}=x. \end{cases}$$

Складывая эти уравнения, находим $C_1'(x)=\frac{1}{2}xe^{-x}$.

Отсюда, интегрируя, получаем

$$C_1(x)=-\frac{1}{2}(x+1)e^{-x}.$$

Произвольную постоянную мы не пишем, так как ищем какое-нибудь частное решение. Подставляя выражение $C'_1(x)$ в первое уравнение системы, находим

$$C_2'(x)=-\frac{1}{2}xe^x,$$

откуда, интегрируя, получаем

$$C_2(x) = -\frac{1}{2}(x-1)e^x.$$

Подставляя найденные выражения $C_1(x)$ и $C_2(x)$ в равенство (16), получим частное решение $\tilde{y}$ данного неоднородного уравнения

$$\tilde{y}(x) = \left[-\frac{1}{2}(x+1)e^{-x}\right]e^x + \left[-\frac{1}{2}(x-1)e^x\right]e^{-x} = -x.$$

Заметим, что, найдя частное решение неоднородного уравнения и зная общее решение соответствующего однородного уравнения, мы на основании равенства (10) можем написать общее решение данного неоднородного уравнения

$$y = \tilde{y}(x) + Y(x) = -x + C_1e^x + C_2e^{-x},$$

где C_1 и C_2 — произвольные постоянные.

§ 4. Линейные дифференциальные уравнения второго порядка с постоянными коэффициентами

В данном параграфе мы рассмотрим важный частный случай линейных дифференциальных уравнений второго порядка, в которых функции $p(x)$ и $q(x)$ являются постоянными величинами. Такие уравнения называются *линейными уравнениями с постоянными коэффициентами.*

1. Линейные однородные дифференциальные уравнения второго порядка с постоянными коэффициентами. Рассмотрим линейное однородное уравнение второго порядка

$$y'' + py' + qy = 0, \tag{1}$$

где p и q — вещественные числа.

Теорема 15.8. *Если число k является корнем уравнения*

$$k^2 + pk + q = 0, \tag{2}$$

то функция $y = e^{kx}$ является решением уравнения (1).

Доказательство. Пусть $y = e^{kx}$ ($k = \text{const}$). Тогда $y' = ke^{kx}$, $y'' = k^2e^{kx}$. Подставляя y, y' и y'' в уравнение (1), получим

$$e^{kx}(k^2 + pk + q) = 0.$$

Так как $e^{kx}\neq 0$, то, сокращая на e^{kx}, получим уравнение

$$k^2+pk+q=0.$$

Следовательно, если k будет корнем уравнения (2), то функция $y=e^{kx}$ будет решением уравнения (1) ■

Уравнение (2) называется *характеристическим* уравнением данного уравнения (1).

Характеристическое уравнение (2) является квадратным уравнением, имеющим два корня. Обозначим их через k_1 и k_2.

Теорема 15.9. *Общее решение уравнения* (1) *может быть записано следующим образом:*

1) *если корни характеристического уравнения вещественные и различные* $(k_1\neq k_2)$, *то общее решение имеет вид*

$$y=C_1e^{k_1x}+C_2e^{k_2x};$$

2) *если корни характеристического уравнения вещественные и равные* $(k_1=k_2)$, *то общее решение имеет вид*

$$y=C_1e^{k_1x}+C_2xe^{k_1x};$$

3) *если корни характеристического уравнения комплексные* $(k_1=\alpha+i\beta,\ k_2=\alpha-i\beta,\ \beta\neq 0)$, *то общее решение имеет вид*

$$y=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x).$$

Доказательство. 1) Пусть корни k_1 и k_2 различны. По теореме 15.8 функции

$$y_1=e^{k_1x} \text{ и } y_2=e^{k_2x}$$

являются частными решениями уравнения (1). Эти решения линейно-независимы, так как

$$\frac{y_2}{y_1}=e^{(k_2-k_1)x}\neq \text{const}\ (k_1\neq k_2).$$

Следовательно, по теореме 15.6 общее решение уравнения (1) имеет вид

$$y=C_1e^{k_1x}+C_2e^{k_2x}.$$

2) Пусть корни k_1 и k_2 равны: $k_1=k_2$. По теореме 15.8 функция $y_1=e^{k_1x}$ является частным решением

уравнения (1). Требуется найти второе частное решение, линейно-независимое с первым. Положим $y_2 = xe^{k_1x}$ и покажем, что эта функция является вторым частным решением уравнения (1). Для этого найдем ее производные: $y_2' = e^{k_1x} + k_1xe^{k_1x}$, $y_2'' = 2k_1e^{k_1x} + k_1^2xe^{k_1x}$. Подставляя y_2, y_2' и y_2'' в левую часть уравнения (1), получаем

$$2k_1e^{k_1x} + k_1^2xe^{k_1x} + p(e^{k_1x} + k_1xe^{k_1x}) + qxe^{k_1x} =$$

$$= e^{k_1x}\left[x(k_1^2 + pk_1 + q) + 2\left(k_1 + \frac{p}{2}\right)\right].$$

По условию $k_1^2 + pk_1 + q = 0$. Кроме того, $k_1 = k_2 = -\frac{p}{2}$, поэтому $k_1 + \frac{p}{2} = 0$. Следовательно, функция $y = xe^{k_1x}$ является вторым частным решением уравнения (1). Решения y_1 и y_2 линейно-независимы, так как

$$\frac{y_1}{y_2} = \frac{1}{x} \neq \text{const},$$

и по теореме 15.6 общее решение уравнения (1) имеет вид

$$y = C_1e^{k_1x} + C_2xe^{k_1x}.$$

3) Пусть k_1 и k_2 — комплексно-сопряженные корни, т. е. имеют вид $k_1 = \alpha + i\beta$, $k_2 = \alpha - i\beta$ ($\beta \neq 0$). Тогда по теореме 15.8 функции

$$y_1 = e^{k_1x} = e^{(\alpha+i\beta)x} \text{ и } y_2 = e^{k_2x} = e^{(\alpha-i\beta)x}$$

будут частными решениями уравнения (1). Общее же решение уравнения (1) будет иметь вид

$$y = \tilde{C}_1e^{(\alpha+i\beta)x} + \tilde{C}_2e^{(\alpha-i\beta)x},$$

где $\tilde{C}_1$ и $\tilde{C}_2$ — произвольные постоянные. Отсюда

$$y = e^{\alpha x}(\tilde{C}_1e^{i\beta x} + \tilde{C}_2e^{-i\beta x}).$$

Применяя формулы Эйлера (см. п. 4, § 7, гл. XIV), последнее равенство можно записать в виде

$$y = e^{\alpha x}[\tilde{C}_1(\cos\beta x + i\sin\beta x) + \tilde{C}_2(\cos\beta x - i\sin\beta x)],$$

откуда

$$y=e^{\alpha x}\left[(\tilde{C}_1+\tilde{C}_2)\cos\beta x+i(\tilde{C}_1-\tilde{C}_2)\sin\beta x\right].$$

Обозначив $\tilde{C}_1+\tilde{C}_2=C_1$, $i(\tilde{C}_1-\tilde{C}_2)=C_2$, получим окончательный вид общего решения $y=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x)$ ■

Пример 1. Найти общее решение уравнения $y''+y'-2y=0$.

Решение. Характеристическое уравнение имеет вид $k^2+k-2=0$, его корни $k_1=1$, $k_2=-2$ — вещественные и различные. Соответствующие частные решения уравнения: $y_1=e^x$, $y_2=e^{-2x}$. Общее решение уравнения имеет вид $y=C_1e^x+C_2e^{-2x}$.

Пример 2. Найти общее решение уравнения $y''-2y'+y=0$.

Решение. Характеристическое уравнение имеет вид $k^2-2k+1=0$, его корни $k_1=k_2=1$ — вещественные и равные. Соответствующие частные решения уравнения: $y_1=e^x$, $y_2=xe^x$. Общее решение уравнения имеет вид $y=C_1e^x+C_2xe^x=e^x(C_1+C_2x)$.

Пример 3. Найти общее решение уравнения $y''-4y'+13y=0$.

Решение. Характеристическое уравнение имеет вид $k^2-4k+13=0$, его корни $k_1=2+i3$, $k_2=2-i3$ — комплексные. Соответствующие частные решения уравнения: $y_1=e^{2x}\cos 3x$, $y_2=e^{2x}\sin 3x$. Общее решение уравнения имеет вид $y=e^{2x}(C_1\cos 3x+C_2\sin 3x)$.

2. Линейные неоднородные дифференциальные уравнения второго порядка с постоянными коэффициентами. Рассмотрим линейное неоднородное уравнение второго порядка

$$y''+py'+qy=f(x), \tag{3}$$

где p и q — вещественные числа, $f(x)$ — непрерывная функция.

Мы уже знаем, что общее решение такого уравнения представляет собой сумму частного решения неоднородного уравнения и общего решения соответствующего однородного уравнения. Поскольку общее решение однородного уравнения мы находить умеем, то остается рассмотреть вопрос о нахождении частного решения. Для нахождения частного решения можно применить метод вариации произвольных постоянных. Однако если в правой части уравнения (3) стоит многочлен, или показательная функция, или тригонометрическая функция $\cos\beta x$ или $\sin\beta x$, или, наконец, линейная комбинация из названных функций, то частное решение может быть найдено методом неопределенных коэффициентов, не содержащим процесса интегрирования.

Рассмотрим различные виды правых частей уравнения (3).

1°. *Пусть правая часть имеет вид*

$$f(x)=P_n(x),$$

где $P_n(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ — *многочлен степени n.* Тогда частное решение $\tilde{y}$ можно искать в виде

$$\tilde{y}=Q_n(x)x^r,$$

где $Q_n(x)$ — многочлен той же степени, что и $P_n(x)$, а r — число корней характеристического уравнения, равных нулю.

Пример 4. Найти общее решение уравнения $y''-2y'+y=x+1$.

Решение. Общее решение соответствующего однородного уравнения имеет вид $Y=(C_1+C_2x)e^x$ (см. пример 2). Так как правая часть уравнения является многочленом первой степени и ни один из корней характеристического уравнения $k^2-2k+1=0$ не равен нулю ($k_2=k_1=1$), то частное решение ищем в виде

$$\tilde{y}=(Ax+B)x^0=Ax+B,$$

где A и B — неизвестные коэффициенты. Дифференцируя дважды $\tilde{y}=Ax+B$ и подставляя $\tilde{y}$, $\tilde{y}'$ и $\tilde{y}''$ в данное уравнение, находим

$$-2A+Ax+B=x+1.$$

Приравнивая коэффициенты при одинаковых степенях x в обеих частях равенства

$$A=1,\quad -2A+B=1,$$

находим $A=1$, $B=3$. Итак, частное решение данного уравнения имеет вид $\tilde{y}=x+3$, а его общее решение —

$$y=(C_1+C_2x)e^x+x+3.$$

2°. *Пусть правая часть имеет вид*

$$f(x)=e^{\alpha x}P_n(x),$$

где $P_n(x)$ — многочлен степени n. Тогда частное решение $\tilde{y}$ следует искать в виде

$$\tilde{y}=Q_n(x)x^r e^{\alpha x},$$

где $Q_n(x)$ — многочлен той же степени, что и $P_n(x)$, а r — число, показывающее, сколько раз α является корнем характеристического уравнения. При $\alpha=0$ $f(x)=P_n(x)$, т. е. имеет место случай 1°.

Пример 5. Найти общее решение уравнения $y''-4y'+3y=xe^x$.

Решение. Здесь характеристическое уравнение $k^2-4k+3=0$ имеет корни $k_1=1$, $k_2=3$. Значит, общее решение соответствующего однородного уравнения имеет вид $Y=C_1e^x+C_2e^{3x}$. В правой части данного уравнения стоит произведение многочлена первой степени на показательную функцию $e^{\alpha x}$ при $\alpha=1$. Так как среди корней характеристического уравнения имеется только один корень $k_1=\alpha=1$, то $r=1$. $P_n(x)$ — многочлен первой степени. Поэтому частное решение данного уравнения ищем в виде

$$\tilde{y}=(Ax+B)xe^x=(Ax^2+Bx)e^x.$$

Дифференцируя и подставляя в уравнение, получаем

$$-4Ax+2A-2B=x.$$

Приравнивая коэффициенты при одинаковых степенях x в обеих частях равенства

$$-4A=1,\ 2A-2B=0,$$

находим $A=-\frac{1}{4}$, $B=-\frac{1}{4}$. Подставляя найденные значения A и B в выражение для $\tilde{y}$, найдем частное решение данного уравнения $\tilde{y}=-\frac{1}{4}(x^2+x)e^x$, и общее решение запишется в виде

$$y=\tilde{y}+Y=C_1e^x+C_2e^{3x}-\frac{1}{4}(x^2+x)e^x.$$

3°. *Пусть правая часть имеет вид*

$$f(x)=a\cos\beta x+b\sin\beta x,$$

где a, b и β — известные числа. Тогда частное решение $\tilde{y}$ надо искать в виде

$$\tilde{y}=(A\cos\beta x+B\sin\beta x)x^r,$$

где A и B — неизвестные коэффициенты, а r — число, равное числу корней характеристического уравнения, совпадающих с $i\beta$.

Пример 6. Найти общее решение уравнения $y''+y=\sin x$.

Решение. Здесь характеристическое уравнение $k^2+1=0$ имеет корни $k_1=i$, $k_2=-i$. Поэтому общее решение соответствующего однородного уравнения будет $Y=C_1\cos x+C_2\sin x$. В правой части стоит тригонометрическая функция $\sin x$, т. е. $a=0$, $b=1$, $\beta=1$. Так как $i\beta=i$ является корнем характеристического уравнения, то $r=1$, и частное решение надо искать в виде

$$\tilde{y}=(A\cos x+B\sin x)x.$$

Дифференцируя и подставляя в уравнение, получим:

$$2(-A\sin x+B\cos x)=\sin x,$$

откуда $A=-\frac{1}{2}$, $B=0$. Таким образом, частное решение $\tilde{y}=-\frac{1}{2}x\cos x$, а общее решение уравнения

$$y=\tilde{y}+Y=C_1\cos x+C_2\sin x-\frac{1}{2}x\cos x.$$

Пример 7. Найти общее решение уравнения $y''+y=\sin 2x$.

Решение. Данное уравнение отличается от предыдущего только тем, что $\beta=2$. Так как $i\beta=i2$ не является корнем характеристического уравнения, то $r=0$, и частное решение надо искать в виде

$$\tilde{y}=A\cos 2x+B\sin 2x.$$

Дифференцируя и подставляя в уравнение, получим

$$-3A\cos 2x-3B\sin 2x=\sin 2x,$$

откуда $A=0$, $B=-\frac{1}{3}$, т. е. частное решение $\tilde{y}=-\frac{1}{3}\sin 2x$, а общее решение уравнения

$$y=\tilde{y}+Y=C_1\cos x+C_2\sin x-\frac{1}{3}\sin 2x.$$

4°. *Пусть правая часть имеет вид*

$$f(x)=e^{\alpha x}[P_n(x)\cos\beta x+P_m(x)\sin\beta x],$$

где $P_n(x)$ — многочлен степени n, а $P_m(x)$ — многочлен степени m. Тогда частное решение следует искать в виде

$$\tilde{y}=x^r e^{\alpha x}[Q_1(x)\cos\beta x+Q_2(x)\sin\beta x],$$

где $Q_1(x)$ и $Q_2(x)$ — многочлены степени s, $s=\max\{n, m\}$, а r — число, равное числу корней характеристического уравнения, совпадающих с $\alpha+i\beta$.

Пример 8. Найти общее решение уравнения $y''-y=3e^{2x}\cos x$.

Решение. Здесь характеристическое уравнение $k^2-1=0$ имеет корни $k_1=1$, $k_2=-1$. Общее решение однородного уравнения имеет вид $Y=C_1e^x+C_2e^{-x}$. В правой части уравнения стоит произведение многочлена нулевой степени, показательной и тригонометрической функции, так что $P_n(x)=3$, $P_m(x)=0$, $s=0$. Так как число $\alpha+i\beta=2+i1$ не является корнем характеристического уравнения, то $r=0$, и частное решение ищем в виде

$$\tilde{y}=e^{2x}(A\cos x+B\sin x).$$

Дифференцируя и подставляя в уравнение, получим

$$(2A+4B)\cos x+(2B-4A)\sin x=3\cos x.$$

Приравнивая коэффициенты при $\cos x$ и $\sin x$, находим

$$2A+4B=3,\quad -4A+2B=0,$$

отсюда $A=\frac{3}{10}$, $B=\frac{3}{5}$. Таким образом, частное решение $\tilde{y}=e^{2x}\left(\frac{3}{10}\cos x+\frac{3}{5}\sin x\right)$, а общее решение уравнения

$$y=\tilde{y}+Y=e^{2x}\left(\frac{3}{10}\cos x+\frac{3}{5}\sin x\right)+C_1e^x+C_2e^{-x}.$$

При решении линейных дифференциальных уравнений обычно возникают трудности при составлении частных решений. В связи с этим рассмотрим

Пример 9. По данным корням характеристического уравнения и данной правой части $f(x)$ написать вид частного решения $\tilde{y}$ линейного неоднородного уравнения:

а) $k_1=3+i2$, $k_2=3-i2$, $f(x)=8e^{3x}\sin 2x$;

б) $k_1=k_2=-3$, $f(x)=2xe^{-3x}\sin x$;

в) $k_1=1$, $k_2=-3$, $f(x)=e^x(1-x)\cos 3x$;

г) $k_1=1+i2$, $k_2=1-i2$, $f(x)=e^x(\cos 2x-3\sin 2x)$;

д) $k_1=2+i\frac{1}{2}$, $k_2=2-i\frac{1}{2}$, $f(x)=e^{2x}\left[(x^3-\right.$

$$\left.-1)\cos\frac{x}{2}+x\sin\frac{x}{2}\right].$$

Решение: а) Имеем: $\alpha=3$, $\beta=2$, $P_n(x)=0$, $P_m(x)=8$, $s=0$. Так как число $\alpha+i\beta=3+i2$ является корнем характеристического уравнения, то $r=1$. Поэтому $\tilde{y}=xe^{3x}(A\cos 2x+B\sin 2x)$.

б) Имеем: $\alpha=-3$, $\beta=1$, $P_n(x)=0$, $P_m(x)=2x$, $m=1$, $s=1$. Так как число $\alpha+i\beta=-3+i$ не совпадает с корнями характеристического уравнения, то $r=0$. Поэтому $\tilde{y}=e^{-3x}[(Ax+B)\cos x+(Cx+D)\sin x]$.

в) Имеем: $\alpha=1$, $\beta=3$, $P_n(x)=1-x$, $P_m(x)=0$, $n=1$, $s=1$. Так как число $\alpha+i\beta=1+i3$ не является корнем характеристического уравнения, то $r=0$. Поэтому $\tilde{y}=e^x[(Ax+B)\cos 3x+(Cx+D)\sin 3x]$.

г) Имеем: $\alpha=1$, $\beta=2$, $P_n(x)=1$, $P_m(x)=-3$, $s=0$. Так как число $\alpha+i\beta=1+i2$ является корнем характеристического уравнения, то $r=1$. Поэтому $\tilde{y}=xe^x(A\cos 2x+B\sin 2x)$.

д) Имеем: $\alpha=2, \beta=\frac{1}{2}$, $P_n(x)=(x^3-1)$, $P_m(x)=x$, $n=3$, $m=1$, $s=3$. Число $\alpha+i\beta=2+i\frac{1}{2}$ является корнем характеристического уравнения, значит, $r=1$. Следовательно,

$$\tilde{y}=xe^{2x}\left[(Ax^3+Bx^2+Cx+D)\cos\frac{x}{2}+\right.$$
$$\left.+(Ex^3+Fx^2+Gx+H)\sin\frac{x}{2}\right].$$

В заключение докажем теорему, которая часто применяется при решении линейных неоднородных уравнений, в правой части которых стоит сумма нескольких слагаемых.

Теорема 15.10. *Если $\tilde{y}_1$ — частное решение уравнения*

$$y''+py'+qy=f_1(x), \tag{4}$$

а $\tilde{y}_2$ — частное решение уравнения

$$y''+py'+qy=f_2(x), \tag{5}$$

то сумма $\tilde{y}_1+\tilde{y}_2$ является частным решением уравнения

$$y''+py'+qy=f_1(x)+f_2(x). \tag{6}$$

Доказательство. Пусть $\tilde{y}_1$ и $\tilde{y}_2$ — частные решения соответственно уравнений (4) и (5). Составим сумму $\tilde{y}_1+\tilde{y}_2$ и подставим ее вместе с производными в левую часть уравнения (6). Получим

$$(\tilde{y}_1+\tilde{y}_2)''+p(\tilde{y}_1+\tilde{y}_2)'+q(\tilde{y}_1+\tilde{y}_2)=$$
$$=(\tilde{y}_1''+p\tilde{y}_1'+q\tilde{y}_1)+(\tilde{y}_2''+p\tilde{y}_2'+q\tilde{y}_2)=f_1+f_2,$$

так как по условию выражение в первой скобке равно $f_1(x)$, а выражение во второй скобке равно $f_2(x)$. Следовательно, $\tilde{y}_1+\tilde{y}_2$ есть частное решение уравнения (6) ■

Пример 10. Найти общее решение уравнения $y''-2y'+y=\sin x+e^{-x}$.

Решение. Характеристическое уравнение $k^2-2k+1=0$ имеет корни $k_1=k_2=1$, поэтому общее решение соответствующего однородного уравнения будет

$$Y=C_1e^x+C_2xe^x=e^x(C_1+C_2x).$$

Так как правая часть уравнения состоит из суммы двух функций $\sin x$ и e^{-x}, то по теореме 15.10 частное решение данного уравнения будем искать в виде

$$\tilde{y}=\tilde{y}_1+\tilde{y}_2,$$

где $\tilde{y}_1$ — частное решение уравнения $y''-2y'+y=\sin x$, а $\tilde{y}_2$ — частное решение уравнения $y''-2y'+y=e^{-x}$.

Сначала найдем частное решение $\tilde{y}_1$. Так как число $i\beta=i$ не является корнем характеристического уравнения ($r=0$), то частное решение $\tilde{y}_1$ будем искать в виде $\tilde{y}_1 = A\sin x + B\cos x$. Подставляя $\tilde{y}_1$, $\tilde{y}'_1$ и $\tilde{y}''_1$ в уравнение $y''-2y'+y=\sin x$ и сравнивая коэффициенты при $\sin x$ и $\cos x$, получим $-2A=0$, $2B=1$, отсюда $A=0$, $B=\frac{1}{2}$ и, следовательно, $\tilde{y}_1 = \frac{1}{2}\cos x$.

Теперь найдем частное решение $\tilde{y}_2$. Будем его искать в виде $\tilde{y}_2 = Ae^{-x}$, так как число $\alpha=-1$ не является корнем характеристического уравнения. Подставляя $\tilde{y}_2$, $\tilde{y}'_2$ и $\tilde{y}''_2$ в уравнение $y''-2y'+y=e^{-x}$, получим $A=\frac{1}{4}$. Следовательно, $\tilde{y}_2 = \frac{1}{4}e^{-x}$.

Таким образом, частное решение данного уравнения будет

$$\tilde{y} = \tilde{y}_1 + \tilde{y}_2 = \frac{1}{2}\cos x + \frac{1}{4}e^{-x},$$

а общее решение этого уравнения запишется в виде

$$y = \tilde{y} + Y = \frac{1}{2}\cos x + \frac{1}{4}e^{-x} + C_1e^x + C_2xe^x.$$

§ 5. Применение линейных дифференциальных уравнений к изучению колебательных явлений

Линейные дифференциальные уравнения являются мощным аппаратом в решении задач о колебаниях, занимающих значительное место в современной технике и физике. Мы познакомимся с одной из них — с задачей о колебании груза, подвешенного на вертикальной пружине.

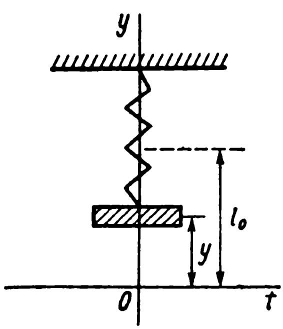


Рис. 106

Постановка задачи. Пусть груз массы m, подвешенный на пружине, движется по вертикальной прямой. Если пружину с грузом оттянуть или сжать, то груз начнет совершать колебания около положения равновесия. Установим закон движения груза, т. е. найдем формулу, выражающую отклонения груза от положения равновесия в любой момент времени t.

Совместим начало координат с положением равновесия груза, а ось Oy направим вертикально вверх. Обозначим через l_0 расстояние от конца нерастянутой пружины без груза до положения равновесия груза, через y_0 — начальное отклонение, а через y — отклонение груза от положения равновесия в момент времени t (рис. 106). Тогда на груз будет действовать сила, равная сумме трех сил:

1) сила тяжести груза mg, направленная в отрицательном направлении; 2) сила сопротивления среды, направленная в сторону, противоположную движению груза, и по величине, как показывает опыт, пропорциональная скорости движения груза: $\lambda \frac{dy}{dt}$, где λ — коэффициент пропорциональности; 3) упругая сила пружины, направленная в положительном направлении, величина которой, по закону Гука*, пропорциональна ее деформации, т. е. равна $c(l_0-y)$, где c — коэффициент пропорциональности, называемый коэффициентом жесткости пружины (масса пружины не учитывается).

Переходим к составлению дифференциального уравнения движения. По второму закону Ньютона получаем следующее уравнение движения груза:

$$m\frac{d^2y}{dt^2} = c(l_0 - y) - mg - \lambda \frac{dy}{dt}.$$

Так как в положении равновесия ($y=0$) вес груза mg уравновешивается упругой силой пружины, то $mg = cl_0$. Поэтому

$$m\frac{d^2y}{dt^2} = -cy - \lambda \frac{dy}{dt},$$

или

$$m\frac{d^2y}{dt^2} + \lambda \frac{dy}{dt} + cy = 0. \qquad (1)$$

Получили искомое дифференциальное уравнение (1), которое называется *уравнением свободных колебаний груза, подвешенного на пружине.*

Если на груз действует внешняя сила, направленная вертикально вдоль оси Oy, величина которой $F(t)$ зависит от времени t, например сила, заставляющая пе-

---

* Гук Роберт (1635—1703) — английский естествоиспытатель.

риодически колебаться всю систему, то уравнение (1) запишется в виде

$$m\frac{d^2y}{dt^2}+\lambda\frac{dy}{dt}+cy=F(t). \qquad (2)$$

Уравнение (2) называется *уравнением вынужденных колебаний груза, подвешенного на пружине.*

Поделив все члены уравнения (2) на m и обозначив $\frac{c}{m}=\omega^2$, $\frac{\lambda}{m}=2\mu$, $\frac{F(t)}{m}=f(t)$, получим окончательный вид уравнения вынужденных колебаний:

$$\frac{d^2y}{dt^2}+2\mu\frac{dy}{dt}+\omega^2y=f(t). \qquad (3)$$

Уравнение (3) представляет собой *линейное неоднородное дифференциальное уравнение второго порядка с постоянными коэффициентами.*

Перейдем теперь к исследованию колебаний с применением известных нам решений линейных уравнений с постоянными коэффициентами.

Свободные колебания. Пусть сначала отсутствуют внешняя сила $f(t)$ и сопротивление среды ($\mu=0$). Тогда уравнение (3) примет вид

$$\frac{d^2y}{dt^2}+\omega^2y=0.$$

Это линейное однородное уравнение. Характеристическое уравнение $k^2+\omega^2=0$ имеет корни $k_1=i\omega$, $k_2=-i\omega$, и общее решение уравнения запишется по формуле

$$y=C_1\cos\omega t+C_2\sin\omega t,$$

где C_1 и C_2 — произвольные постоянные. Для дальнейших рассуждений заменим произвольные постоянные C_1 и C_2 другими: $A>0$ и φ, полагая $C_1=A\sin\varphi$, $C_2=A\cos\varphi$ (отсюда $A=\sqrt{C_1^2+C_2^2}$ и $\operatorname{tg}\varphi=C_1/C_2$). Тогда

$$y=A\sin\varphi\cos\omega t+A\cos\varphi\sin\omega t,$$

и общее решение можно записать в виде

$$y=A\sin(\omega t+\varphi). \qquad (4)$$

Формула (4) устанавливает закон движения груза, подвешенного на пружине, т. е. выражает отклонение y груза от положения равновесия в любой момент времени t. По этой формуле груз совершает, как говорят, свободные гармонические колебания около положения равновесия. Величина A называется *амплитудой колебаний*, ω — *частотой колебаний* и φ — *начальной фазой*.

Для того чтобы выделить из общего решения частное, необходимо задать начальные условия движения. Пусть в начальный момент времени $t=0$ отклонение и скорость груза известны:

$$y|_{t=0}=y_0,\ y'|_{t=0}=y'_0. \tag{5}$$

Тогда, найдя из (4)

$$y'=A\omega\cos(\omega t+\varphi) \tag{6}$$

и подставляя начальные условия (5) в (4) и (6), получим

$$\begin{cases} y_0=A\sin\varphi, \\ y'_0=A\omega\cos\varphi. \end{cases} \tag{7}$$

Отсюда, выражая произвольные постоянные A и φ через ω, y_0 и y'_0 и подставляя их значения в (4), получим искомое частное решение, удовлетворяющее заданным начальным условиям (5).

Из формул (7), в частности, следует, что постоянные A и φ зависят от частоты колебаний ω и от начальных условий движения. Частота колебаний не зависит от начальных условий, а зависит от отношения коэффициента жесткости пружины к массе груза $\left(\omega=\sqrt{\frac{c}{m}}\right)$.

Пусть теперь отсутствует внешняя сила $f(t)$ и имеет место сопротивление среды ($\mu\neq0$), например сопротивление воздуха. В этом случае уравнение (3) примет вид

$$\frac{d^2y}{dt^2}+2\mu\frac{dy}{dt}+\omega^2y=0. \tag{8}$$

Характеристическое уравнение $k^2+2\mu k+\omega^2=0$ имеет корни $k_{1,2}=-\mu\pm\sqrt{\mu^2-\omega^2}$. Здесь возможны три случая:

1) $\mu>\omega$. Корни

$$k_1=-\mu+\sqrt{\mu^2-\omega^2} \text{ и } k_2=-\mu-\sqrt{\mu^2-\omega^2} -$$

вещественные, различные и отрицательные. Общее решение уравнения (8) запишется в виде

$$y=C_1e^{(-\mu+\sqrt{\mu^2-\omega^2})t}+C_2e^{(-\mu-\sqrt{\mu^2-\omega^2})t}.$$

Из полученной формулы следует, что груз колебаний не совершает, при неограниченном возрастании t отклонение груза y бесконечно долго приближается к положению равновесия ($y\to 0$ при $t\to+\infty$). В этом случае говорят, что груз совершает непериодическое затухающее движение.

2) $\mu=\omega$. Корни $k_1=k_2=-\mu$ — вещественные, равные и отрицательные. Общее решение уравнения (8) примет вид

$$y=C_1e^{-\mu t}+C_2te^{-\mu t}=e^{-\mu t}(C_1+C_2t).$$

В этом случае груз совершает движение, аналогичное предыдущему.

3) $\mu<\omega$. Корни

$$k_1=-\mu+i\sqrt{\omega^2-\mu^2} \text{ и } k_2=-\mu-i\sqrt{\omega^2-\mu^2}$$

— комплексные. Общее решение уравнения (8) будет иметь вид

$$y=e^{-\mu t}(C_1\cos\tilde{\omega}t+C_2\sin\tilde{\omega}t),$$

где $\tilde{\omega}=\sqrt{\omega^2-\mu^2}$. Совершив преобразование, аналогичное (4), получим решение уравнения (8) в виде

$$y=Ae^{-\mu t}\sin(\tilde{\omega}t+\varphi).$$

Здесь, в отличие от формулы (4), амплитуда $Ae^{-\mu t}$ зависит от времени t. Так как $\mu<0$, то она стремится к нулю при $t\to+\infty$. Поэтому в данном случае груз совершает так называемые свободные затухающие гармонические колебания около положения равновесия.

Вынужденные колебания. Резонанс. Рассмотрим теперь случай, когда на колебательную систему действует периодическая внешняя сила $f(t)=a\sin\omega_1 t$, предположив для простоты, что сопротивление среды отсутствует ($\mu=0$). В этом случае уравнение (8) примет вид

$$\frac{d^2y}{dt^2}+\omega^2 y=a\sin\omega_1 t. \qquad (9)$$

Это линейное неоднородное уравнение с постоянными коэффициентами. Известно, что общее решение данного уравнения представляет собой сумму общего решения Y соответствующего однородного уравнения, которое выше было найдено (см. формулу (4)), и частного решения $\tilde{y}$ неоднородного уравнения, которое надо найти.

Здесь нужно рассмотреть отдельно два случая:

1) $\omega\neq\omega_1$, т. е. частота внешней периодической силы отлична от частоты свободных колебаний груза. Так как число $i\omega_1$ не совпадает с корнем характеристического уравнения $k^2+\omega^2=0$, то частное решение можно найти в виде

$$\tilde{y}=A\cos\omega_1 t+B\sin\omega_1 t.$$

Дифференцируя $\tilde{y}$ дважды и подставляя $\tilde{y}$ и $\tilde{y}''$ в уравнение (9), найдем $B=\dfrac{a}{\omega^2-\omega_1^2}$, $A=0$. Таким образом,

$$\tilde{y}=\frac{a}{\omega^2-\omega_1^2}\sin\omega_1 t,$$

и общее решение уравнения (9) имеет вид

$$y=\tilde{y}+Y=\frac{a}{\omega^2-\omega_1^2}\sin\omega_1 t+A\sin(\omega t+\varphi). \qquad (10)$$

Формула (10) показывает, что частное решение $\tilde{y}$ определяет колебание системы, создаваемое внешней силой, общее решение Y — свободное колебание груза, а общее решение y — сложное колебательное движение, складывающееся из двух колебаний с разными частотами ω и ω_1.

В этом случае амплитуда постоянна, поэтому если ω и ω_1 близки по величине, то груз совершает при движении колебания около положения равновесия с большой амплитудой.

2) $\omega=\omega_1$, т. е. частота внешней периодической силы совпадает с частотой груза. Так как $i\omega_1$ является корнем характеристического уравнения $k^2+\omega^2=0$, то частное решение в этом случае следует искать в виде

$$\tilde{y}=t(A\cos\omega_1 t+B\sin\omega_1 t).$$

Дифференцируя $\tilde{y}$ дважды и подставляя $\tilde{y}$ и $\tilde{y}''$ в уравнение (9), найдем $A=-\frac{a}{2\omega}$, $B=0$. Таким образом,

$$\tilde{y}=-\frac{at}{2\omega}\cos\omega t,$$

и общее решение уравнения (9) запишется в виде

$$y=\tilde{y}+Y=A\sin(\omega t+\varphi)-\frac{at}{2\omega}\cos\omega t.$$

Полученная формула показывает, что в данном случае, как и в предыдущем, имеет место сложное колебательное движение, складывающееся из двух колебаний с одинаковыми частотами.

Присутствие множителя t во втором члене говорит о том, что амплитуда колебаний неограниченно возрастает при неограниченном возрастании времени t, т. е. груз будет совершать через некоторый промежуток времени колебания с очень большой амплитудой, даже если амплитуда внешней силы мала. Это явление в технике и физике называется *резонансом*. Иногда оно приводит к разрушению колеблющихся систем.

На примере груза, подвешенного на пружине, мы рассмотрели случай так называемых механических колебаний упругих систем (колебание на рессорах вагонов, автомобилей и т. п.). Аналогичное исследование проводится и при изучении электрических колебаний, звуковых колебаний и многих других. И главное место в этих исследованиях принадлежит дифференциальным уравнениям.

В заключение отметим, что изложенная выше теория линейных дифференциальных уравнений второго порядка переносится и на линейные дифференциальные уравнения любого порядка.

ОГЛАВЛЕНИЕ

Учебник состоит из трех частей:

часть I — **«Анализ функций одной переменной»**,
часть II — **«Анализ функций нескольких переменных»**,
часть III — **«Ряды. Дифференциальные уравнения»**.

Материал учебника изложен в доступной форме и сопровождается большим количеством примеров и задач. Для студентов нематематических факультетов, а также для аспирантов и преподавателей вузов

ОНИКС

ISBN 978-5-488-02067-2

9 785488 020672

В.С. Шипачев

КУРС ВЫСШЕЙ МАТЕМАТИКИ