# ВЫСШАЯ МАТЕМАТИКА

А. А. ГУСАК

1 том

ББК 22.1я73
Г 96
УДК 51(075.8)

Рецензенты:

*кафедра общей математики ЛГУ им. А. А. Жданова,*
*Л. Д. Кудрявцев,*
*доктор физико-математических наук*

**Гусак А. А.**

Г 96 Высшая математика. Т. 1: [Учеб. пособие для естеств. спец. университетов.— 2-е изд., перераб. и доп.— Минск: Изд-во БГУ, 1983.—462 с., черт.

В пер.: 90 к.

Второе издание пособия для студентов естественных специальностей вузов переработано и дополнено в соответствии с новой программой по высшей математике (1975 г.), в частности, предусмотрено изучение линейной алгебры и элементов теории групп.

Г $\frac{1702010000-074}{М\,317-83}$ 27—83 ББК 22.1я73

# ПРЕДИСЛОВИЕ

Второе издание этой книги является переработанным и дополненным. Оно подготовлено в соответствии с новыми программами по высшей математике для студентов химических, биологических, геологических и географических специальностей университетов.

Первый том включает программный материал по разделу «Аналитическая геометрия и линейная алгебра» и некоторой части второго раздела «Математический анализ».

В первом томе наиболее существенной переработке подверглись начальные главы обоих разделов. Переработана первая глава: объединены два параграфа, посвященные выводу формул для расстояния между двумя точками на плоскости и в пространстве (при этом существенно сокращен первый из них), написан новый параграф «Центр тяжести системы масс», добавлен параграф «Параметрические уравнения линии» (перенесен из гл. 3), исключен параграф «Площадь треугольника». Глава «Прямая линия» и глава «Кривые» объединены в одну главу «Алгебраические линии первого и второго порядка», причем первая из указанных глав значительно сокращена. Глава «Элементы линейной алгебры» разбита на две: «Матрицы и определители», «Системы линейных уравнений». В первую из этих глав включены новые параграфы: «Обратная матрица», «Ранг матрицы», «Комплексные числа. Комплексные матрицы», вторая из них содержит новый параграф «Исследование систем линейных уравнений».

В одну главу «Функции и пределы» объединены две главы: «Функция», «Пределы и бесконечно малые», причем исключены параграфы, относящиеся к свойствам пределов последовательностей. Главы «Производная», «Дифференциал функции», «Теоремы о конечных приращениях» во втором издании представлены одной главой «Производные и дифференциалы».

Указанная переработка обусловлена содержанием школьного курса математики и новыми программами по высшей математике для студентов естественных специальностей вузов.

Во втором издании первого тома имеются новые главы: «Линейные пространства. Линейные преобразования», «Квадратичные формы», «Группы».

Считаю своим приятным долгом выразить искреннюю признательность и сердечную благодарность докторам физико-математических наук, профессорам Л. Д. Кудрявцеву, Н. М. Матвееву, М. М. Смирнову, а также всем сотрудникам кафедры общей математики Ленинградского государственного университета за критические замечания и полезные советы.

*Автор*

## ВВЕДЕНИЕ

Математика — одна из самых древних наук. Первые математические сведения и понятия восходят к доисторическим временам. Они возникли в процессе практической деятельности людей. Из самой природы человек заимствовал геометрические формы, в процессе решения жизненных задач зарождались арифметика и геометрия.

Основные особенности математики — абстрактность, логическая строгость, исключительная широта ее приложений. Разумеется, абстракция свойственна не только математике, но и другим наукам, всему человеческому мышлению. Но если в других науках для доказательства своих утверждений исследователи постоянно обращаются к опыту, то в математике справедливость рассматриваемого факта доказывается не проверкой его на примерах, не проведением экспериментов, а чисто логическим путем; теоремы доказываются только путем рассуждений и выкладок. Следовательно, не только понятия, но и методы математики являются отвлеченными.

Чрезвычайная широта приложений математики — одна из самых характерных ее особенностей. В повседневной жизни и практике постоянно пользуются наиболее распространенными понятиями арифметики и геометрии. Без применения математических методов была бы невозможна современная техника, так как без надлежащих расчетов не обходится ни одно техническое усовершенствование, тем более создание новых областей техники. Точные науки (астрономия, механика, физика, химия) развивают свои теории, используя математический аппарат, их прогресс был бы немыслим без математики. С другой стороны, потребности этих наук всегда оказывали и оказывают решающее влияние на развитие математики, способствуют разработке новых математических методов и теорий.

Благодаря созданию быстродействующих электронных вычислительных машин математические методы начали не только еще шире использоваться в точных науках, но и применяться в тех областях, где математика раньше применялась мало, либо ее применение даже не представлялось возможным (биология, медицина, экономика, лингвистика, социология и т. п.).

Абстрактный характер математики порождает (и порождал еще в древности) идеалистические представления о ее независимости от материальной действительности. Видя крайнюю абстрактность математических понятий и убедительность ее выводов, идеалисты считают, что математика не зависит от опыта, происходит из чистого мышления, является продуктом собственного творчества ученых. Возражая против такой точки зрения, Ф. Энгельс писал: «Но совершенно неверно, будто в чистой математике разум имеет дело только с продуктами своего собственного творчества и воображения. Понятия числа и фигуры взяты не откуда-нибудь, а только из действительного мира. <...> Чистая математика имеет своим объектом пространственные формы и количественные отношения действительного мира, стало быть — весьма реальный материал. Тот факт, что этот материал принимает чрезвычайно абстрактную форму, может лишь слабо затушевать его происхождение из внешнего мира» [1].

Основы теоретической арифметики и элементарная геометрия сложились еще в Древней Греции. «Начала» Евклида, написанные в III в. до н. э., служат основой для школьных учебников геометрии во всех странах в течение более двух тысяч лет.

К концу XVI в. математика складывалась из арифметики, алгебры, геометрии и тригонометрии. Это была математика постоянных величин.

В XVII столетии в связи с запросами практики математические исследования необычайно расширяются и возникает несколько новых направлений: аналитическая геометрия, анализ бесконечно малых, теория вероятностей и др. Создание аналитической геометрии и анализа явилось подлинной революцией в математике. В центре исследований оказались новые объекты и методы. Через анализ в математику проникает идея движения, измене-

---

[1] *Маркс К., Энгельс Ф.* Соч.— М., 1961, т. 20, с. 37.

ния и, стало быть, диалектика. Математика перешла к изучению переменных величин и функций, как аналогов механического движения и любого изменения вообще.

Одновременно с дифференциальным и интегральным исчислением возникли другие разделы анализа: теория рядов, теория дифференциальных уравнений, приложение анализа к геометрии, выделившееся позже в особую область геометрии, называемую дифференциальной геометрией.

Крупнейшие достижения анализа всегда были связаны с решением задач точного естествознания. Начиная с Ньютона, величайшие ученые Д. Бернулли (1700—1782), Л. Эйлер (1707—1783), Ж. Лагранж (1736—1813), П. Лаплас (1749—1827), Ж. Фурье (1768—1830), А. Пуанкаре (1854—1912), М. В. Остроградский (1801—1861), А. М. Ляпунов (1857—1918), как и многие другие, в своих математических трудах исходили из насущных задач механики и физики. В связи с запросами механики Эйлер и Лагранж создают вариационное исчисление. Пуанкаре и Ляпунов — качественную теорию дифференциальных уравнений. Эта теория решает задачи об устойчивости движения, работы механизмов, электроколебательных систем и т. п.

В XIX в. анализ обогатился новой важной ветвью — теорией функций комплексной переменной, оформленной в трудах французского математика О. Коши (1789—1857). Зачатки этой теории имелись уже в сочинениях Эйлера и некоторых других математиков. Эта теория нашла существенные приложения к решению важных задач самой математики, физики и техники. Так, основная теорема Н. Е. Жуковского (1847—1921) о подъемной силе крыла самолета доказана средствами этой теории. В. И. Ленин назвал Н. Е. Жуковского «отцом русской авиации».

В 1826 г. гениальный русский математик Н. И. Лобачевский (1792—1856) открыл неевклидову геометрию; с этого времени началось принципиально новое развитие геометрии.

В связи с уточнением основ анализа возникла новая область математики — теория множеств, созданная немецким математиком Г. Кантором (1845—1918). Общие идеи теории множеств проникли во все области математики. Эти идеи явились основой новой главы анализа — теории функций действительной переменной, развитие которой многим обязано Н. Н. Лузину (1883—1950) и его школе, а также французским математикам Э. Борелю (1871—1956) и А. Лебегу (1875—1941).

В анализе возникли и другие новые теории, например теория приближения функций, основоположником которой является великий русский математик П. Л. Чебышев (1821—1894). Возникновение этой теории было обусловлено необходимостью решения важных практических задач, связанных с вопросами конструирования паровых машин. П. Л. Чебышев подчеркивал роль практики в развитии теории, их полезные взаимосвязи.

До Великой Октябрьской социалистической революции математика в России концентрировалась лишь в трех-четырех научных центрах. Славу русской математике создавали Петербургская математическая школа, у колыбели которой стоял Эйлер и которая переживала расцвет в период, когда творили П. Л. Чебышев и его ученики, и Казанская школа, основанная Н. И. Лобачевским.

В короткое время после революции были созданы новые научные школы, новые направления, обеспечившие успехи советской математики и получившие мировое признание. Ряд новых математических школ был создан в союзных республиках, в городах, бывших до революции окраинами царской России.

Советская математика занимает передовое место в мировой математической науке. Наши ученые имеют блестящие достижения во всех основных областях современной математики, а во многих из них результаты советских математиков играют определяющую роль.

Коммунистическая партия и Советское правительство уделяют большое внимание развитию науки, в частности математики, и внедрению ее достижений в производство. Ученые призваны обеспечить «дальнейшую разработку проблем теоретической и прикладной математики и кибернетики для более широкого применения в народном хозяйстве математических методов и электронно-вычислительной техники, автоматизации процессов производства и совершенствования управления» [1)].

В «Основных направлениях экономического и социального развития СССР на 1981—1985 годы и на период до 1990 года» указано на необходимость развития математической теории, повышения эффективности ее использования в прикладных целях.

1) Материалы XXIV съезда КПСС.— М., 1971, с. 244.

РАЗДЕЛ I

# АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ И ЛИНЕЙНАЯ АЛГЕБРА

**Линейная алгебра исследует системы линейных уравнений, т. е. уравнений, содержащих неизвестные только в первой степени. Аналитическая геометрия — раздел математики, в котором изучаются геометрические объекты с помощью алгебраических методов. Основным методом аналитической геометрии является метод координат.**

## Глава 1

## МЕТОД КООРДИНАТ

*Координатами* точки называются числа, заданием которых определяется положение точки на прямой, плоскости или в пространстве. Метод координат позволяет геометрические задачи сводить к алгебраическим. Заслуга введения метода координат принадлежит Декарту[1] и Ферма[2].

### § 1.1. Координаты на прямой

Рассмотрим некоторую прямую. Фиксируем одно из двух определяемых ею направлений и назовем его положительным, другое — отрицательным. Прямая, на которой указано положительное направление, называется

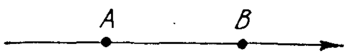



Рис. 1.1

осью (на чертеже положительное направление обычно обозначается стрелкой, рис. 1.1). Выберем масштаб, т. е.

1) *Рене Декарт* (*René Descartes*, 1596—1650) — французский философ, математик.

2) *Пьер Ферма* (*Pierre Fermat*, 1601—1665) — французский математик, по профессии юрист.

единицу длины для измерения длин любых отрезков этой оси.

Отрезок, ограниченный точками $A$ и $B$, называется *направленным отрезком,* или *вектором,* если указано, какая из данных точек является началом, а какая — концом. Направлением отрезка считается направление от начала к концу. Направленный отрезок с началом в точке $A$ и концом в точке $B$ обозначается символом $\overline{AB}$. Две различные точки $A$ и $B$ определяют два направленных отрезка $\overline{AB}$ и $\overline{BA}$. Если точки $A$ и $B$ совпадают, отрезок $AA$ называется *нулевым;* нулевой отрезок не имеет определенного направления.

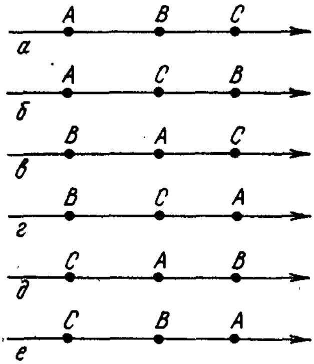


Рис. 1.2

*Величиной направленного отрезка* $\overline{AB}$ некоторой оси называется его длина, взятая со знаком плюс, когда направление этого отрезка совпадает с положительным направлением оси, и со знаком минус, когда оно совпадает с отрицательным направлением оси. Величину направленного отрезка $\overline{AB}$ будем обозначать через $AB$; очевидно,

$$AB = -BA. \qquad (1.1)$$

При любом расположении трех точек $A$, $B$, $C$ на оси (рис. 1.2) величины направленных отрезков $\overline{AB}$, $\overline{BC}$, $\overline{AC}$ связаны соотношением

$$AB + BC = AC, \qquad (1.2)$$

которое называется *основным тождеством.*

Действительно, пусть точки $A$, $B$, $C$ расположены на оси в порядке, указанном на рис. 1.2, *а*. По свойству длин отрезков $|AB| + |BC| = |AC|$. Но в данном случае $|AB| = AB$, $|BC| = BC$, $|AC| = AC$, поэтому $AB + BC = AC$.

В случае, изображенном на рис. 1.2, *б*, $|AC| + |CB| = |AB|$, $AC + CB = AB$. Принимая во внимание формулу

(1.1), из последнего равенства получаем $AC=AB-CB=$ $=AB+BC$, $AB+BC=AC$, т. е. равенство (1.2) выполняется. Все остальные случаи (рис. 1.2, *б—е*) рассматриваются аналогично. Отметим, что равенство (1.2) выполняется и тогда, когда один из отрезков является нулевым или все три отрезка — нулевые.

Фиксируем на прямой $x$ две различные точки $O$ и $E$; точку $O$ назовем ***началом координат***, точку $E$ — ***единичной точкой***. В качестве масштаба примем отрезок $OE$, с помощью этого отрезка будем измерять длины всех других отрезков оси. Выберем на прямой положительное направление, совпадающее с направлением отрезка $OE$ (рис. 1.3).

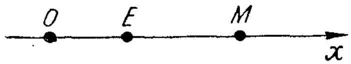



Рис. 1.3

Прямая $Ox$, на которой фиксировано положительное направление, начало отсчета и выбран масштаб для измерения длин, называется *координатной осью*.

*Координатой* произвольной точки $M$ координатной оси называется величина направленного отрезка $\overline{OM}$ этой оси. Если обозначить координату точки $M$ через $x$, то по определению

$$x=OM, \tag{1.3}$$

где $OM$ — величина направленного отрезка $\overline{OM}$ оси $Ox$. Запись $M(x)$ означает, что $M$ имеет координату $x$.

Таким образом, если дана точка координатной оси, можно указать число, являющееся ее координатой на этой оси; если дано число, на координатной оси можно построить единственную точку, для которой это число является координатой указанной точки. Следовательно, между множеством точек координатной оси и множеством действительных чисел установлено взаимно-однозначное соответствие. Это дает возможность вместо точки рассматривать соответствующее число — ее координату.

Величину направленного отрезка и его длину можно выразить через координаты начала и конца этого отрезка.

**Теорема 1.1.** Если $M_1(x_1)$, $M_2(x_2)$ — две заданные точки оси, то величина направленного отрезка $\overline{M_1M_2}$ и его длина выражаются соответственно формулами

$$M_1M_2 = x_2 - x_1; \tag{1.4}$$

$$|M_1M_2| = |x_2 - x_1|. \tag{1.5}$$

Доказательство. Основное тождество (1.2) для трех точек $O$, $M_1$, $M_2$ координатной оси принимает вид $OM_1 + M_1M_2 = OM_2$, откуда $M_1M_2 = OM_2 - OM_1$. В соответствии с определением (1.3) $x_2 = OM_2$, $x_1 = OM_1$, поэтому $M_1M_2 = x_2 - x_1$; формула (1.4) доказана. Так как длина направленного отрезка равна модулю его величины, то $|M_1M_2| = |x_2 - x_1|$, что и требовалось доказать.

Сущность формулы (1.4) можно выразить следующими словами: чтобы найти величину направленного отрезка, необходимо из координаты его конца вычесть координату начала.

Замечание. Формулу (1.5) можно записать в виде

$$\rho(M_1, M_2) = |x_2 - x_1|, \tag{1.5'}$$

где $\rho(M_1, M_2)$ — расстояние между точками $M_1$ и $M_2$.

Поскольку $|x_2 - x_1| = |x_1 - x_2|$, то смысл формул (1.5) и (1.5′) выражается так: чтобы вычислить расстояние между двумя точками оси, нужно из координаты одной точки вычесть координату другой и взять модуль полученной разности.

Пример. Даны две точки $M_1$ (1), $M_2$ (−5). Найти величину направленного отрезка $\overline{M_1M_2}$ и расстояние между точками.

В данном случае $x_1 = 1$, $x_2 = -5$; по формулам (1.4) и (1.5) находим $M_1M_2 = -5 - 1 = -6$, $\rho(M_1, M_2) = |-5-1| = |-6| = 6$.

## § 1.2. Декартовы прямоугольные координаты на плоскости

Будем говорить, что на плоскости введена система координат, если указан способ, позволяющий устанавливать положение точки заданием чисел. Существуют различные способы определения положения точки на плоскости, в зависимости от которых рассматривают разные системы координат: декартовы прямоугольные, полярные и др.

Декартова прямоугольная система координат на плоскости вводится следующим образом. Выберем масштаб для измерения длин любых отрезков на плоскости. В данной плоскости проведем две взаимно перпендикулярные координатные оси: горизонтальную $Ox$ (ось абсцисс) и вертикальную $Oy$ (ось ординат). Точку $O$ пересечения координатных осей назовем *началом координат* (рис. 1.4).

Отметим, что если $E_1$ и $E_2$ — единичные точки координатных осей, то

$$|OE_1| = |OE_2| = 1. \tag{1.6}$$

Равенство (1.6) означает, что для измерения длин координатных осей выбрана одна и та же единица масштаба (совпадающая с единицей масштаба для измерения длин любых отрезков на плоскости).

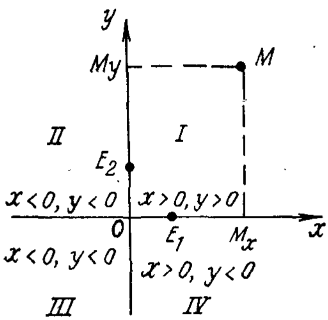



Рис. 1.4

Пусть дана произвольная точка $M$ рассматриваемой плоскости. Проведем через точку $M$ прямую, параллельную координатной оси $Oy$, точку пересечения этой прямой с осью $Ox$ обозначим через $M_x$, проведем через $M$ также прямую, параллельную оси $Ox$, до пересечения с осью $Oy$ в точке $M_y$.

*Декартовыми прямоугольными координатами* точки $M$ на плоскости называются числа, определяемые формулами

$$x = OM_x, \quad y = OM_y, \tag{1.7}$$

где $OM_x$, $OM_y$ — величины направленных отрезков $\overline{OM_x}$, $\overline{OM_y}$ соответствующих координатных осей.

Первая координата называется *абсциссой*, вторая — *ординатой* данной точки $M$. Запись $M(x, y)$ означает, что точка $M$ имеет координаты $x$, $y$; если дано несколько точек, то их координаты записываются с соответствующими индексами, например, $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ и т. д. Отметим, что для точек, лежащих на оси $Oy$, $x=0$; для точек, лежащих на оси $Ox$, $y=0$; начало координат — точка $O$ имеет координаты $x=0$, $y=0$. Координатные оси разбивают на четыре части множество тех точек плоскости, которые не лежат на осях. Каждая из этих частей называется *координатной четвертью*, или *квадрантом*. На рис. 1.4 указаны знаки координат точек в каждой четверти.

Каждой точке $M$ плоскости поставлена в соответствие упорядоченная пара чисел $(x, y)$, определяемых формулами (1.7). С другой стороны, если дана упорядоченная пара чисел $(x, y)$, то можно построить единственную точку, для которой $x$ и $y$ будут декартовыми прямоугольными координатами.

## § 1.3. Деление отрезка в данном отношении

Рассмотрим две различные точки $M_1$ и $M_2$ некоторой плоскости (точка $M_1$ считается первой, точка $M_2$ — второй). Проведем через данные точки прямую, зафиксируем на ней положительное направление (т. е. сделаем ее осью) и выберем масштаб. Пусть $M$ — любая, принадлежащая этой прямой точка, отличная от $M_2$.

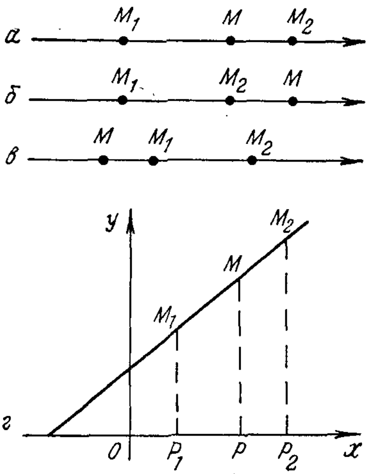


Рис. 1.5

Отношением, в котором точка $M$ делит направленный отрезок $\overline{M_1M_2}$, ограниченный точками $M_1$ и $M_2$ (рис. 1.5), называется число

$$\lambda = \frac{M_1M}{MM_2}, \qquad (1.8)$$

где $M_1M$, $MM_2$ — величины направленных отрезков $\overline{M_1M}$, $\overline{MM_2}$ указанной оси. Отношение (1.8) не изменится, если по-другому фиксировать положительное направление оси или выбрать другой масштаб, так как в обоих случаях числитель и знаменатель умножатся на одно и то же число (в первом случае — на $(-1)$).

Если точка $M$ принадлежит отрезку $[M_1M_2]$ (рис. 1.5, *а*), то $\lambda > 0$ (так как $M_1M$ и $MM_2$ одного знака); в этом случае говорят, что $M$ делит отрезок $\overline{M_1M_2}$ внутренним образом. Если точка $M$ лежит вне отрезка $[M_1M_2]$ (рис. 1.5, *б*, *в*), то $\lambda < 0$ (числа $M_1M$ и $MM_2$ имеют противоположные знаки); в этом случае говорят, что точка $M$ делит отрезок $\overline{M_1M_2}$ внешним образом. Отметим, что

$$\lambda \neq -1. \tag{1.9}$$

Действительно, предположив противное, т. е. $\lambda = -1$, получим $\frac{M_1M}{MM_2} = -1$, $M_1M = -MM_2$, $M_1M + MM_2 = 0$, $M_1M_2 = 0$, что возможно лишь в случае, когда $M_1$ совпадает с $M_2$, а это противоречит условию (точки $M_1$ и $M_2$ различны).

Если точка $M$ совпадает с $M_1$, то $\lambda = 0$; если $M$ неограниченно приближается к точке $M_2$, то $|\lambda|$ неограниченно возрастает.

Задача о делении отрезка в данном отношении формулируется следующим образом. Дано отношение $\lambda$, в котором точка $M$ (неизвестная) делит отрезок $\overline{M_1M_2}$, ограниченный данными точками $M_1$ и $M_2$. Требуется найти точку $M$ (т. е. определить ее координаты). Решение поставленной задачи дает следующая теорема.

**Теорема 1.2.** Если точка $M(x, y)$ делит отрезок $\overline{M_1M_2}$, ограниченный точками $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, в отношении $\lambda$, ее координаты определяются формулами

$$x = \frac{x_1 + \lambda x_2}{1+\lambda}, \quad y = \frac{y_1 + \lambda y_2}{1+\lambda}. \tag{1.10}$$

В частности, координаты середины отрезка $\overline{M_1M_2}$ вычисляются по формулам

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}. \tag{1.11}$$

Доказательство. Точки $M_1$, $M$, $M_2$ спроектируем на ось $Ox$, обозначим их проекции соответственно через $P_1$, $P$, $P_2$ (рис. 1.5, *г*). Используя условие, теорему элементарной геометрии о пропорциональности отрезков прямых, заключенных между параллельными прямыми, и формулу (1.4), получаем

$$\frac{M_1M}{MM_2} = \lambda, \quad \frac{P_1P}{PP_2} = \frac{M_1M}{MM_2}, \quad \frac{P_1P}{PP_2} = \lambda, \quad P_1P = x - x_1,$$

$$PP_2 = x_2 - x, \quad \frac{x - x_1}{x_2 - x} = \lambda, \quad x - x_1 = \lambda(x_2 - x),$$

$$x(1+\lambda) = x_1 + \lambda x_2. \tag{1.12}$$

Принимая во внимание условие (1.9), из равенства (1.12) находим первую из формул (1.10). Вторая из этих формул устанавливается аналогично (точки $M_1$, $M$ и $M_2$ необходимо предварительно спроектировать на ось $Oy$).

Если точка $M$ находится в середине отрезка $\overline{M_1M_2}$, то $M_1M = MM_2$ и $\lambda = 1$. Полагая в формулах (1.10) $\lambda = 1$, получаем формулы (1.11).

Пример. Даны две точки $M_1(-1, -2)$, $M_2(3, 4)$. На прямой $(M_1M_2)$ найти точку $M$, которая в три раза ближе к $M_1$, чем к $M_2$, и находится вне отрезка $\overline{M_1M_2}$. Найти середину этого отрезка.

Искомая точка $M$ делит отрезок $\overline{M_1M_2}$ в отношении $\lambda = -\frac{1}{3}$. По формулам (1.10), считая в них $x_1 = -1$, $y_1 = -2$, $x_2 = 3$, $y_2 = 4$, находим

$$x = \frac{-1 + \left(-\frac{1}{3}\right)\cdot 3}{1 + \left(-\frac{1}{3}\right)} = -3, \quad y = \frac{-2 + \left(-\frac{1}{3}\right)\cdot 4}{1 + \left(-\frac{1}{3}\right)} = -5,$$

$$M(-3, -5).$$

С помощью формул (1.11) находим точку $N(1, 1)$ — середину отрезка $\overline{M_1M_2}$.

## § 1.4. Центр тяжести системы масс

Дана система масс $m_1, m_2, \ldots, m_n$, помещенных соответственно в точках $M_1(x_1, y_1), M_2(x_2, y_2), \ldots, M_n(x_n, y_n)$ некоторой плоскости. Требуется найти формулы, выражающие координаты центра тяжести этой системы масс.

Рассмотрим сначала случай $n=2$: массы $m_1$, $m_2$ помещены в точках $M_1(x_1, y_1)$, $M_2(x_2, y_2)$. Согласно известному принципу механики центр тяжести $M(x, y)$ системы данных масс делит отрезок $\overline{M_1M_2}$ на части, обратно пропорциональные массам $m_1$, $m_2$, т. е. в отношении $\lambda = m_2 : m_1$. По формулам (1.10) получаем координаты центра тяжести $M(x, y)$ этой системы:

$$x = \frac{x_1 + \frac{m_2}{m_1}x_2}{1 + \frac{m_2}{m_1}}, \quad y = \frac{y_1 + \frac{m_2}{m_1}y_2}{1 + \frac{m_2}{m_1}};$$

$$x = \frac{x_1m_1 + x_2m_2}{m_1 + m_2}, \quad y = \frac{y_1m_1 + y_2m_2}{m_1 + m_2}. \tag{1.13}$$

Аналогичные формулы находим и для координат центра тяжести $M(x, y)$ системы масс $m_1$, $m_2$, $m_3$, помещенных соответственно в точках $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$. Положение точки $M(x, y)$ не изменится, если массы

$m_1$ и $m_2$ сосредоточить в центре тяжести $M'(x', y')$ этой системы. Точку $M(x, y)$ будем рассматривать как центр тяжести системы двух масс: массы $m_3$, находящейся в точке $M_3$, и массы $m_1+m_2$, сосредоточенной в точке $M'(x', y')$ — центре системы масс $m_1$, $m_2$, где $x'$ и $y'$ определяются соответственно правыми частями формул (1.13). Находим центр тяжести $M(x, y)$ как точку, делящую отрезок $\overline{M'M_3}$ в отношении $\lambda = m_3 : (m_1+m_2)$. Применив формулы (1.10) в этом случае, найдем

$$x = \frac{x' + \frac{m_3}{m_1+m_2}x_3}{1+\frac{m_3}{m_1+m_2}} = \frac{\frac{x_1m_1+x_2m_2}{m_1+m_2}+\frac{m_3x_3}{m_1+m_2}}{\frac{m_1+m_2+m_3}{m_1+m_2}},$$

$$y = \frac{\frac{y_1m_1+y_2m_2}{m_1+m_2}+\frac{m_3y_3}{m_1+m_2}}{\frac{m_1+m_2+m_3}{m_1+m_2}},$$

$$x = \frac{x_1m_1+x_2m_2+x_3m_3}{m_1+m_2+m_3}, \quad y = \frac{y_1m_1+y_2m_2+y_3m_3}{m_1+m_2+m_3}. \qquad (1.14)$$

Методом математической индукции можно доказать, что искомые формулы имеют вид

$$x = \frac{x_1m_1+x_2m_2+\ldots+x_nm_n}{m_1+m_2+\ldots+m_n}, \quad y = \frac{y_1m_1+y_2m_2+\ldots+y_nm_n}{m_1+m_2+\ldots+m_n}, \qquad (1.15)$$

или

$$x = \frac{\sum_{k=1}^{n} x_km_k}{\sum_{k=1}^{n} m_k}, \quad y = \frac{\sum_{k=1}^{n} y_km_k}{\sum_{k=1}^{n} m_k}, \qquad (1.16)$$

где $\Sigma$ — сумма однотипных слагаемых.

## § 1.5. Полярные координаты

Пусть дана некоторая плоскость. Зафиксируем на этой плоскости точку, обозначим ее буквой $O$ и назовем *полюсом*. Луч $[OP)$, исходящий из полюса, назовем *полярной осью*. Выберем масштаб для измерения длин отрезков и условимся, какие повороты вокруг точки $O$ будем считать положительными; обычно считают положительными те повороты, которые совершаются против часовой стрелки.

Рассмотрим произвольную точку $M$ заданной плоскости, обозначим через $\rho$ ее расстояние до полюса и назовем *полярным радиусом* (рис. 1.6); угол, на который нужно повернуть полярную ось $[OP)$, чтобы она совпала с $[OM)$, обозначим через $\varphi$ и назовем *полярным углом*.

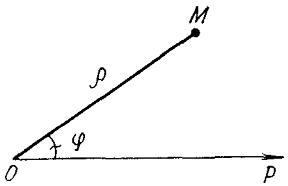


Рис. 1.6

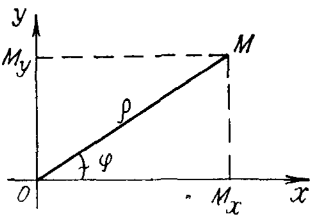


Рис. 1.7

Полярными координатами точки $M$ называются ее полярный радиус $\rho$ и полярный угол $\varphi$. Каждой точке плоскости соответствует вполне определенное значение $\rho \geqslant 0$. Значение $\varphi$ для точек, отличных от полюса, определено с точностью до слагаемого $2k\pi$, где $k$ — любое целое число. Для полюса $\rho=0$, а значение $\varphi$ не определено. Чтобы каждая точка плоскости получила вполне определенные значения полярных координат, достаточно считать, что $0 \leqslant \varphi < 2\pi$, а в полюсе $\varphi=0$. Указанные значения $\varphi$ назовем *главными*.

Наряду с введенной полярной системой координат рассмотрим прямоугольную декартову систему такую, что полюс совпадает с началом, а полярная ось — с положительной полуосью $Ox$ (рис. 1.7). Если $M$ — произвольная точка плоскости, $(x, y)$ — ее декартовы, а $(\rho, \varphi)$ — полярные координаты, то, очевидно,

$$x=\rho \cos \varphi, \; y=\rho \sin \varphi. \tag{1.17}$$

Формулы (1.17) выражают прямоугольные декартовы координаты точки плоскости через ее полярные координаты.

Полярные координаты точки выражаются через ее декартовы координаты формулами

$$\rho=\sqrt{x^2+y^2},\ \cos\varphi=\frac{x}{\sqrt{x^2+y^2}},\ \sin\varphi=\frac{y}{\sqrt{x^2+y^2}},\quad (1.18)$$

которые следуют из формул (1.17). Чтобы получить первую из формул (1.18), достаточно возвести в квадрат обе формулы (1.17) и почленно сложить (первая из формул (1.18) имеет простой геометрический смысл, она выражает теорему Пифагора для треугольника $OMM_x$).

## § 1.6. Прямоугольные декартовы координаты в пространстве

Прямоугольная декартова система координат в пространстве определяется заданием масштаба (отрезка для измерения длин) и трех пересекающихся в одной точке взаимно перпендикулярных осей, занумерованных в определенном порядке.

Точка пересечения осей называется *началом координат*, а сами оси — *координатными осями*, причем первая из них — *осью абсцисс*, вторая — *осью ординат*, третья — *осью аппликат*. Обозначим начало координат буквой $O$; координатные оси — соответственно через $Ox$, $Oy$, $Oz$ (рис. 1.8).

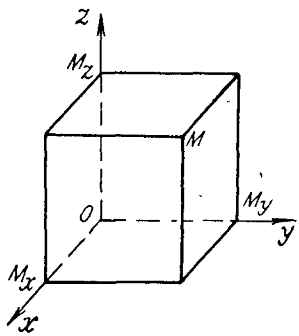



Рис. 1.8

Пусть $M$ — произвольная точка пространства. Проведем через нее три плоскости, перпендикулярные координатным осям, и точки пересечения с осями обозначим соответственно через $M_x$, $M_y$, $M_z$. Прямоугольными декартовыми координатами точки $M$ называются числа, определяемые формулами

$$x=OM_x,\ y=OM_y,\ z=OM_z,\quad (1.19)$$

где $OM_x$, $OM_y$, $OM_z$ — величины направленных отрезков $\overline{OM_x}$, $\overline{OM_y}$, $\overline{OM_z}$ соответствующих координатных осей. Число $x$ называется первой координатой, или абсциссой, число $y$ — второй координатой, или ординатой, число $z$ — третьей координатой, или аппликатой, точки $M$. Запись $M(x, y, z)$ означает, что точка $M$ имеет координаты $x$, $y$, $z$.

Каждая пара координатных осей определяет координатную плоскость. Координатные плоскости обозначают соответственно $Oxy$, $Oxz$, $Oyz$. Легко видеть, что прямоугольные декартовы координаты точки $M$ представляют собой расстояния этой точки до соответствующих координатных плоскостей, взятые с надлежащими знаками (абсцисса $x$ — расстояние до координатной плоскости $Oyz$, взятое со знаком плюс, если $M$ — впереди плоскости $Oyz$, со знаком минус, если $M$ — сзади плоскости $Oyz$, когда положительное направление оси $Ox$ и взаимное расположение осей выбраны так, как показано на рис. 1.8).

На основании сказанного заключаем, что для точек, лежащих в плоскости $Oyz$,

$$x=0; \tag{1.20}$$

для точек, лежащих в плоскости $Oxz$,

$$y=0; \tag{1.21}$$

для точек, лежащих в плоскости $Oxy$,

$$z=0. \tag{1.22}$$

Точки оси $Ox$ характеризуются двумя равенствами

$$y=0, \ z=0; \tag{1.23}$$

точки оси $Oy$ — равенствами

$$x=0, \ z=0; \tag{1.24}$$

точки оси $Oz$ — равенствами

$$x=0, \ y=0. \tag{1.25}$$

Начало координат имеет все нулевые координаты: $x=y==z=0$.

Координатные плоскости $Oxy$, $Oxz$, $Oyz$ делят все точки пространства, не принадлежащие этим плоскостям, на восемь частей, называемых *октантами*. Исходя из октанта I, в котором все координаты положительны, перенумеруем октанты (I, II, III, IV) верхнего полупространства ($z>0$) против часовой стрелки (для наблюдателя со стороны положительной оси $Oz$). В нижнем полупространстве ($z<0$) произведем соответствующую нумерацию октантов (V, VI, VII, VIII) так, чтобы V находился под I, VI — под II, VII — под III, VIII — под IV. Знаки координат точек в различных октантах приведены в табл. 1.

Таблица 1

| Координаты | Октанты | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI | VII | VIII |
| $x$ | + | − | − | + | + | − | − | + |
| $y$ | + | + | − | − | + | + | − | − |
| $z$ | + | + | + | + | − | − | − | − |

Очевидно, знаки координат однозначно определяют октант пространства.

Итак, если задана система прямоугольных декартовых координат в пространстве, то каждой точке пространства можно поставить в соответствие упорядоченную тройку чисел $(x, y, z)$ — ее координаты. Обратно, если задана упорядоченная тройка вещественных чисел $(x, y, z)$, можно указать точку, для которой $x$ будет абсциссой, $y$ — ординатой, $z$ — аппликатой. Чтобы построить эту точку, необходимо на оси $Ox$ отложить направленный отрезок $\overline{OM_x}$, величина которого равна $x$, на оси $Oy$ — отрезок $\overline{OM_y}$, величина которого равна $y$, на оси $Oz$ — отрезок $\overline{OM_z}$, величина которого равна $z$. Через точки $M_x$, $M_y$, $M_z$ провести плоскости, перпендикулярные соответственно осям $Ox$, $Oy$, $Oz$; точка пересечения указанных плоскостей и будет искомой. Следовательно, между точками пространства и упорядоченными тройками чисел установлено взаимно-однозначное соответствие.

## § 1.7. Расстояние между двумя точками

Пусть относительно прямоугольной декартовой системы координат в пространстве заданы две точки, т. е. указаны их координаты. Требуется вывести формулу, по которой можно вычислить расстояние между данными точками. Решение этой задачи дает следующая теорема.

**Теорема 1.3.** Если $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ — две любые точки пространства, то расстояние между ними определяется формулой

$$\rho(M_1, M_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \quad (1.26)$$

Доказательство. Через точки $M_1$ и $M_2$ проведем плоскости, параллельные координатным плоскостям, и рассмотрим получившийся прямоугольный параллелепипед (рис. 1.9). Точки $M_1$ и $M_2$ являются противоположными вершинами этого параллелепипеда. Концы трех взаимно перпендикулярных ребер, исходящих из верши-

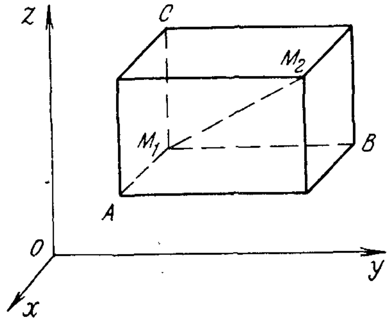



Рис. 1.9

ны $M_1$, обозначим через $A$, $B$, $C$, причем $A(x_2, y_1, z_1)$, $B(x_1, y_2, z_1)$, $C(x_1, y_1, z_2)$. На основании известной теоремы стереометрии получаем $|M_1M_2|^2=|M_1A|^2+|M_1B|^2+|M_1C|^2$, где $|M_1M_2|$, $|M_1A|$, $|M_1B|$, $|M_1C|$ — длины диагонали и боковых ребер.

Обозначив $\rho(M_1, M_2)=|M_1M_2|$, найдем $\rho(M_1, M_2)=\sqrt{|M_1A|^2+|M_1B|^2+|M_1C|^2}$. Подставляя в это равенство выражения $|M_1A|=|x_2-x_1|$, $|M_1B|=|y_2-y_1|$, $|M_1C|=|z_2-z_1|$, получаем формулу (1.26). В частном случае, если точка $M_1$ лежит в начале координат, т. е. если $x_1=y_1=z_1=0$, то

$$\rho(O, M_2)=\sqrt{x_2^2+y_2^2+z_2^2}. \tag{1.27}$$

Пример 1. Вычислить расстояние между точками $M_1(2, 1, -2)$ и $M_2(4, -3, -6)$, а также расстояние от точки $M_1$ до начала координат.

По формулам (1.26) и (1.27) соответственно получаем

$$\rho(M_1, M_2)=\sqrt{(4-2)^2+(-3-1)^2+(-6+2)^2}=6,$$

$$\rho(O, M_1)=\sqrt{2^2+1^2+(-2)^2}=3.$$

Формулы (1.26) и (1.27) упрощаются, когда точки $M_1$ и $M_2$ лежат в плоскости, параллельной одной из координатных плоскостей или в самой этой плоскости. Пусть точки $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ лежат в плоскости $Oxy$, тогда

$$\rho(M_1, M_2) = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}; \tag{1.28}$$

$$\rho(O, M_2) = \sqrt{x_2^2+y_2^2}. \tag{1.29}$$

Эти формулы можно получить и непосредственно, рассмотрев прямоугольные треугольники $M_1NM_2$ (рис. 1.10) и $OM_2P_2$ (рис. 1.11).

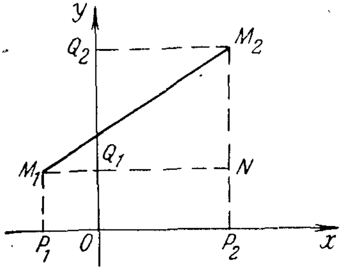



Рис. 1.10

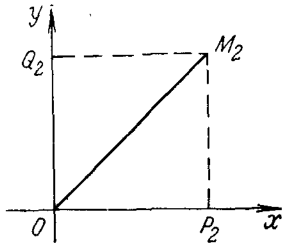



Рис. 1.11

Пример 2. Вычислить расстояние между точками $M_1(5, -2)$, $M_2(8, -6)$ и расстояние от точки $M_2$ до начала координат.

Так как $x_1=5$, $y_1=-2$, $x_2=8$, $y_2=-6$, то по формулам (1.28) и (1.29) соответственно находим

$$\rho(M_1, M_2) = \sqrt{(8-5)^2+[-6-(-2)]^2}=5, \quad \rho(O, M_2) = \sqrt{8^2+(-6)^2}=10.$$

## § 1.8. Цилиндрические и сферические координаты

В некоторой плоскости П фиксируем точку $O$ (рис. 1.12) и исходящий из нее луч $[OP)$. Через точку $O$ проведем прямую, перпендикулярную плоскости П, и укажем на ней положительное направление; полученную ось обозначим $Oz$. Выберем масштаб для измерения длин. Пусть $M$ — произвольная точка пространства, $N$ — ее проекция на плоскость П, $M_z$ — проекция на ось $Oz$. Обозначим через $\rho$ и $\varphi$ полярные координаты точки $N$ в плоскости П относительно полюса $O$ и полярной оси $OP$. *Цилиндрическими* координатами точки $M$ называются числа $\rho$, $\varphi$, $z$, где $\rho$, $\varphi$ — полярные координаты точки $N$ ($\rho \geqslant 0$, $0 \leqslant \varphi < 2\pi$), а $z = OM_z$ — величина направленного отрезка $\overline{OM_z}$ оси $Oz$. Запись $M(\rho, \varphi, z)$ означает, что точка $M$ имеет цилиндрические координаты $\rho$, $\varphi$, $z$. Наименование «цилиндрические координаты» объясняется тем, что координатная поверхность $\rho = \text{const}$ (т. е. множество точек, имеющих одну и ту же первую координату $\rho$) является цилиндром (на рис. 1.12 он изображен пунктиром).

Если выбрать систему прямоугольных декартовых координат так, как показано на рис 1.12, то декартовы координаты $x$, $y$, $z$ точки $M$ будут связаны с ее цилиндрическими координатами $\rho$, $\varphi$, $z$ формулами

$$x=\rho\cos\varphi,\quad y=\rho\sin\varphi,\quad z=z. \tag{1.30}$$

Сферические координаты вводятся следующим образом. Выберем масштаб для измерения длин отрезков, за-

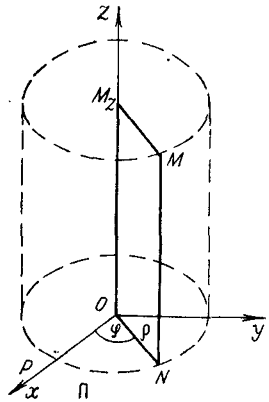



Рис. 1.12

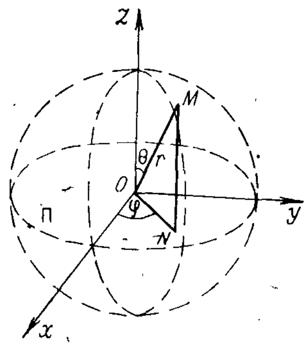



Рис. 1.13

фиксируем плоскость П с точкой $O$ и полуосью $Ox$, ось $Oz$, перпендикулярную плоскости П (рис. 1.13). Пусть $M$ — произвольная точка пространства (отличная от $O$), $N$ — проекция ее на плоскость П, $r$ — расстояние точки $M$ до начала координат, $\Theta$ — угол, образуемый отрезком $\overline{OM}$ с осью $Oz$, $\varphi$ — угол, на который нужно повернуть ось $Ox$ против часовой стрелки (если смотреть со стороны положительного направления оси $Oz$), чтобы она совпала с лучом $ON$; $\Theta$ называется *широтой*, $\varphi$ — *долготой*.

*Сферическими координатами* точки $M$ называются три числа $r$, $\Theta$, $\varphi$, определенные выше. Если точка $M$ имеет сферические координаты $r$, $\Theta$, $\varphi$, то пишут $M(r, \Theta, \varphi)$.

Наименование «сферические координаты» связано с тем, что координатная поверхность $r=\text{const}$ (т. е. множество точек, имеющих одну и ту же координату $r$) является сферой (на рис. 1.13 одна из таких сфер изображена пунктиром; фиксировав другое значение $r$, получим другую сферу).

Для того чтобы соответствие между точками пространства и тройками сферических координат $(r, \Theta, \varphi)$ было взаимно-однозначным, обычно считают, что $r$, $\Theta$, $\varphi$ изменяются в следующих границах: $0 \leqslant r < +\infty$, $0 \leqslant \Theta \leqslant \pi$, $0 \leqslant \varphi < 2\pi$. Если выбрать оси прямоугольной декартовой системы координат так, как указано на рис. 1.13, то декартовы координаты $x$, $y$, $z$ точки $M$ связаны с ее сферическими координатами $r$, $\Theta$, $\varphi$ формулами

$$x = r \sin \Theta \cos \varphi, \quad y = r \sin \Theta \sin \varphi, \quad z = r \cos \Theta. \qquad (1.31)$$

## § 1.9. Геометрическое значение уравнений между координатами на плоскости

В § 1.2, 1.5, 1.6 было установлено соответствие между точками плоскости (пространства) и упорядоченными парами (тройками) действительных чисел, что позволило вместо точек рассматривать их координаты. Решение простейших геометрических задач, рассмотренных в § 1.3, 1.4, сводилось к нахождению некоторых чисел по определенным формулам, т. е. осуществлялось алгебраическими методами.

Выясним, какой геометрический образ соответствует уравнению между координатами на плоскости. Уравнение с двумя переменными $x$ и $y$ [1] в общем виде записывается так:

$$F(x, y) = 0, \qquad (1.32)$$

где $F(x, y)$ — некоторая зависимость между $x$ и $y$, например, $F(x, y) = x - y$, $F(x, y) = x^2 + y^2 - 1$, $F(x, y) = y - \sin x$ и т. п.

Рассмотрим множество [2] всех точек плоскости, координаты которых удовлетворяют уравнению (1.32). Будем говорить, что уравнение (1.32) определяет указанное множество (фигуру). Например, уравнение $x - y = 0$ определяет прямую линию — биссектрису первого и третьего координатных углов (декартовы прямоугольные координаты любой точки этой биссектрисы удовлетворяют данному уравнению; это следует из определения декарто-

[1] Напомним, что указанным уравнением называется соотношение вида (1.32), которому удовлетворяют не все пары действительных чисел $x$ и $y$, в отличие от тождества, справедливого для всех пар значений $x$ и $y$; примеры уравнений: $x + y - 3 = 0$, $x^2 + y^2 - 9 = 0$; примеры тождеств: $(x + y) - x - y = 0$, $(x + y)^2 - x^2 - 2xy - y^2 = 0$.

[2] Множество точек плоскости называют *фигурой*.

вых прямоугольных координат и свойства биссектрисы — множества точек, равноудаленных от сторон угла). Уравнение $x^2+y^2-1=0$ определяет окружность радиуса $R=1$ с центром в начале координат, уравнение $y-\sin x=0$ — линию, называемую синусоидой, уравнение $\frac{|x|}{x}+\frac{|y|}{y}=$ $=0$ — множество точек, лежащих внутри второго и четвертого координатных углов.

Множество, определяемое некоторым уравнением, может состоять из одной точки (например, уравнению $(x-2)^2+(y+3)^2=0$ удовлетворяют координаты только одной точки $M(2, -3)$). Такое множество может оказаться пустым, т. е. не содержащим ни одной точки (например, уравнению $x^2+y^2+4=0$ не удовлетворяют координаты ни одной точки плоскости).

Аналитическая геометрия на плоскости изучает геометрические объекты, соответствующие алгебраическим уравнениям первой и второй степени относительно декартовых координат. В дальнейшем будет показано, что уравнение первой степени определяет прямую линию, а уравнение второй степени — одну из линий: окружность, эллипс, гиперболу, параболу, пару прямых, точку или пустое множество.

Будем исходить из наглядного представления о линии (кривой), известного из курса математики средней школы. Введем понятие уравнения линии.

**Уравнение линии на плоскости.** Уравнением линии на плоскости (относительно выбранной системы координат) называется такое уравнение с двумя переменными[1)]

$$F(x, y)=0, \tag{1.33}$$

которому удовлетворяют координаты $x$, $y$[2)] любой точки данной линии и не удовлетворяют координаты ни одной точки, не лежащей на этой линии.

Из определения уравнения линии следует решение простой задачи: выяснить, лежит ли данная точка на заданной линии. Если при подстановке координат точки в уравнение линии получается числовое равенство (тождество), то точка лежит на данной линии; если тождество не получается, то точка линии не принадлежит.

---

[1)] В частных случаях уравнение может не содержать одной из переменных.

[2)] Если речь идет о полярных координатах $\rho$, $\varphi$, то уравнение линии записывают в виде $F(\rho, \varphi)=0$.

Пример. Составить уравнение окружности радиуса $R$ с центром в точке $C(a, b)$.

Пусть $M$ — произвольная точка данной окружности (рис. 1.14), $x$, $y$ — ее декартовы прямоугольные координаты (их называют текущими координатами, так как они меняются при переходе от одной точки окружности к другой). Поскольку окружность — множество точек плоскости, равноудаленных от данной точки (центра), и расстояние между любой точкой окружности и ее центром равно $R$, то

$$\rho(C, M)=R. \qquad (1.34)$$

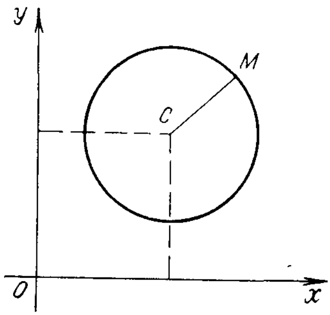



Рис. 1.14

На основании формулы (1.28)

$$\rho(C, M)=\sqrt{(x-a)^2+(y-b)^2}. \qquad (1.35)$$

Из двух последних равенств получаем

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

или

$$(x-a)^2+(y-b)^2=R^2. \qquad (1.36)$$

Итак, координаты любой точки данной окружности удовлетворяют уравнению (1.36). Если точка $N(x, y)$ не лежит на рассматриваемой окружности, то

$$\rho(C, N)<R \text{ или } \rho(C, N)>R,$$

поэтому ее координаты не будут удовлетворять уравнению (1.36). Следовательно, в соответствии с определением уравнение (1.36) и является уравнением данной окружности. Уравнение

$$x^2+y^2=R^2, \qquad (1.37)$$

полученное из уравнения (1.36) при $a=b=0$, называется *каноническим уравнением окружности.*

Замечание. Неравенству

$$(x-a)^2+(y-b)^2<R^2 \qquad (1.38)$$

удовлетворяют координаты всех точек, лежащих внутри окружности, определяемой уравнением (1.36); неравенству

$$(x-a)^2+(y-b)^2>R^2 \qquad (1.39)$$

удовлетворяют координаты всех точек, лежащих вне этой окружности.

Данные утверждения следуют из простых геометрических соображений, неравенств (1.38), (1.39) и формулы (1.28).

**Пересечение линий.** Рассмотрим две линии, определяемые уравнениями

$$F(x, y)=0; \qquad (1.40)$$

$$\Phi(x, y) = 0. \tag{1.41}$$

Требуется найти точку пересечения этих линий.

Точка пересечения данных линий принадлежит первой и второй линии, поэтому ее координаты удовлетворяют как первому, так и второму уравнению, т. е. удовлетворяют системе уравнений

$$\left.\begin{aligned} F(x, y) &= 0, \\ \Phi(x, y) &= 0. \end{aligned}\right\} \tag{1.42}$$

Обратно, если координаты $x_0$, $y_0$ некоторой точки удовлетворяют системе (1.42), то точка $M_0(x_0, y_0)$ лежит на линии (1.40) и на линии (1.41), т. е. является точкой их пересечения.

Следовательно, чтобы найти точки пересечения линий, необходимо решить систему их уравнений. Число решений равно числу точек пересечения. Если система (1.42) действительных решений не имеет, линии (1.40) и (1.41) не пересекаются.

Пример. Найти точки пересечения линий $x^2+y^2=1$, $x^2+(y-1)^2=2$. Решим систему уравнений

$$\left.\begin{aligned} x^2+y^2-1 &= 0, \\ x^2+(y-1)^2-2 &= 0. \end{aligned}\right\}$$

Вычитая почленно второе уравнение из первого, получаем $2y=0$, $y=0$. При $y=0$ из первого уравнения следует, что $x_1=-1$, $x_2=1$. Найдены две точки пересечения (рис. 1.15): $M_1(-1, 0)$, $M_2(1, 0)$.

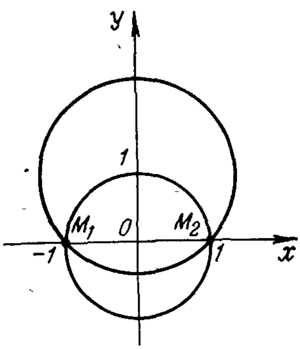


Рис. 1.15

Две основные задачи аналитической геометрии на плоскости состоят в следующем.

1. На плоскости задана линия как множество всех точек, удовлетворяющих определенному условию. Требуется составить уравнение линии относительно какой-либо системы координат.

2. Дано уравнение с двумя переменными. Требуется выяснить вид и расположение фигуры, определяемой этим уравнением относительно некоторой системы координат.

Следовательно, по геометрическим условиям, задающим линию, находят ее уравнение и затем по алгебраическим свойствам уравнения исследуют геометрические свойства этой линии. Таким путем геометрические задачи можно сводить к алгебраическим.

## § 1.10. Параметрические уравнения линии

Пусть относительно системы декартовых прямоугольных координат на плоскости задана некоторая линия. Координаты $(x, y)$ ее произвольной точки часто можно выразить через некоторую новую переменную (параметр) $t$:

$$x=\varphi_1(t),\ y=\varphi_2(t). \tag{1.43}$$

Уравнения (1.43) называются *параметрическими уравнениями* линии, если при изменении $t$ в конечном или бесконечном промежутке формулы (1.43) дают координаты любой точки данной линии и не дают координаты ни одной точки, не лежащей на этой линии.

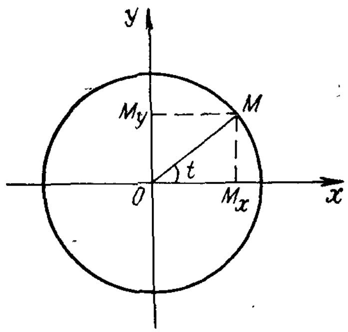


Рис. 1.16

Если дана окружность радиуса $R$ с центром в начале координат (рис. 1.16), то, взяв в качестве параметра $t$ угол, образуемый отрезком $OM$ с осью $Ox$, получим ее параметрические уравнения

$$x=R\cos t,\ y=R\sin t\ \ (0\leqslant t<2\pi). \tag{1.44}$$

Исключив из этих уравнений $t$ (для чего можно возвести в квадрат оба равенства и почленно сложить), получим уравнение (1.37).

Рассмотрим линию, являющуюся траекторией фиксированной точки окружности радиуса $R$, катящейся по прямой. Линия эта называется *циклоидой*. Указанную прямую примем за ось $Ox$ декартовой прямоугольной системы координат (рис. 1.17). Предположим, что фиксированная точка при начальном положении окружности находилась в начале координат, а после того, как окружность повернулась на угол $t$, заняла положение $M$.

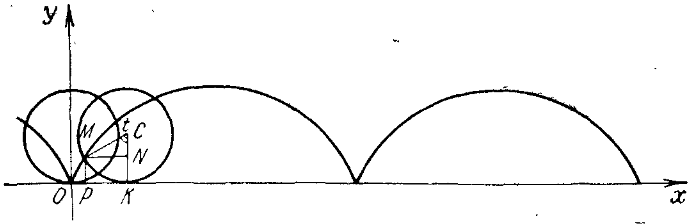



Рис. 1.17

Поскольку $x=OP=OK-PK$, $y=MP=CK-CN$ и $OK=$ $=MK=Rt$, $PK=MN=R\sin t$, $CK=R$, $CN=R\cos t$, то $x=$ $=Rt-R\sin t$, $y=R-R\cos t$, или

$$x=R(t-\sin t),\ y=R(1-\cos t). \qquad (1.45)$$

Уравнения (1.45) называются *параметрическими уравнениями циклоиды.*

## Глава 2

## АЛГЕБРАИЧЕСКИЕ ЛИНИИ ПЕРВОГО И ВТОРОГО ПОРЯДКА

*Алгебраической линией (кривой)* $n$-го порядка называется линия, определяемая алгебраическим уравнением $n$-й степени относительно декартовых координат. В этой главе изучаются линии первого порядка, т. е. линии, определяемые уравнением первой степени $Ax+$ $+By+C=0$, где $A^2+B^2\neq 0$, и линии второго порядка, т. е. линии, определяемые уравнением второй степени

$$Ax^2+Bxy+Cy^2+Dx+Ey+F=0 \quad (A^2+B^2+C^2\neq 0).$$

Будет показано, что линии первого порядка являются прямыми, а к важнейшим линиям второго порядка относятся окружность, эллипс, гипербола, парабола.

### § 2.1. Прямая на плоскости

Прямую линию на плоскости относительно системы декартовых прямоугольных координат можно задать различными способами. Прямая однозначно определяется углом, образуемым ею с осью $Ox$, и величиной направленного отрезка, отсекаемого на оси $Oy$, координатами двух точек и т. п. В зависимости от способа задания прямой рассматривают различные виды ее уравнения.

**Различные виды уравнения прямой на плоскости.** Рассмотрим прямую, параллельную оси $Oy$ прямоугольной декартовой системы координат (рис. 2.1). Обозначим буквой $A$ точку пересечения этой прямой с осью $Ox$, $(a, 0)$ — ее координаты, где $a = OA$. Уравнение

$$x = a \qquad (2.1)$$

является уравнением данной прямой. Действительно, уравнению (2.1) удовлетворяет координата $x$ любой точки $M(x, y)$ этой прямой и не удовлетворяет координата $x$ ни одной точки, не лежащей на прямой. Если $a = 0$, то прямая совпадает с осью $Oy$, которая, следовательно, имеет уравнение

$$x = 0. \qquad (2.2)$$

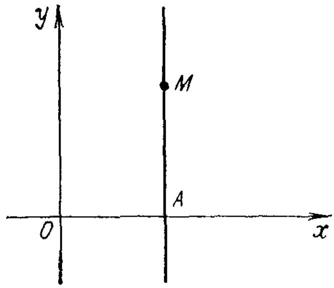


Рис. 2.1

Из курса математики средней школы известно уравнение прямой, пересекающей ось $Oy$:

$$y = kx + b, \qquad (2.3)$$

в котором $k$ — угловой коэффициент, определяемый формулой

$$k = \operatorname{tg} \alpha, \qquad (2.4)$$

где $\alpha$ — угол между прямой и осью $Ox$; $b = OB$ — величина направленного отрезка $\overline{OB}$, отсекаемого прямой на оси $Oy$ (рис. 2.2). Уравнение (2.3) называется *уравнением прямой с угловым коэффициентом.* Если прямая параллельна оси $Ox$, т. е. $\alpha = 0$, $k = 0$, то уравнение (2.3) принимает вид

$$y = b. \qquad (2.5)$$

Отметим, что ось $Ox$ определяется уравнением

$$y = 0. \qquad (2.6)$$

Выразим угловой коэффициент прямой (2.3) через координаты ее двух различных точек $M_1(x_1, y_1)$, $M_2(x_2, y_2)$. Так как эти точки лежат на прямой (2.3), то их координаты удовлетворяют данному уравнению, т. е. $y_1=kx_1+b$, $y_2=kx_2+b$. Вычитая первое равенство из второго, получаем $y_2-y_1=k(x_2-x_1)$, откуда

$$k=\frac{y_2-y_1}{x_2-x_1}, \tag{2.7}$$

поскольку $x_2-x_1\neq 0$ (прямая пересекает ось $Oy$, поэтому $x_1\neq x_2$). Равенство (2.7) имеет простой геометрический смысл: оно определяет тангенс угла $M_1$ в прямоугольном треугольнике $M_1M_2N$ (рис. 2.2).

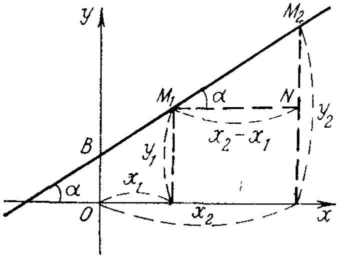



Рис. 2.2

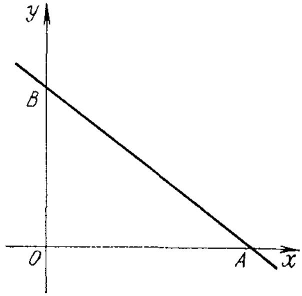



Рис. 2.3

Пусть заданы угловой коэффициент $k$ прямой и ее точка $M_1(x_1, y_1)$. Составим уравнение этой прямой. Зафиксируем произвольную точку $M(x, y)$ данной прямой и найдем выражение для ее углового коэффициента по формуле (2.7), положив в ней $y_2=y$, $x_2=x$: $k=\dfrac{y-y_1}{x-x_1}$ откуда

$$y-y_1=k(x-x_1). \tag{2.8}$$

Уравнение (2.8) называется *уравнением прямой, проходящей через данную точку в данном направлении.*

*Пучком прямых* на плоскости называется множество всех прямых этой плоскости, проходящих через данную точку (центр пучка). Будем считать, что в уравнении вида (2.8) $k$ может принимать любые действительные значения. Зафиксировав значение $k=k_1$, получим уравнение некоторой прямой; взяв другое значение $k=k_2$, получим уравнение другой прямой и т. д. Уравнение (2.8) в этом случае будет определять множество всех прямых, проходящих через точку $M_1(x_1, y_1)$, за исключением одной прямой, параллельной оси $Oy$, уравнение которой имеет вид

$x=x_1$ (в соответствии с (2.1)). Следовательно, пучок прямых с центром в точке $M_1(x_1, y_1)$ определяется уравнениями

$$y-y_1=k(x-x_1),\quad x=x_1, \tag{2.9}$$

где $k$ может принимать любые действительные значения.

Составим уравнение прямой, проходящей через две данные различные точки $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, где $x_1\neq x_2$, $y_1\neq y_2$. Поскольку эта прямая проходит через точку $M_1(x_1, y_1)$, уравнение (2.8) с учетом формулы (2.7) запишется так:

$$y-y_1=\frac{y_2-y_1}{x_2-x_1}(x-x_1), \text{ или } \frac{y-y_1}{y_2-y_1}=\frac{x-x_1}{x_2-x_1}. \tag{2.10}$$

Уравнение (2.10) называется *уравнением прямой, проходящей через две данные точки.*

Обозначив равные отношения буквой $t$, получим

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

$$x=x_1+(x_2-x_1)t,\quad y=y_1+(y_2-y_1)t. \tag{2.11}$$

Отметим, что при $t=0$ из уравнений (2.11) получаем координаты точки $M_1(x_1, y_1)$, при $t=1$ — координаты точки $M_2(x_2, y_2)$, при $0<t<1$ — координаты любой внутренней точки отрезка $[M_1M_2]$; когда $t$ меняется в бесконечном промежутке $]-\infty, +\infty[$, точка $M(x, y)$ описывает рассматриваемую прямую. Уравнения (2.11) называются *параметрическими уравнениями прямой.*

Пусть прямая $(AB)$ (рис. 2.3) отсекает на координатных осях отрезки, величины которых соответственно равны $a$ и $b$, т. е. $OA=a$, $OB=b$, $A(a, 0)$, $B(0, b)$. Применяя уравнение (2.10) для этого случая, т. е., полагая $x_1=a$, $y_1=0$, $x_2=0$, $y_2=b$, получаем уравнение в отрезках на осях координат:

$$\frac{y-0}{b-0}=\frac{x-a}{-a},\quad \frac{x}{a}+\frac{y}{b}=1. \tag{2.12}$$

Рассмотрим общее уравнение первой степени относительно прямоугольных декартовых координат $x$, $y$

$$Ax+By+C=0, \tag{2.13}$$

где $A$ и $B$ одновременно не равны нулю, т. е.

$$A^2+B^2\neq 0. \tag{2.14}$$

Вопрос о том, какая линия определяется этим уравнением, решает

**Теорема 2.1.** Каждое уравнение первой степени относительно декартовых координат определяет прямую. Обратно, каждая прямая на плоскости в фиксированной декартовой прямоугольной системе координат определяется уравнением первой степени.

Доказательство. Если $B \neq 0$, то, разрешив (2.13) относительно $y$, получим уравнение

$$y = -\frac{A}{B}x - \frac{C}{B}, \text{ или } y = kx + b,$$

где

$$k = -\frac{A}{B}, \quad b = -\frac{C}{B}, \tag{2.15}$$

которое, как известно, является уравнением прямой с угловым коэффициентом. Если $B=0$, то в силу условия (2.14) $A \neq 0$, поэтому уравнение (2.13) можно привести к виду $x=a$, где $a = -\frac{C}{A}$. В этом случае уравнение определяет прямую, параллельную оси $Oy$. Следовательно, уравнение (2.13) определяет некоторую прямую на плоскости.

Обратно, если данная прямая пересекает ось $Oy$, то она имеет уравнение $y=kx+b$ или $kx-y+b=0$; если прямая параллельна оси $Oy$, то она определяется уравнением $x=a$, или $x-a=0$. Каждое из этих уравнений является уравнением первой степени относительно декартовых координат.

Доказанная теорема означает, что линии первого порядка — прямые.

Уравнение (2.13) называется общим уравнением прямой. Первая из формул (2.15) определяет угловой коэффициент этой прямой, если $B \neq 0$. В случае $B=0$ уравнение (2.13) принимает вид $x = a$ $\left(a = -\frac{C}{A}\right)$ и определяет прямую, параллельную оси $Oy$ $(\alpha = 90°,\ \operatorname{tg}\alpha = \infty)$; такая прямая не имеет углового коэффициента.

Замечание. Коэффициенты $A$, $B$ в общем уравнении прямой (2.13), где $A$ и $B$ не равны нулю одновременно, являются координатами ненулевого вектора $\mathbf{n} = (A, B)$, перпендикулярного этой прямой. Действительно, пусть $M_0(x_0, y_0)$ — фиксированная точка прямой (2.13), тогда $Ax_0+By_0+C=0$. Вычитая это равенство из уравнения (2.13), получаем равносильное уравнение $A(x-x_0)+B(y-y_0)=0$, или $\frac{x-x_0}{B} = \frac{y-y_0}{-A}$.

Рассмотрим вектор $\overline{M_0M}=(x-x_0,\ y-y_0)$, где $M(x,\ y)$ — произвольная точка прямой $Ax+By+C=0$ (рис. 2.4), векторы $\mathbf{n}=(A,\ B)$ и $\mathbf{a}=(B,\ -A)$. Первое из полученных равенств означает перпендикулярность векторов $\mathbf{n}$ и $\overline{MM_0}$ (их скалярное произведение равно нулю), а второе — параллельность векторов $\mathbf{a}$ и $\overline{MM_0}$ (их координаты пропорциональны). Отметим, что $\mathbf{n}\perp\mathbf{a}$, т. е. $\mathbf{na}=0$, так как $A\cdot B+B(-A)=0$.

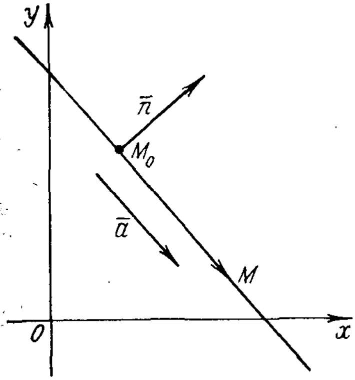



Рис. 2.4

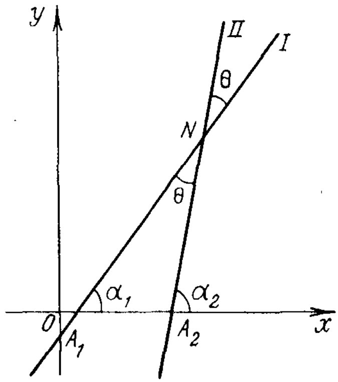



Рис. 2.5

**Угол между двумя прямыми. Условия параллельности и перпендикулярности двух прямых.** Рассмотрим две прямые; предположим, что ни одна из них не параллельна оси $Oy$ (рис. 2.5). В этом случае прямые можно задать их уравнениями с угловыми коэффициентами

$$y=k_1x+b_1, \tag{2.16}$$

$$y=k_2x+b_2, \tag{2.17}$$

где

$$k_1=\operatorname{tg}\alpha_1,\ k_2=\operatorname{tg}\alpha_2 \tag{2.18}$$

(в силу предположения $\alpha_1\neq 90°$, $\alpha_2\neq 90°$, $k_1\neq\infty$, $k_2\neq\infty$).

Обозначим через $\Theta$ угол наклона второй прямой к первой, т. е. угол, на который нужно повернуть вокруг точки пересечения первую из них, чтобы она совпала со второй. Выведем формулу, по которой можно найти угол $\Theta$, если известны $k_1$ и $k_2$. Из треугольника $A_1A_2N$ (рис. 2.5) следует, что $\alpha_1+\Theta=\alpha_2$, $\Theta=\alpha_2-\alpha_1$, поэтому

$$\operatorname{tg}\Theta=\operatorname{tg}(\alpha_2-\alpha_1)=\frac{\operatorname{tg}\alpha_2-\operatorname{tg}\alpha_1}{1+\operatorname{tg}\alpha_1\cdot\operatorname{tg}\alpha_2}.$$

Подставив выражения (2.18) в последнее равенство, получим искомую формулу

$$\operatorname{tg}\Theta = \frac{k_2 - k_1}{1 + k_1 k_2}. \qquad (2.19)$$

Если одна из прямых параллельна оси $Oy$ (рис. 2.6), то угол $\Theta$ находится непосредственно. (Если $\alpha_1 = \alpha$, $\alpha_2 = \frac{\pi}{2}$, то $\Theta = \frac{\pi}{2} - \alpha$.)

Замечание. При другой нумерации прямых (замене $k_1$ на $k_2$ и $k_2$ на $k_1$) правая часть формулы (2.19) меняет знак; в этом случае формула определяет тангенс другого угла $\Theta'$ между двумя прямыми, причем $\Theta + \Theta' = 180°$.

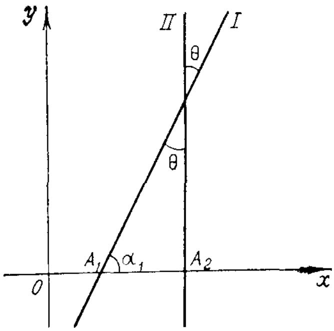



Рис. 2.6

Необходимое и достаточное условие параллельности прямых (2.16) и (2.17) выражается равенством

$$k_1 = k_2. \qquad (2.20)$$

Действительно, если прямые параллельны, то $\alpha_1 = \alpha_2$, $\operatorname{tg}\alpha_1 = \operatorname{tg}\alpha_2$, т. е. выполнено равенство (2.20). Обратно, если выполнено равенство (2.20), то $\operatorname{tg}\alpha_1 = \operatorname{tg}\alpha_2$ и, так как углы заключены между 0 и $\pi$, $\alpha_1 = \alpha_2$, т. е. прямые параллельны.

Пусть прямые, заданные уравнениями (2.16) и (2.17), перпендикулярны, т. е. $\Theta = \frac{\pi}{2}$, в этом случае $\operatorname{ctg}\Theta = 0$, следовательно,

$$\operatorname{ctg}\Theta = \frac{1}{\operatorname{tg}\Theta} = \frac{1 + k_1 k_2}{k_2 - k_1} = 0, \quad 1 + k_1 k_2 = 0,$$

откуда

$$k_1 = -\frac{1}{k_2}. \qquad (2.21)$$

Обратно, если выполнено условие (2.21), то $k_1 k_2 + 1 = 0$, $\operatorname{ctg}\Theta = 0$, $\Theta = \frac{\pi}{2}$, т. е. прямые перпендикулярны. Итак, равенство (2.21) выражает необходимое и достаточное условие перпендикулярности прямых (2.16) и (2.17). Это равенство формулируется следующим образом: угловые коэффициенты перпендикулярных прямых обратны по величине и противоположны по знаку.

Если прямые заданы общими уравнениями

$$A_1x+B_1y+C_1=0; \tag{2.22}$$

$$A_2x+B_2y+C_2=0, \tag{2.23}$$

то тангенс угла между ними определяется формулой

$$\operatorname{tg}\Theta=\frac{A_1B_2-A_2B_1}{A_1A_2+B_1B_2}. \tag{2.24}$$

В самом деле, разрешив уравнения (2.22), (2.23) относительно $y$ и сравнив их соответственно с уравнениями (2.16), (2.17), получим выражения для угловых коэффициентов

$$k_1=-\frac{A_1}{B_1},\quad k_2=-\frac{A_2}{B_2}. \tag{2.25}$$

Формула (2.24) следует из формулы (2.19) и равенств (2.25).

Необходимое и достаточное условие параллельности прямых (2.22) и (2.23) выражается равенством

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}, \tag{2.26}$$

или

$$A_1=lA_2,\ B_1=lB_2, \tag{2.27}$$

а условие их перпендикулярности — равенством

$$-\frac{A_1}{B_1}=\frac{B_2}{A_2},\ \text{или}\ A_1A_2+B_1B_2=0. \tag{2.28}$$

Отметим, что прямые, заданные соответственно уравнениями

$$Ax+By+C=0,\ Bx-Ay+C=0,$$

взаимно перпендикулярны, так как для них выполнено условие (2.28).

Пример 2. Найти угол между прямыми, определяемыми уравнениями $3x-5y+15=0$, $8x-2y-1=0$.

Согласно условию $A_1=3$, $B_1=-5$, $A_2=8$, $B_2=-2$; по формуле (2.24) находим

$$\operatorname{tg}\Theta=\frac{3(-2)-8(-5)}{3\cdot 8+(-5)(-2)}=\frac{34}{34}=1,\ \Theta=45^\circ. \tag{2.29}$$

Замечание. При другой нумерации прямых ($A_1=8$, $B_1=-2$, $A_2=3$, $B_2=-5$) получаем $\operatorname{tg}\Theta'=-1$, $\Theta'=135^\circ$. Очевидно, $\Theta+\Theta'=180^\circ$ (сделайте чертеж).

**Расстояние от точки до прямой.** Дана прямая своим общим уравнением (2.13), т. е. уравнением $Ax+By+C=0$, и точка $M_1(x_1, y_1)$ относительно прямоугольной декартовой системы координат (рис. 2.7). Требуется вывести формулу для вычисления расстояния от точки $M_1(x_1, y_1)$ до прямой $Ax+By+C=0$. Это расстояние равно длине отрезка перпендикуляра, опущенного из точки $M_1$ на данную прямую. Пусть $M_2(x_2, y_2)$ — основание этого перпендикуляра, тогда искомое расстояние $d$ по определению выражается формулой

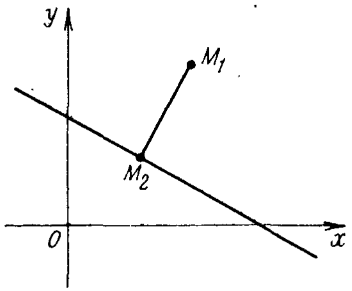


Рис. 2.7

$$d=|M_1M_2|. \qquad (2.30)$$

Прямая (2.13) имеет угловой коэффициент $k=-\frac{A}{B}$ (см. первую из формул (2.15)), а прямая, перпендикулярная ей,— угловой коэффициент $k_1=\frac{B}{A}$, что следует из условия (2.21). Выразим угловой коэффициент прямой, проходящей через точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$, по формуле (2.7), обозначим равные отношения буквой $t$ и найдем $x_2$, $y_2$:

$$\frac{y_2-y_1}{x_2-x_1}=\frac{B}{A}, \quad \frac{y_2-y_1}{B}=\frac{x_2-x_1}{A}=t; \ x_2=x_1+At,$$

$$y_2=y_1+Bt.$$

Запишем формулу (1.28) для точек $M_1(x_1, y_1)$, $M_2(x_1+At, y_1+Bt)$:

$$d=|M_1M_2|=\sqrt{(At)^2+(Bt)^2}=\sqrt{A^2+B^2}\,|t|,$$

$$|M_1M_2|=\sqrt{A^2+B^2}\,|t|. \qquad (2.31)$$

Определим значение $t$. Так как точка $M_2$ лежит на данной прямой $Ax+By+C=0$, то ее координаты удовлетворяют данному уравнению, т. е. $Ax_2+By_2+C=0$, или $A(x_1+At)+B(y_1+Bt)+C=0$, откуда $Ax_1+By_1+C+A^2t+B^2t=0$, $Ax_1+By_1+C+(A^2+B^2)t=0$,

$$t=-\frac{Ax_1+By_1+C}{A^2+B^2}, \qquad (2.32)$$

поскольку $A^2+B^2\neq 0$ в силу условия (2.14). Из равенств (2.30)—(2.32) следует искомая формула для вычисления расстояния от точки $M_1(x_1, y_1)$ до прямой $Ax+By+C=0$:

$$d=\frac{|Ax_1+By_1+C|}{\sqrt{A^2+B^2}}. \tag{2.33}$$

Пример. Найти расстояние от точки $M_1(-6, 3)$ до прямой, заданной уравнением $3x-4y+15=0$.

По формуле (2.33) получаем

$$d=\frac{|3\cdot(-6)-4\cdot 3+15|}{\sqrt{3^2+(-4)^2}}=\frac{|-18-12+15|}{5}=\frac{|-15|}{5}=3.$$

## § 2.2. Общее уравнение окружности

Рассмотрим алгебраическое уравнение второй степени относительно декартовых координат $x$ и $y$:

$$Ax^2+2Bxy+Cy^2+Dx+Ey+F=0; \tag{2.34}$$

предполагается, что коэффициенты $A$, $B$, $C$ одновременно в нуль не обращаются, т. е.

$$A^2+B^2+C^2\neq 0. \tag{2.35}$$

В § 1.9 было записано каноническое уравнение окружности (см. уравнение (1.36)). Если в этом уравнении раскроем скобки, то получим уравнение второй степени вида (2.34), в котором $A=C=1$, $B=0$.

Рассмотрим теперь уравнение (2.34) при $A=C$ и $B=0$ (т. е. уравнение второй степени, имеющее равные коэффициенты при квадратах координат и не содержащее члена с произведением координат). Вопрос в том, какую линию определяет указанное уравнение, решает

**Теорема 2.2.** Если уравнение

$$Ax^2+Ay^2+Dx+Ey+F=0 \tag{2.36}$$

относительно декартовых прямоугольных координат $x$ и $y$ определяет некоторую линию на плоскости, то этой линией является окружность.

Доказательство. Разделив почленно уравнение (2.36) на $A$ ($A\neq 0$ в силу условия (2.35)), получим

$$x^2+y^2+dx+ey+f=0, \tag{2.37}$$

где $d=\frac{D}{A}$, $e=\frac{E}{A}$, $f=\frac{F}{A}$.

Преобразуем уравнение (2.37), выделив в левой его части полные квадраты:

$$\left(x^2 + 2\frac{d}{2}x + \frac{d^2}{4}\right) + \left(y^2 + 2\frac{e}{2}y + \frac{e^2}{4}\right) +$$

$$+ f - \frac{d^2}{4} - \frac{e^2}{4} = 0,$$

$$\left(x + \frac{d}{2}\right)^2 + \left(y + \frac{e}{2}\right)^2 = \frac{d^2}{4} + \frac{e^2}{4} - f. \qquad (2.38)$$

Алгебраическая сумма в правой части уравнения (2.38) может быть положительной, равной нулю или отрицательной. Исследуем каждый случай в отдельности.

1. Если $\frac{d^2}{4} + \frac{e^2}{4} - f > 0$, то, введя обозначения

$$\frac{d^2}{4} + \frac{e^2}{4} - f = R^2, \ \frac{d}{2} = -a, \ \frac{e}{2} = -b,$$

получим уравнение

$$(x-a)^2 + (y-b)^2 = R^2, \qquad (2.39)$$

определяющее окружность радиуса $R$ с центром в точке $C(a, b)$.

2. Если $\frac{d^2}{4} + \frac{e^2}{4} - f = 0$, уравнение (2.38) принимает вид

$$(x-a)^2 + (y-b)^2 = 0. \qquad (2.40)$$

Этому уравнению удовлетворяют координаты единственной точки $x = a$, $y = b \left(a = -\frac{d}{2}, \ b = -\frac{e}{2}\right)$.

3. Если $\frac{d^2}{4} + \frac{e^2}{4} - f < 0$, то, положив $\frac{e}{2}\Big)^{2}_{1} + \frac{e^2}{4} - f = -R^2$, получим

$$(x-a)^2 + (y-b)^2 = -R^2. \qquad (2.41)$$

Уравнению (2.41) не удовлетворяют координаты ни одной точки плоскости, линии оно не определяет.

Итак, уравнение (2.36) либо не определяет никакой линии, либо определяет точку, либо определяет окружность. Теорема доказана.

Пример 1. Найти центр и радиус окружности, определяемой уравнением $4x^2 + 4y^2 - 8x + 16y + 19 = 0$.

Преобразуем данное уравнение:

$$x^2+y^2-2x+4y+\frac{19}{4}=0,\ (x^2-2x+1)+(y^2+4y+4)-1-$$
$$-4+\frac{19}{4}=0,\ (x-1)^2+(y+2)^2-\frac{1}{4}=0,\ (x-1)^2+(y+2)^2=$$
$$=\frac{1}{4}.$$

Сравнивая последиее уравнение с уравнением (2.39), заключаем, что центр окружности находится в точке $C(1,\ -2)$ и ее радиус $R=\frac{1}{2}$.

Пример 2. Какое множество точек плоскости определяет уравнение $x^2+y^2+6x-8y+25=0$?

Так как это уравнение сводится к уравнению $(x+3)^2+(y-4)^2=0$, которому удовлетворяют лишь координаты $x=-3$, $y=4$, то оно определяет единственную точку $C(-3,\ 4)$.

## § 2.3. Эллипс

**Определение эллипса и вывод канонического уравнения.** *Эллипсом* называется множество всех точек плоскости, для каждой из которых сумма расстояний до двух данных точек (фокусов) той же плоскости есть величина постоянная (большая, чем расстояние между фокусами). Обозначим фокусы буквами $F_1$ и $F_2$, расстояние между ними — через $2c$, т. е.

$$|F_1F_2|=2c, \qquad (2.42)$$

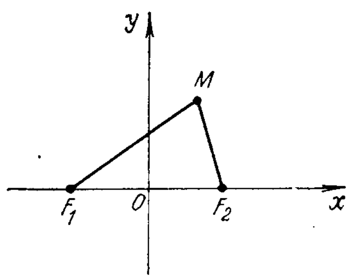


Рис. 2.8

и назовем *фокусным расстоянием*. Постоянную величину, о которой идет речь в определении эллипса, обозначим через $2a$. Пусть $M$ — произвольная точка эллипса, тогда по определению

$$|F_1M|+|F_2M|=2a, \qquad (2.43)$$

где $|F_1M|$, $|F_2M|$ — длины соответствующих отрезков.

Из треугольника $F_1MF_2$ (рис. 2.8) следует, что $|F_1M|+|F_2M|>|F_1F_2|$. Это неравенство, с учетом (2.42) и (2.43), принимает вид

$$2a>2c,\ a>c. \qquad (2.44)$$

Составим уравнение эллипса относительно некоторой системы декартовых прямоугольных координат. Выберем ось $Ox$ так, чтобы она проходила через фокусы (рис. 2.8) и имела положительное направление от $F_1$ к $F_2$. Начало координат поместим в середине отрезка $[F_1F_2]$, тогда $F_1(-c, 0)$, $F_2(c, 0)$; текущие координаты точки $M$ обозначим через $x$, $y$.

На основании формулы (1.28)

$$|F_1M|=\sqrt{(x+c)^2+y^2},\quad |F_2M|=\sqrt{(x-c)^2+y^2}. \tag{2.45}$$

Подставив выражения (2.45) в равенство (2.43), получим

$$\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2}=2a. \tag{2.46}$$

Уравнение (2.46) является уравнением эллипса, так как ему удовлетворяют координаты любой точки эллипса и только они. Упростим это уравнение. Перенесем один из корней в правую часть, возведем в квадрат полученное уравнение и приведем подобные члены:

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

Возводя в квадрат последнее уравнение, получаем

$$a^2(x^2-2cx+c^2+y^2)=a^4-2a^2cx+c^2x^2,$$
$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2). \tag{2.47}$$

Так как $a^2-c^2>0$ (см. условие (2.44)), то можно ввести обозначение

$$b^2=a^2-c^2; \tag{2.48}$$

уравнение (2.47) принимает вид $b^2x^2+a^2y^2=a^2b^2$ или

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{2.49}$$

Таким образом, координаты любой точки эллипса удовлетворяют уравнению (2.49). Покажем обратное: если координаты точки $M$ удовлетворяют уравнению (2.49), то точка $M$ лежит на эллипсе, т. е. выполняется равенство (2.43). Из уравнения (2.49) найдем $y^2$, подставим его выражение в формулы (2.45) и воспользуемся равенством (2.48):

$$y^2 = b^2\left(1 - \frac{x^2}{a^2}\right), \ |F_1M| = \sqrt{(x+c)^2 + y^2} =$$
$$= \sqrt{x^2 + 2cx + c^2 + b^2 - \frac{b^2x^2}{a^2}} = \sqrt{\frac{c^2x^2}{a^2} + 2cx + a^2} =$$
$$= \sqrt{\left(\frac{cx}{a} + a\right)^2} = \pm\left(a + \frac{c}{a}x\right).$$

Знак нужно выбрать так, чтобы правая часть была положительной. Поскольку $c < a$ и $|x| \leqslant a$ (последнее следует из уравнения (2.49)), необходимо выбрать знак плюс, т. е.

$$|F_1M| = a + \frac{c}{a}x. \tag{2.50}$$

Аналогично можно получить формулу

$$|F_2M| = a - \frac{c}{a}x. \tag{2.51}$$

Из последних двух равенств следует равенство (2.43), а это означает, что точка $M$ лежит на эллипсе. Следовательно, уравнение (2.49) является уравнением эллипса; оно называется *каноническим уравнением эллипса*. Поскольку это уравнение второй степени, то эллипс — линия второго порядка.

**Исследование формы эллипса.** Из уравнения (2.49) следует, что $x^2 \leqslant a^2$, т. е. $|x| \leqslant a$, $-a \leqslant x \leqslant a$ и $y^2 \leqslant b^2$, т. е. $-b \leqslant y \leqslant b$, а это означает, что эллипс расположен целиком в прямоугольнике, основание которого равно $2a$, высота — $2b$, центр находится в начале координат (рис. 2.9). Поскольку в уравнение (2.49) $x$ входит только во второй (четной) степени, то эллипс симметричен относительно оси $Oy$. Действительно, если точка $M_1(x_1, y_1)$ лежит на эллипсе, то точка $M_2(-x_1, y_1)$ также лежит на эллипсе, ибо

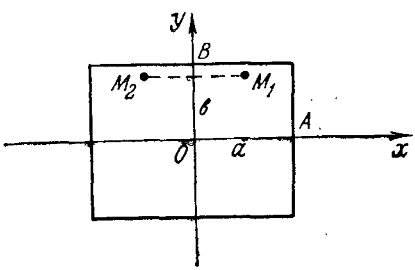


Рис. 2.9

$$\frac{(-x_1)^2}{a^2} + \frac{y_1^2}{b^2} = \frac{x_1^2}{a^2} + \frac{y_1^2}{b^2} = 1,$$

а точки $M_1$ и $M_2$ симметричны относительно оси $Oy$ (рис. 2.9). Аналогично доказывается, что эллипс симметричен относительно оси $Ox$, так как в уравнение (2.49) $y$ входит только в четной степени.

В силу симметрии эллипса относительно координатных осей достаточно исследовать его форму в первой четверти. Разрешив уравнение (2.49) относительно $y$, получим

$$y = -\frac{b}{a}\sqrt{a^2 - x^2}, \quad y = +\frac{b}{a}\sqrt{a^2 - x^2}. \qquad (2.52)$$

Первое из уравнений определяет половину эллипса, расположенную ниже оси $Ox$, второе — другую полови-

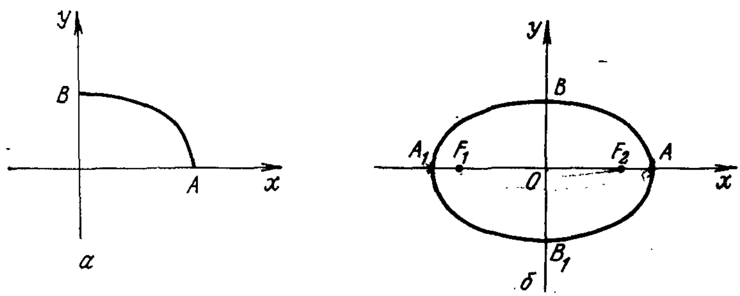



Рис. 2.10

ну, расположенную выше этой оси. Рассмотрим второе из уравнений (2.52). Если $x$ возрастает от 0 до $a$, то $y$ убывает от $b$ до 0; при $x=0$, $y=b$ получаем точку $B(0, b)$, при $x=a$, $y=0$ — точку $A(a, 0)$. Соответствующая дуга эллипса, расположенная в первой четверти, изображена на рис. 2.10, *а*, а весь эллипс — на рис. 2.10, *б*.

Точки пересечения эллипса с координатными осями называются *вершинами эллипса* (точки $A$, $A_1$, $B$, $B_1$ на рис. 2.10). Оси симметрии эллипса (оси $Ox$ и $Oy$) называются просто *осями*, точка пересечения осей — центром эллипса. Осями называются также отрезки $A_1A=2a$, $B_1B=2b$, полуосями — отрезки $OA=a$, $OB=b$ и их длины. В случае, когда фокусы расположены на оси $Ox$, $a>b$ (это следует из равенства (2.48)); отрезок $OA=a$ называют *большой полуосью,* отрезок $OB=b$ — *малой полуосью.*

Уравнение (2.49) можно рассматривать и в случае

$b > a$, оно определяет эллипс с большой полуосью $OB = b$, фокусы такого эллипса лежат на оси $Oy$.

В случае $b = a = R$ уравнение (2.49) принимает вид $x^2 + y^2 = R^2$ и определяет окружность радиуса $R$ с центром в начале координат.

**Эксцентриситет эллипса.** *Эксцентриситетом* эллипса называется отношение фокусного расстояния к длине большой оси. В случае $a > b$ эксцентриситет эллипса (2.49) выражается формулой

$$\varepsilon = \frac{c}{a}. \tag{2.53}$$

Так как для эллипса $0 < c < a$, то $0 < \varepsilon < 1$ (для окружности $\varepsilon = 0$, так как $c = 0$).

Принимая во внимание равенство (2.48), получаем

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

Из этих формул следует, что эксцентриситет характеризует форму эллипса; чем больше эксцентриситет, тем более вытянут эллипс.

**Фокальные радиусы.** *Фокальными* радиусами точки $M$ эллипса называются отрезки прямых, соединяющих эту точку с фокусами $F_1$ и $F_2$ данного эллипса. Их длины $r_1$ и $r_2$ определяются формулами (2.50) и (2.51); с учетом (2.53) эти формулы принимают вид

$$r_1 = a + \varepsilon x, \quad r_2 = a - \varepsilon x. \tag{2.54}$$

**Эллипс как проекция окружности на плоскость.** Окружность радиуса $a$, лежащую в плоскости $\Pi$, спроектируем на плоскость $\Pi'$, острый угол между плоскостями обозначим через $\varphi$ (рис. 2.11). Предположим, что пло-

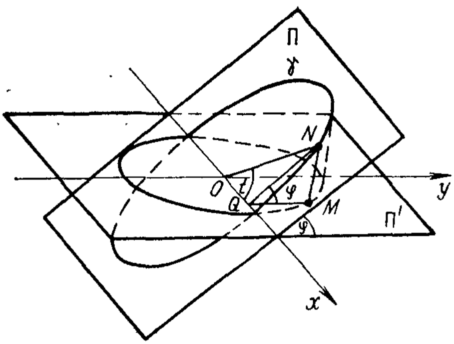



Рис. 2.11

Аналогично находим $|F_2M| = \pm\left(\frac{c}{a}x - a\right)$. В этих формулах знаки нужно выбрать так, чтобы правые части были неотрицательны. Поскольку $|x| \geqslant a$ (это следует из уравнения (2.59)) и $a < c$ (неравенство (2.57)), то для $x > a$

$$|F_1M| = \frac{c}{a}x + a, \quad |F_2M| = \frac{c}{a}x - a, \tag{2.61}$$

поэтому $|F_1M| - |F_2M| = 2a$. Для $x < -a$

$$\left.\begin{aligned} |F_1M| &= -\left(\frac{c}{a}x + a\right) = -\frac{c}{a}x - a, \\ |F_2M| &= -\left(\frac{c}{a}x - a\right) = -\frac{c}{a}x + a, \end{aligned}\right\} \tag{2.62}$$

следовательно, $|F_1M| - |F_2M| = -2a$.

Итак, уравнения (2.58) и (2.59) равносильны. Уравнение (2.59) называется каноническим уравнением гиперболы. Так как это уравнение второй степени, то гипербола — линия второго порядка.

**Исследование формы гиперболы. Асимптоты.** Из уравнения (2.59) следует, что $x^2 \geqslant a^2$, т. е. $x < -a$ и $x > a$. Это означает, что в полосе между прямыми $x = -a$ и $x = a$ нет ни одной точки гиперболы. Поскольку в уравнение (2.59) $x$ и $y$ входят только во второй (четной) степени, то гипербола симметрична относительно координатных осей; достаточно изучить ее форму в первой четверти, в которой она определяется уравнением $y = \frac{b}{a}\sqrt{x^2 - a^2}$. При $x = a$ получаем $y = 0$, точка $A(a, 0)$ принадлежит гиперболе (рис. 2.13). Из уравнения видно, что $y$ неограниченно возрастает при неограниченном возрастании $x$. Покажем, что точка дуги гиперболы неограниченно приближается к прямой, определяемой уравнением

$$y = \frac{b}{a}x, \tag{2.63}$$

при ее неограниченном удалении от начала координат.

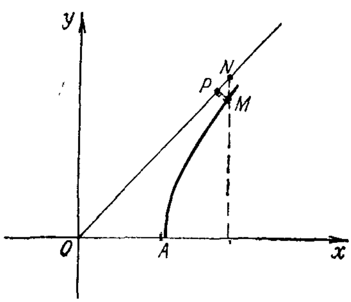


Рис. 2.13

Фиксируем некоторое значение $x$, ему будет соответствовать точка $M(x, y)$ гиперболы и точка $N(x, Y)$ прямой (2.63). Из точки $M$ опустим перпендикуляр $MP$ на прямую. Отметим, что дуга гиперболы подходит к прямой снизу, так как

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

длина отрезка $MN$ выразится формулой

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

Выясним, как изменяется длина отрезка $MN$, когда $x$ неограниченно возрастает, для чего предварительно преобразуем правую часть последнего равенства:

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

$$= \frac{b}{a} \frac{[x^2 - (x^2 - a^2)]}{x + \sqrt{x^2 - a^2}}, \quad |MN| = \frac{ab}{x + \sqrt{x^2 - a^2}}.$$

Из этого равенства видно, что длина отрезка $MN$ стремится к нулю, когда $x$ неограниченно возрастает. Так как $|MP| < |MN|$, то $|MP|$ также стремится к нулю, т. е. расстояние от точки $M$ до прямой стремится к нулю, когда эта точка неограниченно удаляется от начала координат. Отметим, что гипербола (2.59) и прямая (2.63) общих точек не имеют, так как система их уравнений не имеет решений.

Прямая, определяемая уравнением (2.63), называется *асимптотой гиперболы*. Легко видеть, что гипербола (2.59) имеет две асимптоты:

$$y = \frac{b}{a} x, \quad y = -\frac{b}{a} x.$$

Изобразим гиперболу на чертеже. Построим сначала так называемый основной прямоугольник, центр которого совпадает с началом координат, стороны равны $2a$ и $2b$, параллельны соответственно осям $Ox$ и $Oy$. Прямые, проходящие через противоположные вершины основного прямоугольника, являются асимптотами гиперболы. Построив асимптоты, строим саму гиперболу (рис. 2.14); она состоит из двух частей, называемых ветвями (левой и правой). Разумеется, на рис. 2.14 изображены лишь конечные дуги ветвей (сами ветви простираются неограниченно). Центр симметрии гиперболы называется ее *центром*. Оси симметрии гиперболы называются просто ее *осями*; одна ось пересекает гиперболу в двух точках, называемых *вершинами* (на рис. 2.14 точки $A_1$ и $A_2$), эта ось называется *действительной осью гиперболы*, другая ось — *мнимой осью*, она не имеет общих точек с гиперболой. Длины отрезков $|A_1A_2|=2a$ и $|B_1B_2|=2b$ также называ-

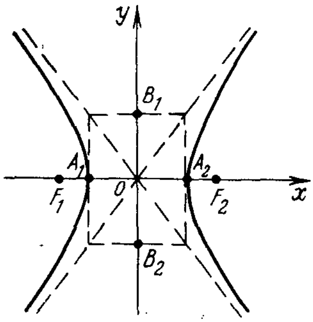



Рис. 2.14

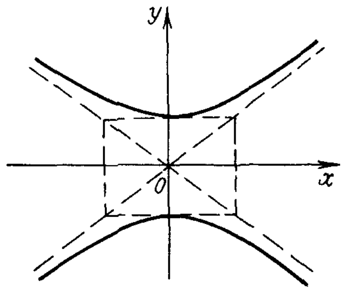



Рис. 2.15

ются *осями*. Величины $a$ и $b$ называются *полуосями гиперболы*. Если $a=b$, гипербола называется *равносторонней*, ее уравнение

$$x^2-y^2=a^2. \tag{2.64}$$

Уравнение

$$-\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \tag{2.65}$$

определяет гиперболу с действительной осью $Oy$ (рис. 2.15).

Гиперболы, определяемые уравнениями (2.59) и (2.65) в одной и той же системе координат с одинаковыми значениями $a$ и $b$, называются *сопряженными*.

**Эксцентриситет гиперболы.** *Эксцентриситетом* гиперболы называется отношение ее фокусного расстояния к расстоянию между ее вершинами. Если действительной осью является ось $Ox$, то по определению

$$\varepsilon=\frac{c}{a}. \tag{2.66}$$

Так как для гиперболы $c>a$, то $\varepsilon>1$. Приняв во внимание формулу (2.60), получим

$$\varepsilon=\sqrt{1+\left(\frac{b}{a}\right)^2},\quad \frac{b}{a}=\sqrt{\varepsilon^2-1}.$$

Следовательно, эксцентриситет гиперболы характеризует форму основного прямоугольника и форму самой гиперболы.

**Фокальные радиусы гиперболы.** *Фокальными радиусами* точки $M$ гиперболы называются отрезки, соединяющие эту точку с фокусами данной гиперболы. Их длины $r_1=|F_1M|$, $r_2=|F_2M|$ выражаются формулами (2.61), (2.62). С учетом (2.66) эти формулы для правой ветви гиперболы принимают вид

$$r_1=\varepsilon x+a,\ r_2=\varepsilon x-a, \tag{2.67}$$

для левой ветви

$$r_1=-\varepsilon x-a,\ r_2=-\varepsilon x+a. \tag{2.68}$$

Пример. Найти полуоси, координаты фокусов и эксцентриситет гиперболы, заданной уравнением $5x^2-4y^2=20$. Вычислить длины фокальных радиусов точки $M(-4, \sqrt{15})$.

Разделив обе части уравнения на 20, получим $\frac{x^2}{4}-\frac{y^2}{5}=1$. Сравнивая это уравнение с уравнением (2.59), заключаем, что $a^2=4$, $b^2=5$, т. е. $a=2$, $b=\sqrt{5}$. Так как $b^2=c^2-a^2$, то $c^2=a^2+b^2=9$, $c=3$, $F_1(-3,\ 0)$, $F_2(3,\ 0)$, $\varepsilon=\frac{c}{a}=\frac{3}{2}$. Поскольку точка $M$ лежит на левой ветви гиперболы, при вычислении $r_1$ и $r_2$ необходимо пользоваться формулами (2.68): $r_1=-\frac{3}{2}\times(-4)-2=4$, $r_2=-\frac{3}{2}(-4)+2=8$. Отметим, что $r_2-r_1=8-4=4=2a$.

## § 2.5. Директрисы эллипса и гиперболы

*Директрисами эллипса* называются две прямые, перпендикулярные большой оси эллипса и расположенные симметрично относительно центра на расстоянии $a:\varepsilon$ от него ($a$ — большая полуось, $\varepsilon$ — эксцентриситет эллипса). Если эллипс задан каноническим уравнением (2.59), причем $a>b$, то в выбранной системе координат его директрисы определяются уравнениями

$$x=-\frac{a}{\varepsilon},\ x=\frac{a}{\varepsilon}. \tag{2.69}$$

Так как для эллипса $0<\varepsilon<1$, то $\frac{a}{\varepsilon}>a$. Это означает, что директрисы эллипса не имеют с ним общих точек (рис. 2.16).

*Директрисами гиперболы* называются прямые, перпендикулярные действительной оси гиперболы и расположенные симметрично относительно центра на расстоянии $a:\varepsilon$ от него ($a$ — действительная полуось, $\varepsilon$ — эксцентриситет гиперболы). Если гипербола задана каноническим уравнением (2.59), то в данной системе координат ее директрисы определяются уравнениями (2.69). Поскольку для гиперболы $\varepsilon > 1$, то $\frac{a}{\varepsilon} < a$; это означает, что директрисы гиперболы не имеют с ней общих точек (рис. 2.17).

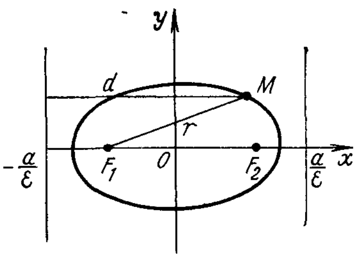



Рис. 2.16

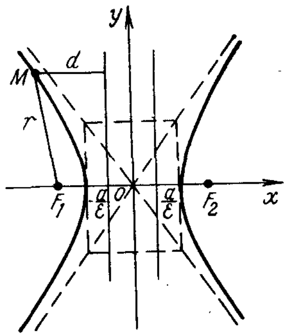



Рис. 2.17

Свойство директрис эллипса и гиперболы выражается следующей теоремой.

**Теорема 2.3.** Отношение расстояния $r$ произвольной точки эллипса (гиперболы) до фокуса к расстоянию $d$ этой точки до соответствующей директрисы есть постоянная величина, равная эксцентриситету эллипса (гиперболы), т. е. $r/d = \varepsilon$.

Д о к а з а т е л ь с т в о. Рассмотрим, например, левый фокус и левую директрису. Пусть $M(x, y)$ — произвольная точка эллипса (рис. 2.16), тогда

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

$$\frac{r}{d} = \frac{a + \varepsilon x}{x + \frac{a}{\varepsilon}} = \frac{a + \varepsilon x}{\frac{a + \varepsilon x}{\varepsilon}} = \varepsilon.$$

Если $M(x, y)$ — произвольная точка левой ветви ги-

перболы (рис. 2.17), то $r=-\varepsilon x-a$, $d=-x-\frac{a}{\varepsilon}$, $\frac{r}{d}=$
$=\frac{-\varepsilon x-a}{-x-\frac{a}{\varepsilon}}=\varepsilon$.

Все остальные случаи рассматриваются аналогично.

Замечание. Свойство, выраженное теоремой 2.3, можно положить в основу определения эллипса и гиперболы. Множество точек, для которых отношение расстояния $r$ до фиксированной точки (фокуса) к расстоянию $d$ до фиксированной прямой (директрисы) есть постоянная $\varepsilon$, оказывается эллипсом, если $\varepsilon<1$, гиперболой, если $\varepsilon>1$. Чтобы убедиться в справедливости этого утверждения, необходимо определенным образом выбрать систему декартовых координат, получить уравнение указанного множества точек и установить, что оно является уравнением эллипса при $\varepsilon<1$ и уравнением гиперболы при $\varepsilon>1$. Вопрос о том, что представляет собой это множество точек в случае $\varepsilon=1$, рассматривается в § 2.6.

## § 2.6. Парабола

Параболой называется множество всех точек плоскости, равноудаленных от данной точки (фокуса) и данной прямой (директрисы).

Пусть $p$ — расстояние от фокуса $F$ до директрисы $\Delta$. Ось $Ox$ декартовой прямоугольной системы координат выберем так, чтобы она проходила через $F$ перпендикулярно $\Delta$, ее положительное направление — от $\Delta$ к $F$ (рис. 2.18), начало координат поместим в середине отрезка $BF$, где $B$ — точка пересечения $Ox$ и $\Delta$. В этой системе координат точки $F$ и $B$ имеют следующие координаты: $F(p/2;\ 0)$, $B(-p/2;\ 0)$.

Возьмем произвольную точку $M(x,\ y)$ параболы, обозначим через $r$ расстояние до фокуса, через $d$ — расстояние до директрисы ($r=|MF|$, $d=|NF|$), по определению параболы $r=d$. Поскольку

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

то

$$\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}=x+\frac{p}{2}. \qquad (2.70)$$

Уравнение (2.70) является уравнением параболы. Возведя почленно в квадрат это уравнение и приводя подобные члены, получаем

$$y^2=2px. \tag{2.71}$$

Легко видеть, что уравнения (2.70) и (2.71) равносильны.

Уравнение (2.71) называется каноническим уравнением параболы. Из уравнения (2.71) следует, что $x \geqslant 0$ (так как $p>0$), т. е. парабола лежит целиком справа от оси $Oy$ (при указанном выборе оси $Ox$). Парабола симметрична относительно оси $Ox$ (так как $y$ входит в урав-

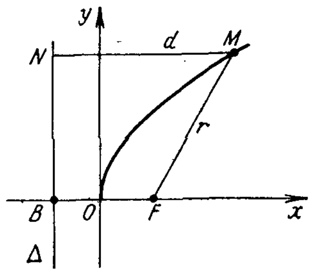



Рис. 2.18

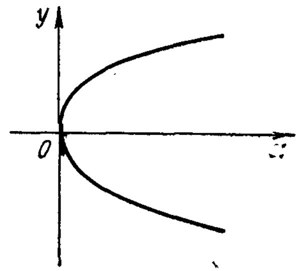



Рис. 2.19

нение только во второй степени); когда $x$ неограниченно возрастает, $y$ также неограниченно возрастает.

Дуга параболы, для которой $y=\sqrt{2px}$, изображена на рис. 2.18, дуга всей параболы — на рис. 2.19.

Уравнение директрисы $\left(\text{как прямой, параллельной оси } Oy \text{ и проходящей через точку } B\left(-\frac{p}{2}, 0\right)\right)$ имеет вид $x=-p/2$.

Замечание. Каждое из уравнений $y^2=-2px$, $x^2=2qy$, $x^2=-2qy$ определяет параболу.

## § 2.7. Полярное уравнение эллипса, гиперболы, параболы

Пусть $\gamma$ — дуга эллипса, гиперболы или параболы (рис. 2.20). Если $r$ — расстояние произвольной точки $M$ этой дуги до фокуса $F$, $d$ — расстояние ее до соответствующей директрисы $\Delta$, то в силу теоремы 2.3 и определения параболы

$$\frac{r}{d}=\varepsilon. \tag{2.72}$$

Проведем через фокус $F$ прямую, перпендикулярную директрисе $\Delta$. Пусть $A$ — точка их пересечения, $N$ — проекция точки $M$ на эту прямую. Через точку $F$ проведем перпендикуляр к прямой $AN$ (оси линии $\gamma$) до пересечения с дугой $\gamma$ в точке $P$. Длину отрезка $FP$ обозначим буквой $p$, т. е.

$$|FP|=p, \tag{2.73}$$

и назовем ее *фокальным параметром линии* $\gamma$.

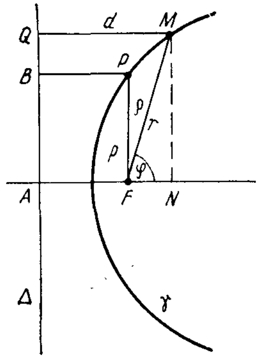



Рис. 2.20

Пусть $\rho$ и $\varphi$ — полярные координаты точки $M$ в системе координат с полюсом в точке $F$ и полярной осью $FN$, тогда

$$r=\rho, \tag{2.74}$$

$$d=QM=AN=AF+\rho\cos\varphi. \tag{2.75}$$

Равенство (2.72) выполняется для всех точек линии $\gamma$, в том числе и для точки $P$, поэтому

$$\frac{FP}{BP}=\varepsilon,\ \frac{p}{AF}=\varepsilon,\ AF=\frac{p}{\varepsilon},$$

следовательно,

$$d=\frac{p}{\varepsilon}+\rho\cos\varphi. \tag{2.76}$$

Подставляя выражения (2.74), (2.76) в равенство (2.72) и преобразуя его, получаем

$$\rho=\frac{p}{1-\varepsilon\cos\varphi}. \tag{2.77}$$

Уравнение (2.77) называется *полярным уравнением* эллипса, гиперболы, параболы (это уравнение определяет одну из двух ветвей гиперболы).

Отметим, что для параболы фокальный параметр совпадает с параметром $p$, входящим в уравнение (2.71). Для эллипса и гиперболы, заданных соответственно уравнениями (2.49) и (2.59), он выражается формулой

$$p=\frac{b^2}{a}. \tag{2.78}$$

Действительно, по определению фокальный параметр равен модулю ординаты точки $P$ (считаем, что ось $AN$ является осью $Ox$). Подставив координаты точки $P(-c, y)$ эллипса в уравнение (2.49), получим

$$\frac{(-c)^2}{a^2}+\frac{y^2}{b^2}=1,\ y^2=b^2\left(1-\frac{c^2}{a^2}\right)=\frac{b^2(a^2-c^2)}{a^2}=\frac{b^4}{a^2},$$

откуда и следует формула (2.78). Тот же результат будет при подстановке координат точки $P(-c, y)$ в уравнение (2.59). Подстановкой координат точки $P\left(-\frac{p}{2}, y\right)$ в уравнение (2.71) проверяется тот факт, что фокальный параметр параболы равен расстоянию от фокуса до директрисы.

Пример. Какую линию определяет уравнение $\rho=\frac{9}{4-5\cos\varphi}$ в полярных координатах?

Разделив числитель и знаменатель правой части на 4, приведем это уравнение к виду (2.77): $\rho=\frac{\frac{9}{4}}{1-\frac{5}{4}\cos\varphi}$, следовательно, $p=\frac{b^2}{a}=\frac{9}{4}$, $\varepsilon=\frac{c}{a}=\frac{5}{4}>1$. Данное уравнение определяет гиперболу с полуосями $a=4$, $b=3$.

## § 2.8. Преобразования декартовых прямоугольных координат

Одна и та же точка имеет различные координаты в разных системах декартовых координат. Установим связь между координатами точки в разных системах координат, рассмотрев предварительно простейшие случаи взаимного расположения двух систем.

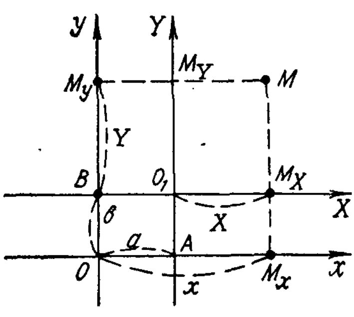


Рис. 2.21

**Параллельный перенос.** Пусть даны две системы декартовых прямоугольных координат с общим масштабным отрезком: $Oxy$ (старая) и $O_1XY$ (новая), соответствующие оси которых параллельны (рис. 2.21), положительные полуоси имеют одинаковые направления. Начало новой системы находится в точке $O_1(a, b)$, старые координаты которой $x=a$, $y=b$ (новые координаты ее равны нулю). Относительно таких систем говорят, что одна получена из другой путем параллельного переноса.

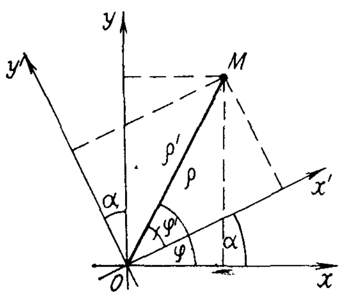



Рис. 2.22

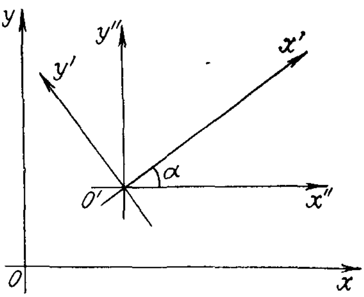



Рис. 2.23

Из основного тождества для трех точек $O$, $A$, $M_x$ (оси $Ox$), $O$, $B$, $M_y$ (оси $Oy$) и определения координат получаем выражения старых координат $x$, $y$ точки $M$ через ее новые координаты и старые координаты нового начала:

$$x=X+a, \quad y=Y+b. \tag{2.79}$$

Очевидно, новые координаты через старые выражаются формулами

$$X=x-a, \quad Y=y-b. \tag{2.80}$$

**Поворот координатных осей.** Новая система $Ox'y'$ получена путем поворота старой на угол $\alpha$ вокруг точки $O$. С каждой из них свяжем полярную систему координат, как показано на рис. 2.22, тогда $\rho'=\rho$, $\varphi=\alpha+\varphi'$.

Принимая во внимание эти равенства и формулы (1.17), выражаем старые декартовы прямоугольные координаты $x$, $y$ точки $M$ через ее новые координаты $x'$, $y'$:

$$x=\rho\cos\varphi=\rho'\cos(\alpha+\varphi')=(\rho'\cos\varphi')\cos\alpha-$$
$$-(\rho'\sin\varphi')\sin\alpha=x'\cos\alpha-y'\sin\alpha,$$
$$y=\rho\sin\varphi=\rho'\sin(\alpha+\varphi')=(\rho'\cos\varphi')\sin\alpha+$$
$$+(\rho'\sin\varphi')\cos\alpha=x'\sin\alpha+y'\cos\alpha,$$

т. е.

$$\left.\begin{aligned} x &= x'\cos\alpha - y'\sin\alpha, \\ y &= x'\sin\alpha + y'\cos\alpha. \end{aligned}\right\} \tag{2.81}$$

Чтобы выразить $x'$, $y'$ через $x$, $y$, необходимо разрешить систему (2.81) относительно $x'$, $y'$. Можно сделать проще: считать систему $Ox'y'$ старой, тогда переход к новой системе $Oxy$ совершается поворотом на угол $(-\alpha)$, поэтому в формулах (2.81) достаточно поменять местами $x$ и $x'$, $y$ и $y'$, записать $(-\alpha)$ вместо $\alpha$.

В общем случае, когда даны две системы $Oxy$ и $O'x'y'$ (рис. 2.23), вводя промежуточную систему $O'x''y''$ и применяя последовательно формулы (2.79) и (2.81), получаем

$$\left.\begin{aligned} x &= x'\cos\alpha - y'\sin\alpha + a, \\ y &= x'\sin\alpha + y'\cos\alpha + b. \end{aligned}\right\} \tag{2.82}$$

Замечание. Система координат $Oxy$, в которой кратчайший поворот положительной полуоси $Ox$ до совпадения с положительной полуосью $Oy$ совершается против часовой стрелки, называется *правой*; если указанный поворот совершается по часовой стрелке, система называется *левой*. До сих пор рассматривались только правые системы. Формулы (2.82) остаются прежними, если обе системы координат являются левыми. Если одна система правая, другая — левая, то в формулах (2.82) изменится знак перед $y'$, так как в случае простейшего преобразования координат разноименных систем формулы имеют вид $x=x'$, $y=-y'$.

## § 2.9. Простейшие приложения формул преобразования координат к упрощению уравнений линий второго порядка

**Линии, определяемые уравнениями** $x=Ay^2+By+C$, $y=ax^2+bx+c$. Преобразуем уравнение $x=Ay^2+By+C$, выделив в правой части его полный квадрат:

$$x = A\left(y^2 + 2\frac{B}{2A}y + \frac{B^2}{4A^2}\right) - \frac{B^2}{4A} + C;$$

$$x + \frac{B^2}{4A} - C = A\left(y + \frac{B}{2A}\right)^2.$$

Переходя к новым координатам по формулам

$$X = x + \left(\frac{B^2}{4A} - C\right), \quad Y = y + \frac{B}{2A}$$

(т. е. осуществляя параллельный перенос, так как это формулы вида (2.80)), перепишем преобразованное уравнение в виде $X=AY^2$ или $Y^2=2pX$ $\left(\frac{1}{A}=2p\right)$. Поскольку последнее уравнение определяет параболу в системе $O_1XY$, уравнение $x=Ay^2+By+C$ также является уравнением параболы. Отметим, что ось этой параболы параллельна оси $Ox$.

Уравнение $y=ax^2+bx+c$ аналогичным образом приводится к уравнению $X^2=2qY$; следовательно, оно определяет параболу с осью, параллельной оси $Oy$.

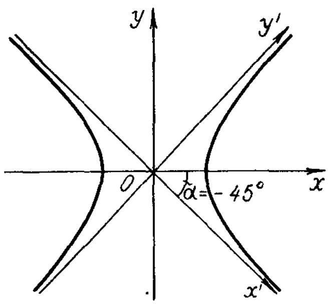


Рис. 2.24

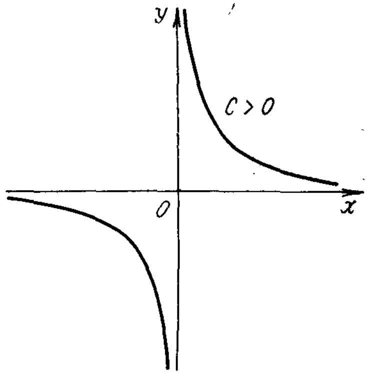


Рис. 2.25

**Уравнение равносторонней гиперболы относительно ее асимптот.** Асимптоты равносторонней гиперболы, определяемой уравнением (2.64), взаимно перпендикулярны (это, в частности, следует из простых геометрических соображений: основной прямоугольник является квадратом).

Возьмем эти асимптоты в качестве осей новой системы координат $Ox'y'$, фиксировав положительные направления так, как указано на рис. 2.24. Формулы преобразования (2.81) в данном случае ($\alpha=-45°$) примут вид $x=\frac{\sqrt{2}}{2}(x'+y')$, $y=\frac{\sqrt{2}}{2}(y'-x')$. Подставляя эти выражения в уравнение (2.64), получаем $\left[\frac{\sqrt{2}}{2}(x'+y')\right]^2-$ $-\left[\frac{\sqrt{2}}{2}(y'-x')\right]^2=a^2$, $\frac{1}{2}(x'^2+2x'y'+y'^2)-\frac{1}{2}(x'^2-$

$-2x'y'+y'^2)=a^2,\ 2x'y'=a^2,\ x'y'=\frac{a^2}{2},\ x'y'=C\Big(C=$
$=\frac{a^2}{2}\neq 0\Big).$

Если взять $\alpha=45°$, то $x'y'=-C\left(C=\frac{a^2}{2}\neq 0\right)$. Следовательно, уравнение

$$xy=C\quad (C\neq 0) \tag{2.83}$$

определяет равностороннюю гиперболу в системе $Oxy$, для которой координатными осями являются ее асимптоты (рис. 2.25).

Замечание. Уравнение $xy=C$ рассматривалось в курсе математики средней школы при изучении обратной пропорциональной зависимости $y=\frac{C}{x}$.

**Линия, определяемая уравнением** $y=\frac{ax+b}{cx+d}$. Предположим, что в уравнении

$$y=\frac{ax+b}{cx+d} \tag{2.84}$$

$c\neq 0$ (в случае $c=0$ уравнение принимает вид $y=kx+m$ и определяет прямую),

$$ad-bc\neq 0 \tag{2.85}$$

(в случае $ad-bc=0$ уравнение принимает вид $y=l$ и также определяет прямую).

Уравнение (2.84) преобразуем следующим образом:

$$y=\frac{ax+b}{cx+d}=\frac{a\left(x+\frac{b}{a}\right)}{c\left(x+\frac{d}{c}\right)}=\frac{a\left(x+\frac{d}{c}-\frac{d}{c}+\frac{b}{a}\right)}{c\left(x+\frac{d}{c}\right)}=$$

$$=\frac{a\left(x+\frac{d}{c}\right)+\frac{bc-ad}{c}}{c\left(x+\frac{d}{c}\right)}=\frac{a}{c}+\frac{\frac{bc-ad}{c^2}}{x+\frac{d}{c}};$$

$$y=\frac{a}{c}+\frac{C}{x+\frac{d}{c}}, \tag{2.86}$$

где $C=\frac{bc-ad}{c^2}$, причем $C\neq 0$ в силу условия (2.85).

Запишем уравнение (2.86) в виде

$$y-\frac{a}{c}=\frac{C}{x+\frac{d}{c}} \tag{2.87}$$

и перейдем к новым координатам по формулам

$$X=x+\frac{d}{c},\ Y=y-\frac{a}{c}. \tag{2.88}$$

В новой системе координат $O_1XY$ уравнение (2.87) принимает вид

$$Y=\frac{C}{X},\ \text{или}\ XY=C\quad (C\neq 0). \tag{2.89}$$

Уравнение (2.89) определяет равностороннюю гиперболу, поэтому уравнение (2.84) также определяет гиперболу.

Пример 1. Построить линию, определяемую уравнением $2y=x^2-2x+5$. Преобразуя это уравнение, получаем $y=\frac{1}{2}x^2-x+\frac{5}{2}$, $y=\frac{1}{2}(x^2-2x+1)-\frac{1}{2}+\frac{5}{2}$, $y=\frac{1}{2}(x-1)^2+2$, $y-2=\frac{1}{2}(x-1)^2$.

В новой системе координат $O_1XY$, где $X=x-1$, $Y=y-2$, уравнение принимает вид $Y=\frac{1}{2}X^2$ и определяет параболу. Строим системы координат $Oxy$, $O_1XY$ и параболу в новой системе по ее каноническому уравнению (рис. 2.26).

Пример 2. Построить линию, определяемую уравнением $y=\frac{4x+14}{x+2}$.

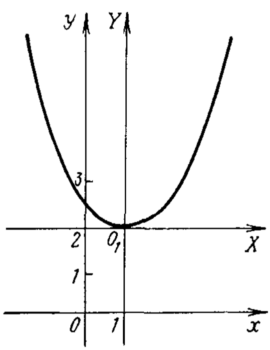



Рис. 2.26

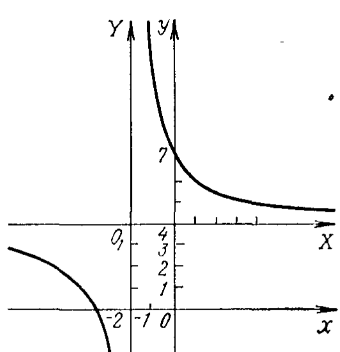



Рис. 2.27

Преобразуем уравнение: $y(x+2)-4x-14=0$, $y(x+2)-4x-8-6=0$, $y(x+2)-4(x+2)-6=0$, $(x+2)(y-4)=6$.

Переходя к новым координатам по формулам $X=x+2$, $Y=y-4$, получаем уравнение $XY=6$. Строим гиперболу в системе $O_1XY$ (рис. 2.27).

## § 2.10. Упрощение уравнения второй степени, не содержащего члена с произведением координат

Пусть дано уравнение

$$Ax^2+Cy^2+2Dx+2Ey+F=0 \tag{2.90}$$

относительно декартовых координат, для которого

$$A^2+C^2\neq 0. \tag{2.91}$$

Рассмотрим три случая: 1) коэффициенты $A$ и $C$ одного знака ($AC>0$, уравнение в этом случае называется уравнением эллиптического типа); 2) коэффициенты $A$ и $C$ разных знаков ($AC<0$, уравнение гиперболического типа); 3) один из коэффициентов равен нулю ($AC=0$, другой коэффициент отличен от нуля в силу условия (2.91); уравнение параболического типа).

1) Выделяя полные квадраты в левой части уравнения (2.90), получаем

$$A\left(x^2+2\frac{D}{A}x+\frac{D^2}{A^2}\right)+C\left(y^2+2\frac{E}{C}y+\frac{E^2}{C^2}\right)-\frac{D^2}{A}-\frac{E^2}{C}+F=0,$$

$$A\left(x+\frac{D}{A}\right)^2+C\left(y+\frac{E}{C}\right)^2=\frac{D^2}{A}+\frac{E^2}{C}-F. \tag{2.92}$$

Переходя к новым координатам по формулам

$$X=x+\frac{D}{A},\quad Y=y+\frac{E}{C} \tag{2.93}$$

и обозначая правую часть уравнения (2.92) через $K$, т. е. $K=\frac{D^2}{A}+\frac{E^2}{C}-F$, перепишем его в виде

$$AX^2+CY^2=K. \tag{2.94}$$

В зависимости от того, каково $K$ ($KA>0$, $KA<0$, $K=0$), уравнение (2.94) можно привести к одному из уравнений:

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=1, \tag{2.95}$$

где $\frac{1}{a^2}=\frac{A}{K}$, $\frac{1}{b^2}=\frac{C}{K}$;

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=-1, \tag{2.96}$$

$a$ и $b$ выражаются предыдущими формулами:

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=0, \tag{2.97}$$

$$\frac{1}{a^2}=A,\ \frac{1}{b^2}=C \quad (\text{при } A>0). \tag{2.98}$$

Уравнение (2.95) определяет эллипс с полуосями $a$ и $b$ (окружность в случае $a=b$), уравнение (2.97) — точку ($X=0$, $Y=0$ — в новой системе координат, ее старые координаты находятся с помощью формул (2.93)), уравнение (2.96) линии не определяет.

2) Поскольку $AC<0$, уравнение (2.94) приводится к одному из уравнений

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=1; \tag{2.99}$$

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=-1,\ -\frac{X^2}{a^2}+\frac{Y^2}{b^2}=1; \tag{2.100}$$

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=0. \tag{2.101}$$

Уравнение (2.99) определяет гиперболу с действительной осью $O_1X$, уравнение (2.100) — гиперболу с действительной осью $O_1Y$, уравнение (2.101) — пару пересекающихся прямых, так как оно распадается на два уравнения: $Y=-\frac{b}{a}X$, $Y=\frac{b}{a}X$.

3) Пусть $A=0$, $C\neq 0$. Предполагая, что $D\neq 0$, получаем

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

$$C\left(y+\frac{E}{C}\right)^2=-2D\left(x+\frac{F-\frac{E^2}{C}}{2D}\right),$$

$$Y^2=2pX, \tag{2.102}$$

где $X=x+\frac{FC-E^2}{2CD}$, $Y=y+\frac{E}{C}$, $p=-\frac{D}{C}$.

Если $D=0$, то $C\left(y+\frac{E}{C}\right)^2=\frac{E^2}{C}-F$ или

$$CY^2=L, \qquad (2.103)$$

где $L=\frac{E^2}{C}-F$.

Уравнение (2.103) сводится к одному из уравнений

$$Y^2=b^2, \qquad (2.104)$$

$$Y^2=-b^2, \qquad (2.105)$$

$$Y^2=0, \qquad (2.106)$$

в зависимости от того, каково $L$ ($LC>0$, $LC<0$, $L=0$).

Уравнение (2.102) определяет параболу с осью $O_1X$, уравнение (2.106) — пару совпавших прямых, уравнение (2.104) — пару параллельных прямых, уравнение (2.105) линии не определяет.

Замечание. Если $C=0$, $A\neq 0$, то уравнение (2.90) приводится к виду $X^2=2qY$, когда $E\neq 0$, и к одному из уравнений $X^2=a^2$, $X^2=-a^2$, $X^2=0$, когда $E=0$.

Пример. Построить линию, определяемую уравнением $9y^2-16x^2+18y+32x-151=0$.

Преобразуем это уравнение: $9(y^2+2y+1)-16(x^2-2x+1)-9+16-151=0$, $9(y+1)^2-16(x-1)^2=144$,

$$\frac{(y+1)^2}{16}-\frac{(x-1)^2}{9}=1, \quad -\frac{(x-1)^2}{9}+\frac{(y+1)^2}{16}=1.$$

Перейдем к новой системе координат $O_1XY$, полученной из старой путем параллельного переноса начала в точку $O_1(1, -1)$. Новые координаты через старые при этом преобразовании выражаются формулами $X=x-1$, $Y=y+1$, а само уравнение принимает вид $-\frac{X^2}{9}+\frac{Y^2}{16}=1$. Это уравнение определяет гиперболу с действительной осью $O_1Y$ и параметрами $a=3$, $b=4$. Строим старую и новую системы координат и гиперболу относительно системы $O_1XY$ (рис. 2.28). Тем самым построена линия, заданная исходным уравнением, в старой системе координат.

## § 2.11. Упрощение общего уравнения второй степени

Рассмотрим общее уравнение второй степени относительно декартовых прямоугольных координат $x$ и $y$, т. е. уравнение $Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0$. Будем считать, что в этом уравнении $B\neq 0$ (случай, когда $B=0$, изучен в § 2.10).

Перейдем к новым координатам $x'$, $y'$ по формулам (2.81). Подставляя эти формулы в данное уравнение, получаем

$$A_1x'^2+2B_1x'y'+C_1y'^2+2D_1x'+2E_1y'+F=0, \quad (2.107)$$

где

$$A_1=A\cos^2\alpha+2B\cos\alpha\sin\alpha+C\sin^2\alpha,$$

$$B_1=(C-A)\cos\alpha\sin\alpha+B(\cos^2\alpha-\sin^2\alpha),$$

$$C_1=A\sin^2\alpha-2B\sin\alpha\cos\alpha+C\cos^2\alpha.$$

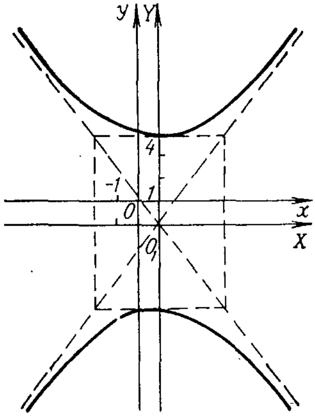



Рис. 2.28

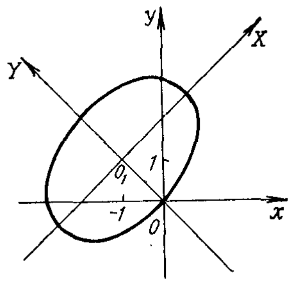



Рис. 2.29

Угол поворота $\alpha$ выберем так, чтобы в уравнении (2.107) обратился в нуль коэффициент $B_1$, т. е. $(C-A)\cos\alpha\sin\alpha+B(\cos^2\alpha-\sin^2\alpha)=0$.

Из этого уравнения находим

$$\operatorname{ctg} 2\alpha = \frac{A-C}{2B}. \quad (2.108)$$

Докажем, что коэффициенты $A_1$ и $C_1$ при этом одновременно в нуль обратиться не могут. Действительно, предполагая противное, т. е. $A_1=0$, $C_1=0$, получаем уравнения $A\cos^2\alpha+2B\cos\alpha\sin\alpha+C\sin^2\alpha=0$, $A\sin^2\alpha-2B\cos\alpha\sin\alpha+C\cos^2\alpha=0$, из которых следует, что $\operatorname{ctg} 2\alpha=-\frac{2B}{A-C}$. Сравнивая это равенство с равенством (2.108), заключаем, что $-\frac{2B}{A-C}=\frac{A-C}{2B}$, откуда $4B^2=$

$= -(A - C)^2$. Последнее возможно лишь в случае $B = 0$, что противоречит условию ($B \neq 0$).

Тем самым доказано, что при переходе к новой системе координат, полученной из старой путем поворота осей на угол $\alpha$, степень уравнения понизиться не может. Степень уравнения не может понизиться при параллельном переносе, а также при общем преобразовании координат (формулы (2.82)). Если данная линия в некоторой декартовой системе координат определяется уравнением второй степени, то в любой другой декартовой системе она также определяется уравнением второй степени.

Замечание. В формулы преобразования (2.81) входят $\sin\alpha$ и $\cos\alpha$, чтобы найти их, необходимо воспользоваться известиыми тригонометрическими формулами

$$\cos 2\alpha = \frac{\operatorname{ctg} 2\alpha}{\pm\sqrt{1+\operatorname{ctg}^2 2\alpha}}, \quad \sin\alpha = \pm\sqrt{\frac{1-\cos 2\alpha}{2}},$$

$$\cos\alpha = \pm\sqrt{\frac{1+\cos 2\alpha}{2}}.$$

Зиаки в этих формулах можно выбрать по своему усмотреиию. Посредством найденных зиачеиий $\sin\alpha$ н $\cos\alpha$ определяются коэффициенты уравнения (2.107).

Уравнение (2.107) принимает вид

$$A_1x'^2 + C_1y'^2 + 2D_1x' + 2E_1y' + F_1 = 0. \qquad (2.109)$$

Приведение уравнения (2.109) к каноническому виду осуществляется способами, указанными в § 2.10.

Выводы § 2.10 и 2.11 можно сформулировать следующей теоремой.

**Теорема 2.4.** Общее уравнение второй степени относительно декартовых прямоугольных координат определяет или пустое множество, или точку, или пару прямых (пересекающихся, параллельных, совпавших), или одну из линий: эллипс (окружность), гиперболу, параболу.

Пример. Определить вид, параметры и расположеиие линии, заданной уравнением $5x^2 - 6xy + 5y^2 + 16x - 16y = 0$.

По формуле (2.108) получаем $\operatorname{ctg} 2\alpha = \frac{A-C}{2B} = \frac{5-5}{-6} = 0$. Взяв $2\alpha = 90°$, формулы (2.81) запишем в виде

$$x = \frac{\sqrt{2}}{2}(x' - y'), \quad y = \frac{\sqrt{2}}{2}(x' + y'). \qquad (I)$$

Подставим эти выражения в даиное уравнение и преобразуем его:

$\frac{5}{2}(x'^2 - 2x'y' + y'^2) - 3(x'^2 - y'^2) + \frac{5}{2}(x'^2 + 2x'y' + y'^2) + 8\sqrt{2}(x' - - y') - 8\sqrt{2}(x' + y') = 0$, $2x'^2 + 8y'^2 - 16\sqrt{2}y' = 0$, $x'^2 + 4y'^2 - - 8\sqrt{2}y' = 0$, $x'^2 + 4(y'^2 - 2\sqrt{2}y' + 2) - 8 = 0$, $x'^2 + 4(y' - - \sqrt{2})^2 = 8$. Переходя к новым координатам по формулам

$$X = x', \ Y = y' - \sqrt{2}, \tag{II}$$

преобразуем последнее уравнение к виду

$$X^2 + 4Y^2 = 8, \quad \text{или} \quad \frac{X^2}{8} + \frac{Y^2}{2} = 1. \tag{III}$$

Уравнение (III) определяет эллипс с полуосями $a=2\sqrt{2}$, $b=\sqrt{2}$ (эти полуоси легко построить, $\sqrt{2}$ выражает длину диагонали квадрата, сторона которого равна единице). Центр эллипса находится в точке, для которой $X=0$, $Y=0$. С помощью формул преобразования (II) и (I) находим $x'=0$, $y'=\sqrt{2}$, $x=-1$, $y=1$; центр эллипса и начало новой системы $O_1XY$ находятся в точке $O_1(-1, 1)$. Строим эллипс, определяемый уравнением (III), относительно системы $O_1XY$ (рис. 2.29), тем самым линия построена относительно старой системы $Oxy$. Отметим, что эллипс проходит через начало старой системы координат (исходное уравнение не содержит свободного члена, ему удовлетворяют координаты $x=0$, $y=0$).

# Глава 3

# МАТРИЦЫ И ОПРЕДЕЛИТЕЛИ

В этой главе рассматриваются матрицы и действия над ними, изучаются свойства определителей. Матрицы и определители широко используются при решении систем линейных алгебраических уравнений.

## § 3.1. Матрицы. Основные определения

*Матрицей* называется система $m \cdot n$ чисел, расположенных в прямоугольной таблице из $m$ строк и $n$ столбцов. Числа этой таблицы называются элементами матрицы[1]. Обозначения матрицы:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

[1] Элементами матрицы могут быть и другие объекты.

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}.$$

Элементы $a_{i1}, a_{i2}, \dots, a_{in}$ составляют $i$-ю строку ($i=1, 2, \dots, m$), элементы $a_{1k}, a_{2k}, \dots, a_{mk}$ — $k$-й столбец ($k=1, 2, \dots, n$); $a_{ik}$ — элемент, принадлежащий $i$-й строке и $k$-му столбцу матрицы, числа $i$, $k$ называют *индексами элемента*. Строки и столбцы матрицы называют ее *рядами*; под двумя параллельными рядами понимают две строки или два столбца матрицы. Матрицу, имеющую $m$ строк и $n$ столбцов, называют матрицей размеров $m \times n$ (читается $m$ на $n$). Употребляются и более краткие обозначения матрицы размеров $m \times n$: $[a_{ik}]_{mn}$, $(a_{ik})_{mn}$, $\|a_{ik}\|_{mn}$. Матрицу обозначают также одной заглавной буквой, например,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}. \quad (3.1)$$

Если необходимо отметить, что матрица $A$ имеет $m$ строк и $n$ столбцов, т. е. необходимо указать ее размеры, то пишут $A_{m \times n}$, или $A_{mn}$.

Две матрицы $A_{mn} = (a_{ik})_{mn}$, $B_{pq} = (b_{ik})_{pq}$ называются равными, если $p=m$, $q=n$ и $a_{ik}=b_{ik}$ ($i=1, 2, \dots, m$; $k=1, 2, \dots, n$), другими словами, если они одинаковых размеров и их соответствующие элементы равны.

Матрица, состоящая лишь из одной строки, называется *строчной матрицей*, или *матрицей-строкой*. Строчная матрица имеет вид

$$[a_{11}, a_{12}, \dots, a_{1n}].$$

Матрица

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix},$$

имеющая лишь один столбец, называется *столбцевой матрицей*, или *матрицей-столбцом*.

Матрица, все элементы которой равны нулю, называется *нулевой* матрицей. Обозначим нулевую матрицу буквой $O$, тогда по определению

$$O=\begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 \end{bmatrix}. \tag{3.2}$$

*Квадратной* матрицей называется матрица, у которой число строк равно числу столбцов ($m=n$), т. е. матрица вида

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}. \tag{3.3}$$

*Порядком* квадратной матрицы называется число ее строк (или столбцов). Квадратная матрица первого порядка отождествляется со своим единственным элементом. Выпишем квадратные матрицы первых трех порядков:

$$[a_{11}],\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Будем говорить, что элементы $a_{11}, a_{22}, \ldots, a_{nn}$ квадратной матрицы (3.3) образуют ее *главную диагональ*, а элементы $a_{1n}, a_{2,n-1}, \ldots, a_{n1}$ — вторую диагональ.

Квадратная матрица называется симметрической, если $a_{ik}=a_{ki}$, т. е. равны ее элементы, симметричные относительно главной диагонали. Например, симметрическими являются матрицы

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix},\ \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix},\ \begin{bmatrix} 2 & 4 & 3 \\ 4 & 1 & 5 \\ 3 & 5 & 7 \end{bmatrix},\ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}.$$

*Диагональной матрицей* называется квадратная матрица, у которой все элементы, не принадлежащие главной диагонали, равны нулю, т. е. матрица

$$\operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})=\begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix}. \tag{3.4}$$

*Единичной матрицей* называется диагональная матрица, у которой все элементы главной диагонали равны единице. Обозначим единичную матрицу буквой $E$, тогда

$$E=\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}. \qquad (3.5)$$

Единичную матрицу кратко обозначают так: $E=[\delta_{ik}]$, где $\delta_{ik}$ — символ Кронекера [1], т. е.

$$\delta_{ik}=\begin{cases} 1, & \text{если } i=k, \\ 0, & \text{если } i\neq k. \end{cases}$$

Желая отметить, что единичная матрица имеет порядок $n$, пишут $E_n$.

*Треугольной матрицей* называется квадратная матрица, все элементы которой, расположенные по одну сторону от главной диагонали, равны нулю. Различают соответственно верхнюю и нижнюю треугольные матрицы:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}. \qquad (3.6)$$

Матрица произвольных размеров

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1r} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2r} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3r} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{rr} & \dots & a_{rn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix}, \qquad (3.7)$$

где $a_{ii}\neq 0$ $(i=1, 2, \dots, r)$, называется *квазитреугольной* (ступенчатой или трапециевидной).

## § 3.2. Действия над матрицами

**Линейные действия над матрицами.** *Линейными действиями* над матрицами называются сложение и вычитание матриц, умножение матрицы на число. Сложение и

[1] *Леопольд Кронекер* (*Leopold Kronecker*, 1823—1891) — немецкий математик.

вычитание матриц определяются только для матриц одинаковых размеров.

Суммой двух матриц $A=(a_{ik})_{mn}$, $B=(b_{ik})_{mn}$ называется такая матрица $C=(c_{ik})_{mn}$, что

$$c_{ik}=a_{ik}+b_{ik} \quad (i=1, 2, \ldots, m;\ k=1, 2, \ldots, n), \qquad (3.8)$$

т. е. матрица, элементы которой равны суммам соответствующих элементов матриц слагаемых. Сумма двух матриц $A$ и $B$ обозначается $A+B$.

Пример 1.

$$A=\begin{bmatrix} 3 & -2 & 4 \\ 5 & -1 & -6 \end{bmatrix},\quad B=\begin{bmatrix} 1 & 8 & -3 \\ 2 & -4 & 9 \end{bmatrix},\quad A+B=\begin{bmatrix} 4 & 6 & 1 \\ 7 & -5 & 3 \end{bmatrix}.$$

Под суммой $A+B+C$ трех матриц $A$, $B$, $C$ понимается матрица, полученная в результате последовательного сложения этих матриц, т. е.

$$A+B+C=(A+B)+C.$$

Аналогично определяется сумма матриц для большего числа слагаемых.

Разностью $A-B$ двух матриц $A=(a_{ik})_{mn}$, $B=(b_{ik})_{mn}$ называется матрица $D=(d_{ik})_{mn}$, для которой

$$d_{ik}=a_{ik}-b_{ik} \quad (i=1, 2, \ldots, m;\ k=1, 2, \ldots, n). \qquad (3.9)$$

Пример 2.

$$A=\begin{bmatrix} 3 & 4 & 2 \\ 9 & 8 & -5 \end{bmatrix},\quad B=\begin{bmatrix} 1 & 5 & -1 \\ 6 & 3 & -9 \end{bmatrix},\quad A-B=\begin{bmatrix} 2 & -1 & 3 \\ 3 & 5 & 4 \end{bmatrix}.$$

Произведением матрицы $A=(a_{ik})_{mn}$ на число $\alpha$ (или числа $\alpha$ на матрицу $A$) называется матрица $B=(b_{ik})_{mn}$, для которой

$$b_{ik}=\alpha a_{ik} \quad (i=1, 2, \ldots, m;\ k=1, 2, \ldots, n), \qquad (3.10)$$

т. е. матрица, полученная из данной умножением всех ее элементов на число $\alpha$. Произведение матрицы $A$ на число $\alpha$ обозначается $A\alpha$, или $\alpha A$.

Матрицу $(-1)A$ будем называть матрицей, *противоположной* матрице $A$, и обозначать $-A$.

Замечание. Разиость $A-B$ двух матриц можно определить так:

$$A-B=A+(-B). \qquad (3.11)$$

Линейные операции над матрицами обладают следующими свойствами: 1. $A+B=B+A$. 2. $(A+B)+C=A+(B+C)$. 3. $A+O=A$. 4. $A+(-A)=O$. 5. $1\cdot A=A$.

6. $\alpha(\beta A)=(\alpha\beta)A$. 7. $\alpha(A+B)=\alpha A+\alpha B$. 8. $(\alpha+\beta)A=\alpha A+\beta A$, где $A$, $B$, $C$ — матрицы одних и тех же размеров; $O$ — нулевая матрица; $(-A)$ — матрица, противоположная матрице $A$; $\alpha$, $\beta$ — любые действительные числа.

Докажем, например, первое из них. Так как $(A+B)=(a_{ik}+b_{ik})_{mn}$, $(B+A)=(b_{ik}+a_{ik})_{mn}$, $a_{ik}+b_{ik}=b_{ik}+a_{ik}$ ($i=1, 2, \ldots, m$; $k=1, 2, \ldots, n$), то $A+B=B+A$. Доказательства остальных свойств предоставляются читателю.

**Умножение матриц.** Это действие определяется для согласованных матриц. Матрица $A$ называется *согласованной* с матрицей $B$, если число столбцов матрицы $A$ равно числу строк матрицы $B$: матрица $A_{mn}$ согласована с матрицей $B_{nl}$ («ширина» матрицы $A$ равна «высоте» матрицы $B$). Отметим следующее: 1) из согласованности матрицы $A$ с матрицей $B$ не следует согласованность матрицы $B$ с матрицей $A$; 2) если $A$ и $B$ — квадратные матрицы одного порядка, то они взаимно согласованы ($A$ согласована с $B$, $B$ согласована с $A$). Поясним это понятие на примере. Даны три матрицы

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix},\quad B=\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix},\quad C=\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}.$$

Матрица $A$ согласована с матрицей $B$ ($A$ имеет 3 столбца, $B$ — 3 строки), матрица $B$ не согласована с матрицей $A$ ($B$ имеет 3 столбца, $A$ — 2 строки), но согласована с матрицей $C$, матрица $C$ согласована с матрицей $B$ и не согласована с матрицей $A$.

Произведением матрицы $A_{mn}=(a_{ik})_{mn}$ на матрицу $B_{nl}=(b_{ik})_{nl}$ называется такая матрица $C_{ml}=(c_{ik})_{ml}$, для которой

$$c_{ik}=a_{i1}b_{1k}+a_{i2}b_{2k}+\ldots+a_{in}b_{nk},\quad c_{ik}=\sum_{j=1}^{n} a_{ij}b_{jk}, \tag{3.12}$$

т. е. элемент $c_{ik}$ матрицы $C_{ml}$ равен сумме произведений элементов $i$-й строки матрицы $A_{mn}$ на соответствующие элементы $k$-го столбца матрицы $B_{nl}$. Матрица $C_{ml}$ имеет $m$ строк (как и матрица $A_{mn}$) и $l$ столбцов (как и матрица $B_{nl}$). Произведение матрицы $A$ на матрицу $B$ обозначается $AB$.

Замечание. Из того, что матрицу $A$ можно умножить на матрицу $B$, не следует, что матрицу $B$ можно умножать на $A$ (так как из согласованности $A$ с $B$ не следует согласованность $B$ с $A$).

Произведение $AB$ рассматривают как результат умножения матрицы $A$ на матрицу $B$ слева или умножения матрицы $B$ на матрицу $A$ справа.

Пример 4. Даны две матрицы второго порядка

$$A=\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix},\quad B=\begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix}.$$

Найти произведения $AB$ и $BA$.

Пользуясь формулой (3.12), для матриц $A=(a_{ik})_{22}$, $B=(b_{ik})_{22}$ иаходим общие выражеиия указаниых произведений:

$$AB=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\cdot\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}=\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix},$$

$$BA=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\cdot\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}=\begin{bmatrix} b_{11}a_{11}+b_{12}a_{21} & b_{11}a_{12}+b_{12}a_{22} \\ b_{21}a_{11}+b_{22}a_{21} & b_{21}a_{12}+b_{22}a_{22} \end{bmatrix}.$$

В случае даиных матриц $A$ и $B$ получаем

$$AB=\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix}=\begin{bmatrix} 3\cdot0+0\cdot4 & 3\cdot0+0\cdot0 \\ 0\cdot0+0\cdot4 & 0\cdot0+0\cdot0 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$BA=\begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix}\cdot\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}=\begin{bmatrix} 0\cdot3+0\cdot0 & 0\cdot0+0\cdot0 \\ 4\cdot3+0\cdot0 & 4\cdot0+0\cdot0 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 12 & 0 \end{bmatrix}.$$

Для этих произведений $AB \neq BA$. Отметим, что $AB$ — нулевая матрица. Этот пример показывает, что произведеиие матриц может быть нулевой матрицей, когда оба множителя не являются нулевыми матрицами.

Пример 5. Даны две матрицы

3×4 4×2

$$A=\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 4 & 5 & -2 \\ 7 & 1 & 0 & -3 \end{bmatrix},\quad B=\begin{bmatrix} 5 & 4 \\ 3 & 2 \\ 1 & -5 \\ 9 & -7 \end{bmatrix}.$$

Найти произведение $AB$. Можио ли получить произведение $BA$?

Число столбцов матрицы $A$ равио числу строк матрнцы $B$ («шириоа» матрицы $A$ равна «высоте» матрицы $B$), поэтому произведение $AB$ определено.

В соответствии с формулой (3.12), умножая «строку матрицы $A$ на столбец матрицы $B$», получаем

$$AB=\begin{bmatrix} 1\cdot5+2\cdot3+3\cdot1+(-1)\cdot9 & 1\cdot4+2\cdot2+3(-5)+(-1)(-7) \\ 0\cdot5+4\cdot3+5\cdot1+(-2)\cdot9 & 0\cdot4+4\cdot2+5(-5)+(-2)(-7) \\ 7\cdot5+1\cdot3+0\cdot1+(-3)\cdot9 & 7\cdot4+1\cdot2+0(-5)+(-3)(-7) \end{bmatrix}=$$

$$=\begin{bmatrix} 5 & 0 \\ -1 & -3 \\ 11 & 51 \end{bmatrix}.$$

Произведеиие $BA$ ие определено, так как число столбцов матрицы $B$ не равно числу строк матрицы $A$.

Если для матриц $A$ и $B$ определены произведения $AB$ и $BA$, то не всегда $AB=BA$. В случае, когда

$$AB=BA, \tag{3.13}$$

матрицы $A$ и $B$ называются *перестановочными,* или *коммутативными.*

Легко видеть, что

$$AE=EA=A; \tag{3.14}$$

$$AO=OA=O, \tag{3.15}$$

т. е. при умножении матриц единичная матрица $E$ играет роль единицы, нулевая матрица $O$ — роль нуля.

Умножение матриц обладает следующими свойствами. Если имеют смысл соответствующие действия, то 1. $(AB)C=A(BC)$. 2. $\alpha(AB)=(\alpha A)B=A(\alpha B)$. 3. $(A+B)C=AC+BC$. 4. $C(A+B)=CA+CB$. Первое свойство называется свойством сочетательности, или ассоциативности, третье — свойством распределительности (дистрибутивности) умножения справа относительно сложения матриц.

Докажем первое из этих свойств: $(AB)C=A(BC)$. Пусть даны три матрицы: $A=(a_{ik})_{mn}$, $B=(b_{ik})_{nq}$, $C=(c_{ik})_{ql}$; очевидно, матрица $A$ согласована с матрицей $B$, матрица $B$ согласована с матрицей $C$. Так как матрица $AB$ согласована с матрицей $C$ и матрица $A$ согласована с матрицей $BC$, то обе части равенства имеют смысл. Матрицы $(AB)C$ и $A(BC)$ имеют одинаковые размеры $m\times l$. Введем обозначения: $AB=D=(d_{i\nu})_{mq}$, $BC=G=(g_{\mu k})_{nl}$, $U=(AB)C=(u_{ik})_{ml}$, $V=A(BC)=(v_{ik})_{ml}$. Поскольку $d_{i\nu}=\sum\limits_{\mu=1}^{n}a_{i\mu}b_{\mu\nu}$, $g_{\mu k}=\sum\limits_{\nu=1}^{q}b_{\mu\nu}c_{\nu k}$, то $u_{ik}=\sum\limits_{\nu=1}^{q}d_{i\nu}c_{\nu k}=$ $=\sum\limits_{\nu=1}^{q}\left(\sum\limits_{\mu=1}^{n}a_{i\mu}b_{\mu\nu}\right)c_{\nu k}=\sum\limits_{\nu=1}^{q}\sum\limits_{\mu=1}^{n}a_{i\mu}b_{\mu\nu}c_{\nu k}$, $v_{ik}=\sum\limits_{\mu=1}^{n}a_{i\mu}g_{\mu k}=$ $=\sum\limits_{\mu=1}^{n}a_{i\mu}\left(\sum\limits_{\nu=1}^{q}b_{\mu\nu}c_{\nu k}\right)=\sum\limits_{\mu=1}^{n}\sum\limits_{\nu=1}^{q}a_{i\mu}b_{\mu\nu}c_{\nu k}=\sum\limits_{\nu=1}^{q}\sum\limits_{\mu=1}^{n}a_{i\mu}b_{\mu\nu}c_{\nu k}$. Следовательно, $u_{ik}=v_{ik}$ $(i=1, 2, \ldots, m;\ k=1, 2, \ldots, l)$, т. е. $U=V$, или $(AB)C=A(BC)$.

**Многочлены от матриц.** Целой положительной степенью $A^k$ $(k>1)$ квадратной матрицы $A$ называется произведение $k$ матриц, каждая из которых равна $A$, т. е. $A^k=$

$=\underbrace{A\cdot A\ldots A}_{k\text{ раз}}$. Матрица $A^k$ имеет тот же порядок, что и матрица $A$. *Нулевой степенью* квадратной матрицы $A$ $(A\neq O)$ называется единичная матрица того же порядка, что и $A$, т. е. $A^0=E$. Первой степенью $A^1$ матрицы $A$ называется сама матрица $A$, т. е. $A^1=A$. *Многочленом* (или *полиномом*) степени $k$ ($k$ — целое неотрицательное число) от квадратной матрицы $A$ называется выражение вида

$$a_kA^k+a_{k-1}A^{k-1}+\ldots+a_2A^2+a_1A+a_0A^0,$$

где $a_i$ $(i=0, 1, 2, \ldots, k)$ — любые числа, причем $a_k\neq 0$. Обозначим многочлен от матрицы $A$ через $P(A)$, тогда по определению $P(A)=a_kA^k+a_{k-1}A^{k-1}+\ldots+a_2A^2+a_1A+$ $+a_0A^0$, или

$$P(A)=a_kA^k+a_{k-1}A^{k-1}+\ldots+a_2A^2+a_1A+a_0E.$$

Из определения следует, что многочлен от матрицы можно получить, если в обычный многочлен $P(x)=$ $=a_kx^k+a_{k-1}x^{k-1}+\ldots+a_1x+a_0$ вместо $x$ подставить квадратную матрицу (и учесть, что $a_0=a_0x^0$).

Пусть дан многочлен $P(x)$. Если $P(A)$ является нулевой матрицей, т. е. $P(A)=O$, *то матрица $A$ называется корнем многочлена $P(x)$*, а многочлен $P(x)$ — *аннулирующим многочленом для матрицы $A$*.

Пример 6. Показать, что матрица $A=\begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix}$ является корнем многочлена $P(x)=x^2+2x-1$.

Подставив в данный многочлен вместо $x$ матрицу $A$, получим

$$P(A)=A^2+2A-E=\begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix}^2+2\begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix}-\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=$$

$$=\begin{bmatrix} -1 & 2 \\ -4 & 7 \end{bmatrix}+\begin{bmatrix} 2 & -2 \\ 4 & -6 \end{bmatrix}-\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}=O.$$

Так как $P(A)=A^2+2A-E=O$, то матрица $A$ — корень данного многочлена.

**Транспонирование матрицы.** Матрица, полученная из данной заменой каждой ее строки столбцом с тем же номером, называется матрицей, *транспонированной* относительно данной. Матрицу, транспонированную относительно матрицы $A$, обозначим через $A'$ или $A^T$. Таким образом, если

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \text{ то } A'=\begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}. \quad (3.16)$$

Отметим, что если $A$ — матрица размеров $m\times n$, то матрица $A'$ имеет размеры $n\times m$. Операция нахождения матрицы, транспонированной к данной, называется *транспонированием* матрицы. Справедливы следующие свойства: 1. $(A')'=A$. 2. $(\alpha A)'=\alpha A'$. 3. $(A+B)'=A'+B'$. 4. $(AB)'=B'A'$.

В справедливости первых трех свойств читатель легко может убедиться самостоятельно. Докажем свойство 4. Прежде всего, заметим, что левая часть равенства $(AB)'=B'A'$ имеет смысл только в том случае, когда матрица $A$ согласована с матрицей $B$, т. е. в случае $A=$ $=A_{mn}$ и $B=B_{nl}$. Так как $B'=B'_{ln}$, $A'=A'_{nm}$, то матрица $B'$ согласована с $A'$ и произведение $B'A'$ определено. Чтобы доказать равенство $(AB)'=B'A'$, необходимо убедиться в том, что матрицы $(AB)'$ и $B'A'$ имеют одинаковые размеры и их соответствующие элементы равны между собой. Произведение $AB$ — матрица размеров $m\times l$, поэтому транспонированная к ней матрица $(AB)'$ имеет размеры $l\times m$. Матрица $B'A'$ имеет также размеры $l\times m$, поскольку $B'=B'_{ln}$, $A'=A'_{nm}$. Элементы матриц $A$, $B$, $AB$ обозначим буквами $a$, $b$, $c$ с соответствующими индексами, а элементы матриц $A'$, $B'$, $(AB)'$ — теми же буквами $a$, $b$, $c$ со штрихом. Пусть $c'_{ik}$ — элемент матрицы $(AB)'$, стоящий в $i$-й строке и $k$-м столбце матрицы $AB$. Элемент $c'_{ik}$ равен элементу $c_{ki}$ матрицы $AB$, принадлежащему ее $k$-й строке и $i$-му столбцу, т. е. $c'_{ik}=c_{ki}$. В соответствии с правилом умножения матриц получаем $c'_{ik}=c_{ki}=$ $=\sum_{j=1}^{n} a_{kj}b_{ji}$, где $a_{kj}$, $b_{ji}$ — элементы матриц $A$ и $B$ соответственно. Поскольку $a_{kj}=a'_{jk}$, $b_{ji}=b'_{ij}$, то $c'_{ik}=$ $=\sum_{j=1}^{n} a'_{jk}b'_{ij}=\sum_{j=1}^{n} b'_{ij}a'_{jk}$. Последнее выражение, представляющее собой сумму произведений элементов $i$-й строки матрицы $B'$ на соответствующие элементы $k$-го столбца матрицы $A'$, является элементом матрицы $B'A'$, принадлежащим $i$-й строке и $k$-му столбцу. Следовательно, соответствующие элементы матриц $(AB)'$ и $B'A'$ равны между собой. Свойство 4 доказано.

## § 3.3. Определители второго и третьего порядков и их свойства

Определителем квадратной матрицы второго порядка

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

называется число, равное $a_{11}a_{22}-a_{12}a_{21}$ и обозначаемое символом

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \text{ т. е. } \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}. \qquad (3.17)$$

Числа $a_{11}$, $a_{12}$, $a_{21}$, $a_{22}$ называются *элементами определителя* матрицы второго порядка. Каждый элемент определителя обозначается буквой $a$ с двумя индексами: первый (1) обозначает номер строки, второй (2) — номер столбца, на пересечении которых находится соответствующий элемент (например, элемент $a_{21}$ принадлежит второй строке и первому столбцу определителя).

Определитель матрицы называют также *детерминантом*. Для определителя матрицы $A$ употребляются следующие обозначения: $|A|$, $\Delta$, $\det A$, $\det(a_{ik})$.

Определителем квадратной матрицы третьего порядка

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

называется число

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{21}a_{32}a_{13}-a_{13}a_{22}a_{31}-$$

$$-a_{12}a_{21}a_{33}-a_{23}a_{32}a_{11}. \qquad (3.18)$$

Заметим, что каждое слагаемое алгебраической суммы в правой части формулы (3.18) представляет собой произведение элементов определителя, взятых по одному и только по одному из каждой строки и каждого столбца. Этому произведению приписывается соответствующий знак. Чтобы запомнить, какие произведения следует брать со знаком плюс, какие со знаком минус, полезно правило, схематически изображенное на рис. 3.1.

*Минором* какого-либо элемента определителя называется определитель, полученный из данного вычеркиванием той строки и того столбца, которым принадлежит данный элемент. Минор элемента $a_{ik}$ обозначим $M_{ik}$. Например, минором элемента $a_{21}$ определителя (3.18) является определитель второго порядка

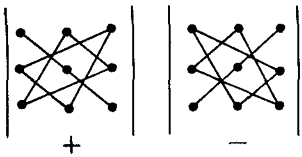


Рис. 3.1

$$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \qquad (3.19)$$

минором элемента $a_{21}$ определителя (3.17) — элемент $a_{12}$ (определитель первого порядка).

*Алгебраическим дополнением* элемента $a_{ik}$ определителя называется его минор, взятый со знаком $(-1)^{i+k}$. Например, алгебраическим дополнением элемента $a_{21}$ определителя (3.18) является определитель (3.19), взятый со знаком минус. Алгебраическое дополнение элемента $a_{ik}$ будем обозначать через $A_{ik}$. В соответствии с определением

$$A_{ik} = (-1)^{i+k} M_{ik}. \qquad (3.20)$$

Определители матриц второго и третьего порядка будем короче называть определителями второго и третьего порядка. Свойства определителей второго и третьего порядка выражаются следующими теоремами.

**Теорема 3.1.** 1) Определитель не изменится при замене всех его строк соответствующими столбцами; 2) при перестановке двух столбцов (строк) определитель меняет знак; 3) определитель с двумя одинаковыми столбцами (строками) равен нулю; 4) множитель, общий для элементов некоторого столбца (строки), можно выносить за знак определителя; 5) определитель равен нулю, если все элементы некоторого столбца (строки) равны нулю.

Доказательство проведем для определителя третьего порядка.

1. В определителе (3.18) каждую строку заменим столбцом с тем же номером, получим новый определитель

$$\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{12}a_{23}a_{31} - a_{31}a_{22}a_{13} - \\ - a_{21}a_{12}a_{33} - a_{32}a_{23}a_{11}.$$

Сравнивая это равенство с равенством (3.18), заключаем, что определители равны, так как равны правые части указанных равенств.

2. В определителе (3.18) переставим, например, второй и третий столбцы, тогда

$$\begin{vmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{vmatrix} = \left\{ \begin{array}{c} a_{11}a_{23}a_{32} + a_{13}a_{22}a_{31} + a_{21}a_{33}a_{12} - a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32} - \\ - a_{22}a_{33}a_{11} = -(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}). \end{array} \right.$$

Алгебраическая сумма в скобке равна правой части формулы (3.18), новый определитель отличается от исходного только знаком.

Другие случаи рассматриваются аналогично.

3. Определитель (3.18) обозначим через $\Delta$. Пусть он содержит два одинаковых столбца. Переставив эти столбцы, получим тот же определитель $\Delta$. По свойству 2 определитель при этом изменит знак, т. е. $\Delta = -\Delta$, откуда $\Delta = 0$.

4. Пусть в определителе (3.18) элементы второго столбца имеют общий множитель $\lambda$, тогда

$$\begin{vmatrix} a_{11} & \lambda a_{12} & a_{13} \\ a_{21} & \lambda a_{22} & a_{23} \\ a_{31} & \lambda a_{32} & a_{33} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

так как $a_{11}\lambda a_{22}a_{33} + \lambda a_{12}a_{23}a_{31} + a_{21}\lambda a_{32}a_{13} - a_{13}\lambda a_{22}a_{31} - \lambda a_{12}a_{21}a_{33} - \lambda a_{32}a_{23}a_{11} = \lambda(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{32}a_{23}a_{11})$.

Другие случаи рассматриваются аналогично.

Следствие. Определитель с двумя пропорциональными столбцами (строками) равен нулю.

Действительно, выделяя общий множитель элементов одного из этих столбцов (коэффициент пропорциональности) и вынося его за знак определителя, получаем определитель с двумя одинаковыми столбцами, равный нулю.

5. Если все элементы некоторого столбца (строки) равны нулю, то каждое слагаемое алгебраической суммы в правой части (3.18) равно нулю, как произведение, содержащее нулевой множитель.

Все свойства определителя второго порядка доказываются аналогично, при этом используется формула (3.17).

**Теорема 3.2.** Определитель не изменится, если к элементам некоторого столбца (строки) прибавить соответственные элементы другого столбца (строки), предварительно умножив их на один и тот же множитель.

Пусть, например, к элементам третьего столбца определителя (3.18) прибавлены соответственные элементы второго столбца, умноженные на множитель $\lambda$, тогда

$$\begin{vmatrix} a_{11} & a_{12} & a_{13}+\lambda a_{12} \\ a_{21} & a_{22} & a_{23}+\lambda a_{22} \\ a_{31} & a_{32} & a_{33}+\lambda a_{32} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \lambda a_{12} \\ a_{21} & a_{22} & \lambda a_{22} \\ a_{31} & a_{32} & \lambda a_{32} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \lambda \begin{vmatrix} a_{11} & a_{12} & a_{12} \\ a_{21} & a_{22} & a_{22} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{33} & a_{33} \end{vmatrix},$$

поскольку $a_{11}a_{22}(a_{33}+\lambda a_{32})+a_{12}(a_{23}+\lambda a_{22})a_{31}+a_{21}a_{32}(a_{13}+$ $+\lambda a_{12}) - (a_{13}+\lambda a_{12})a_{22}a_{31} - a_{12}a_{21}(a_{33}+\lambda a_{32}) - (a_{23}+$ $+\lambda a_{22})a_{32}a_{11} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{21}a_{32}a_{13}-a_{13}a_{22}a_{31}-$ $-a_{12}a_{21}a_{33}-a_{23}a_{32}a_{11}+\lambda(a_{11}a_{22}a_{32}+a_{12}a_{22}a_{31}+a_{21}a_{32}a_{12}-$ $-a_{12}a_{22}a_{31}-a_{12}a_{21}a_{32}-a_{22}a_{32}a_{11})$.

Замечание. Одновременно доказано, что если в определителе все элементы некоторого столбца (строки) равны суммам двух слагаемых, то такой определитель равен сумме двух соответствующих определителей.

**Теорема 3.3.** Определитель равен сумме произведений элементов любой строки (столбца) на их алгебраические дополнения.

Доказательство. Для определителя второго порядка теорема очевидна, утверждение ее следует из формулы (3.17).

Преобразуем правую часть формулы (3.18). Так как

$$a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{21}a_{32}a_{13}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{32}a_{23}a_{11} =$$
$$= a_{11}(a_{22}a_{33}-a_{23}a_{32})+a_{12}(a_{23}a_{31}-a_{21}a_{33})+a_{13}(a_{21}a_{32}-a_{22}a_{31}) =$$
$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}\begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

и

$$A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad A_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad A_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}, \qquad (3.21)$$

то

$$\Delta = a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}, \qquad (3.22)$$

где через $\Delta$ обозначен определитель (3.18); $A_{11}$, $A_{12}$, $A_{13}$ — алгебраические дополнения элементов $a_{11}$, $a_{12}$, $a_{13}$.

Формула (3.22) называется разложением определителя по элементам первой строки. Аналогично получается разложение по элементам других строк и столбцов.

**Теорема 3.4.** Пусть $\Delta$ — некоторый определитель третьего порядка. Сумма произведений алгебраических дополнений элементов какой-нибудь строки (столбца) на любые числа $q_1$, $q_2$, $q_3$ равна определителю $\Delta'$, который получается из данного $\Delta$ заменой упомянутой строки (столбца) строкой (столбцом) из чисел $q_1$, $q_2$, $q_3$.

Доказательство. Рассмотрим определитель

$$\Delta' = \begin{vmatrix} q_1 & q_2 & q_3 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

полученный из определителя $\Delta$ (см. (3.18)) заменой первой строки строкой из чисел $q_1$, $q_2$, $q_3$. На основании теоремы (3.3) $\Delta' = q_1Q_1 + q_2Q_2 + q_3Q_3$, где $Q_1$, $Q_2$, $Q_3$ — алгебраические дополнения элементов $q_1$, $q_2$, $q_3$. Так как $Q_1 = A_{11}$, $Q_2 = A_{12}$, $Q_3 = A_{13}$, где $A_{11}$, $A_{12}$, $A_{13}$ определяются формулами (3.21), то $\Delta' = q_1A_{11} + q_2A_{12} + q_3A_{13}$, что и требовалось доказать.

**Теорема 3.5.** Сумма произведений элементов какой-нибудь строки (столбца) на алгебраические дополнения соответствующих элементов другой строки (столбца) равна нулю.

Доказательство. Для определителя второго порядка теорема очевидна (получаем определитель с двумя одинаковыми строками).

Пусть дан определитель $\Delta$ (3.18). Покажем, например, что $a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0$. По теореме 3.4

$$a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{23} \end{vmatrix}.$$

Поскольку этот определитель равен нулю (как содержащий две одинаковые строки), то $a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0$. Остальные случаи рассматриваются аналогично.

Пример. Вычислить определитель третьего порядка

$$\Delta = \begin{vmatrix} 1 & -2 & 3 \\ 4 & -1 & 5 \\ 6 & -8 & 7 \end{vmatrix}$$

тремя способами (по определению, по теореме разложения, преобразованием его с помощью свойств).

1. $\Delta=-7-60-96+18+56+40=-49$.

2. $\Delta = 1\cdot\begin{vmatrix} -1 & 5 \\ -8 & 7 \end{vmatrix} + 2\cdot\begin{vmatrix} 4 & 5 \\ 6 & 7 \end{vmatrix} + 3\cdot\begin{vmatrix} 4 & -1 \\ 6 & -8 \end{vmatrix} = 1\cdot 33 + 2(-2) + 3(-26) = -49$.

3. Умножая первую строку на $(-4)$ и прибавляя ко второй, затем умножая первую строку на $(-6)$ и прибавляя к третьей, получаем

$$\Delta = \begin{vmatrix} 1 & -2 & 3 \\ 0 & 7 & -7 \\ 0 & 4 & -11 \end{vmatrix}.$$

Разлагая этот определитель по элементам первого столбца, находим $\Delta = 1\cdot\begin{vmatrix} 7 & -7 \\ 4 & -11 \end{vmatrix} = (-77+28) = -49$.

## § 3.4. Определители $n$-го порядка

Перейдем к выяснению понятия определителя любого порядка $n$, где $n \geqslant 3$.

Рассмотрим квадратную матрицу $n$-го порядка

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}. \qquad (3.23)$$

Понятие определителя этой матрицы или определителя порядка $n$ введем индуктивно, считая, что уже введено понятие определителя порядка $n-1$, соответствующего квадратной матрице $(n-1)$-го порядка.

*Минором* элемента $a_{ik}$ матрицы $n$-го порядка (3.23) называют определитель порядка $n-1$, соответствующий той матрице, которая получается из матрицы (3.23) в результате вычеркивания $i$-й строки и $k$-го столбца. Минор элемента $a_{ik}$ обозначим через $M_{ik}$. Алгебраическим дополнением элемента $a_{ik}$ назовем его минор, взятый со знаком $(-1)^{i+k}$, и обозначим через $A_{ik}$, т. е. $A_{ik} = (-1)^{i+k}M_{ik}$.

*Определителем порядка $n$*, соответствующим матрице (3.23), называют число, равное $\sum_{k=1}^{n}(-1)^{1+k}a_{1k}M_{1k}$ и обозначаемое одним из символов

$$\Delta,\ \det A,\ \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Итак, по определению

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{k=1}^{n} (-1)^{1+k} a_{1k} M_{1k},$$

или

$$\det A = \sum_{k=1}^{n} a_{1k} A_{1k}. \tag{3.24}$$

Последняя формула выражает правило составления определителя порядка $n$ по элементам первой строки соответствующей ему матрицы и по алгебраическим дополнениям этих элементов, являющимся определителями порядка $n-1$, взятыми с надлежащими знаками. Из формулы (3.24) при $n=2$ получаем формулу (3.17), при $n=3$ — формулу (3.22).

Естественно возникает вопрос, нельзя ли использовать для получения величины определителя элементы и соответствующие им миноры не первой, а любой строки матрицы, а также вопрос о разложении определителя по элементам любого столбца. Ответ на эти вопросы дают две следующие основные теоремы, приводимые здесь без доказательств.

**Теорема 3.6.** Каков бы ни был номер строки $i$ ($i=1, 2, \dots, n$), для определителя $n$-го порядка (3.24) справедлива формула

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik},$$

или

$$\det A = \sum_{k=1}^{n} a_{ik} A_{ik} \, (i = 1, 2, \dots, n), \tag{3.25}$$

называемая разложением этого определителя по $i$-й строке.

**Теорема 3.7.** Каков бы ни был номер столбца $k$ ($k=1, 2, \dots, n$), для определителя $n$-го порядка (3.23) справедлива формула

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} M_{ik},$$

или

$$\det A=\sum_{i=1}^{n}a_{ik}A_{ik}\,(k=1,\ 2,\ \ldots,\ n), \tag{3.26}$$

называемая разложением этого определителя по $k$-му столбцу.

Пример. Вычислить определитель матрицы

$$A=\begin{bmatrix}1&-1&2&2\\1&-2&3&2\\4&-1&5&15\\6&-8&7&9\end{bmatrix},\ \text{т. е.}\quad \det A=\begin{vmatrix}1&-1&2&2\\1&-2&3&2\\4&-1&5&15\\6&-8&7&9\end{vmatrix}.$$

Применяя формулу (3.24), получаем

$$\det A=\begin{vmatrix}-2&3&2\\-1&5&15\\-8&7&9\end{vmatrix}+\begin{vmatrix}1&3&2\\4&5&15\\6&7&9\end{vmatrix}+2\cdot\begin{vmatrix}1&-2&2\\4&-1&15\\6&-8&9\end{vmatrix}-2\cdot\begin{vmatrix}1&-2&3\\4&-1&5\\6&-8&7\end{vmatrix}.$$

Вычисляя определители третьего порядка, находим $\det A=-147+98+2(-49)-2(-49)=-49$.

Замечание. Этот определитель можно вычислить путем его преобразований на основании свойств:

$$\begin{vmatrix}1&-1&2&2\\1&-2&3&2\\4&-1&5&15\\6&-8&7&9\end{vmatrix}=\begin{vmatrix}1&0&0&0\\1&-1&1&0\\4&3&-3&7\\6&-2&-5&-3\end{vmatrix}=\begin{vmatrix}-1&1&0\\3&-3&7\\-2&-5&-3\end{vmatrix}=$$

$$=\begin{vmatrix}-1&0&0\\3&0&7\\-2&-7&-3\end{vmatrix}=(-1)\ 49=-49.$$

(Второй определитель получен из первого путем прибавления первого столбца ко второму, умножения на (—2) первого столбца и прибавления к третьему, а затем к четвертому столбцу.)

Отметим, что для определителей $n$-го порядка ($n>3$) также верны теоремы 3.1—3.5. Сформулируем теорему 3.4 (теорему замещения) и теорему 3.5 (теорему аннулирования) для определителей $n$-го порядка.

**Теорема 3.8 (теорема замещения).** Сумма произведений произвольных чисел $b_1, b_2, \ldots, b_n$ соответственно на алгебраические дополнения элементов некоторого столбца (строки) матрицы порядка $n$ равна определителю матрицы, которая получается из данной заменой элементов этого столбца (строки) числами $b_1, b_2, \ldots, b_n$.

Например, если $A_{1k}, A_{2k}, \ldots, A_{nk}$ — алгебраические дополнения элементов $k$-го столбца определителя

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2k} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nk} & \ldots & a_{nn} \end{vmatrix}, \text{ то } \begin{vmatrix} a_{11} & a_{12} & \ldots & b_1 & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & b_2 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & b_n & \ldots & a_{nn} \end{vmatrix} =$$

$$= b_1A_{1k} + b_2A_{2k} + \ldots + b_nA_{nk}. \qquad (3.27)$$

**Теорема 3.9 (теорема аннулирования).** Сумма произведений элементов одного из столбцов (строк) матрицы на соответствующие алгебраические дополнения элементов другого столбца (строки) равна нулю.

Например, если дана квадратная матрица $A=(a_{ik})$ порядка $n$, $a_{1k}, a_{2k}, \ldots, a_{nk}$ — элементы $k$-го столбца, $A_{1l}, A_{2l}, \ldots, A_{nl}$ — алгебраические дополнения элементов $l$-го столбца $(l \neq k)$, то

$$a_{1k}A_{1l} + a_{2k}A_{2l} + \ldots + a_{nk}A_{nl} = 0, \qquad (3.28)$$

где $k \neq l$, $1 \leqslant k \leqslant n$, $1 \leqslant l \leqslant n$.

Приведем без доказательства теорему об определителе произведения матриц.

**Теорема 3.10.** Определитель произведения двух квадратных матриц одного порядка равен произведению определителей перемножаемых матриц.

Если $A=(a_{ik})_{nn}$, $B=(b_{ik})_{nn}$, то

$$\det(AB) = \det A \cdot \det B. \qquad (3.29)$$

## § 3.5. Обратная матрица

Матрицей, обратной квадратной матрице $A$, называется квадратная матрица $B$, удовлетворяющая равенствам

$$AB = BA = E, \qquad (3.30)$$

где $E$ — единичная матрица. Из определения следует, что обратная матрица может существовать только для квадратной матрицы и обе матрицы имеют один и тот же порядок. Матрицу, обратную матрице $A$, будем обозначать через $A^{-1}$.

Рассмотрим квадратную матрицу $n$-го порядка

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}. \tag{3.31}$$

Матрицей, присоединенной к матрице $A$, называется матрица

$$C=\begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}, \tag{3.32}$$

где $A_{ik}$ — алгебраическое дополнение элемента $a_{ik}$ матрицы $A$. Отметим, что алгебраические дополнения элементов $i$-й строки матрицы $A$ расположены в $i$-м столбце матрицы $C$.

**Теорема 3.11.** Если $A$ — квадратная матрица порядка $n$, а $C$ — присоединенная к ней матрица, то

$$AC=CA=E\det A, \tag{3.33}$$

где $E$ — единичная матрица $n$-го порядка.

Обозначим произведение $AC$ матриц (3.31) и (3.32) буквой $D$, т. е.

$$D=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}\cdot\begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}.$$

Согласно определению произведения матриц элемент $d_{ik}$ матрицы $D$ равен сумме произведений элементов $i$-й строки матрицы $A$ на соответствующие элементы $k$-го столбца матрицы $C$, присоединенной к матрице $A$. Для элементов $d_{ii}$ $(i=1, 2, \ldots, n)$, принадлежащих главной диагонали, эта сумма является определителем матрицы $A$ (по теореме о разложении определителя по элементам любой строки), для всех остальных элементов $d_{ik}$ $(i\neq k)$ равна нулю (по теореме аннулирования). Следовательно,

$$D=AC=\begin{bmatrix} \det A & 0 & \ldots & 0 \\ 0 & \det A & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \det A \end{bmatrix}=\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{bmatrix}\det A=E\det A.$$

Аналогично доказывается, что $CA=E\cdot\det A$.

Квадратная матрица называется *невырожденной*, или *неособенной*, если определитель ее отличен от нуля. Если определитель матрицы равен нулю, матрица называется *вырожденной*, или *особенной*.

**Теорема 3.12.** Для невырожденной матрицы $A$ существует единственная обратная матрица $A^{-1}$, определяемая формулой

$$A^{-1} = \frac{1}{\det A} C, \tag{3.34}$$

где $C$ — матрица, присоединенная к матрице $A$.

Так как $A$ — невырожденная матрица, то $\det A \neq 0$. Из равенств (3.33) находим

$$A\left(\frac{1}{\det A} C\right) = \left(\frac{1}{\det A} C\right) A = E.$$

Это означает, что выполнены равенства (3.30) и матрица (3.34) является обратной к матрице $A$.

Матрица $A^{-1}$ — единственная матрица, удовлетворяющая условиям (3.30). Действительно, если матрица $A_1^{-1}$ такова, что $AA_1^{-1} = A_1^{-1}A = E$, то $A_1^{-1}AA^{-1} = A_1^{-1}(AA^{-1}) = A_1^{-1}E = A_1^{-1}$, $A_1^{-1}AA^{-1} = (A_1^{-1}A)A^{-1} = EA^{-1} = A^{-1}$, откуда $A_1^{-1} = A^{-1}$.

Для невырожденных матриц выполняются следующие свойства: 1. $\det A^{-1} = \frac{1}{\det A}$. 2. $(A^{-1})^{-1} = A$. 3. $(A^k)^{-1} = (A^{-1})^k$. 4. $(AB)^{-1} = B^{-1}A^{-1}$.

Замечание. Вырожденная матрица не имеет обратной.

Пример. Найти матрицу $A^{-1}$, обратную матрице

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -3 & 1 & 1 \\ 2 & -1 & 0 \end{bmatrix}.$$

Вычислим определитель матрицы $A$ и алгебраические дополнения ее элементов:

$$\det A = \begin{vmatrix} 2 & 0 & -1 \\ -3 & 1 & 1 \\ 2 & -1 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 2 & -1 & 0 \end{vmatrix} =$$

$$= (-1)\begin{vmatrix} -1 & 1 \\ 2 & -1 \end{vmatrix} = 1, \quad A_{11} = \begin{vmatrix} 1 & 1 \\ -1 & 0 \end{vmatrix} = 1,$$

$$A_{12}=-\begin{vmatrix}-3 & 1\\ 2 & 0\end{vmatrix}=2,\quad A_{13}=\begin{vmatrix}-3 & 1\\ 2 & -1\end{vmatrix}=1,\quad A_{21}=-\begin{vmatrix}0 & -1\\ -1 & 0\end{vmatrix}=1,$$

$$A_{22}=2,\ A_{23}=2,\ A_{31}=1,\ A_{32}=1,\ A_{33}=2.$$

В соответствии с формулой (3.34) получаем

$$A^{-1}=\frac{1}{1}\begin{bmatrix}1 & 1 & 1\\ 2 & 2 & 1\\ 1 & 2 & 2\end{bmatrix},\quad A^{-1}=\begin{bmatrix}1 & 1 & 1\\ 2 & 2 & 1\\ 1 & 2 & 2\end{bmatrix}.$$

## § 3.6. Ранг матрицы

Рассмотрим матрицу размеров $m\times n$

$$A=\begin{bmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{bmatrix}.$$

Выберем в ней произвольно $s$ различных строк и $s$ различных столбцов, причем $1 < s \leqslant \min(m, n)$, где $\min(m, n)$ — меньшее из чисел $m$ и $n$. Элементы, стоящие на пересечении выбранных строк и столбцов, образуют матрицу порядка $s$. Определитель этой матрицы называется минором порядка $s$ матрицы $A$. Например, если дана матрица

$$\begin{bmatrix}1 & -2 & 3 & 4 & 5\\ 0 & -4 & 1 & 2 & 7\\ 1 & -7 & 3 & 6 & 8\end{bmatrix},$$

то, взяв первую и вторую строку, третий и четвертый столбец, получим матрицу второго порядка и ее определитель

$$\begin{bmatrix}3 & 4\\ 1 & 2\end{bmatrix},\quad \begin{vmatrix}3 & 4\\ 1 & 2\end{vmatrix}=2.$$

Этот определитель является минором второго порядка для исходной матрицы. Аналогично можно получить и другие миноры второго порядка, а также миноры третьего порядка. Некоторые из миноров могут оказаться равными нулю.

*Рангом* матрицы называется наибольший из порядков ее миноров, отличных от нуля. Ранг матрицы будем обозначать буквой $r$. Если все миноры матрицы равны иулю, ранг ее считается равным нулю.

Из определения ранга матрицы получаем следующие утверждения.

1. Ранг матрицы выражается целым числом, заключенным между 0 и меньшим из чисел $m$, $n$, т. е. $0 \leqslant r \leqslant \min(m, n)$.

2. Ранг матрицы равен нулю тогда и только тогда, когда матрица является нулевой.

3. Для квадратной матрицы $n$-го порядка $r=n$ тогда и только тогда, когда матрица — невырожденная.

При нахождении ранга матрицы можно пользоваться свойствами миноров. Если все миноры порядка $k$ данной матрицы равны нулю, то все миноры более высокого порядка также равны нулю. Это следует, например, из теоремы о разложении определителя по элементам любой строки (столбца). Таким образом, если среди миноров порядка $k$ данной матрицы есть отличные от нуля, а все миноры порядка $k+1$ равны нулю или не существуют, то $r=k$.

Итак, ранг матрицы может быть найден следующим образом. Если все миноры первого порядка (элементы матрицы) равны нулю, то $r=0$. Если хотя бы один из миноров первого порядка отличен от нуля, а все миноры второго порядка равны нулю, то $r=1$. В случае, когда имеется минор второго порядка, отличный от нуля, исследуют миноры третьего порядка. Так поступают до тех пор, пока не обнаружится одно из двух: либо все миноры порядка $k$ равны нулю, либо миноры порядка $k$ не существуют, тогда $r=k-1$.

Пример 1. Найти ранг матрицы

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 3 & 0 & -6 & 0 \\ 4 & 0 & -8 & 0 \end{bmatrix}.$$

Среди миноров первого порядка этой матрицы (ее элементов) есть отличный от нуля, поэтому $r>0$. Из элементов данной матрицы можно составить миноры второго и третьего порядка, но все они равны нулю. Следовательно, $r=1$.

Указанный выше способ нахождения ранга матрицы не всегда бывает удобным, так как во многих случаях он связан с вычислением большого количества определителей.

Отметим некоторые очевидные свойства ранга матрицы.

1. Ранг матрицы, полученной из данной транспонированием, равен рангу исходной матрицы.

2. Ранг матрицы не изменится, если вычеркнуть или приписать нулевой ряд, т. е. ряд, все элементы которого равны нулю.

Элементарными преобразованиями матрицы называются следующие преобразования: 1) умножение некоторого ряда матрицы на отличное от нуля число; 2) прибавление к одному ряду матрицы другого, параллельного ему ряда, умноженного на любое число; 3) перестановка местами двух параллельных рядов матрицы. Можно доказать, что при элементарных преобразованиях ранг матрицы не меняется. Это следует, например, из определения ранга матрицы и свойств определителей. С помощью элементарных преобразований матрицу можно привести к квазитреугольной форме (см. (3.7)). Ранг квазитреугольной матрицы (3.7) равен $r$, поскольку ее минор с главной диагональю $a_{11}$, $a_{22}$, ..., $a_{rr}$ равен произведению $a_{11}a_{22}\ldots a_{rr}\neq 0$, а все миноры более высокого порядка равны нулю (как содержащие нулевые строки).

Пример 2. Найти ранг матрицы

$$\begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 4 & -2 & 5 & 0 & 3 \\ 2 & -3 & 0 & 6 & 1 \\ 7 & -6 & 5 & 9 & 6 \end{bmatrix}.$$

Применяя элементарные преобразования, получаем

$$\begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 4 & -2 & 5 & 0 & 3 \\ 2 & -3 & 0 & 6 & 1 \\ 7 & -6 & 5 & 9 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 0 & 2 & 5 & -12 & -5 \\ 0 & -1 & 0 & 0 & -3 \\ 0 & 1 & 5 & -12 & -8 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 0 & 1 & 5 & -12 & -8 \\ 0 & -1 & 0 & 0 & -3 \\ 0 & 2 & 5 & -12 & -5 \end{bmatrix} \to$$

$$\to \begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 0 & 1 & 5 & -12 & -8 \\ 0 & 0 & 5 & -12 & -11 \\ 0 & 0 & -5 & 12 & 11 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 0 & 1 & 5 & -12 & -8 \\ 0 & 0 & 5 & -12 & -11 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(Вторая матрица получена из первой путем поочередного умножения первой строки на (—4), (—2), (—7) и соответственного прибавления ко второй, третьей и четвертой строке; поменяв местами вторую и четвертую строки во второй матрице, получили третью матрицу; четвертая матрица получена из третьей путем умножения второй строки на $(-2)$ и прибавления к четвертой строке, сложения второй и третьей строк.)

Так как ранг последней матрицы равен трем, для исходной матрицы также $r=3$.

## § 3.7. Комплексные числа. Комплексные матрицы

Введем в рассмотрение множество чисел, каждое из которых определяется упорядоченной парой действительных чисел. Действительные числа будем обозначать буквами $a, b, c, \ldots$, а упорядоченные пары действительных чисел — буквами $\alpha, \beta, \gamma, \ldots$ и соответственно писать $\alpha=(a, b)$, $\beta=(c, d)$ и т. д. *Комплексным числом* $\alpha$ назовем упорядоченную пару действительных чисел $(a, b)$. Комплексное число $\alpha=(a, b)$ равно нулю, если $a=0$, $b=0$. Две упорядоченные пары $\alpha_1=(a_1, b_1)$, $\alpha_2=(a_2, b_2)$ считаются равными, если $a_1=a_2$, $b_1=b_2$.

Определим операции над упорядоченными парами действительных чисел. Суммой двух упорядоченных пар $\alpha=(a, b)$, $\beta=(c, d)$ называется упорядоченная пара $\gamma=(a+c, b+d)$:

$$(a, b)+(c, d)=(a+c, b+d); \tag{3.35}$$

произведением — упорядоченная пара $\delta=(ac-bd, ad+bc)$:

$$(a, b)(c, d)=(ac-bd, ad+bc). \tag{3.36}$$

Можно показать, что для этих операций существуют обратные операции — вычитание и деление (кроме деления на нуль). Разностью упорядоченных пар $\alpha=(a, b)$ и $\beta=(c, d)$ служит упорядоченная пара $(a-c, b-d)$, т. е.

$$\alpha-\beta=(a-c, b-d). \tag{3.37}$$

Упорядоченной парой, противоположной для упорядоченной пары $\alpha=(a, b)$, будет пара $-\alpha=(-a, -b)$, так как $\alpha+(-\alpha)=(0, 0)$.

Если $\beta\neq 0$, то частное $\frac{\alpha}{\beta}$ существует и однозначно определяется формулой

$$\frac{\alpha}{\beta}=\left(\frac{ac+bd}{c^2+d^2}, \frac{bc-ad}{c^2+d^2}\right). \tag{3.38}$$

Полагая в этой формуле $\alpha=\beta$ (т. е. $a=c$, $b=d$), находим, что единицей при умножении упорядоченных пар служит упорядоченная пара $(1, 0)$. Положив $\alpha=1=(1, 0)$, из формулы (3.38) получим, что при $\beta \neq 0$ упорядоченной парой, обратной для $\beta$, будет упорядоченная пара

$$\frac{1}{\beta} = \beta^{-1} = \left(\frac{c}{c^2+d^2}, \frac{-d}{c^2+d^2}\right). \tag{3.39}$$

Итак, построено множество чисел, операции над которыми определяются по формулам (3.35) и (3.36). Это множество чисел называется множеством комплексных чисел.

Множество комплексных чисел в качестве своего подмножества содержит все действительные числа. Чтобы показать это, рассмотрим упорядоченные пары вида $(a, 0)$. Каждой паре $(a, 0)$ поставим в соответствие действительное число $a$, в результате получим взаимно-однозначное соответствие между множеством упорядоченных пар и множеством всех действительных чисел. Применяя к указанным упорядоченным парам формулы (3.35) и (3.36), имеем $(a, 0)+(b, 0)=(a+b, 0)$, $(a, 0)\cdot(b, 0)= =(ab, 0)$. Следовательно, упорядоченные пары вида $(a, 0)$ складываются и умножаются так же, как соответствующие действительные числа. Таким образом, множество упорядоченных пар действительных чисел, рассматриваемое как часть множества комплексных чисел, по своим алгебраическим свойствам не отличается от множества действительных чисел. Это позволяет положить

$$(a, 0) = a, \tag{3.40}$$

т. е. не различать упорядоченную пару $(a, 0)$ действительных чисел и действительное число $a$. В частности, нуль $(0, 0)$ и единица $(1, 0)$ множества комплексных чисел оказываются обычными действительными числами 0 и 1.

Покажем, что среди комплексных чисел содержится корень уравнения $x^2+1=0$. Корнем уравнения $x^2+1=0$ является такое число, квадрат которого равен действительному числу $-1$. Это будет упорядоченная пара $(0, 1)$. Действительно, применяя формулу (3.36), получаем $(0, 1)(0, 1)=(-1, 0)=-1$. Обозначим эту упорядоченную пару через $i$, т. е. $i=(0, 1)$, тогда

$$i^2 = -1, \quad i = \sqrt{-1}. \tag{3.41}$$

Найдем произведение действительного числа $b$ на упорядоченную пару $(0, 1)=i$:

$$bi=(b, 0)(0, 1)=(0, b). \tag{3.42}$$

Если $(a, b)$ — произвольная упорядоченная пара, то из очевидного равенства $(a, b)=(a, 0)+(0, b)$ и формул (3.40), (3.42) следует

$$(a, b)=a+bi. \tag{3.43}$$

Итак, комплексное число $\alpha=(a, b)$ может быть записано в виде

$$a+bi\equiv a+ib, \tag{3.44}$$

где $a$ и $b$ — действительные числа; $i$ — мнимая единица, определяемая соотношением (3.41). Выражение $a+bi$ называется алгебраической формой комплексного числа. Число $a$ называется действительной, число $b$ — мнимой частью комплексного числа $a+bi$. Обозначая комплексное число $a+bi$ одной буквой $\alpha$, пишут

$$a=\mathrm{Re}\,\alpha, \quad b=\mathrm{Im}\,\alpha, \tag{3.45}$$

где Re — начальные буквы латинского слова realis (действительный); Im — начальные буквы латинского слова imaginarius (мнимый). Числа вида $bi$ называются чисто мнимыми числами, или просто мнимыми числами.

Два комплексных числа $a+bi$, $c+di$ считаются равными тогда и только тогда, когда равны между собой соответственно их действительные и мнимые части, т. е. $a=c$, $b=d$. Отметим, что из равенства $a+bi=0$ следуют равенства $a=0$, $b=0$.

Если дано комплексное число $\alpha=a+bi$, то число $a-bi$, отличающееся от $\alpha$ только знаком при мнимой части, называется числом, сопряженным с $\alpha$, и обозначается $\bar{\alpha}$. Числом, сопряженным с $\bar{\alpha}$, будет, очевидно, число $\alpha$, поэтому можно говорить о паре сопряженных чисел. Действительные числа, и только они, сопряжены сами себе.

Сложение, вычитание, умножение и деление комплексных чисел $a+bi$, $c+di$ производятся по следующим формулам:

$$(a+bi)+(c+di)=(a+c)+(b+d)i, \tag{3.46}$$

$$(a+bi)-(c+di)=(a-c)+(b-d)i, \tag{3.47}$$

$$(a+bi)(c+di)=(ac-bd)+(ad+bc)i, \tag{3.48}$$

$$\frac{a+bi}{c+di}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i. \tag{3.49}$$

Эти формулы следуют из формул (3.35)—(3.38). Первая из них означает, что при сложении комплексных чисел складываются отдельно их действительные и отдельно их мнимые части. Формула (3.47) выражает правило вычитания комплексных чисел: при вычитании комплексных чисел вычитаются отдельно их действительные и отдельно их мнимые части. Формулу (3.48) можно получить, приняв во внимание свойства умножения и равенство $i^2=-1$. Действительно,

$$(a+bi)(c+di)=ac+adi+bci+bdi^2=$$
$$=ac+(ad+bc)i-bd=(ac-bd)+(ad+bc)i.$$

Формулу (3.49) можно вывести, умножив числитель и знаменатель дроби на комплексное число, сопряженное знаменателю. В самом деле,

$$\frac{a+bi}{c+di}=\frac{(a+bi)(c-di)}{(c+di)(c-di)}=\frac{(ac+bd)+(bc-ad)i}{c^2+d^2}=$$
$$=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.$$

Таким образом, действия над комплексными числами производятся по обычным правилам алгебры; для получения окончательного результата необходимо принять во внимание соотношение $i^2=-1$ и привести подобные члены.

Отметим, что сумма и произведение сопряженных комплексных чисел являются действительными числами. Так, если $\alpha=a+bi$ и $\bar{\alpha}=a-bi$, то $\alpha+\bar{\alpha}=2a$, $\alpha\bar{\alpha}=a^2+b^2$.

Можно показать, что для комплексных чисел $\alpha=a+bi$, $\beta=c+di$, $\bar{\alpha}=a-bi$, $\bar{\beta}=c-di$ выполняются следующие равенства:

$$\overline{\alpha+\beta}=\bar{\alpha}+\bar{\beta},\quad \overline{\alpha\beta}=\bar{\alpha}\bar{\beta},\quad \overline{\left(\frac{\alpha}{\beta}\right)}=\frac{\bar{\alpha}}{\bar{\beta}}. \qquad (3.50)$$

Комплексные числа можно изображать точками плоскости. Фиксируем некоторую плоскость и выберем на ней декартову прямоугольную систему координат. Каждому комплексному числу $\alpha=a+bi$ поставим в соответствие точку плоскости с прямоугольными декартовыми координатами $(a, b)$; обратно, каждой точке $(a, b)$ ставится в соответствие комплексное число $\alpha=a+bi$. Плоскость, точки которой отождествлены с комплексными числами, называется комплексной плоскостью. Действительные числа $\alpha=a+0i=a$, определяемые упорядоченными парами $(a, 0)=a$, изображаются точками на оси $Ox$ (рис. 3.2), поэтому ее называют действительной осью. Чисто мнимые числа $\alpha=0+bi=bi$, определяемые упорядоченными парами $(0, b)$, изображаются точками оси $Oy$, которую называют мнимой осью. Мнимая единица $i$, определяе-

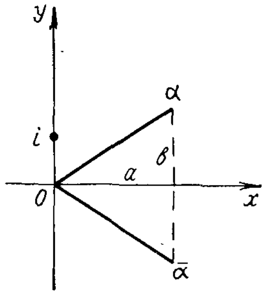



Рис. 3.2

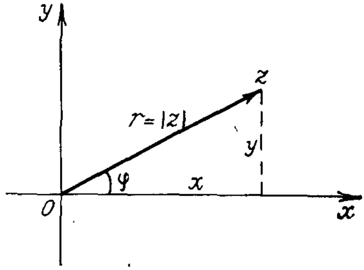



Рис. 3.3

мая упорядоченной парой (0, 1), изображается точкой оси $Oy$, расположенной на расстоянии, равном единице, вверх от начала координат. Любое комплексное число $\alpha=a+ib$ $(a\neq0,\ b\neq0)$ изображается точкой комплексной плоскости, не лежащей на осях координат. Сопряженные комплексные числа $\alpha$ и $\tilde{\alpha}$ изображаются точками комплексной плоскости, симметричными относительно действительной оси.

Комплексные числа и точки комплексной плоскости часто обозначают буквой $z$ и пишут $z=x+iy$, где $x$ — действительная часть $(x=\operatorname{Re} z)$, $y$ — мнимая часть $(y=\operatorname{Im} z)$. Комплексную плоскость называют также плоскостью комплексного переменного $z$. С каждой точкой $z$ комплексной плоскости можно связать радиус-вектор $\mathbf{r}=\overline{Oz}$ этой точки (рис. 3.3). Комплексное число $z=x+iy$ изображается точкой $z$ с прямоугольными декартовыми координатами $(x, y)$ или той же точкой с полярными координатами $(r, \varphi)$, где $r$ — расстояние точки $z$ до полюса $O$, или длина радиуса-вектора $\overline{Oz}$, а $\varphi$ — величина угла наклона радиуса-вектора $\mathbf{r}$ к оси $Ox$.

Модулем комплексного числа $z=x+iy$ называется длина $r$ радиуса-вектора соответствующей точки $z$. Модуль комплексного числа $z$ обозначается символом $|z|$. Следовательно, по определению $r=|z|$ ($|z|\geqslant 0$); если $z=x+iy$, то $|z|=\sqrt{x^2+y^2}$. Аргументом комплексного числа $z$ называется величина угла $\varphi$ наклона радиуса-вектора $\bar{r}=\overline{Oz}$ к оси $Ox$. Аргумент комплексного числа обозначают символом $\operatorname{Arg} z$. Отметим, что $\operatorname{Arg} z=\arg z+2k\pi$, $k=0, \pm 1, \pm 2, \ldots$, где $\arg z$ — главное значение аргумента, определяемое условиями $-\pi<\arg z\leqslant\pi$. Другими словами, модуль и аргумент комплексного числа равны соответственно полярным координатам точки, изображающей данное комплексное число. Так как $x=r\cos\varphi$, $y=r\sin\varphi$, то комплексное число $z=x+iy$ можно представить в тригонометрической форме

$$z=r(\cos\varphi+i\sin\varphi). \tag{3.51}$$

Очевидно, $r(\cos\varphi+i\sin\varphi)=r(\cos(\varphi+2k\pi)+i\sin(\varphi+2k\pi))$. Два комплексных числа, заданных в тригонометрической форме, равны тогда и только тогда, когда их модули равны, а аргументы отличаются на величину, кратную $2\pi$. Число $\bar{z}$, сопряженное числу (3.51), имеет тригонометрическую форму $\bar{z}=r(\cos(-\varphi)+i\sin(-\varphi))$, поэтому $|\bar{z}|=|z|$, $\arg\bar{z}=-\arg z$.

Можно показать, что произведение и частное двух комплексных чисел

$$z_1=r_1(\cos\varphi_1+i\sin\varphi_1),\quad z_2=r_2(\cos\varphi_2+i\sin\varphi_2) \tag{3.52}$$

выражаются соответственно формулами

$$z_1z_2=r_1r_2(\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)), \tag{3.53}$$

$$\frac{z_1}{z_2}=\frac{r_1}{r_2}(\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)). \tag{3.54}$$

Комплексное число (3.51) возводится в степень $n$ ($n$ — натуральное число) по формуле

$$z^n=(r(\cos\varphi+i\sin\varphi))^n=r^n(\cos n\varphi+i\sin n\varphi), \tag{3.55}$$

которая называется формулой Муавра. При $r=1$ получаем формулу

$$(\cos\varphi+i\sin\varphi)^n=\cos n\varphi+i\sin n\varphi. \tag{3.56}$$

Матрица, элементами которой являются комплексные числа, называется комплексной матрицей. Над комплексными матрицами определяются действия, аналогичные действиям над действительными матрицами: сложение, вычитание, умножение матрицы на число, умножение матриц. Определителем квадратной комплексной матрицы будет комплексное число.

Если элементами матрицы $Z$ являются комплексные числа $z_{kj}=x_{kj}+iy_{kj}$, то ее можно представить в виде $Z=X+iY$, где матрицы $X$ и $Y$ составлены из элементов $x_{kj}$, $y_{kj}$ соответственно; матрица $X$ называется действительной частью, матрица $Y$ — мнимой частью матрицы $Z$. Две комплексные матрицы $Z_1=X_1+iY_1$, $Z_2=X_2+iY_2$ считаются равными тогда и только тогда, когда равны их действительные и мнимые части, т. е. $X_1=X_2$, $Y_1=Y_2$.

Матрица $\bar{Z}$, составленная из элементов $\bar{z}_{kj}=x_{kj}-iy_{kj}$, называется комплексно-сопряженной матрице $Z$, элементы которой — комплексные числа $x_{kj}+iy_{kj}$. Отметим, что

$$\overline{Z_1+Z_2}=\bar{Z}_1+\bar{Z}_2,\quad \overline{Z_1Z_2}=\bar{Z}_1\cdot\bar{Z}_2,\quad \overline{\det Z}=\det\bar{Z}. \tag{3.57}$$

# Глава 4

# СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

*В этой главе изучаются линейные алгебраические уравнения*, т. е. уравнения, содержащие неизвестные только в первой степени. Рассматриваются методы исследования систем линейных уравнений.

## § 4.1. Системы $m$ линейных уравнений с $n$ неизвестными

*Системой $m$ линейных уравнений* с $n$ неизвестными $x_1, x_2, \ldots, x_n$ (или *линейной системой*) называется система вида

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=b_m, \end{aligned} \tag{4.1}$$

где $a_{ik}$, $b_i$ — числа. Числа $a_{ik}$ $(i=1, 2, \ldots, m;\ k=1, 2, \ldots, n)$ называются коэффициентами, числа $b_i$ $(i=1, 2, \ldots, m)$ — свободными членами. Коэффициенты обозначены буквой $a$ с двумя индексами $i$ и $k$; первый индекс ($i$) указывает номер уравнения, второй — ($k$) — номер неизвестной, к которой относится данный коэффициент. Число $m$ уравнений может быть больше, равно или меньше числа $n$ неизвестных.

Линейная система называется *неоднородной*, если среди свободных членов имеются отличные от нуля. Если все свободные члены равны нулю, линейная система называется *однородной*. Однородная линейная система имеет вид

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=0,\\ \ldots\ldots\ldots\ldots&\ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=0.\end{aligned} \tag{4.2}$$

*Решением* линейной системы (4.1) называется упорядоченная совокупность $n$ чисел[1]:

$$c_1,\ c_2,\ \ldots,\ c_n, \tag{4.3}$$

подстановка которых вместо $x_1, x_2, \ldots, x_n$ соответственно ($x_1=c_1$, $x_2=c_2, \ldots, x_n=c_n$) обращает в тождество каждое из уравнений этой системы.

Система, имеющая хотя бы одно решение, называется *совместной*, а система, не имеющая ни одного решения,— *несовместной*. Отметим, что однородная система всегда совместна, так как она имеет нулевое решение: $x_1=0$, $x_2=0, \ldots, x_n=0$.

Совместная система называется *определенной*, если она имеет единственное решение, и *неопределенной*, если имеет более одного решения.

Две системы называются *эквивалентными* или *равносильными*, если любое решение одной из них является также решением другой и обратно, т. е. если они имеют одно и то же множество решений. Любые две несовместные системы считаются эквивалентными.

Элементарными преобразованиями линейной системы называются следующие преобразования: 1) умножение уравнения системы на число, отличное от нуля; 2) прибавление к одному уравнению системы другого ее уравнения, умноженного на любое число; 3) перестановка местами двух уравнений системы. Докажем, что при элементарных преобразованиях системы получаем систему, эквивалентную данной.

[1] Числа (4.3) образуют одно решение, а не $n$ решений.

Умножим обе части одного из уравнений системы (4.1) на число $\lambda \neq 0$ и прибавим к соответствующим частям другого уравнения той же системы, в результате получим новое уравнение. Например, если умножить на $\lambda \neq 0$ первое уравнение и прибавить ко второму, то получим уравнение

$$\lambda(a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n)+ \\ +(a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n)=\lambda b_1+b_2 \tag{4.4}$$

или

$$a'_{21}x_1+a'_{22}x_2+\ldots+a'_{2n}x_n=b'_2, \tag{4.5}$$

где

$$a'_{2k}=\lambda a_{1k}+a_{2k} \ (k=1,\ 2,\ \ldots,\ n),\ b'_2=\lambda b_1+b_2. \tag{4.6}$$

Рассмотрим систему уравнений

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a'_{21}x_1+a'_{22}x_2+\ldots+a'_{2n}x_n&=b'_2,\\ a_{31}x_1+a_{32}x_2+\ldots+a_{3n}x_n&=b_3,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=b_m, \end{aligned} \tag{4.7}$$

отличающуюся от системы (4.1) лишь вторым уравнением — уравнением (4.5).

Если $c_1, c_2, \ldots, c_n$ — решение системы (4.1), то оно будет решением системы (4.7), т. е. если значения $x_k=c_k$ ($k=1, 2, \ldots, n$) удовлетворяют каждому уравнению (4.1), то они будут удовлетворять и каждому уравнению (4.7), так как все они, кроме второго, те же, что и в системе (4.1), а второе уравнение — это уравнение (4.4). Обратное также верно: если $c_1, c_2, \ldots, c_n$ — решение системы (4.7), оно будет решением системы (4.1), поскольку второе уравнение (4.1) получается путем умножения первого (4.7) на $(-\lambda)$ и прибавления ко второму (4.7), а остальные уравнения одинаковы в обеих системах. Следовательно, системы (4.1) и (4.7) эквивалентны.

Понятно, что если к системе (4.1) будут несколько раз применены указанные преобразования, новая система будет эквивалентна исходной системе (4.1).

В процессе преобразований указанного типа может оказаться, что в новой системе появится уравнение, все коэффициенты которого равны нулю. Если и свободный член такого уравнения равен нулю, то уравнению будут удовлетворять любые значения неизвестных; отбрасывая это уравнение[1]), получаем систему уравнений, эквивалентную исходной. Если же в рассматриваемом уравнении свободный член отличен от нуля, то уравнение не может быть удовлетворено ни при каких значениях неизвестных, поэтому полученная система уравнений, как и эквивалентная ей исходная система, будет несовместной.

Линейную систему (4.1) можно записать в матричном виде. Матрица

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \tag{4.8}$$

составленная из коэффициентов линейных уравнений системы (4.1), называется *основной матрицей* системы (или матрицей системы). Матрица

$$\tilde{A}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}, \tag{4.9}$$

полученная из основной присоединением столбца из свободных членов, называется *расширенной матрицей* системы (4.1).

Рассмотрим столбцевые матрицы, составленные из неизвестных и свободных членов

$$X=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad B=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \tag{4.10}$$

Поскольку матрица $A$ согласована с матрицей $X$, то можно найти произведение

$$AX=\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}. \tag{4.11}$$

[1]) Таких уравнений может оказаться несколько. Следовательно, число уравнений преобразованной системы в некоторых случаях оказывается меньше числа уравнений исходной системы.

Элементами этой столбцевой матрицы являются левые части уравнений системы (4.1), поэтому на основании определения равенства матриц

$$AX = B. \tag{4.12}$$

Таким образом, система линейных уравнений (4.1) записана в виде одного матричного уравнения (4.12), где $A$, $X$, $B$ определяются формулами (4.8) и (4.10); эта запись системы называется *матричной*.

Каждой линейной системе (4.1) соответствует единственная пара матриц $A$, $B$ и обратно: каждой паре матриц — единственная линейная система. Система (4.1) может быть записана и в таком виде:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{bmatrix} x_2 + \dots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}. \tag{4.13}$$

Если $(c_1, c_2, \dots, c_n)$ — решение системы (4.1), то матрица

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix} \tag{4.14}$$

называется вектором-решением этой системы. Матрица (4.14) удовлетворяет уравнению (4.12).

Выражение «решить систему» означает выяснить, совместна она или несовместна; в случае совместности — найти все ее решения.

## § 4.2. Решение систем линейных уравнений с помощью определителей

Рассмотрим систему $n$ линейных уравнений с $n$ неизвестными $x_1, x_2, \dots, x_n$:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n. \end{aligned}\right\} \tag{4.15}$$

Определителем системы (4.15) называется определитель матрицы $A$ из коэффициентов уравнений этой системы, обозначим его через $\Delta$ ($\Delta = \det A$).

Обозначим через $\Delta_k$ определитель, полученный заменой в определителе $\Delta$ столбца из коэффициентов при неизвестной $x_k$ столбцом свободных членов системы (4.15), т. е.

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nk} & \dots & a_{nn} \end{vmatrix}, \quad \Delta_k = \begin{vmatrix} a_{11} & a_{12} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_2 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & b_n & \dots & a_{nn} \end{vmatrix}, \quad (4.16)$$

где $k$ — одно из чисел $1, 2, \dots, n$.

**Теорема 4.1.** Если определитель системы (4.15) отличен от нуля, то система имеет единственное решение

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad \dots, \quad x_n = \frac{\Delta_n}{\Delta}, \qquad (4.17)$$

где $\Delta$ и $\Delta_k$ $(k=1, 2, \dots, n)$ определены формулами (4.16).

Доказательство. Систему (4.15) запишем в матричной форме $AX=B$. По условию $\Delta=\det A \neq 0$, поэтому для матрицы $A$ существует обратная матрица $A^{-1}$, определяемая формулой (3.34). Умножив обе части равенства $AX=B$ слева на матрицу $A^{-1}$, получим $A^{-1}AX=$ $=A^{-1}B$. Так как $A^{-1}A=E$, $EX=X$, то равенство $A^{-1}AX=$ $=A^{-1}B$ примет вид

$$X=A^{-1}B. \qquad (4.18)$$

Эта формула является матричной записью решения системы (4.15). Матричное равенство (4.18) можно записать так:

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_k \\ \dots \\ x_n \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1k} & A_{2k} & \dots & A_{nk} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_k \\ \dots \\ b_n \end{bmatrix}, \text{ или}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_k \\ \dots \\ x_n \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} A_{11} b_1 + A_{21} b_2 + \dots + A_{n1} b_n \\ A_{12} b_1 + A_{22} b_2 + \dots + A_{n2} b_n \\ \dots \\ A_{1k} b_1 + A_{2k} b_2 + \dots + A_{nk} b_n \\ \dots \\ A_{1n} b_1 + A_{2n} b_2 + \dots + A_{nn} b_n \end{bmatrix}.$$

Из последнего равенства следуют формулы (4.17), поскольку в силу теоремы замещения (см. § 3.4) $A_{1k}b_1+$ $+A_{2k}b_2+\ldots+A_{nk}b_n=\Delta_k$ $(k=1, 2, \ldots, n)$.

Доказанная теорема называется теоремой Крамера, а формулы (4.17) — формулами Крамера[1].

Пример. Решить систему уравнений

$$x-2y+3z=2,$$
$$4x-\ y+5z=15,$$
$$6x-8y+7z=9.$$

Неизвестные здесь обозначены буквами $x$, $y$, $z$, причем $x$ является первой неизвестной ($x_1$), $y$ — второй ($x_2$), $z$ — третьей ($x_3$).

Составим определитель системы $\Delta$ и определители $\Delta_k$ $(k=$ $=1, 2, 3)$:

$$\Delta=\begin{vmatrix} 1 & -2 & 3 \\ 4 & -1 & 5 \\ 6 & -8 & 7 \end{vmatrix}, \quad \Delta_1=\begin{vmatrix} 2 & -2 & 3 \\ 15 & -1 & 5 \\ 9 & -8 & 7 \end{vmatrix},$$

$$\Delta_2=\begin{vmatrix} 1 & 2 & 3 \\ 4 & 15 & 5 \\ 6 & 9 & 7 \end{vmatrix}, \quad \Delta_3=\begin{vmatrix} 1 & -2 & 2 \\ 4 & -1 & 15 \\ 6 & -8 & 9 \end{vmatrix}.$$

Определитель системы $\Delta=-49\neq0$, т. е. выполнено условие теоремы 4.1. Вычисляем определители $\Delta_1$, $\Delta_2$, $\Delta_3$ и пользуемся формулами (4.17), полагая в них $n=3$. Поскольку $\Delta_1=-147$, $\Delta_2=-98$, $\Delta_3=$ $=-49$, то система имеет единственное решение:

$$x=\frac{-147}{-49}=3, \quad y=\frac{-98}{-49}=2, \quad z=\frac{-49}{-49}=1.$$

Следствие из теоремы Крамера. Если однородная система

$$a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,$$
$$a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0$$

имеет ненулевое решение, то ее определитель $\Delta$ равен нулю.

Действительно, если $\Delta\neq0$, то система имеет единственное нулевое решение (так как все $\Delta_k=0$, $k=1, 2, \ldots$ $\ldots, n$, поскольку содержат нулевой столбец).

[1] *Габриэль Крамер* (*Gabriel Cramer*, 1704—1752) — швейцарский математик.

## § 4.3. Метод последовательного исключения неизвестных

Пусть дана система $m$ линейных алгебраических уравнений с $n$ неизвестными $x_1, x_2, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ a_{31}x_1+a_{32}x_2+\ldots+a_{3n}x_n&=b_3,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=b_m. \end{aligned} \tag{4.19}$$

Метод последовательного исключения неизвестных (или метод Гаусса[1]), применяемый для решения системы (4.19), состоит в следующем.

Предполагая, что $a_{11}\neq 0$ (это всегда можно сделать за счет нумерации уравнений), умножая первое уравнение системы (4.19) на $-\frac{a_{21}}{a_{11}}$ и прибавляя ко второму, получаем уравнение, в котором коэффициент при $x_1$ обращается в нуль. Умножая первое уравнение на $-\frac{a_{31}}{a_{11}}$ и прибавляя к третьему, получаем уравнение, также не содержащее члена с $x_1$. Аналогичным путем преобразуем все остальные уравнения, в результате чего придем к системе, эквивалентной исходной системе уравнений:

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a'_{22}x_2+\ldots+a'_{2n}x_n&=b'_2,\\ a'_{32}x_2+\ldots+a'_{3n}x_n&=b'_3,\\ \ldots\ldots\ldots\ldots\ldots&\ldots\\ a'_{m2}x_2+\ldots+a'_{mn}x_n&=b'_m, \end{aligned}\right\} \tag{4.20}$$

где $a'_{ik}$ ($i=2, 3, \ldots, m$; $k=2, 3, \ldots, n$) — некоторые новые коэффициенты.

Предполагая, что $a'_{22}\neq 0$, и оставляя неизменными первые два уравнения системы (4.20), преобразуем ее так, чтобы в каждом из остальных уравнений коэффициент при $x_2$ обратился в нуль. Продолжая этот процесс, систему (4.20) можно привести к одной из следующих систем:

[1] *Карл Фридрих Гаусс* (*Carl Friedrich Gauss*, 1777—1855) — знаменитый немецкий математик.

$$\left.\begin{aligned} c_{11}x_1 + c_{12}x_2 + c_{13}x_3 + \ldots + c_{1n}x_n &= d_1, \\ c_{22}x_2 + c_{23}x_3 + \ldots + c_{2n}x_n &= d_2, \\ c_{33}x_3 + \ldots + c_{3n}x_n &= d_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ c_{nn}x_n &= d_n, \end{aligned}\right\} \quad (4.21)$$

где $c_{kk} \neq 0$ $(k=1, 2, \ldots, n)$; $c_{kk}$ — некоторые коэффициенты; $d_k$ — свободные члены;

$$\left.\begin{aligned} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1k}x_k + \ldots + c_{1n}x_n &= d_1, \\ c_{22}x_2 + \ldots + c_{2k}x_k + \ldots + c_{2n}x_n &= d_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ c_{kk}x_k + \ldots + c_{kn}x_n &= d_n, \end{aligned}\right\} \quad (4.22)$$

где $k < n$;

$$\left.\begin{aligned} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n &= d_1, \\ c_{22}x_2 + \ldots + c_{2n}x_n &= d_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ 0 \cdot x_n &= d_k, \end{aligned}\right\} \quad (4.23)$$

где $k \leqslant n$.

Система (4.21) имеет единственное решение, значение $x_n$ находится из последнего уравнения, значение $x_{n-1}$ — из предпоследнего, значение $x_1$ — из первого.

Система (4.22) имеет бесконечное множество решений. Из последнего уравнения можно выразить одно из неизвестных (например, $x_k$) через остальные $n-k$ неизвестных $(x_{k+1}, x_{k+2}, \ldots, x_n)$, входящих в это уравнение. Из предпоследнего уравнения можно выразить $x_{k-1}$ через эти неизвестные и т. д. В полученных формулах, выражающих $x_1, x_2, \ldots, x_{k-1}, x_k$ через $x_{k+1}, x_{k+2}, \ldots, x_n$, неизвестные $x_{k+1}, x_{k+2}, \ldots, x_n$ могут принимать любые значения.

Система (4.23) несовместна, так как никакие значения неизвестных не могут удовлетворять ее последнему уравнению.

Итак, метод последовательного исключения неизвестных применим к любой системе линейных уравнений. Решая систему этим методом, преобразования совершают не над уравнениями, а над матрицами, составленными из коэффициентов при неизвестных и свободных членов.

Пример 1. Решить систему уравнеиий

$$x_1 + x_2 + 4x_3 + 3x_4 = 2,$$

$$x_1 - x_2 + 12x_3 + 6x_4 = 6,$$
$$4x_1 + 4x_2 - 4x_3 + 3x_4 = 0,$$
$$2x_1 + 2x_2 + 8x_3 - 3x_4 = 1.$$

Составим матрицу из коэффициентов при неизвестных и свободных членов, преобразуем ее:

$$\left[\begin{array}{cccc|c} 1 & 1 & 4 & 3 & 2 \\ 1 & -1 & 12 & 6 & 6 \\ 4 & 4 & -4 & 3 & 0 \\ 2 & 2 & 8 & -3 & 1 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 1 & 4 & 3 & 2 \\ 0 & -2 & 8 & 3 & 4 \\ 0 & 0 & -20 & -9 & -8 \\ 0 & 0 & 0 & -9 & -3 \end{array}\right].$$

Вторая матрица получена из первой путем поочередного умножения первой строки на $(-1)$, $(-4)$, $(-2)$ и прибавления соответственно ко второй, третьей и четвертой строке первой матрицы; вертикальной чертой отделен столбец из свободных членов.

Второй матрице соответствует система уравнений

$$x_1 + x_2 + 4x_3 + 3x_4 = 2,$$
$$-2x_2 + 8x_3 + 3x_4 = 4,$$
$$-20x_3 - 9x_4 = -8,$$
$$-9x_4 = -3,$$

из которой находим $x_4 = \frac{1}{3}$, $20x_3 = 8 - 9x_4$, $20x_3 = 5$, $x_3 = \frac{1}{4}$, $2x_2 = 8x_3 + 3x_4 - 4$, $x_2 = -\frac{1}{2}$, $x_1 = 2 - x_2 - 4x_3 - 3x_4$, $x_1 = \frac{1}{2}$.

Следовательно, исходная система имеет то же решение

$$x_1 = \frac{1}{2}, \quad x_2 = -\frac{1}{2}, \quad x_3 = \frac{1}{4}, \quad x_4 = \frac{1}{3}.$$

Пример 2. Решить систему уравнений

$$x_1 + x_2 + x_3 + x_4 = 5,$$
$$2x_1 - 3x_2 + x_3 - x_4 = 0,$$
$$5x_1 + 4x_3 + 2x_4 = 12,$$
$$3x_1 + 4x_2 - 2x_3 + 6x_4 = -1.$$

Составим матрицу системы и преобразуем ее:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 2 & -3 & 1 & -1 & 0 \\ 5 & 0 & 4 & 2 & 12 \\ 3 & 4 & -2 & 6 & -1 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 0 & -5 & -1 & -3 & -10 \\ 0 & -5 & -1 & -3 & -13 \\ 0 & 1 & -5 & 3 & -16 \end{array}\right] \to$$

$$\rightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 5 \\ 0 & -5 & -1 & -3 & -10 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 1 & -5 & 3 & -16 \end{array}\right].$$

Система несовместна, так как последняя матрица содержнт строку, соответствующую уравнению, в котором все коэффициенты при неизвестных равны нулю, а свободный член отличен от нуля.

Пример 3. Решить систему уравнений

$$\begin{aligned} x_1+x_2+x_3-x_4&=1, \\ 3x_1+x_2+2x_3-2x_4&=2, \\ 2x_1-4x_2-3x_3+6x_4&=7, \\ 7x_1+5x_2+6x_3-6x_4&=6. \end{aligned}$$

Поскольку

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & -1 & 1 \\ 3 & 1 & 2 & -2 & 2 \\ 2 & -4 & -3 & 6 & 7 \\ 7 & 5 & 6 & -6 & 6 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 1 & -1 & 1 \\ 0 & -2 & -1 & 1 & -1 \\ 0 & -6 & -5 & 8 & 5 \\ 0 & -2 & -1 & 1 & -1 \end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 1 & -1 & 1 \\ 0 & -2 & -1 & 1 & -1 \\ 0 & 0 & -2 & 5 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

то данная система сводится к системе трех уравненнй:

$$\begin{aligned} x_1+x_2+x_3-x_4&=1, \\ 2x_2+x_3-x_4&=1, \\ 2x_3-5x_4&=-8, \end{aligned}$$

откуда $x_3=\frac{5}{2}x_4-4$, $2x_2=1-x_3+x_4=5-\frac{3}{2}x_4$, $x_2=\frac{5}{2}-\frac{3}{4}x_4$,

$$x_1=1-x_2-x_3+x_4=1-\left(\frac{5}{2}-\frac{3}{4}x_4\right)-\left(\frac{5}{2}x_4-4\right)+x_4=\frac{5}{2}-\frac{3}{4}x_4.$$

Следовательно, исходная система также нмеет решение

$$x_1=\frac{5}{2}-\frac{3}{4}x_4,\quad x_2=\frac{5}{2}-\frac{3}{4}x_4,\quad x_3=\frac{5}{2}x_4-4,$$

где $x_4$ может принимать любые действнтельные значения.

## § 4.4. Исследование систем линейных уравнений

Как уже отмечалось, метод Гаусса применим к любой системе линейных уравнений. Этот метод позволяет выяснить, совместна ли данная система; с его помощью находятся все решения совместной линейной системы.

Исследование систем линейных уравнений можно осуществлять с помощью других методов, основанных на теоремах, которые приводятся здесь без доказательств.

Пусть дана система $m$ уравнений с $n$ неизвестными $x_1, x_2, \ldots, x_n$:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{array}\right\} \tag{4.24}$$

Рассмотрим основную матрицу $A$ этой системы и ее расширенную матрицу $\tilde{A}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix}. \tag{4.25}$$

**Теорема 4.2 (Кронекера — Капелли).** Для совместности системы (4.24) необходимо и достаточно, чтобы ранг ее матрицы был равен рангу расширенной матрицы.

**Теорема 4.3.** Если ранг матрицы совместной системы равен числу неизвестных, то система имеет единственное решение.

**Теорема 4.4.** Если ранг матрицы совместной системы меньше числа неизвестных, то множество ее решений является бесконечным.

Базисным минором матрицы называется отличный от нуля ее минор, порядок которого равен рангу этой матрицы. Базисными неизвестными совместной системы, ранг матрицы которой равен $r$, назовем $r$ неизвестных, коэффициенты при которых образуют базисный минор. Остальные неизвестные назовем свободными.

Из теорем следует, что решение системы линейных уравнений можно проводить следующим образом.

1. Находят ранг $r$ матрицы $A$ системы и ранг $\tilde{r}$ расширенной матрицы $\tilde{A}$. Если $r \neq \tilde{r}$, то система несовместна.

2. Если $r=\tilde{r}$, то выделяют базисный минор и базисные неизвестные. Исходную систему уравнений заменяют эквивалентной ей системой, состоящей из тех $r$ уравнений, в которые вошли элементы базисного минора.

Отметим, что в случае, когда число базисных неизвестных равно числу неизвестных системы, система имеет единственное решение, которое можно найти по формулам Крамера. Если число базисных неизвестных меньше числа всех неизвестных, то из соответствующей системы находят выражения базисных неизвестных через свободные, используя, например, формулы Крамера. Придавая свободным неизвестным произвольные значения, получают бесконечное множество решений исходной системы.

Пример. Решить систему уравнений

$$x_1-3x_2+2x_3=1,$$
$$2x_1+x_2-4x_3=5,$$
$$5x_1-8x_2+2x_3=8.$$

Поскольку

$$A=\begin{bmatrix}1 & -3 & 2\\ 2 & 1 & -4\\ 5 & -8 & 2\end{bmatrix}\to\begin{bmatrix}1 & -3 & 2\\ 0 & 7 & -8\\ 0 & 7 & -8\end{bmatrix}\to\begin{bmatrix}1 & -3 & 2\\ 0 & 7 & -8\\ 0 & 0 & 0\end{bmatrix},\quad r=2;$$

$$\tilde{A}=\begin{bmatrix}1 & -3 & 2 & 1\\ 2 & 1 & -4 & 5\\ 5 & -8 & 2 & 8\end{bmatrix}\to\begin{bmatrix}1 & -3 & 2 & 1\\ 0 & 7 & -8 & 3\\ 0 & 7 & -8 & 3\end{bmatrix}\to$$

$$\to\begin{bmatrix}1 & -3 & 2 & 1\\ 0 & 7 & -8 & 3\\ 0 & 0 & 0 & 0\end{bmatrix},\quad \tilde{r}=2,\quad r=\tilde{r},$$

то система совместна. В матрице $A$ минор $\begin{vmatrix}1 & -3\\ 2 & 1\end{vmatrix}$ отличен от нуля, ему соответствует система уравнений $x_1-3x_2=1-2x_3$, $2x_1+x_2=5+4x_3$, в которой $x_1$, $x_2$ — базисные неизвестные, $x_3$ — свободная неизвестная. Решая эту систему по формулам Крамера, находим

$$x_1=\frac{10x_3+16}{7},\qquad x_2=\frac{8x_3+3}{7},$$

где $x_3$ может принимать любые действительные значения.

# Глава 5

# ВЕКТОРЫ

Некоторые физические величины (например, температура, масса, работа) могут быть охарактеризованы одним числом, которое выражает отношение этой величины к соответствующей единице измерения; такие величины называются *скалярными*. Другие величины (например, сила, перемещение точки, скорость, ускорение) характеризуются числом и направлением, эти величины называются *векторными*. Для геометрического изображения физических векторных величин служат векторы.

## § 5.1. Основные понятия

*Вектором* называется *направленный отрезок*. На рисунке направление вектора обычно обозначают стрелкой (рис. 5.1). Если начало вектора находится в точке $A$, конец — в точке $B$, то вектор обозначается символом $\overline{AB}$ или $\overrightarrow{AB}$. Начало вектора называют также *точкой его приложения*. Вектор иногда обозначается одной строчной буквой жирного шрифта **a**, **b** и т. д. или такой же буквой светлого шрифта с черточкой наверху $\bar{a}$, $\bar{b}$ и т. д.

*Модулем вектора* **a** называется его длина, он обозначается через $|\mathbf{a}|$ или просто $a$. Модуль вектора — скалярная неотрицательная величина.

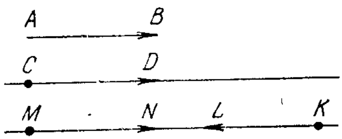


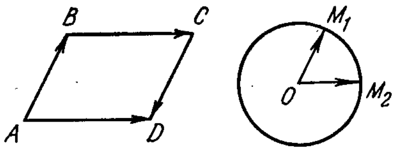

Рис. 5.1

*Нуль-вектором* (или нулевым вектором) называется вектор, начало и конец которого совпадают, обозначается символом **0**. Модуль нуль-вектора равен нулю, а направление не определено.

*Единичным* вектором называется вектор, длина которого равна единице.

Векторы, лежащие на параллельных прямых (или на одной прямой), называются *коллинеарными* (на рис. 5.1 векторы $\overline{CD}$ и $\overline{MN}$, $\overline{KL}$ и $\overline{MN}$, $\overline{CD}$ и $\overline{KL}$).

Коллинеарные векторы, имеющие одинаковые направления и равные длины, называются *равными*. На рис. 5.2, *а* изображен параллелограмм $ABCD$, где векторы $\overline{BC}$ и $\overline{AD}$ равны. Так как векторы $\overline{AB}$ и $\overline{CD}$ имеют противоположные направления, то $\overline{AB} \neq \overline{CD}$, хотя $|\overline{AB}| = |\overline{CD}|$. Отметим, что $\overline{OM_1} \neq \overline{OM_2}$, где $M_1$ и $M_2$ — две различные точки




*а* *б*

Рис. 5.2

окружности радиуса $R$ с центром в точке $O$ (рис. 5.2, *б*), поскольку векторы $\overline{OM_1}$ и $\overline{OM_2}$ имеют разные направления.

Векторы, противоположно направленные и имеющие равные длины, называются *противоположными* (векторы $\overline{AB}$ и $\overline{CD}$ на рис. 5.2, *а*). Вектор, противоположный вектору **a**, обозначается через —**a**.

Векторы, лежащие в параллельных плоскостях (или в одной плоскости), называются *компланарными*.

Из определения равенства векторов следует, что каков бы ни был вектор **a** и точка $A$, всегда можно построить единственный вектор $\overline{AB}$ с началом в точке $A$, равный вектору **a**, т. е. $\overline{AB} = \mathbf{a}$, или, как говорят, перенести вектор **a** в точку $A$. Вектор, точка приложения которого может быть выбрана произвольно, называется *свободным*.

В дальнейшем преимущественно будем рассматривать свободные векторы.

Замечание. Иногда свобода перемещення вектора ограничивается. Если точка приложения вектора строго фиксироваиа, то вектор изывается *связанным*. Примером связаиного вектора может служить *радиус-вектор* точки $M$ (см. § 5.5). Если задана прямая, иа которой должеи быть расположен вектор, то он иазывается *скользящим*. Примером такого вектора является вектор угловой скорости при вращательиом движеиии, который располагается иа оси вращеиия.

## § 5.2. Линейные операции над векторами

*Линейными операциями* над векторами называются сложение, вычитание и умножение вектора на число.

*Суммой* двух векторов **a** и **b** называется третий вектор **c**, начало которого совпадает с началом вектора **a**, а конец — с концом вектора **b** при условии, что вектор **b** отложен из конца вектора **a** (рис. 5.3, *а*). Вектор **c** получается по правилу треугольника (рис. 5.3, *а*) или по правилу параллелограмма (рис. 5.3, *б*).

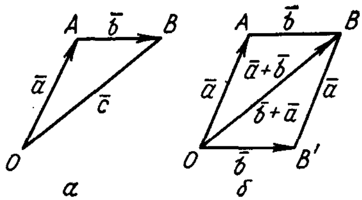



Рис. 5.3

Аналогично определяется сумма трех и более векторов. *Суммой n* векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ называется вектор, начало которого совпадает с началом первого вектора $\mathbf{a}_1$, конец — с концом последнего $\mathbf{a}_n$ при условии, что каждый последующий вектор $\mathbf{a}_{k+1}$ отложен из конца предыдущего $\mathbf{a}_k$ $(k=1, 2, \ldots, n)$. Указанный способ построения суммы называется правилом замыкающей. На рис. 5.4 изображена сумма трех векторов **a**, **b**, **c**. Очевидно, сумма векторов обладает свойством *переместительности (коммутативности)*

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}, \tag{5.1}$$

так как $\overline{OB}=\overline{OA}+\overline{AB}=\mathbf{a}+\mathbf{b}$ и $\overline{OB}=\overline{OB'}+\overline{B'B}=\mathbf{b}+\mathbf{a}$ (см. рис. 5.3, *б*), и свойством *сочетательности (ассоциативности)*

$$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c}), \tag{5.2}$$

поскольку $\overline{OC}=\overline{OB}+\overline{BC}=(\overline{OA}+\overline{AB})+\overline{BC}=(\mathbf{a}+\mathbf{b})+\mathbf{c}$ и $\overline{OC}=\overline{OA}+\overline{AC}=\overline{OA}+(\overline{AB}+\overline{BC})=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ (см. рис. 5.4).

При определении суммы не предполагалось, что векторы являются компланарными. Сумма трех некомпланарных векторов **a**, **b**, **c**, наряду с правилом замыкающей, получается и по *правилу параллелепипеда*: сумма $\mathbf{a}+\mathbf{b}+\mathbf{c}$ равна вектору $\overline{OD}$, где $OD$ — диагональ параллелепипеда, построенного на векторах $\overline{OA}=\mathbf{a}$, $\overline{OB}=\mathbf{b}$, $\overline{OC}=\mathbf{c}$, отложенных из одной точки $O$ (рис. 5.5).

Из определения суммы следует, что

$$\mathbf{a}+\mathbf{0}=\mathbf{a}, \tag{5.3}$$

т. е. нуль-вектор при сложении векторов играет ту же роль, что и число 0 при сложении чисел,

$$\mathbf{a}+(-\mathbf{a})=\mathbf{0}, \tag{5.4}$$

т. е. сумма противоположных векторов равна нуль-вектору.

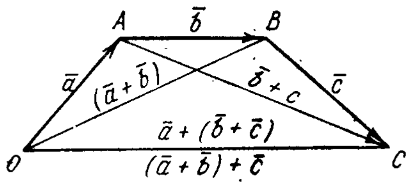



Рис. 5.4

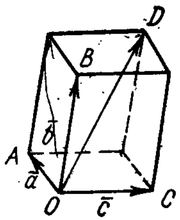



Рис. 5.5

*Разностью* $\mathbf{a}-\mathbf{b}$ двух векторов $\mathbf{a}$ и $\mathbf{b}$ называется такой вектор $\mathbf{d}$, который в сумме с вектором $\mathbf{b}$ дает вектор $\mathbf{a}$:

$$\mathbf{a}-\mathbf{b}=\mathbf{d}, \text{ если } \mathbf{b}+\mathbf{d}=\mathbf{a}. \tag{5.5}$$

Чтобы получить разность $\mathbf{a}-\mathbf{b}$ двух векторов $\mathbf{a}$ и $\mathbf{b}$, необходимо отложить их из одной точки и соединить конец второго вектора с концом первого (рис. 5.6, *а*).

Отметим, что

$$\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b}), \tag{5.6}$$

т. е. разность $\mathbf{a}-\mathbf{b}$ равна сумме двух векторов $\mathbf{a}$ и $(-\mathbf{b})$, где $(-\mathbf{b})$ — вектор, противоположный вектору $\mathbf{b}$ (рис. 5.6, *б*).

Векторы-диагонали параллелограмма $OABC$ (рис. 5.6, *в*), построенного на векторах $\overline{OA}=\mathbf{a}$, $\overline{OB}=\mathbf{b}$, являются соответственно суммой и разностью этих векторов.

*Произведением* вектора $\mathbf{a}$ на число $\alpha$ называется новый вектор

$$\mathbf{b}=\alpha\mathbf{a}, \tag{5.7}$$

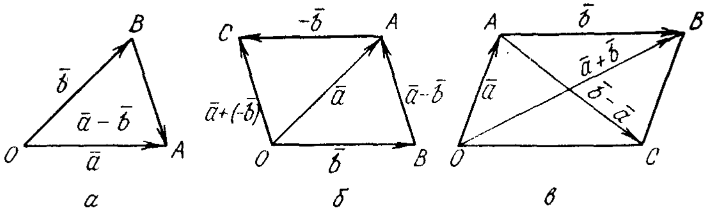



Рис. 5.6

удовлетворяющий условиям: 1) $|\mathbf{b}|=|\alpha|\,|\mathbf{a}|$; 2) $\mathbf{b}$ и $\mathbf{a}$ одинаково направлены при $\alpha>0$; 3) $\mathbf{b}$ и $\mathbf{a}$ имеют противоположные направления при $\alpha<0$ (на рис. 5.7, *а* изображены векторы $\mathbf{a}$, $-2\mathbf{a}$, $3{,}5\,\mathbf{a}$); очевидно, $\mathbf{b}=\mathbf{0}$, если $\alpha=0$ или $\mathbf{a}=\mathbf{0}$.

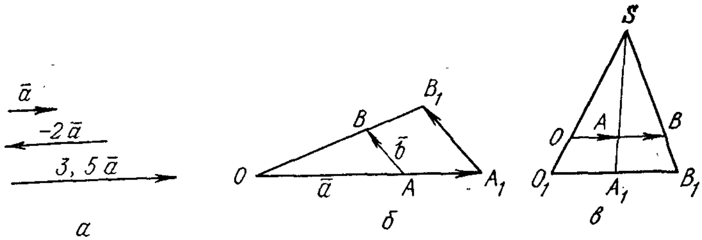



**Рис. 5.7**

Произведение вектора на число обладает следующими свойствами:

$$\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a}; \tag{5.8}$$

$$\alpha(\mathbf{a}+\mathbf{b})=\alpha\mathbf{a}+\alpha\mathbf{b}; \tag{5.9}$$

$$(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}. \tag{5.10}$$

Докажем эти свойства. Векторы $\alpha(\beta\mathbf{a})$ и $(\alpha\beta)\mathbf{a}$ имеют равные длины, так как

$$|\alpha(\beta\mathbf{a})|=|\alpha|\,|\beta\mathbf{a}|=|\alpha|\,|\beta|\,|\mathbf{a}|,$$

$$|(\alpha\beta)\mathbf{a}|=|\alpha\beta|\,|\mathbf{a}|=|\alpha|\,|\beta|\,|\mathbf{a}|,$$

и одинаковые направления, поскольку эти направления совпадают с направлением вектора $\mathbf{a}$ при $\alpha\beta>0$ и противоположны ему при $\alpha\beta<0$, следовательно, $\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a}$, т. е. выполнено равенство (5.8).

Если $\alpha>0$, то равенство (5.9) следует из подобия треугольников $OAB$ и $OA_1B_1$ (рис. 5.7, *б*), где $\overline{OA}=\mathbf{a}$, $\overline{AB}=\mathbf{b}$, $\overline{OA_1}=\alpha\mathbf{a}$, $\overline{OB_1}=\alpha\mathbf{b}$, когда векторы $\mathbf{a}$ и $\mathbf{b}$ не коллинеарны, или треугольников $SOB$, $SO_1B_1$ (рис. 5.7, *в*), где $\overline{SO_1}=\alpha\overline{SO}$, $\overline{OA}=\mathbf{a}$, $\overline{AB}=\mathbf{b}$, когда $\mathbf{a}$ и $\mathbf{b}$ коллинеарны. Случай $\alpha<0$ рассматривается аналогично. При $\alpha=0$ равенство (5.9) очевидно.

Предположим, что $\alpha\beta>0$. Очевидно, векторы, стоящие в обеих частях равенства (5.10), имеют одинаковые направления. Поскольку $|\alpha\mathbf{a}+\beta\mathbf{a}|=|\alpha\mathbf{a}|+|\beta\mathbf{a}|=|\alpha|\,|\mathbf{a}|+|\beta|\,|\mathbf{a}|=(|\alpha|+|\beta|)\mathbf{a}=|\alpha+\beta|\,|\mathbf{a}|=|(\alpha+\beta)\mathbf{a}|$, то они имеют также и равные длины. Следовательно, $(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}$.

Если $\alpha\beta<0$ и, например, $|\beta|>|\alpha|$, то $\alpha+\beta$ и $(-\alpha)$ имеют одинаковые знаки; на основании доказанного $(\alpha+\beta)\mathbf{a}+(-\alpha)\mathbf{a}=(\alpha+\beta-\alpha)\mathbf{a}=\beta\mathbf{a}$, $(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}$.

С помощью метода математической индукции можно доказать следующие равенства:

$$\alpha(\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_n)=\alpha\mathbf{a}_1+\alpha\mathbf{a}_2+\ldots+\alpha\mathbf{a}_n; \quad (5.11)$$

$$(\alpha_1+\alpha_2+\ldots+\alpha_n)\mathbf{a}=\alpha\mathbf{a}_1+\alpha\mathbf{a}_2+\ldots+\alpha\mathbf{a}_n. \quad (5.12)$$

Докажем, например, равенство (5.11). Для $n=2$ оно доказано (равенство (5.9)). Пусть оно верно для $n-1$ слагаемых ($n>1$), тогда оно будет верным и для $n$ слагаемых, ибо

$$\alpha[\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_{n-1}+\mathbf{a}_n]=$$
$$=\alpha[(\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_{n-1})+\mathbf{a}_n]=$$
$$=\alpha(\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_{n-1})+\alpha\mathbf{a}_n=$$
$$=\alpha\mathbf{a}_1+\alpha\mathbf{a}_2+\ldots+\alpha\mathbf{a}_{n-1}+\alpha\mathbf{a}_n.$$

## § 5.3. Условие коллинеарности двух векторов

2

Сделаем предварительное замечание. Если $\mathbf{a}$ — некоторый ненулевой вектор и $\mathbf{a}_0$ — единичный вектор того же направления (рис. 5.8), то из определения произведения вектора на число следует равенство

$$\mathbf{a}=|\mathbf{a}|\,\mathbf{a}_0. \quad (5.13)$$

Умножая обе части этого равенства на число $\alpha = \frac{1}{|\mathbf{a}|}$ ($|\mathbf{a}| \neq 0$), получаем

$$\mathbf{a}_0 = \frac{1}{|\mathbf{a}|}\mathbf{a}, \text{ или } \mathbf{a}_0 = \frac{\mathbf{a}}{|\mathbf{a}|}. \qquad (5.14)$$

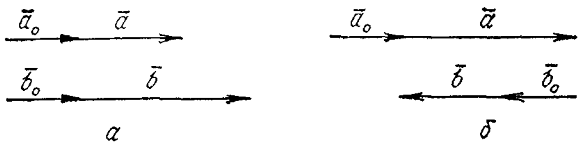


Рис. 5.8

Докажем, что необходимое и достаточное условие коллинеарности двух векторов $\mathbf{a}$ и $\mathbf{b}$ выражается равенством

$$\mathbf{b} = \alpha\mathbf{a}. \qquad (5.15)$$

Действительно, если выполнено равенство (5.15), то векторы $\mathbf{b}$ и $\mathbf{a}$ коллинеарны, что следует из определения произведения вектора на число. Обратно, если $\mathbf{b}$ и $\mathbf{a}$ коллинеарны, то единичные векторы $\mathbf{a}_0$ и $\mathbf{b}_0$ одинаково направлены (рис. 5.8, *а*) или имеют противоположные направления (рис. 5.8, *б*), т. е.

$$\mathbf{a}_0 = \mathbf{b}_0 \text{ или } \mathbf{a}_0 = -\mathbf{b}_0. \qquad (5.16)$$

Принимая во внимание формулу (5.14), последние равенства можно записать так:

$$\frac{\mathbf{a}}{|\mathbf{a}|} = \frac{\mathbf{b}}{|\mathbf{b}|} \text{ или } \frac{\mathbf{a}}{|\mathbf{a}|} = -\frac{\mathbf{b}}{|\mathbf{b}|}.$$

Из этих равенств следует, что $\mathbf{b} = \frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}$ или $\mathbf{b} = -\frac{|\mathbf{b}|}{|\mathbf{a}|}\mathbf{a}$, т. е. $\mathbf{b} = \alpha\mathbf{a}$, где $\alpha = \frac{|\mathbf{b}|}{|\mathbf{a}|}$ или $\alpha = -\frac{|\mathbf{b}|}{|\mathbf{a}|}$.

## § 5.4. Проекция вектора на ось

В пространстве заданы вектор $\overline{AB}$ и ось $n$ (рис. 5.9). Пусть $A_1$ — проекция точки $A$ на ось $n$, $B_1$ — проекция точки $B$, т. е. основания перпендикуляров, опущенных из данных точек на эту ось.

*Проекцией*[1] вектора на ось $n$ называется величина[2] направленного отрезка (вектора) $A_1B_1$ оси $n$. Проекция вектора $\overline{AB}$ на ось $n$ обозначается через $\text{пр}_n\overline{AB}$, т. е.

$$A_1B_1 = \text{пр}_n\overline{AB}. \tag{5.17}$$

Очевидно,

$$\text{пр}_n\overline{AB} = |\overline{AB}| \cos\varphi, \tag{5.18}$$

где $\varphi$ — угол между вектором $\overline{AB}$ и осью $n$.

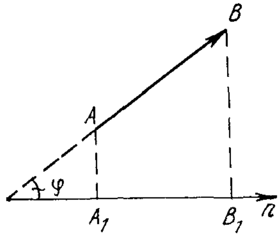



Рис. 5.9

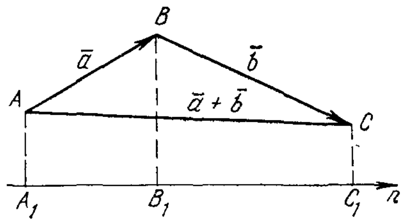



Рис. 5.10

Из равенства (5.18) следует, что если $\mathbf{a}=\mathbf{b}$, то

$$\text{пр}_n\mathbf{a} = \text{пр}_n\mathbf{b}, \tag{5.19}$$

т. е. равные векторы имеют равные проекции (на одну и ту же ось).

Проекция вектора на ось обладает следующими свойствами:

$$\text{пр}_n(\mathbf{a}+\mathbf{b}) = \text{пр}_n\mathbf{a} + \text{пр}_n\mathbf{b}; \tag{5.20}$$

$$\text{пр}_n(\alpha\mathbf{a}) = \alpha\,\text{пр}_n\mathbf{a}. \tag{5.21}$$

Пусть $\overline{AB}=\mathbf{a}$, $\overline{BC}=\mathbf{b}$, $\overline{AC}=\mathbf{a}+\mathbf{b}$, $A_1$, $B_1$, $C_1$ — проекции точек $A$, $B$, $C$ на ось $n$ (рис. 5.10). Запишем основное тождество (см. § 1.1) для трех точек $A_1$, $B_1$, $C_1$ оси $n$: $A_1B_1+B_1C_1=A_1C_1$.

Далее, по определению $A_1B_1=\text{пр}_n\mathbf{a}$, $B_1C_1=\text{пр}_n\mathbf{b}$, $\overline{AC}= =\text{пр}_n(\mathbf{a}+\mathbf{b})$. Подставив три последних равенства в пре-

---

[1] Эту проекцию иногда называют *алгебраической*, в отличие от *геометрической* проекции — вектора $\overline{A_1B_1}$. В дальнейшем, если это не оговорено особо, будем рассматривать алгебраические проекции.

[2] См. § 1.1.

дыдущее, получим равенство (5.20). Равенство (5.21) следует из подобия треугольников $OAA_1$, $OBB_1$ (рис. 5.11). С помощью метода математической индукции можно доказать, что

$$\text{пр}_n(\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_n)=$$
$$=\text{пр}_n\mathbf{a}_1+\text{пр}_n\mathbf{a}_2+\ldots+\text{пр}_n\mathbf{a}_n. \qquad (5.22)$$

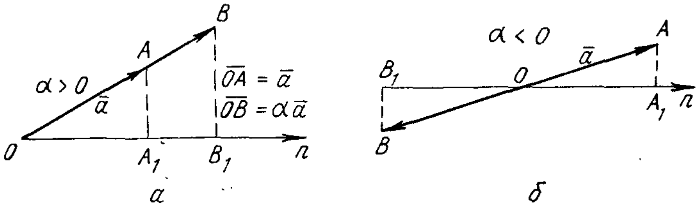



Рис. 5.11

Это доказательство предоставляется читателю.

Если

$$\mathbf{a}_1,\ \mathbf{a}_2,\ \ldots,\ \mathbf{a}_n \qquad (5.23)$$

— произвольная конечная система векторов, $\alpha_1$, $\alpha_2$, … …, $\alpha_n$ — произвольная система действительных чисел, то вектор

$$\mathbf{a}=\alpha_1\mathbf{a}_1+\alpha\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n \qquad (5.24)$$

называется *линейной комбинацией* векторов системы (5.23).

Из равенств (5.21) и (5.22) следует:

$$\text{пр}_n(\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n)=$$
$$\alpha_1\text{пр}_n\mathbf{a}_1+\alpha_2\text{пр}_n\mathbf{a}_2+\ldots+\alpha_n\text{пр}_n\mathbf{a}_n. \qquad (5.25)$$

## § 5.5. Декартовы прямоугольные координаты вектора в пространстве. Длина вектора. Направляющие косинусы вектора

Рассмотрим в пространстве декартову прямоугольную систему координат. *Радиусом-вектором* точки $M$ называется вектор

$$\mathbf{r}=\overline{OM}, \qquad (5.26)$$

точка приложения которого совпадает с началом координат, а конец находится в точке $M$ (рис. 5.12).

*Декартовыми прямоугольными координатами* $X$, $Y$, $Z$ вектора $\mathbf{r}$ называются его проекции на координатные оси

$$X = \text{пр}_x\mathbf{r},\quad Y = \text{пр}_y\mathbf{r},\quad Z = \text{пр}_z\mathbf{r}. \tag{5.27}$$

Каждая из записей

$$\mathbf{r}(X, Y, Z),\ \mathbf{r} = \{X, Y, Z\},\ \mathbf{r} = (X, Y, Z) \tag{5.28}$$

означает, что вектор $\mathbf{r}$ имеет координаты $X$, $Y$, $Z$.

Если $x$, $y$, $z$ — декартовы прямоугольные координаты точки $M$, то

$$X = x,\ Y = y,\ Z = z, \tag{5.29}$$

т. е. координаты радиуса-вектора $\overline{OM}$ равны координатам точки (это следует из определения координат вектора $\mathbf{r}$ и координат точки $M$).

Введем в рассмотрение единичные векторы $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ координатных осей (их называют *ортами*) и векторы

$$\overline{OA} = X\mathbf{i},\quad \overline{OB} = Y\mathbf{j},\quad \overline{OC} = Z\mathbf{k}, \tag{5.30}$$

где $A$, $B$, $C$ — вершины прямоугольного параллелепипеда, для которого $OM$ является диагональю ($A$, $B$, $C$ — проекции точки $M$ на координатные оси, $OA = X$, $OB = Y$, $OC = Z$ — проекции вектора $\overline{OM}$ на координатные оси). По определению суммы $\overline{OM} = \overline{OA} + \overline{OB} + \overline{OC}$, поэтому

$$\mathbf{r} = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}. \tag{5.31}$$

Формула (5.31) выражает разложение вектора $\mathbf{r}$ по *базисным векторам* $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. Векторы, стоящие в правой части формулы (5.31), называются *составляющими*, или *компонентами*, вектора $\mathbf{r}$.

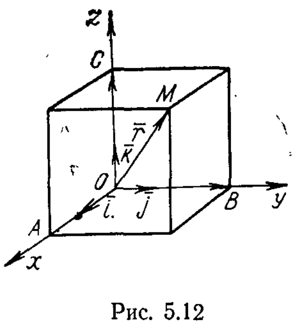


Рис. 5.12

На основании теоремы о квадрате диагонали прямоугольного параллелепипеда получаем формулу, выражающую длину вектора (5.28) (или (5.31)) через его координаты:

$$|\mathbf{r}| = \sqrt{X^2 + Y^2 + Z^2}. \tag{5.32}$$

Из равенства (5.19) следует, что равные векторы имеют равные координаты, поэтому координаты вектора не зависят от его точки приложения.

Координатами любого вектора называются его проекции на координатные оси.

Направляющими косинусами вектора называются косинусы углов $\alpha$, $\beta$, $\gamma$, образуемых им с координатными осями.

Принимая во внимание формулу (5.18), для вектора (5.28) получаем

$$X = |\mathbf{r}| \cos\alpha, \quad Y = |\mathbf{r}| \cos\beta, \quad Z = |\mathbf{r}| \cos\gamma. \tag{5.33}$$

Из равенств (5.32) и (5.33) находим формулы для направляющих косинусов вектора $\mathbf{r}$:

$$\cos\alpha = \frac{X}{\sqrt{X^2+Y^2+Z^2}}, \quad \cos\beta = \frac{Y}{\sqrt{X^2+Y^2+Z^2}},$$

$$\cos\gamma = \frac{Z}{\sqrt{X^2+Y^2+Z^2}}. \tag{5.34}$$

Возводя в квадрат обе части каждого равенства (5.34) и почленно складывая, имеем

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1; \tag{5.35}$$

следовательно, сумма квадратов направляющих косинусов вектора равна единице.

Из формул (5.33) следует, что координаты единичного вектора $\mathbf{e}$ равны его направляющим косинусам, т. е.

$$\mathbf{e} = (\cos\alpha, \cos\beta, \cos\gamma). \tag{5.36}$$

Пример. Дан вектор $\mathbf{a} = (1, -2, 2)$. Найти его длину и единичный вектор $\mathbf{a}_0$ направления вектора $\mathbf{a}$.

По формуле (5.32) находим длину вектора $\mathbf{a}$: $|\mathbf{a}| = \sqrt{1^2+(-2)^2+2^2} = 3$, а по формулам (5.34) — его направляющие косинусы: $\cos\alpha = \frac{1}{3}$, $\cos\beta = -\frac{2}{3}$, $\cos\gamma = \frac{2}{3}$, $\mathbf{a}_0 = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$.

## § 5.6. Переход от векторных соотношений к координатным

При рассмотрении некоторых вопросов теории и решении практических задач часто бывает необходимым переход от векторных соотношений к координатным. Если даны векторы (т. е. известны их координаты) и указаны определенные соотношения между ними, то они равносильны аналогичным числовым соотношениям между координатами. Приведем простейшие примеры.

**Координаты произведения вектора на число.** Пусть дан вектор $\mathbf{a}=(X, Y, Z)$ и число $\alpha\neq 0$. Найти координаты вектора $\mathbf{b}=\alpha\mathbf{a}$.

На основании определений (см. (5.7), (5.27)) и свойств проекций (см. (5.21)) заключаем, что искомые координаты $X_2$, $Y_2$, $Z_2$ вектора $\mathbf{b}$ выражаются формулами

$$X_2=\alpha X_1,\ Y_2=\alpha Y_1,\ Z_2=\alpha Z_1, \qquad (5.37)$$

так как $X_2=\text{пр}_x\mathbf{b}=\text{пр}_x(\alpha\mathbf{a})=\alpha\,\text{пр}_x\mathbf{a}=\alpha X_1$, $Y_2=\text{пр}_y\mathbf{b}$, $Z_2=\text{пр}_z\mathbf{b}$.

Равенства (5.37) выражают необходимое и достаточное условие коллинеарности двух векторов: $\mathbf{a}=(X_1, Y_1, Z_1)$, $\mathbf{b}=(X_2, Y_2, Z_2)$. Если ни одно из чисел $X_1$, $Y_1$, $Z_1$ не равно нулю, то эти равенства можно записать так:

$$\frac{X_2}{X_1}=\frac{Y_2}{Y_1}=\frac{Z_2}{Z_1}. \qquad (5.38)$$

Итак, векторы коллинеарны тогда и только тогда, когда пропорциональны их одноименные координаты.

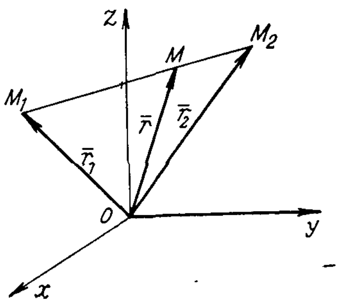



Рис. 5.13

**Координаты суммы (разности) двух векторов.** Пусть даны два вектора $\mathbf{a}=(X_1, Y_1, Z_1)$, $\mathbf{b}=(X_2, Y_2, Z_2)$. На основании формул (5.20) и (5.27) получаем координаты $X$, $Y$, $Z$ вектора суммы $\mathbf{a}+\mathbf{b}$:

$$X=X_1+X_2,\ Y=Y_1+Y_2,\ Z=Z_1+Z_2. \qquad (5.39)$$

Так как $\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b})$, то

$$X'=X_1-X_2,\ Y'=Y_1-Y_2,\ Z'=Z_1-Z_2, \qquad (5.40)$$

где $X'$, $Y'$, $Z'$ — координаты разности $\mathbf{a}-\mathbf{b}$.

**Координаты вектора, заданного двумя точками.** Начало вектора $\overline{M_1M_2}$ находится в точке $M_1(x_1, y_1, z_1)$, конец — в точке $M_2(x_2, y_2, z_2)$. Найдем выражения для его координат через координаты точек $M_1$ и $M_2$.

Введем в рассмотрение радиусы-векторы точек $M_1$ и $M_2$, т. е. $\mathbf{r}_1=\overline{OM_1}$, $\mathbf{r}_2=\overline{OM_2}$ (рис. 5.13). Очевидно, $\overline{M_1M_2}=\mathbf{r}_2-\mathbf{r}_1$. В силу равенств (5.29) $\mathbf{r}_1=(x_1, y_1, z_1)$, $\mathbf{r}_2=$

$=(x_2, y_2, z_2)$. С помощью формул (5.40) получаем координаты $X$, $Y$, $Z$ вектора $\overline{M_1M_2}$

$$X=x_2-x_1,\ Y=y_2-y_1,\ Z=z_2-z_1. \qquad (5.41)$$

Следовательно, чтобы получить координаты вектора, необходимо из координат его конца вычесть соответствующие координаты начала.

**Координаты линейной комбинации векторов.** Заданы $n$ векторов $\mathbf{a}_1=\{X_1, Y_1, Z_1\}$, $\mathbf{a}_2=\{X_2, Y_2, Z_2\}$, ..., $\mathbf{a}_n=\{X_n, Y_n, Z_n\}$ и их линейная комбинация

$$\mathbf{a}=\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n. \qquad (5.42)$$

Принимая во внимание формулы (5.25) и (5.27), заключаем, что координаты вектора (5.42) определяются формулами

$$\left.\begin{aligned} X &= \alpha_1X_1+\alpha_2X_2+\ldots+\alpha_nX_n;\\ Y &= \alpha_1Y_1+\alpha_2Y_2+\ldots+\alpha_nY_n;\\ Z &= \alpha_1Z_1+\alpha_2Z_2+\ldots+\alpha_nZ_n.\end{aligned}\right\} \qquad (5.43)$$

**Деление отрезка в данном отношении.** Даны две точки в пространстве $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$. Найти координаты точки $M$, делящей отрезок $M_1M_2$ в отношении $\lambda$.

По определению (см. § 1.3) $\dfrac{M_1M}{MM_2}=\lambda$, где $M_1M$ и $MM_2$ — величины направленных отрезков $\overline{M_1M}$, $\overline{MM_2}$ оси, проходящей через точки $M_1$ и $M_2$, поэтому $\overline{M_1M}=\lambda\overline{MM_2}$.

Введем в рассмотрение радиусы-векторы точек $M_1$, $M$, $M_2$, т. е. векторы $\mathbf{r}_1=\overline{OM_1}$, $\mathbf{r}=\overline{OM}$, $\mathbf{r}_2=\overline{OM_2}$ (рис. 5.13). Поскольку $\mathbf{r}_1+\overline{M_1M}=\mathbf{r}$ и $\overline{M_1M}=\lambda\overline{MM_2}$, $\overline{MM_2}=\mathbf{r}_2-\mathbf{r}$, т. е. $\overline{M_1M}=\lambda(\mathbf{r}_2-\mathbf{r})$, то $\mathbf{r}_1+\lambda(\mathbf{r}_2-\mathbf{r})=\mathbf{r}$; $(1+\lambda)\mathbf{r}=\mathbf{r}_1+\lambda\mathbf{r}_2$, откуда

$$\mathbf{r}=\frac{1}{1+\lambda}(\mathbf{r}_1+\lambda\mathbf{r}_2),\quad \mathbf{r}=\frac{\mathbf{r}_1+\lambda\mathbf{r}_2}{1+\lambda}. \qquad (5.44)$$

Так как координаты линейной комбинации векторов равны таким же линейным комбинациям их координат (см. формулы (5.43)), то

$$x=\frac{x_1+\lambda x_2}{1+\lambda},\quad y=\frac{y_1+\lambda y_2}{1+\lambda},\quad z=\frac{z_1+\lambda z_2}{1+\lambda}, \qquad (5.45)$$

где $x$, $y$, $z$ — координаты точки $M$ (и координаты ее радиуса-вектора $\mathbf{r}$).

В частности, координаты середины отрезка определяются формулами

$$x=\frac{x_1+x_2}{2},\quad y=\frac{y_1+y_2}{2},\quad z=\frac{z_1+z_2}{2}. \qquad (5.46)$$

**Преобразование декартовых прямоугольных координат при параллельном переносе.** Рассмотрим две декартовы прямоугольные системы координат с одним и тем же масштабным отрезком и одинаковыми направлениями одноименных координатных осей (рис. 5.14). Начало новой системы находится в точке $O_1(a, b, c)$. Пусть $M$ — произвольная точка пространства; $(x, y, z)$ — ее координаты в старой системе, $(X, Y, Z)$ — в новой; $\mathbf{r}=\overline{OM}$, $\mathbf{r}_1=\overline{O_1M}$ — ее радиусы-векторы.

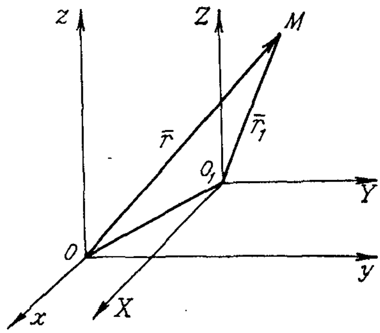


Рис. 5.14

Поскольку $\mathbf{r}=\overline{OO_1}+\mathbf{r}_1$, или $\mathbf{r}=\mathbf{r}_1+\overline{OO_1}$, и $\mathbf{r}=(x, y, z)$, $\mathbf{r}_1=(X, Y, Z)$, $\overline{OO_1}=(a, b, c)$, то

$$x=X+a,\ y=Y+b,\ z=Z+c. \qquad (5.47)$$

Очевидно,

$$X=x-a,\ Y=y-b,\ Z=z-c. \qquad (5.48)$$

## § 5.7. Скалярное произведение двух векторов

*Скалярным произведением* двух векторов $\mathbf{a}$ и $\mathbf{b}$ называется число, равное произведению их длин на косинус угла между ними. Если обозначить скалярное произведение через $\mathbf{ab}$, то

$$\mathbf{ab}=|\mathbf{a}|\,|\mathbf{b}|\cos\varphi. \qquad (5.49)$$

Так как $|\mathbf{b}|\cos\varphi=\text{пр}_{\mathbf{a}}\mathbf{b}$ и $|\mathbf{a}|\cos\varphi=\text{пр}_{\mathbf{b}}\mathbf{a}$ (рис. 5.15), то равенство (5.49) можно представить в виде

$$\mathbf{ab}=|\mathbf{a}|\,\text{пр}_{\mathbf{a}}\mathbf{b} \qquad (5.50)$$

и

$$\mathbf{ab}=|\mathbf{b}|\,\text{пр}_{\mathbf{b}}\mathbf{a}. \qquad (5.51)$$

Понятие скалярного произведения происходит из механики. Если вектор **a** изображает силу, точка приложения которой перемещается из начала в конец вектора **b**, то работа $w$ указанной силы определяется равенством $w=|\mathbf{a}|\,|\mathbf{b}|\cos\varphi$.

*Скалярным квадратом* вектора **a** называется скалярное произведение вектора **a** на себя:

$$\mathbf{a}^2=\mathbf{a}\mathbf{a}=|\mathbf{a}|\,|\mathbf{a}|\cos 0^0=|\mathbf{a}|^2,\ \mathbf{a}^2=|\mathbf{a}|^2. \qquad (5.52)$$

Итак, скалярный квадрат вектора равен квадрату его длины, поэтому $\mathbf{a}^2>0$, когда $|\mathbf{a}|\neq 0$, $\mathbf{a}^2=0$, когда $|\mathbf{a}|=0$.

Пусть векторы **a** и **b** перпендикулярны, т. е. $\varphi=90°$, тогда $\cos\varphi=0$ и

$$\mathbf{a}\mathbf{b}=0. \qquad (5.53)$$

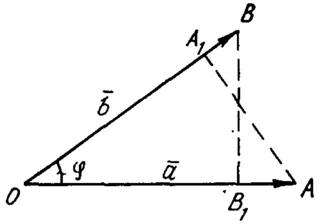


Рис. 5.15

Обратно, если выполнено равенство (5.53), то $\varphi=90°$, когда **a** и **b** — ненулевые векторы, т. е. $\mathbf{a}\perp\mathbf{b}$. Если один из векторов нулевой, то его можно считать перпендикулярным другому (так как нулевой вектор не имеет определенного направления).

Скалярное произведение обладает свойствами:

1) переместительности (коммутативности)

$$\mathbf{a}\mathbf{b}=\mathbf{b}\mathbf{a}; \qquad (5.54)$$

2) сочетательности (ассоциативности) относительно числового множителя

$$(\alpha\mathbf{a})\mathbf{b}=\alpha\mathbf{a}\mathbf{b}; \qquad (5.55)$$

3) распределительности (дистрибутивности) относительно суммы векторов

$$\mathbf{a}(\mathbf{b}+\mathbf{c})=\mathbf{a}\mathbf{b}+\mathbf{a}\mathbf{c}. \qquad (5.56)$$

Докажем справедливость формул (5.54)—(5.56). Формула (5.54) следует из (5.49). Применяя формулы (5.51) и (5.21), получаем (5.55):

$$(\alpha\mathbf{a})\mathbf{b}=|\mathbf{b}|\,\text{пр}_{\mathbf{b}}(\alpha\mathbf{a})=|\mathbf{b}|\,\alpha\,\text{пр}_{\mathbf{b}}\mathbf{a}=$$
$$=\alpha|\mathbf{b}|\,\text{пр}_{\mathbf{b}}\mathbf{a}=\alpha(\mathbf{a}\mathbf{b})=\alpha\mathbf{a}\mathbf{b}.$$

Аналогично доказывается формула (5.56), при этом используются формулы (5.50) и (5.20):

$$\mathbf{a}(\mathbf{b}+\mathbf{c}) = |\mathbf{a}|\,\text{пр}_{\mathbf{a}}(\mathbf{b}+\mathbf{c}) = |\mathbf{a}|(\text{пр}_{\mathbf{a}}\mathbf{b}+\text{пр}_{\mathbf{a}}\mathbf{c}) =$$
$$= |\mathbf{a}|\,\text{пр}_{\mathbf{a}}\mathbf{b}+|\mathbf{a}|\,\text{пр}_{\mathbf{a}}\mathbf{c} = \mathbf{ab}+\mathbf{ac}.$$

Из формул (5.54) и (5.55) следует, что

$$(\alpha\mathbf{a})(\beta\mathbf{b}) = (\alpha\beta)(\mathbf{ab}). \tag{5.57}$$

Действительно,

$$(\alpha\mathbf{a})(\beta\mathbf{b}) = \alpha(\mathbf{a}(\beta\mathbf{b})) = \alpha((\beta\mathbf{b})\mathbf{a}) = \alpha(\beta(\mathbf{ba})) =$$
$$= \alpha\beta(\mathbf{ba}) = \alpha\beta(\mathbf{ab}).$$

Доказанные свойства дают возможность при скалярном умножении векторных многочленов выполнять действия почленно. Например, $(3\mathbf{a}-4\mathbf{b})(5\mathbf{c}+2\mathbf{d}) = 15\mathbf{ac} - 20\mathbf{bc}+6\mathbf{ad}-8\mathbf{bd}$.

Вопрос о выражении скалярного произведения в координатах решает следующая теорема.

**Теорема 5.1.** Скалярное произведение двух векторов

$$\mathbf{a} = (X_1,\ Y_1,\ Z_1),\ \mathbf{b} = (X_2,\ Y_2,\ Z_2) \tag{5.58}$$

выражается формулой

$$\mathbf{ab} = X_1X_2+Y_1Y_2+Z_1Z_2. \tag{5.59}$$

Доказательство. С помощью формул (5.52) и (5.53) получаем «таблицу скалярного умножения» базисных векторов:

$$\left.\begin{array}{lll} \mathbf{i}^2 = 1, & \mathbf{ij} = 0, & \mathbf{ik} = 0; \\ \mathbf{ji} = 0, & \mathbf{j}^2 = 1, & \mathbf{jk} = 0; \\ \mathbf{ki} = 0, & \mathbf{kj} = 0, & \mathbf{k}^2 = 1. \end{array}\right\} \tag{5.60}$$

Воспользуемся разложениями векторов $\mathbf{a}$ и $\mathbf{b}$ по базисным векторам (см. (5.31)):

$$\mathbf{a} = X_1\mathbf{i}+Y_1\mathbf{j}+Z_1\mathbf{k};\quad \mathbf{b} = X_2\mathbf{i}+Y_2\mathbf{j}+Z_2\mathbf{k}.$$

Принимая во внимание формулы (5.54)—(5.57) и таблицу (5.60), находим

$$\mathbf{ab} = (X_1\mathbf{i}+Y_1\mathbf{j}+Z_1\mathbf{k})(X_2\mathbf{i}+Y_2\mathbf{j}+Z_2\mathbf{k}) =$$
$$= X_1X_2\mathbf{i}^2+X_1Y_2\mathbf{ij}+X_1Z_2\mathbf{ik}+Y_1X_2\mathbf{ji}+Y_1Y_2\mathbf{jj}+$$
$$+Y_1Z_2\mathbf{jk}+Z_1X_2\mathbf{ki}+Z_1Y_2\mathbf{kj}+Z_1Z_2\mathbf{k}^2 =$$
$$= X_1X_2+Y_1Y_2+Z_1Z_2,$$

что и требовалось доказать. Итак, скалярное произведение двух векторов равно сумме произведений одноименных координат.

Замечание. Если $\mathbf{b}=\mathbf{a}$, формула (5.59) принимает вид $\mathbf{aa}=X_1^2+Y_1^2+Z_1^2$. Поскольку $\mathbf{aa}=\mathbf{a}^2=|\mathbf{a}|^2$, то

$$|\mathbf{a}|^2=X_1^2+Y_1^2+Z_1^2,\quad |\mathbf{a}|=\sqrt{X_1^2+Y_1^2+Z_1^2}. \tag{5.61}$$

Мы получили формулу (5.32) другим путем. Читателю предлагается получить самостоятельно формулу для расстояния между точками $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$

$$\rho(M_1, M_2)=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}.$$

Следствие 1. Косинус угла между векторами (5.58) определяется формулой

$$\cos\varphi=\frac{X_1X_2+Y_1Y_2+Z_1Z_2}{\sqrt{X_1^2+Y_1^2+Z_1^2}\sqrt{X_2^2+Y_2^2+Z_2^2}}. \tag{5.62}$$

Формула (5.62) следует из формул (5.49), (5.59) и (5.61).

Следствие 2. Необходимое и достаточное условие перпендикулярности двух векторов (5.58) выражается равенством

$$X_1X_2+Y_1Y_2+Z_1Z_2=0. \tag{5.63}$$

Формула (5.63) следует из формул (5.53) и (5.59).

Следствие 3. Если ось $n$ образует с координатными осями углы $\alpha$, $\beta$, $\gamma$ соответственно, то проекция вектора $\mathbf{s}=(X, Y, Z)$ на эту ось определяется равенством

$$\text{пр}_n\bar{s}=X\cos\alpha+Y\cos\beta+Z\cos\gamma. \tag{5.64}$$

Действительно, если $\mathbf{e}$ — единичный вектор оси $n$, то с помощью формулы (5.50) находим

$$\mathbf{es}=|\mathbf{e}|\,\text{пр}_{\mathbf{e}}\mathbf{s}=1\cdot\text{пр}_n\mathbf{s}=\text{пр}_n\mathbf{s},\quad \text{пр}_n\mathbf{s}=\mathbf{es}.$$

Приняв во внимание формулы (5.36) и (5.59), из последнего равенства получим формулу (5.64).

Пример. Даны векторы $\mathbf{a}=(7, 2, -8)$, $\mathbf{b}=(11, -8, -7)$. Найти угол между ними.

По формуле (5.62) получаем

$$\cos\varphi=\frac{7\cdot11+2(-8)+(-8)(-7)}{\sqrt{7^2+2^2+(-8)^2}\sqrt{11^2+(-8)^2+(-7)^2}}=$$

$$=\frac{117}{117\sqrt{2}}=\frac{\sqrt{2}}{2};\quad \varphi=45^\circ.$$

## § 5.8. Правые и левые тройки векторов. Правые и левые системы координат

Три некомпланарных вектора $\overline{OA}=\mathbf{a}$, $\overline{OB}=\mathbf{b}$, $\overline{OC}=\mathbf{c}$, взятых в указанном порядке ($\mathbf{a}$ — первый вектор, $\mathbf{b}$ — второй, $\mathbf{c}$ — третий) и приложенных в одной точке (рис. 5.16, *а*, *b*), называют *тройкой векторов* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$. Будем смотреть с конца вектора $\mathbf{c}$ на плоскость, определяемую векторами $\mathbf{a}$ и $\mathbf{b}$. Если кратчайший поворот от вектора $\mathbf{a}$ к вектору $\mathbf{b}$ совершается против часовой стрелки, то тройка векторов $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ называется *правой*[1] (рис. 5.16, *а*),

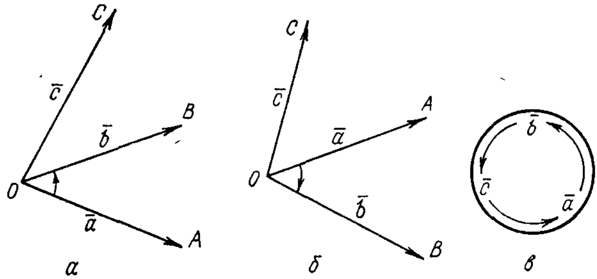



Рис. 5.16.

если указанный поворот совершается по часовой стрелке, тройка $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ называется *левой* (рис. 5.16, *b*).

Две тройки, обе правые или обе левые, называются тройками *одной ориентации*; если одна тройка является правой, а другая — левой, они называются тройками *различной ориентации*. При круговой перестановке векторов (первый заменяется вторым, второй — третьим, третий — первым), схематически изображенной на рис. 5.16, *в*, ориентация тройки не меняется (см. рис. 5.16, *а*, *б*).

Если поменять местами два вектора, то ориентация тройки меняется, например, если $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ — правая тройка (см. рис. 5.16, *а*), то тройка $\mathbf{b}$, $\mathbf{a}$, $\mathbf{c}$ (тех же векторов, взятых в данном порядке) будет левой.

Следовательно, если даны три некомпланарных вектора $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$, то они образуют шесть троек, из которых тройки $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{a}$, $\mathbf{c}$, $\mathbf{a}$, $\mathbf{b}$ являются тройками одной

[1] В случае правой тройки $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ векторы $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ располагаются так, как большой, указательный и средний пальцы правой руки; если тройка $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ является левой, то векторы $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ располагаются так, как указанные пальцы левой руки.

ориентации, тройки **b**, **a**, **c**, **a**, **c**, **b**, **c**, **b**, **a** — тройками другой ориентации.

Декартова прямоугольная система координат называется *правой*, если тройка базисных векторов **i**, **j**, **k** является правой; если эта тройка левая, то система координат называется *левой*.

В данной книге будем пользоваться правыми системами координат.

## § 5.9. Векторное произведение двух векторов

*Векторным произведением* вектора **a** на вектор **b** называется третий вектор, обозначаемый символом [**a**, **b**] и удовлетворяющий следующим условиям:

1) $|[\mathbf{a}, \mathbf{b}]| = |\mathbf{a}|\ |\mathbf{b}| \sin \varphi$, где $\varphi$ — угол между векторами **a** и **b**;

2) вектор [**a**, **b**] перпендикулярен каждому из векторов **a** и **b**;

3) тройки (**a**, **b**, [**a**, **b**]) (**i**, **j**, **k**) являются тройками одной ориентации.

Замечание 1. Для векторного произведения применяются и другие обозначения, например $\mathbf{a} \times \mathbf{b}$.

Замечание 2. Поскольку будем пользоваться только правыми системами координат, условие 3) можно заменить другим — тройка (**a**, **b**, [**a**, **b**]) является правой.

Понятие векторного произведения имеет свой источник в механике. Если вектор **b** изображает силу, приложенную в точке $M$, а вектор $\mathbf{a} = \overline{OM}$, то [**a**, **b**] выражает момент силы **b** относительно точки $O$.

Из условия 1) следует, что модуль векторного произведения [**a**, **b**] равен площади $S$ параллелограмма, построенного на векторах **a** и **b** (рис. 5.17), т. е.

$$|[\mathbf{a}, \mathbf{b}]| = S, \tag{5.65}$$

поэтому (см. (5.13))

$$[\mathbf{a}, \mathbf{b}] = S\mathbf{e}, \tag{5.66}$$

где **e** — единичный вектор направления вектора [**a**, **b**].

Пусть векторы **a** и **b** коллинеарны, т. е. $\varphi = 0$ или $\varphi = \pi$, тогда $\sin \varphi = 0$, $|[\mathbf{a}, \mathbf{b}]| = 0$; следовательно,

$$[\mathbf{a}, \mathbf{b}] = \mathbf{0}. \tag{5.67}$$

Если выполнено равенство (5.67), то в случае ненулевых векторов $\sin\varphi=0$, откуда $\varphi=0$, $\varphi=\pi$, т. е. векторы **a** и **b** коллинеарны. Если один из векторов есть нуль-вектор, его можно считать коллинеарным другому вектору. Итак, равенство (5.67) выражает необходимое и

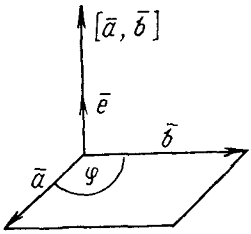



Рис. 5.17

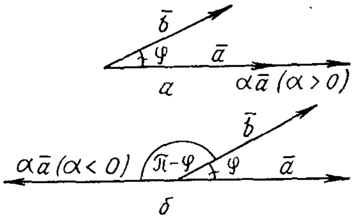



Рис. 5.18

достаточное условие коллинеарности двух векторов **a** и **b**, в частности, для всякого вектора **a**

$$[\mathbf{a}, \mathbf{a}]=0. \qquad (5.68)$$

Векторное произведение двух векторов обладает свойствами:

1) антиперестановочности множителей

$$[\mathbf{a}, \mathbf{b}]=-[\mathbf{b}, \mathbf{a}]; \qquad (5.69)$$

2) сочетательности относительно скалярного множителя

$$[(\alpha\mathbf{a}), \mathbf{b}]=\alpha[\mathbf{b}, \mathbf{a}]; \qquad (5.70)$$

$$[\mathbf{a}, (\beta\mathbf{b})]=\beta[\mathbf{a}, \mathbf{b}]; \qquad (5.71)$$

3) распределительности относительно сложения

$$[(\mathbf{a}+\mathbf{b}), \mathbf{c}]=[\mathbf{a}, \mathbf{c}]+[\mathbf{b}, \mathbf{c}]; \qquad (5.72)$$

$$[\mathbf{a}, (\mathbf{b}+\mathbf{c})]=[\mathbf{a}, \mathbf{b}]+[\mathbf{a}, \mathbf{c}]. \qquad (5.73)$$

Равенство (5.69) следует из определения векторного произведения.

Докажем равенство (5.70). Векторы $[(\alpha\mathbf{a}), \mathbf{b}]$ и $\alpha[\mathbf{a}, \mathbf{b}]$ имеют одинаковые длины, так как при $\alpha>0$ (рис. 5.18)

$$|[(\alpha\mathbf{a}), \mathbf{b}]|=|\alpha\mathbf{a}|\,|\mathbf{b}|\sin\varphi=|\alpha|\,|\mathbf{a}|\,|\mathbf{b}|\sin\varphi,$$

при $\alpha<0$

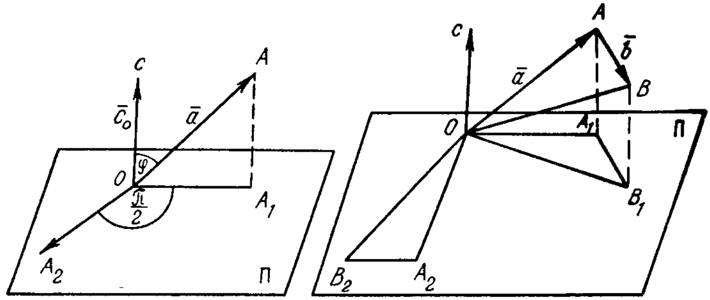
$$|[(\alpha\mathbf{a}), \mathbf{b}]|=|\alpha\mathbf{a}|\,|\mathbf{b}|\sin(\pi-\varphi)=|\alpha|\,|\mathbf{a}|\,|\mathbf{b}|\sin\varphi.$$

Эти векторы также одинаково направлены, поскольку при $\alpha>0$ имеют направление вектора $[\mathbf{a}, \mathbf{b}]$, а прн $\alpha<0$ — направление, противоположное направлению вектора $[\mathbf{a}, \mathbf{b}]$. Формула (5.71) следует из формул (5.69) и (5.70). Действительно,

$$[\mathbf{a}, (\beta\mathbf{b})] = -[(\beta\mathbf{b}), \mathbf{a}] = -\beta[\mathbf{b}, \mathbf{a}] = \beta[\mathbf{a}, \mathbf{b}].$$

Равенство (5.72) докажем сначала для случая, когда вектор $\mathbf{c}$ — единичный, т. е.

$$[(\mathbf{a}+\mathbf{b}), \mathbf{c}_0] = [\mathbf{a}, \mathbf{c}_0] + [\mathbf{b}, \mathbf{c}_0] \quad (|\mathbf{c}_0| = 1). \qquad (5.72')$$




*а* *б*

Рис. 5.19

Предварительно укажем способ построения векторного произведения произвольного вектора $\mathbf{a}$ на единичный вектор $\mathbf{c}_0$. Отложим от фиксированной точки $O$ вектор $\overline{OC} = \mathbf{c}_0$ и вектор $\overline{OA} = \mathbf{a}$, через точку $O$ проведем плоскость П, перпендикулярную вектору $\mathbf{c}_0$ (рис. 5.19). Пусть $A_1$ — ортогональная проекция точки $A$ на плоскость П. Вектор $\overline{OA_1}$ повернем на угол 90° вокруг точки $O$ в плоскости П по часовой стрелке, если смотреть со стороны вектора $\mathbf{c}_0$. Вектор $\overline{OA_2}$, полученный в результате этого поворота, и будет векторным произведением $[\mathbf{a}, \mathbf{c}_0]$. В самом деле,

1) $|\overline{OA_2}| = |\overline{OA_1}| = |\mathbf{a}|\cos\left(\frac{\pi}{2} - \varphi\right) = |\mathbf{a}|\sin\varphi = |\mathbf{a}|\,|\mathbf{c}_0|\sin\varphi$;

2) вектор $\overline{OA_2}$ перпендикулярен каждому из векторов $\mathbf{a}$, $\mathbf{c}_0$;

3) векторы $\mathbf{a}$, $\mathbf{c}_0$, $\overline{OA_2}$ образуют правую тройку.

Проведем построение (рис. 5.19, *б*), аналогичное построению, указанному на рис. 5.19, *а*. Из точки $A$ отложим вектор $\overline{AB}=\mathbf{b}$, тогда $\overline{OB}=\mathbf{a}+\mathbf{b}$. Пусть $A_1$ и $B_1$ — ортогональные проекции точек $A$ и $B$ на плоскость П. Повернув треугольник $OA_1B_1$ на угол $\alpha=90°$ вокруг точки $O$ в плоскости П, получим треугольник $OA_2B_2$. Поскольку $\overline{OB_2}=\overline{OA_2}+\overline{A_2B_2}$ и $\overline{OB_2}=[(\mathbf{a}+\mathbf{b}), \mathbf{c}_0]$, $\overline{OA_2}=[\mathbf{a}, \mathbf{c}_0]$, $\overline{A_2B_2}=[\mathbf{b}, \mathbf{c}_0]$, то из этих равенств и следует равенство (5.72′). Умножив его почленно на $|\mathbf{c}|$ и приняв во внимание формулу $\mathbf{c}=|\mathbf{c}|\mathbf{c}_0$ (см. равенство (5.13)), получим равенство (5.72). Равенство (5.73) следует из равенств (5.69) и (5.72): $[\mathbf{a}, (\mathbf{b}+\mathbf{c})]=-[(\mathbf{b}+\mathbf{c}), \mathbf{a}]=-[\mathbf{b}, \mathbf{a}]-[\mathbf{c}, \mathbf{a}]=[\mathbf{a}, \mathbf{b}]+[\mathbf{a}, \mathbf{c}]$.

Вопрос о выражении векторного произведения через координаты перемножаемых векторов решает следующая теорема.

**Теорема 5.2.** Векторное произведение $[\mathbf{a}, \mathbf{b}]$ двух векторов

$$\mathbf{a}=(X_1, Y_1, Z_1), \mathbf{b}=(X_2, Y_2, Z_2) \tag{5.74}$$

выражается формулой

$$[\mathbf{a}, \mathbf{b}]=\begin{vmatrix} Y_1 Z_1 \\ Y_2 Z_2 \end{vmatrix} \mathbf{i}-\begin{vmatrix} X_1 Z_1 \\ X_2 Z_2 \end{vmatrix} \mathbf{j}+\begin{vmatrix} X_1 Y_1 \\ X_2 Y_2 \end{vmatrix} \mathbf{k}. \tag{5.75}$$

Доказательство. Из определения векторного произведения и равенства (5.68) следует «таблица векторного умножения» ортов $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ в правой декартовой прямоугольной системе координат:

$$[\mathbf{i}, \mathbf{i}]=0, [\mathbf{i}, \mathbf{j}]=\mathbf{k}, [\mathbf{i}, \mathbf{k}]=-\mathbf{j};$$
$$[\mathbf{j}, \mathbf{i}]=-\mathbf{k}, [\mathbf{j}, \mathbf{j}]=0, [\mathbf{j}, \mathbf{k}]=\mathbf{i};$$
$$[\mathbf{k}, \mathbf{i}]=\mathbf{j}, [\mathbf{k}, \mathbf{j}]=-\mathbf{i}, [\mathbf{k}, \mathbf{k}]=0.$$

Принимая во внимание свойства векторного произведения (формулы (5.69)—(5.73)) и указанную таблицу, находим

$$[\mathbf{a}, \mathbf{b}]=[(X_1\mathbf{i}+Y_1\mathbf{j}+Z_1\mathbf{k}), (X_2\mathbf{i}+Y_2\mathbf{j}+Z_2\mathbf{k})]=$$
$$=(X_1Y_2-Y_1X_2)[\mathbf{i}, \mathbf{j}]+(X_1Z_2-Z_1X_2)[\mathbf{i}, \mathbf{k}]+$$
$$+(Y_1Z_2-Z_1Y_2)[\mathbf{j}, \mathbf{k}],$$

т. е.

$$[\mathbf{a}, \mathbf{b}]=(Y_1Z_2-Z_1Y_2)\mathbf{i}-(X_1Z_2-Z_1X_2)\mathbf{j}+(X_1Y_2-Y_1X_2)\mathbf{k}.$$

Переходя к определителям второго порядка, из последней формулы получаем формулу (5.75). Эту формулу можно выразить через символический определитель третьего порядка

$$[\mathbf{a}, \mathbf{b}] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{vmatrix}. \tag{5.76}$$

Замечание. Составим матрицу из координат векторов **a** и **b**:

$$\begin{pmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{pmatrix}.$$

*Координаты векторного произведения* [**a**, **b**] *равны минорам* второго порядка этой матрицы, полученным путем поочередного вычеркивания первого, второго и третьего столбцов, причем второй минор нужно взять со знаком минус.

Следствие 1. Площадь параллелограмма, построенного на векторах (5.74), вычисляется по формуле

$$S = \sqrt{\begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}^2 + \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}^2 + \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix}^2}. \tag{5.77}$$

Эта формула следует из формул (5.32) и (5.75).

Следствие 2. Площадь треугольника $ABC$ определяется формулой

$$S = \frac{1}{2} \left| [\overline{AB}, \overline{AC}] \right|, \tag{5.78}$$

которая следует из формулы (5.65), так как площадь треугольника $ABC$ составляет половину площади параллелограмма, построенного на векторах $\overline{AB}$ и $\overline{AC}$.

Пример 1. Даны два вектора $\mathbf{a} = (7, -5, -6)$, $\mathbf{b} = (1, -2, -3)$. Найти координаты векторного произведения [**a**, **b**].

По формуле (5.76) получаем

$$[\mathbf{a}, \mathbf{b}] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 7 & -5 & -6 \\ 1 & -2 & -3 \end{vmatrix} = \begin{vmatrix} -5 & -6 \\ -2 & -3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 7 & -6 \\ 1 & -3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 7 & -5 \\ 1 & -2 \end{vmatrix} \mathbf{k},$$

$$[\mathbf{a}, \mathbf{b}] = 3\mathbf{i} + 15\mathbf{j} - 9\mathbf{k}, \quad [\mathbf{a}, \mathbf{b}] = \{3, 15, -9\}.$$

Пример 2. Вершины треугольника находятся в точках $A(-1, -1, 1)$, $B(1, -3, 4)$, $C(3, -1, -5)$. Вычислить его площадь.

С помощью формул (5.41) находим координаты векторов $\overline{AB}$ и $\overline{AC}$: $\overline{AB} = (2, -2, 3)$, $\overline{AC} = (4, 0, -6)$. Так как

$$[\overline{AB}, \overline{AC}] = \left( \begin{vmatrix} -2 & 3 \\ 0 & -6 \end{vmatrix}, -\begin{vmatrix} 2 & 3 \\ 4 & -6 \end{vmatrix}, \begin{vmatrix} 2 & -2 \\ 4 & 0 \end{vmatrix} \right) = (12, 24, 8),$$

то

$$S=\frac{1}{2}\left|[\overline{AB},\ \overline{AC}]\right|=\frac{1}{2}\sqrt{12^2+24^2+8^2}=$$
$$=\frac{1}{2}\sqrt{4^2(3^2+6^2+2^2)}=\frac{1}{2}7\cdot 4=14\ (\text{кв. ед.}).$$

## § 5.10. Смешанное произведение трех векторов

Пусть даны три вектора **a**, **b**, **c**. Вектор **a** умножим векторно на **b**, векторное произведение [**a**, **b**] умножим скалярно на **c**, в результате получаем число, которое называют *векторно-скалярным* произведением, или *смешанным произведением* [**a**, **b**] **c** трех векторов **a**, **b**, **c**.

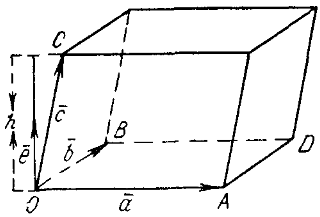


Рис. 5.20

Геометрический смысл смешанного произведения трех векторов выясняет следующая теорема.

**Теорема 5.3.** Смешанное произведение [**a**, **b**] **c** трех некомпланарных векторов равно объему параллелепипеда, построенного на векторах $\overline{OA}=\mathbf{a}$, $\overline{OB}=\mathbf{b}$, $\overline{OC}=\mathbf{c}$, взятому со знаком плюс, если тройка (**a**, **b**, **c**) — правая, со знаком минус, если эта тройка — левая.

Доказательство. Рассмотрим параллелограмм, построенный на векторах $\overline{OA}=\mathbf{a}$, $\overline{OB}=\mathbf{b}$ (рис. 5.20), лежащий в основании указанного параллелепипеда. Его площадь $S$ выражается формулой (5.65), а векторное произведение [**a**, **b**] — формулой (5.66), поэтому $[\mathbf{a},\ \mathbf{b}]\,\mathbf{c}=$ $=(S\mathbf{e})\,\mathbf{c}=S(\mathbf{e}\mathbf{c})$. Применяя формулу (5.50), получаем $(\mathbf{e}\mathbf{c})=|\mathbf{e}|\,\text{пр}_{\mathbf{e}}\,\mathbf{c}=\text{пр}_{\mathbf{e}}\,\mathbf{c}$. С другой стороны, $\text{пр}_{\mathbf{e}}\,\mathbf{c}=\pm h$, где $h$ — высота параллелепипеда, опущенная на основание $OADB$ (рис. 5.20), знак плюс получается в случае, если (**a**, **b**, **c**) — правая тройка, знак минус — если эта тройка левая.

Из трех последних равенств получаем

$$[\mathbf{a},\ \mathbf{b}]\,\mathbf{c}=\pm Sh,\quad [\mathbf{a},\ \mathbf{b}]\,\mathbf{c}=\pm V, \tag{5.79}$$

что и требовалось доказать.

Отметим, что

$$V=\operatorname{mod}([\bar{a}\bar{b}]\bar{c}). \tag{5.80}$$

Следствие 1. Векторы **a**, **b**, **c** компланарны тогда и только тогда, когда их смешанное произведение равно нулю, т. е.

$$[\mathbf{a}, \mathbf{b}]\,\mathbf{c}=0. \qquad (5.81)$$

Действительно, если векторы компланарны, то $[\mathbf{a}, \mathbf{b}]\perp\mathbf{c}$; равенство (5.81) выполняется. Обратное также верно. Если выполнено равенство (5.81), то векторы компланарны. Предположив противное, т. е. что векторы **a**, **b**, **c** некомпланарны, построим на них параллелепипед, тогда по теореме 5.3 $[\mathbf{a}, \mathbf{b}]\,\mathbf{c}=\pm V\neq 0$, что противоречит условию.

Следствие 2. Справедливо равенство $[\mathbf{a}, \mathbf{b}]\,\mathbf{c}=\mathbf{a}\,[\mathbf{b}, \mathbf{c}]$.

В самом деле, так как скалярное произведение не зависит от порядка множителей, то $\mathbf{a}\,[\mathbf{b}, \mathbf{c}]=[\mathbf{b}, \mathbf{c}]\,\mathbf{a}$. По теореме 5.3 $[\mathbf{a}, \mathbf{b}]\,\mathbf{c}=\pm V$, $[\mathbf{b}, \mathbf{c}]\,\mathbf{a}=\pm V$, поскольку речь идет об одном и том же параллелепипеде. Согласно § 5.8 (**a**, **b**, **c**) и (**b**, **c**, **a**) — тройки одной ориентации, поэтому в двух последних равенствах нужно брать один и тот же знак. Следовательно,

$$[\mathbf{a}, \mathbf{b}]\,\mathbf{c}=[\mathbf{b}, \mathbf{c}]\,\mathbf{a}=\mathbf{a}\,[\mathbf{b}, \mathbf{c}].$$

Принимая во внимание эти равенства, смешанное произведение $[\mathbf{a}, \mathbf{b}]\,\mathbf{c}$ и $\mathbf{a}\,[\mathbf{b}, \mathbf{c}]$ обозначают **abc**, т. е.

$$\mathbf{abc}=[\mathbf{a}, \mathbf{b}]\,\mathbf{c}=\mathbf{a}\,[\mathbf{b}, \mathbf{c}]. \qquad (5.82)$$

Замечание. Если поменять местами два вектора, то смешанное произведение изменит лишь знак (см. § 5.8). Для трех векторов **a**, **b**, **c**

$$\mathbf{abc}=\mathbf{bca}=\mathbf{cab}=-\mathbf{bac}=-\mathbf{cba}=-\mathbf{acb}.$$

**Теорема 5.4.** Смешанное произведение трех векторов

$$\mathbf{a}=(X_1, Y_1, Z_1),\ \mathbf{b}=(X_2, Y_2, Z_2),\ \mathbf{c}=(X_3, Y_3, Z_3) \qquad (5.83)$$

определяется формулой

$$\mathbf{abc}=\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}. \qquad (5.84)$$

Доказательство. Поскольку $\mathbf{abc}=[\mathbf{ab}]\,\mathbf{c}$, то с помощью формул (5.75) и (5.59) получаем

$$\mathbf{abc}=X_3\begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}-Y_3\begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}+Z_3\begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix},$$

что равносильно формуле (5.84), так как правая часть последнего равенства является разложением определителя третьего порядка по элементам третьей строки.

Следствие 1 из теоремы 5.3 теперь можно сформулировать следующим образом. Необходимое и достаточное условие компланарности трех векторов (5.83) выражается равенством

$$\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix} = 0. \qquad (5.85)$$

Из формул (5.80), (5.82) и (5.84) следует, что

$$V = \mathrm{mod}\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}. \qquad (5.86)$$

Пример. Найти объем параллелепипеда, построенного на векторах $\mathbf{a}=\{3, 1, 2\}$, $\mathbf{b}=\{2, 2, 3\}$, $\mathbf{c}=\{1, 3, 1\}$.

По формуле (5.86) получаем

$$V = \mathrm{mod}\begin{vmatrix} 3 & 1 & 2 \\ 2 & 2 & 3 \\ 1 & 3 & 1 \end{vmatrix} = \mathrm{mod}\begin{vmatrix} 0 & -8 & -1 \\ 0 & -4 & 1 \\ 1 & 3 & 1 \end{vmatrix} = \mathrm{mod}\begin{vmatrix} -8 & -1 \\ -4 & 1 \end{vmatrix} =$$

$$= |-12| = 12.$$

## § 5.11. Линейная зависимость векторов

Векторы $\mathbf{a}_1, \ldots, \mathbf{a}_n$ называются *линейно зависимыми*, если существуют действительные числа $\alpha_1, \alpha_2, \ldots, \alpha_n$, из которых по меньшей мере одно отлично от нуля, такие, что

$$\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n=\mathbf{0}. \qquad (5.87)$$

В противном случае (т. е. когда таких чисел не существует) векторы называются *линейно независимыми;* другими словами, векторы линейно независимы, если равенство (5.87) выполняется лишь при

$$\alpha_1=\alpha_2=\ldots=\alpha_n=0. \qquad (5.88)$$

Если один из векторов, например $\mathbf{a}_1$, является нулевым, то система $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ окажется линейно зависимой, так как равенство (5.87) будет выполнено при $\alpha_1=1$, $\alpha_2=\alpha_3=\ldots=\alpha_n=0$. Если часть векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно зависима, то и вся система $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ линейно зависима, поскольку из равенства $\alpha_1\mathbf{a}_1+\ldots+\alpha_k\mathbf{a}_k=\mathbf{0}$ $(\alpha_1^2+\ldots+\alpha_k^2\neq 0)$ следует равенство (5.87), в котором $\alpha_{k+1}=\alpha_{k+2}=\ldots=\alpha_n=0$.

**Теорема 5.5.** Чтобы векторы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ $(n>1)$ были линейно зависимы, необходимо и достаточно, чтобы по меньшей мере один из них был линейной комбинацией остальных.

Доказательство. *Необходимость.* Пусть данные векторы линейно зависимы, тогда выполняется равенство (5.87), причем хотя бы одно из чисел $\alpha_k$ $(k=1, 2, \ldots, n)$ отлично от нуля. Предположим, что $\alpha_n\neq 0$, тогда

$$\mathbf{a}_n=-\frac{\alpha_1}{\alpha_n}\mathbf{a}_1-\frac{\alpha_2}{\alpha_n}\mathbf{a}_2-\ldots-\frac{\alpha_{n-1}}{\alpha_n}\mathbf{a}_{n-1}, \qquad (5.89)$$

т. е. $\mathbf{a}_n$ — линейная комбинация остальных векторов.

*Достаточность.* Пусть один из векторов, например $\mathbf{a}_n$, является линейной комбинацией остальных:

$$\mathbf{a}_n=\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_{n-1}\mathbf{a}_{n-1},$$

откуда

$$\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_{n-1}\mathbf{a}_{n-1}+(-1)\mathbf{a}_n=\mathbf{0},$$

т. е. выполнено равенство (5.87), в котором $\alpha_n=-1\neq 0$; это означает, что векторы линейно зависимы.

Отметим, что система из одного ненулевого вектора $\mathbf{a}$ линейно независима, так как равенство $\alpha\mathbf{a}=\mathbf{0}$ возможно только тогда, когда $\alpha=0$; система из одного вектора линейно зависима тогда и только тогда, когда $\mathbf{a}=\mathbf{0}$.

Выясним, что представляют собой линейно зависимые системы из двух и трех векторов.

**Теорема 5.6.** Два вектора $\mathbf{a}_1$ и $\mathbf{a}_2$ линейно зависимы тогда и только тогда, когда они коллинеарны.

Доказательство. Пусть данные векторы коллинеарны, тогда (см. (5.15)) $\mathbf{a}_2=\alpha\mathbf{a}_1$, откуда $\alpha\mathbf{a}_1+(-1)\mathbf{a}_2=\mathbf{0}$, $\alpha_1\mathbf{a}_2+\alpha_2\mathbf{a}_2=\mathbf{0}$, т. е. векторы линейно зависимы, так как выполнено равенство (5.87), где $n=2$, $\alpha_2=-1\neq 0$.

Обратно, если векторы $\mathbf{a}_1$ и $\mathbf{a}_2$ линейно зависимы, т. е. $\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2=\mathbf{0}$, где хотя бы одно из чисел отлично от нуля, например $\alpha_2\neq 0$, то

$$\mathbf{a}_2=-\frac{\alpha_1}{\alpha_2}\mathbf{a}_1, \text{ или } \mathbf{a}_2=\alpha\mathbf{a}_1\left(\alpha=-\frac{\alpha_1}{\alpha_2}\right),$$

а это означает, что векторы коллинеарны.

**Теорема 5.7.** Если $\mathbf{e}_1$ и $\mathbf{e}_2$ — два неколлинеарных вектора некоторой плоскости, то любой третий вектор $\mathbf{a}$ той же плоскости можно единственным образом разложить по ним, т. е. представить в виде

$$\mathbf{a} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2. \tag{5.90}$$

Доказательство. Отложим из одной точки $O$ все три вектора: $\overline{OE_1} = \mathbf{e}_1$, $\overline{OE_2} = \mathbf{e}_2$, $\overline{OA} = \mathbf{a}$ (рис. 5.21). Через точку $A$ проведем две прямые, параллельные соот-

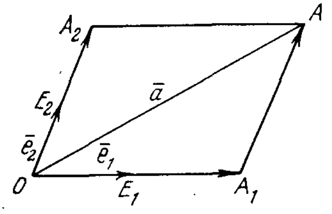



Рис. 5.21

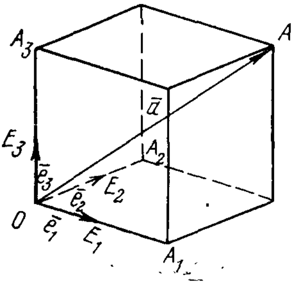



Рис. 5.22

ветственно векторам $\mathbf{e}_1$ и $\mathbf{e}_2$; обозначим через $A_1$ и $A_2$ точки пересечения этих прямых с прямыми, на которых лежат векторы $\mathbf{e}_1$ и $\mathbf{e}_2$.

По определению суммы $\overline{OA} = \overline{OA_1} + \overline{A_1A}$, но $\overline{A_1A} = \overline{OA_2}$, поэтому

$$\mathbf{a} = \overline{OA} = \overline{OA_1} + \overline{OA_2}. \tag{5.91}$$

Поскольку векторы $\mathbf{e}_1$ и $\overline{OA_1}$ коллинеарны, то $\overline{OA_1} = x_1\mathbf{e}_1$ (см. (5.15)), по той же причине $\overline{OA_2} = x_2\mathbf{e}_2$. Подставив эти выражения в равенство (5.91), получим формулу (5.90).

Коэффициенты $x_1$ и $x_2$ разложения (5.90) определяются однозначно. Предположим, что существует разложение

$$\mathbf{a} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2. \tag{5.92}$$

Вычитая почленно равенство (5.92) из равенства (5.90), получаем $(x_1 - y_1)\mathbf{e}_1 + (x_2 - y_2)\mathbf{e}_2 = 0$. По условию векторы $\mathbf{e}_1$ и $\mathbf{e}_2$ не коллинеарны. Из теоремы 5.6 следует, что они линейно независимы, поэтому последнее равенство выполняется лишь при $x_1 - y_1 = 0$, $x_2 - y_2 = 0$, откуда $x_1 = y_1$, $x_2 = y_2$.

**Теорема 5.8.** Три вектора $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ линейно зависимы тогда и только тогда, когда они компланарны.

Доказательство. Пусть векторы $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ линейно зависимы, тогда по теореме 5.5 один из них, например $\mathbf{a}_3$, есть линейная комбинация остальных: $\mathbf{a}_3=\alpha\mathbf{a}_1+\alpha_2\mathbf{a}_2$. Следовательно, вектор $\mathbf{a}_3$ лежит в той плоскости, что и векторы $\mathbf{a}_1$ и $\mathbf{a}_2$, т. е. векторы $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ компланарны.

Обратно, если данные векторы компланарны и векторы $\mathbf{a}_1$, $\mathbf{a}_2$ не коллинеарны, то по теореме 5.7 $\mathbf{a}_3=\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2$, откуда $\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+(-1)\mathbf{a}_3=\mathbf{0}$, т. е. векторы линейно зависимы, так как выполнено равенство (5.87) при $\alpha_3=-1\neq 0$; если $\mathbf{a}_1$ и $\mathbf{a}_2$ коллинеарны, то $\mathbf{a}_2=\alpha\mathbf{a}_1$. Следовательно, $\mathbf{a}_2=\alpha\mathbf{a}_1+0\cdot\mathbf{a}_3$, поэтому векторы линейно зависимы (по теореме 5.5).

**Теорема 5.9.** Если векторы $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ не компланарны, то любой вектор $\mathbf{a}$ можно единственным образом разложить по ним, т. е.

$$\mathbf{a}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3. \tag{5.93}$$

Доказательство. Из некоторой точки $O$ отложим все четыре вектора: $\overline{OE_1}=\mathbf{e}_1$, $\overline{OE_2}=\mathbf{e}_2$, $\overline{OE_3}=\mathbf{e}_3$, $\overline{OA}=\mathbf{a}$. Через точку $A$ проведем три плоскости, параллельные соответственно плоскостям, определяемым каждой парой векторов $(\mathbf{e}_2, \mathbf{e}_3)$, $(\mathbf{e}_1, \mathbf{e}_3)$, $(\mathbf{e}_1, \mathbf{e}_2)$. Пусть $A_1$, $A_2$, $A_3$ — точки пересечения указанных плоскостей с прямыми, на которых соответственно лежат векторы $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. В результате этого построения получим параллелепипед (рис. 5.22), для которого $\overline{OA}$ — диагональ.

По определению суммы трех векторов

$$\mathbf{a}=\overline{OA}=\overline{OA_1}+\overline{OA_2}+\overline{OA_3}. \tag{5.94}$$

Подставляя в это равенство выражения $\overline{OA_1}=x_1\mathbf{e}_1$, $\overline{OA_2}=x_2\mathbf{e}_2$, $\overline{OA_3}=x_3\mathbf{e}_3$, получаем разложение (5.93). Единственность этого разложения доказывается так же, как и в теореме (5.7).

**Теорема 5.10.** Всякие четыре вектора линейно зависимы.

Доказательство. Пусть среди векторов $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ имеется три некомпланарных вектора, например $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, тогда $\mathbf{a}_4=\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\alpha_3\mathbf{a}_3$ (теорема 5.9), т. е. векторы линейно зависимы (по теореме 5.5).

Если векторы $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ линейно зависимы, то $\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\alpha_3\mathbf{a}_3=\mathbf{0}$ $(\alpha_1^2+\alpha_2^2+\alpha_3^2\neq 0)$, следовательно, $\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\alpha_3\mathbf{a}_3+0\mathbf{a}_4=\mathbf{0}$, т. е. векторы $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ линейно зависимы.

## § 5.12. Понятие об аффинных координатах

Зафиксируем некоторую точку $O$ заданной плоскости и выберем два неколлинеарных вектора $\mathbf{e}_1$, $\mathbf{e}_2$. Назовем эту точку *началом координат*, векторы $\mathbf{e}_1$, $\mathbf{e}_2$ — базисными. От точки $O$ отложим векторы $\overline{OE_1}=\mathbf{e}_1$ и $\overline{OE_2}=\mathbf{e}_2$, проведем прямые, которым принадлежат векторы $\overline{OE_1}$ и $\overline{OE_2}$. Фиксируем на них положительные направления, совпадающие с направлениями $\overline{OE_1}$ и $\overline{OE_2}$ соответственно. Получим две координатные оси $Ox$ и $Oy$ (рис. 5.23). Будем говорить, что построена общая *декартова, или аффинная,* система координат $\{O, \mathbf{e}_1, \mathbf{e}_2\}$.

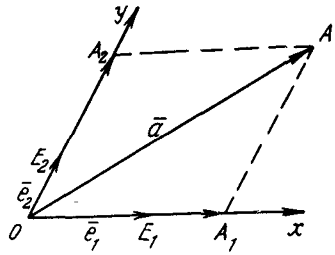



Рис. 5.23

Пусть $\mathbf{a}$ — любой вектор данной плоскости. Отложим из точки $O$ вектор $\overline{OA}=\mathbf{a}$, тогда по теореме 5.7

$$\mathbf{a}=x\mathbf{e}_1+y\mathbf{e}_2. \qquad (5.95)$$

Числа $x$ и $y$ формулы (5.95) называются общими декартовыми, или аффинными, координатами вектора $\mathbf{a}$ в системе $(O, \mathbf{e}_1, \mathbf{e}_2)$, они называются также *аффинными* координатами точки $A$ в той же системе, т. е. $\mathbf{a}=(x, y)$, $A(x, y)$.

Так как $\overline{OA_1}=x\mathbf{e}_1$, $\overline{OA_2}=y\mathbf{e}_2$, то $x$ и $y$ — величины направленных *отрезков $\overline{OA_1}$ и $\overline{OA_2}$ координатных осей; $|x|$ — длина отрезка $\overline{OA_1}$,* измеренная с помощью масштабного отрезка $OE_1$; $|y|$ — длина отрезка $OA_2$, измеренная с помощью масштабного отрезка $OE_2$. Другими словами, аффинными координатами точки $A$ (и вектора $\mathbf{a}=\overline{OA}$) называются числа $x$ и $y$, определяемые формулами

$$x=OA_1,\quad y=OA_2, \qquad (5.96)$$

где $OA_1$, $OA_2$ — величины направленных отрезков $\overline{OA_1}$ и $\overline{OA_2}$ координатных осей ($A_1$ — проекция точки $A$ на ось $Ox$, взятая параллельно оси $Oy$, $A_2$—проекция точки $A$ на ось $Oy$, взятая параллельно оси $Ox$; длины отрезков на каждой оси измеряются с помощью своего масштабного отрезка).

В частном случае, *когда векторы $\mathbf{e}_1$ и $\mathbf{e}_2$ перпендикулярны,* причем $|\mathbf{e}_1|=|\mathbf{e}_2|=1$ (их называют ортами и обозначают соответственно через $\mathbf{i}$, $\mathbf{j}$), получаем декартову прямоугольную систему координат.

Аналогично вводится аффинная система координат в пространстве. Фиксируем начало координат — точку $O$, базис — три некомпланарных вектора $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, отложим из точки $O$ векторы $\overline{OE_1}=\mathbf{e}_1$, $\overline{OE_2}=\mathbf{e}_2$, $\overline{OE_3}=\mathbf{e}_3$, координатные оси $Ox$, $Oy$, $Oz$. Если $\mathbf{a}$ — любой вектор, то, отложив из точки $O$ вектор $\overline{OA}=\mathbf{a}$, по теореме 5.9 получим

$$\mathbf{a}=x\mathbf{e}_1+y\mathbf{e}_2+z\mathbf{e}_3. \qquad (5.97)$$

*Общими декартовыми,* или *аффинными,* координатами вектора $\mathbf{a}$ (и точки $A$) называются числа $x$, $y$, $z$ в разложении (5.97).

Пусть $A_1$ — проекция точки $A$ на ось $x$, взятая параллельно координатной плоскости $Oyz$ (определяемой векторами $\mathbf{e}_2$, $\mathbf{e}_3$), т. е. точка пересечения оси $Ox$ и плоскости, проходящей через точку $A$ и параллельной плоскости $Oyz$; $A_2$ — проекция точки $A$ на ось $Oy$, взятая параллельно плоскости $Oxz$; $A_3$ — проекция точки $A$ на ось $Oz$, взятая параллельно плоскости $Oxy$, тогда $\overline{OA_1}=x\mathbf{e}_1$, $\overline{OA_2}=y\mathbf{e}_2$, $\overline{OA_3}=z\mathbf{e}_3$.

Следовательно, $x$, $y$, $z$ — проекции вектора $\overline{OA}$ на координатные оси, т. е. величины направленных отрезков $\overline{OA_1}$, $\overline{OA_2}$, $\overline{OA_3}$, длины отрезков на каждой координатной оси измеряются с помощью своего масштабного отрезка ($\mathbf{e}_1$ — на оси $Ox$, $\mathbf{e}_2$ — на оси $Oy$, $\mathbf{e}_3$ — на оси $Oz$).

# Глава 6

# ЭЛЕМЕНТЫ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ В ПРОСТРАНСТВЕ

В главе рассматриваются некоторые линии и поверхности в пространстве. Будем исходить из наглядного представления о линии и поверхности, известного из курса математики средней школы. Вводятся понятия уравнений поверхности и линии в пространстве. Как будет показано, поверхность в пространстве определяется одним уравнением с тремя переменными, а линия — двумя такими уравнениями. Основное внимание уделено случаям, когда данные уравнения являются алгебраическими уравнениями первой или второй степени.

Отметим, что в § 6.1—6.10 будут использоваться только декартовы прямоугольные координаты.

## § 6.1. Уравнение поверхности. Уравнения линии в пространстве

*Уравнением поверхности* (в фиксированной системе координат) называется такое уравнение с тремя переменными, которому удовлетворяют координаты любой точки данной поверхности и только они.

Из этого определения вытекает способ решения следующей простой задачи: выяснить, лежит ли данная точка на поверхности, определяемой заданным уравнением. Для решения задачи необходимо подставить ее координаты в данное уравнение, если получается числовое равенство, то точка лежит на поверхности, в противном случае точка поверхности не принадлежит.

Всякое уравнение с тремя переменными $x$, $y$, $z$ можно записать как

$$F(x, y, z)=0^{1)}, \tag{6.1}$$

[1] В некоторых случаях уравнение может не содержать одной или двух переменных.

где $F(x, y, z)$ — некоторая зависимость между $x, y, z$; в частных случаях, например, $F(x, y, z) = x - y + z$, $F(x, y, z) = x^2+y^2+z^2-16$, $F(x, y, z) = x^2+y^2-z^2-1$ и т. д.

В качестве примера составим уравнение сферы радиуса $R$ с центром в точке $C(a, b, c)$. Исходя из определения сферы как множества точек пространства, равноудаленных от данной точки (центра), для произвольной ее точки $M(x, y, z)$ получаем

$$\rho(C, M) = R. \tag{6.2}$$

Так как

$$\rho(C, M) = \sqrt{(x-a)^2+(y-b)^2+(z-c)^2},$$

то

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

или

$$(x-a)^2+(y-b)^2+(z-c)^2=R^2. \tag{6.3}$$

Для точки $N$, не лежащей на данной сфере, равенство (6.2) не будет выполнено, поэтому ее координаты не удовлетворяют уравнению (6.3). Следовательно, уравнение (6.3) и является уравнением сферы радиуса $R$ с центром в точке $C(a, b, c)$.

В частном случае, когда центр сферы находится в начале координат ($a=b=c=0$), уравнение (6.3) принимает вид

$$x^2+y^2+z^2=R^2. \tag{6.4}$$

Уравнение (6.4) называется *каноническим* уравнением сферы.

Приведем еще один пример. Из определения декартовых прямоугольных координат точки (см. § 1.6) следует, что координатные плоскости $Oyz$, $Oxz$, $Oxy$ определяются соответственно уравнениями $x=0$, $y=0$, $z=0$ ($x=0$ — уравнение плоскости $Oyz$ и т. д.).

Поверхность, определяемая алгебраическим уравнением $n$-й степени относительно декартовых координат, называется *поверхностью $n$-го порядка.* В § 6.4, 6.9—6.10 рассматриваются поверхности первого и второго порядка. В § 6.4 показано, что поверхностью первого порядка является плоскость. Отметим, что сфера — поверхность второго порядка (см. уравнения (6.4) и (6.7)).

Линию в пространстве можно рассматривать как пересечение двух поверхностей, поэтому она определяется двумя уравнениями. Пусть $l$ — линия, по которой пересекаются поверхности, определяемые уравнениями $F_1(x, y, z)=0$ и $F_2(x, y, z)=0$, т. е. множество общих точек этих поверхностей, тогда координаты любой точки линии $l$ одновременно удовлетворяют обоим уравнениям

$$\left.\begin{aligned} F_1(x, y, z) &= 0; \\ F_2(x, y, z) &= 0. \end{aligned}\right\} \tag{6.5}$$

Уравнения (6.5) и являются уравнениями указанной линии. Например, уравнения

$$\left.\begin{aligned} x^2+y^2+z^2 &= 9; \\ z &= 0 \end{aligned}\right\}$$

определяют окружность радиуса $R=3$, лежащую в плоскости $Oxy$ (первое уравнение определяет сферу радиуса $R=3$ с центром в начале координат, второе — координатную плоскость $Oxy$). Запишем уравнения координатных осей декартовой системы координат. Ось $Ox$, как линия пересечения плоскостей $Oxz$ и $Oxy$, определяется уравнениями $y=0$, $z=0$, ось $Oy$, как линия пересечения плоскостей $Oyz$, $Oxy$,— уравнениями $x=0$, $z=0$; ось $Oz$ — уравнениями $x=0$, $y=0$.

## § 6.2. Геометрическое значение одного и двух уравнений между координатами в пространстве

Пусть дано уравнение относительно декартовых прямоугольных координат $x$, $y$, $z$

$$\Phi(x, y, z)=0. \tag{6.6}$$

Рассмотрим множество точек пространства, координаты которых удовлетворяют уравнению (6.6). Всякое множество точек пространства называется фигурой; мы не исключаем случая, когда это множество содержит лишь одну точку или оказывается пустым. Следовательно, всякое уравнение (6.6) определяет некоторую фигуру в пространстве.

Выясним, какое множество точек пространства определяет уравнение

$$Ax^2+Ay^2+Az^2+2Bx+2Cy+2Dz+E=0. \tag{6.7}$$

Это уравнение второй степени, имеющее равные коэффициенты при квадратах координат ($A \neq 0$) и не содержащее членов с произведением координат.

**Теорема 6.1.** Если уравнение (6.7) определяет некоторую поверхность в пространстве, то этой поверхностью является сфера.

Доказательство. Так как $A \neq 0$, то обе части уравнения можно разделить на $A$, после чего оно примет вид

$$x^2+y^2+z^2+2bx+2cy+2dz+e=0.$$

Выделяя полные квадраты, получаем

$$(x+b)^2+(y+c)^2+(z+d)^2=f, \tag{6.8}$$

где $f=b^2+c^2+d^2-e$. В зависимости от того, каково $f$ ($f>0$, $f=0$, $f<0$), уравнение (6.8) сводится к одному из уравнений

$$(x+b)^2+(y+c)^2+(z+d)^2=R^2; \tag{6.9}$$

$$(x+b)^2+(y+c)^2+(z+d)^2=0; \tag{6.10}$$

$$(x+b)^2+(y+c)^2+(z+d)^2=-R^2. \tag{6.11}$$

Уравнение (6.9) определяет сферу радиуса $R$ с центром в точке $C(-b, -c, -d)$, уравнение (6.10) — указанную точку, уравнение (6.11) — пустое множество. Теорема доказана.

Отметим, что уравнение $x=a$ определяет плоскость, перпендикулярную оси $Ox$, уравнение $x^2-a^2=0$ — две плоскости (уравнения которых $x=a$, $x=-a$); уравнению

$$\frac{x}{|x|}-\frac{2y}{|y|}+\frac{z}{|z|}=0$$

удовлетворяют координаты точек пространства, заполняющих два октанта.

Пусть заданы два уравнения

$$\left.\begin{aligned} F(x, y, z)=0; \\ \Phi(x, y, z)=0. \end{aligned}\right\} \tag{6.12}$$

Если каждое из этих уравнений определяет поверхность и поверхности пересекаются по некоторой линии, то система (6.12) задает линию их пересечения; если эти поверхности имеют только одну общую точку, система (6.12) определяет эту точку; если поверхности не пересекаются, система определяет пустое множество. В общем случае эта система задает некоторое множество точек пространства. Например, система уравнений

$$\left.\begin{aligned} x^2+y^2+z^2=1; \\ \frac{x}{|x|}+\frac{y}{|y|}-\frac{2z}{|z|}=0 \end{aligned}\right\}$$

определяет множество точек сферы $x^2+y^2+z^2=1$, принадлежащих первому и седьмому октанту.

## § 6.3. Параметрические уравнения линии и поверхности

*Параметрическими уравнениями линии* в пространстве называются уравнения вида

$$x=\varphi_1(t), \quad y=\varphi_2(t), \quad z=\varphi_3(t), \tag{6.13}$$

где $\varphi_1(t)$, $\varphi_2(t)$, $\varphi_3(t)$ — функции некоторой переменной $t$ (параметра), если при каждом значении $t$ из конечного или бесконечного промежутка они дают координаты точек данной линии и только таких точек.

Параметрические уравнения часто применяются в механике для описания траектории движущейся точки; роль параметра $t$ в таких случаях играет время.

В качестве примера приведем параметрические уравнения винтовой линии. *Винтовой линией* называется линия, описываемая точкой, равномерно движущейся по образующей кругового цилиндра, который при этом вращается вокруг своей оси с постоянной угловой скоростью.

Выберем ось вращения цилиндра в качестве оси $Oz$ декартовой прямоугольной системы координат в пространстве (рис. 6.1). Обозначим через $v$ постоянную скорость прямолинейного движения точки вдоль образующей, через $\omega$ — скорость вращательного движения, через $R$ — радиус цилиндра. Пусть в начальный момент точка находилась на оси $Ox$ (совпадала с точкой $A$), а в момент времени $t$ — в положении $M$. Обозначим буквой $N$ проекцию точки $M$ на плоскость $Oxy$, буквой $P$ — проекцию точки $N$ на ось $Ox$, буквой $Q$ — проекцию точки $N$ на ось $Oy$. Обозначая через $\varphi$ угол между $OP$ и $ON$, получаем $x=OP=$ $=R\cos\varphi$, $y=OQ=R\sin\varphi$, $z=NM=vt$.

Поскольку $\varphi=\omega t$, то

$$x=R\cos\omega t, \quad y=R\sin\omega t, \quad z=vt. \tag{6.14}$$

Уравнения (6.14) являются параметрическими уравнениями винтовой линии.

*Параметрическими уравнениями поверхности* называются уравнения вида

$$x=f_1(u, v), \quad y=f_2(u, v), \quad z=f_3(u, v), \tag{6.15}$$

где $f_1(u, v)$, $f_2(u, v)$, $f_3(u, v)$ — функции двух переменных $u$ и $v$ (параметров), если при любых значениях $u$ и $v$ (меняющихся в некоторой области) они дают координаты всех точек данной поверхности и только таких точек[1].

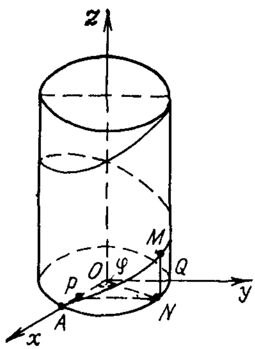



Рис. 6.1

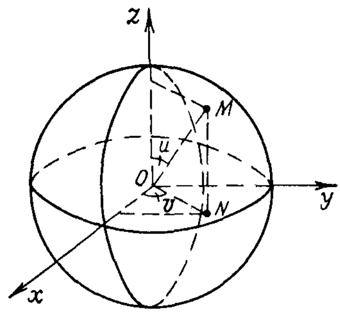



Рис. 6.2

Рассмотрим примеры параметрических уравнений поверхностей.

1. Составим параметрические уравнения сферы радиуса $R$.

Введем в рассмотрение систему декартовых прямоугольных координат с началом в центре сферы и систему сферических координат с началом в той же точке (рис. 6.2). Пусть $M$ — произвольная точка сферы, $N$ — ее проекция на плоскость $Oxy$. Обозначим угол, образуемый вектором $\overline{OM}$ с осью $Oz$, через $u$ *(широта)*; угол, образуемый вектором $\overline{ON}$ с осью $Ox$,— через $v$ *(долгота)*. Принимая во внимание определение декартовых координат (или связь между декартовыми и сферическими координатами, см. § 1.8), получаем параметрические уравнения сферы

[1] Правые части уравнений (6.15) содержат два параметра, а правые части уравнений (6.13) — только один параметр.

$$x = R \sin u \cos v,\ y = R \sin u \sin v,\ z = R \cos u, \quad (6.16)$$

где $0 \leqslant u \leqslant \pi$, $0 \leqslant v < 2\pi$.

Исключив из этих уравнений параметры $u$ и $v$ (для чего нужно возвести в квадрат обе части каждого уравнения и почленно сложить), получим уравнение (6.4), выведенное ранее другим путем.

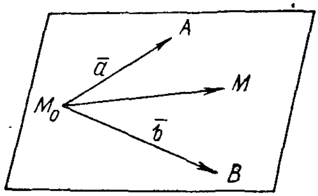



Рис. 6.3

2. Составим параметрические уравнения плоскости, проходящей через точку $M_0(x_0, y_0, z_0)$ и параллельной двум неколлинеарным векторам $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$.

Отложим от точки $M_0$ два вектора: $\overline{M_0A} = \mathbf{a}$ и $\overline{M_0B} = \mathbf{b}$ (рис. 6.3). Пусть $M(x, y, z)$ — произвольная точка плоскости, тогда вектор $\overline{M_0M}$ принадлежит плоскости, определяемой векторами $\overline{M_0A}$ и $\overline{M_0B}$. Этот вектор можно разложить по векторам $\mathbf{a}$ и $\mathbf{b}$ (см. теорему 5.7), т. е. представить в виде

$$\overline{M_0M} = u\mathbf{a} + v\mathbf{b}, \quad (6.17)$$

где $u$ и $v$ — некоторые числа. Меняя точку $M$, получаем новые значения $u$, $v$.

Заметив, что $\overline{M_0M} = (x - x_0,\ y - y_0,\ z - z_0)$, от векторного соотношения (6.17) перейдем к координатным

$$\left.\begin{aligned} x - x_0 &= ua_1 + vb_1; \\ y - y_0 &= ua_2 + vb_2; \\ z - z_0 &= ua_3 + vb_3 \end{aligned}\right\}, \text{ или } \left.\begin{aligned} x &= x_0 + ua_1 + vb_1; \\ y &= y_0 + ua_2 + vb_2; \\ z &= z_0 + ua_3 + vb_3. \end{aligned}\right\} \quad (6.18)$$

Уравнения (6.18) называются параметрическими уравнениями плоскости.

## § 6.4. Плоскость в пространстве

Плоскость в пространстве можно задать различными способами (тремя точками, точкой и вектором, перпендикулярным плоскости, и т. п.). В зависимости от этого рассматривают различные виды ее уравнения.

**Уравнение плоскости, проходящей через данную точку перпендикулярно данному вектору.** Дана точка $M_0(x_0, y_0, z_0)$ и ненулевой вектор

$$\mathbf{n}=(A,\ B,\ C). \qquad (6.19)$$

Требуется составить уравнение плоскости, проходящей через точку $M_0$ перпендикулярно к указанному вектору $\mathbf{n}$ (этот вектор называют *нормальным* вектором плоскости).

Рассмотрим произвольную точку $M(x,\ y,\ z)$ данной плоскости. Так как вектор $\overline{M_0M}=(x-x_0,\ y-y_0,\ z-z_0)$ лежит на плоскости, то он перпендикулярен вектору $\mathbf{n}$ (рис. 6.4). Следовательно, их скалярное произведение равно нулю (см. условие (5.53)), т. е.

$$\mathbf{n}\overline{MM_0}=0. \qquad (6.20)$$

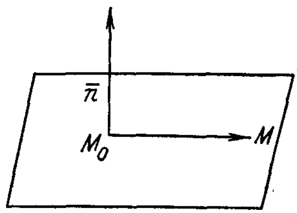



Рис. 6.4

Пользуясь выражением скалярного произведения в координатах (см. формулу (5.59)), получаем искомое уравнение

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0, \qquad (6.21)$$

которому можно придать вид

$$Ax+By+Cz+D=0, \qquad (6.21')$$

где $D=-(Ax_0+By_0+Cz_0)$.

Подчеркнем, что в уравнениях (6.21) и (6.21′) коэффициентами при $x$, $y$, $z$ являются координаты нормального вектора плоскости.

**Общее уравнение плоскости.** Уравнение первой степени относительно декартовых координат

$$Ax+By+Cz+D=0, \qquad (6.22)$$

где $A$, $B$, $C$ одновременно в нуль не обращаются, т. е.

$$A^2+B^2+C^2\neq 0 \qquad (6.23)$$

определяет плоскость в пространстве. Действительно, пусть $x_0$, $y_0$, $z_0$ — три числа, удовлетворяющие уравнению (6.22), т. е. $Ax_0+By_0+Cz_0+D\equiv 0$. Вычитая почленно это тождество из уравнения (6.22), получаем

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0,$$

Это уравнение вида (6.21), оно определяет плоскость, проходящую через точку $M_0(x_0, y_0, z_0)$ и имеющую нормальный вектор $\mathbf{n}=(A, B, C)$ (в силу условия (6.23) этот вектор не является нулевым). Следовательно, уравнение (6.22) тоже определяет плоскость. Уравнение (6.22) называется общим уравнением плоскости. Поскольку это уравнение первого порядка относительно декартовых координат, плоскость является поверхностью первого порядка (в соответствии с определением, данным в § 6.1).

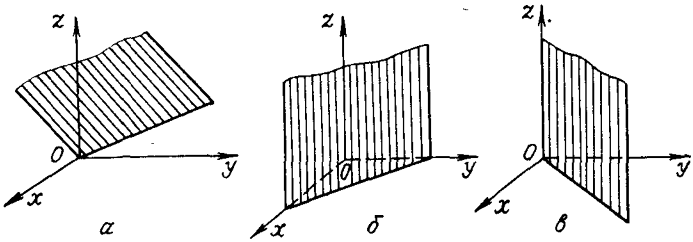



Рис. 6.5

Рассмотрим частные случаи уравнения (6.22), когда один или несколько коэффициентов его обращаются в нуль.

1) $D=0$. Уравнение (6.22) принимает вид

$$Ax+By+Cz=0. \tag{6.22а}$$

Уравнение (6.22а) определяет плоскость, проходящую через начало координат (рис. 6.5, *а*), так как координаты $x=0$, $y=0$, $z=0$ удовлетворяют ему.

2) $C=0$. Уравнение (6.22) принимает вид

$$Ax+By+D=0. \tag{6.22б}$$

Вектор $\mathbf{n}=(A, B, 0)$ перпендикулярен оси $Oz$ (см. определение координат вектора в § 5.5) и данной плоскости (как ее нормальный вектор), поэтому плоскость, определяемая уравнением (6.22б), параллельна оси $Oz$ (рис. 6.5, *б*).

3) $C=0$, $D=0$. Уравнение

$$Ax+By=0 \tag{6.22в}$$

определяет плоскость, проходящую через ось $Oz$ (рис. 6.5, *в*); это следует из двух рассмотренных случаев.

4) $B=0$, $C=0$. Уравнение (6.22) сводится к уравнению

$$Ax+D=0, \text{ или } x=a\left(a=-\frac{D}{A}\right), \qquad (6.22\text{г})$$

которое определяет плоскость, параллельную координатной плоскости $Oyz$, поскольку вектор $\mathbf{n}=(A, 0, 0)$ перпендикулярен оси $Oy$ и оси $Oz$ (см. определение координат вектора), а также перпендикулярен данной плоскости (как ее нормальный вектор). Другими словами, уравнение $x=a$ определяет плоскость, перпендикулярную оси $Ox$ и отсекающую на ней отрезок $OP$, величина которого равна $a$, т. е. $OP=a$.

5) $B=0$, $C=0$, $D=0$. Уравнение (6.22) сводится к уравнению

$$Ax=0, \text{ или } x=0 \qquad (6.22\text{д})$$

(поскольку $A\neq0$ в силу условия (6.23)), и определяет координатную плоскость $Oyz$: это частный случай предыдущего ($D=0$, поэтому $a=0$).

Отметим, что уравнения (6.22г) и (6.22д) можно получить непосредственно, использовав определение декартовых прямоугольных координат точки (см., например, § 6.1).

Читателю предлагается аналогичным образом рассмотреть другие возможные случаи: 6) $B=0$; 7) $B=0$, $D=0$; 8) $A=0$; 9) $A=0$, $D=0$; 10) $A=0$; $C=0$ и т. д. и сделать соответствующие рисунки.

**Нормальное уравнение плоскости.** Пусть дана некоторая плоскость. Через начало координат проведем прямую, перпендикулярную этой плоскости, и обозначим буквой $P$ точку их пересечения (рис. 6.6). Установим на указанной прямой (нормали $n$) положительное направление, совпадающее с направлением вектора $\overline{OP}$. Обозначим через $\alpha$, $\beta$, $\gamma$ углы, образованные нормалью $n$ с координатными осями $Ox$, $Oy$, $Oz$ соответственно, через $p$ — величину вектора $\overline{OP}$, т. е.

$$p=OP. \qquad (6.24)$$

Если $M(x, y, z)$ — произвольная точка плоскости и $\mathbf{r}=\overline{OM}=(x, y, z)$ — ее радиус-вектор, то

$$\text{пр}_n\overline{OM}=p. \qquad (6.25)$$

На основании следствия 3 из теоремы 5.1 (см. § 5.7)

$$\text{пр}_n\overline{OM}=x\cos\alpha+y\cos\beta+z\cos\gamma. \qquad (6.26)$$

Из соотношений (6.25) и (6.26) получаем уравнение

$$x\cos\alpha+y\cos\beta+z\cos\gamma-p=0, \tag{6.27}$$

которое называется *нормальным уравнением* плоскости.

Общее уравнение плоскости (6.22) можно привести к виду (6.27). Умножив обе части уравнения (6.22) на число $\mu\neq 0$, получим

$$\mu Ax+\mu By+\mu Cz+\mu D=0, \tag{6.28}$$

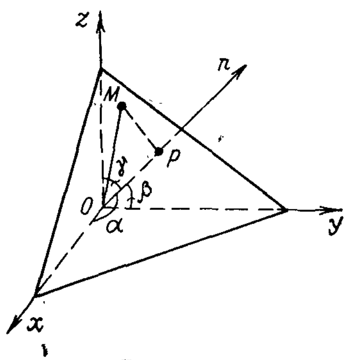


Рис. 6.6

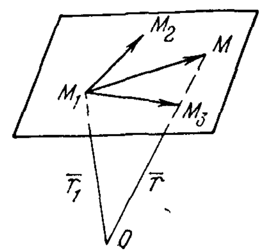


Рис. 6.7

определяющее ту же плоскость, что и исходное уравнение. Выберем $\mu$ так, чтобы

$$\mu A=\cos\alpha,\ \mu B=\cos\beta,\ \mu C=\cos\gamma,\ \mu D=-p. \tag{6.29}$$

Возводя в квадрат обе части каждого из первых трех уравнений (6.29) и почленно складывая, находим

$$\mu^2(A^2+B^2+C^2)=\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1$$

(здесь использовано равенство (5.35)). Так как выполнено условие (6.23), то

$$\mu=\frac{1}{\pm\sqrt{A^2+B^2+C^2}}. \tag{6.30}$$

Число $\mu$, определяемое равенством (6.30), называется *нормирующим множителем*. Принимая во внимание четвертое равенство (6.29), заключаем, что в формуле (6.30) нужно выбрать знак, противоположный знаку $D$ (ибо $p>0$).

Следовательно, уравнение (6.28) принимает *нормальный* вид

$$\frac{Ax+By+Cz+D}{\pm\sqrt{A^2+B^2+C^2}}=0, \qquad (6.31)$$

где следует взять знак, противоположный знаку $D$ в уравнении (6.22).

**Уравнение плоскости, проходящей через три точки.** Составим уравнение плоскости, проходящей через три точки $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, $M_3(x_3, y_3, z_3)$, не лежащие на одной прямой. Пусть $M(x, y, z)$ — произвольная точка этой плоскости (рис. 6.7). Рассмотрим векторы

$$\overline{M_1M}=(x-x_1, y-y_1, z-z_1),$$

$$\overline{M_1M_2}=(x_2-x_1, y_2-y_1, z_2-z_1),$$

$$\overline{M_1M_3}=(x_3-x_1, y_3-y_1, z_3-z_1).$$

Эти векторы компланарны, поэтому их смешанное произведение равно нулю: $\overline{M_1M}\cdot\overline{M_1M_2}\cdot\overline{M_1M_3}=0$ (см. условие (5.81)). Принимая во внимание выражение смешанного произведения в координатах, получаем искомое уравнение

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix}=0. \qquad (6.32)$$

Замечание. Разлагая определитель по элементам первой строки, уравнение (6.32) можно представить в виде $A(x-x_1)+B(y-y_1)+C(z-z_1)=0$, где через $A$, $B$, $C$ обозначены соответствующие определители; эти определители одновременно в нуль обратиться не могут (векторы $\overline{M_1M_2}$ и $\overline{M_1M_3}$ не коллинеарны), так что условие (6.23) будет выполнено.

Читателю предлагается составить уравнение плоскости, проходящей через две точки $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ и параллельной вектору $\mathbf{a}=(a_1, a_2, a_3)$, а также уравнение плоскости, проходящей через точку $M_1(x_1, y_1, z_1)$ и параллельной двум неколлинеарным векторам $\mathbf{a}=(a_1, a_2, a_3)$, $\mathbf{b}=(b_1, b_2, b_3)$.

**Взаимное расположение двух плоскостей.** Рассмотрим две плоскости, заданные уравнениями

$$A_1x+B_1y+C_1z+D_1=0; \qquad (6.33)$$

$$A_2x+B_2y+C_2z+D_2=0. \qquad (6.34)$$

Первая из них имеет нормальный вектор $\mathbf{n}_1=(A_1, B_1, C_1)$, вторая — нормальный вектор $\mathbf{n}_2=(A_2, B_2, C_2)$. Плоскости параллельны, когда векторы $\mathbf{n}_1$ и $\mathbf{n}_2$ коллинеарны, поэтому из условий (5.37) и (5.38) следует, что необходимое и достаточное условие параллельности двух плоскостей (6.33), (6.34) выражается равенствами

$$A_2 = \lambda A_1, \quad B_2 = \lambda B_1, \quad C_2 = \lambda C_1$$

или

$$\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1}. \qquad (6.35)$$

Условие совпадения двух плоскостей выражается равенствами

$$A_2=\lambda A_1,\ B_2=\lambda B_1,\ C_2=\lambda C_1,\ D_2=\lambda D_1,$$

или

$$\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1} = \frac{D_2}{D_1}, \qquad (6.36)$$

так как в этом случае коэффициенты уравнения второй плоскости получаются из соответствующих коэффициентов уравнения первой путем умножения на одно и то же число, отличное от нуля (такие уравнения определяют одну и ту же плоскость).

Если условие (6.35) не выполнено, плоскости пересекаются. В частности, когда плоскости перпендикулярны, то перпендикулярны и векторы $\mathbf{n}_1$, $\mathbf{n}_2$, поэтому их скалярное произведение равно нулю, т. е. $\mathbf{n}_1\mathbf{n}_2=0$, или

$$A_1A_2+B_1B_2+C_1C_2=0. \qquad (6.37)$$

Равенство (6.37) выражает необходимое и достаточное условие перпендикулярности двух плоскостей, заданных уравнениями (6.33) и (6.34).

**Угол между двумя плоскостями.** Пусть даны две плоскости уравнениями (6.33) и (6.34). Угол $\varphi$ между этими плоскостями равен углу между их нормальными векторами $\mathbf{n}_1$ и $\mathbf{n}_2$, поэтому

$$\cos\varphi = \frac{\mathbf{n}_1\mathbf{n}_2}{|\mathbf{n}_1||\mathbf{n}_2|} = \frac{A_1A_2+B_1B_2+C_1C_2}{\sqrt{A_1^2+B_1^2+C_1^2}\sqrt{A_2^2+B_2^2+C_2^2}}. \qquad (6.38)$$

Пример 1. Составить уравнение плоскости, проходящей через точку $M_0(4, -5, 6)$ и имеющей нормальный вектор $\mathbf{n}=\{1, 2, -3\}$.

Так как в данном случае $x_0=4$, $y_0=-5$, $z_0=6$, $A=1$, $B=2$, $C=-3$, то уравнение (6.21) принимает вид $1(x-4)+2(y+5)-3(z-6)=0$, или $x+2y-3z+24=0$.

Пример 2. Составить уравнение плоскости, проходящей через три точки $M_1(1, -4, 9)$, $M_2(-2, -5, 7)$, $M_3(3, -6, 8)$.

В соответствии с формулой (6.32) получаем

$$\begin{vmatrix} x-1 & y+4 & z-9 \\ -2-1 & -5+4 & 7-9 \\ 3-1 & -6+4 & 8-9 \end{vmatrix} = 0, \quad \begin{vmatrix} x-1 & y+4 & z-9 \\ -3 & -1 & -2 \\ 2 & -2 & -1 \end{vmatrix} = 0;$$

$$(x-1)\begin{vmatrix} -1 & -2 \\ -2 & -1 \end{vmatrix} - (y+4)\begin{vmatrix} -3 & -2 \\ 2 & -1 \end{vmatrix} +$$

$$+(z-9)\begin{vmatrix} -3 & -1 \\ 2 & -2 \end{vmatrix} = 0; \quad 3x+7y-8z+97=0.$$

Пример 3. Уравнение $2x-y+2z-15=0$ привести к нормальному виду.

По формуле (6.30) находим

$$\mu = \frac{1}{\sqrt{2^2+(-1)^2+2^2}} = \frac{1}{3}$$

(взят знак плюс, так как $D=-15$). Умножив на $\mu$ исходное уравнение, получим $\frac{2}{3}x - \frac{1}{3}y + \frac{2}{3}z - 5 = 0$.

## § 6.5. Прямая в пространстве

**Векторно-параметрическое уравнение прямой.** *Направляющим вектором прямой* называется любой вектор, лежащий на прямой или параллельный ей.

*Направляющими косинусами прямой* называются направляющие косинусы ее направляющего вектора.

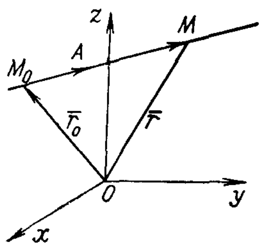



Рис. 6.8

Составим уравнение прямой, проходящей через точку $M_0(x_0, y_0, z_0)$ и имеющей направляющий вектор $\mathbf{a}=(a_1, a_2, a_3)$. Отложим из точки $M_0$ вектор $\overline{M_0A} = \mathbf{a}$ (рис. 6.8). Пусть $M(x, y, z)$ — произвольная точка данной прямой, $\mathbf{r} = \overline{OM}$ — ее радиус-вектор, а $\mathbf{r}_0 = \overline{OM_0}$ — радиус-вектор точки $M_0$, тогда $\overline{OM_0}+\overline{M_0M}=\overline{OM}$, $\overline{M_0M}=\mathbf{a}t$, поэтому

$$\mathbf{r}=\mathbf{r}_0+\mathbf{a}t. \tag{6.39}$$

Уравнение (6.39) называется *векторно-параметрическим* уравнением прямой.

**Параметрические уравнения прямой.** Заметив, что $\mathbf{r}=\overline{OM}=(x, y, z)$, $\mathbf{r}_0=\overline{OM_0}=(x_0, y_0, z_0)$, $\mathbf{a}=(a_1, a_2, a_3)$, от векторного соотношения (6.39) перейдем к координатным соотношениям

$$x=x_0+a_1t, \quad y=y_0+a_2t, \quad z=z_0+a_3t. \tag{6.40}$$

Уравнения (6.40) называются *параметрическими уравнениями прямой* (см. § 6.3).

**Канонические уравнения прямой.** Каждое из уравнений (6.40) разрешим относительно $t$:

$$\frac{x-x_0}{a_1}=t, \quad \frac{y-y_0}{a_2}=t, \quad \frac{z-z_0}{a_3}=t,$$

откуда

$$\frac{x-x_0}{a_1}=\frac{y-y_0}{a_2}=\frac{z-z_0}{a_3}. \tag{6.41}$$

Уравнения (6.41) называются *каноническими уравнениями* прямой.

**Уравнение прямой, проходящей через две данные точки.** Пусть даны две различные точки $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$. В качестве направляющего вектора прямой можно взять вектор $\mathbf{a}=\overline{M_1M_2}=(x_2-x_1, y_2-y_1, z_2-z_1)$. Поскольку прямая проходит через точку $M_1(x_1, y_1, z_1)$, ее канонические уравнения в соответствии с (6.41) запишутся так:

$$\frac{x-x_1}{x_2-x_1}=\frac{y-y_1}{y_2-y_1}=\frac{z-z_1}{z_2-z_1}. \tag{6.42}$$

Уравнения (6.42) являются уравнениями прямой, проходящей через две указанные точки.

**Угол между двумя прямыми.** Рассмотрим две прямые, заданные параметрическими уравнениями

$$x=x_1+a_1t, \quad y=y_1+a_2t, \quad z=z_1+a_3t; \tag{6.43}$$

$$x=x_2+b_1t, \quad y=y_2+b_2t, \quad z=z_2+b_3t; \tag{6.44}$$

первая из этих прямых проходит через точку $M_1(x_1, y_1, z_1)$, вторая — через точку $M_2(x_2, y_2, z_2)$, прямые имеют соответственно направляющие векторы

$$\mathbf{a}=(a_1, a_2, a_3), \quad \mathbf{b}=(b_1, b_2, b_3). \tag{6.45}$$

Поскольку угол между прямыми равен углу между их направляющими векторами, то

$$\cos\varphi = \frac{\mathbf{ab}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}}, \quad (6.46)$$

Очевидно, условие перпендикулярности прямых (6.43) и (6.44) выражается равенством

$$a_1b_1 + a_2b_2 + a_3b_3 = 0. \quad (6.47)$$

Принимая во внимание формулы (5.37), (5.38), заключаем, что необходимое и достаточное условие параллельности прямых (6.43) и (6.44) выражается равенствами

$$b_1 = \lambda a_1, \quad b_2 = \lambda a_2, \quad b_3 = \lambda a_3, \quad (6.48)$$

или

$$\frac{b_1}{a_1} = \frac{b_2}{a_2} = \frac{b_3}{a_3}. \quad (6.49)$$

**Взаимное расположение прямых в пространстве.** Исследуем вопрос о взаимном расположении двух прямых в пространстве, заданных уравнениями (6.43) и (6.44). Кроме векторов (6.45) рассмотрим еще вектор $\overline{M_1M_2} = (x_2 - x_1,\ y_2 - y_1,\ z_2 - z_1)$.

Очевидно, данные прямые являются скрещивающимися тогда и только тогда, когда векторы $\mathbf{a}$, $\mathbf{b}$ и $\overline{M_1M_2}$ некомпланарны, в этом случае их смешанное произведение отлично от нуля, т. е. $[\mathbf{a}, \mathbf{b}]\,\overline{M_1M_2} \neq 0$, или

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{vmatrix} \neq 0. \quad (6.50)$$

Данные прямые лежат в одной плоскости тогда и только тогда, когда векторы $\mathbf{a}$, $\mathbf{b}$, $\overline{M_1M_2}$ компланарны, а в этом случае $[\mathbf{a}, \mathbf{b}]\,\overline{M_1M_2} = 0$, или

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{vmatrix} = 0. \quad (6.51)$$

Прямые будут пересекаться, если условие (6.49) не выполнено (т. е. если первые две строки определителя в формуле (6.51) не пропорциональны); они будут параллельными, если условие (6.49) выполняется (т. е. первые две строки указанного определителя пропорциональны), прямые совпадут, если все три вектора $\mathbf{a}$, $\mathbf{b}$, $\overline{M_1M_2}$ коллинеарны (все строки этого определителя пропорциональны).

Пример 1. Составить параметрические уравнения прямой, проходящей через точку $M_0(2, -3, -7)$ и имеющей направляющий вектор $\mathbf{a}=(4, -6, 5)$.

Поскольку по условию $x_0=2$, $y_0=-3$, $z_0=-7$, $a_1=4$, $a_2=-6$, $a_3=5$, в соответствии с формулами (6.40) получаем искомые уравнения $x=2+4t$, $y=-3-6t$, $z=-7+5t$.

Пример 2. Исследовать взаимное расположение двух прямых, заданных уравнениями:

$$x = 1 - 9t, \qquad y = 2 + 8t, \qquad z = 3 - 7t;$$
$$x = 6 - 2t, \qquad y = 5 + 3t, \qquad z = 4 + t.$$

Так как

$$\begin{vmatrix} -9 & 8 & -7 \\ -2 & 3 & 1 \\ 6-1 & 5-2 & 4-3 \end{vmatrix} = \begin{vmatrix} -9 & 8 & -7 \\ -2 & 3 & 1 \\ 5 & 3 & 1 \end{vmatrix} =$$
$$= \begin{vmatrix} -9 & 8 & -7 \\ -7 & 0 & 0 \\ 5 & 3 & 1 \end{vmatrix} = 7 \begin{vmatrix} 8 & -7 \\ 3 & 1 \end{vmatrix} = 7(8+21) \neq 0,$$

то прямые скрещиваются.

## § 6.6. Задачи на прямую и плоскость в пространстве

**Прямая как пересечение двух плоскостей.** Рассмотрим две плоскости, заданные общими уравнениями (6.33) и (6.34). Предположим, что условие (6.35) для этих плоскостей не выполнено. Пусть

$$\frac{A_1}{A_2} \neq \frac{B_1}{B_2}, \text{ т. е. } \begin{vmatrix} A_1 B_1 \\ A_2 B_2 \end{vmatrix} \neq 0.$$

В этом случае плоскости пересекаются по прямой, определяемой уравнениями

$$\left.\begin{aligned} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{aligned}\right\} \qquad (6.52)$$

Уравнения (6.52) приведем к параметрическому виду, для чего нужно выбрать точку этой прямой и направляющий вектор. В качестве последнего можно взять вектор $\mathbf{a}=[\mathbf{n}_1, \mathbf{n}_2]$, где $\mathbf{n}_1$ и $\mathbf{n}_2$ — нормальные векторы рассматриваемых плоскостей. Так как $\mathbf{n}_1=(A_1, B_1, C_1)$, $\mathbf{n}_2=(A_2, B_2, C_2)$ (см. формулу (6.19) § 6.4), то (см. § 5.9)

$$\mathbf{a} = \left\{ \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}, \; -\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \; \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \right\}.$$

Чтобы выбрать точку, фиксируем произвольное значение $z=z_0$ в уравнениях (6.52), получим систему двух уравнений

$$\left.\begin{aligned} A_1x+B_1y&=-C_1z_0-D_1;\\ A_2x+B_2y&=-C_2z_0-D_2.\end{aligned}\right\}$$

Решая эту систему (определитель ее отличен от нуля), получаем значения $x=x_0$ и $y=y_0$. Итак, фиксирована точка $M(x_0, y_0, z_0)$.

Искомые параметрические уравнения принимают вид

$$x=x_0+\begin{vmatrix} B_1 & C_1\\ B_2 & C_2\end{vmatrix}t,\qquad y=y_0-\begin{vmatrix} A_1 & C_1\\ A_2 & C_2\end{vmatrix}t,$$
$$z=z_0+\begin{vmatrix} A_1 & B_1\\ A_2 & B_2\end{vmatrix}t. \tag{6.53}$$

Пример 1. Составить параметрические уравнения прямой

$$\left.\begin{aligned} 3x-4y+z-2&=0;\\ 7x-5y+5z-9&=0.\end{aligned}\right\}$$

Положив, например, $z_0=0$, получим систему уравнений

$$3x-4y=2;$$
$$7x-5y=9,$$

из которой находим $x_0=2$, $y_0=1$; фиксирована точка $M_0(2, 1, 0)$ этой прямой.

Поскольку $\mathbf{n}_1=(3, -4, 1)$, $\mathbf{n}_2=(7, -5, 5)$, то

$$\mathbf{a}=[\mathbf{n}_1, \mathbf{n}_2]=\left(\begin{vmatrix} -4 & 1\\ -5 & 5\end{vmatrix}, -\begin{vmatrix} 3 & 1\\ 7 & 5\end{vmatrix}, \begin{vmatrix} 3 & -4\\ 7 & -5\end{vmatrix}\right)=\{-15, -8, 13\}.$$

Следовательно, искомые уравнения имеют вид $x=2-15t$, $y=1-8t$, $z=13t$.

**Взаимное расположение прямой и плоскости.** Рассмотрим прямую, заданную параметрическими уравнениями (6.40), и плоскость, определяемую общим уравнением (6.22).

Чтобы найти общие точки прямой и плоскости, необходимо решить систему их уравнений. Подставляя выражения (6.40) в уравнение (6.22), получаем

$$A(x_0+a_1t)+B(y_0+a_2t)+C(z_0+a_3t)+D=0$$

или

$$(Aa_1+Ba_2+Ca_3)t+(Ax_0+By_0+Cz_0+D)=0. \tag{6.54}$$

Исследуем возможные случаи, касающиеся коэффициентов этого уравнения. Если

$$Aa_1+Ba_2+Ca_3\neq 0, \qquad (6.55)$$

то уравнение (6.54) имеет единственное решение

$$t=-\frac{Ax_0+By_0+Cz_0+D}{Aa_1+Ba_2+Ca_3}.$$

Прямая и плоскость пересекаются в единственной точке, координаты которой находятся подстановкой значения $t$ в уравнения (6.40). Если

$$Aa_1+Ba_2+Ca_3=0, \quad Ax_0+By_0+Cz_0+D\neq 0, \qquad (6.56)$$

уравнению (6.54) не удовлетворяет ни одно значение $t$; прямая и плоскость общих точек не имеют. Если

$$Aa_1+Ba_2+Ca_3=0, \quad Ax_0+By_0+Cz_0+D=0, \qquad (6.57)$$

уравнению (6.54) удовлетворяют любые значения $t$; прямая целиком лежит в плоскости.

Замечание. Первое из равенств в формулах (6.56), (6.57), т. е.

$$Aa_1+Ba_2+Ca_3=0, \qquad (6.58)$$

выражает необходимое и достаточное условие параллельности прямой и плоскости; это условие можно получить непосредственно, так как векторы $\mathbf{n}=(A, B, C)$ и $\mathbf{a}=(a_1, a_2, a_3)$ в этом случае перпендикулярны и их скалярное произведение равно нулю.

Пример 2. Исследовать взаимное расположение плоскости $x+y+z-3=0$ и прямой $x=1-t$, $y=2+3t$, $z=6-4t$.

Поскольку условие (6. 58) не выполнено, то прямая и плоскость пересекаются. Подставив выражения для $x$, $y$, $z$ через $t$ в уравнение плоскости, получим $(1-t)+(2+3t)+(6-4t)-3=0$, $-2t+6=0$, $t=3$. Из уравнений прямой находим координаты точки пересечения $x=1-3=-2$, $y=2+9=11$, $z=6-12=-6$, $M(-2, 11, -6)$.

**Угол между прямой и плоскостью.** Найдем угол $\varphi$ между прямой $x=x_0+a_1t$, $y=y_0+a_2t$, $z=z_0+a_3t$ и плоскостью $Ax+By+Cz+D=0$.

Поскольку вектор $\mathbf{n}=(A, B, C)$ перпендикулярен данной плоскости, направляющий вектор $\mathbf{a}=(a_1, a_2, a_3)$ прямой образует с вектором $\mathbf{n}$ угол $\psi=\frac{\pi}{2}-\varphi$ или $\psi=\frac{\pi}{2}+\varphi$ (рис. 6.9). Так как под $\varphi$ понимаем острый положительный угол, то $\sin\varphi=|\cos\psi|$.

По формуле (5.62) находим

$$\sin\varphi=\frac{|Aa_1+Ba_2+Ca_3|}{\sqrt{A^2+B^2+C^2}\sqrt{a_1^2+a_2^2+a_3^2}}. \qquad (6.59)$$

**Расстояние точки до плоскости.** Пусть дана плоскость нормальным уравнением (6.27) и произвольная точка $M_0(x_0, y_0, z_0)$ пространства. *Отклонением* $\delta$ точки $M_0$ от плоскости назовем ее расстояние $d$ до этой плоскости,

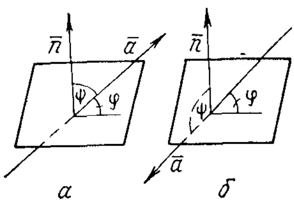



Рис. 6.9

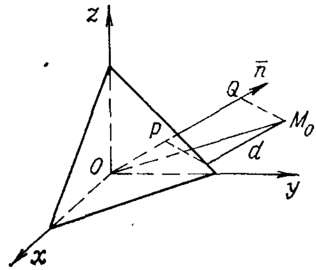



Рис. 6.10

взятое со знаком плюс, если $M_0$ и начало координат лежат по разные стороны от плоскости (рис. 6.10), и со знаком минус, если они лежат по одну сторону от плоскости, т. е.

$$\delta=\pm d,\ d=|\delta|. \qquad (6.60)$$

Спроектируем точку $M_0$ на нормаль $n$ к плоскости. Пусть $Q$ — ее проекция, тогда $\delta=PQ=OQ-OP$, где $PQ$, $OQ$ и $OP$ — величины соответствующих направленных отрезков нормали.

Поскольку $OQ=\text{пр}_n\ \overline{OM_0}$, $OP=p$ и $\text{пр}_n\ \overline{OM_0}=x_0\cos\alpha+y_0\cos\beta+z_0\cos\gamma$ (см. следствие 3 из теоремы 5.1), то

$$\delta=x_0\cos\alpha+y_0\cos\beta+z_0\cos\gamma-p. \qquad (6.61)$$

Следовательно, расстояние точки $M_0(x_0, y_0, z_0)$ до плоскости (6.27) определяется формулой

$$d=|x_0\cos\alpha+y_0\cos\beta+z_0\cos\gamma-p|. \qquad (6.62)$$

Замечание. Если плоскость задана общим уравнением (6.22), то с учетом ее нормального уравнения (6.31) получаем

$$d=\frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}. \qquad (6.63)$$

Пример 3. Вычислить расстояние от точки $M_0(7, -8, 3)$ до плоскости $x+2y-2z-12=0$.

По формуле (6.63) получаем

$$d=\frac{|1\cdot 7+(-8)2-2\cdot 3-12|}{\sqrt{1^2+2^2+(-2)^2}}=\frac{|-27|}{3}=9.$$

## § 6.7. Уравнение цилиндрической поверхности с образующей, параллельной координатной оси. Цилиндры второго порядка

*Цилиндрической поверхностью* называется поверхность, описываемая прямой (*образующей*), движущейся вдоль некоторой линии (*направляющей*) и остающейся параллельной исходному направлению (рис. 6.11).

Рассмотрим уравнение

$$F(x, y)=0, \qquad (6.64)$$

особенность которого состоит в том, что оно не содержит переменной $z$.

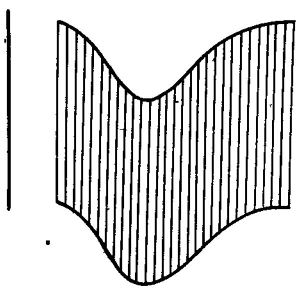


Рис. 6.11

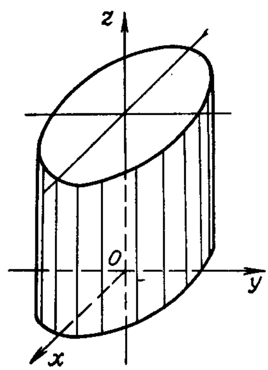



Рис. 6.12

Покажем, что если это уравнение определяет некоторую поверхность, то она является цилиндрической поверхностью с образующей, параллельной оси $Oz$. Если $x_0$, $y_0$ — значения переменных $x$ и $y$, удовлетворяющие уравнению (6.64), т. е. точка $M_0(x_0, y_0, z_0)$ принадлежит поверхности, определяемой уравнением, то ей будет принадлежать и точка $M(x_0, y_0, z)$, так как $z$ в уравнение (6.64) не входит. Следовательно, поверхности целиком принадлежит прямая, проходящая через точки $M_0$, $M$, а эта прямая параллельна оси $Oz$. Итак, поверхность, определяемая уравнением (6.64), является цилиндрической поверхностью с образующей, параллельной оси $Oz$.

Отметим, что в плоскости $Oxy$ уравнение (6.64) определяет линию (направляющую), которая в пространстве задается уже двумя уравнениями

$$\left.\begin{aligned} F(x, y) = 0; \\ z = 0 \end{aligned}\right\} \tag{6.65}$$

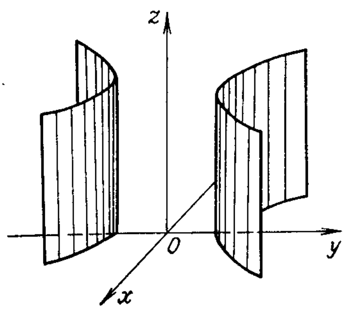



Рис. 6.13

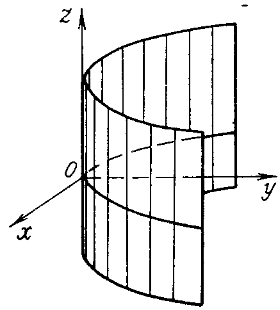



Рис. 6.14

(уравнения (6.65) задают эту линию как пересечение цилиндрической поверхности (6.64) и координатной плоскости $Oxy$).

*Цилиндром второго порядка* называется цилиндрическая поверхность, направляющей которой является эллипс (окружность), гипербола или парабола.

К цилиндрам второго порядка относятся *эллиптический цилиндр* (рис. 6.12), определяемый уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{6.66}$$

в некоторой системе декартовых прямоугольных координат; *гиперболический цилиндр* (рис. 6.13), определяемый уравнением

$$-\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1; \tag{6.67}$$

*параболический цилиндр* (рис. 6.14), определяемый уравнением

$$x^2 = 2qy \tag{6.68}$$

или уравнением

$$y^2 = 2px. \tag{6.69}$$

Замечание 1. Если уравнение

$$F(x, z)=0 \tag{6.70}$$

определяет некоторую поверхность, то ею является цилиндрическая поверхность с образующей, параллельной оси $Oy$.

Замечание 2. Если уравнение

$$F(y, z)=0 \tag{6.71}$$

определяет некоторую поверхность, то ею является цилиндрическая поверхность с образующей, параллельной оси $Ox$.

## § 6.8. Уравнение поверхности вращения

Рассмотрим поверхность, полученную вращением линии $l$, заданной уравнениями

$$x=\varphi_1(z), \quad y=\varphi_2(z)^{1)}, \tag{6.72}$$

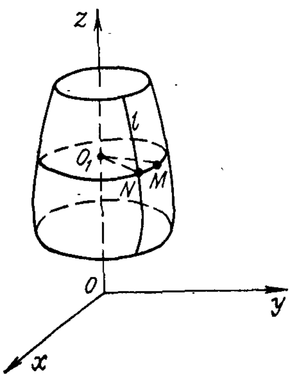



Рис. 6.15

вокруг оси $Oz$. Пусть $M$ — произвольная точка этой поверхности, $X$, $Y$, $Z$ — ее текущие координаты. Через точку $M$ проведем плоскость, перпендикулярную оси $Oz$; эта плоскость (ее уравнение $z=Z$) пересечет поверхность по некоторой окружности с центром в точке $O_1(0, 0, z)$ на оси $Oz$ (рис. 6.15). Обозначим буквой $N$ точку пересечения указанной окружности и линии $l$, а буквами $x$, $y$, $z$ — ее координаты.

Поскольку длины отрезков $O_1M$ и $O_1N$ равны между собой (как радиусы одной и той же окружности), т. е. $\rho(O_1, M)=\rho(O_1, N)$ и

$$\rho(O_1, M)=\sqrt{X^2+Y^2}, \quad \rho(O_1, N)=\sqrt{x^2+y^2},$$

то

$$\sqrt{X^2+Y^2}=\sqrt{x^2+y^2}, \quad X^2+Y^2=x^2+y^2. \tag{6.73}$$

Для координат точки $N$ $x=\varphi_1(z)$, $y=\varphi_2(z)$, $z=Z$, поэтому уравнение (6.73) принимает вид

$$X^2+Y^2=\varphi_1^2(Z)+\varphi_2^2(Z). \tag{6.74}$$

---

1) Здесь предполагается, что уравнения (6.5) линии можно представить в таком виде, т. е. выразить $x$ и $y$ только через $z$.

Уравнение (6.74) является уравнением поверхности, полученной вращением линии (6.72) вокруг оси $Oz$.

Если обозначить текущие координаты точки поверхности буквами $x$, $y$, $z$, уравнение (6.74) можно переписать так:

$$x^2 + y^2 = \varphi_1^2(z) + \varphi_2^2(z). \tag{6.75}$$

Обратно, если задано уравнение

$$x^2 + y^2 = f(z), \tag{6.76}$$

где $f(z) \geqslant 0$, то оно определяет поверхность вращения, так как в сечении ее с плоскостью $z=h$ будет окружность

$$x^2 + y^2 = f(h), \quad z = h. \tag{6.77}$$

Замечание. Уравнение (6.75) поверхиости вращения получается из уравнений линии $l$ в результате следующих формальиых действий: возведеиия в квадрат обеих частей каждого из уравнений (6.72) и почленного их сложения.

## § 6.9. Поверхности вращения второго порядка

*Поверхностью вращения второго порядка* называется поверхность, образованная вращением линии второго порядка вокруг ее оси.

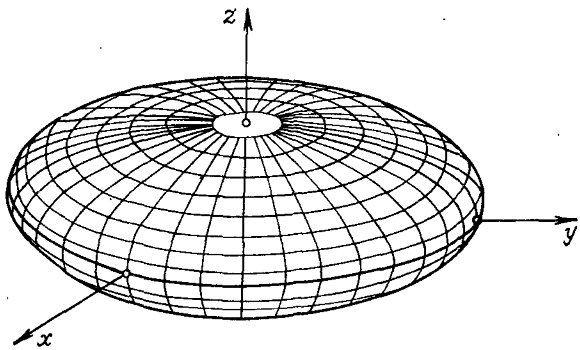



Рис. 6.16

*Эллипсоидом вращения* называется поверхность, полученная вращением эллипса вокруг одной из его осей (рис. 6.16).

Пусть в плоскости $Oyz$ задан эллипс уравнениями

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad x = 0^{1)}. \tag{6.78}$$

Составим уравнение эллипсоида вращения, полученного вращением эллипса (6.78) вокруг оси $Oz$. Уравнения (6.78) приведем к виду (6.72). Поскольку в уравнение поверхности вращения входят квадраты $x$ и $y$, достаточно выразить их из уравнений (6.78):

$$y^2 = b^2\left(1 - \frac{z^2}{c^2}\right), \quad x^2 = 0.$$

На основании (6.75) получаем $x^2 + y^2 = b^2\left(1 - \frac{z^2}{c^2}\right)$, или

$$\frac{x^2}{b^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{6.79}$$

Уравнение (6.79) является уравнением эллипсоида вращения, полученного вращением эллипса (6.78) вокруг оси $Oz$.

*Однополостным гиперболоидом* вращения называется поверхность, полученная вращением гиперболы вокруг ее мнимой оси (рис. 6.17).

Составим уравнение однополостного гиперболоида вращения, полученного вращением гиперболы

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad x = 0 \tag{6.80}$$

вокруг оси $Oz$:

$$y^2 = b^2\left(1 + \frac{z^2}{c^2}\right), \quad x^2 = 0, \quad x^2 + y^2 = b^2\left(1 + \frac{z^2}{c^2}\right),$$

откуда

$$\frac{x^2}{b^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \tag{6.81}$$

Уравнение (6.81) является уравнением указанного однополостного гиперболоида вращения.

---

1) Линия в пространстве определяется двумя уравнениями относительно декартовых координат (см. § 6.1, систему (6.5)); в данном случае первое уравнение определяет эллиптический цилиндр с осью $Ox$ (см. уравнения (6.66) и (6.71)), второе — координатную плоскость $Oyz$; пересечение указанных поверхностей является эллипсом с полуосями $b$ и $c$.

*Двуполостным гиперболоидом вращения* называется поверхность, полученная вращением гиперболы вокруг ее действительной оси (рис. 6.18).

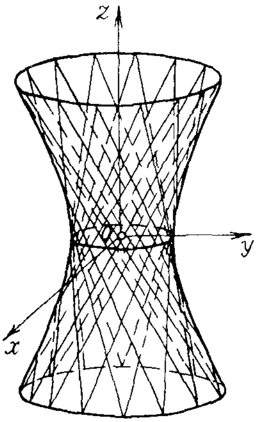



Рис. 6.17

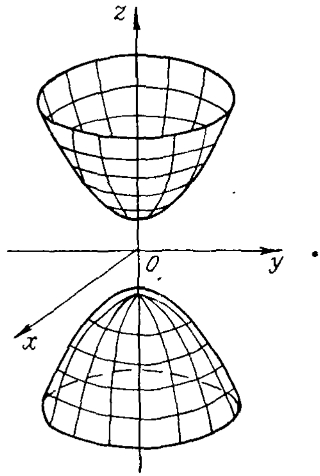



Рис. 6.18

Составим уравнение двуполостного гиперболоида, полученного вращением гиперболы

$$\frac{z^2}{c^2} - \frac{y^2}{b^2} = 1, \quad x = 0 \tag{6.82}$$

вокруг оси $Oz$:

$$y^2 = b^2\left(\frac{z^2}{c^2} - 1\right), \; x^2 = 0, \; y^2 + x^2 = b^2\left(\frac{z^2}{c^2} - 1\right).$$

Отсюда следует искомое уравнение

$$\frac{x^2}{b^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \tag{6.83}$$

*Конусом вращения* называется поверхность, полученная вращением двух пересекающихся прямых вокруг биссектрисы угла между ними.

Составим уравнение конуса, полученного вращением прямых

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 0, \quad x = 0$$

вокруг оси $Oz$ (рис. 6.19). Поскольку

$$y^2 = \frac{b^2}{c^2} z^2, \quad x^2 = 0,$$

то

$$x^2 + y^2 = \frac{b^2}{c^2} z^2, \text{ или } \frac{x^2}{b^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0. \quad (6.84)$$

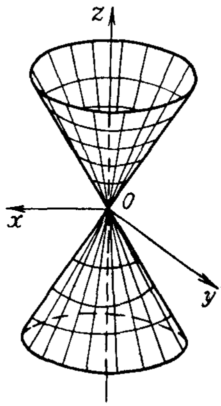



Рис. 6.19

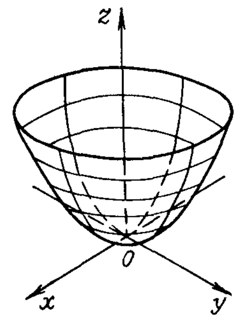



Рис. 6.20

Уравнение (6.84) является уравнением указанного конуса вращения.

*Параболоидом вращения* называется поверхность, полученная вращением параболы вокруг ее оси (рис. 6.20).

Уравнение параболоида вращения, полученного вращением параболы

$$y^2 = 2pz, \quad x = 0 \quad (6.85)$$

вокруг оси $Oz$, имеет вид

$$x^2 + y^2 = 2pz, \quad \text{или} \quad \frac{x^2}{p} + \frac{y^2}{p} = 2z. \quad (6.86)$$

## § 6.10. Поверхности второго порядка, заданные каноническими уравнениями

Введем сначала понятие одного преобразования точек пространства, которым будем пользоваться в дальнейшем. «Сжатием к плоскости» П называется такое преобразование, которое каждой точке $M$ пространства ставит в соответствие точку $M'$, лежащую на перпендикуляре $MP$ к плоскости П (рис. 6.21), такую, что

$$\frac{MP}{M'P} = k, \tag{6.87}$$

где $MP$, $M'P$ — величины направленных отрезков перпендикуляра, опущенного на плоскость П ($P$ — основание перпендикуляра), число $k$ называется коэффициентом «сжатия», $k > 0$. Отметим, что при $k > 1$ будет сжатие, при $k < 1$ — растяжение, при $k = 1$ — тождественное преобразование (каждая точка остается на месте).

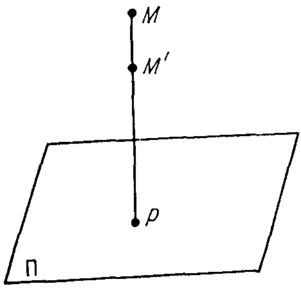



Рис. 6.21

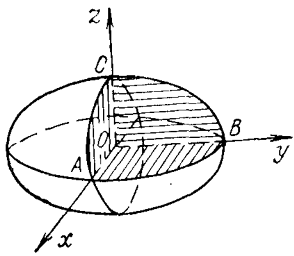



Рис. 6.22

Если плоскость П — одна из координатных плоскостей, то формулы преобразования имеют простой вид, так как две координаты остаются прежними, меняется лишь третья из них. Например, если П — плоскость $Oyz$, то формулы преобразования имеют вид

$$x = kX, \quad y = Y, \quad z = Z, \tag{6.88}$$

где $(x, y, z)$ — координаты точки $M$; $(X, Y, Z)$ — координаты точки $M'$.

Эллипсоидом называется поверхность, полученная в результате сжатия эллипсоида вращения к плоскости, содержащей ось вращения.

Эллипсоид вращения, определяемый уравнением (6.79), подвергнем преобразованию «сжатия» к плоскости $Oyz$ по формулам

$$x = \frac{b}{a} X, \quad y = Y, \quad z = Z. \tag{6.89}$$

Подставив формулы (6.89) в уравнение (6.79), получим уравнение эллипсоида

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}+\frac{Z^2}{c^2}=1. \tag{6.90}$$

Если текущие координаты точки эллипсоида обозначить $x$, $y$, $z$, то уравнение примет вид

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1. \tag{6.90a}$$

Эллипсоид отсекает на осях координат $Ox$, $Oy$, $Oz$ отрезки $2a$, $2b$, $2c$, которые называются осями, $a=OA$, $b=OB$, $c=OC$ — полуоси эллипсоида (рис. 6.22). Если $a\neq b$, $b\neq c$, $a\neq c$, то поверхность называется *трехосным эллипсоидом.*

Замечание. В случае равенства двух полуосей получаем эллипсоид вращения, если все полуоси равны $a=b=c=R$, имеем сферу радиуса $R$, уравнение которой $x^2+y^2+z^2=R^2$. Это уравнение было выведено ранее другим путем (см. § 6.1 уравнение (6.4)).

Подвергая преобразованию «сжатия» к плоскости $Oyz$ по формулам (6.89) поверхности (6.81), (6.83), (6.84), получаем соответственно *однополостный гиперболоид,* определяемый уравнением

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1, \tag{6.91}$$

*двуполостный гиперболоид,* имеющий уравнение

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1, \tag{6.92}$$

*конус второго порядка,* определяемый уравнением

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0. \tag{6.93}$$

Подвергнув преобразованию «сжатия» к плоскости $Oxz$ поверхность параболоида вращения (6.86) по формулам

$$x=X,\ y=\sqrt{\frac{p}{q}}\,Y,\ \ z=Z\,(pq>0), \tag{6.94}$$

получим поверхность, называемую *эллиптическим параболоидом.*

Эллиптический параболоид имеет уравнение

$$\frac{x^2}{p}+\frac{y^2}{q}=2z. \tag{6.95}$$

Поверхность эллиптического параболоида можно получить и другим способом.

Рассмотрим две параболы

$$\left.\begin{aligned} y^2 &= 2qz; \\ x &= 0 \end{aligned}\right\} \tag{6.96}$$

и

$$\left.\begin{aligned} x^2 &= 2pz; \\ y &= 0, \end{aligned}\right\} \tag{6.97}$$

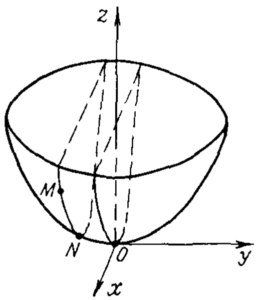



Рис. 6.23

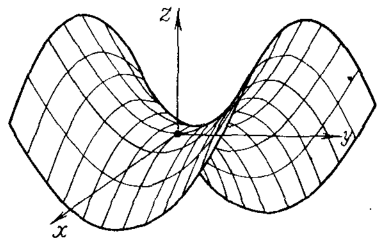



Рис. 6.24

имеющие общую вершину, общую ось и расположенные в перпендикулярных плоскостях $Oyz$ и $Oxz$ (рис. 6.23). Найдем поверхность, описанную параболой (6.97) при таком непрерывном перемещении ее, когда ось не меняет направления, а вершина все время находится на параболе (6.96), причем плоскости, в которых лежат подвижная и неподвижная параболы, остаются взаимно перпендикулярными. Возьмем произвольную точку $M(X, Y, Z)$ искомой поверхности и проведем через нее плоскость, перпендикулярную оси $Oy$ (рис. 6.23). Эта плоскость имеет уравнение $y=Y$, где $Y$ — ордината точки $M$.

При своем движении парабола (6.97) в некоторый момент времени займет положение, при котором ее плоскость совпадет с проведенной плоскостью $y=Y$. Пусть $N$ — вершина параболы, тогда $N(0, Y, z)$, так как эта точка лежит в плоскости $Oyz$ и в плоскости $y=Y$. Уравнение параболы с вершиной в точке $N(0, Y, z)$ имеет вид

$$\left.\begin{aligned} &(X-0)^2 = 2p(Z-z) \\ &y = Y \end{aligned}\right\}, \quad \text{или} \quad \left.\begin{aligned} &X^2 = 2p(Z-z) \\ &y = Y \end{aligned}\right\}.$$

Поскольку точка $N$ лежит на параболе (6.96), то $Y^2=2qz$, откуда $z=\frac{Y^2}{2q}$.

Следовательно, уравнение искомой поверхности имеет вид

$$X^2=2p\left(Z-\frac{Y^2}{2q}\right)$$

или

$$\frac{X^2}{p}+\frac{Y^2}{q}=2Z. \tag{6.98}$$

Возможны два случая: 1) числа $p$ и $q$ одного знака ($pq>0$), т. е. направления осей парабол (6.96) и (6.97) одинаковы, например, $p=a^2$, $q=b^2$, тогда искомая поверхность имеет уравнение

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=2z, \tag{6.99}$$

следовательно, является поверхностью эллиптического параболоида; 2) числа $p$ и $q$ имеют разные знаки ($pq<0$), т. е. направления осей данных парабол противоположны, например, $p=-a^2$, $q=b^2$, тогда уравнение поверхности примет вид

$$-\frac{x^2}{a^2}+\frac{y^2}{b^2}=2z. \tag{6.100}$$

Эта поверхность называется *гиперболическим параболоидом* (рис 6.24).

Замечание. При рассмотрении поверхиостей второго порядка мы исходили из наглядного геометрического представления об их образовании, затем составили уравнения. Можно поступать по-другому. Задать поверхности уравнениями, рассмотреть линии, получающиеся в сечении их координатными плоскостями, и по этим линиям составить представление о виде поверхности.

# Глава 7

## ЛИНЕЙНЫЕ ПРОСТРАНСТВА. ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ

Понятие линейного пространства (многомерного векторного пространства) является одним из основных в современной математике. Это понятие возникает в результате обобщения некоторых свойств векторов.

## § 7.1. Понятие линейного пространства. Подпространство

Рассмотрим множества $V$ элементов $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ и множество $R$ действительных чисел. Определим операцию сложения элементов множества $V$ (ее называют внутренней операцией): любой упорядоченной паре элементов $\mathbf{x} \in V$, $\mathbf{y} \in V$ поставим в соответствие третий элемент $\mathbf{z} \in V$, называемый их *суммой*; будем писать в этом случае $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Введем также операцию умножения элементов множества $V$ на действительные числа (эту операцию называют внешней); каждому элементу $\mathbf{x} \in V$ и действительному числу $\alpha \in R$ поставим в соответствие элемент $z = \alpha\mathbf{x} = \mathbf{x}\alpha$, где $\mathbf{z} \in V$. Потребуем, чтобы операция сложения элементов множества $V$ и операция умножения элементов этого множества на действительные числа удовлетворяли следующим аксиомам:

I. Сложение коммутативно, т. е. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ для любых $\mathbf{x} \in V$, $\mathbf{y} \in V$.

II. Сложение ассоциативно, т. е. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ для любых $\mathbf{x} \in V$, $\mathbf{y} \in V$, $\mathbf{z} \in V$.

III. Существует нулевой элемент, т. е. такой элемент, который в сумме с любым элементом $\mathbf{x}$ дает тот же элемент $\mathbf{x}$; обозначим этот элемент символом $\mathbf{0}$, тогда $\mathbf{x} + \mathbf{0} = \mathbf{x}$ для любого $\mathbf{x} \in V$.

IV. Для каждого элемента $\mathbf{x} \in V$ существует противоположный элемент, т. е. такой элемент, который в сумме с данным дает нулевой элемент; элемент, противоположный элементу $\mathbf{x}$, обозначим через $-\mathbf{x}$, тогда $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ для любого $\mathbf{x} \in V$.

V. Умножение на число 1 не меняет элемента, т. е. $1 \cdot \mathbf{x} = \mathbf{x}$ для любого $\mathbf{x} \in V$.

Для любых $\mathbf{x}, \mathbf{y} \in V$, $\alpha, \beta \in R$:

VI. $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$.

VII. $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.

VIII. $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.

Множество $V$ элементов $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, ..., в котором определены операции сложения элементов и умножения элемента на действительное число, удовлетворяющие аксиомам I—VIII, называется *действительным линейным пространством*, или *действительным векторным пространством*. Элементы действительного линейного пространства называют *векторами*.

Замечание. Аналогично определяется комплексное линейное пространство: вместо множества $R$ действительных чисел рассматривается множество $C$ комплексных чисел. В дальнейшем будем изучать преимущественно действительное линейное пространство и называть его просто линейным пространством.

Из определения линейного пространства вытекают следующие утверждения.

1. В линейном пространстве имеется единственный нулевой элемент.

2. Для любого элемента $\mathbf{x}$ линейного пространства существует единственный противоположный элемент $-\mathbf{x}$.

3. Для элемента $-\mathbf{x}$ противоположным будет элемент $\mathbf{x}$.

4. Для любого элемента $\mathbf{x}$ произведение $0\mathbf{x}=\mathbf{0}$, где $0$ — число нуль, $\mathbf{0}$ — нулевой элемент.

5. Для любого элемента $\mathbf{x}$ $-1\cdot\mathbf{x}=-\mathbf{x}$, где $(-\mathbf{x})$ — элемент, противоположный элементу $\mathbf{x}$.

6. Для любого числа $\alpha$ произведение $\alpha\mathbf{0}=\mathbf{0}$, где $\mathbf{0}$ — нулевой элемент.

7. Если $\alpha\mathbf{x}=\mathbf{0}$ и $\alpha\neq 0$, то $\mathbf{x}=\mathbf{0}$.

8. Если $\alpha\mathbf{x}=\mathbf{0}$ и $\mathbf{x}\neq\mathbf{0}$, то $\alpha=0$.

Докажем эти утверждения. 1. Предположим, что в $V$ имеются два нулевых элемента $\mathbf{0}_1$ и $\mathbf{0}_2$, тогда $\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_1$ и $\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_2$, поэтому $\mathbf{0}_1=\mathbf{0}_2$. 2. Предположим, что для элемента $\mathbf{x}$ существует два противоположных элемента $\mathbf{x}_1$ и $\mathbf{x}_2$, т. е. $\mathbf{x}+\mathbf{x}_1=\mathbf{0}$ и $\mathbf{x}+\mathbf{x}_2=\mathbf{0}$, тогда $\mathbf{x}_1=(\mathbf{x}_2+\mathbf{x})+\mathbf{x}_1=$ $=\mathbf{x}_2+(\mathbf{x}+\mathbf{x}_1)=\mathbf{x}_2$, или $\mathbf{x}_1=\mathbf{x}_2$. 3. Поскольку $-\mathbf{x}+\mathbf{x}=$ $=\mathbf{x}+(-\mathbf{x})$ (по аксиоме I) и $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$ (по аксиоме III), то $-\mathbf{x}+\mathbf{x}=\mathbf{0}$, а это означает, что $\mathbf{x}$ — элемент, противоположный элементу $-\mathbf{x}$. 4. Так как $0\mathbf{x}+\mathbf{0}=0\mathbf{x}+(\mathbf{x}+(-\mathbf{x}))=$ $=(0\mathbf{x}+\mathbf{x})+(-\mathbf{x})=(0+1)\mathbf{x}+(-\mathbf{x})=\mathbf{x}+(-\mathbf{x})=\mathbf{0}$, т. е. $0\mathbf{x}+\mathbf{0}=\mathbf{0}$, то $0\mathbf{x}=\mathbf{0}$. 5. Поскольку $-1\cdot\mathbf{x}+\mathbf{x}=(-1+1)\mathbf{x}=$ $=0\mathbf{x}=\mathbf{0}$, или $-1\cdot\mathbf{x}+\mathbf{x}=\mathbf{0}$, то $-1\cdot\mathbf{x}$ — элемент, противоположный элементу $\mathbf{x}$, т. е. $(-1)\mathbf{x}=-\mathbf{x}$. 6. $\alpha\mathbf{0}=$ $=\alpha(\mathbf{x}+(-\mathbf{x}))=\alpha(\mathbf{x}+(-1)\mathbf{x})=\alpha\mathbf{x}+\alpha(-1)\mathbf{x}=\alpha\mathbf{x}+$ $+(-\alpha\mathbf{x})=\mathbf{0}$, $\alpha\mathbf{0}=\mathbf{0}$. 7. Пусть $\alpha\mathbf{x}=\mathbf{0}$ и $\alpha\neq 0$, тогда $\frac{1}{\alpha}\cdot\alpha\mathbf{x}=\mathbf{0}$, или $\mathbf{x}=\mathbf{0}$. 8. Пусть $\alpha\mathbf{x}=\mathbf{0}$ и $\mathbf{x}\neq\mathbf{0}$, тогда $\alpha=0$. Действительно, предположив противное, т. е. $\alpha\neq 0$, получим $\frac{1}{\alpha}(\alpha\mathbf{x})=\frac{1}{\alpha}\cdot\mathbf{0}=\mathbf{0}$, или $\frac{1}{\alpha}\cdot\alpha\mathbf{x}=\mathbf{x}=\mathbf{0}$, что противоречит условию.

С л е д с т в и е. Равенство $\alpha\mathbf{x}=\mathbf{0}$ выполняется тогда и только тогда, когда $\alpha=0$ или $\mathbf{x}=\mathbf{0}$.

Это непосредственно вытекает из 4, 6—8.

З а м е ч а н и е. Сумму $\mathbf{x}+(-\mathbf{y})$ обозначают $\mathbf{x}-\mathbf{y}$ и называют *разностью* элементов $\mathbf{x}$ и $\mathbf{y}$.

Приведем примеры линейных пространств.

1. Множество всех свободных векторов $\mathbf{a}(a_1, a_2, a_3)$, где $a_1, a_2, a_3$ могут принимать любые действительные значения, для которых определены сложение и умножение вектора на число так, как в § 5.2, является линейным пространством. Обозначим это линейное пространство символом $V_3$. Отметим, что роль нулевого элемента здесь играет нуль-вектор; для любого вектора $\mathbf{a}$ противоположным является $-\mathbf{a}$. Аксиомы I—VIII выполняются, о чем свидетельствуют формулы (5.1)—(5.4), (5.7)—(5.10).

2. Множество всех матриц размеров $m\times n$, для которых определены сложение матриц и умножение матрицы на число соответственно формулами (3.8), (3.10), является линейным пространством. Роль нулевого элемента здесь играет нулевая матрица; для матрицы $(a_{ik})_{mn}$ противоположной будет матрица $(-a_{ik})_{mn}$. Легко видеть, что аксиомы I—VIII выполняются (см. § 3.2, свойства 1—8 линейных операций над матрицами).

3. Множество $\{P_n(x)\}$ всех алгебраических многочленов степени, не превышающей натурального числа $n$, для которых операции сложения многочленов и умножения многочлена на действительное число определены обычными правилами, является линейным пространством. Роль нулевого элемента играет многочлен, все коэффициенты которого равны нулю; для многочлена $P_n(x)=a_0x^n+ + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ противоположным будет $-P_n(x)=-a_0x^n-a_1x^{n-1}- \ldots -a_{n-1}x-a_n$. Легко видеть, что аксиомы I—VIII здесь выполняются.

З а м е ч а н и е. Множество всех многочленов степени, точно равной натуральному числу $n$, не является линейным пространством, так как сумма двух таких многочленов может оказаться многочленом степени ниже $n$ (т. е. не принадлежать рассматриваемому множеству).

4. Рассмотрим множество $A_n$, элементами которого являются упорядоченные совокупности $n$ действительных чисел $(x_1, x_2, \ldots, x_n)$. Каждый элемент этого множества будем обозначать одним символом, например, $\mathbf{x}$, $\mathbf{y}$, ..., и

писать $\mathbf{x}=(x_1, x_2, \ldots, x_n)$, $\mathbf{y}=(y_1, y_2, \ldots, y_n)$; действительные числа $x_1, x_2, \ldots, x_n$ называют координатами элемента $\mathbf{x}$. Линейные операции над элементами $A_n$ определяются формулами

$$\mathbf{x}+\mathbf{y}=((x_1+y_1), (x_2+y_2), \ldots, (x_n+y_n)),$$
$$\alpha\mathbf{x}=(\alpha x_1, \alpha x_2, \ldots, \alpha x_n).$$

Отметим, что элемент $\mathbf{0}=(0, 0, \ldots, 0)$ является нулевым, элемент $-\mathbf{x}=(-x_1, -x_2, \ldots, -x_n)$ — противоположным элементу $\mathbf{x}=(x_1, x_2, \ldots x_n)$.

Множество $A_n$ — линейное пространство, так как все аксиомы I—VIII выполнены (проверка выполнения аксиом предоставляется читателю).

5. Рассмотрим множество $C[a, b]$ всех функций $x=x(t)$, определенных и непрерывных на отрезке $[a, b]$. Операции сложения этих функций и умножения функции на число определяются обычными правилами. Нулевым элементом будет функция $x(t)\equiv 0$ для всех $t\in[a, b]$. Элементом, противоположным элементу $x(t)$, будет $-x(t)$. Легко проверить, что аксиомы I—VIII удовлетворяются, поэтому $C[a, b]$ — линейное пространство.

Введем понятие подпространства. Множество $W\subset V$ называется подпространством линейного пространства $V$, если выполняются следующие условия: 1) в множестве $W$ определены те же операции, что и в множестве $V$; 2) если $\mathbf{x}, \mathbf{y}\in W$, то $\mathbf{x}+\mathbf{y}\in W$; 3) если $\mathbf{x}\in W$, то $\alpha\mathbf{x}\in W$. Очевидно, всякое подпространство $W$ линейного пространства $V$ является линейным пространством, т. е. в $W$ выполняются аксиомы I—VIII. Прежде всего, в $W$ имеется нулевой элемент $\mathbf{0}$: если $\mathbf{x}\in W$, то $0\mathbf{x}=\mathbf{0}\in W$. Для любого элемента $\mathbf{x}\in W$ имеется противоположный элемент $-\mathbf{x}$: если $\mathbf{x}\in W$, то $(-1)\mathbf{x}=-\mathbf{x}\in W$. Легко видеть, что аксиомы I—VIII для элементов множества $W$ будут выполнены. Отметим, что нулевой элемент $\mathbf{0}$ линейного пространства $V$ образует подпространство данного пространства, которое называют нулевым подпространством. Само линейное пространство $V$ можно рассматривать как подпространство этого пространства. Эти подпространства называют тривиальными, а все другие, если они имеются,— нетривиальными.

Приведем примеры нетривиальных подпространств.

1. Множество $V_2$ всех свободных векторов $\mathbf{a}(a_1, a_2)$, параллельных некоторой плоскости, для которых обычным образом определены операции сложения векторов и умножения вектора на число, является подпространством линейного пространства $V_3$ (см. пример 1).

2. Множество $V_1$ всех свободных векторов $\mathbf{a}(a_1)$, параллельных некоторой прямой, представляет подпространство линейного пространства $V_3$.

3. Множество $\{P_{n-1}(x)\}$ всех алгебраических многочленов степени, не превышающей натурального числа $n-1$, является подпространством линейного пространства $\{P_n(x)\}$ (см. пример 3 линейных пространств).

## § 7.2. Линейная зависимость и линейная независимость векторов линейного пространства

Рассмотрим векторы (элементы) $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ линейного пространства. Вектор $\mathbf{y}=\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\ldots+\alpha_n\mathbf{x}_n$, где $\alpha_1, \alpha_2, \ldots, \alpha_n$ — некоторые числа, называется линейной комбинацией векторов $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, а числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ — коэффициентами этой линейной комбинации. Если все числа $\alpha_i$ $(i=1, 2, \ldots, n)$ равны нулю, линейная комбинация называется тривиальной. Если хотя бы одно из чисел $\alpha_i$ отлично от нуля, линейная комбинация называется нетривиальной.

Система векторов

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \tag{7.1}$$

называется линейно зависимой, если существуют числа

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \tag{7.2}$$

не все равные нулю, такие, что

$$\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\ldots+\alpha_n\mathbf{x}_n=0. \tag{7.3}$$

Если таких чисел не существует, т. е. равенство (7.3) выполняется только в случае

$$\alpha_1=\alpha_2=\ldots=\alpha_n=0, \tag{7.4}$$

система векторов (7.1) называется линейно независимой.

Другими словами, векторы (7.1) называются линейно зависимыми, если существует их нетривиальная линейная комбинация, равная нуль-вектору, и линейно независимыми, если только их тривиальная линейная комбинация является нуль-вектором.

Из определения линейной зависимости и линейной независимости векторов вытекают следующие утверждения.

1. Всякая система векторов, содержащая нуль-вектор, линейно зависима.

2. Если $k$ $(k<n)$ векторов системы (7.1) линейно зависимы, то и вся система линейно зависима.

3. Если из системы линейно независимых векторов $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ отбросить $r (r<n)$ векторов, то оставшиеся векторы образуют также линейно независимую систему.

4. Если среди векторов системы (7.1) имеются такие векторы $\mathbf{x}_k$ и $\mathbf{x}_m$, что $\mathbf{x}_k=\lambda\mathbf{x}_m$, где $\lambda$ — некоторое число, то система (7.1) линейно зависима.

Докажем эти утверждения. 1. Пусть, например, $\mathbf{x}_1=\mathbf{0}$, тогда $1\cdot\mathbf{0}+0\cdot\mathbf{x}_2+\ldots+0\cdot\mathbf{x}_n=\mathbf{0}$, т. е. выполнено равенство (7.3), в котором не все числа $\alpha_i$ равны нулю, так как $\alpha_1=1\neq0$. 2. В этом случае $\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\ldots+\alpha_k\mathbf{x}_k=\mathbf{0}$, где среди чисел $\alpha_1, \alpha_2, \ldots, \alpha_k$ имеются отличные от нуля, поэтому выполняется равенство (7.3), т. е. $\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\ldots+\alpha_k\mathbf{x}_k+0\cdot\mathbf{x}_{k+1}+\ldots+0\cdot\mathbf{x}_n=\mathbf{0}$. 3. Это следует из предыдущего. Действительно, допустим противное, т. е. оставшаяся система линейно зависима, тогда и вся система должна быть линейно зависимой, что противоречит условию. 4. Так как $\mathbf{x}_k=\lambda\mathbf{x}_m$, то $\mathbf{x}_k-\lambda\mathbf{x}_m=\mathbf{0}$, т. е. $\mathbf{x}_k$ и $\mathbf{x}_m$ линейно зависимы, ибо $\alpha_k=1\neq0$, поэтому и вся система линейно зависима.

Отметим частный случай системы (7.1) при $n=1$, т. е. случай, когда имеется только один вектор, который обозначим буквой $\mathbf{x}$ без индекса. Если $\mathbf{x}\neq\mathbf{0}$, то система линейно независима, так как равенство $\alpha\mathbf{x}=\mathbf{0}$ выполняется лишь при $\alpha=0$. Если $\mathbf{x}=\mathbf{0}$, то система линейно зависима, поскольку равенство $\alpha\mathbf{0}=\mathbf{0}$ выполняется при $\alpha\neq0$.

**Теорема 7.1.** Векторы $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ линейно зависимы тогда и только тогда, когда хотя бы один из них является линейной комбинацией всех остальных.

Доказательство теоремы аналогично доказательству теоремы 5.5.

Введем понятия коллинеарности и компланарности векторов линейного пространства. Два вектора линейного пространства называются коллинеарными, если они линейно зависимы, и неколлинеарными, если линейно независимы. Три вектора линейного пространства называются компланарными, если они линейно зависимы, и не-компланарными, если линейно независимы. Введенные понятия коллинеарности и компланарности векторов линейного пространства совпадают с известными из аналитической геометрии понятиями коллинеарности и компланарности обычных векторов.

## § 7.3. Размерность и базис линейного пространства. Изоморфизм линейных пространств

Линейное пространство $V$ называется $n$-мерным, если в нем существует $n$ линейно независимых векторов, а любые $n+1$ векторы являются линейно зависимыми. Число $n$ называется в этом случае размерностью линейного пространства $V$. Другими словами, число $n$ называется *размерностью* линейного пространства $V$, если выполняются следующие условия: 1) в $V$ существует $n$ линейно независимых векторов; 2) любая система $n+1$ векторов из $V$ линейно зависима. Размерность линейного пространства $V$ обозначают $\dim V$ (от французского слова dimension — размерность). Если пространство состоит из одного нулевого элемента, его размерность считают равной нулю. Итак, размерность линейного пространства — это наибольшее возможное количество линейно независимых элементов в нем. Введенное понятие размерности согласуется с наглядным представлением о ней; так, пространство $V_3$ всех свободных векторов является трехмерным ($\dim V_3=3$), пространство $V_2$ — двумерным, пространство $V_1$ — одномерным.

*Базисом* $n$-мерного линейного пространства $V_n$ называется любая упорядоченная система $n$ линейно независимых векторов этого пространства. Приведем примеры базисов некоторых линейных пространств. Базис пространства $V_3$ образует любая тройка некомпланарных векторов, так как эти векторы линейно независимы (см. теорему 5.8) и любая четверка векторов линейно зависима (см. теорему 5.10). Базис пространства $V_2$ образуют два любых неколлинеарных вектора, поскольку они линейно независимы (см. теорему 5.6) и любой вектор плоскости, определяемой этими двумя векторами, можно разложить по ним (см. теорему 5.7). Базисом линейного пространства $V_1$ является любой ненулевой вектор, коллинеарный данной прямой.

Рассмотрим вопрос о базисе и размерности линейного пространства $A_n$ (см. § 7.1, пример 4). Докажем, что система $n$ векторов

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots$$
$$\ldots, \mathbf{e}_n = (0, 0, \ldots, 0, 1) \qquad (7.5)$$

линейно независима, а совокупность $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{x}$, где $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ — любой элемент этого пространства, образует линейно зависимую систему. Действительно, линейная комбинация векторов (7.5) представляет вектор $\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, который является нулевым лишь при $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Это означает, что векторы (7.5) линейно независимы. Поскольку $\mathbf{x} = (x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ есть линейная комбинация векторов (7.5), то система $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{x}$ линейно зависима (в силу теоремы 7.1). Следовательно, линейное пространство $A_n$ будет $n$-мерным, а система векторов (7.5) образует базис этого линейного пространства.

Линейное пространство, в котором имеется базис, состоящий из конечного числа векторов, называется конечномерным. Примерами конечномерных пространств являются пространства $V_1$, $V_2$, $V_3$, $A_n$.

Линейное пространство называется бесконечномерным, если при любом натуральном числе $m$ в нем найдется $m$ линейно независимых векторов. Примером бесконечномерного пространства может служить линейное пространство $C[a, b]$ всех функций $x = x(t)$, определенных и непрерывных на отрезке $[a, b]$. Действительно, для любого натурального числа $m$ система $m+1$ элементов (векторов) этого пространства $1, t, t^2, \ldots, t^m$ линейно независима, в противном случае многочлен $\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_m t^m$, не все коэффициенты которого равны нулю, оказался бы тождественно равным нулю на отрезке $[a, b]$.

Пусть даны два линейных пространства $V$ и $U$. Если между элементами $x \in V$ и $y \in U$ этих пространств установлено взаимно-однозначное соответствие, будем писать $x \leftrightarrow y$. Два линейных пространства $V$ и $U$ называются *изоморфными,* когда между их элементами можно установить взаимно однозначное соответствие такое, что если $x_1 \leftrightarrow y_1$, $x_2 \leftrightarrow y_2$, где $x_1, x_2 \in V$, $y_1, y_2 \in U$, то $(x_1 + x_2) \leftrightarrow (y_1 + y_2)$, $\alpha x_1 \leftrightarrow \alpha y_1$, где $\alpha$ — действительное число. Приведем без доказательства теорему о необходимом и достаточном условии изоморфизма двух линейных пространств.

**Теорема 7.2.** Два линейных пространства изоморфны тогда и только тогда, когда они имеют одинаковую размерность.

В частности, пространство $V_3$ (всех свободных векторов) и пространство $A_3$ (всех упорядоченных троек действительных чисел) изоморфны. Отметим также, что каждое конечномерное линейное пространство размерности $n$ изоморфно линейному пространству $A_n$.

## § 7.4. Координаты вектора линейного пространства

**Теорема 7.3.** Если $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ — базис линейного $n$-мерного пространства $V_n$, то любой вектор $\mathbf{x}$ этого пространства линейно выражается через базисные векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, т. е.

$$\mathbf{x}=\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n. \qquad (7.6)$$

Коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_n$ этого разложения определяются однозначно.

Доказательство. Пусть $\mathbf{x}$ — произвольный вектор пространства $V_n$. Поскольку это пространство является $n$-мерным, то система векторов $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{x}$ линейно зависима, т. е. существуют числа $\beta_1, \beta_2, \ldots, \beta_n, \beta_{n+1}$, не все равные нулю, такие, что

$$\beta_1\mathbf{e}_1+\beta_2\mathbf{e}_2+\ldots+\beta_n\mathbf{e}_n+\beta_{n+1}\mathbf{x}=\mathbf{0}.$$

Можно утверждать, что $\beta_{n+1}\neq 0$. Действительно, предположив противное, т. е. $\beta_{n+1}=0$, получим, что $\beta_1, \beta_2, \ldots, \beta_n$ не все равны нулю, поэтому векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно зависимы, а это противоречит условию. Поскольку $\beta_{n+1}\neq 0$, то из этого равенства следует (7.6), где $\alpha_i=-\beta_i|\beta_{n+1}$ $(i=1, 2, \ldots, n)$.

Докажем единственность разложения (7.6). Предположим, что существует другая система чисел $\gamma_1, \gamma_2, \ldots, \gamma_n$ такая, что $\mathbf{x}=\gamma_1\mathbf{e}_1+\gamma_2\mathbf{e}_2+\ldots+\gamma_n\mathbf{e}_n$, тогда из последнего равенства и (7.6) следует, что

$$\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n=\gamma_1\mathbf{e}_1+\gamma_2\mathbf{e}_2+\ldots+\gamma_n\mathbf{e}_n,$$

или

$$(\alpha_1-\gamma_1)\mathbf{e}_1+(\alpha_2-\gamma_2)\mathbf{e}_2+\ldots+(\alpha_n-\gamma_n)\mathbf{e}_n=\mathbf{0}.$$

Так как $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно независимы, то отсюда следует, что $\alpha_1-\gamma_1=0$, $\alpha_2-\gamma_2=0, \ldots, \alpha_n-\gamma_n=0$, или $\alpha_1=\gamma_1$, $\alpha_2=\gamma_2, \ldots, \alpha_n=\gamma_n$.

Выражение (7.6) называется разложением вектора $\mathbf{x}$ по базису $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. *Координатами вектора* $\mathbf{x}$ в базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ называются коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_n$ в разложении этого вектора по данному базису, т. е. в формуле (7.6). Если вектор $\mathbf{x}$ в некотором базисе имеет координаты $\alpha_1, \alpha_2, \ldots, \alpha_n$, то пишут $\mathbf{x}=(\alpha_1, \alpha_2, \ldots, \alpha_n)$ или $\mathbf{x}(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Отметим, что если в линейном пространстве $V_3$ в качестве базиса взять орты $\mathbf{i}, \mathbf{j}, \mathbf{k}$, то координатами вектора $\mathbf{x}$ в данном базисе будут декартовы прямоугольные координаты этого вектора, известные из аналитической геометрии (см. формулу (5.31)).

Пример 1. Пусть $A_4$ — четырехмерное линейное пространство с базисом $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$. Найти координаты векторов $\mathbf{e}_3$ и $\mathbf{x}=2\mathbf{e}_1+4\mathbf{e}_3-5\mathbf{e}_4$ в этом базисе.

Представим каждый из векторов $\mathbf{e}_3$ и $\mathbf{x}$ в виде (7.6). Поскольку $\mathbf{e}_3=0\cdot\mathbf{e}_1+0\cdot\mathbf{e}_2+1\cdot\mathbf{e}_3+0\cdot\mathbf{e}_4$, то вектор $\mathbf{e}_3$ имеет координаты $(0, 0, 1, 0)$. Так как $\mathbf{x}=2\mathbf{e}_1+0\cdot\mathbf{e}_2+4\mathbf{e}_3-5\mathbf{e}_4$, то вектор $\mathbf{x}$ имеет координаты $(2, 0, 4, -5)$ или $\mathbf{x}\ (2, 0, 4, -5)$.

Операции над векторами сводятся к операциям над их координатами на основании следующих свойств.

1. Вектор является нулевым вектором линейного пространства тогда и только тогда, когда все его координаты в любом базисе равны нулю.

2. Координаты суммы двух векторов в некотором базисе равны сумме соответствующих координат данных векторов в том же базисе.

3. Координаты произведения вектора на число равны произведению соответствующих координат на это число (в одном и том же базисе).

4. Два вектора равны тогда и только тогда, когда равны их соответствующие координаты в одном и том же базисе.

5. Вектор $\mathbf{y}$ является линейной комбинацией векторов $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ тогда и только тогда, когда каждая координата вектора $\mathbf{y}$ является такой же линейной комбинацией соответствующих координат этих векторов в одном и том же базисе.

Докажем, например, первые два свойства. 1. Если в некотором базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ все координаты вектора $\mathbf{x}$ равны нулю, то $\mathbf{x}=0\cdot\mathbf{e}_1+0\cdot\mathbf{e}_2+\ldots+0\cdot\mathbf{e}_n=\mathbf{0}$, т. е. $\mathbf{x}$ — нулевой вектор. Обратно, если $\mathbf{x}=\mathbf{0}$ и в некотором базисе $\mathbf{x}=\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n$, то $\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n=\mathbf{0}$. Поскольку векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно независимы, то $\alpha_1=\alpha_2=\ldots=\alpha_n=0$, т. е. все координаты вектора $\mathbf{x}$ равны нулю. 2. Если в некотором базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ $\mathbf{x}=\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n$ и $\mathbf{y}=\beta_1\mathbf{e}_1+\beta_2\mathbf{e}_2+\ldots+\beta_n\mathbf{e}_n$, то $\mathbf{x}+\mathbf{y}=(\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n)+(\beta_1\mathbf{e}_1+\beta_2\mathbf{e}_2+\ldots+\beta_n\mathbf{e}_n)$. На основании аксиом I, II, VIII линейного пространства

$$\mathbf{x}+\mathbf{y}=(\alpha_1+\beta_1)\mathbf{e}_1+(\alpha_2+\beta_2)\mathbf{e}_2+\ldots+(\alpha_n+\beta_n)\mathbf{e}_n.$$

Поскольку разложение вектора $\mathbf{x}+\mathbf{y}$ по базису $\mathbf{e}_1, \mathbf{e}_2, \ldots \ldots, \mathbf{e}_n$ определяется однозначно, то вектор $\mathbf{x}+\mathbf{y}$ имеет координаты

$$(\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_n+\beta_n).$$

Пример 2. В некотором базисе даны векторы $\mathbf{x}(1, -2, 3, -1, 0)$, $\mathbf{y}(2, 1, -3, -1, 4)$. *Найти координаты вектора* $4\mathbf{x}-3\mathbf{y}$.

Поскольку $4\mathbf{x}=(4, -8, 12, -4, 0)$ и $-3\mathbf{y}(-6, -3, 9, 3, -12)$, то вектор $4\mathbf{x}-3\mathbf{y}=4\mathbf{x}+(-3\mathbf{y})$ имеет координаты $(-2, -11, 21, -1, -12)$.

## § 7.5. Ранг системы векторов линейного пространства

Рассмотрим систему $m$ векторов

$$\left.\begin{aligned} \mathbf{a}_1 &= (a_{11}, \ a_{21}, \ \ldots, \ a_{n1}), \\ \mathbf{a}_2 &= (a_{12}, \ a_{22}, \ \ldots, \ a_{n2}), \\ &\ldots\ldots\ldots\ldots\ldots \\ \mathbf{a}_m &= (a_{1m}, \ a_{2m}, \ \ldots, \ a_{nm}) \end{aligned}\right\} \tag{7.7}$$

линейного $n$-мерного пространства, координаты которых заданы в одном и том же базисе. Системе векторов (7.7) поставим в соответствие матрицу

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix}, \tag{7.8}$$

в $k$-м столбце которой записаны координаты вектора $\mathbf{a}_k$ $(k=1, 2, \ldots, m)$. Матрицу (7.8) называют матрицей системы векторов (7.7) в данном базисе, а ранг этой матрицы — рангом системы векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. Обратно, если дана матрица (7.8), ей можно поставить в соответствие систему (7.7) $m$ векторов линейного $n$-мерного пространства. Согласно свойству 5 § 7.4 будем говорить, что столбцы матрицы (7.8) линейно зависимы, если векторы (7.7) линейно зависимы и обратно.

Приведем без доказательства теорему, которая позволяет судить о линейной независимости векторов, заданных своими координатами.

**Теорема 7.4.** Для того чтобы $m$ векторов $n$-мерного линейного пространства были линейно независимыми, необходимо и достаточно, чтобы ранг матрицы этой системы был равен $m$.

Следствие 1. Система $n$ векторов $n$-мерного линейного пространства линейно независима тогда и только тогда, когда матрица этой системы векторов является невырожденной.

Следствие 2. Если ранг матрицы системы $m$ векторов линейного пространства равен $r$, максимальное число линейно независимых векторов этой системы равно $r$.

Пример. Найти максимальное число линейно независимых векторов в системе $\mathbf{a}_1(1, 4, 2, 7)$, $\mathbf{a}_2(-1, -2, -3, -6)$, $\mathbf{a}_3(0, 5, 0, 5)$, $\mathbf{a}_4(3, 0, 6, 9)$, $\mathbf{a}_5(2, 3, 1, 6)$.

Матрица данной системы векторов имеет вид

$$A = \begin{bmatrix} 1 & -1 & 0 & 3 & 2 \\ 4 & -2 & 5 & 0 & 3 \\ 2 & -3 & 0 & 6 & 1 \\ 7 & -6 & 5 & 9 & 6 \end{bmatrix}.$$

Так как ранг этой матрицы равен трем (см. § 3.6, пример 2), то максимальное число линейно независимых векторов этой системы равно трем.

Замечание. Можно доказать, что максимальное число линейно независимых строк всякой матрицы равно максимальному числу ее линейно независимых столбцов, т. е. равно рангу этой матрицы.

## § 7.6. Преобразование координат вектора при изменении базиса

Координаты вектора, очевидно, зависят от выбора базиса. Координаты одного и того же вектора будут различными в разных базисах. Задача преобразования координат вектора состоит в нахождении зависимости между координатами данного вектора в разных базисах, связь между которыми известна. Формулами преобразования координат называют формулы, связывающие координаты вектора в разных базисах.

В линейном $n$-мерном пространстве $V_n$ фиксируем два базиса

$$\mathbf{e}_1, \quad \mathbf{e}_2, \quad \ldots, \quad \mathbf{e}_n, \tag{7.9}$$

$$\mathbf{e}_1', \mathbf{e}_2', \ \ldots, \ \mathbf{e}_n'. \tag{7.10}$$

Матрицей перехода от базиса (7.9) к базису (7.10) называется матрица системы векторов (7.10) в базисе (7.9). Каждый вектор системы (7.10) можно разложить по базису (7.9). Пусть

$$\left.\begin{array}{l} \mathbf{e}_1' = t_{11}\mathbf{e}_1 + t_{21}\mathbf{e}_2 + \ldots + t_{n1}\mathbf{e}_n, \\ \mathbf{e}_2' = t_{12}\mathbf{e}_1 + t_{22}\mathbf{e}_2 + \ldots + t_{n2}\mathbf{e}_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{e}_n' = t_{1n}\mathbf{e}_1 + t_{2n}\mathbf{e}_2 + \ldots + t_{nn}\mathbf{e}_n, \end{array}\right\} \tag{7.11}$$

тогда матрица перехода от базиса (7.9) к базису (7.10) имеет вид

$$T = \begin{bmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{bmatrix}. \tag{7.12}$$

На основании следствия 1 из теоремы 7.4 (см. § 7.5) заключаем, что матрица перехода от одного базиса к другому будет невырожденной (так как базисные векторы линейно независимы), всякую невырожденную матрицу $n$-го порядка можно рассматривать как матрицу перехода от одного базиса $n$-мерного линейного пространства к другому базису этого пространства. Очевидно, матрица $T^{-1}$, обратная матрице (7.12), является матрицей перехода от базиса (7.10) к базису (7.9).

Докажем теорему, устанавливающую связь между координатами вектора в различных базисах.

**Теорема 7.5.** Если $x_1, x_2, \ldots, x_n$ — координаты вектора $\mathbf{x}$ в базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, $x_1', x_2', \ldots, x_n'$ — координаты того же вектора в базисе $\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n'$, то

$$X = TX', \tag{7.13}$$

где

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix}, \quad X' = \begin{bmatrix} x_1' \\ x_2' \\ \ldots \\ x_n' \end{bmatrix}, \tag{7.14}$$

$T$ — матрица перехода от базиса $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ к базису $\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n'$.

Доказательство. Из условия теоремы следует, что

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n, \quad \mathbf{x} = x_1'\mathbf{e}_1' + x_2'\mathbf{e}_2' + \ldots + x_n'\mathbf{e}_n'. \tag{7.15}$$

Если матрица $T$ имеет вид (7.12), т. е. выполнены равенства (7.11), то второе из равенств (7.15) можно записать так:

$$\mathbf{x} = x_1'(t_{11}\mathbf{e}_1 + t_{21}\mathbf{e}_2 + \ldots + t_{n1}\mathbf{e}_n) + x_2'(t_{12}\mathbf{e}_1 + t_{22}\mathbf{e}_2 + \ldots$$
$$\ldots + t_{n2}\mathbf{e}_n) + \ldots + x_n'(t_{1n}\mathbf{e}_1 + t_{2n}\mathbf{e}_2 + \ldots + t_{nn}\mathbf{e}_n) =$$
$$= (t_{11}x_1' + t_{12}x_2' + \ldots + t_{1n}x_n')\mathbf{e}_1 + (t_{21}x_1' + t_{22}x_2' + \ldots$$
$$\ldots + t_{2n}x_n')\mathbf{e}_2 + \ldots + (t_{n1}x_1' + t_{n2}x_2' + \ldots + t_{nn}x_n')\mathbf{e}_n.$$

Сравнивая полученное разложение вектора $\mathbf{x}$ по базису $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ с первым из равенств (7.15), получаем

$$\left.\begin{array}{l} x_1 = t_{11}x_1' + t_{12}x_2' + \ldots + t_{1n}x_n', \\ x_2 = t_{21}x_1' + t_{22}x_2' + \ldots + t_{2n}x_n', \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = t_{n1}x_1' + t_{n2}x_2' + \ldots + t_{nn}x_n', \end{array}\right\} \tag{7.16}$$

или в матричной форме $X = TX'$, где $X$ и $X'$ определяются формулами (7.14), а $T$ — формулой (7.12).

Замечание. Формулы (7.13) и (7.16) выражают старые координаты $x_1, x_2, \ldots, x_n$ вектора $\mathbf{x}$ через его новые координаты. Чтобы получить формулы, выражающие новые координаты через старые, умножим слева равенство (7.13) на матрицу $T^{-1}$, обратную матрице $T$; $T^{-1}X = T^{-1}TX'$, $T^{-1}X = X'$, или $X' = T^{-1}X$.

Пример. В пространстве $V_2$ рассмотрим базис $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$, где $\mathbf{i}$, $\mathbf{j}$ — орты, и базис $\mathbf{e}_1' = \mathbf{i}'$, $\mathbf{e}_2' = \mathbf{j}'$, где $\mathbf{i}'$, $\mathbf{j}'$ — орты, причем $\mathbf{i}'$ образует с $\mathbf{i}$ угол $\varphi$ (рис. 7.1). В данном случае $\mathbf{i}' = \mathbf{i}\cos\varphi + \mathbf{j}\sin\varphi$, $\mathbf{j}' = -\mathbf{i}\sin\varphi + \mathbf{j}\cos\varphi$. Матрица перехода от базиса $\mathbf{i}$, $\mathbf{j}$ к базису $\mathbf{i}'$, $\mathbf{j}'$ имеет вид

$$T = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$

Если вектор $\mathbf{a}$ имеет координаты $x$, $y$ в базисе $\mathbf{i}$, $\mathbf{j}$, $x'$, $y'$ — в базисе $\mathbf{i}'$, $\mathbf{j}'$, то $x = x'\cos\varphi - y'\sin\varphi$, $y = x'\sin\varphi + y'\cos\varphi$.

## § 7.7. Евклидово пространство

**Определение евклидова пространства.** В линейном пространстве $V$ кроме операций сложения векторов и умножения вектора на действительное число введем еще одну операцию, которую назовем *скалярным умножением* векторов. Каждой упорядоченной паре векторов $\mathbf{x}$, $\mathbf{y} \in V$ поставим в соответствие действительное число, которое назовем их *скалярным произведением* и обозначим $(\mathbf{x}, \mathbf{y})$. Потребуем, чтобы для любых $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in V$ и любого числа $\alpha \in R$ выполнялись следующие аксиомы:

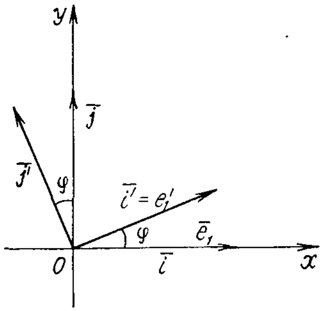



Рис. 7.1

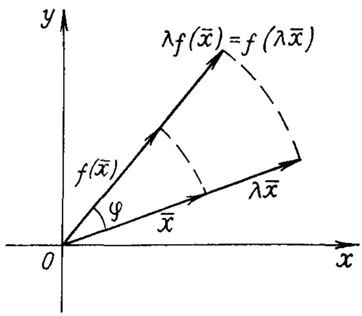



Рис. 7.2

I. $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$. II. $(\mathbf{x}+\mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$.

III. $(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y})$. IV. $(\mathbf{x}, \mathbf{x}) > 0$ для всех $\mathbf{x} \neq \mathbf{0}$, $(\mathbf{x}, \mathbf{x}) = 0$ для $\mathbf{x} = \mathbf{0}$.

Очевидно, скалярное произведение равно нулю, если хотя бы один из векторов нулевой: $(\mathbf{0}, \mathbf{y}) = (0\mathbf{x}, \mathbf{y}) = 0(\mathbf{x}, \mathbf{y}) = 0$.

Скалярное произведение $(\mathbf{x}, \mathbf{x})$ вектора $\mathbf{x}$ на себя называется *скалярным квадратом* этого вектора и обозначается $\mathbf{x}^2$, т. е.

$$(\mathbf{x}, \mathbf{x}) = \mathbf{x}^2. \tag{7.17}$$

*Евклидовым пространством* называется линейное действительное пространство, в котором задана операция скалярного умножения векторов, удовлетворяющая аксиомам I — IV. Если $n$-мерное линейное пространство является евклидовым, будем называть его евклидовым $n$-мерным пространством, а базис этого линейного пространства — базисом евклидова пространства.

Приведем примеры евклидовых пространств.

1. В линейном пространстве $V_3$ скалярное произведение двух векторов **a** и **b** определим так, как в § 5.7; аксиомы I — IV для него будут выполнены (см. формулы (5.54)—(5.56) и (5.52)).

Следовательно, линейное пространство $V_3$ всех свободных векторов с обычным определением скалярного произведения является евклидовым пространством.

2. Рассмотрим $n$-мерное линейное пространство $A_n$ упорядоченных совокупностей $n$ действительных чисел. Скалярное произведение двух его элементов $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ по аналогии с формулой (5.59) определим соотношением

$$(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n. \tag{7.18}$$

Легко видеть, что все аксиомы I—IV скалярного произведения при этом выполняются. Таким образом, рассматриваемое линейное пространство со скалярным произведением (7.18) является евклидовым пространством, его обозначают $E_n$.

3. В бесконечномерном линейном пространстве $C[a, b]$ всех функций, непрерывных на отрезке $[a, b]$, скалярное произведение двух его функций $x(t)$, $y(t)$ определим формулой

$$(x, y) = \int_a^b x(t)\, y(t)\, dt. \tag{7.19}$$

Непосредственной проверкой можно убедиться в том, что аксиомы I — IV скалярного произведения будут выполнены, в частности,

$$(x, x) = \int_a^b x^2(t)\, dt > 0 \text{ при } x(t) \neq 0, \ (x, x) = 0 \text{ при } x(t) \equiv 0.$$

Следовательно, линейное пространство $C[a, b]$ с указанным определением скалярного произведения любых двух его элементов будет евклидовым пространством.

**Норма вектора евклидова пространства.** Нормой вектора евклидова пространства называется арифметическое значение корня из скалярного квадрата этого вектора. Норму вектора **x** обозначим $\|\mathbf{x}\|$, тогда по определению

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\mathbf{x}^2}. \tag{7.20}$$

Норма вектора обладает следующими свойствами.

1. $\|\mathbf{x}\|=0$ тогда и только тогда, когда $\mathbf{x}=\mathbf{0}$.
2. $\|\alpha\mathbf{x}\|=|\alpha|\cdot\|\mathbf{x}\|$, где $\alpha$ — действительное число.
3. $|(\mathbf{x}, \mathbf{y})|\leqslant\|\mathbf{x}\|\cdot\|\mathbf{y}\|$. 4. $\|\mathbf{x}+\mathbf{y}\|\leqslant\|\mathbf{x}\|+\|\mathbf{y}\|$.

Докажем эти свойства. 1. Из аксиомы IV скалярного произведения следует, что $\|\mathbf{x}\|>0$ для $\mathbf{x}\neq\mathbf{0}$ и $\|\mathbf{x}\|=0$ при $\mathbf{x}=\mathbf{0}$. 2. На основании аксиом I и III скалярного произведения получаем $\|\alpha\mathbf{x}\|=\sqrt{(\alpha\mathbf{x}, \alpha\mathbf{x})}=\sqrt{\alpha^2(\mathbf{x}, \mathbf{x})}=\sqrt{\alpha^2}\cdot\sqrt{(\mathbf{x}, \mathbf{x})}=|\alpha|\cdot\|\mathbf{x}\|$, $\|\alpha\mathbf{x}\|=|\alpha|\cdot\|\mathbf{x}\|$. 3. В соответствии с аксиомой IV скалярного произведения $\|\mathbf{x}+\mathbf{y}\|^2\geqslant 0$, поэтому $0\leqslant\|\mathbf{x}+\alpha\mathbf{y}\|^2=(\mathbf{x}+\alpha\mathbf{y}, \ \mathbf{x}+\alpha\mathbf{y})=(\mathbf{y}, \mathbf{y})\alpha^2+2(\mathbf{x}, \mathbf{y})\alpha+(\mathbf{x}, \mathbf{x})$. Эту сумму рассматриваем как квадратный трехчлен относительно $\alpha$. Поскольку указанный трехчлен сохраняет знак, его дискриминант неположителен, т. е.

$$(\mathbf{x}, \mathbf{y})^2-(\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})\leqslant 0, \quad (\mathbf{x}, \mathbf{y})^2\leqslant(\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y}).$$

Так как $\sqrt{(\mathbf{x}, \mathbf{y})^2}=|(\mathbf{x}, \mathbf{y})|$, $\sqrt{(\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})}=\sqrt{(\mathbf{x}, \mathbf{x})}\times\sqrt{(\mathbf{y}, \mathbf{y})}=\|\mathbf{x}\|\cdot\|\mathbf{y}\|$, из последнего неравенства следует неравенство

$$|(\mathbf{x}, \mathbf{y})|\leqslant\|\mathbf{x}\|\cdot\|\mathbf{y}\|, \tag{7.21}$$

которое называют неравенством Коши[1] — Буняковского[2]. 4. Используя неравенство (7.21), получаем

$$\|\mathbf{x}+\mathbf{y}\|^2=(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y})=(\mathbf{x}, \mathbf{x})+2(\mathbf{x}, \mathbf{y})+(\mathbf{y}, \mathbf{y})\leqslant$$
$$\leqslant\|\mathbf{x}\|^2+2\|\mathbf{x}\|\cdot\|\mathbf{y}\|+\|\mathbf{y}\|^2=(\|\mathbf{x}\|+\|\mathbf{y}\|)^2,$$

откуда

$$\|\mathbf{x}+\mathbf{y}\|\leqslant\|\mathbf{x}\|+\|\mathbf{y}\|. \tag{7.22}$$

Неравенство (7.22) называется неравенством треугольника.

Запишем норму и неравенства (7.21), (7.22) для векторов (элементов) каждого из рассмотренных выше евклидовых пространств.

В евклидовом пространстве $V_3$ с обычным определением скалярного произведения норма вектора совпадает с его длиной, т. е. $\|\mathbf{a}\|=|\mathbf{a}|$; это следует из формул (5.52)

[1] *Огюстен Коши* (*Augustin Cauchy*, 1789—1857) — французский математик.

[2] *В. Я. Буняковский* (1804—1889) — русский математик, академик Петербургской академии наук.

и (7.20). Неравенства (7.21) и (7.22) принимают соответственно вид $|(\mathbf{a}, \mathbf{b})| \leqslant |\mathbf{a}| \cdot |\mathbf{b}|$, $|\mathbf{a}+\mathbf{b}| \leqslant |\mathbf{a}|+|\mathbf{b}|$. Отметим, что неравенство $|(\mathbf{a}, \mathbf{b})| \leqslant |\mathbf{a}| \cdot |\mathbf{b}|$ следует из формулы (5.49), а неравенство $|\mathbf{a}+\mathbf{b}| \leqslant |\mathbf{a}|+|\mathbf{b}|$ — из определений суммы векторов и длины вектора; оно имеет простой геометрический смысл (в треугольнике сумма длин двух сторон больше длины третьей стороны).

В евклидовом пространстве $C[a, b]$ норма элемента $x(t)$ определяется формулой

$$\|x(t)\| = \sqrt{\int_a^b x^2(t)\,dt},$$

неравенства (7.21) и (7.22) принимают вид

$$\int_a^b x(t)\,y(t)\,dt \leqslant \sqrt{\int_a^b x^2(t)\,dt} \cdot \sqrt{\int_a^b y^2(t)\,dt},$$

$$\sqrt{\int_a^b [x(t)+y(t)]^2\,dt} \leqslant \sqrt{\int_a^b x^2(t)\,dt} + \sqrt{\int_a^b y^2(t)\,dt}.$$

В евклидовом пространстве $E_n$ со скалярным произведением (7.18) норма элемента $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ определяется формулой

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2},$$

а неравенства (7.21) и (7.22) принимают вид

$$|x_1y_1 + x_2y_2 + \ldots + x_ny_n| \leqslant$$
$$\leqslant \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \cdot \sqrt{y_1^2 + y_2^2 + \ldots + y_n^2},$$

$$\sqrt{(x_1+y_1)^2 + (x_2+y_2)^2 + \ldots + (x_n+y_n)^2} \leqslant$$
$$\leqslant \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} + \sqrt{y_1^2 + y_2^2 + \ldots + y_n^2}.$$

**Угол между двумя векторами евклидова пространства.** Из неравенства Коши — Буняковского получаем

$$\frac{|(\mathbf{x}, \mathbf{y})|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \leqslant 1, \text{ или } -1 \leqslant \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \leqslant 1;$$

следовательно, отношение $\dfrac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$ можно рассматривать как косинус некоторого угла. Углом между двумя векторами **x** и **y** евклидова пространства называется угол $\varphi$, для которого

$$\cos\varphi = \frac{(\mathbf{x},\ \mathbf{y})}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}\ (0\leqslant\varphi<2\pi). \tag{7.23}$$

Отметим, что в пространстве $V_3$ всех свободных векторов введенное понятие угла совпадает с понятием угла, рассматриваемого в векторной алгебре.

Два вектора евклидова пространства называются *ортогональными,* если их скалярное произведение равно нулю. Очевидно, нулевой вектор ортогонален любому другому вектору. В пространстве $V_3$ ортогональность векторов означает их перпендикулярность.

Из определений следует, что ненулевые векторы **x** и **y** ортогональны тогда и только тогда, когда $\cos\varphi=0$.

Для того чтобы ненулевые векторы **x** и **y** были коллинеарны, т. е. $\mathbf{y}=\alpha\mathbf{x}$, необходимо и достаточно, чтобы $\cos\varphi=\pm1$, т. е. $\varphi=0$ (при $\alpha>0$) и $\varphi=\pi$ (при $\alpha<0$). Действительно, пусть векторы **x** и **y** коллинеарны. Поскольку в данном случае $\mathbf{y}=\alpha\mathbf{x}$, то

$$|(\mathbf{x},\ \mathbf{y})| = |(\mathbf{x},\ \alpha\mathbf{x})| = |\alpha|\,(\mathbf{x},\ \mathbf{x}) = |\alpha|\cdot\|\mathbf{x}\|^2,$$

$$\|\mathbf{x}\|\cdot\|\mathbf{y}\| = \|\mathbf{x}\|\cdot|\alpha|\cdot\|\mathbf{x}\| = |\alpha|\cdot\|\mathbf{x}\|^2,$$

откуда $|(\mathbf{x},\ \mathbf{y})| = \|\mathbf{x}\|\cdot\|\mathbf{y}\|$, поэтому $\cos\varphi=\pm1$. Обратно, пусть $\cos\varphi=\pm1$, тогда $|(\mathbf{x},\ \mathbf{y})| = \|\mathbf{x}\|\cdot\|\mathbf{y}\|$. Предположим, что **x** и **y** не коллинеарны, т. е. $\mathbf{y}-\alpha\mathbf{x}\neq0$ при любом $\alpha$. В этом случае $(\mathbf{y}-\alpha\mathbf{x})^2>0$, $\alpha^2\mathbf{x}^2-2\alpha(\mathbf{x},\ \mathbf{y})+\mathbf{y}^2>0$, откуда $(\mathbf{x},\ \mathbf{y})^2<\mathbf{x}^2\mathbf{y}^2$, или $|(\mathbf{x},\ \mathbf{y})|<\|\mathbf{x}\|\cdot\|\mathbf{y}\|$, что противоречит условию. Следовательно, векторы **x** и **y** коллинеарны.

Замечание. Одновременно доказано, что равенство

$$|(\mathbf{x},\ \mathbf{y})| = \|\mathbf{x}\|\cdot\|\mathbf{y}\| \tag{7.24}$$

выполняется тогда и только тогда, когда **x** и **y** коллинеарны. Другими словами, в формуле (7.21) равенство достигается лишь в случае коллинеарности векторов **x** и **y**.

**Ортонормированный базис.** Система векторов $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ называется *ортогональной,* если эти векторы попарно ортогональны, т. е. $(\mathbf{a}_i,\ \mathbf{a}_k)=0$ при $i\neq k$.

**Теорема 7.6.** Ортогональная система ненулевых векторов линейно независима.

Доказательство. Пусть $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ — ортогональная система ненулевых векторов. Допустим, что эта система линейно зависима, тогда существуют числа

$\alpha_1, \alpha_2, \ldots, \alpha_n$, не все равные нулю, такие, что

$$\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n=\mathbf{0}.$$

Предположим, что $\alpha_i \neq 0$, тогда скалярное произведение вектора $\mathbf{a}_i$ и линейной комбинации $\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n$ равно нулю:

$$(\mathbf{a}_i, \alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2+\ldots+\alpha_n\mathbf{a}_n)=0,$$

или

$$\alpha_1(\mathbf{a}_i, \mathbf{a}_1)+\alpha_2(\mathbf{a}_i, \mathbf{a}_2)+\ldots+\alpha_i(\mathbf{a}_i, \mathbf{a}_i)+\ldots+\alpha_n(\mathbf{a}_i, \mathbf{a}_n)=0.$$

Поскольку $(\mathbf{a}_i, \mathbf{a}_k)=0$ при $i \neq k$, то $\alpha_i(\mathbf{a}_i, \mathbf{a}_i)=0$, или $\alpha_i\|\mathbf{a}_i\|^2=0$, откуда $\|\mathbf{a}_i\|=0$, т. е. $\mathbf{a}_i$ — нулевой вектор, что противоречит условию. Следовательно, ортогональная система ненулевых векторов линейно независима.

Вектор $\mathbf{a}$ называется *нормированным*, или *единичным*, если $\|\mathbf{a}\|=1$. Если $\mathbf{a}$ — ненулевой вектор, то каждый из векторов

$$\mathbf{a}_1^0=\frac{\mathbf{a}}{\|\mathbf{a}\|}, \quad \mathbf{a}_2^0=-\frac{\mathbf{a}}{\|\mathbf{a}\|} \tag{7.25}$$

будет нормированным. Нахождение для данного вектора нормированного вектора по формулам (7.25) называется нормированием данного вектора, а множитель $\mu=\dfrac{1}{\pm\|\mathbf{a}\|}$ — нормирующим множителем.

Система векторов $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ называется ортонормированной, если она ортогональна и каждый вектор является нормированным, т. е.

$$(\mathbf{e}_i, \mathbf{e}_k)=\begin{cases}0, & \text{при } i \neq k,\\ 1, & \text{при } i=k,\end{cases} \tag{7.26}$$

где $i, k=1, 2, \ldots, n$.

Базис $n$-мерного евклидова пространства называется *ортонормированным,* если базисные векторы образуют ортонормированную систему. Приведем без доказательства теорему о возможности выбора ортонормированного базиса в евклидовом пространстве.

**Теорема 7.7.** Во всяком евклидовом $n$-мерном пространстве ($n \geqslant 2$) существует ортонормированный базис.

Отметим, что в трехмерном евклидовом пространстве $V_3$ ортонормированным базисом является базис $\mathbf{i}, \mathbf{j}, \mathbf{k}$, где $\mathbf{i}, \mathbf{j}, \mathbf{k}$ — орты.

Замечание. Если $e_1, e_2, \ldots, e_n$ — ортогональный базис $n$-мерного евклидова пространства, то любой вектор $\mathbf{x}$ этого пространства можно представить в виде

$$\mathbf{x}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+\ldots+x_n\mathbf{e}_n,$$

откуда

$$x_k=\frac{(\mathbf{x},\ \mathbf{e}_k)}{(\mathbf{e}_k,\ \mathbf{e}_k)}\quad (k=1,\ 2,\ \ldots,\ n).$$

Последняя формула упрощается в случае ортонормированного базиса; при этом $(\mathbf{e}_k,\ \mathbf{e}_k)=1$.

**Выражение скалярного произведения через координаты векторов в ортонормированном базисе.** Пусть в $n$-мерном евклидовом пространстве фиксирован ортонормированный базис $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ и даны векторы этого пространства

$$\mathbf{x}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+\ldots+x_n\mathbf{e}_n,$$
$$\mathbf{y}=y_1\mathbf{e}_1+y_2\mathbf{e}_2+\ldots+y_n\mathbf{e}_n.$$

Найдем скалярное произведение этих векторов. Принимая во внимание аксиомы скалярного произведения и формулу (7.26), получаем

$$(\mathbf{x},\ \mathbf{y})=(x_1\mathbf{e}_1+x_2\mathbf{e}_2+\ldots+x_n\mathbf{e}_n,\ y_1\mathbf{e}_1+y_2\mathbf{e}_2+\ldots+y_n\mathbf{e}_n)=$$
$$=\left(x_1\mathbf{e}_1,\ \sum_{i=1}^{n}y_i\mathbf{e}_i\right)+\left(x_2\mathbf{e}_2,\ \sum_{i=1}^{n}y_i\mathbf{e}_i\right)+\ldots+\left(x_n\mathbf{e}_n,\ \sum_{i=1}^{n}y_i\mathbf{e}_i\right)=$$
$$=(x_1\mathbf{e}_1,\ y_1\mathbf{e}_1)+(x_2\mathbf{e}_2,\ y_2\mathbf{e}_2)+\ldots+(x_n\mathbf{e}_n,\ y_n\mathbf{e}_n)=$$
$$=x_1y_1(\mathbf{e}_1,\ \mathbf{e}_1)+x_2y_2(\mathbf{e}_2,\ \mathbf{e}_2)+\ldots+x_ny_n(\mathbf{e}_n,\mathbf{e}_n)=x_1y_1+$$
$$+x_2y_2+\ldots+x_ny_n,$$
$$(\mathbf{x},\ \mathbf{y})=x_1y_1+x_2y_2+\ldots+x_ny_n.$$

Итак, скалярное произведение двух векторов, заданных координатами в ортонормированном базисе, равно сумме произведений одноименных координат.

Очевидно,

$$\|\mathbf{x}\|=\sqrt{(\mathbf{x},\ \mathbf{x})}=\sqrt{x_1^2+x_2^2+\ldots+x_n^2},$$

где $x_1, x_2, \ldots, x_n$ — координаты вектора $\mathbf{x}$ в ортонормированном базисе.

## § 7.8. Линейное преобразование и его матрица

Если указано правило, по которому каждому вектору $\mathbf{x}$ линейного пространства $V$ ставится в соответствие единственный вектор $\mathbf{y}$ этого пространства, то будем говорить, что в нем задано преобразование (отображение, оператор) $f$ или задано преобразование $f$ пространства $V$ в себя, и писать $f: V \to V$. Говорят также, что преобразование $f$ переводит вектор $\mathbf{x}$ в вектор $\mathbf{y}$, и пишут $\mathbf{y}=f(\mathbf{x})$. Вектор $\mathbf{y}$ называют образом вектора $\mathbf{x}$, а $\mathbf{x}$ — прообразом вектора $\mathbf{y}$.

Замечание. Из сказанного следует, что каждый вектор имеет единственный образ, но не следует, что каждый вектор имеет прообраз или этот прообраз единственный. Преобразование, при котором каждый вектор имеет единственный прообраз, называется взаимно-однозначным (или биективным).

Преобразование $f$ линейного пространства $V$ называется *линейным преобразованием (линейным оператором)*, если для любых векторов этого пространства $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}$ и любого действительного числа $\lambda$ выполняются условия:

$$1)\ f(\mathbf{x}_1+\mathbf{x}_2)=f(\mathbf{x}_1)+f(\mathbf{x}_2);\quad 2)\ f(\lambda\mathbf{x})=\lambda f(\mathbf{x}).$$

(Если рассматривается комплексное пространство, то $\lambda$ — любое комплексное число.)

Из этих условий следует, что

$$f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2)=\alpha f(\mathbf{x}_1)+\beta f(\mathbf{x}_2), \tag{7.27}$$

где $\alpha$, $\beta$ — любые числа (действительные или комплексные). Обратно, из равенства (7.27) следуют условия 1) и 2). Итак, линейное преобразование (линейный оператор) определяется равенством (7.27).

Отметим, что линейное преобразование переводит нулевой вектор в нулевой, так как согласно условию 2) $f(\mathbf{0})=f(0\mathbf{x})=0f(\mathbf{x})=\mathbf{0}$.

Простейшим примером линейного преобразования является тождественное преобразование, или преобразование $f(\mathbf{x})=\mathbf{x}$, т. е. преобразование, которое каждому вектору линейного пространства ставит в соответствие тот же вектор.

Пусть линейное преобразование $f$ $n$-мерного пространства переводит базисные векторы $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ соответственно в векторы $\mathbf{e}'_1$, $\mathbf{e}'_2$, ..., $\mathbf{e}'_n$, т. е.

$$\mathbf{e}'_1=f(\mathbf{e}_1),\ \mathbf{e}'_2=f(\mathbf{e}_2),\ \ldots,\ \mathbf{e}'_n=f(\mathbf{e}_n).$$

Образ любого вектора $\mathbf{x}$ данного линейного пространства можно выразить через образы базисных векторов $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$. Действительно, если $\mathbf{x}=\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\ldots+\alpha_n\mathbf{e}_n$,

то $f(\mathbf{x}) = f(\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n) = \alpha_1 f(\mathbf{e}_1) + \alpha_2 f(\mathbf{e}_2) + \ldots$
$\ldots + \alpha_n f(\mathbf{e}_n) = \alpha_1\mathbf{e}_1' + \alpha_2\mathbf{e}_2' + \ldots + \alpha_n\mathbf{e}_n'$, т. е.

$$f(\mathbf{x}) = \alpha_1\mathbf{e}_1' + \alpha_2\mathbf{e}_2' + \ldots + \alpha_n\mathbf{e}_n'.$$

Следовательно, линейное преобразование будет вполне определено, если заданы образы базисных векторов рассматриваемого пространства.

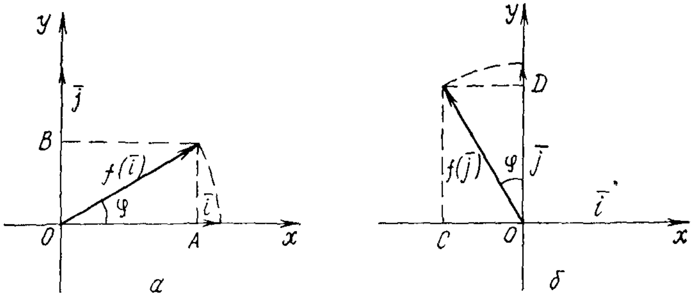



Рис. 7.3

Пусть $f$ — линейное преобразование $n$-мерного линейного пространства, переводящее базисные векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ в векторы $\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n'$. Каждый из последних векторов разложим по базису:

$$\begin{aligned}
\mathbf{e}_1' &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + \ldots + a_{n1}\mathbf{e}_n,\\
\mathbf{e}_2' &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + \ldots + a_{n2}\mathbf{e}_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\mathbf{e}_n' &= a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + \ldots + a_{nn}\mathbf{e}_n.
\end{aligned}$$

Матрица

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix},$$

в которой $k$-й столбец состоит из координат вектора $\mathbf{e}_k'$ ($k = 1, 2, \ldots, n$), называется *матрицей линейного преобразования* $f$ в базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$; ранг $r$ матрицы $A$ называется *рангом преобразования* $f$, а число $n-r$ — *дефектом* этого преобразования. Итак, каждому линейному преобразованию $n$-мерного линейного пространства соответствует матрица порядка $n$ в данном базисе; обратно, каждой матрице порядка $n$ соответствует линейное преобразование $n$-мерного пространства.

Отметим, что матрица тождественного преобразования в любом базисе будет единичной; обратно, любой единичной матрице $n$-го порядка соответствует тождественное преобразование линейного $n$-мерного пространства.

Пример. В пространстве $V_2$ всех свободных векторов на плоскости определим преобразование поворота всех векторов вокруг начала координат на угол $\varphi$. Каждому вектору $\mathbf{x}$ (рис. 7.2) этой плоскости ставим в соответствие вектор $\mathbf{y}=f(\mathbf{x})$, полученный вращением вектора $\mathbf{x}$ на один и тот же угол $\varphi$. Это преобразование является линейным, поскольку условия 1) и 2), определяющие линейное преобразование, будут выполнены. Найдем матрицу этого линейного преобразования в базисе $\mathbf{i}$, $\mathbf{j}$ (рис. 7.3). Так как $f(\mathbf{i})=\overline{OA}+\overline{OB}=\mathbf{i}\cos\varphi+$ $+\mathbf{j}\sin\varphi$, $f(\mathbf{j})=\overline{OC}+\overline{OD}=-\mathbf{i}\sin\varphi+\mathbf{j}\cos\varphi$, то

$$A=\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}.$$

## § 7.9. Линейное преобразование в координатах

Рассмотрим линейное преобразование $f$ $n$-мерного линейного пространства, заданное в некотором базисе $\mathbf{e}_1$, $\mathbf{e}_2, \ldots, \mathbf{e}_n$ матрицей

$$A=\begin{bmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{bmatrix}. \tag{7.28}$$

Координаты вектора $\mathbf{x}$ и его образа $\mathbf{y}=f(\mathbf{x})$ известны:

$$\begin{aligned}\mathbf{x}&=x_1\mathbf{e}_1+x_2\mathbf{e}_2+\ldots+x_n\mathbf{e}_n,\\ f(\mathbf{x})&=y_1\mathbf{e}_1+y_2\mathbf{e}_2+\ldots+y_n\mathbf{e}_n.\end{aligned} \tag{7.29}$$

Найдем зависимость между координатами векторов $\mathbf{x}$ и $\mathbf{y}$. Принимая во внимание определение матрицы линейного преобразования (см. § 7.8), получаем

$$\begin{aligned}f(\mathbf{x})&=f(x_1\mathbf{e}_1+x_2\mathbf{e}_2+\ldots+x_n\mathbf{e}_n)=x_1f(\mathbf{e}_1)+x_2f(\mathbf{e}_2)+\ldots\\ &\ldots+x_nf(\mathbf{e}_n)=x_1(a_{11}\mathbf{e}_1+a_{21}\mathbf{e}_2+\ldots+a_{n1}\mathbf{e}_n)+\\ &+x_2(a_{12}\mathbf{e}_1+a_{22}\mathbf{e}_2+\ldots+a_{n2}\mathbf{e}_n)+\ldots\\ &\ldots+x_n(a_{1n}\mathbf{e}_1+a_{2n}\mathbf{e}_2+\ldots+a_{nn}\mathbf{e}_n),\end{aligned}$$

или

$$f(\mathbf{x}) = (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)\mathbf{e}_1 + (a_{21}x_1 + a_{22}x_2 + \dots$$
$$\dots + a_{2n}x_n)\mathbf{e}_2 + \dots + (a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n)\mathbf{e}_n.$$

Поскольку каждый вектор разлагается по базису единственным образом, то из этого равенства и второго из равенств (7.29) следуют искомые формулы:

$$\left.\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ &\dots\dots\dots\dots\dots\dots\dots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n. \end{aligned}\right\} \qquad (7.30)$$

Формулы (7.30) можно записать в матричном виде

$$Y = AX, \qquad (7.31)$$

где $A$ определяется формулой (7.28), а $X$ и $Y$ — формулами

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}.$$

Если переменные $y_1, y_2, \dots, y_n$ связаны с переменными $x_1, x_2, \dots, x_n$ формулами (7.30), будем говорить, что задано линейное однородное преобразование переменных с матрицей $A$, переводящее переменные $x_1, x_2, \dots, x_n$ в переменные $y_1, y_2, \dots, y_n$. Оно обладает теми же свойствами, что и линейное преобразование $n$-мерного линейного пространства. Линейное однородное преобразование переменных (7.30) или (7.31) называется *невырожденным,* если $\det A \neq 0$.

Замечание. При рассмотрении линейных преобразований (линейных операторов) пользуются и другими обозначениями. Если $\mathbf{y} = f(\mathbf{x})$, где $f$ — линейное преобразование (линейный оператор) с матрицей $A$ в некотором базисе, то пишут $\mathbf{y} = A\mathbf{x}$. Условия 1) и 2), определяющие линейное преобразование, можно записать в виде

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2, \quad A(\lambda\mathbf{x}) = \lambda A\mathbf{x}.$$

## § 7.10. Зависимость между матрицами одного и того же преобразования в различных базисах. Подобные матрицы

В $n$-мерном линейном пространстве фиксируем два базиса: $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ и $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$; первый из них назовем старым, второй — новым. Предположим, что известно преобразование, переводящее старый базис в новый.

**Теорема 7.8.** Если $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ и $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ — два базиса линейного пространства, $A$—матрица линейного преобразования в старом базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, то матрица $B$ этого преобразования в новом базисе $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ имеет вид $B=T^{-1}AT$, где $T$ — матрица перехода от старого базиса к новому.

Доказательство. Координаты векторов $\mathbf{x}$, $\mathbf{y}$ в старом базисе обозначим соответственно $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$, а в новом базисе — теми же буквами со штрихом, т. е.

$$\mathbf{x}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+\ldots+x_n\mathbf{e}_n, \quad \mathbf{y}=y_1\mathbf{e}_1+y_2\mathbf{e}_2+\ldots+y_n\mathbf{e}_n,$$

$$\mathbf{x}=x'_1\mathbf{e}'_1+x'_2\mathbf{e}'_2+\ldots+x'_n\mathbf{e}'_n, \quad \mathbf{y}=y'_1\mathbf{e}'_1+y'_2\mathbf{e}'_2+\ldots+y'_n\mathbf{e}'_n.$$

На основании теоремы 7.5 (см. § 7.6) получаем $X=TX'$, $Y=TY'$, а в соответствии с формулой (7.31) имеем $Y=AX$, $Y'=BX'$. Умножив равенство $X=TX'$ слева на матрицу $A$, получим $AX=ATX'$, или $TY'=ATX'$, откуда $Y'=T^{-1}ATX'$. Сравнивая это равенство с равенством $Y'=BX'$, находим доказываемую формулу

$$B=T^{-1}AT. \qquad (7.32)$$

Следствие. Если линейное преобразование имеет невырожденную матрицу в некотором базисе, то матрица этого преобразования будет невырожденной в любом другом базисе.

Действительно, пусть $A$ и $B$ — матрицы линейного преобразования $f$ в двух различных базисах, причем $\det A\neq 0$. Так как $B=T^{-1}AT$, где $T$ — невырожденная матрица (см. § 7.6), то $\det B=\det(T^{-1}AT)=\det T^{-1}\times\det A\cdot\det T\neq 0$.

Пример. В базисе $\mathbf{e}_1, \mathbf{e}_2$ преобразование $f$ имеет матрицу

$$A=\begin{bmatrix} 8 & 4 \\ 5 & 2 \end{bmatrix}.$$

Найти матрицу преобразования $f$ в базисе $\mathbf{e}'_1=2\mathbf{e}_1+\mathbf{e}_2$, $\mathbf{e}'_2=6\mathbf{e}_1+4\mathbf{e}_2$.

Так как

$$T=\begin{bmatrix} 2 & 6 \\ 1 & 4 \end{bmatrix}, \quad T^{-1}=\frac{1}{2}\begin{bmatrix} 4 & -6 \\ -1 & 2 \end{bmatrix}=\begin{bmatrix} 2 & -3 \\ -0{,}5 & 1 \end{bmatrix},$$

то

$$B=T^{-1}AT=\begin{bmatrix} 2 & -3 \\ -0,5 & 1 \end{bmatrix}\cdot\begin{bmatrix} 8 & 4 \\ 5 & 2 \end{bmatrix}\cdot\begin{bmatrix} 2 & 6 \\ 1 & 4 \end{bmatrix}=\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}\cdot\begin{bmatrix} 2 & 6 \\ 1 & 4 \end{bmatrix}=$$

$$=\begin{bmatrix} 4 & 14 \\ 2 & 6 \end{bmatrix}.$$

Введем понятие подобия матриц. Матрица $B$ называется *подобной* матрице $A$, если существует невырожденная квадратная матрица $C$, удовлетворяющая равенству $B=C^{-1}AC$.

Из теоремы 7.8 следует, что матрица $B$ линейного преобразования $f$ в базисе $\mathbf{e}_1'$, $\mathbf{e}_2'$, $\mathbf{e}_n'$, ..., $\mathbf{e}_n'$ подобна матрице $A$ того же преобразования $f$ в другом базисе $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$, так как эти матрицы связаны формулой $B=C^{-1}AC$, где $C$ — невырожденная матрица перехода от базиса $\mathbf{e}_1$, $e_2$, ..., $e_n$ к базису $\mathbf{e}_1'$, $\mathbf{e}_2'$, ..., $\mathbf{e}_n'$.

Можно доказать, что две квадратные матрицы $A$ и $B$ порядка $n$ тогда и только тогда являются матрицами одного и того же линейного преобразования пространства $V_n$ в соответствующих базисах, когда матрица $B$ подобна матрице $A$.

## § 7.11. Характеристическое уравнение линейного преобразования

**Теорема 7.9.** Если линейное преобразование $f$ в базисе $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ имеет матрицу $A$ и в базисе $\mathbf{e}_1'$, $\mathbf{e}_2'$, ..., $\mathbf{e}_n'$ матрицу $B$, то $\det(A-\lambda E)=\det(B-\lambda E)$, где $\lambda$ — любое действительное число; $E$ — единичная матрица $n$-го порядка.

Доказательство. Обозначим через $T$ матрицу перехода от базиса $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ к базису $\mathbf{e}_1'$, $\mathbf{e}_2'$, ..., $\mathbf{e}_n'$, тогда $B=T^{-1}AT$ (см. § 7.10). Следовательно, $\det(B-\lambda E)=$ $=\det(T^{-1}AT-\lambda E)=\det(T^{-1}AT-\lambda T^{-1}ET)=\det(T^{-1}\times$ $\times(A-\lambda E)T)=\det T^{-1}\cdot\det(A-\lambda E)\cdot\det T=\det(A-\lambda E)$, так как $\det T^{-1}\cdot\det T=1$. Итак, доказано, что

$$\det(B-\lambda E)=\det(A-\lambda E). \tag{7.33}$$

Отметим, что $\det(A-\lambda E)$ является многочленом степени $n$ относительно $\lambda$. Многочлен $\det(A-\lambda E)$ называется *характеристическим многочленом* матрицы $A$, или характеристическим многочленом линейного преобразования $f$.

Замечание. Равенство (7.33) означает, что характеристический многочлен линейного преобразования остается неизменным при переходе к новому базису; матрица линейного преобразования меняется.

*Характеристическим уравнением* линейного преобразования называется уравнение

$$\det(A-\lambda E)=0, \qquad (7.34)$$

где $A$ — матрица этого преобразования в некотором базисе. Очевидно, характеристическое уравнение не зависит от выбора базиса. Уравнение (7.34) называют также характеристическим уравнением матрицы $A$. Корни уравнения (7.34) называются *характеристическими числами* линейного преобразования $f$, или характеристическими числами матрицы $A$.

Если линейное преобразование $f$ в некотором базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ имеет квадратную матрицу $n$-го порядка $A=(a_{ik})$, то характеристическое уравнение (7.34) запишется так:

$$\det\left[\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{bmatrix}\right] = 0,$$

или

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix} = 0. \qquad (7.35)$$

Левая часть равенства (7.35) является характеристическим многочленом матрицы $A$; обозначим его $P_n(\lambda)$, тогда характеристическое уравнение (7.35) примет вид $P_n(\lambda)=0$.

Пример. Найти характеристический многочлен и характеристические числа матрицы

$$A=\begin{bmatrix} 4 & -1 & -2 \\ 2 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}.$$

В соответствии с определением характеристического многочлена получаем

$$P_n(\lambda) = \begin{vmatrix} 4-\lambda & -1 & -2 \\ 2 & 1-\lambda & -2 \\ 1 & -1 & 1-\lambda \end{vmatrix},$$

$$P_n(\lambda) = (4-\lambda)(1-\lambda)^2+4+2+2(1-\lambda)+2(1-\lambda)-2(4-\lambda) = \\ = -\lambda^3+6\lambda^2-11\lambda+6.$$

Приравнивая этот многочлен нулю, находим характеристическое уравнение $-\lambda^3+6\lambda^2-11\lambda+6=0$, или $\lambda^3-6\lambda^2+11\lambda-6=0$. Разлагая левую часть этого уравнения на множители $\lambda^3-6\lambda^2+11\lambda-6=\lambda^3-\lambda^2-$ $-5\lambda^2+5\lambda+6\lambda-6 = \lambda^2(\lambda-1)-5\lambda(\lambda-1)+6(\lambda-1)=(\lambda-1)(\lambda^2-5\lambda+6)$, приводим данное уравнение к виду $(\lambda-1)(\lambda^2-5\lambda+6)=0$, откуда $\lambda_1=1$, $\lambda_2=2$, $\lambda_3=3$. Эти корни — характеристические числа данной матрицы.

## § 7.12. Собственные векторы линейного преобразования

Ненулевой вектор **x** линейного пространства называется *собственным вектором* линейного преобразования $f$ этого пространства, если существует число $k$ такое, что

$$f(\mathbf{x}) = k\mathbf{x}, \tag{7.36}$$

причем $k$ — действительное число для действительного линейного пространства и $k$ — комплексное число в случае комплексного пространства. Число $k$ называется *собственным значением* вектора **x** относительно преобразования $f$. Равенство (7.36) можно записать в матричном виде

$$AX=kX, \tag{7.37}$$

где $A$ — матрица преобразования $f$ в некотором базисе; $X$ — матрица-столбец из координат собственного вектора **x** в том же базисе. Ненулевая матрица-столбец $X$, удовлетворяющая уравнению (7.37), называется собственным вектором-столбцом матрицы $A$ с собственным значением $k$.

Собственные векторы и собственные значения обладают следующими свойствами:

1. Собственный вектор линейного преобразования имеет единственное значение $k$.

2. Если **x** — собственный вектор линейного преобразования $f$ с собственным числом $k$ и $\lambda$ — любое отличное от нуля число, то $\lambda$**x** — также собственный вектор преобразования $f$ с собственным значением $k$.

3. Если **x** и **y** — линейно независимые собственные

векторы линейного преобразования $f$ с одним и тем же собственным значением $k$, то $\mathbf{x}+\mathbf{y}$ — также собственный вектор этого преобразования с собственным значением $k$.

4. Если $\mathbf{x}$ и $\mathbf{y}$ — собственные векторы линейного преобразования $f$ с собственными числами $k$ и $m$, причем $k \neq m$, то $\mathbf{x}$ и $\mathbf{y}$ — линейно независимы.

Докажем эти свойства. 1. Пусть $k$ и $m$ — собственные значения собственного вектора $\mathbf{x}$ относительно линейного преобразования $f$, тогда $f(\mathbf{x}) = k\mathbf{x}$, $f(\mathbf{x}) = m\mathbf{x}$, откуда $k\mathbf{x} = m\mathbf{x}$, $(k-m)\mathbf{x} = \mathbf{0}$. Поскольку $\mathbf{x} \neq \mathbf{0}$, то $k-m=0$, или $k=m$. 2. Если $\mathbf{x}$ — собственный вектор с собственным значением $k$, то $f(\lambda\mathbf{x}) = \lambda f(\mathbf{x}) = \lambda(k\mathbf{x}) = k(\lambda\mathbf{x})$, или $f(\lambda\mathbf{x}) = k(\lambda\mathbf{x})$. Это равенство означает, что $\lambda\mathbf{x}$ — собственный вектор линейного преобразования $f$ с собственным значением $k$. 3. Если векторы $\mathbf{x}$ и $\mathbf{y}$ линейно независимы, то $\mathbf{x}+\mathbf{y}$ — ненулевой вектор и $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) = k\mathbf{x} + k\mathbf{y} = k(\mathbf{x}+\mathbf{y})$, или $f(\mathbf{x}+\mathbf{y}) = k(\mathbf{x}+\mathbf{y})$; $(\mathbf{x}+\mathbf{y})$ — собственный вектор с собственным значением $k$. 4. Предположим противное, т. е. векторы $\mathbf{x}$ и $\mathbf{y}$ линейно зависимы, тогда $\mathbf{y} = \lambda\mathbf{x}$, причем $\lambda \neq 0$, так как $\mathbf{y} \neq \mathbf{0}$. Согласно свойству 2 вектор $\lambda\mathbf{x}$ является собственным вектором с собственным числом $k$. Учитывая свойство 1, из равенства $\mathbf{y} = \lambda\mathbf{x}$ заключаем, что $k=m$, а это противоречит условию. Следовательно, векторы $\mathbf{x}$ и $\mathbf{y}$ — линейно независимы. Отметим, что свойство 4 справедливо и для $m$ $(m > 2)$ векторов, т. е. собственные векторы линейного преобразования с попарно различными собственными числами линейно независимы.

Следствие. Если $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ — линейно независимые собственные векторы линейного преобразования $f$ с одним и тем же собственным значением $k$, то любая нетривиальная линейная комбинация этих векторов есть собственный вектор этого преобразования с собственным значением $k$.

Это утверждение следует из свойств 2 и 3.

**Теорема 7.10.** В комплексном линейном пространстве все корни характеристического уравнения и только они являются собственными значениями линейного преобразования.

Доказательство. Если ненулевой вектор $\mathbf{x}$ будет собственным вектором линейного преобразования $f$ комплексного линейного пространства, то по определению $AX = kX$, или $(A-kE)X = 0$, где $A$ — матрица линейного преобразования $f$ в некотором базисе, $E$ — единичная матрица, $X$ — матрица-столбец из координат вектора

$\mathbf{x}$ в том же базисе. Пусть $A=(a_{ik})_{nn}$, тогда уравнение $(A-kE)X=0$ представляет матричную запись однородной системы уравнений

$$\left.\begin{aligned} &(a_{11}-k)x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ &a_{21}x_1+(a_{22}-k)x_2+\ldots+a_{2n}x_n=0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &a_{n1}x_1+a_{n2}x_2+\ldots+(a_{nn}-k)x_n=0 \end{aligned}\right\} \qquad (7.38)$$

относительно координат $x_1, x_2, \ldots, x_n$ собственного вектора $\mathbf{x}\neq\mathbf{0}$. Эта система имеет ненулевое решение тогда и только тогда, когда ее определитель равен нулю (см. следствие из теоремы Крамера), т. е.

$$\det(A-kE)=0, \quad \begin{vmatrix} a_{11}-k & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22}-k & \ldots & a_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn}-k \end{vmatrix}=0. \qquad (7.39)$$

Это означает, что число $k$ — корень характеристического уравнения. Обратно, если $k$ — корень характеристического уравнения (7.39), то система (7.38) или матричное уравнение $(A-kE)X=0$ имеет ненулевое решение $x_1, x_2, \ldots, x_n$, причем $AX=kX$ или $f(\mathbf{x})=k\mathbf{x}$, т. е. $\mathbf{x}(x_1, x_2, \ldots, x_n)$ — собственный вектор.

Замечание 1. Уравнение (7.39) является алгебраическим уравнением $n$-й степени относительно $k$. Такое уравнение имеет ровно $n$ корней, считая равные и комплексные. Среди корней этого уравнения может и не оказаться действительных.

Замечание 2. Собственными значениями линейного преобразования действительного пространства будут только действительные корни характеристического уравнения.

Собственные значения линейного преобразования называются также собственными значениями матрицы этого преобразования. Собственное значение называется $m$-кратным, если оно является $m$-кратным корнем характеристического уравнения.

## § 7.13. Собственные значения и собственные векторы симметрической матрицы

**Теорема 7.11.** Корни характеристического уравнения действительной симметрической матрицы являются действительными числами.

Доказательство. Если $A$ — действительная симметрическая матрица, то $A^T=A$ и $\bar{A}=A$, где $\bar{A}$ — матрица, сопряженная матрице $\bar{A}$; $A^T$ — матрица, полученная транспонированием матрицы $A$. Пусть $k$ — корень характеристического уравнения, $\mathbf{x}(x_1, x_2, \ldots, x_n)$ — собственный вектор данной квадратной матрицы порядка $n$. Найдем произведение $X^T A \bar{X}$, где $X$ — матрица-столбец из координат вектора $\mathbf{x}(x_1, x_2, \ldots, x_n)$, которые могут быть и комплексными числами; $\bar{X}$ — матрица, сопряженная матрице $X$; $X^T$ — матрица-строка, полученная транспонированием матрицы $X$. Применяя свойства транспонированных и сопряженных матриц (см. § 3.2 и 3.7), получаем

$$X^T A\overline{X} = X^T \overline{\overline{A}X} = X^T \overline{AX} = X^T \overline{kX} = X^T \bar{k}\overline{X} = \bar{k}(X^T \overline{X}),$$

$$X^T A\overline{X} = X^T A^T \overline{X} = (AX)^T \overline{X} = (kX)^T \overline{X} = k(X^T \overline{X}).$$

Из этих равенств следует, что

$$\bar{k}(X^T\bar{X}) = k(X^T\bar{X}), \text{ или } (\bar{k}-k)(X^T\bar{X}) = 0.$$

Можно утверждать, что $X^T\bar{X} \neq 0$. Действительно,

$$X^T \overline{X} = [x_1 x_2 \ \ldots \ x_n] \cdot \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \ldots \\ \bar{x}_n \end{bmatrix} = [x_1\bar{x}_1 + x_2\bar{x}_2 + \ldots + x_n\bar{x}_n],$$

или $X^T\bar{X} = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2$, где $|x_i|$ — модуль числа $x_i$ $(i = 1, 2, \ldots, n)$. Поскольку $X$ — ненулевой столбец (см. определение собственного вектора), то $X^T\bar{X} \neq 0$. Из равенства $(\bar{k}-k)(X^T\bar{X}) = 0$ получаем $\bar{k}-k = 0$, или $\bar{k} = k$, т. е. $k$ — действительное число.

Следствие. Действительная симметрическая матрица имеет только действительные собственные векторы.

Система (7.38) для определения координат собственного вектора в этом случае имеет только действительные решения, так как $a_{ij}$ и $k$ — действительные числа.

**Теорема 7.12.** Собственные векторы действительной симметрической матрицы, соответствующие различным собственным значениям, ортогональны.

Доказательство. Пусть $\mathbf{x}(x_1, x_2, \ldots, x_n)$, $\mathbf{y}(y_1, y_2, \ldots, y_n)$ — собственные векторы симметрической матрицы $A$ порядка $n$, собственные значения которых $k$ и $m$ различны, $X$, $Y$ — соответствующие векторы-столбцы из их координат. Найдем произведение $X^T A Y$, пользуясь ассоциативным свойством умножения матриц:

$$X^TAY = X^T(AY) = X^T(mY) = mX^TY,$$

$$X^TAY = (X^TA)Y = (X^TA^T)Y = (AX)^TY = (kX)^TY = k(X^TY).$$

Из этих двух равенств следует, что

$$mX^TY = kX^TY, \text{ или } (m-k)X^TY = 0.$$

Так как по условию $k \neq m$ и $(m-k) \neq 0$, то $X^TY = 0$, или $x_1y_1 + x_2y_2 + \ldots + x_ny_n = 0$. Это означает, что векторы **x** и **y** ортогональны.

## § 7.14. Приведение матрицы линейного преобразования к диагональному виду

**Теорема 7.13.** Матрица линейного преобразования имеет диагональный вид *тогда и только тогда, когда каждый* базисный вектор является собственным вектором этого преобразования.

Доказательство. Пусть матрица $A$ линейного преобразования в базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ имеет диагональный вид, т. е.

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}, \tag{7.40}$$

тогда по определению этой матрицы (см. § 7.8) $f(\mathbf{e}_i) = \lambda_i\mathbf{e}_i$ $(i = 1, 2, \ldots, n)$. Полученное равенство означает, что каждый базисный вектор $\mathbf{e}_i$ $(i = 1, 2, \ldots, n)$ будет собственным вектором линейного преобразования $f$. Обратно, пусть базисные векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ будут собственными векторами линейного преобразования $f$ с соответствующими собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_n$. В данном случае $f(\mathbf{e}_i) = \lambda_i\mathbf{e}_i (i = 1, 2, \ldots, n)$, поэтому матрица $A$ линейного преобразования имеет вид (7.40), т. е. является диагональной.

Матрица $A$ называется *приводимой к диагональному виду*, если существует невырожденная матрица $T$ такая, что матрица $T^{-1}AT = B$ будет диагональной. Характеристические многочлены матриц $A$ и $B = T^{-1}AT$ совпадают (см. § 7.11); характеристические числа диагональной матрицы равны ее диагональным элементам (это следует из определения, см. § 7.11, уравнение (7.35)). Следовательно, если матрица $A$ приводима к диагональному виду, то

$$B=\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ . & . & . & .\\ 0 & 0 & \dots & \lambda_n\end{bmatrix},$$

где $\lambda_1, \lambda_2, \dots, \lambda_n$ — характеристические числа матрицы $A$.

**Теорема 7.14.** Матрица $A$ линейного преобразования $f$ $n$-мерного линейного пространства приводима к диагональному виду тогда и только тогда, когда существует базис этого пространства, состоящий из собственных векторов данного преобразования.

Доказательство. Пусть линейное преобразование $f$ в базисе $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ имеет матрицу $A$, которая приводима к диагональному виду, т. е. существует невырожденная матрица $T$ такая, что $B=T^{-1}AT$ — диагональная матрица. Эту матрицу $T$ можно рассматривать как матрицу перехода от базиса $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ к базису $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$ в котором матрица $B$ является диагональной. На основании предыдущей теоремы заключаем, что $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$ — собственные векторы преобразования $f$.

Обратно, пусть базис $\mathbf{e}''_1, \mathbf{e}''_2, \dots, \mathbf{e}''_n$ данного $n$-мерного пространства состоит из собственных векторов линейного преобразования $f$. По теореме 7.13 в этом базисе матрица $D=T^{-1}AT$ — диагональная, где $T$ — матрица перехода от базиса $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ к базису $\mathbf{e}''_1, \mathbf{e}''_2, \dots, \mathbf{e}''_n$. Следовательно, матрица $A$ приводима к диагональному виду.

Замечание 1. Для построения матрицы $T$ достаточно найти собственные векторы матрицы $A$ с помощью собственных чисел $\lambda_1, \lambda_2, \dots, \lambda_n$, которые определяются из характеристического уравнения.

Замечание 2. Можно доказать, что если все собственные числа матрицы $A$ попарно различны, то матрица приводится к диагональному виду.

## § 7.15. Действия над линейными преобразованиями

**Произведение преобразований.** Рассмотрим преобразование $f$, переводящее вектор $\mathbf{x}$ в вектор $\mathbf{y}$, т. е. $\mathbf{y}=f(\mathbf{x})$. К вектору $\mathbf{y}$ применим преобразование $g$, переводящее вектор $\mathbf{y}$ в вектор $\mathbf{z}$, т. е. $\mathbf{z}=g(\mathbf{y})$. Так как $\mathbf{y}=f(\mathbf{x})$, то имеем преобразование $\mathbf{z}=g(f(\mathbf{x}))$, переводящее вектор $\mathbf{x}$ в вектор $\mathbf{z}$, причем вектор $\mathbf{z}$ получен в результате последовательного применения преобразований $f$ и $g$. Преобразование, заключающееся в последовательном применении преобразований $f$ и $g$, называется *произведением* преобразования $f$ на преобразование $g$, или *композицией* этих преобразований, и обозначается $g \circ f$ (или просто $gf$); отметим, что справа записывается первое преобразование. Таким образом,

$$g \circ f(\mathbf{x})=g(f(\mathbf{x})). \tag{7.41}$$

Докажем, что произведение линейных преобразований является линейным произведением. Пусть $f$ и $g$ — линейные преобразования некоторого линейного пространства. В соответствии с определением (см. формулу (7.27)) для любых векторов $\mathbf{x}_1$, $\mathbf{x}_2$ этого пространства

$$f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2)=\alpha f(\mathbf{x}_1)+\beta f(\mathbf{x}_2),\ g(\alpha\mathbf{x}_1+\beta\mathbf{x}_2)= \\ =\alpha g(\mathbf{x}_1)+\beta g(\mathbf{x}_2),$$

где $\alpha$ и $\beta$ — любые числа. Принимая во внимание формулу (7.41), получаем

$$g \circ f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2)=g(f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2))=g(\alpha f(\mathbf{x}_1)+\beta f(\mathbf{x}_2))= \\ =\alpha g(f(\mathbf{x}_1))+\beta g(f(\mathbf{x}_2))=\alpha g \circ f(\mathbf{x}_1)+\beta g \circ f(\mathbf{x}_2),$$

а это означает, что $g \circ f$ — линейное преобразование.

**Теорема 7.15.** Если в некотором базисе линейные преобразования $f$ и $g$ имеют соответственно матрицы $A$ и $B$, то их произведение $g \circ f$ в том же базисе имеет матрицу $BA$.

Доказательство. Пусть

$$\mathbf{y}=f(\mathbf{x}),\ \mathbf{z}=g(\mathbf{y}),\ \mathbf{z}=g \circ f(\mathbf{x}),$$

тогда (см. формулу (7.31))

$$Y=AX,\ Z=BY,\ Z=CX,$$

где $C$ — матрица преобразования $g \circ f$ в рассматриваемом базисе; $X$, $Y$ — матрицы-столбцы соответственно из координат векторов $\mathbf{x}$, $\mathbf{y}$.

Уравнения $Z=BY$, $Y=AX$ приводят к уравнению $Z=$ $=BAX$. Сравнивая это уравнение с уравнением $Z=CX$, заключаем, что $C=BA$.

**Сумма преобразований.** Суммой преобразований $f$ и $g$

некоторого пространства называется преобразование $h$ такое, что для любого вектора $\mathbf{x}$ этого пространства

$$h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}). \tag{7.42}$$

Сумму преобразований $f$ и $g$ будем обозначать $f+g$. Очевидно, $f+g=g+f$.

Докажем, что сумма линейных преобразований является линейным преобразованием. Пусть $f$ и $g$ — линейные преобразования некоторого пространства, т. е. для любых двух векторов $\mathbf{x}_1$, $\mathbf{x}_2$ этого пространства

$$f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) = \alpha f(\mathbf{x}_1) + \beta f(\mathbf{x}_2),\quad g(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) = \\ = \alpha g(\mathbf{x}_1) + \beta g(\mathbf{x}_2),$$

где $\alpha$ и $\beta$ — произвольные числа. Для преобразования $f+g$ имеем $(f+g)(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) = h(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) = f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) + g(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) = \alpha f(\mathbf{x}_1)+\beta f(\mathbf{x}_2) + \alpha g(\mathbf{x}_1)+\beta g(\mathbf{x}_2) = \alpha(f(\mathbf{x}_1)+g(\mathbf{x}_1)) + \beta(f(\mathbf{x}_2)+g(\mathbf{x}_2)) = \alpha h(\mathbf{x}_1)+\beta h(\mathbf{x}_2)$, или $h(\alpha\mathbf{x}_1+\beta\mathbf{x}_2) = \alpha h(\mathbf{x}_1)+\beta h(\mathbf{x}_2)$. Последнее равенство означает, что $h=f+g$ — линейное преобразование.

**Теорема 7.16.** Если линейные преобразования $f$ и $g$ в некотором базисе имеют соответственно матрицы $A$ и $B$, то преобразование $f+g$ в том же базисе имеет матрицу $A+B$.

Доказательство. Пусть

$$\mathbf{y}=f(\mathbf{x}),\ \mathbf{z}=g(\mathbf{x}),\ \mathbf{y}+\mathbf{z}=(f+g)\mathbf{x}=h(\mathbf{x}),$$

тогда

$$Y=AX,\ Z=BX,\ Y+Z=CX,$$

где $C$ — матрица преобразования $h$ в данном базисе.

Из первых двух матричных уравнений следует, что $Y+Z=(A+B)X$. Сравнивая это уравнение с третьим уравнением, получаем $C=A+B$.

## § 7.16. Невырожденные линейные преобразования. Преобразование, обратное данному

Линейное преобразование называется *невырожденным*, если его матрица является невырожденной, и *вырожденным* — в противном случае.

**Теорема 7.17.** Линейное преобразование является невырожденным тогда и только тогда, когда оно взаимнооднозначно.

Доказательство. Пусть линейное преобразование $f$ имеет невырожденную матрицу $A$, т. е. $\det A \neq 0$, и $\mathbf{y} = f(x)$, тогда $Y = AX$. Фиксируем произвольный вектор $\mathbf{y}_0(y_1^0, y_2^0, \ldots, y_n^0)$ и рассмотрим соответствующую систему линейных уравнений в матричной форме $AX = Y_0$, где $Y_0$ — матрица-столбец из координат вектора $y_0$; $X$ — матрица-столбец из координат вектора $\mathbf{x}(x_1, x_2, \ldots, x_n)$. Так как $\det A \neq 0$, то линейная система имеет единственное решение $\mathbf{x}_0(x_1^0, x_2^0, \ldots, x_n^0)$. Это означает, что любой вектор $\mathbf{y}_0$ имеет единственный прообраз $\mathbf{x}_0$, т. е. $f$ — взаимно-однозначное преобразование.

Обратно, пусть линейное преобразование $f$ — взаимно-однозначно и в некотором базисе имеет матрицу $A$. Поскольку каждому вектору $\mathbf{y}_0(y_1^0, y_2^0, \ldots, y_n^0)$ соответствует единственный прообраз $\mathbf{x}_0(x_1^0, x_2^0, \ldots, x_n^0)$, то система линейных уравнений $AX_0 = Y_0$ имеет единственное решение. Следовательно, $\det A \neq 0$, т. е. линейное преобразование $f$ является невырожденным.

Следствие. Линейное невырожденное преобразование ненулевой вектор переводит в ненулевой; обратное также верно.

Действительно, всякое линейное преобразование нулевой вектор переводит в нулевой (см. § 7.8). По теореме 7.17 линейное невырожденное преобразование — взаимно-однозначно, поэтому оно ненулевой вектор переводит в ненулевой. В справедливости обратного утверждения можно убедиться методом от противного.

**Теорема 7.18.** Произведение двух линейных невырожденных преобразований есть невырожденное линейное преобразование.

Доказательство. Пусть $f$ и $g$ — линейное преобразование соответственно с матрицами $A$ и $B$ в некотором базисе. Так как эти преобразования являются невырожденными, то $\det A \neq 0$, $\det B \neq 0$. Произведение двух линейных преобразований является линейным (см. § 7.15). Произведение $g \circ f$ этих преобразований будет также невырожденным, поскольку оно имеет матрицу $BA$ (см. теорему 7.15) и $\det BA = \det B \cdot \det A \neq 0$.

Введем понятие преобразования, обратного данному преобразованию. Преобразование $\varphi$ называется *обратным* преобразованию $f$, если для любого вектора $\mathbf{x}$

$$f \circ \varphi(\mathbf{x}) = \varphi \circ f(\mathbf{x}) = \mathbf{x}, \tag{7.43}$$

т. е. произведение этих преобразований — тождественное преобразование. Из определения следует, что если $\varphi$ — преобразование, обратное преобразованию $f$, то $f$ — преобразование, обратное $\varphi$. Преобразования $f$ и $\varphi$, удовлетворяющие условию (7.43), называют *взаимно-обратными.*

Если преобразования $f$ и $\varphi$ в некотором базисе имеют соответственно матрицы $A$ и $B$, то из (7.43) и теоремы 7.15 следуют равенства

$$AB = BA = E,$$

где $A$ и $B$ — взаимно-обратные матрицы, т. е. $B = A^{-1}$.

Сказанное выше позволяет сделать следующие утверждения.

1. Линейное преобразование имеет обратное преобразование тогда и только тогда, когда оно является невырожденным.

2. Для любого невырожденного линейного преобразования с матрицей $A$ в некотором базисе существует единственное обратное преобразование с матрицей $A^{-1}$ в том же базисе.

## § 7.17. Ортогональные матрицы

Рассмотрим евклидово $n$-мерное пространство $E_n$. Для любых двух векторов $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$, $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \ldots + y_n\mathbf{e}_n$ в ортонормированном базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ скалярное произведение выражается формулой (см. § 7.7)

$$(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n. \tag{7.44}$$

Матрица

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \tag{7.45}$$

называется *ортогональной,* если соответствующая ей система векторов

$$\mathbf{a}_1(a_{11}, a_{21}, \ldots, a_{n1}),\ \mathbf{a}_2(a_{12}, a_{22}, \ldots, a_{n2}), \ldots$$
$$\ldots, \mathbf{a}_n(a_{1n}, a_{2n}, \ldots, a_{nn}) \tag{7.46}$$

является ортонормированной.

В соответствии с формулой (7.44) получаем

$$(\mathrm{a}_i,\ \mathrm{a}_j)=\sum_{k=1}^{n} a_{ki}\,a_{kj}. \tag{7.47}$$

Векторы (7.46) будут ортонормированными (см. § 7.7), если

$$\sum_{k=1}^{n} a_{ki}\,a_{kj}=\begin{cases}1 \text{ при } i=j,\\ 0 \text{ при } i\neq j\end{cases} \tag{7.48}$$

для любых $i, j$ $(i=1, 2, \ldots, n, j=1, 2, \ldots, n)$.

Примеры ортогональных матриц:

$$\begin{bmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{bmatrix},\quad \begin{bmatrix}0{,}8 & -0{,}6\\ -0{,}6 & -0{,}8\end{bmatrix},\quad \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.$$

Отметим, что единичная матрица любого порядка является ортогональной.

**Теорема 7.19.** Необходимое и достаточное условие ортогональности матрицы $A$ выражается равенством

$$A^TA=E, \tag{7.49}$$

где $A^T$ — матрица, полученная из матрицы $A$ транспонированием; $E$ — единичная матрица того же порядка, что и $A$.

Доказательство. *Необходимость.* Пусть матрица $A$, определяемая формулой (7.45), ортогональна, т. е. выполнены условия (7.48). Для транспонированной матрицы $A^T=(a'_{ik})$ по определению $a'_{ik}=a_{ki}$. Найдем произведение $A^TA=C=(c_{ij})$. Принимая во внимание (7.48), получаем

$$c_{ij}=\sum_{k=1}^{n} a'_{ik}a_{kj}=\sum_{k=1}^{n} a_{ki}\,a_{kj}=\begin{cases}1, & i=j,\\ 0, & i\neq j,\end{cases}\quad C=E=\begin{bmatrix}1 & 0 & \ldots & 0\\ 0 & 1 & \ldots & 0\\ . & . & . & .\\ 0 & 0 & \ldots & 1\end{bmatrix},$$

$$A^TA=E.$$

*Достаточность.* Пусть $A^TA=E$, тогда

$$\sum_{k=1}^{n} a'_{ik}a_{kj}=\sum_{k=1}^{n} a_{ki}\,a_{kj}=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

Это означает, что $A$ — ортогональная матрица.

С л е д с т в и е 1. Модуль определителя ортогональной матрицы равен единице.

Действительно, из (7.49) получаем $\det(A^TA)=\det E$, $\det A^T\cdot\det A=\det E$, $\det A\cdot\det A=\det E$, $(\det A)^2=1$, $\det A=\pm1$, $|\det A|=1$.

С л е д с т в и е 2. Ортогональная матрица является невырожденной матрицей. Это следует из предыдущего утверждения.

С л е д с т в и е 3. Произведение двух ортогональных матриц есть ортогональная матрица.

Пусть $A$ и $B$ — ортогональные матрицы одного порядка. Так как $(AB)^T=B^TA^T$, то $(AB)^T(AB)=B^TA^TAB=$ $=B^T(A^TA)B=B^TEB=B^TB=E$, т. е. $(AB)^T(AB)=E$. На основании (7.49) заключаем, что $AB$ — ортогональная матрица.

С л е д с т в и е 4. Равенство $A^T=A^{-1}$ выражает необходимое и достаточное условие ортогональности матрицы $A$.

В самом деле, если $A$ — ортогональная матрица, то $A^TA=E$. С другой стороны, $A^{-1}A=E$. Из этих двух равенств следует, что $A^T=A^{-1}$. Обратно, пусть $A^T=A^{-1}$, тогда отсюда и из равенства $A^{-1}A=E$ следует $A^TA=E$, т. е. $A$ — ортогональная матрица.

С л е д с т в и е 5. Матрица, полученная транспонированием ортогональной матрицы,— ортогональна.

Если $A$ — ортогональная матрица, то $(A^T)^TA^T=AA^T=$ $=AA^{-1}=E$, или $(A^T)^TA^T=E$. В силу (7.49) это означает, что $A^T$ — ортогональная матрица.

С л е д с т в и е 6. Матрица, обратная ортогональной матрице,— ортогональна.

Если $A$ — ортогональная матрица, то $A^T=A^{-1}$, $A^{-1}A=$ $=E$. Далее, $(A^{-1})^{-1}A^{-1}=AA^{-1}=E$, или $(A^{-1})^{-1}A^{-1}=E$. В силу (7.49) и следствия 4 это означает, что $A^{-1}$ — ортогональная матрица.

З а м е ч а н и е 1. Из условия $\det A=\pm1$ не следует, что $A$ — ортогональная матрица. Например, матрица $A=\begin{bmatrix}2&3\\1&2\end{bmatrix}$, для которой $\det A=1$, не будет ортогональной, так как $A^TA\neq E$.

З а м е ч а н и е 2. Сумма ортогональных матриц не является ортогональной матрицей.

З а м е ч а н и е 3. Необходимое и достаточное условие ортогональности матрицы $A$ можно выразить равенством $AA^T=E$.

**Теорема 7.20.** Матрица перехода от одного ортонормированного базиса к другому ортонормированному базису является ортогональной.

Пусть $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ — два ортонормированных базиса евклидова $n$-мерного пространства и $T = (t_{ij})$ — матрица перехода от первого базиса ко второму, т. е.

$$\begin{aligned}
e'_1 &= t_{11}e_1 + t_{21}e_2 + \ldots + t_{n1}e_n,\\
e'_2 &= t_{12}e_1 + t_{22}e_2 + \ldots + t_{n2}e_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
e'_n &= t_{1n}e_1 + t_{2n}e_2 + \ldots + t_{nn}e_n.
\end{aligned}$$

Так как $e'_1, e'_2, \ldots, e'_n$ — ортонормированный базис, т. е.

$$(e'_i, e'_j) = \begin{cases} 1, & i = j, \quad i = 1, 2, \ldots, n,\\ 0, & i \neq j, \quad j = 1, 2, \ldots, n \end{cases}$$

и

$$(e'_i, e'_j) = \sum_{k=1}^{n} t_{ki} t_{kj},$$

то

$$\sum_{k=1}^{n} t_{ki}t_{kj} = \begin{cases} 1, & i = j,\\ 0, & i \neq j. \end{cases}$$

Это означает, что $T = (t_{ij})$ — ортогональная матрица.

## § 7.18. Ортогональные преобразования

Линейное преобразование евклидова пространства называется *ортогональным*, если в некотором ортонормированном базисе его матрица ортогональна.

**Теорема 7.21.** Линейное преобразование евклидова пространства является ортогональным тогда и только тогда, когда оно ортонормированный базис переводит в ортонормированный.

Доказательство. Пусть $f$ — линейное преобразование, имеющее в ортонормированном базисе $e_1, e_2, \ldots$ $\ldots, e_n$ ортогональную матрицу

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix},$$

в которой $k$-й столбец состоит из координат вектора $\mathbf{e}'_k = f(\mathbf{e}_k) = a_{1k}\mathbf{e}_1 + a_{2k}\mathbf{e}_2 + \ldots + a_{nk}\mathbf{e}_n$ $(k = 1, 2, \ldots, n)$. Поскольку матрица $A$ ортогональна, то для всех $i, j = 1, 2, \ldots, n$

$$(\mathbf{e}'_i, \mathbf{e}'_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

т. е. базис $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ является ортонормированным. Обратно, пусть линейное преобразование $f$ переводит ортонормированный базис $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ в ортонормированный базис $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$, тогда по теореме 7.20 матрица перехода является ортогональной и $f$ — ортогональное преобразование.

**Теорема 7.22.** Ортогональное преобразование не меняет скалярного произведения векторов.

Доказательство. Пусть $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ — ортонормированный базис евклидова пространства. Для любых двух векторов этого пространства $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ и $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \ldots + y_n\mathbf{e}_n$ скалярное произведение выражается формулой (7.44). Если $f$ — ортогональное преобразование данного евклидова пространства, то

$$f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \ldots + x_n f(\mathbf{e}_n),$$
$$f(\mathbf{y}) = y_1 f(\mathbf{e}_1) + y_2 f(\mathbf{e}_2) + \ldots + y_n f(\mathbf{e}_n),$$

где $f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n)$ — ортонормированный базис. Следовательно,

$$(f(\mathbf{x}), f(\mathbf{y})) = x_1y_1 + x_2y_2 + \ldots + x_ny_n,$$

т. е. скалярное преобразование образов совпадает со скалярным произведением прообразов.

Следствие 1. При ортогональном преобразовании $f$ остается неизменной норма вектора, т. е. $\|\mathbf{x}\| = \|f(\mathbf{x})\|$.

Следствие 2. При ортогональном преобразовании $f$ остается неизменным угол между векторами, т. е.

$$\frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} = \frac{(f(\mathbf{x}), f(\mathbf{y}))}{\|f(\mathbf{x})\| \cdot \|f(\mathbf{y})\|}.$$

Ортогональные преобразования обладают следующими свойствами:

1. Ортогональное преобразование является невырожденным.

2. Для любого ортогонального преобразования суще-

ствует обратное преобразование, являющееся ортогональным.

3. Если ортогональное преобразование имеет матрицу $A$, то обратное ему преобразование имеет матрицу $A^T$.

4. Произведение двух ортогональных преобразований — ортогональное преобразование.

# Глава 8

# КВАДРАТИЧНЫЕ ФОРМЫ

В этой главе рассматриваются квадратичные формы и их приложения к упрощению уравнений линий второго порядка на плоскости и уравнений поверхностей второго порядка в пространстве.

## § 8.1. Основные определения

Рассмотрим $n$ переменных $x_1, x_2, \ldots, x_n$, принимающих любые действительные значения. Составим сумму их произведений $x_i x_j$, умноженных на некоторые числа $a_{ij}$, т. е. сумму

$$\left.\begin{array}{l} a_{11}x_1^2 + a_{12}x_1x_2 + \ldots + a_{1n}x_1x_n + \\ + a_{21}x_2x_1 + a_{22}x_2^2 + \ldots + a_{2n}x_2x_n + \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \ldots + a_{nn}x_n^2 = \\ = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\,x_i\,x_j. \end{array}\right\} \qquad (8.1)$$

*Квадратичной формой* $n$ переменных $x_1, x_2, \ldots, x_n$ называется сумма вида (8.1). Обозначим эту сумму $f(x_1, x_2, \ldots, x_n)$, тогда по определению

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\,x_i\,x_j, \qquad (8.2)$$

числа $a_{ij}$ называются *коэффициентами* квадратичной формы.

Квадратичная форма называется действительной или комплексной в зависимости от того, являются ли ее коэффициенты соответственно действительными или комплексными числами. В дальнейшем будем рассматривать действительные квадратичные формы.

Так как $a_{ij}x_ix_j + a_{ji}x_jx_i = (a_{ij} + a_{ji})x_ix_j = \frac{1}{2}(a_{ij} + a_{ji})x_ix_j + \frac{1}{2}(a_{ij} + a_{ji})x_jx_i = a'_{ij}x_ix_j + a'_{ji}x_jx_i$, то коэффициенты $a_{ij}$ и $a_{ji}$ можно считать равными, т. е.

$$a_{ij} = a_{ji}. \tag{8.3}$$

Составим матрицу из коэффициентов квадратичной формы

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

В силу условия (8.3) матрица $A$ — симметрическая (ее элементы, симметричные относительно главной диагонали, равны между собой). Очевидно, действительная матрица $A$ является симметрической тогда и только тогда, когда она совпадает с матрицей $A^T$, полученной транспонированием, т. е.

$$A^T = A. \tag{8.4}$$

*Матрицей квадратичной формы* называется матрица, составленная из ее коэффициентов. Квадратичной форме (8.2) $n$ переменных $x_1, x_2, \dots, x_n$ соответствует единственная симметрическая матрица

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}. \tag{8.5}$$

Обратно, всякой симметрической матрице (8.5) соответствует единственная квадратичная форма с точностью до обозначения переменных.

*Рангом квадратичной формы* называется ранг $r$ ее матрицы. Квадратичная форма $n$ переменных называется невырожденной, если ее матрица — невырожденная, т. е. $r = n$, и вырожденной, если $r < n$.

Пример. Записать матрицу квадратичной формы $f(x_1, x_2, x_3) = x_1^2 - 6x_1x_2 - 8x_1x_3 + 7x_2^2 + 4x_2x_3 - 5x_3^2$ и найти ее ранг.

В данном случае $a_{11} = 1$, $a_{12} = a_{21} = -3$, $a_{13} = a_{31} = -4$, $a_{22} = 7$, $a_{23} = a_{32} = 2$, $a_{33} = -5$, поэтому

$$A = \begin{bmatrix} 1 & -3 & -4 \\ -3 & 7 & 2 \\ -4 & 2 & -5 \end{bmatrix}.$$

Вычислим определитель этой матрицы:

$$\det A = \begin{vmatrix} 1 & -3 & -4 \\ -3 & 7 & 2 \\ -4 & 2 & -5 \end{vmatrix} = -35+24+24-112+45-4 = -58.$$

Так как $\det A \neq 0$, то ранг матрицы равен трем, т. е. $r=3$.

Квадратичную форму (8.2) $n$ переменных $x_1, x_2, \ldots, x_n$ можно записать в матричном виде. Действительно, если $X$ — матрица-столбец из переменных $x_1, x_2, \ldots, x_n$, $X^T$ — матрица, полученная транспонированием матрицы $X$, т. е. строка из тех же переменных, то

$$X^T A X = [x_1,\ x_2,\ \ldots,\ x_n] \cdot \begin{bmatrix} a_{11} & a_{12}, & \ldots, & a_{1n} \\ a_{21} & a_{22}, & \ldots, & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2}, & \ldots, & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} =$$

$$= [a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n \quad a_{12}x_1 + a_{22}x_2 + \ldots + a_{n2}x_n \ \ldots \ a_{1n}x_1 +$$

$$+ a_{2n}x_2 + \ldots + a_{nn}x_n] \cdot \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

т. е.

$$X^T A X = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad \text{или} \quad f(x_1,\ x_2,\ \ldots,\ x_n) = X^T A X. \qquad (8.6)$$

## § 8.2. Преобразование квадратичной формы при линейном однородном преобразовании переменных

Рассмотрим квадратичную форму (8.6). Перейдем к новым переменным $y_1, y_2, \ldots, y_n$ по формулам

$$\left.\begin{aligned} x_1 &= b_{11}y_1 + b_{12}y_2 + \ldots + b_{1n}y_n, \\ x_2 &= b_{21}y_1 + b_{22}y_2 + \ldots + b_{2n}y_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= b_{n1}y_1 + b_{n2}y_2 + \ldots + b_{nn}y_n \end{aligned}\right\}, \qquad (8.7)$$

или в матричном виде

$$X = BY, \tag{8.8}$$

где

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}. \tag{8.9}$$

В квадратичной форме (8.6) вместо $x_1, x_2, \dots, x_n$ подставим их выражения через $y_1, y_2, \dots, y_n$, определяемые формулами (8.7), получим квадратичную форму $\varphi(y_1, y_2, \dots, y_n)$ $n$ переменных $y_1, y_2, \dots, y_n$ с некоторой матрицей $C$. В этом случае говорят, что квадратичная форма $f(x_1, x_2, \dots, x_n)$ переводится в квадратичную форму $\varphi(y_1, y_2, \dots, y_n)$ линейным однородным преобразованием (8.7). Линейное однородное преобразование (8.8) называется *невырожденным*, если $\det B \neq 0$.

Две квадратичные формы называются *конгруэнтными*, если существует невырожденное линейное однородное преобразование, переводящее одну из них в другую. Если $f(x_1, x_2, \dots, x_n)$ и $\varphi(y_1, y_2, \dots, y_n)$ конгруэнтны, будем писать $f(x_1, x_2, \dots, x_n) \sim \varphi(y_1, y_2, \dots, y_n)$. Отметим следующие свойства конгруэнтности квадратичных форм:

1. $f(x_1, x_2, \dots, x_n) \sim f(x_1, x_2, \dots, x_n)$.
2. Если $f(x_1, x_2, \dots, x_n) \sim \varphi(y_1, y_2, \dots, y_n)$, $\varphi(y_1, y_2, \dots, y_n) \sim \psi(z_1, z_2, \dots, z_n)$, то $f(x_1, x_2, \dots, x_n) \sim \psi(z_1, z_2, \dots, z_n)$.

Справедливость первого свойства очевидна: квадратичная форма переводится в себя тождественным преобразованием (невырожденным). Второе свойство следует из того, что произведение двух невырожденных линейных однородных преобразований является невырожденным.

Выясним, какой вид имеет матрица $C$ квадратичной формы $\varphi(y_1, y_2, \dots, y_n)$, в которую переходит квадратичная форма $f(x_1, x_2, \dots, x_n)$ при преобразовании (8.7).

**Теорема 8.1.** Квадратичная форма $f(x_1, x_2, \dots, x_n)$ с матрицей $A$ линейным однородным преобразованием $X = BY$ переводится в квадратичную форму $\varphi(y_1, y_2, \dots, y_n)$ с матрицей $C = B^T AB$.

Доказательство. В квадратичную форму $f(x_1, x_2, \dots, x_n) = X^T AX$ подставим $X = BY$, получим $f(x_1, x_2, \dots, x_n) = X^T AX = (BY)^T A(BY) = (Y^T B^T)ABY =$

$= Y^T(B^TAB)Y$, или $f(x_1, x_2, \ldots, x_n) = Y^TCY$, где $C = B^TAB$. Поскольку

$C^T = (B^TAB)^T = (B^T(AB))^T = (AB)^T(B^T)^T = (B^TA^T)B = B^TAB = C$, т. е. $C^T = C$, то $C$ — симметрическая матрица. Следовательно, $C = B^TAB$ является матрицей квадратичной формы $\varphi(y_1, y_2, \ldots, y_n)$.

Следствие 1. Определители матриц конгруэнтных невырожденных действительных квадратичных форм имеют одинаковые знаки.

Поскольку матрицы $A$ и $C$ двух конгруэнтных квадратичных форм связаны соотношением $C = B^TAB$, где $B$ — невырожденная матрица, то $\det C = \det(B^TAB) = \det B^T \times \times \det A \cdot \det B = (\det B)^2 \cdot \det A$, или $\det C = (\det B)^2 \cdot \det A$, откуда следует $\det C \cdot \det A > 0$, ибо $(\det B)^2 > 0$.

Следствие 2. Конгруэнтные квадратичные формы имеют одинаковые ранги. Поскольку ранг матрицы не меняется при умножении ее слева или справа на невырожденную матрицу и $C = B^TAB$, то $\operatorname{rang} C = \operatorname{rang} A$.

## § 8.3. Приведение действительной квадратичной формы к нормальному виду

Квадратичная форма $f(x_1, x_2, \ldots, x_n)$ называется *канонической*, если она не содержит произведений различных переменных, т. е.

$$f(x_1, x_2, \ldots x_n) = \sum_{i=1}^{r} \alpha_{ii} x_i^2 \quad (r \leqslant n). \tag{8.10}$$

Каноническая квадратичная форма называется *нормальной* (или имеет нормальный вид), если $|\alpha_{ii}| = 1$ $(i = 1, 2, \ldots, r)$, т. е. отличные от нуля коэффициенты при квадратах переменных равны $+1$ или $-1$. Например, квадратичная форма $f(x_1, x_2, x_3, x_4) = 6x_1^2 + 4x_3^2 - 3x_4^2$, для которой $\alpha_{11} = 6$, $\alpha_{22} = 0$, $\alpha_{33} = 4$, $\alpha_{44} = -3$, имеет канонический вид; квадратичная форма $f(x_1, x_2, x_3, x_4) = x_1^2 - x_3^2 + x_4^2$ является нормальной, так как $\alpha_{11} = 1$, $\alpha_{22} = 0$, $\alpha_{33} = -1$, $\alpha_{44} = 1$.

Пусть квадратичная форма $f(x_1, x_2, \ldots, x_n)$ линейным невырожденным преобразованием приведена к каноническому виду

$$\varphi(y_1, y_2, \ldots, y_n) = b_{11}y_1^2 + b_{22}y_2^2 + \ldots + b_{nn}y_n^2,$$

где $y_1, y_2, \ldots, y_n$ — новые переменные. Некоторые из коэффициентов $b_{ii}$ могут оказаться равными нулю; докажем, что число отличных от нуля коэффициентов в этой формуле равно рангу $r$ матрицы квадратичной формы $f$. На основании следствия 2 из теоремы 8.1 ранги квадратичных форм $f$ и $\varphi$ равны между собой. Поскольку матрица квадратичной формы $\varphi$ является диагональной и ранг ее также равен $r$, то это означает, что среди диагональных коэффициентов имеется $r$ отличных от нуля.

**Теорема 8.2.** Любая квадратичная форма некоторым невырожденным линейным преобразованием может быть приведена к каноническому виду.

Доказательство проведем индукцией по числу переменных. При $n=1$ теорема верна, так как в этом случае квадратичная форма $f(x_1)=a_{11}x_1^2$ имеет канонический вид. Докажем теорему для квадратичных форм от $n$ переменных, считая ее уже доказанной для форм с меньшим числом переменных. Предположим, что среди коэффициентов $a_{ii}$ $(i=1, 2, \ldots, n)$ данной квадратичной формы

$$f(x_1, x_2, \ldots, x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

имеются отличные от нуля. Пусть $a_{11}\neq 0$, тогда выражение $a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n)^2$, являющееся квадратичной формой, содержит такие же члены с переменной $x_1$, что и форма $f$, поэтому разность $f-a_{11}^{-1}(a_{11}x_1+$ $+a_{12}x_2+\ldots+a_{1n}x_n)^2=f_1$ будет квадратичной формой, не зависящей от $x_1$, а только от $x_2, x_3, \ldots, x_n$, т. е. $f_1=$ $=f_1(x_2, x_3, \ldots, x_n)$, откуда $f=a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\ldots$ $\ldots+a_{1n}x_n)^2+f_1$.

Перейдя к новым переменным $y_1, y_2, \ldots, y_n$ по формулам

$$y_1=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n,\ y_i=x_i\ (i=2, 3, \ldots, n), \tag{8.11}$$

получим квадратичную форму

$$f=a_{11}^{-1}y_1^2+f_2,\ f_2=f_2(y_2, y_3, \ldots, y_n). \tag{8.12}$$

Так как квадратичная форма $f_2$ зависит от меньшего числа переменных, по предположению ее можно привести к каноническому виду. Следовательно, квадратичная форма $f$ также приведена к каноническому виду (8.12). Отметим, что преобразование (8.11) является невырожденным, так как его определитель равен $a_{11}\neq 0$. Преобразование, обратное преобразованию (8.11), также будет невырожденным, оно приводит исходную форму $f$ к виду (8.12).

Случай, когда все коэффициенты $a_{ii}=0$ $(i=1, 2, \ldots \ldots, n)$, но, например, $a_{12}\neq 0$, с помощью преобразования

$$x_1=z_1-z_2,\ x_2=z_1+z_2,\ x_i=z_i \quad (i=3, 4, \ldots, n)$$

сводится к предыдущему. Действительно, это преобразование является невырожденным, ибо его определитель отличен от нуля:

$$\begin{vmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix}=2.$$

В результате этого преобразования член $2a_{12}x_1x_2$ принимает вид $2a_{12}x_1x_2=2a_{12}(z_1-z_2)(z_1+z_2)=2a_{12}z_1^2-2a_{12}z_2^2$; в конгруэнтной квадратичной форме $\varphi(z_1, z_2, \ldots, z_n)$ будут отличными от нуля коэффициенты при квадратах переменных $z_1$ и $z_2$.

Таким образом, квадратичная форма несколькими невырожденными преобразованиями, которые можно заменить одним невырожденным преобразованием — их произведением, приводится к каноническому виду, т. е. к виду суммы квадратов переменных с некоторыми коэффициентами. Число этих квадратов с коэффициентами, отличными от нуля, равно рангу формы.

Замечание. Если исходная квадратичная форма является действительной, то коэффициенты невырожденного линейного преобразования и коэффициенты в ее каноническом виде будут также действительными.

**Теорема 8.3.** Любую действительную квадратичную форму линейным невырожденным преобразованием можно привести к нормальному виду.

Доказательство. На основании теоремы 8.2 любую квадратичную форму $f(x_1, x_2, \ldots, x_n)$ можно привести к каноническому виду. Представим этот канонический вид следующим образом: $\varphi(y_1, y_2, \ldots, y_n)=c_1y_1^2+$ $+c_2y_2^2+\ldots+c_ky_k^2-c_{k+1}y_{k+1}^2-\ldots-c_ry_r^2$ $(0\leqslant k\leqslant r)$, где все числа $c_i$ $(i=1, 2, \ldots, k, k+1, \ldots, r)$ положительны.

Применяя невырожденное преобразование $z_i = \sqrt{c_i} y_i$ ($i = 1, 2, \ldots, r$), $z_j = y_j$ ($j = r+1, \ldots, n$), получаем квадратичную форму

$$\psi(z_1, z_2, \ldots, z_n) = z_1^2 + z_2^2 + \ldots + z_k^2 - z_{k+1}^2 - \ldots - z_r^2$$

в нормальном виде. Число входящих сюда квадратов равно рангу формы.

## § 8.4. Закон инерции квадратичных форм

Канонический вид, к которому приводится данная квадратичная форма, не является для нее однозначно определенным, так как любая квадратичная форма может быть приведена к этому виду различными способами.

Действительная квадратичная форма также может быть приведена к нормальному виду различными преобразованиями, однако с точностью до нумерации и обозначения переменных она приводится только к одному нормальному виду. Об этом свидетельствует следующая важная теорема, выражающая закон инерции квадратичных форм.

**Теорема 8.4.** Число положительных и число отрицательных квадратов в нормальном виде, к которому приводится данная действительная квадратичная форма невырожденным действительным линейным преобразованием, не зависит от выбора преобразования.

Доказательство. Предположим, что квадратичная форма $f(x_1, x_2, \ldots, x_n)$ ранга $r$ двумя способами приведена к нормальному виду

$$y_1^2 + y_2^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 =$$
$$= z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2. \qquad (8.13)$$

Поскольку переход от переменных $x_1, x_2, \ldots, x_n$ к переменным $y_1, y_2, \ldots, y_n$ осуществлен действительным линейным невырожденным преобразованием, то и обратное преобразование

$$y_i = \sum_{s=1}^{n} b_{is} x_s \quad (i = 1, 2, \ldots, n) \qquad (8.14)$$

также является действительным невырожденным. Точно так же будет действительным невырожденным и линейное преобразование

$$z_j = \sum_{i=1}^{n} c_{ji} x_i \quad (j = 1, 2, \ldots, n). \tag{8.15}$$

Предположим, что $k<l$, и напишем систему равенств

$$y_1=0, y_2=0, \ldots, y_k=0,$$
$$z_{l+1}=0, \ldots, z_r=0, \ldots, z_n=0. \tag{8.16}$$

Если левые части этих равенств заменить соответственно их выражениями (8.14) и (8.15), получим систему $n-l+k$ линейных однородных уравнений с $n$ неизвестными $x_1, x_2, \ldots, x_n$. Поскольку $n-l+k<n$ (ибо $k<l$), т. е. число уравнений однородной системы меньше числа неизвестных, система имеет ненулевое действительное решение $\alpha_1, \alpha_2, \ldots, \alpha_n$.

В равенство (8.13) вместо $y_i$, $z_j$ подставим соответственно их выражения (8.14) и (8.15), а вместо неизвестных $x_1, x_2, \ldots, x_n$ — числа $\alpha_1, \alpha_2, \ldots, \alpha_n$. Принимая во внимание (8.16), получаем

$$-y_{k+1}^2(\alpha) - \ldots - y_r^2(\alpha) = z_1^2(\alpha) + \ldots + z_l^2(\alpha). \tag{8.17}$$

Поскольку все коэффициенты линейных преобразований (8.14) и (8.15) — действительные числа, то все квадраты в равенстве (8.17) неотрицательны, что влечет за собой равенство их нулю, в частности,

$$z_1(\alpha)=0, \ z_2(\alpha)=0, \ \ldots, \ z_l(\alpha)=0. \tag{8.18}$$

С другой стороны, по самому выбору чисел $\alpha_1, \alpha_2, \ldots$ $\ldots, \alpha_n$

$$z_{l+1}(\alpha)=0, \ \ldots, \ z_n(\alpha)=0. \tag{8.19}$$

Следовательно, система $n$ линейных уравнений $z_i(x)=0$ $(i=1, 2, \ldots, n)$ с $n$ неизвестными $x_1, x_2, \ldots, x_n$ имеет ненулевое решение $\alpha_1, \alpha_2, \ldots, \alpha_n$. Определитель такой системы должен быть равен нулю (см. следствие из теоремы Крамера). Это противоречит тому, что преобразование (8.15) является невырожденным. Предположение, что $l<k$, приводит к такому же противоречию. Отсюда следует равенство $k=l$, которое и доказывает теорему.

Число положительных квадратов в нормальной форме, к которой приводится данная действительная квадратичная форма, называют *положительным индексом* инерции этой формы, число отрицательных квадратов — *отрицательным индексом* инерции, разность между положительным и отрицательным индексами инерции — *сигнатурой* формы $f$. Очевидно, если известен ранг формы, то задание любого из трех указанных выше чисел определяет два других.

**Теорема 8.5.** Две действительные квадратичные формы от $n$ переменных тогда и только тогда конгруэнтны, когда они имеют одинаковые ранги и одинаковые сигнатуры.

Доказательство. Пусть форма $f$ переводится в форму $\varphi$ невырожденным линейным действительным преобразованием. При этом, как известно, ранг не меняется. Это преобразование не меняет и сигнатуры, поскольку в противном случае формы $f$ и $\varphi$ приводились бы к различным нормальным видам, а тогда форма $f$ приводилась бы к обоим этим нормальным видам, что противоречит закону инерции квадратичных форм.

Обратно, если квадратичные формы имеют одинаковые ранги и одинаковые сигнатуры, они приводятся к одному и тому же нормальному виду и поэтому могут быть переведены одна в другую.

## § 8.5. Знакоопределенные квадратичные формы

Действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ называется *положительно определенной,* если она приводится к нормальному виду, состоящему из $n$ положительных квадратов; $f(x_1, x_2, \ldots, x_n) \sim \varphi(y_1, y_2, \ldots, y_n)$, где

$$\varphi(y_1, y_2, \ldots, y_n) = y_1^2 + y_2^2 + \ldots + y_n^2, \tag{8.20}$$

т. е. если ранг и положительный индекс инерции равен числу неизвестных.

Систему значений $x_1, x_2, \ldots, x_n$ назовем *нулевой,* если $x_1 = x_2 = \ldots = x_n = 0$, и *ненулевой,* если хотя бы одно из них отлично от нуля.

**Теорема 8.6.** Действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ является положительно определенной тогда и только тогда, когда она принимает положительные значения при любой ненулевой системе значений переменных $x_1, x_2, \ldots, x_n$.

Доказательство. Пусть $f(x_1, x_2, \ldots, x_n)$ — положительно определенная форма, т. е. преобразуется к виду (8.20) невырожденным линейным преобразованием

$$y_i = \sum_{j=1}^{n} b_{ij} x_j \quad (i=1,\ 2,\ \ldots,\ n) \tag{8.21}$$

(с матрицей $B=(b_{ij})$, для которой $\det B \neq 0$). Фиксируем любую ненулевую систему значений $x_1, x_2, \ldots, x_n$. Подставляя эту систему значений в (8.21), получаем соответствующую систему значений $y_1, y_2, \ldots, y_n$, которая также будет ненулевой. Действительно, в противном случае имели бы линейную однородную систему с отличным от нуля определителем, обладающую ненулевым решением, что противоречит следствию из теоремы Крамера. При подстановке ненулевой системы значений $y_1, y_2, \ldots, y_n$ в (8.20) получим $\varphi(y_1, y_2, \ldots, y_n) > 0$.

Обратно, пусть квадратичная форма $f$ не является положительно определенной, т. е. ее ранг или положительный индекс инерции меньше $n$; линейным невырожденным преобразованием (8.21) эта форма приводится, в частности, к одному из видов:

$$\varphi(y_1,\ y_2,\ \ldots,\ y_n) = y_1^2 + y_2^2 + \cdots + y_{n-1}^2,$$

$$\varphi(y_1,\ y_2,\ \ldots,\ y_n) = y_1^2 + y_2^2 + \ldots + y_{n-1}^2 - y_n^2.$$

В этом случае существует ненулевая система значений $x_1, x_2, \ldots, x_n$, при которой $\varphi(x_1, x_2, \ldots, x_n) = 0$ или $\varphi(x_1, x_2, \ldots, x_n) < 0$; такими значениями будут значения, полученные из системы (8.21) при $y_1 = y_2 = \ldots = y_{n-1} = 0$, $y_n = 1$.

Докажем теорему, которая позволит по коэффициентам квадратичной формы установить, будет ли эта форма положительно определенной. Предварительно введем вспомогательное понятие. Пусть дана квадратичная форма $f(x_1, x_2, \ldots, x_n)$ с матрицей $A = (a_{ij})$. *Главными минорами* квадратичной формы $f$ называются миноры

$$a_{11},\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\quad \ldots,\quad \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix},\quad \ldots,\quad \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

т. е. миноры порядка $1, 2, \ldots, n$ матрицы $A$, расположенные в левом верхнем углу; последний из них совпадает с определителем матрицы.

**Теорема 8.7.** Квадратичная форма $f(x_1, x_2, \ldots, x_n)$ с действительной матрицей является положительно определенной тогда и только тогда, когда все ее главные миноры положительны.

Докажем теорему для случая $n$ переменных, предположив, что для квадратичных форм $n-1$ переменных она верна. Действительно, при $n=1$ теорема верна, поскольку квадратичная форма $a_{11}x_1^2$ — положительно определенная тогда и только тогда, когда $a_{11}>0$.

Квадратичную форму $n$ переменных $f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$ (см. (8.4)) можно записать в виде

$$f = f_1(x_1, x_2, \ldots, x_{n-1}) + 2\sum_{i=1}^{n-1} a_{in}x_ix_n + a_{nn}x_n^2, \quad (8.22)$$

где $f_1 = f_1(x_1, x_2, \ldots, x_{n-1})$ — квадратичная форма $n-1$ переменных $x_1, x_2, \ldots, x_{n-1}$, составленная из тех членов формы $f$, которые не содержат переменной $x_n$. Главные миноры формы $f_1$ совпадают с главными минорами формы $f$, кроме последнего.

Пусть квадратичная форма $f$ является положительно определенной, тогда форма $f_1$ будет также положительно определенной. Действительно, если бы существовала ненулевая система значений $x_1^0, x_2^0, \ldots, x_{n-1}^0$, при которой $f_1(x_1^0, x_2^0, \ldots, x_{n-1}^0) \leqslant 0$, то в силу (8.22) имелась бы и ненулевая система значений $x_1^0, x_2^0, \ldots, x_{n-1}^0, x_n^0$, где $x_n^0 = 0$, для которой $f(x_1^0, x_2^0, \ldots, x_n^0) \leqslant 0$, что противоречит тому, что $f$ — положительно определенная форма. Следовательно, по индуктивному предположению все главные миноры формы $f_1$, т. е. все главные миноры формы $f$, кроме последнего, положительны. Последний главный минор формы $f$, или определитель матрицы $A$, также положителен. В самом деле, форма $f(x_1, x_2, \ldots, x_n)$ может быть приведена к виду $\varphi(y_1, y_2, \ldots, y_n) = y_1^2 + y_2^2 + \ldots + y_n^2$, определитель которой положителен (он равен единице). На основании следствия 1 из теоремы 8.1 определители квадратичных форм $f$ и $\varphi$ имеют одинаковые знаки, т. е. определитель формы $f$ также будет положительным. Итак, все главные миноры положительно определенной квадратичной формы положительны.

Обратно, пусть все главные миноры квадратичной формы $f(x_1, x_2, \ldots, x_n)$ положительны, тогда положительными будут и главные миноры формы $f_1$ в равенстве (8.22). Форма $f_1$ по индуктивному предположению является положительно определенной, поэтому существует такое невырожденное линейное преобразование переменных $x_1, x_2, \ldots, x_{n-1}$, которое приводит форму $f_1$ к нормальному виду $f_2 = y_1^2 + y_2^2 + \ldots + y_{n-1}^2$. Это преобразование можно дополнить до невырожденного линейного преобразования всех переменных $x_1, x_2, \ldots, x_n$, положив $x_n = y_n$. Указанным преобразованием форма $f$ в силу (8.22) приводится к виду

$$f = \sum_{i=1}^{n} y_i^2 + 2\sum_{i=1}^{n-1} b_{in} y_i y_n + b_{nn} y_n^2, \tag{8.23}$$

где $b_{in}$ — некоторые коэффициенты. Поскольку $y_i^2 + 2b_{in}y_iy_n = (y_i + b_{in}y_n)^2 - b_{in}^2 y_n^2$, то невырожденное линейное преобразование

$$z_i = y_i + b_{in} y_n \quad (i = 1, 2, \ldots, n-1), \quad z_n = y_n$$

приводит форму (8.23) к каноническому виду:

$$f_3 = \sum_{i=1}^{n-1} z_i^2 + c z_n^2. \tag{8.24}$$

Убедимся в том, что $c > 0$. Определитель квадратичной формы $f_3$ равен $c$, где $c$ — положительное число. Действительно, форма $f_3$ получена из формы $f$ двумя невырожденными линейными преобразованиями, а определитель формы $f$, как последний из главных миноров этой формы, был положительным. На основании следствия из теоремы 8.1 заключаем, что определитель формы $f_3$ также положителен, т. е. $c > 0$. Итак, форма (8.24) — положительно определенная.

Действительная квадратичная форма называется *отрицательно определенной*, если она является невырожденной и приводится к нормальному виду, содержащему только отрицательные квадраты всех переменных; эту форму можно привести к виду

$$\varphi(y_1, y_2, \ldots, y_n) = -y_1^2 - y_2^2 - \ldots - y_n^2. \tag{8.25}$$

Положительно определенные и отрицательно определенные квадратичные формы называются *знакоопределенными* квадратичными формами.

Вырожденные квадратичные формы, нормальный вид которых состоит из квадратов одного знака, называются *полуопределенными.*

*Неопределенными* называются квадратичные формы, нормальный вид которых содержит как положительные, так и отрицательные квадраты переменных.

Очевидно, для отрицательно определенной квадратичной формы $n$ переменных ее определитель будет положительным при четном $n$ и отрицательным — при нечетном $n$.

Приведем без доказательства теорему о необходимом и достаточном условии того, что квадратичная форма является отрицательно определенной.

**Теорема 8.8.** Квадратичная форма является отрицательно определенной тогда и только тогда, когда ее главные миноры четного порядка положительны, а нечетного — отрицательны.

Отметим, что необходимость следует непосредственно из вида матрицы.

## § 8.6. Приведение квадратичной формы к каноническому виду ортогональным преобразованием переменных

Докажем, что действительную квадратичную форму можно привести к каноническому виду с помощью ортогонального преобразования.

**Теорема 8.9.** Если существует ортогональное преобразование с матрицей $C$, приводящее действительную квадратичную форму $f=f(x_1, x_2, \ldots, x_n)$ к каноническому виду

$$\varphi(y_1, y_2, \ldots, y_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2, \qquad (8.26)$$

то $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы $A$ квадратичной формы $f$, причем столбцами матрицы $C$ являются собственные векторы-столбцы матрицы $A$ с собственными числами $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Доказательство. Пусть ортогональное преобразование $X=CY$, где $C=(c_{ij})$, приводит квадратичную форму $f$ к каноническому виду (8.26), тогда матрицы формы $\varphi$ имеют вид

$$D=\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}. \tag{8.27}$$

Поскольку $D=C^TAC$ (см. теорему 8.1) и $C$ — ортогональная матрица, то $C^T=C^{-1}$; следовательно, $D=C^{-1}AC$, откуда получаем, что $\lambda_i$ — характеристические числа матрицы $A$ (см. § 7.14).

Принимая во внимание выражения $D=C^TAC$, $CC^T=E$ и умножая обе части равенства (8.27) на $C$ слева, получаем (так как $CD=CC^{-1}AC=AC$)

$$AC=C\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}. \tag{8.28}$$

Для элементов $b_{ij}$ $(i=1,\ 2,\ \dots,\ n)$ матрицы $B=AC$ согласно правилу умножения матриц имеем $b_{ij}=a_{i1}c_{1j}+$ $+a_{i2}c_{2j}+\dots+a_{in}c_{nj}$. С другой стороны, из равенства (8.28) следует, что $b_{ij}=c_{ij}\lambda_j$, поэтому

$$a_{i1}c_{1j}+a_{i2}c_{2j}+\dots+a_{in}c_{nj}=\lambda_j c_{ij}\quad (i=1,\ 2,\ \dots,\ n),$$

или

$$A\begin{bmatrix}c_{1j}\\ c_{2j}\\ \dots\\ c_{nj}\end{bmatrix}=\lambda_j\begin{bmatrix}c_{1j}\\ c_{2j}\\ \dots\\ c_{nj}\end{bmatrix}.$$

Итак, $j$-й столбец $(j=1,\ 2,\ \dots,\ n)$ матрицы $C$ является собственным вектором-столбцом матрицы $A$ с собственным числом $\lambda_j$.

**Теорема 8.10.** Для любой действительной квадратичной формы существует ортогональное преобразование, приводящее ее к каноническому виду.

Доказательство. Теорема верна для квадратичной формы одной переменной $f(x_1)=a_{11}x_1^2$, которая имеет канонический вид; искомым ортогональным преобразованием будет тождественное преобразование. Предположим, что теорема верна для квадратичных форм $n-1$ переменных $x_1,\ x_2,\ \dots,\ x_{n-1}$, и докажем ее для квадратичных форм $n$ переменных.

Пусть $f(x_1, x_2, \ldots, x_n)$ — квадратичная форма с матрицей $A=(a_{ij})$, $\lambda$ — одно из собственных чисел этой матрицы и $c_{i1}$ $(i=1, 2, \ldots, n)$ — нормированный собственный вектор-столбец матрицы $A$ с собственным значением $\lambda$. Построим ортогональную матрицу $C=(c_{ij})$, первым столбцом которой является указанный столбец $c_{i1}$ $(i=1, 2, \ldots, n)$. Рассмотрим преобразование

$$X=CY, \tag{8.29}$$

переводящее форму $f(x_1, x_2, \ldots, x_n)$ в форму $\varphi(y_1, y_2, \ldots$ $\ldots, y_n)$. Если $D$ — матрица формы $\varphi$, то, как известно, $D=C^TAC$. Найдем эту матрицу, определив сначала $B=$ $=C^TA$, потом $D=BC$. Для членов $b_{1j}$ $(j=1, 2, \ldots, n)$ первой строки матрицы $B$ имеем

$$b_{1j}=c_{11}a_{1j}+c_{21}a_{2j}+\ldots+c_{n1}a_{nj} \quad (j=1, 2, \ldots, n).$$

С другой стороны, по теореме 8.9 $c_{i1}$ $(i=1, 2, \ldots, n)$ — собственный вектор-столбец симметрической матрицы $A$ с собственным числом $\lambda$, т. е.

$$c_{11}a_{1j}+c_{21}a_{2j}+\ldots+c_{n1}a_{nj}=\lambda c_{j1}.$$

Из этих двух равенств следует, что $b_{1j}=\lambda c_{j1}$. Находим элементы $d_{1j}$ первой строки матрицы $D=BC$: $d_{1j}=$ $=\lambda c_{11}c_{1j}+\lambda c_{21}c_{2j}+\ldots+\lambda c_{n1}c_{nj}$. Поскольку $C$ — ортогональная матрица, то $d_{1j}=\lambda$ при $j=1$ и $d_{1j}=0$ при $j\neq 1$. Следовательно, симметрическая матрица $D$ имеет вид

$$D=\begin{bmatrix} \lambda & 0 & \ldots & 0 \\ 0 & d_{22} & \ldots & d_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & d_{2n} & \ldots & d_{nn} \end{bmatrix},$$

поэтому квадратичную форму $\varphi(y_1, y_2, \ldots, y_n)$ можно записать так:

$$\varphi(y_1, y_2, \ldots, y_n)=\lambda y_1^2+\varphi_1(y_2, y_3, \ldots, y_n).$$

Согласно индуктивному предположению для квадратичной формы $\varphi_1(y_2, y_3, \ldots, y_n)$ $n-1$ переменных $y_2$, $y_3, \ldots, y_n$ существует ортогональное преобразование

$$z_2=k_{22}y_2+k_{23}y_3+\ldots+k_{2n}y_n,$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$z_n=k_{n2}y_2+k_{n3}y_3+\ldots+k_{nn}y_n,$$

приводящее её к каноническому виду $\varphi_2(z_2, z_3, \ldots, z_n) = \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2$. Преобразование

$$\left.\begin{aligned} z_1 &= y_1, \\ z_2 &= k_{22}y_2 + k_{23}y_3 + \ldots + k_{2n}y_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ z_n &= k_{n2}y_2 + k_{n3}y_3 + \ldots + k_{nn}y_n \end{aligned}\right\} \qquad (8.30)$$

является ортогональным; оно приводит квадратичную форму $\varphi(y_1, y_2, \ldots, y_n)$ к каноническому виду $\lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2$.

Итак, квадратичная форма $f(x_1, x_2, \ldots, x_n)$ приводится к каноническому виду $\lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2$ ортогональным преобразованием, равным произведению преобразований (8.29) и (8.30).

**Теорема 8.11.** Для любой действительной симметрической матрицы $A$ существует такая ортогональная матрица $T$, что $T^{-1}AT$ — диагональная матрица.

Доказательство. Пусть $A$ — действительная симметрическая матрица порядка $n$, $f(x_1, x_2, \ldots, x_n)$ — квадратичная форма с матрицей $A$. По теореме 8.10 существует ортогональное преобразование, приводящее эту форму к каноническому виду. Матрицу этого преобразования обозначим через $T$, тогда в силу теоремы 8.1 $T^{\mathrm{T}}AT = D$, где $D$ — диагональная матрица. Поскольку $T$ — ортогональная матрица, т. е. $T^{\mathrm{T}} = T^{-1}$, то $T^{-1}AT = D$ — диагональная матрица.

Следствие. Любая действительная симметрическая матрица может быть приведена к диагональному виду.

**Теорема 8.12.** Если линейное преобразование действительного линейного пространства имеет действительную симметрическую матрицу в некотором ортонормированном базисе, то существует ортонормированный базис, состоящий из собственных векторов этого преобразования.

Доказательство. Пусть $A$ — действительная симметрическая матрица линейного преобразования $f$ в ортонормированном базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Согласно теореме 8.11 существует такая ортогональная матрица $T$, что $T^{-1}AT$ — диагональная матрица. Как известно (см. теорему 7.21), ортогональное преобразование $T$ переводит ортонормированный базис $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ в ортонормированный базис $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$, в котором матрица $f$ является диагональной. Это означает, что $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ — собственные векторы преобразования $f$.

Из доказанных теорем получаем правило нахождения ортогонального преобразования, приводящего квадратичную форму $n$ переменных к каноническому виду. Это правило состоит в следующем: 1) записать матрицу данной квадратичной формы, найти ее собственные значения и $n$ попарно ортогональных собственных векторов, пронормировать их; 2) составить матрицу из ортонормированных собственных векторов-столбцов; 3) записать искомое ортогональное преобразование с помощью последней матрицы.

Пример. Найти ортогональное преобразование, приводящее к каноническому виду квадратичную форму двух переменных $x_1, x_2$:

$$f(x_1, x_2) = 11x_1^2 - 16x_1x_2 - x_2^2.$$

Поскольку в данном случае $a_{11} = 11$, $a_{12} = -8$, $a_{22} = -1$, то матрица $A$ этой квадратичной формы и ее характеристическое уравнение $\det(A - \lambda E) = 0$ запишутся так:

$$A = \begin{bmatrix} 11 & -8 \\ -8 & -1 \end{bmatrix},$$

$$\begin{vmatrix} 11-\lambda & -8 \\ -8 & -1-\lambda \end{vmatrix} = 0.$$

Характеристическое уравнение $-(11-\lambda)(1+\lambda) - 64 = 0$, или $\lambda^2 - 10\lambda + 75 = 0$, имеет корни $\lambda_1 = -5$, $\lambda_2 = 15$, которые являются собственными значениями матрицы $A$.

Найдем собственные векторы, соответствующие полученным собственным значениям. Координаты $(s, t)$ этих векторов определяются из системы уравнений (7.38), которая в данном случае имеет вид

$$\left.\begin{aligned} (11-\lambda)s - 8t &= 0, \\ -8s - (1+\lambda)t &= 0. \end{aligned}\right\}$$

При $\lambda_1 = -5$ и $\lambda_2 = 15$ имеем две системы:

$$\left.\begin{aligned} 16s - 8t &= 0, \\ -8s + 4t &= 0 \end{aligned}\right\}, \quad \left.\begin{aligned} -4s - 8t &= 0, \\ -8s - 16t &= 0. \end{aligned}\right\}$$

Из этих систем соответственно находим собственные векторы $\mathbf{u} = (s, 2s)$, где $s \neq 0$, $\mathbf{v} = (-2t, t)$, где $t \neq 0$. Положив $s = 1$, $t = 1$, получим $\mathbf{a}(1, 2)$, $\mathbf{b}(-2, 1)$. Пронормировав эти векторы, запишем их координаты в столбцы и составим матрицу $B$:

$$\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}, \quad \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

С помощью матрицы $B$ записываем искомое ортогональное преобразование

$$\left.\begin{aligned} x_1 &= \frac{1}{\sqrt{5}} y_1 - \frac{2}{\sqrt{5}} y_2, \\ x_2 &= \frac{2}{\sqrt{5}} y_1 + \frac{1}{\sqrt{5}} y_2 \end{aligned}\right\}, \text{ или } \left.\begin{aligned} x_1 &= \frac{1}{\sqrt{5}} (y_1 - 2y_2), \\ x_2 &= \frac{1}{\sqrt{5}} (2y_1 + y_2) \end{aligned}\right\}.$$

Это преобразование приводит данную квадратичную форму к каноническому виду $\varphi(y_1, y_2) = -5y_1^2 + 15y_2^2$.

## § 8.7. Упрощение уравнений фигур второго порядка на плоскости

*Фигурой второго порядка* на плоскости называется множество точек этой плоскости, декартовы координаты которых удовлетворяют уравнению второй степени

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + a_{13}x + a_{23}y + a_{33} = 0, \qquad (8.31)$$

где $a_{11}$, $a_{12}$, $a_{22}$ одновременно в нуль не обращаются. Отметим, что это множество, в частности, может состоять из единственной точки или оказаться пустым.

Первые три члена левой части уравнения (8.31) образуют квадратичную форму двух переменных $x_1 = x$, $x_2 = y$:

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 \qquad (8.32)$$

с симметрической матрицей

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}. \qquad (8.33)$$

По теореме 8.11 эту квадратичную форму ортогональным преобразованием можно привести к каноническому виду

$$f_1(x', y') = \lambda_1 x'^2 + \lambda_2 y'^2 \qquad (8.34)$$

с матрицей

$$C = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad (8.35)$$

где $\lambda_1$, $\lambda_2$ — характеристические числа матрицы $A$, т. е. корни характеристического уравнения матрицы $A$

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{12} & a_{22}-\lambda \end{vmatrix} = 0. \qquad (8.36)$$

При этом ортогональном преобразовании уравнение (8.31) примет вид

$$\lambda_1 x'^2 + \lambda_2 y'^2 + a'_{13}x' + a'_{23}y' + a'_{33} = 0. \qquad (8.37)$$

Уравнение (8.37) можно привести к каноническому виду путем выделения в левой части полных квадратов.

Фигуру второго порядка, определяемую уравнением (8.31), называют *центральной*, если $\det A \neq 0$, и *нецентральной*, когда $\det A = 0$.

Отметим, что при ортогональном преобразовании переменных определитель матрицы квадратичной формы не меняется. Это следует из того, что $C = B^T AB$, где $B$ — матрица преобразования (см. теорему 8.1), $B^T = B^{-1}$, $|\det B| = 1$ в случае, когда $B$ — ортогональная матрица. Поскольку $C = B^{-1}AB$, то $\det C = \det A$. Так как $\det C = \lambda_1\lambda_2$ (см. (8.35)), то

$$\det A = \lambda_1\lambda_2. \qquad (8.38)$$

Пусть уравнение (8.37) определяет центральную фигуру, т. е. $\det A \neq 0$. Здесь возможны два случая: 1) $\lambda_1\lambda_2 > 0$ (числа $\lambda_1$ и $\lambda_2$ одного знака), фигура называется фигурой эллиптического типа; 2) $\lambda_1\lambda_2 < 0$ (числа $\lambda_1$ и $\lambda_2$ имеют разные знаки), фигура называется фигурой гиперболического типа.

Если $\lambda_1\lambda_2 \neq 0$, то уравнение (8.37), выделив в его левой части полные квадраты, можно привести к виду

$$\lambda_1(x'-h_1)^2 + \lambda_2(y'-h_2)^2 = q$$

или

$$\lambda_1 X^2 + \lambda_2 Y^2 = q, \qquad (8.39)$$

где

$$X = x' - h_1, \quad Y = y' - h_2. \qquad (8.40)$$

Формулы (8.40) выражают зависимость между координатами $(x', y')$ и $(X, Y)$ при параллельном переносе координатных осей в точку $O_1(h_1, h_2)$.

В случае $\lambda_1\lambda_2 > 0$ уравнение (8.39) приводится к одному из канонических видов:

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1; \tag{8.41}$$

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = -1; \tag{8.42}$$

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 0 \tag{8.43}$$

в зависимости от знаков $\lambda_1$ и $q$: 1) $\lambda_1 q > 0$, 2) $\lambda_1 q < 0$, 3) $q = 0$.

Уравнение (8.41) определяет эллипс, уравнению (8.42) не удовлетворяют координаты ни одной точки плоскости (иногда говорят, что это уравнение определяет мнимый эллипс), уравнению (8.43) удовлетворяют координаты одной точки ($X = 0$, $Y = 0$).

В случае $\lambda_1 \lambda_2 < 0$ уравнение (8.39) приводится к одному из канонических видов:

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1; \tag{8.44}$$

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = -1; \tag{8.45}$$

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 0 \tag{8.46}$$

в зависимости от знаков $\lambda_1$ и $q$: 1) $\lambda_1 q > 0$, 2) $\lambda_1 q < 0$, 3) $q = 0$.

Уравнение (8.44) определяет гиперболу с действительной осью $O_1X$, уравнение (8.45) — гиперболу с действительной осью $O_1Y$, уравнение (8.46) — пару пересекающихся прямых, так как оно распадается на два уравнения:

$$\frac{X}{a} - \frac{Y}{b} = 0, \quad \frac{X}{a} + \frac{Y}{b} = 0,$$

$$\text{или } Y = \frac{b}{a} X, \quad Y = -\frac{b}{a} X.$$

Обратимся к нецентральным фигурам, т. е. к случаю, когда $\det A = 0$. В силу (8.38) из равенства $\det A = 0$ следует, что $\lambda_1 \lambda_2 = 0$. Последнее равенство означает, что одно из чисел $\lambda_1$, $\lambda_2$ равно нулю (оба числа $\lambda_1$, $\lambda_2$ в нуль обратиться не могут, так как это означало бы, что квадратичная форма (8.34) является вырожденной, чего быть не может, поскольку $a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$). Если $a'_{23} \neq 0$, то уравнение (8.37) можно привести к виду $\lambda_1 (x' - h_1)^2 + a'_{23} y' + q = 0$ и записать так:

$$\lambda_1 (x' - h_1)^2 = - a'_{23} (y' - h_2). \tag{8.47}$$

Осуществим параллельный перенос репера $(O_1, \mathbf{e}'_1, \mathbf{e}'_2)$ на вектор $\overline{OO_1} = h_1\mathbf{e}'_1 + h_2\mathbf{e}'_2$, получим новую систему координат $O_1XY$, причем $X$ и $Y$ определяются формулами (8.40). Уравнение (8.47) приведем к виду

$$X^2 = 2pY. \tag{8.48}$$

Уравнение (8.48) определяет параболу с осью $O_1Y$.

Если в уравнении (8.37) $a'_{23} = 0$ (и $\lambda_2 = 0$), то, выделив полный квадрат, его можно записать так:

$$\lambda_1 (x' - h_1)^2 + q = 0. \tag{8.49}$$

Осуществив параллельный перенос репера $(O_1, \mathbf{e}'_1, \mathbf{e}'_2)$ на вектор $\overline{OO_1} = h_1\mathbf{e}_1$, т. е. выполнив преобразование $X = x' - h_1$, $Y = y'$, получим новую систему координат $O_1XY$, в которой уравнение (8.49) принимает один из видов:

$$X^2 = a^2; \tag{8.50}$$

$$X^2 = -a^2; \tag{8.51}$$

$$X^2 = 0 \tag{8.52}$$

в зависимости от соотношения знаков чисел $\lambda_1$ и $q$: $\lambda_1 q < 0$, $\lambda_1 q > 0$, $q = 0$. Уравнение (8.50) определяет пару параллельных прямых $X = a$, $X = -a$, уравнению (8.51) не удовлетворяют координаты ни одной точки плоскости (говорят, что это уравнение определяет пару мнимых параллельных прямых), уравнение (8.52) — пару совпавших прямых $X = 0$, $X = 0$.

Операция перехода от уравнения (8.31) к уравнению (8.37) называется отнесением фигуры к главным осям. Новые оси координат параллельны осям симметрии фигуры. Главными направлениями фигуры, заданной уравнением (8.31), называют направления ортогональных собственных векторов матрицы квадратичной формы, соответствующей этому уравнению.

Из теорем § 8.6 следует, что существует декартова прямоугольная система координат, в которой уравнение (8.31) принимает канонический вид. Чтобы выбрать эту систему координат, необходимо сделать следующее: 1. Найти ортогональное преобразование, приводящее к каноническому виду квадратичную форму, соответствующую данному уравнению. 2. С помощью этого преобразования определить главные направления фигуры, т. е. векторы $\mathbf{e}_1'$, $\mathbf{e}_2'$ —ортонормированные собственные векторы матрицы указанной квадратичной формы. 3. Найти уравнение фигуры в репере $(O, \mathbf{e}_1', \mathbf{e}_2')$. 4. Выделить полные квадраты в полученном уравнении. 5. Совершить параллельный перенос системы $(O, \mathbf{e}_1', \mathbf{e}_2')$ на соответствующий вектор $\overline{OO_1}$ и составить каноническое уравнение фигуры в репере $(O_1, \mathbf{e}_1', \mathbf{e}_2')$.

Пример. Привести к каноническому виду уравнение

$$11x^2-16xy-y^2-26x-22y-61=0$$

и построить фигуру, определяемую этим уравнением.

Данному уравнению соответствует квадратичная форма $f(x, y)= =11x^2-16xy-y^2$ переменных $x$, $y$, которая с точностью до обозначений переменных $(x_1=x,\ x_2=y,\ y_1=x',\ y_2=y')$ в примере § 8.6 приведена к каноническому виду $\varphi(x', y')=-5x'^2+15y'^2$ посредством ортогонального преобразования

$$\left.\begin{aligned} x&=\frac{1}{\sqrt{5}}x'-\frac{2}{\sqrt{5}}y' \\ y&=\frac{2}{\sqrt{5}}x'+\frac{1}{\sqrt{5}}y' \end{aligned}\right\},\ \text{или}\ \left.\begin{aligned} x&=\frac{1}{\sqrt{5}}(x'-2y') \\ y&=\frac{1}{\sqrt{5}}(2x'+y') \end{aligned}\right\}. \tag{I}$$

Это преобразование переводит данное уравнение в уравнение

$$11\cdot\frac{1}{5}(x'-2y')^2-16\cdot\frac{1}{5}(x'-2y')(2x'+y')-$$

$$-\frac{1}{5}(2x'+y')^2-26\cdot\frac{1}{\sqrt{5}}(x'-2y')-22\cdot\frac{1}{\sqrt{5}}(2x'+y')-61=0,$$

или

$$-5x'^2+15y'^2-\frac{70}{\sqrt{5}}x'+\frac{30}{\sqrt{5}}y'-61=0.$$

Выделяя полные квадраты, приводим это уравнение к виду

$$-5\left(x'+\frac{7}{\sqrt{5}}\right)^2+15\left(y'+\frac{1}{\sqrt{5}}\right)^2-15=0. \tag{II}$$

Перейдем к новому координатному реперу $(O_1, \mathbf{e}_1', \mathbf{e}_2')$, где

$$\mathbf{e}_1'=\left(\frac{1}{\sqrt{5}},\ \frac{2}{\sqrt{5}}\right),\quad \mathbf{e}_2'=\left(-\frac{2}{\sqrt{5}},\ \frac{1}{\sqrt{5}}\right),$$

полученному параллельным переносом старого репера на вектор $\overline{OO_1}$. Новые координаты $(X, Y)$ связаны со старыми координатами формулами

$$X = x' + \frac{7}{\sqrt{5}}, \quad Y = y' + \frac{1}{\sqrt{5}}. \qquad \text{(III)}$$

В новом репере уравнение (II) примет вид

$$-5X^2+15Y^2=15, \text{ или } 3Y^2-X^2=3. \qquad \text{(IV)}$$

Координаты нового начала (точки $O_1$) найдем из формул (III), (I):

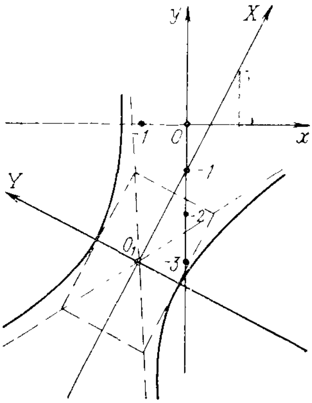



Рис. 8.1

$$X = 0, \quad Y = 0,$$

$$x' + \frac{7}{\sqrt{5}} = 0,$$

$$y' + \frac{1}{\sqrt{5}} = 0,$$

$$x' = -\frac{7}{\sqrt{5}},$$

$$y' = -\frac{1}{\sqrt{5}},$$

$$x = \frac{1}{\sqrt{5}}\left(-\frac{7}{\sqrt{5}} + \frac{2}{\sqrt{5}}\right) = -1,$$

$$y = \frac{1}{\sqrt{5}}\left(-\frac{14}{\sqrt{5}} - \frac{1}{\sqrt{5}}\right) = $$

$$= -3, \quad O_1(-1, -3).$$

Вектор $\mathbf{e}_1'$ определяет направление оси $O_1X$, он образует с осью $Ox$ угол $\alpha$, для которого $\operatorname{tg}\alpha = 2$. Строим систему координат $O_1XY$ (рис. 8.1) и относительно этой системы гиперболу, определяемую уравнением (IV).

## § 8.8. Упрощение уравнений фигур второго порядка в пространстве

*Фигурой второго порядка в пространстве* называется множество точек пространства, декартовы координаты которых удовлетворяют уравнению

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+ \\ +2a_{23}yz+a_{14}x+a_{24}y+a_{34}z+a_{44}=0, \qquad (8.53)$$

где $a_{11}^2 + a_{22}^2 + a_{33}^2 + a_{12}^2 + a_{13}^2 + a_{23}^2 \neq 0$.

Сумма первых шести членов левой части уравнения (8.53) представляет собой квадратичную форму трех переменных $x$, $y$, $z$:

$$f(x, y, z) = a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz \qquad (8.54)$$

с симметрической матрицей

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}. \quad (8.55)$$

Фигура второго порядка называется *центральной,* если $\det A \neq 0$, и *нецентральной,* если $\det A = 0$.

Как известно, с помощью ортогонального преобразования квадратичную форму (8.54) можно привести к каноническому виду $\varphi(x', y', z') = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$, где $\lambda_1, \lambda_2, \lambda_3$ — корни характеристического уравнения $\det(A - \lambda E) = 0$. Матрица квадратичной формы $\varphi = \varphi(x', y', z')$ имеет вид

$$C=\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \quad (8.56)$$

Указанное ортогональное преобразование уравнение (8.53) приводит к уравнению

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + a'_{14} x' + a_{24} y' + a_{34} z' + a'_{44} = 0. \quad (8.57)$$

**Центральные фигуры.** Если $\det A \neq 0$, то $\det C = \lambda_1 \lambda_2 \lambda_3 \neq 0$, ибо $\det A = \det C$. Выделяя полные квадраты в левой части уравнения (8.57), можно привести его к виду

$$\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2 = \mu, \quad (8.58)$$

где $X = x' - h_1$, $Y = y' - h_2$, $Z = z' - h_3$.

Поскольку $\lambda_1 \lambda_2 \lambda_3 \neq 0$, то ни одно из чисел не равно нулю, все эти числа могут иметь один знак ($\lambda_1 \lambda_2 \lambda_3 > 0$) или только два из них одного знака.

1. Если все числа $\lambda_1, \lambda_2, \lambda_3$ одного знака, то уравнение (8.58) можно привести к одному из следующих канонических видов:

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 1; \quad (8.59)$$

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = -1; \quad (8.60)$$

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 0 \quad (8.61)$$

в зависимости от $\lambda_1$ и $\mu$: $\lambda_1 \mu > 0$, $\lambda_1 \mu < 0$, $\mu = 0$.

Уравнение (8.59) определяет эллипсоид, уравнению (8.60) не удовлетворяют координаты ни одной точки пространства (иногда говорят, что это уравнение определяет мнимый эллипсоид), уравнению (8.61) удовлетворяют координаты единственной точки ($X=0$, $Y=0$, $Z=0$).

2. Пусть знак одного из этих чисел противоположен знаку двух других; предположим, что $\lambda_1\lambda_2>0$. Уравнение (8.58) можно привести к одному из канонических видов:

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}-\frac{Z^2}{c^2}=1; \tag{8.62}$$

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}-\frac{Z^2}{c^2}=-1; \tag{8.63}$$

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}-\frac{Z^2}{c^2}=0 \tag{8.64}$$

в зависимости от $\lambda_1$ и $\mu$: $\lambda_1\mu>0$, $\lambda_1\mu<0$, $\mu=0$.

Уравнения (8.62), (8.63), (8.64) определяют соответственно однополостный гиперболоид, двухполостный гиперболоид и конус второго порядка.

**Нецентральные фигуры.** Если $\det A=0$, или $\lambda_1\lambda_2\lambda_3=0$, то одно или два из этих чисел равны нулю.

1. Пусть $\lambda_3=0$, $a'_{34}\neq 0$, тогда уравнение (8.57) приводится к виду

$$\lambda_1X^2+\lambda_2Y^2+\mu Z=0. \tag{8.65}$$

Если $\lambda_1\lambda_2>0$ и $\lambda_1\mu<0$, то имеем

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=2Z; \tag{8.66}$$

в случае $\lambda_1\lambda_2<0$, $\lambda_1\mu<0$ получаем

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=2Z. \tag{8.67}$$

Уравнения (8.66) и (8.67) определяют соответственно эллиптический параболоид и гиперболический параболоид.

2. Пусть $\lambda_3=0$ и $a'_{34}=0$, тогда имеем уравнение

$$\lambda_1X^2+\lambda_2Y^2+\nu=0, \tag{8.68}$$

которое приводится к одному из следующих канонических видов:

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=1; \tag{8.69}$$

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=-1; \tag{8.70}$$

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=1; \tag{8.71}$$

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=-1; \tag{8.72}$$

$$\frac{X^2}{a^2}-\frac{Y^2}{b^2}=0. \tag{8.73}$$

Уравнение (8.69) определяет эллиптический цилиндр, а каждое из уравнений (8.71), (8.72) — гиперболический цилиндр, уравнение (8.73) — пару пересекающихся плоскостей; уравнению (8.70) не удовлетворяют координаты ни одной точки (говорят, что это уравнение определяет мнимый эллиптический цилиндр).

3. Если $\lambda_2=\lambda_3=0$ и $a'_{24}\neq 0$, то уравнение (8.57) приводится к виду $\lambda_1 X^2+\mu Y=0$, или

$$X^2=2pY, \tag{8.74}$$

и определяет параболический цилиндр.

4. Если $\lambda_2=\lambda_3=0$ и $a'_{24}=0$, имеем уравнение $\lambda_1 X^2+$ $+\nu=0$, которое приводится к одному из канонических видов:

$$X^2=a^2; \tag{8.75}$$

$$X^2=-a^2; \tag{8.76}$$

$$X^2=0. \tag{8.77}$$

Уравнение (8.75) определяет пару параллельных плоскостей ($X=a$, $X=-a$), уравнение (8.77) — пару совпавших плоскостей; уравнению (8.76) не удовлетворяют координаты ни одной точки пространства (говорят, что оно определяет пару мнимых параллельных плоскостей).

# Глава 9

# ГРУППЫ

В этой главе рассматриваются простейшие факты, связанные с понятием группы. Понятие группы относится к числу центральных понятий абстрактной алгебры — одной из наиболее важных и быстро развивающихся областей современной математики.

## § 9.1. Некоторые общие свойства операций над числами, векторами, матрицами и другими объектами

Рассмотрим множество $Z$ всех целых чисел. Сумма любых двух целых чисел $a$ и $b$ также является целым числом, причем $a+b=b+a$. Операция сложения обладает следующими свойствами: 1) для любых трех целых чисел $a$, $b$, $c$ $(a+b)+c=a+(b+c)$; 2) если одно из слагаемых равно нулю, то сумма равна другому слагаемому, т. е. $a+0=a$, $0+a=a$; 3) для каждого целого числа $a$ существует противоположное ему, т. е. такое целое число $-a$, что $a+(-a)=0$, $(-a)+a=0$.

Очевидно, указанными свойствами обладают множества $Q$ всех рациональных чисел, множество $R$ всех действительных чисел, множество $C$ всех комплексных чисел.

Рассмотрим множество всех действительных чисел, отличных от нуля. Произведение любых двух таких чисел $a$ и $b$ будет также отличным от нуля, т. е. принадлежит данному множеству, причем $ab=ba$. Операция умножения этих чисел обладает следующими свойствами: 1) для любых трех чисел $a$, $b$, $c$ справедливо равенство $a(bc)=(ab)c$; 2) если один из множителей равен единице, то произведение равно другому множителю, т. е. $a\cdot 1=$ $=a$, $1\cdot a=a$; 3) для любого числа $a$ этого множества $(a\neq 0)$ существует обратное, т. е. такое число $a^{-1}$, что $aa^{-1}=a^{-1}a=1$.

Легко видеть, что аналогичными свойствами обладают и другие множества: всех положительных рациональных чисел, всех рациональных чисел, отличных от нуля, всех положительных действительных чисел, всех комплексных чисел, отличных от нуля.

Рассмотрим множество векторов в трехмерном пространстве. Операция сложения векторов состоит в том, что каждой паре векторов $\mathbf{a}$ и $\mathbf{b}$ ставится в соответствие третий вектор (их сумма), причем $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$. Сложение векторов, как известно, обладает следующими свойствами: 1) $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$; 2) $\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}$; 3) $\mathbf{a}+$ $+(-\mathbf{a})=\mathbf{0}$, $(-\mathbf{a})+\mathbf{a}=\mathbf{0}$.

Аналогичными свойствами обладает операция сложения над векторами $\mathbf{x}=(x_1, x_2, \ldots, x_n)$, $\mathbf{y}=(y_1, y_2, \ldots, y_n)$ линейного $n$-мерного пространства.

Складывать можно не только числа и векторы, но и матрицы. Пусть дано множество всех матриц размеров $m\times n$, где $m$ и $n$ — натуральные числа. Операция сложения таких матриц состоит в том, что каждой паре матриц $A$ и $B$ по определенному закону ставится в соответствие третья матрица — их сумма $A+B$, причем $A+B=B+A$. Сложение матриц обладает следующими свойствами (см. § 3.2): 1) $(A+B)+C=A+(B+C)$; 2) $A+O=O+A=A$;

3) $A+(-A)=O$, где $O$ — нулевая матрица; $(-A)$ — матрица, противоположная матрице $A$.

Матрицы также можно и умножать. Пусть дано множество невырожденных квадратных матриц порядка $n$. Произведение двух таких матриц $A$ и $B$ будет невырожденной квадратной матрицей того же порядка. Хотя в общем случае $AB \neq BA$, операция умножения матриц обладает следующими свойствами (см. § 3.2): 1) $(AB)C=$ $=A(BC)$; 2) $AE=EA=A$; 3) $AA^{-1}=A^{-1}A=E$, где $E$ — единичная матрица; $A^{-1}$ — матрица, обратная матрице $A$.

Операции сложения и умножения определены также для линейных преобразований линейных векторных пространств. Операция сложения линейных преобразований обладает теми же свойствами, что и операция сложения матриц. Операция умножения невырожденных линейных преобразований обладает теми же свойствами, что и операция умножения невырожденных квадратных матриц.

Таким образом, соответствующей операции (умножения или сложения) над элементами каждого из рассмотренных множеств характерны следующие свойства: 1) ассоциативность операции; 2) существование так называемого нейтрального элемента (единицы или нуля), который при данной операции не меняет любого элемента; 3) существование для каждого элемента $a$ элемента $a'$ (обратного или противоположного) такого, что операция над $a$ и $a'$ дает нейтральный элемент.

## § 9.2. Понятие группы. Примеры групп

*Группой* называется множество $G$ элементов $a, b, c, \ldots$ $\ldots$, для которых определена операция (сложения или умножения), каждой упорядоченной паре $(a, b)$ элементов $G$ ставящая в соответствие единственный элемент $c=a \circ b$ данного множества, причем операция обладает следующими свойствами:

1) для любых $a, b, c \in G$

$$a \circ (b \circ c) = (a \circ b) \circ c; \qquad (9.1)$$

2) в $G$ существует нейтральный элемент, т. е. такой элемент $e$, что для любого элемента $a \in G$

$$a \circ e = e \circ a = a; \qquad (9.2)$$

3) для каждого элемента $a \in G$ существует обратный ему элемент, т. е. такой элемент $a^{-1}$, что

$$a \circ a^{-1} = e, \ a^{-1} \circ a = e. \tag{9.3}$$

Если, кроме того, для любых $a, b \in G$ выполняется условие

$$a \circ b = b \circ a, \tag{9.4}$$

то группа называется *коммутативной,* или *абелевой,* группой.

Можно доказать, что в любой группе нейтральный элемент определен однозначно; для каждого элемента существует единственный обратный элемент.

Группа, состоящая из конечного числа элементов, называется *конечной.* Число элементов группы называют ее *порядком.* Группа, не являющаяся конечной, называется *бесконечной.*

Группа называется *аддитивной,* или группой по сложению, когда групповая операция, ставящая упорядоченной паре элементов $(a, b)$ элемент $c = a \circ b$, является сложением. В этом случае символ операции $\circ$ заменяется знаком «+», $c = a + b$, нейтральный элемент называют *нулем* и обозначают символом 0; $a + 0 = 0 + a = a$. Элемент, «обратный» к элементу $a$, называют *противоположным* и обозначают $-a$: $a + (-a) = (-a) + a = 0$.

Группа называется *мультипликативной,* или группой по умножению, когда групповая операция, ставящая упорядоченной паре $(a, b)$ элемент $c = a \circ b$, является умножением. В данном случае произведение $a \circ b$ обозначается $a \cdot b$, или $ab$; нейтральный элемент называется *единицей* и обозначается символом 1: $a \cdot 1 = 1 \cdot a = a$.

Пользуясь ассоциативным свойством, можно определить произведение (сумму) любого числа элементов. Поскольку $a(bc) = (ab)c$, то имеет смысл говорить о произведении $abc$ трех элементов, равном по определению $(ab)c = a(bc)$.

Произведение $n$ элементов, равных $a$, называют $n$-й степенью элемента $a$ и обозначают $a^n$. Отрицательные степени элемента $a$ можно определить или как элементы группы $G$, обратные положительным степеням этого элемента, или как произведения соответствующего числа множителей, равных элементу $a^{-1}$. Указанные определения совпадают, так как верно равенство $(a^n)^{-1} = (a^{-1})^n$ $(n > 0)$. Действительно, $a^n (a^{-1})^n = a \cdot a \ldots a \cdot a^{-1} \cdot a^{-1} \ldots a^{-1} =$

$=e$, т. е. $a^n(a^{-1})^n=e$, откуда следует, что $(a^{-1})^n$ — элемент, обратный элементу $a^n$, или $(a^{-1})^n=(a^n)^{-1}$. Эти равные элементы обозначим $a^{-n}$. Условимся под нулевой степенью $a^0$ элемента $a$ понимать элемент 1, т. е. $a^0=1$. Очевидно, в любой группе $G$ для степеней каждого элемента $a$ при любых показателях $m$ и $n$ (положительных, отрицательных или нулевых) выполняются равенства $a^na^m=$ $=a^ma^n=a^{n+m}$, $(a^n)^m=a^{nm}$.

Отметим, что если операция в группе называется сложением, то вместо степеней элемента $a$ говорят о кратных этого элемента и пишут $na$.

В каждой мультипликативной группе однозначно разрешимы уравнения $ax=b$, $ya=b$, первое из них имеет решение $x=a^{-1}b$, второе — $y=ba^{-1}$. В самом деле, умножив первое уравнение слева на $a^{-1}$, получим $a^{-1}ax=a^{-1}b$. Поскольку $a^{-1}a=e$ и $ex=x$, то $x=a^{-1}b$. Умножая второе уравнение справа на $a^{-1}$, находим $yaa^{-1}=ba^{-1}$, $y=ba^{-1}$. Если группа является коммутативной, то эти уравнения не различаются, они имеют одинаковые решения $x=y=$ $=a^{-1}b$.

Принимая во внимание определение группы, свойства операций над элементами множеств, рассмотренных в § 9.1, получаем следующие примеры групп.

1. Множество всех целых чисел с операцией сложения образует аддитивную группу. Действительно, сумма $a+b$ двух целых чисел $a$ и $b$ также будет целым числом. В этом случае говорят, что множество целых чисел замкнуто относительно операции сложения. Сложение целых чисел коммутативно: $(a+b)+c=a+(b+c)$. В данном множестве имеется нейтральный элемент, т. е. число 0 такое, что $a+0=a$ при любом целом числе $a$. Для каждого элемента — целого числа $a$ существует обратный элемент — противоположное число, т. е. такое число $-a$, что $a+(-a)=0$. Рассматриваемая группа коммутативна, так как $a+b=b+a$.

Замечание 1. Множество всех целых чисел не образует группу по умножению, так как обратные для целых чисел (отличных от $-1$ и 1) не являются целыми числами. Например, для числа 2 обратное число $2^{-1}$ не принадлежит множеству целых чисел.

2. Множество всех действительных чисел, отличных от нуля, с операцией умножения образует мультипликативную группу. Эта группа коммутативна, так как $ab=ba$.

Замечание 2. Множество всех действительных чисел не образует группу по умножению, поскольку для числа 0 нет обратного.

3. Множество всех векторов трехмерного пространства образует группу по сложению. Эта группа коммутативна ($\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$).

4. Множество матриц размеров $m\times n$ образует коммутативную группу по сложению ($A+B=B+A$). Для матрицы $A$ «обратным» элементом является матрица $(-A)$; нейтральный элемент — нулевая матрица $O$.

5. Множество всех невырожденных квадратных матриц порядка $n$ образует группу по умножению. Эта группа не является коммутативной (в общем случае $AB\neq BA$).

Замечание 3. Множество всех квадратных матриц порядка $n$ не образует группу по умножению, так как для некоторых его элементов нет обратных (вырожденная матрица не имеет обратной).

6. Множество, состоящее только из двух чисел $+1$, $-1$, образует группу по умножению. Действительно, каждое из произведений $(+1)(-1)=-1$, $(+1)(+1)=+1$, $(-1)(-1)=+1$ принадлежит данному множеству. Можно показать, что умножение ассоциативно. Существует единица — число $+1$, которое удовлетворяет условию $(-1)(+1)=-1$, $(+1)(+1)=+1$. Для каждого элемента существует обратный: каждое из этих двух чисел совпадает со своим обратным.

Замечание 4. Множество, состоящее из двух чисел $+1$, $-1$, не образует группу по сложению, так как сумма $(+1)+(-1)=0$, а число 0 не принадлежит данному множеству. (В таком случае говорят, что данное множество не является замкнутым относительно операции сложения.)

7. Множество, состоящее из одного элемента 0, образует группу по сложению. Действительно, $0+0=0$, сумма принадлежит данному множеству. Свойства операции сложения очевидны.

Замечание 5. Множество, состоящее из одного элемента 1, образует группу по умножению.

Замечание 6. Группа, образованная одним элементом, называется *единичной*.

## § 9.3. Подгруппа

*Подгруппой* группы $G$ называется подмножество $H$ ее элементов, образующее группу относительно операции, определенной в $G$. Чтобы убедиться в том, что подмножество $H$ группы $G$ является ее подгруппой, необходимо проверить, что: 1) произведение (сумма) любых двух элементов $a, b \in H$ принадлежит $H$; 2) для любого элемента $a \in H$ обратный элемент принадлежит $H$. Это будет и достаточно, так как ассоциативный закон выполняется для любых трех элементов $G$, в том числе и для элементов $H$, а нейтральный элемент $e$ (1 или 0) будет принадлежать $H$ (как произведение $aa^{-1}$ или сумма $a+(-a)$).

Приведем примеры подгрупп некоторых групп.

I. Множество всех действительных чисел является группой по сложению, так как: 1) сумма двух действительных чисел будет действительным числом; 2) для любых трех действительных чисел $a$, $b$, $c$ справедливо равенство $(a+b)+c=a+(b+c)$; 3) существует нейтральный элемент — число 0 такое, что $a+0=a$ при любом $a$; 4) для любого элемента — действительного числа $a$ существует «обратный» элемент — противоположное число $-a$ такое, что $a+(-a)=0$.

Подгруппами аддитивной группы всех действительных чисел являются, в частности, следующие: 1) аддитивная группа рациональных чисел; 2) аддитивная группа целых чисел; 3) аддитивная группа целых чисел, кратных числу $k$, в частности, аддитивная группа четных чисел. Действительно, каждое из этих множеств (всех рациональных; всех целых; всех целых чисел, кратных $k$; всех четных чисел) является подмножеством множества всех действительных чисел, кроме того: 1) сумма двух рациональных (целых; целых, кратных числу $k$; четных) чисел будет рациональным (целым; целым, кратным числу $k$; четным) числом; 2) для каждого рационального (целого, кратного $k$, четного) числа $a$ существует соответствующее противоположное число $(-a)$.

Отметим, что группа целых чисел — подгруппа группы рациональных чисел.

Замечание 1. Множество нечетных чисел не образует группу по сложению, так как сумма двух нечетных чисел будет четным числом (и не принадлежит данному множеству).

II. Мультипликативная группа всех действительных чисел, отличных от нуля, имеет, в частности, следующие подгруппы: 1) мультипликативную группу положительных действительных чисел; 2) мультипликативную группу рациональных чисел, отличных от нуля; 3) множество, состоящее из двух чисел $+1$, $-1$ с операцией умножения.

Замечание 2. Мультипликативная группа положительных действительных чисел не является подгруппой аддитивной группы всех действительных чисел, так как групповые операции в рассматриваемых множествах разные (соответственно умножение, сложение).

III. Мультипликативная группа невырожденных матриц порядка $n$ (которую называют *полной линейной группой*) имеет, в частности, подгруппы: 1) группу ортогональных матриц; 2) группу диагональных матриц; 3) группу верхних треугольных матриц; 4) группу матриц с положительным определителем; 5) группу матриц с определителем, равным единице.

Легко видеть, что пересечение двух подгрупп группы $G$ есть подгруппа в $G$. Например, в аддитивной группе целых чисел пересечение подгруппы четных чисел и подгруппы чисел, кратных трем, будет подгруппой чисел, кратных шести.

Очевидно, каждая группа является своей подгруппой. Далее, каждая группа имеет единичную подгруппу (см. § 9.2, пример 7, замечания 5 и 6), состоящую из одного нейтрального элемента (единицы или нуля). Эти две подгруппы называются *несобственными* (или *тривиальными*) подгруппами. Остальные подгруппы называются *собственными* (или *истинными*) подгруппами. Отметим, что в любой группе все подгруппы каждой из подгрупп будут в то же время подгруппами исходной группы. Например, аддитивная группа целых чисел является подгруппой аддитивной группы рациональных чисел, которая, в свою очередь, есть подгруппа аддитивной группы всех действительных чисел; аддитивная группа целых чисел — подгруппа аддитивной группы всех действительных чисел.

## § 9.4. Группы преобразований. Симметрическая группа $n$-й степени

*Преобразованием* множества $X$ называется взаимно-однозначное отображение этого множества на себя. Преобразование множества $X$ обозначим буквой $P$. Определение преобразования $P$ множества $X$ означает следующее: любому элементу $x \in X$ ставится в соответствие единственный элемент $x' = Px$ того же множества; $x'$ называют образом элемента $x$, а $x$ — прообразом $x'$. Каждый элемент $x' \in X$ имеет единственный прообраз $x \in X$.

*Умножением преобразований* называется последовательное их выполнение. Произведение двух преобразований $P$, $Q$ обозначается $QP$ (справа записано то преобразование, которое выполняется первым); по определению $(QP)x=Q(Px)$. Очевидно, произведение двух преобразований данного множества есть преобразование этого множества. Отметим, что в общем случае умножение не является коммутативным, т. е. $QP \neq PQ$. Можно показать, что произведение преобразований подчиняется ассоциативному закону. Роль единицы при умножении преобразований выполняет тождественное преобразование $E$, ставящее в соответствие каждому элементу множества его самого. Для каждого преобразования $P$ существует обратное преобразование $P^{-1}$, которое каждому элементу $x' \in X$ ставит в соответствие его единственный прообраз $x \in X$, причем $PP^{-1}=P^{-1}P=E$. Следовательно, множество преобразований $P$ данного множества $X$ образует группу.

Если множество $X$ конечно и состоит из $n$ элементов, то всевозможные взаимно-однозначные отображения этого множества на себя называются *подстановками*. Подстановку из $n$ элементов можно обозначить так:

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \end{pmatrix},$$

где $\alpha_1, \alpha_2, \ldots, \alpha_n$ — те же числа $1, 2, 3, \ldots, n$, обозначающие данные элементы и записанные в другом порядке.

Приведем некоторые конкретные примеры подстановок при $n=5$:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 4 & 5 & 3 \\ 4 & 3 & 2 & 1 & 5 \end{pmatrix}.$$

Первая подстановка означает такое взаимно-однозначное отображение множества (1, 2, 3, 4, 5) на себя, при котором 1 переходит в 5, 2 — в 4 и т. д. Вторая подстановка называется *тождественной,* каждый элемент соответствует сам себе. Равенство двух других подстановок показывает, что расположение столбцов в записи подстановки не играет роли. Подстановки, отличающиеся только порядком следования столбцов, не считаются различными.

*Умножением подстановок* называют последовательное их выполнение (сначала правого сомножителя, затем левого). Умножение подстановок ассоциативно, но не коммутативно. Например, если

$$P=\begin{pmatrix}1&2&3&4\\2&4&1&3\end{pmatrix},\quad Q=\begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix},$$

то

$$PQ=\begin{pmatrix}1&2&3&4\\4&1&3&2\end{pmatrix},\quad QP=\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix},\quad PQ\neq QP.$$

Единицей при умножении подстановок из $n$ элементов служит тождественная подстановка

$$\begin{pmatrix}1&2&3&\dots&n\\1&2&3&\dots&n\end{pmatrix}.$$

Каждая подстановка из $n$ элементов имеет обратную:

$$\begin{pmatrix}1&2&3&\dots&n\\\alpha_1&\alpha_2&\alpha_3&\dots&\alpha_n\end{pmatrix}^{-1}=\begin{pmatrix}\alpha_1&\alpha_2&\alpha_3&\dots&\alpha_n\\1&2&3&\dots&n\end{pmatrix}.$$

Чтобы получить подстановку, обратную данной, необходимо поменять местами строки.

Множество подстановок из $n$ элементов относительно введенной операции умножения образует группу. Группа подстановок из $n$ элементов называется *симметрической группой* $n$-й степени и обозначается $S_n$. Очевидно, число подстановок из $n$ элементов равно $n!$, поэтому группа $S_n$ имеет порядок $n!$

Рассмотрим группу подстановок из трех элементов $a$, $b$, $c$. Поскольку из трех элементов можно составить всего шесть различных перестановок $abc$, $acb$, $bac$, $bca$, $cab$, $cba$, то и число различных подстановок для них равно шести ($n=3$, $3!=1\cdot2\cdot3=6$).

Обозначим эти подстановки следующим образом:

$$P_1=\begin{pmatrix}a&b&c\\a&b&c\end{pmatrix},\quad P_2=\begin{pmatrix}a&b&c\\a&c&b\end{pmatrix},\quad P_3=\begin{pmatrix}a&b&c\\c&b&a\end{pmatrix},$$

$$P_4=\begin{pmatrix}a&b&c\\b&a&c\end{pmatrix},\quad P_5=\begin{pmatrix}a&b&c\\b&c&a\end{pmatrix},\quad P_6=\begin{pmatrix}a&b&c\\c&a&b\end{pmatrix}.$$

Отметим, что $P_1$ — тождественная подстановка; для каждой подстановки существует обратная ей: $P_1^{-1}=P_1$, $P_2^{-1}=P_2$, $P_3^{-1}=P_3$, $P_4^{-1}=P_4$, $P_5^{-1}=P_6$, $P_6^{-1}=P_5$.

Группа $S_3$ (симметрическая группа подстановок из трех элементов) некоммутативна, поскольку, например, $P_4P_5=P_2$, $P_5P_4=P_3$, т. е. $P_4P_5\neq P_5P_4$.

Группу $S_3$ можно представить следующей таблицей умножения, в которой слева стоят левые множители $P_i$, сверху — правые $P_i$, а на пересечении соответствующих строки и столбца — их произведение. Таблицы такого рода называют таблицами Кэли (табл. 2).

Таблица 2

| | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
|---|---|---|---|---|---|---|
| $P_1$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| $P_2$ | $P_2$ | $P_1$ | $P_5$ | $P_6$ | $P_3$ | $P_4$ |
| $P_3$ | $P_3$ | $P_6$ | $P_1$ | $P_5$ | $P_4$ | $P_2$ |
| $P_4$ | $P_4$ | $P_5$ | $P_6$ | $P_1$ | $P_2$ | $P_3$ |
| $P_5$ | $P_5$ | $P_4$ | $P_2$ | $P_3$ | $P_6$ | $P_1$ |
| $P_6$ | $P_6$ | $P_3$ | $P_4$ | $P_2$ | $P_1$ | $P_5$ |

## § 9.5. Группа вращений правильного многоугольника. Циклические группы. Группа симметрий правильного треугольника

Пусть дан правильный $n$-угольник $A_1A_2\ldots A_n$ с центром в точке $O$ (рис. 9.1 соответствует случаю $n=6$). Рассмотрим всевозможные повороты плоскости вокруг точки $O$, при которых правильный $n$-угольник совмещается сам с собой. Таких поворотов будет $n$: $a_0$ — поворот на угол $\varphi_0=0$ (тождественное преобразование); $a_1$ — поворот на угол $\varphi_1 = 2\pi : n$; $a_2$ — поворот на угол $\varphi_2 = \frac{2\pi}{n}\cdot 2$, $\ldots$; $a_{n-1}$ — поворот на угол $\varphi_{n-1} = \frac{2\pi}{n}(n-1)$. Под *умножением поворотов* будем понимать последовательное их выполнение одного за другим: $a_k \circ a_l = a_{k+l}$, причем $a_{k+n} = a_k$ при любом $k$ $(k=0, 1, 2, \ldots, n)$, в частности, $a_n = a_0$. Умножение поворотов ассоциативно (и коммутативно). Множество указанных поворотов правильного многоугольника образует группу по умножению, роль единицы играет тождественное преобразование — поворот $a_0$. Для каждого элемента $a_k$ существует обратный элемент

$a_k^{-1} = a_{n-k}$ при всех $k$ ($k = 0, 1, 2, \ldots, n-1$), так как $a_k \circ a_{n-k} = a_n = a_0$, т. е. $a_k \circ a_{n-k} = a_0$, где $a_0$ — единичный элемент.

Положим $a_1 = a$, тогда $a_2 = a^2$, $a_3 = a^3$, ..., $a_{n-1} = a^{n-1}$, $a_n = a^n = a_0$. В этом случае говорят, что группа образована степенями одного из своих элементов (или что она «по-

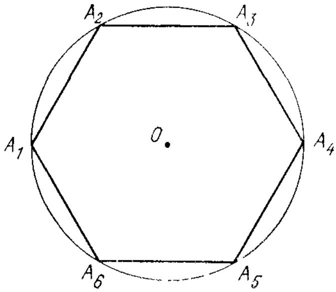



Рис. 9.1

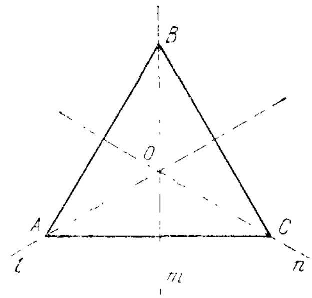



Рис. 9.2

рождается» одним из своих элементов); таким элементом здесь будет элемент $a = a_1$. Группы, образованные степенями одного из своих элементов, называются *циклическими*. Таким образом, группа вращений правильного $n$-угольника является циклической группой порядка $n$, эта группа обозначается $C_n$. Отметим, что аддитивная группа целых чисел также будет циклической, она порождается одним из своих элементов — числом 1: $2 = 1+1$, $3 = (1+1)+1$ и т. д. Эта группа является бесконечной циклической группой, ее обозначают $C_\infty$.

Пусть дан правильный треугольник $ABC$ с центром в точке $O$ (рис. 9.2). Рассмотрим все симметрии данной фигуры, т. е. те преобразования плоскости, при которых этот треугольник переходит в себя (или самосовмещается). К ним относятся: три поворота $\varphi_0$, $\varphi_1$, $\varphi_2$ плоскости вокруг точки $O$ соответственно на углы $0$, $2\pi/3$, $4\pi/3$ (частный случай рассмотренных выше вращений правильного $n$-угольника при $n=3$), три осевые симметрии $\varphi_3$, $\varphi_4$, $\varphi_5$, определяемые соответственно осями симметрии $l$, $m$, $n$, проходящими через вершину правильного треугольника и середину его противоположной стороны (рис. 9.2).

Будем характеризовать каждое самосовмещение $\varphi$ подстановкой на множестве вершин $A$, $B$, $C$ правильного треугольника

$$\begin{pmatrix} A & B & C \\ \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix},$$

где $\alpha_1$, $\alpha_2$, $\alpha_3$ — те же буквы $A$, $B$, $C$, взятые в некотором порядке. Принятое нами соответствие между самосовмещениями треугольника и подстановками множества его вершин дает

$$\varphi_0 = \begin{pmatrix} A & B & C \\ A & B & C \end{pmatrix}, \quad \varphi_1 = \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}, \quad \varphi_2 = \begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix},$$

$$\varphi_3 = \begin{pmatrix} A & B & C \\ A & C & B \end{pmatrix}, \quad \varphi_4 = \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}, \quad \varphi_5 = \begin{pmatrix} A & B & C \\ B & A & C \end{pmatrix}.$$

Очевидно, множество самосовмещений $\varphi_0$, $\varphi_1$, $\varphi_2$, $\varphi_3$, $\varphi_4$, $\varphi_5$ образует группу относительно умножения (последовательного выполнения двух самосовмещений). Роль единицы играет тождественное преобразование, каждый элемент данного множества имеет обратный. Эта группа называется *группой симметрий треугольника.* (Сравните ее с симметрической группой $S_3$.)

## § 9.6. Изоморфизм групп

Вернемся к некоторым примерам, приведенным в двух предыдущих параграфах. Рассмотрим симметрическую группу $S_3$ подстановок из трех элементов $a$, $b$, $c$ (см. § 9.4) и группу симметрий правильного треугольника с вершинами $A$, $B$, $C$ (см. § 9.5). Эти группы отличаются только обозначениями элементов и названиями соответствующих преобразований. Такие группы называют изоморфными, их можно считать одинаковыми, так как они не различаются между собой с точки зрения их «групповых» свойств (т. е. свойств, изучаемых в теории групп). Введем определение изоморфизма групп.

Группы $G_1$ и $G_2$ называются *изоморфными,* если между их элементами можно установить взаимно-однозначное соответствие, сохраняющее групповую операцию, т. е. такое, что если $x_1, y_1 \in G_1$, $x_2, y_2 \in G_2$ и $x_1 \leftrightarrow x_2$, $y_1 \leftrightarrow y_2$, то $x_1 \circ y_1 \leftrightarrow x_2 \circ y_2$.

Если $G_1$ и $G_2$ — мультипликативные группы, то соответствие $\leftrightarrow$ можно рассматривать как такое взаимно-однозначное отображение $f$ группы $G_1$ на $G_2$, при котором $f(x\ y) = f(x) f(y)$ для всех $x, y \in G_1$.

Если $f$ — изоморфное отображение группы $G_1$ на $G_2$, то $f(e_1)=e_2$, где $e_1$, $e_2$ — единичные элементы групп $G_1$ и $G_2$ соответственно, и для любого $x_1 \in G_1$ $f(x_1^{-1})=[f(x_1)]^{-1}$. В самом деле, пусть $x_2$ — произвольный элемент группы $G_2$ и $x_1$ — такой элемент группы $G_1$, что $f(x_1)=x_2$, тогда $x_2=f(x_1)=f(x_1 e_1)=f(x_1)f(e_1)=x_2 f(e_1)$, $x_2=f(x_1)=f(e_1 x_1)=$ $=f(e_1)f(x_1)=f(e_1)x_2$, т. е. $f(e_1)=e_2$ — единица группы $G_2$. Так как $f(x_1)f(x_1^{-1})=f(x_1 x_1^{-1})=f(e_1)=e_2$ и $f(x_1^{-1})f(x_1)=$ $=f(x_1^{-1}x_1)=f(e_1)=e_2$, то $[f(x_1)]^{-1}=f(x_1^{-1})$.

Отметим, что циклическая группа порядка $n$ — это группа, изоморфная группе вращений правильного $n$-угольника; бесконечная циклическая группа — это группа, изоморфная аддитивной группе целых чисел.

Замечание 1. Операции в изоморфных группах могут называться по-разному. Например, мультипликативная группа положительных действительных чисел изоморфна аддитивной группе действительных чисел. Изоморфное соответствие между ними устаиавливается отображением $f(a)=\log_c a$, где $c \neq 1$ — произвольное положительное число.

Замечание 2. Множество групп, изоморфных данной, называется абстрактной группой.

## § 9.7. Разложение группы по подгруппе

Пусть в группе $G$ фиксированы подмножества $P$, $Q$. Произведением $PQ$ этих подмножеств называется множество элементов группы $G$, каждый из которых хотя бы одним способом представлен в виде произведения элемента $p \in P$ на элемент $q \in Q$. Одно из множеств $P$, $Q$, разумеется, может состоять лишь из одного элемента $a$. В этом случае получим произведение $aP$ элемента $a$ на множество $P$ или произведение $Qa$ множества $Q$ на элемент $a$.

Из ассоциативности групповой операции следует ассоциативность умножения подмножеств группы, т. е. $(PQ)S=P(QS)$.

Пусть дана группа $G$ и некоторая ее подгруппа $H$. Фиксировав любой элемент $x \in G$, рассмотрим множество элементов $x \circ h$, где $h$ — любой элемент $H$. Это множество $x \circ H$ называется *левым смежным классом* группы $G$ по подгруппе $H$, порожденным элементом $x$. Очевидно, элемент $x$ принадлежит смежному классу $x \circ H$, поскольку подгруппа $H$ содержит нейтральный элемент $e$ и $x \circ e = x$.

Покажем, что всякий левый смежный класс порождается любым из своих элементов, т. е. если $y$ содержится в смежном классе $x \circ H$, то

$$y \circ H = x \circ H. \tag{9.5}$$

В самом деле, так как $y \in x \circ H$, то его можно представить в виде $y = x \circ h$, где $h$ — элемент подгруппы $H$. Следовательно, для любых элементов $h'$, $h''$ из $H$ будет $y \circ h' = x \circ (h \circ h')$, $x \circ h'' = y \circ (h^{-1} \circ h'')$, откуда и следует (9.5).

Таким образом, два любых смежных класса группы $G$ по подгруппе $H$ или совпадают, или не имеют ни одного общего элемента (не пересекаются). Действительно, если смежные классы $x \circ H$ и $y \circ H$ содержат общий элемент $z$, то $x \circ H = z \circ H = y \circ H$.

Итак, вся группа $G$ распадается на непересекающиеся левые смежные классы по подгруппе $H$. Такое разложение называется левосторонним разложением группы $G$ по подгруппе $H$. Очевидно, одним из левых смежных классов этого разложения будет сама подгруппа $H$, этот смежный класс порождается элементом $e$ (или любым элементом $h \in H$, поскольку $h \circ H = H$).

Аналогично вводится понятие правого смежного класса группы $G$ по подгруппе $H$, порождаемого элементом $x$; это множество $H \circ x$, т. е. множество всех элементов вида $h \circ x$, где $x$ — фиксированный элемент $G$; $h$ — любой элемент из $H$. Аналогичным образом получается правостороннее разложение группы $G$ по подгруппе $H$. Если группа $G$ является абелевой, то оба ее разложения по любой подгруппе (левостороннее и правостороннее) будут совпадать. В этом случае говорят просто о разложении группы по подгруппе. Приведем пример такого разложения.

Пусть $G$ — аддитивная группа целых чисел и $H$ — ее подгруппа, состоящая из всех чисел, кратных числу $k$. Разобьем группу $G$ на классы, относя к одному классу все те числа, которые при делении на $k$ дают одинаковые остатки. Очевидно, два числа $x$ и $y$ попадают в один и тот же класс тогда и только тогда, когда их разность делится на $k$, следовательно, принадлежит подгруппе $H$: $x - y = km$ или $x = y + h$, где $h \in H$. Итак, получено $k$ классов, считая одним из классов и подгруппу $H$. Разложение данной группы $G$ по указанной подгруппе $H$ состоит из $k$ различных смежных классов, порождаемых соответственно числами $0, 1, 2, \ldots, k-1$. В классе, порождаемом числом $l$, где $0 \leqslant l \leqslant k-1$, собраны все те числа, которые при делении на число $k$ дают остаток $l$.

Полученное разложение группы целых чисел по подгруппе чисел, кратных $k$, при $k=3$ можно представить следующим образом:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $H$ | ... | $-9$ | $-6$ | $-3$ | 0 | 3 | 6 | 9 | ... |
| $1+H$ | ... | $-8$ | $-5$ | $-2$ | 1 | 4 | 7 | 10 | ... |
| $2+H$ | ... | $-7$ | $-4$ | $-1$ | 2 | 5 | 8 | 11 | ... |

Замечание. В некоммутативной группе левостороннее и правостороннее разложения могут оказаться различными. Обратимся к симметрической группе $S_3$ (см. § 9.4). Из таблицы Кэли для этой группы видно, что множество элементов $P_1$, $P_2$ образует подгруппу, обозначим ее $B=\{P_1, P_2\}$. Левостороннее разложение группы $S_3$ по подгруппе $B$ состоит из классов $B$, $P_5B=P_4B=\{P_4, P_5\}$, $P_3B=P_6B=\{P_3, P_6\}$, а правостороннее — из классов $B$, $BP_6=BP_4=\{P_4, P_6\}$, $BP_5=BP_3=\{P_3, P_5\}$, т. е. эти разложения различны. Отметим, что левостороннее и правостороннее разложения этой группы по ее подгруппе третьего порядка $A=\{P_1, P_5, P_6\}$ совпадают; каждое из них состоит из двух классов: $A=\{P_1, P_5, P_6\}$, $AP_2=P_2A=\{P_2, P_3, P_4\}$.

В случае конечных групп существование разложения группы по ее подгруппе приводит к важной теореме, устанавливающей связь между порядками группы и ее подгруппы.

**Теорема Лагранжа.** Порядок подгруппы конечной группы является делителем порядка группы.

Доказательство. Пусть $G$ — конечная группа порядка $n$, $H$ — ее подгруппа порядка $k$. Разложим группу $G$ на левые смежные классы по подгруппе $H$. Обозначим буквой $m$ число полученных классов. Каждый левый класс $x \circ H$ состоит ровно из $k$ элементов. Действительно, если $x \circ h_1 = x \circ h_2$, где $h_1, h_2 \in H$, то $h_1=h_2$. Следовательно, общее число элементов группы $n=km$. Это означает, что $n$ делится на $k$. Число $m$ (делитель $n$) называется *индексом* подгруппы $H$ в группе $G$.

Следствие 1. Порядок любого элемента конечной группы будет делителем порядка группы.

Действительно, каждый элемент $g \in G$ порождает в ней циклическую группу $\{g\}$, состоящую из всех степеней этого элемента. Порядок подгруппы $\{g\}$ совпадает с порядком элемента $g$ в группе $G$.

Следствие 2. Всякая конечная группа, порядок которой есть простое число, является циклической.

В самом деле, эта группа должна совпадать с циклической подгруппой, порожденной любым ее элементом, отличным от нейтрального.

## § 9.8. Нормальный делитель

Подгруппа $H$ группы $G$ называется *нормальным делителем* этой группы (или *инвариантной подгруппой*), если левосторонние и правосторонние разложения данной группы по указанной подгруппе совпадают.

Из определения следует, что любая подгруппа коммутативной группы является в ней нормальным делителем. Далее, в любой группе $G$ сама группа и ее единичная подгруппа будут нормальными делителями: разложение группы $G$ по самой этой группе состоит из одного элемента $G$; разложения группы $G$ по единичной подгруппе совпадают с разложением группы на отдельные элементы. Приведем примеры нормальных делителей в некоммутативных группах.

1. В симметрической группе $S_3$ (см. § 9.4) подгруппа $H=\{P_1, P_5, P_6\}$ будет нормальным делителем, так как левостороннее и правостороннее разложения совпадают, они состоят из двух классов: $H=\{P_1, P_5, P_6\}$, $HP_2=$ $=P_2H=\{P_2, P_3, P_4\}$.

2. В мультипликативной группе невырожденных квадратных матриц порядка $n$ подгруппа матриц с определителем, равным единице, будет нормальным делителем. Действительно, левый и правый смежные классы по этой подгруппе, порождаемые матрицей $M$, совпадают с классом всех матриц, определитель которых равен определителю матрицы $M$ (как известно, определитель произведения двух матриц равен произведению определителей этих матриц).

Можно дать и другие определения нормального делителя, равносильные исходному. Приведем одно из них.

Подгруппа $H$ группы $G$ называется нормальным делителем этой группы, если для любого $x\in G$

$$x\circ H=H\circ x, \tag{9.6}$$

т. е. для любого $x\in G$ и элемента $h\in H$ существуют $h_1$, $h_2\in H$ такие, что

$$x\circ h=h_1\circ x,\ h\circ x=x\circ h_2. \tag{9.7}$$

## § 9.9. Классы сопряженных элементов

Чтобы дать еще одно определение нормального делителя, введем понятие сопряженных элементов группы.

Элементы $a$ и $b$ группы $G$ называют *сопряженными,* если в $G$ существует хотя бы один такой элемент $x$, что

$$b = x^{-1}ax. \tag{9.8}$$

В этом случае говорят также, что элемент $b$ получается из элемента $a$ *трансформированием* с помощью элемента $x$. Из равенства (9.8) находим $a = xbx^{-1} = (x^{-1})^{-1}bx^{-1}$, $a = (x^{-1})^{-1}bx^{-1}$, т. е. элемент $a$ при этом получается из элемента $b$ трансформированием элементом $x^{-1}$.

**Теорема 9.1.** Подгруппа $H$ группы $G$ тогда и только тогда будет нормальным делителем в $G$, если вместе с любым своим элементом $h$ она содержит все элементы, сопряженные с ним в $G$.

Доказательство. Если $H$ — нормальный делитель в $G$, то для фиксированного элемента $h \in H$ и любого элемента $x \in G$ существует (в силу 9.7) такой элемент $h_2$, что $hx = xh_2$. Из этого равенства получаем $x^{-1}hx = h_2$, т. е. всякий элемент, сопряженный с $h$, содержится в $H$.

Обратно, если подгруппа $H$ вместе с любым своим элементом содержит и все сопряженные ему элементы, то $H$ принадлежит, в частности, и элемент $x^{-1}hx = h_2$. Из этого равенства следует второе из равенств (9.7). Очевидно, $H$ содержит и элемент $(x^{-1})^{-1}hx^{-1} = xhx^{-1} = h_1$, откуда следует первое из равенств (9.7), т. е. $H$ — нормальный делитель.

Замечание. Как отмечалось, нормальный делитель называют также *инвариантной подгруппой.* Из теоремы 9.1 следует происхождение этого названия. Если $H$ — нормальный делитель группы $G$, то трансформирование любого элемента подгруппы $H$ с помощью элемента группы $G$ дает снова элемент подгруппы $H$ (подгруппа $H$ остается неизменяемой по отношению к операции трансформирования элементов $H$).

**Теорема 9.2.** Пересечение двух нормальных делителей группы $G$ является нормальным делителем этой группы.

Доказательство. Если $H_1$ и $H_2$ — нормальные делители группы $G$, то их пересечение $H_1 \cap H_2$ — подгруппа этой группы (как пересечение двух подгрупп). Пусть $h$ — любой элемент, принадлежащий пересечению $H_1 \cap H_2$, $x$ — любой элемент группы $G$, тогда элемент $x^{-1}hx$ должен содержаться и в $H_1$ и в $H_2$, поскольку оба нормальных делителя содержат элемент $h$. Следовательно, элемент $x^{-1}hx$ входит в пересечение $H_1 \cap H_2$, а это означает, что $H_1 \cap H_2$ — нормальный делитель группы $G$.

## § 9.10. Фактор-группа

Если $H$ — произвольная подгруппа группы $G$, то

$$HH=H, \tag{9.10}$$

поскольку произведение любых двух элементов из $H$ принадлежит $H$; умножив все элементы из $H$ на нейтральный элемент, получим всю подгруппу $H$.

Пусть дана группа $G$ и ее нормальный делитель $H$. Множество всех смежных классов (в этом случае левые и правые смежные классы совпадают) обозначим буквой $S$ и введем в нем операцию, которую будем называть «умножением», положив

$$xH\cdot yH=xyH. \tag{9.11}$$

Равенство следует из ассоциативности умножения подгрупп данной группы и того, что $yH=Hy$; следовательно, $xH\cdot yH=xyHH=xyH$.

Равенство (9.11) означает, что для того, чтобы найти произведение двух смежных классов группы $G$ по нормальному делителю $H$, следует произвольно выбрать в этих классах по одному представителю и взять тот смежный класс, которому принадлежит произведение этих представителей.

Покажем, что множество $S$ смежных классов образует группу.

1. Ассоциативность умножения классов следует из ассоциативности умножения в группе $G$: $(xH\cdot yH)\cdot zH=$ $=(xy)H\cdot zH=(xy)zH=x(yz)H=xH\cdot(yz)H=$ $=xH\cdot(yH\cdot zH)$.

2. Нейтральным элементом является сама подгруппа $H$, так как $H\cdot xH=eH\cdot xH=exH=xH$, $xH\cdot H=xH\cdot eH=$ $=xeH=xH$.

3. Обратным классу $xH$ будет класс $x^{-1}H$, где $x^{-1}$ — элемент, обратный элементу $x$, ибо $xH\cdot x^{-1}H=xx^{-1}H=$ $=eH=H$, $x^{-1}H\cdot xH=x^{-1}xH=eH=H$.

Полученная группа называется фактор-группой группы $G$ по нормальному делителю $H$ и обозначается $G/H$. Более подробно: *фактор-группой* группы $G$ по нормальному делителю $H$ называется группа всех смежных классов этой группы $G$ по подгруппе $H$.

Таким образом, с группой $G$ можно связать некоторый набор новых групп — ее фактор-групп по различным нормальным делителям.

Отметим, что фактор-группа абелевой группы является абелевой, поскольку из условия $xy=yx$ следует $xH\cdot yH=xyH=yxH=yH\cdot xH$. Фактор-группа циклической группы является циклической группой. В самом деле, если группа $G$ порождается элементом $g$ и если дан произвольный смежный класс $xH$, то существует такое целое число $k$, что $x=g^k$, поэтому $xH=(gH)^k$.

Приведем примеры фактор-групп. Пусть $G$ — аддитивная группа целых чисел, $H$ — подгруппа чисел, делящихся на 3. Найдем фактор-группу $G/H$. Групповой операцией в данном случае будет сложение. Число смежных классов равно трем (см. пример в § 9.7): множество чисел, делящихся на 3, два множества чисел, дающих при делении на 3 соответственно остатки 1 и 2; обозначим эти смежные классы [0], [1], [2]. В этом множестве введем операцию сложения следующим образом: сложив соответствующие числа в квадратных скобках, определим, какой остаток при делении на 3 дает их сумма, и будем считать суммой смежных классов тот, которому принадлежит полученный остаток. «Таблица умножения» для фактор-группы имеет вид $[0]+[0]=[1]+[2]=[2]+[1]=[0]$, $[0]+[1]=[1]+[0]=[2]+[2]=[1]$, $[0]+[2]=[2]+[0]=[1]+[1]=[2]$. Отсюда видно, что фактор-группа является абелевой. Кроме того, все смежные классы порождаются классом [1], они совпадают со «степенями» этого класса: [1], $[1]+[1]=[2]$, $[1]+[1]+[1]=[0]$. Поскольку фактор-группа порождена одним элементом, то она является циклической.

2. Пусть $G$ — аддитивная группа целых чисел, $H$ — подгруппа целых чисел, кратных натуральному числу $k$. Фактор-группой $G/H$ будет конечная группа порядка $k$, состоящая из классов [0], [1], [2], ..., $[k-1]$. Эта фактор-группа является циклической, как и сама группа $G$.

3. Пусть $G$ — мультипликативная группа всех невырожденных матриц порядка $n$, $H$ — подгруппа матриц с определителем, равным единице. В примере 2 § 9.8 было показано, что $H$ — нормальный делитель группы $G$. Покажем, что две матрицы $B$ и $C$ принадлежат одному и тому же смежному классу группы $G$ по нормальному делителю $H$ тогда и только тогда, когда они имеют равные определители. Действительно, пусть $B$ и $C$ принадлежат одному классу, тогда $B=CA$, где $A\in H$, т. е. $\det A=1$, поэтому $\det B=\det(CA)=\det C\cdot\det A=\det C$, или $\det B=\det C$. Обратно, пусть $\det B=\det C$. Так как $B=C(C^{-1}B)$ и $\det B=\det C\cdot\det(C^{-1}B)$, то $\det C^{-1}B=1$, а это означает,

что $C^{-1}B \in H$, т. е. $C^{-1}B = A$, где $A \in H$, откуда $CC^{-1}B = CA$, $B = CA$. Следовательно, каждый смежный класс $G$ по $H$ вполне характеризуется определителем входящих в него матриц. Умножению классов соответствует умножение произвольных представителей этих классов; произведение классов $B$ (матриц с определителем $b$) и $C$ (матриц с определителем $c$) является класс $BC$ (матриц с определителем $bc$). Таким образом, фактор-группа $G/H$ изоморфна мультипликативной группе отличных от нуля действительных чисел.

## § 9.11. Гомоморфизм групп

Введем понятие гомоморфизма групп, частным случаем которого является изоморфизм (см. § 9.6).

Если $G$ и $G'$ — группы и $f\colon G \to G'$ — такое отображение, при котором для любых элементов $x$, $y$ группы $G$

$$f(xy) = f(x) f(y), \tag{9.12}$$

то $f$ называется *гомоморфным отображением*, или *гомоморфизмом* группы $G$ в группу $G'$. Отметим, что в определении гомоморфизма имеется в виду, что каждый элемент группы $G$ поставлен в соответствие хотя бы одному элементу группы $G'$, но разным элементам из $G$ может соответствовать один и тот же элемент из $G'$. Другими словами, отображение группы $G$ в группу $G'$ не предполагается взаимно-однозначным, как в случае изоморфизма.

Очевидно, каждая группа гомоморфна себе самой, так как можно положить $f(x) = x$ для всех $x \in G$. Далее, каждая группа гомоморфна единичной группе (состоящей из одного нейтрального элемента $e$). Чтобы убедиться в этом, достаточно положить $f(x) = e$ для всех $x \in G$. Приведем другой пример гомоморфного отображения групп. Циклическая группа $C_6$ шестого порядка с элементами $e$, $a$, $a^2$, $a^3$, $a^4$, $a^5$ гомоморфна циклической группе $C_2$ второго порядка с элементами $E$, $A$: $f(e) = f(a^2) = f(a^4) = E$, $f(a) = f(a^3) = f(a^5) = A$.

Равенство (9.12) означает, что образ произведения двух элементов равен произведению образов этих элементов, в связи с чем говорят, что гомоморфизм «сохраняет групповую операцию» (которые, впрочем, могут называться по-разному в группах $G$ и $G'$). Гомоморфизм групп сохраняет не только групповую операцию, но также нейтральный и обратный элементы: если $f$ — гомоморфное отображение группы $G$ в группу $G'$, то $f(e)=e'$, где $e$, $e'$ — нейтральные элементы групп $G$ и $G'$ соответственно, $[f(x)]^{-1}=f(x^{-1})$. Доказательство этого утверждения аналогично доказательству соответствующего факта для изоморфного отображения (см. § 9.6).

Приведем без доказательства теорему, которая дает ответ на вопрос, каким вообще группам может быть гомоморфна данная группа.

**Теорема о гомоморфизмах.** Каждая группа гомоморфна любой своей фактор-группе. Обратно, если группа $G$ гомоморфна группе $G'$, то $G'$ изоморфна фактор-группе $G$ по некоторому нормальному делителю $H$.

## § 9.12. Представления групп

С точки зрения алгебры изоморфные группы не считаются различными. О группе, изоморфной некоторой подгруппе группы подстановок, говорят, что она представлена подстановками. Верна теорема Кэли, которая утверждает, что каждая конечная группа изоморфна некоторой группе подстановок (т. е. всякую конечную группу можно представить подстановками). Следовательно, при описании любой конечной группы можно воспользоваться преимуществами группы подстановок.

Для теории и приложений наиболее интересными являются линейные представления конечных групп. Говоря о линейном представлении конечной группы $G$, предполагают, что дано векторное пространство $V_n$, в котором действуют линейные невырожденные преобразования. Эти преобразования образуют группу $G'$, которой гомоморфна исходная группа $G$; при этом говорят, что группа $G'$ представляет группу $G$. Введем следующее определение.

Гомоморфное отображение $\Gamma$ группы $G$ в группу $G'$ невырожденных линейных преобразований $n$-мерного векторного пространства $V_n$ называется *линейным представлением* группы $G$.

Следовательно, если $\Gamma$ — линейное представление группы $G$ группой $G'$, то каждому элементу $a \in G$ поставлено в соответствие невырожденное линейное преобразование $\Gamma(a) \in G'$ пространства $V_n$ так, что для любых $a$, $b \in G$ справедливо соотношение $\Gamma(ab)=\Gamma(a)\Gamma(b)$. Как известно, при этом $\Gamma(e)=E$, где $e$, $E$ — нейтральные элементы групп $G$, $G'$ соответственно, и $\Gamma(a^{-1})=[\Gamma(a)]^{-1}$ для любого $a\in G$.

Пространство $V_n$, в котором действуют преобразования из группы $G'$, называется *пространством представления* группы $G$. Размерность пространства $V_n$ называют размерностью (или степенью) рассматриваемого представления.

Вместо линейных преобразований часто рассматривают соответствующие им матрицы. Если в пространстве $V_n$ фиксирован базис, то каждому преобразованию $\Gamma(a)$ будет соответствовать определенная матрица, которую обозначим тем же символом $\Gamma(a)$. Следовательно, каждому элементу $a\in G$ ставится в соответствие невырожденная квадратная матрица $\Gamma(a)$ порядка $n$ такая, что $\Gamma(ab)=\Gamma(a)\Gamma(b)$. При $n=1$, т. е. в случае одномерного линейного пространства, эта матрица будет матрицей первого порядка. Каждому элементу $a\in G$ ставится в соответствие отличное от нуля число $\Gamma(a)$, причем $\Gamma(ab)=$ $=\Gamma(a)\Gamma(b)$. Нейтральному элементу $e$ группы $G$ ставится в соответствие число 1.

Важным (хотя и тривиальным) примером является единичное представление группы $G$, т. е. такое одномерное представление, при котором каждому элементу $a\in G$ поставлено в соответствие число 1, т. е. $\Gamma(a)=1$ для любого $a\in G$.

Если группа $G$ изоморфна группе $G'$, то представление $\Gamma$ группы $G$ группой $G'$ называется *точным*, в противном случае — *неточным*.

Если $G$ — группа линейных преобразований линейного пространства, то она — одно из своих линейных представлений (причем точное); это представление называют *основным представлением* группы $G$.

В качестве примера найдем одномерные представления циклической группы $C_2$ второго порядка, состоящей из двух элементов $e$, $a$, причем $a^2=e$. Если $\Gamma$ — одномерное представление группы $C_2$, то $\Gamma(e)=1$. Предположим, что $\Gamma(a)=\alpha$, тогда $\Gamma(a^2)=[\Gamma(a)]^2=\alpha^2$. Поскольку $a^2=e$, то $\Gamma(a^2)=\Gamma(e)=1$. Следовательно, $\alpha^2=1$, $\alpha=\pm 1$; получено два одномерных представления: $\Gamma_1(e)=1$, $\Gamma_1(a)=$ $=1$ — единичное (неточное); $\Gamma_2(e)=1$, $\Gamma_2(a)=-1$ (точное).

РАЗДЕЛ II

# МАТЕМАТИЧЕСКИЙ АНАЛИЗ

**Математический анализ — раздел математики, в котором изучаются функции. Основу математического анализа составляют дифференциальное и интегральное исчисление, теория рядов. Заслуга открытия дифференциального и интегрального исчисления (или анализа бесконечно малых) принадлежит Ньютону[1] и Лейбницу[2]. Развитие анализа бесконечно малых оказало огромное влияние на прогресс науки и техники**

## Глава 10

## ФУНКЦИИ И ПРЕДЕЛЫ

Понятие функции — одно из основных понятий современной математики. Это понятие прошло сложный путь развития, на каждом этапе которого определялось по-разному. Важным понятием является понятие предела функции, на нем основаны многие другие фундаментальные понятия.

### § 10.1. Числовые множества. Отрезок, интервал, промежуток

Совокупность всех рациональных и иррациональных чисел называется множеством всех действительных (или вещественных) чисел (рациональные числа включают в себя и целые, как частный случай). Действительные числа упорядочены по величине, т. е. для любых двух действительных чисел $a$ и $b$ справедливо одно и только одно из соотношений: $a<b$, $a=b$, $a>b$. Действительные числа можно изображать точками на координатной оси. Между

---

[1] *Исаак Ньютон* (*Isaak Newton,* 1643—1727) — английский математик и физик.

[2] *Готфрид Вильгельм Лейбниц* (*Gottfried Wilhelm Leibniz,* 1646—1716) — немецкий математик и философ.

всеми действительными числами и всеми точками координатной оси устанавливается взаимно-однозначное соответствие. Это дает возможность в некотором смысле равнозначно употреблять понятия «число $x$» и «точка $x$», а также называть числовой прямой множество всех действительных чисел и говорить, что число $a$ лежит левее числа $b$, когда $a<b$.

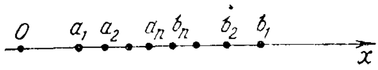



Рис. 10.1

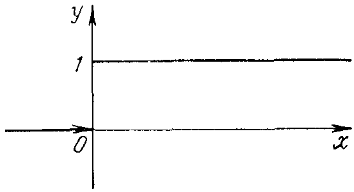



Рис. 10.2

Если заданы два действительных числа $a$ и $b$ $(a<b)$, то множество чисел, удовлетворяющих соотношениям $a\leqslant \leqslant x\leqslant b$, называется *замкнутым промежутком* $[a, b]$ (*отрезком* $[a, b]$, или *сегментом* $[a, b]$), т. е. $[a, b]=\{x\colon a\leqslant \leqslant x\leqslant b\}$.

Система отрезков $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$ называется *системой вложенных отрезков* (рис. 10.1), если $a_1\leqslant \leqslant a_2\leqslant \ldots \leqslant a_n\leqslant \ldots \leqslant b_n\leqslant \ldots \leqslant b_2\leqslant b_1$.

**Теорема 10.1.** Всякая система вложенных отрезков, длина которых стремится к нулю, имеет единственное число, принадлежащее всем этим отрезкам.

Доказательство. Очевидно, отрезки указанной системы имеют общие точки. По условию длина $n$-го отрезка стремится к нулю, т. е. для любого числа $\varepsilon>0$ существует такое натуральное число $N$, что $b_n-a_n<\varepsilon$ для всех $n>N$.

Предположим, что имеется два числа $x$ и $y$, принадлежащих всем отрезкам, причем $x\neq y$. Для определенности будем считать $x<y$, тогда $a_n\leqslant x\leqslant b_n$ и $a_n\leqslant y\leqslant b_n$ $(n=1, 2, 3, \ldots)$. Из неравенств $y\leqslant b_n$ и $x\geqslant a_n$ получаем $y-x\leqslant b_n-a_n$. Задав $\varepsilon>0$, можно указать такое $N$, что $b_n-a_n<\varepsilon$ для всех $n>N$; следовательно, $y-x<\varepsilon$. Поскольку $\varepsilon$ — любое число, то, взяв $\varepsilon=y-x$ ($\varepsilon>0$, так как $y>x$, $y-x>0$), получим $y-x<y-x$, что противоречиво. Наличие противоречия доказывает теорему.

Если $a<b$, то множество действительных чисел

$$]a, b[=\{x\colon a<x<b\}$$

называется *открытым промежутком,* или *интервалом*[1].

*Полуоткрытыми промежутками* (или *полуинтервалами*) называются следующие множества действительных чисел:

$$[a, b[=\{x: a\leqslant x<b\}, ]a, b]=\{x: a<x\leqslant b\}.$$

Множества $[a, b]$, $]a, b]$, $[a, b[$, $]a, b[$ называются *промежутками,* точки $a$ и $b$ — концами, а все остальные их точки — внутренними точками.

Отметим, что отрезок $[a, b]$ включает в себя концы, а интервал $]a, b[$ не содержит их. Докажем, что интервал не содержит также наименьшего и наибольшего числа. Предположим противное, пусть $b_0$ — наибольшее число из интервала $]a, b[$. Поскольку $b_0\neq b$ ($b$ не принадлежит интервалу) и $b_0<b$, то $b-b_0>0$. Обозначим $\varepsilon=\frac{b-b_0}{2}$, тогда число $b_0+\varepsilon$ также принадлежит интервалу $]a, b[$ (ибо $b_0+\varepsilon<b$) и $b_0+\varepsilon>b_0$ (так как $\varepsilon>0$), что противоречит предположению. Аналогично доказывается, что интервал не содержит наименьшего числа.

Всякий интервал $[x-\varepsilon, x+\varepsilon[$, где $\varepsilon>0$, называется *$\varepsilon$-окрестностью* (или просто *окрестностью*) точки (числа) $x$ на числовой прямой; обозначим его через $O(x, \varepsilon)$.

Множество всех действительных чисел называют бесконечным промежутком и обозначают $]-\infty, +\infty[$; считают по определению, что $-\infty<x<+\infty$ для любого действительного числа $x$. К бесконечным промежуткам относятся также следующие множества действительных чисел:

$$]a, +\infty[=\{x: x>a\}, [a, +\infty[=\{x: x\geqslant a\},$$
$$]-\infty, b[=\{x: x<b\}, ]-\infty, b]=\{x: x\leqslant b\}.$$

Множество $X$ действительных чисел $x$ называется *ограниченным сверху,* если существует такое число $b$, что $x\leqslant b$ для всех $x\in X$. В этом случае говорят, что число $b$ ограничивает сверху множество $X$.

Множество $X$ действительных чисел $x$ называется *ограниченным снизу,* если существует такое число $a$, что $x\geqslant a$ для всех $x\in X$. В этом случае говорят, что число $a$ ограничивает снизу множество $X$.

Множество действительных чисел, ограниченное свер-

---

[1] Интервал обозначается иногда и простыми скобками:

$$(a, b)=\{x: a<x<b\}.$$

ху и снизу, называется *ограниченным;* другими словами, множество $X$ действительных чисел $x$ называется ограниченным, если существуют числа $a$ и $b$ такие, что $a \leqslant x \leqslant b$ для всех $x \in X$.

Примерами ограниченных числовых множеств являются отрезок [3, 4], интервал [1, 2], полуинтервалы [0, 1[, ]2, 3]; множество [$-3$, $+\infty$[ ограничено снизу, но не ограничено сверху, множество ]$-\infty$, 5[ ограничено сверху, но не ограничено снизу; множество всех целых чисел не ограничено ни сверху, ни снизу.

Для ограниченного множества существуют различные числа, ограничивающие это множество сверху и снизу, например, любое число $x \geqslant 5$ ограничивает сверху множество $X=]0, 5[$, любое число $x \leqslant 0$ ограничивает это множество снизу.

Среди чисел, ограничивающих сверху данное множество, выделяют наименьшее; среди чисел, ограничивающих множество снизу,— наибольшее; для этого вводят понятия верхней и нижней грани числового множества.

Число $M$ называется *верхней гранью* числового множества $X$, если: 1) $x \leqslant M$ для всех $x \in X$; 2) для любого числа $\varepsilon > 0$ существует такое $x_\varepsilon \in X$, что $x_\varepsilon > M - \varepsilon$.

Верхняя грань множества $X$ обозначается через $\sup X$, или $\sup\limits_{x \in X} x$. Например, для сегмента $X=[1, 2]$ $\sup X=2$; для интервала $X=]2, 3[$ $\sup X=3$.

Число $m$ называется *нижней гранью* числового множества $X$, если: 1) $x \geqslant m$ для всех $x \in X$; 2) для любого числа $\varepsilon > 0$ существует такое $x_\varepsilon \in X$, что $x_\varepsilon < m + \varepsilon$.

Нижняя грань множества $X$ обозначается через $\inf X$, или $\inf\limits_{x \in X} x$. Например, если $X = ]4, 5[$, то $\inf X = 4$; если $X = [6, 7]$, то $\inf X = 6$.

Приведенные примеры показывают, что верхняя и нижняя грани множества могут принадлежать или не принадлежать этому множеству (если $X=[6, 7[$, то $\sup X = 7$, число 7 не принадлежит $X$; $\inf X=6$, это число принадлежит множеству $X$).

Если некоторое множество $X$ действительных чисел не ограничено сверху, полагают по определению $\sup X = +\infty$; если множество не ограничено снизу, полагают $\inf X = -\infty$.

Если верхняя (нижняя) грань множества является числом, то говорят, что она конечна; если соответствующая грань — символ $+\infty$ (или $-\infty$), то говорят, что она бесконечна.

Вопрос о существовании верхней (нижней) грани у ограниченного множества решает следующая теорема, приводимая здесь без доказательства.

**Теорема 10.2.** Всякое непустое множество действительных чисел, ограниченное сверху, имеет конечную верхнюю грань, а ограниченное снизу — конечную нижнюю грань.

Естественно возникает вопрос, будет ли единственной верхняя (нижняя) грань числового множества? Ответ на этот вопрос является положительным.

**Теорема 10.3.** У всякого множества действительных чисел верхняя (нижняя) грань единственна.

Предположим противное: существует множество $X$, имеющее две различные верхние грани $M_1$ и $M_2$. Пусть $M_1 < M_2$, тогда $M_2 - M_1 > 0$; обозначим эту разность через $\varepsilon$, т. е. $M_2 - M_1 = \varepsilon$, $\varepsilon > 0$. Поскольку $M_2$ — верхняя грань множества $X$, то: 1) $x \leqslant M_2$ для всех $x \in X$; 2) для $\varepsilon > 0$, где $\varepsilon = M_2 - M_1$, существует такое $x_\varepsilon$, что $x_\varepsilon > M_2 - \varepsilon$. Но $M_2 - \varepsilon = M_1$, следовательно, $x_\varepsilon > M_1$, а это противоречит тому, что $M_1$ — верхняя грань. Единственность верхней грани доказана. Аналогично доказывается единственность нижней грани.

## § 10.2. Понятие функции

Рассмотрим множество $X$ элементов $x$ и множество $Y$ элементов $y$. Если каждому элементу $x \in X$ поставлен в соответствие единственный элемент $y \in Y$, обозначаемый $y = f(x)$, то говорят, что на множестве $X$ задана *функция* $y = f(x)$ со значениями в множестве $Y$. Элементы $x \in X$ называются *значениями аргумента*, а элементы $y \in Y$ — *значениями функции*. Множество $X$ называется *областью определения функции*, множество всех значений функции — *областью значений* этой функции.

Замечание. Функцию, заданную на множестве $X$ со значениями в множестве $Y$, называют также *отображением множества $X$ в множество* $Y$. Если множество $Y$ является множеством значений функции, рассматриваемую функцию называют *отображением множества $X$ на множество* $Y$.

Функцию, заданную на множестве $X$, называют также *оператором*, заданным на множестве $X$, и обозначают символом $f$.

В случае, когда $X$ и $Y$ — числовые множества, соответствующие функции называют *числовыми функциями*. Если рассматриваются действительные числа, функции называют действительными (вещественными) функциями одной действительной (вещественной) переменной.

Употребляются следующие обозначения функции: $y=f(x)$, $y=F(x)$, $y=\Phi(x)$, $y=\varphi(x)$, $y=y(x)$ и т. п. Значение, которое функция $y=f(x)$ принимает при $x=a$, обозначается через $f(a)$.

Функция и аргумент могут обозначаться также другими буквами, например, $s=f(t)$, $u=f(v)$, $r=r(t)$, $x=x(t)$ и т. д.

К простейшим областям определения функции относится отрезок, интервал, полуинтервалы (см. § 10.1) или совокупность указанных интервалов. Например, для функции $y=-\sqrt{9-x^2}$ областью определения является отрезок $[-3, 3]$, а областью ее значений — отрезок $[-3, 0]$, для функции $y=x^3$ область определения и область значений совпадают с интервалом $]-\infty, +\infty[$.

К традиционным основным способам задания функции относятся: аналитический (с помощью одной или нескольких формул), графический (с помощью графика), табличный (с помощью таблицы значений).

В последние годы получил широкое распространение четвертый способ задания функции — с помощью указания программы для вычисления ее значений на цифровой вычислительной машине.

Функция, заданная формулой

$$y=f(x), \tag{10.1}$$

правая часть которой не содержит $y$, называется *явной функцией.*

Рассмотрим уравнение

$$F(x, y)=0. \tag{10.2}$$

Предположим, что существует непустое множество $X$ значений $x$ таких, что при каждом $x_0 \in X$ уравнение $F(x_0, y)=0$ относительно $y$ имеет действительные решения. Обозначим одно из таких решений через $y_0$, будем считать его соответствующим числу $x_0 \in X$. В результате получим функцию $y=y(x)$, определенную на множестве $X$ и такую, что $F(x, y(x)) \equiv 0$ для всех $x \in X$. Функция $y=y(x)$, удовлетворяющая уравнению вида (10.2), называется функцией, заданной неявно, или *неявной функци-*

*ей.* Это название отражает только способ задания функции, а не характер функциональной зависимости.

Отметим, что одно уравнение вида (10.2) может определять несколько функций. Например, уравнение $x^2+y^2-R^2=0$ определяет две функции: $y_1=f_1(x)=\sqrt{R^2-x^2}$, $y_2=f_2(x)=-\sqrt{R^2-x^2}$.

Обратимся к функции (10.1). Каждому $x\in X$ по определенному закону ставится в соответствие единственное значение $y\in Y$. С другой стороны, каждому $y\in Y$ будет соответствовать одно или несколько значений $x\in X$.

В случае, когда каждому $y\in Y$ по некоторому закону соответствует только одно значение $x\in X$, получаем функцию

$$x=\varphi(y), \tag{10.3}$$

заданную па множестве $Y$ со значениями в множестве $X$. Функцию (10.3) называют *обратной* функцией по отношению к функции (10.1). Функции (10.1) и (10.3) называются *взаимно-обратными.* Для взаимно-обратных функций, очевидно, выполняются тождества $\varphi(f(x))\equiv x$, $x\in X$; $f(\varphi(y))\equiv y$, $y\in Y$. Примерами взаимно-обратных функций являются: $y=3^x$, $x=\log_3 y$; $y=2x-3$, $x=\frac{y+3}{2}$.

Если придерживаться стандартных обозначений ($y$ — функция, $x$ — аргумент), обратную функцию (10.3) следует писать в виде $y=\varphi(x)$. Например, можно говорить, что функции $y=3^x$, $y=\log_3 x$ являются взаимно-обратными.

Функцию, обратную к функции $y=f(x)$, удобно обозначать символом

$$x=f^{-1}(y). \tag{10.4}$$

В этих обозначениях для функций $y=f(x)$ и $x=f^{-1}(y)$ получаем тождества $f^{-1}f(x)\equiv x$, $x\in X$; $ff^{-1}(y)\equiv y$, $y\in Y$, где использованы обозначения $f^{-1}f(x)\equiv f^{-1}(f(x))$, $ff^{-1}(y)\equiv\equiv f(f^{-1}(y))$.

Если $y=f(u)$, $u=\varphi(x)$ — функции своих аргументов, причем область определения функции $f$ содержит область значений $\varphi$, то каждому $x$ из области определения функции $\varphi$ соответствует $y$ такое, что $y=f(u)$, где $u=\varphi(x)$. Эта функция, определяемая соответствием

$$y=f(\varphi(x)), \tag{10.5}$$

называется *функцией от функции,* или *сложной функцией.*

(Применяются также и другие названия: *композиция функций* $\varphi$ и $f$, *суперпозиция функций* $\varphi$ и $f$.) Например, если $y=u^2$, $u=\sin x$, то $y=(\sin x)^2=\sin^2 x$.

Рассматривают также функции, являющиеся композициями более чем двух функций. Например, функция $w=\cos \lg(1+x^2)$ представляет собой композицию следующих функций: $w=\cos u$, $u=\lg v$, $v=1+z$, $z=x^2$.

Функция $y=f(x)$ называется *четной*, если для любых $x$ и $-x$ из области ее определения выполняется равенство $f(-x)=f(x)$. Функция $y=\varphi(x)$ называется *нечетной*, если для любых $x$ и $-x$ из области ее определения выполняется равенство $\varphi(-x)=-\varphi(x)$. Например, $y=x^2$, $y=\cos x$ — четные функции, $y=x^3$, $y=\sin x$ — нечетные функции.

Функция $y=f(x)$ называется *периодической*, если существует число $T\neq 0$ такое, что при всех $x$ и $x+T$ из области ее определения выполняется равенство $f(x+T)=f(x)$. Число $T$ в этом случае называется *периодом функции*. Всякая периодическая функция имеет бесконечное множество периодов. Говоря о периоде функции $y=f(x)$ ($f(x)\neq\text{const}$), обычно имеют в виду наименьший положительный период, так, периодом функции $y=\cos x$ является число $2\pi$, периодом функции $y=\operatorname{tg} x$ — число $\pi$.

Кроме тригонометрических и обратных тригонометрических функций в средней школе изучаются функции: степенная $y=x^a$ ($a=\text{const}$), показательная $y=a^x$ ($a=\text{const}$), логарифмическая $y=\log_a x$ ($a=\text{const}$). Все эти функции называются *основными элементарными функциями*.

*Элементарными функциями* называются функции, которые можно получить из основных элементарных функций с помощью алгебраических действий и образования сложных функций. Например, функции $y=\lg\sin x$, $y=x^2+\cos x$, $y=3^{\lg\cos x+\sin x}$ и т. д. являются элементарными.

## § 10.3. График функции

Рассмотрим плоскость, на которой выбрана декартова прямоугольная система координат.

*Графиком функции* $y=f(x)$ называется множество точек плоскости, координаты которых удовлетворяют данной функциональной зависимости, т. е. точек $M(x, f(x))$. Например, графиком функции $y=-\sqrt{16-x^2}$ будет полуокружность радиуса $R=4$ с центром в начале координат, расположенная ниже оси $Ox$. Действительно, уравнение $x^2+$

$+ y^2 = 16$ является уравнением окружности радиуса $R = 4$ с центром в начале координат. Разрешив его относительно $y$, получим $y = \sqrt{16 - x^2}$, $y = -\sqrt{16 - x^2}$; первое из этих уравнений определяет полуокружность, расположенную выше оси $Ox$ (так как $y \geqslant 0$), второе — полуокружность, расположенную ниже оси $Ox$ (поскольку $y \leqslant 0$).

Из курса математики средней школы известны графики основных элементарных функций. В данной книге приводились графики линейной функции (§ 2.1), дробно-линейной функции (§ 2.9), функции $y = ax^2 + bx + c$ и др. На рис. 10.2 представлен график функции, заданной формулами $f(x) = 0$, если $x < 0$, $f(x) = 1$, если $x \geqslant 0$ (эта функция называется функцией Хевисайда).

Во всех указанных случаях график функции представляет собой некоторую линию на плоскости. График функции от дискретного аргумента не является линией, он состоит из дискретных точек.

Отметим, что график четной функции симметричен относительно оси $Oy$, график нечетной функции симметричен относительно начала координат. Графики взаимно-обратных функций симметричны относительно прямой, на которой лежит биссектриса первого координатного угла.

Вопрос о построении графиков функций детально будет рассмотрен в гл. 12.

## § 10.4. Понятие функции нескольких переменных

Предварительно введем понятие арифметического $n$-мерного пространства. Упорядоченную совокупность $n$ действительных чисел $x_1, x_2, \ldots, x_n$ называют *точкой*, а сами эти числа — ее *координатами*. Запись $M(x_1, x_2, \ldots, x_n)$ означает, что точка $M$ имеет координаты $x_1, x_2, \ldots, x_n$. Множество всевозможных указанных точек называется *арифметическим (координатным)* $n$-мерным пространством.

Обратимся к определению функции, данному в § 10.2. Функция, определенная на некотором множестве $X$ арифметического $n$-мерного пространства, называется функцией $n$ аргументов $y = f(x_1, x_2, \ldots, x_n)$, где $x_1, x_2, \ldots, x_n$ — координаты точки $M(x_1, x_2, \ldots, x_n)$ данного множества. В этом случае говорят, что задана функция точки $M$, и пишут $y = f(M)$.

Рассмотрим случаи, когда $n=2$ и $n=3$. Предположим, что $X$ — некоторое множество точек плоскости, $Y$ — подмножество множества всех действительных чисел. Так как в фиксированной декартовой прямоугольной системе координат $Ox_1x_2$ каждой точке $M$ соответствует упорядоченная пара действительных чисел $x_1$, $x_2$ — ее координаты, то функция, заданная на указанном множестве $X$, является функцией двух аргументов, т. е. $y=f(x_1, x_2)$, где $x_1$, $x_2$ — координаты точки $M(x_1, x_2)$. Если координаты точки $M$ обозначить буквами $x$ и $y$, а функцию — буквой $z$, то $z=f(x, y)$. Переменные $x$ и $y$ при этом называются *аргументами функции $z$,* или *независимыми переменными.* Значение функции $z=f(x, y)$, которое она принимает при $x=a$, $y=b$, обозначается через $f(a, b)$.

Область определения функции двух переменных представляет собой некоторое множество точек плоскости.

*Графиком функции* $z=f(x, y)$ называется множество точек $N(x, y, f(x, y))$, т. е. некоторое множество точек пространства. Например, график функции $z=x-y$ представляет собой плоскость в пространстве, проходящую через начало координат и пересекающую координатную плоскость $Oxy$ по прямой, образующей равные углы с осями $Ox$ и $Oy$; геометрическим изображением функции $z=x^2+y^2$ является поверхность параболоида вращения (см. рис. 6.20), а функции $z=\sqrt{9-x^2-y^2}$ — полусфера радиуса $R=3$ с центром в начале координат, расположенная выше плоскости $Oxy$. Отметим, что первые две функции определены на всей плоскости $Oxy$, третья — в круге радиуса $R=3$ с центром в начале координат, т. е. в области, заданной неравенством $x^2+y^2\leqslant 9$.

Функцию

$$z=f(x, y) \tag{10.6}$$

можно представить так: $z-f(x, y)=0$, или в более общем виде:

$$F(x, y, z)=0. \tag{10.7}$$

Функция, заданная формулой (10.6), называется *явной;* функция, определяемая уравнением (10.7),— *неявной.*

Действительная функция, определенная на некотором множестве $X$ точек пространства, т. е. точек $M(x, y, z)$, где $x$, $y$, $z$ — декартовы координаты, называется функцией трех переменных $x$, $y$, $z$. Функцию трех переменных $x$, $y$, $z$ обозначим буквой $u$, тогда $u=f(x, y, z)$. Значение функции $u=f(x, y, z)$ при $x=a$, $y=b$, $z=c$ обозначается через $f(a, b, c)$. Областью определения функции трех переменных будет некоторое множество точек пространства. Например, область определения функции $u=\sqrt{1-x^2-y^2-z^2}$ представляет собой шар радиуса $R=1$ с центром в начале координат, областью определения функции $u=1:\sqrt{1-x^2-y^2-z^2}$ является множество точек, лежащих внутри указанного шара (граничные точки, т. е. точки сферы $x^2+y^2+z^2=1$, исключаются).

## § 10.5. Предел функции

Рассмотрим функцию $y=f(x)$, определенную в некотором интервале, содержащем точку $x=a$.

Число $A$ называется пределом функции $y=f(x)$ при $x$, стремящемся к $a$ (или в точке $a$), если для любого числа $\varepsilon>0$ существует такое $\delta>0$, что при всех $x$, удовлетворяющих условию

$$0<|x-a|<\delta, \tag{10.8}$$

выполняется неравенство

$$|f(x)-A|<\varepsilon. \tag{10.9}$$

Обозначения предела функции $f(x)$ при $x$, стремящемся к $a$:

$$\lim_{x\to a} f(x)=A; \tag{10.10}$$

$$f(x)\to A \text{ при } x\to a. \tag{10.11}$$

Выясним геометрический смысл этого определения, воспользовавшись графиком функции $y=f(x)$ (рис. 10.3). Неравенство (10.8) означает, что $x$ отстоит от точки $a$ не далее чем на $\delta$, т. е. принадлежит интервалу $]a-\delta, a+\delta[$ или $\delta$-окрестности точки $a$ на оси $Ox$. Неравенство (10.9) означает, что значения функции $y=f(x)$ не выходят из интервала $]A-\varepsilon, A+\varepsilon[$ оси

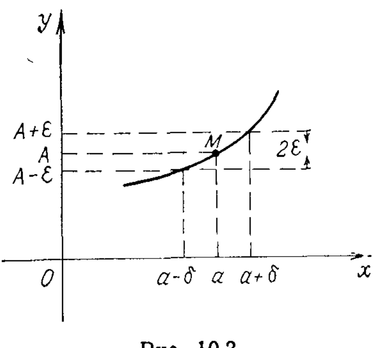



Рис. 10.3

$Oy$, т. е. принадлежат $\varepsilon$-окрестности точки $A$ этой оси. Следовательно, точки $M$ графика функции $y=f(x)$ должны находиться в полоске шириной $2\varepsilon$, ограниченной прямыми $y=A-\varepsilon$, $y=A+\varepsilon$ для всех значений $x$, удаленных от точки $a$ не далее чем на $\delta$. При $\delta \to 0$ в случае непрерывной функции (см. § 10.10) $\varepsilon \to 0$.

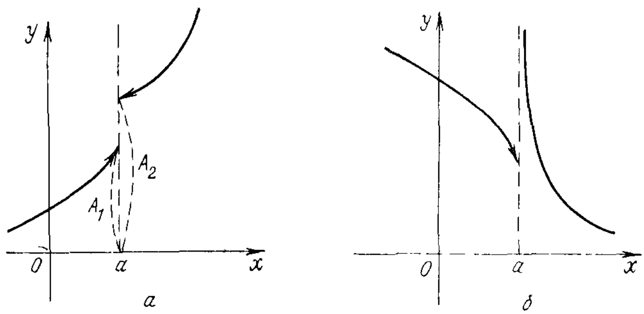



Рис. 10.4

Рассмотрим также односторонние пределы функции: *предел слева* $\lim\limits_{x\to a-0} f(x) = A_1$ ($x$ стремится к $a$, оставаясь меньше $a$: $x < a$) и *предел справа* $\lim\limits_{x\to a+0} f(x) = A_2$ ($x$ стремится к $a$, оставаясь больше $a$: $x > a$).

Когда $a = 0$, то вместо $0 - 0$ пишут $-0$, вместо $0 + 0$ пишут $+0$, поэтому последние формулы принимают вид $\lim\limits_{x\to -0} f(x) = A_1$, $\lim\limits_{x\to +0} f(x) = A_2$.

Можно доказать, что если односторонние пределы равны (рис. 10.3) $\lim\limits_{x\to a-0} f(x) = \lim\limits_{x\to a+0} f(x) = A$, то предел $A$ в точке $x = a$ существует и равен односторонним пределам $\lim\limits_{x\to a} f(x) = A$. Из определения предела функции следует, что предел постоянной равен этой постоянной.

Если односторонние пределы различны (рис. 10.4), $\lim\limits_{x\to a-0} f(x) \neq \lim\limits_{x\to a+0} f(x)$, т. е. $A_1 \neq A_2$, или хотя бы один из них не существует (рис. 10.4, б), то не существует и предел функции в точке $x = a$.

Сформулируем понятие предела функции при $x \to \infty$, т. е. для случая, когда $x$ неограниченно возрастает по модулю.

Число $A$ называется пределом функции $y=f(x)$ при $x \to \infty$, если для любого положительного числа $\varepsilon$ существует положительное число $N$ такое, что для всех значений аргумента $x$, удовлетворяющих условию $|x|>N$, справедливо неравенство $|f(x)-A|<\varepsilon$. Для обозначения предела функции $y=f(x)$ при $x \to \infty$ используется символ $\lim\limits_{x \to \infty} f(x)=A$.

Введем понятие предела функции при стремлении $x$ к бесконечности определенного знака, т. е. для случаев, когда $x$ неограниченно возрастает по модулю, принимая отрицательные $(x \to -\infty)$ или положительные $(x \to +\infty)$ значения.

Число $A$ называется пределом функции $y=f(x)$ при $x \to +\infty$ (при $x \to -\infty$), если для любого положительного числа $\varepsilon$ существует положительное число $N$ такое, что для всех значений аргумента $x$, удовлетворяющих условию $x>N$ $(x<-N)$, выполняется неравенство $|f(x)-A|<\varepsilon$.

Для обозначения введенных понятий используется следующая символика: $\lim\limits_{x \to +\infty} f(x)=A$ $(\lim\limits_{x \to -\infty} f(x)=A)$.

## § 10.6. Бесконечно малые функции и их свойства

Функция $\alpha=\alpha(x)$ называется бесконечно малой при $x \to a$ (или при $x \to \infty$), если $\lim\limits_{x \to a} \alpha(x)=0$ $(\lim\limits_{x \to \infty} \alpha(x)=0)$.

Например, функция $\alpha(x)=(x-3)^2$ есть бесконечно малая при $x \to 3$, так как $\lim\limits_{x \to 3} \alpha(x)=\lim\limits_{x \to 3}(x-3)^2=0$; функция $\alpha(x)=1:x$ является бесконечно малой при $x \to \infty$, поскольку $\lim\limits_{x \to \infty} \alpha(x)=\lim\limits_{x \to \infty} \frac{1}{x}=0$.

Принимая во внимание определение предела функции, определение бесконечно малой функции можно сформулировать следующим образом.

Функция $\alpha=\alpha(x)$ называется бесконечно малой при $x \to a$, если, задав любое число $\varepsilon>0$, можно указать такое число $\delta>0$, что для всех $x$, удовлетворяющих условию

$$0<|x-a|<\delta, \tag{10.12}$$

выполняется неравенство

$$|\alpha(x)|<\varepsilon. \tag{10.13}$$

Свойства бесконечно малых выражаются следующими теоремами.

**Теорема 10.4.** Если функция $y=y(x)$ имеет предел $b$ при $x\to a$, то

$$y(x)=b+\alpha(x),\ y=b+\alpha, \tag{10.14}$$

где $\alpha=\alpha(x)$ — бесконечно малая при $x\to a$. Обратное также верно: если выполнено равенство (10.14), то $\lim\limits_{x\to a} y(x)=b$.

Доказательство. Если $\lim\limits_{x\to a} y(x)=b$, то по определению $|y-b|<\varepsilon$ при $0<|x-a|<\delta$. Обозначив $y-b=\alpha$, получим, что $|\alpha|<\varepsilon$, когда $0<|x-a|<\delta$. Это означает, что $\alpha=\alpha(x)$ — бесконечно малая при $x\to a$ (см. неравенства (10.12) и (10.13)).

Обратно, если $y=b+\alpha$ и $\alpha$ — бесконечно малая при $x\to a$, то $|\alpha|<\varepsilon$, когда $0<|x-a|<\delta$; поскольку $\alpha=y-b$, то $|y-b|<\varepsilon$ при $0<|x-a|<\delta$. Это означает, что предел функции $y=y(x)$ равен $b$, когда $x\to a$.

Отметим, что равенство (10.14) часто используется при нахождении предела функции.

**Теорема 10.5.** Алгебраическая сумма конечного числа бесконечно малых функций при $x\to a$ есть бесконечно малая функция при $x\to a$.

Доказательство. Пусть $u(x)=\alpha(x)-\beta(x)+\gamma(x)$, где $\alpha(x)$, $\beta(x)$, $\gamma(x)$ — бесконечно малые функции при $x\to a$. Следовательно, для любого $\varepsilon>0$ можно указать такое $\delta_1>0$, что $|\alpha(x)|<\frac{\varepsilon}{3}$, если $0<|x-a|<\delta_1$; такое $\delta_2>0$, что $|\beta(x)|<\frac{\varepsilon}{3}$, если $0<|x-a|<\delta_2$; такое $\delta_3>0$, что $|\gamma(x)|<\frac{\varepsilon}{3}$, когда $0<|x-a|<\delta_3$. Обозначим через $\delta$ наименьшее из чисел $\delta_1$, $\delta_2$, $\delta_3$, тогда при $0<|x-a|<\delta$ будут выполнены все указанные неравенства, т. е.

$$|u(x)|\leqslant|\alpha(x)|+|\beta(x)|+|\gamma(x)|<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon,$$

$$|u(x)|<\varepsilon.$$

Итак, $u(x)=\alpha(x)-\beta(x)+\gamma(x)$ — бесконечно малая при $x\to a$.

**Теорема 10.6.** Произведение бесконечно малой на ограниченную функцию есть бесконечно малая функция.

Доказательство. Пусть $\alpha=\alpha(x)$ — бесконечно малая при $x \to a$, а $z=z(x)$ — функция, ограниченная в окрестности точки $x=a$, т. е. для числа $C>0$ найдется окрестность этой точки, в которой $|z(x)|<C$. Для всякого $\varepsilon>0$ можно указать окрестность точки $x=a$, в которой выполняется неравенство $|\alpha(x)|<\frac{\varepsilon}{C}$.

В наименьшей из этих двух окрестностей будет выполнено неравенство

$$|\alpha(x) z(x)|=|\alpha(x)|\,|z(x)|<\frac{\varepsilon}{C}\cdot C=\varepsilon, \quad |\alpha(x) z(x)|<\varepsilon,$$

а это означает, что $\alpha(x)z(x)$ — бесконечно малая функция при $x \to a$.

Следствие 1. Произведение двух бесконечно малых функций есть бесконечно малая функция.

Следствие 2. Произведение бесконечно малой функции на постоянную есть бесконечно малая функция.

Эти следствия получаются непосредственно из теоремы 10.5 с учетом тех фактов, что бесконечно малая функция и постоянная — ограниченные функции.

## § 10.7. Бесконечно большие функции

Функция $y=f(x)$ называется бесконечно большой при $x \to a$, если для любого положительного числа $N$ можно найти такое число $\delta>0$, что при всех значениях $x$, удовлетворяющих условию $0<|x-a|<\delta$, выполняется неравенство $|f(x)|>N$.

Бесконечно большая функция не имеет предела при $x \to a$, но иногда условно говорят, что ее предел равен бесконечности, и пишут $\lim\limits_{x\to a} f(x)=\infty$ или $f(x)\to\infty$ при $x \to a$.

Если $f(x)$ стремится к бесконечности при $x \to a$, принимая только положительные или только отрицательные значения, то соответственно пишут $\lim\limits_{x\to a} f(x)=+\infty$, $\lim\limits_{x\to a} f(x)=-\infty$.

Если функция $f(x)$ стремится к бесконечности при $x \to \infty$, то пишут $\lim\limits_{x\to\infty} f(x)=\infty$.

Примером бесконечно большой функции является функция $f(x)=\frac{1}{x}$ при $x \to 0$. В самом деле, при любом $N>$

$> 0$ неравенство $\left|\frac{1}{x}\right| > N$ будет выполнено, если $|x| = |x - 0| < \frac{1}{N} = \delta$. Эта функция принимает положительные значения при $x > 0$ $(\lim\limits_{x \to +0} f(x) = +\infty)$ и отрицательные — при $x < 0$ $(\lim\limits_{x \to -0} f(x) = -\infty)$.

Функция $f(x) = \frac{1}{(x-2)^2}$ — бесконечно большая при $x \to 2$. Действительно, при любом $N > 0$ неравенство $\frac{1}{(x-2)^2} > N$ будет выполнено, если $(x-2)^2 < \frac{1}{N}$ или $|x - 2| < \frac{1}{\sqrt{N}} = \delta$. Данная функция принимает только положительные значения.

Если функция $\alpha = \alpha(x)$ стремится к нулю при $x \to a$ (или $x \to \infty$) и не обращается в нуль, то функция $y(x) = \frac{1}{\alpha(x)}$ стремится к бесконечности.

Действительно, при любом сколь угодно большом числе $N > 0$ неравенство $\frac{1}{|\alpha|} > N$ будет выполнено, если $|\alpha| < \frac{1}{N}$. Последнее неравенство выполняется для всех значений $\alpha(x)$, начиная с некоторого, так как $\alpha(x) \to 0$.

## § 10.8. Основные теоремы о пределах функций

**Теорема 10.7.** Функция $y = y(x)$ не может иметь более одного предела при $x \to a$.

Доказательство. Предположим противное, пусть функция $y = y(x)$ при $x \to a$ имеет два предела $b_1$ и $b_2$: $\lim\limits_{x \to a} y(x) = b_1$, $\lim\limits_{x \to a} y(x) = b_2$, причем $b_1 \neq b_2$.

Согласно теореме 10.4 из этих равенств следует, что $y = b_1 + \alpha_1$, $y = b_2 + \alpha_2$, где $\alpha_1$ и $\alpha_2$ — бесконечно малые, поэтому $b_1 + \alpha_1 = b_2 + \alpha_2$, или $b_1 - b_2 = \alpha_2 - \alpha_1$. Последнее равенство невозможно, так как в левой части стоит постоянная, отличная от нуля, в правой — бесконечно малая функция.

**Теорема 10.8.** Если каждая из функций $y = y(x)$, $z = z(x)$ имеет предел при $x \to a$, то сумма, разность и произведение этих функций также имеют пределы, причем

$$\lim_{x \to a} [y(x) \pm z(x)] = \lim_{x \to a} y(x) \pm \lim_{x \to a} z(x), \qquad (10.15)$$

$$\lim_{x\to a}[y(x)z(x)] = \lim_{x\to a} y(x) \lim_{x\to a} z(x). \qquad (10.16)$$

Если, кроме того, $\lim\limits_{x\to a} z(x) \neq 0$, то и частное $y(x):z(x)$ имеет предел, причем

$$\lim_{x\to a} \frac{y(x)}{z(x)} = \frac{\lim\limits_{x\to a} y(x)}{\lim\limits_{x\to a} z(x)}. \qquad (10.17)$$

Доказательство. Пусть

$$\lim_{x\to a} y(x) = b, \ \lim_{x\to a} z(x) = c, \qquad (10.18)$$

тогда на основании теоремы 10.4

$$y(x) = b+\beta(x), \ z(x) = c+\gamma(x), \qquad (10.19)$$

где $\beta(x)$ и $\gamma(x)$ — бесконечно малые функции при $x\to a$.

Принимая во внимание равенства (10.19), получаем $y(x)\pm z(x) = (b\pm c)+[\beta(x)\pm\gamma(x)]$, где $(b+c)$ — постоянная; $\beta(x)+\gamma(x)$ — бесконечно малая функция при $x\to a$.

Согласно теореме 10.4 из последнего равенства следует, что $\lim\limits_{x\to a}[y(x)\pm z(x)] = b\pm c = \lim\limits_{x\to a} y(x) \pm \lim z(x)$. Поскольку $y(x)z(x) = [b+\beta(x)][c+\gamma(x)] = bc+[b\gamma(x) + c\beta(x)+\beta(x)\gamma(x)]$, где $bc$ — постоянная; $\delta(x) = b\gamma(x) + c\beta(x) + \beta(x)\gamma(x)$—бесконечно малая (почему?) при $x\to a$, то на основании теоремы 10.4 получаем $\lim\limits_{x\to a}[y(x)z(x)] = bc = \lim\limits_{x\to a} y(x)\cdot\lim\limits_{x\to a} z(x)$.

Предположив, что $\lim\limits_{x\to a} z(x) = c \neq 0$, составим разность

$$\frac{y(x)}{z(x)} - \frac{b}{c} = \frac{b+\beta(x)}{c+\gamma(x)} - \frac{b}{c} = \frac{c\beta(x) - b\gamma(x)}{c[c+\gamma(x)]}.$$

Обозначив $v(x) = \dfrac{1}{c[c+\gamma(x)]}$, $\delta(x) = c\beta(x) - b\gamma(x)$, получим

$$\frac{y(x)}{z(x)} - \frac{b}{c} = v(x)\delta(x), \ \lim_{x\to a} v(x)\delta(x) = 0,$$

так как $v(x)$ — ограниченная функция, а $\delta(x)$ — бесконечно малая при $x\to a$. Следовательно,

$$\lim_{x\to a} \frac{y(x)}{z(x)} = \frac{b}{c} = \frac{\lim\limits_{x\to a} y(x)}{\lim\limits_{x\to a} z(x)}.$$

Следствие 1. Если каждое слагаемое алгебраической суммы функций имеет предел при $x \to a$, то предел суммы существует и равен такой же алгебраической сумме пределов слагаемых.

Следствие 2. Постоянный множитель можно выносить за знак предела:

$$\lim_{x\to a} [cy(x)] = c \cdot \lim_{x\to a} y(x). \tag{10.20}$$

Следствие 3. Если $\lim\limits_{x\to a} y(x) = b$ и $m$ — натуральное число, то

$$\lim_{x\to a} [y(x)^m] = [\lim_{x\to a} y(x)]^m, \tag{10.21}$$

в частности,

$$\lim_{x\to a} (x^m) = (\lim_{x\to a} x)^m = a^m. \tag{10.22}$$

Пример 1. Найти $\lim\limits_{x\to 1} (3x^2 + 8x - 7)$.

На основании формул (10.15), (10.20), (10.22) получаем

$$\lim_{x\to 1} (3x^2 + 8x - 7) = \lim_{x\to 1} (3x^2) + \lim_{x\to 1} (8x) - \lim_{x\to 1} 7 = 3\cdot 1^2 + 8\cdot 1 - 7 = 4.$$

Пример 2. Найти $\lim\limits_{x\to -1} \dfrac{4x^2 - 3x + 1}{2x^2 - 6x - 4}$.

С помощью формулы (10.17) и формул, указанных в примере 1, находим

$$\lim_{x\to -1} \frac{4x^2 - 3x + 1}{2x^2 - 6x - 4} = \frac{\lim\limits_{x\to -1} (4x^2 - 3x + 1)}{\lim\limits_{x\to -1} (2x^2 - 6x - 4)} = \frac{4(-1)^2 - 3(-1) + 1}{2(-1)^2 - 6(-1) - 4} =$$
$$= \frac{8}{4} = 2.$$

Пример 3. Найти $\lim\limits_{x\to 2} \dfrac{x^2 - 3x + 2}{x^2 - 5x + 6}$.

При $x \to 2$ числитель и знаменатель стремятся к нулю, получаем неопределенность вида $\frac{0}{0}$. Чтобы раскрыть ее, предварительно преобразуем данную дробь:

$$\lim_{x\to 2} \frac{x^2 - 3x + 2}{x^2 - 5x + 6} = \lim_{x\to 2} \frac{(x-1)(x-2)}{(x-3)(x-2)} = \lim_{x\to 2} \frac{x-1}{x-3} = \frac{1}{-1} = -1.$$

**Теорема 10.9.** Пусть три функции $u = u(x)$, $y = y(x)$, $v = v(x)$ определены в некотором промежутке, содержащем точку $a$.

Если для любого $x$ из этого промежутка выполняются неравенства

$$u(x) \leqslant y(x) \leqslant v(x) \qquad (10.23)$$

и функции $u=u(x)$, $v=v(x)$ имеют одинаковые пределы при $x \to a$, то функция $y=y(x)$ имеет тот же предел при $x \to a$.

Доказательство. Обозначим через $b$ предел функций $u=u(x)$, $v=v(x)$ при $x \to a$: $\lim\limits_{x \to a} u(x)=b$, $\lim\limits_{x \to a} v(x)=b$. Из неравенств (10.23) следует, что $u(x)-b \leqslant y(x)-b \leqslant v(x)-b$. По определению предела функции для любого числа $\varepsilon>0$ найдутся такие окрестности точки $a$ $O(a, \delta_1)$ и $O(a, \delta_2)$, в которых будут соответственно выполняться неравенства $-\varepsilon<u(x)-b<\varepsilon$, $-\varepsilon<v(x)-b<\varepsilon$.

Обозначим через $\delta$ наименьшее из чисел $\delta_1$, $\delta_2$, тогда для всех $x \in O(a, \delta)$ будут выполнены оба эти неравенства, а поэтому и неравенства $-\varepsilon<y(x)-b<\varepsilon$, $|y(x)-b|<\varepsilon$.

Итак, для любого числа $\varepsilon>0$ можно указать такое $\delta>0$, что при всех $x$, удовлетворяющих условию $0<|x-a|<\delta$, выполняется неравенство $|y(x)-b|<\varepsilon$. Это означает, что $\lim\limits_{x \to a} y(x)=b$.

**Теорема 10.10.** Пусть функция $y=f(x)$ определена в некотором промежутке, содержащем точку $a$. Если при $x \to a$ функция $y=f(x)$ имеет положительный (отрицательный) предел, то найдется $\delta$-окрестность точки $a$ такая, что для всех $x \in O(a, \delta)$ функция положительна (отрицательна).

Доказательство. Пусть $\lim\limits_{x \to a} f(x)=b$. Это означает: для любого числа $\varepsilon>0$ можно указать такое число $\delta>0$, что при всех $x$, удовлетворяющих условию $0<|x-a|<\delta$, выполняется неравенство $|f(x)-b|<\varepsilon$, или $b-\varepsilon<f(x)<b+\varepsilon$.

Если $b>0$, то, взяв $\varepsilon=\frac{1}{2}b$, из неравенства $b-\varepsilon<f(x)$ получим $f(x)>b-\varepsilon=b-\frac{1}{2}b=\frac{1}{2}b>0$, т. е. $f(x)>0$ при $0<|x-a|<\delta$. Если $b<0$, то, взяв $\varepsilon=-\frac{1}{2}b$, из неравенства $f(x)<b+\varepsilon$ получим $f(x)<b-\frac{1}{2}b=\frac{1}{2}b<0$, $f(x)<0$ при $0<|x-a|<\delta$.

Доказанная теорема называется теоремой о сохранении знака функции, имеющей предел.

**Теорема 10.11.** Если функции $u(x)$, $v(x)$ определены в некоторой $\delta$-окрестности точки $a$, для всех $x \in O(a, \delta)$, $x \neq a$, выполняется неравенство $u(x) < v(x)$ и функции имеют пределы при $x \to a$, то $\lim\limits_{x\to a} u(x) \leqslant \lim\limits_{x\to a} v(x)$.

Доказательство. Пусть $\lim\limits_{x\to a} u(x) = b$, $\lim\limits_{x\to a} v(x) = c$. Требуется доказать, что $b \leqslant c$. Предположим противное, т. е. $b > c$.

В силу условия и теоремы 10.8 функция $v(x) - u(x)$ имеет предел $\lim\limits_{x\to a} [v(x) - u(x)] = c - b$, $c - b < 0$, ибо $b > c$.

На основании теоремы 10.10 найдется $\delta$-окрестность $O(a, \delta)$, для всех точек которой $(x \neq a)$ $v(x) - u(x) < 0$, или $v(x) < u(x)$, что противоречит условию.

Следовательно, $b \leqslant c$, т. е. $\lim\limits_{x\to a} u(x) \leqslant \lim\limits_{x\to a} u(x)$.

Замечание. Теорему 10.11 кратко можно сформулировать так: в неравенстве, обе части которого имеют пределы, можно перейти к пределу, присоединив знак равенства. Например,

$$5 + x^2 > 5 - x^2, \quad x \neq 0; \quad \lim_{x\to 0} (5 + x^2) = \lim_{x\to 0} (5 - x^2) = 5.$$

## § 10.9. Предел функции $f(x) = \dfrac{\sin x}{x}$ при $x \to 0$

Данная функция не определена только в точке $x = 0$, так как числитель и знаменатель в ней равны нулю.

Докажем, что эта функция имеет предел при $x \to 0$ и

$$\lim_{x\to 0} \frac{\sin x}{x} = 1. \tag{10.24}$$

Рассмотрим окружность радиуса $R$ (рис. 10.5), радианную меру центрального угла $AOB$ обозначим через $x$; будем считать, что $0 < x < \frac{\pi}{2}$. Поскольку площадь треугольника $AOB$ меньше площади сектора $AOB$, а последняя меньше площади треугольника $AOC$ ($AC$ — касательная к окружности в точке $A$), то

$$\frac{R^2}{2} \sin x < \frac{R^2}{2} x < \frac{R^2}{2} \operatorname{tg} x.$$

Разделив эти неравенства на $\frac{R^2}{2} \sin x$ $\left(\sin x > 0 \text{ при } 0 < x < \frac{\pi}{2}\right)$, получим

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}, \text{ или } \cos x < \frac{\sin x}{x} < 1. \quad (10.25)$$

Неравенства (10.25) в силу четности входящих в них функций верны не только для положительных, но и отрицательных $x$ таких, что $0 < |x| < \frac{\pi}{2}$.

Отметим, что для всех $x$ выполняется неравенство

$$|\sin x| \leqslant |x|. \quad (10.26)$$

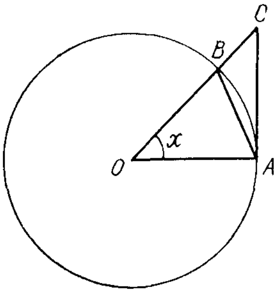



Рис. 10.5

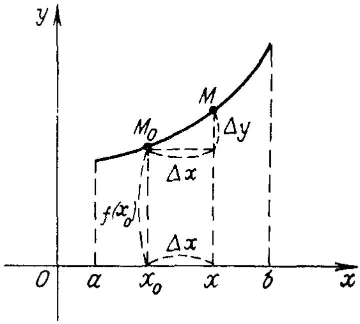



Рис. 10.6

В самом деле, $\sin x = 0$ при $x = 0$; если $0 < |x| < \frac{\pi}{2}$, то неравенство $|\sin x| < |x|$ следует из неравенства (10.25); если $|x| \geqslant \frac{\pi}{2}$, то $|\sin x| \leqslant 1 < \frac{\pi}{2} \leqslant |x|$, т. е. $|\sin x| < < |x|$.

Так как $1 - \cos x = 2\sin^2 \frac{x}{2} = 2\left|\sin \frac{x}{2}\right| \cdot \left|\sin \frac{x}{2}\right| \leqslant \leqslant 2 \cdot 1 \cdot \frac{|x|}{2} = |x|$, $1 - \cos x \leqslant |x|$, то $(1 - \cos x) \to 0$ при $x \to 0$, т. е. $\lim\limits_{x \to 0} \cos x = 1$.

Обращаясь к неравенствам (10.25), заключаем, что функция $f(x) = \sin x : x$ удовлетворяет условиям теоремы 10.9. Согласно этой теореме получаем равенство (10.24), которое и требовалось доказать.

В заключение покажем, что

$$\lim_{x \to 0} \frac{\operatorname{tg} x}{x} = 1. \quad (10.27)$$

Действительно, $\lim\limits_{x\to 0}\frac{\operatorname{tg} x}{x}=\lim\limits_{x\to 0}\left(\frac{\sin x}{x}\ \frac{1}{\cos x}\right)=\lim\limits_{x\to 0}\frac{\sin x}{x}\times$ $\times\lim\limits_{x\to 0}\frac{1}{\cos x}=1.$

## § 10.10. Непрерывность функции в точке

Функция $y=f(x)$, определенная на интервале $]a, b[$, называется непрерывной в точке $x_0\in]a, b[$, если

$$\lim_{x\to x_0} f(x)=f(x_0) \qquad (10.28)$$

(т. е. предел функции равен ее значению при предельном значении аргумента).

Согласно определению предела функции условие (10.28) равносильно следующему: для любого числа $\varepsilon>0$ существует такое число $\delta>0$, что при всех $x$, удовлетворяющих условию $|x-a|<\delta$, выполняется неравенство $|f(x)-f(x_0)|<\varepsilon$.

Если $x_0\in]a, b[$ и $x\in]a, b[$, то разность

$$\Delta x=x-x_0 \qquad (10.29)$$

называется приращением аргумента в точке $x_0$, а разность

$$\Delta y=f(x)-f(x_0),\ \text{или}\ \Delta y=f(x_0+\Delta x)-f(x_0), \qquad (10.30)$$

— приращением функции в той же точке (рис. 10.6); $\Delta y$ — функция $\Delta x$.

Необходимое и достаточное условие непрерывности функции $y=f(x)$ в точке $x_0$ выражается равенством

$$\lim_{\Delta x\to 0}\Delta y=0,\ \text{или}\ \lim_{\Delta x\to 0}[f(x_0+\Delta x)-f(x_0)]=0. \qquad (10.31)$$

Действительно, из равенства (10.28), полагая в нем $x=x_0+\Delta x$, получаем $\lim\limits_{\Delta x\to 0} f(x_0+\Delta x)=f(x_0)$, $\lim\limits_{\Delta x\to 0}[f(x_0+$ $+\Delta x)-f(x_0)]=0$.

Обратно, из равенства (10.31) следует, что $\lim\limits_{x\to x_0}[f(x)-$ $-f(x_0)]=0$, $\lim\limits_{x\to x_0} f(x)=f(x_0)$.

Итак, функция непрерывна в точке, если бесконечно малому приращению аргумента в этой точке соответствует бесконечно малое приращение функции.

Замечание. Поскольку $x_0 = \lim_{x \to x_0} x$, то формулу (10.28) можно переписать так:

$$\lim_{x \to x_0} f(x) = f(\lim_{x \to x_0} x), \tag{10.32}$$

т. е. для непрерывной функции символы предела и функции перестановочны.

Пример. Доказать, что функция $y = \cos x$ непрерывна в каждой точке области ее определения.

Фиксируем произвольное значение $x_0 \in ]-\infty, +\infty[$, тогда $\Delta y = \cos(x_0 + \Delta x) - \cos x_0 = -2\sin\left(x_0 + \frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}$.

Принимая во внимание неравенство $|\sin\alpha| \leqslant |\alpha|$ (см. формулу (10.26)), получаем $|\Delta y| = 2\left|\sin\left(x_0 + \frac{\Delta x}{2}\right)\right| \left|\sin\frac{\Delta x}{2}\right| \leqslant 2 \cdot 1 \times \times \frac{|\Delta x|}{2} = |\Delta x|$, $|\Delta y| \leqslant |\Delta x|$.

Из последнего неравенства следует, что $\lim_{\Delta x \to 0} \Delta y = 0$. Это означает, что функция $y = \cos x$ непрерывна в точке $x_0$.

**Теорема 10.12.** Если функции $f(x)$ и $\varphi(x)$ непрерывны в точке $x_0$, то также непрерывны в этой точке их сумма $f(x)+\varphi(x)$, разность $f(x)-\varphi(x)$, произведение $f(x)\varphi(x)$, а также частное $f(x)/\varphi(x)$ при условии, что $\varphi(x) \neq 0$.

Доказательство непосредственно следует из определения непрерывности функции в точке и теоремы 10.8.

Следствие 1. Целая рациональная функция $P_n(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ непрерывна при всех $x$.

Это следует из теоремы и непрерывности функций $y = x^m$, $m = 1, 2, \ldots, n$ (см. (10.22)), $y = c$ ($c = \text{const}$) ($\Delta y = 0$, $\lim_{\Delta x \to 0} \Delta y = 0$).

Следствие 2. Дробная рациональная функция

$$R(x) = \frac{P_n(x)}{Q_m(x)} = \frac{a_0 + a_1x + a_2x^2 + \ldots + a_nx^n}{b_0 + b_1x + b_2x^2 + \ldots + b_mx^m}$$

непрерывна при всех $x$, для которых знаменатель не обращается в нуль.

**Теорема 10.13.** Если функция $\varphi(x)$ непрерывна в точке $x_0$, а функция $f(y)$ непрерывна в точке $y_0 = \varphi(x_0)$, то сложная функция $F(x) = f[\varphi(x)]$ непрерывна в точке $x_0$.

Доказательство. Зафиксируем число $\varepsilon > 0$. Поскольку функция $f(y)$ непрерывна в точке $y_0 = \varphi(x_0)$, то найдется такое число $\eta > 0$, что $f(y)$ определена на интервале $]y_0 - \eta, y_0 + \eta[$ и выполняется неравенство $|f(y) - f(y_0)| < \varepsilon$ для $0 < |y - y_0| < \eta$. Так как функция $\varphi(x)$

непрерывна в точке $x_0$, то найдется такое число $\delta>0$, что функция определена на интервале $]x_0-\delta, x_0+\delta[$ и $|\varphi(x)-\varphi(x_0)|<\eta$ для $0<|x-x_0|<\delta$.

Из полученных соотношений следует, что для всех $x$, удовлетворяющих неравенству $|x-x_0|<\delta$, функция $f(\varphi(x))$ определена и выполняется неравенство $|f(\varphi(x))-f(\varphi(x_0))|<\varepsilon$ или $|F(x)-F(x_0)|<\varepsilon$, что и требовалось доказать. Например, функция $y=\cos^2 x$ непрерывна при всех $x$, как функция от функции $y=v^2$, где $v=\cos x$, каждая из которых — непрерывная функция своего аргумента.

## § 10.11. Точки разрыва функции

Рассмотрим функцию $y=f(x)$, определенную на интервале $]a, b[$, кроме, быть может, точки $x_0\in]a, b[$. Точка $x_0$ называется точкой разрыва данной функции, если в ней функция определена, но не является непрерывной, или не определена в этой точке.

Если $x_0$ — точка разрыва функции $f(x)$ и существуют конечные пределы $f(x_0-0)=\lim\limits_{x\to x_0-0} f(x)$, $f(x_0+0)=\lim\limits_{x\to x_0+0} f(x)$, то она называется *точкой разрыва первого рода.*

Величина $f(x_0+0)-f(x_0-0)$ называется скачком функции $f(x)$ в точке $x_0$.

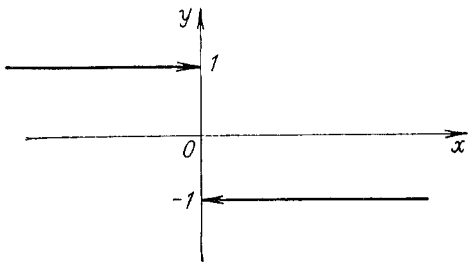



Рис. 10.7

Пример. Функция $f(x)=-\dfrac{x}{|x|}$ в точке $x_0=0$ имеет разрыв первого рода.

Действительно, $f(x)=1$ при $x<0$ и $f(x)=-1$ при $x>0$, в точке $x_0=0$ функция не определена; $\lim\limits_{x\to-0} f(x)=1$, $\lim\limits_{x\to+0} f(x)=-1$. Скачок функции в точке $x_0=0$ равен $-2$ (рис. 10.7).

Пусть функция $y=f(x)$ имеет разрыв в точке $x_0$ и $f(x_0+0)=f(x_0-0)$, тогда $x_0$ называется точкой устранимого разрыва. Это название оправдано тем, что если видоизменить или доопределить такую функцию (если она не была определена в точке $x_0$), положив $f(x_0)=\lim\limits_{x\to x_0+0} f(x)=\lim\limits_{x\to x_0-0} f(x)$, то получится функция, непрерывная в точке $x_0$.

Например, для функции $f(x)=\frac{\sin x}{x}$ (см. § 10.9) точка $x_0=0$ является точкой устранимого разрыва.

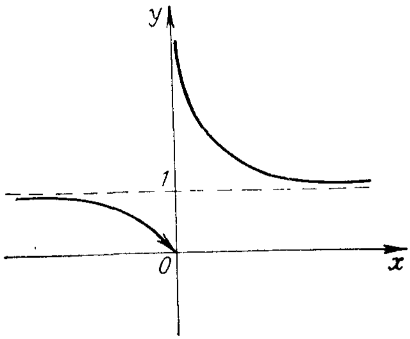



Рис. 10.8

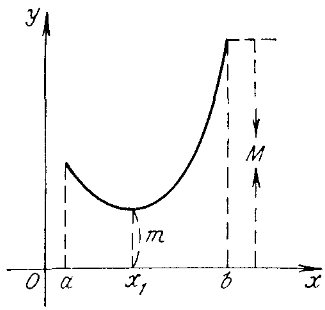



Рис. 10.9

Если $x_0$ — точка разрыва и по крайней мере один из пределов $f(x_0-0)$, $f(x_0+0)$ является бесконечным или не существует, то $x_0$ называется *точкой разрыва второго рода.*

Так, $x_0=0$ — точка разрыва второго рода для функции $f(x)=\frac{1}{x}$, поскольку $f(x_0-0)=-\infty$, $f(x_0+0)=+\infty$; для функции $f(x)=3^{\frac{1}{x}}$ (рис. 10.8) точка $x_0=0$ является точкой разрыва второго рода, в этом случае $\lim f(x_0+0)=+\infty$.

## § 10.12. Непрерывность функции на промежутке

Функция называется непрерывной на интервале, если она непрерывна в каждой точке этого интервала.

Если функция определена при $x=a$ и при этом $\lim\limits_{x\to a+0} f(x)=$

$= f(a)$, то говорят, что $f(x)$ в точке $a$ *непрерывна справа*. Аналогично, если $\lim\limits_{x\to b-0} f(x) = f(b)$, то говорят, что в точке $b$ эта функция *непрерывна слева*.

Функция называется непрерывной на отрезке $[a, b]$, если она непрерывна в каждой его точке (в точке $a$ — непрерывна справа, в точке $b$ — непрерывна слева).

*Наибольшим значением* функции $y=f(x)$ на отрезке $[a, b]$ называется такое ее значение $f(x_1)$, что $f(x) \leqslant f(x_1)$ при всех $x \in [a, b]$.

*Наименьшим значением* функции $y=f(x)$ на отрезке $[a, b]$ называется такое ее значение $f(x_2)$, что $f(x_2) \leqslant f(x)$ для всех $x \in [a, b]$.

Функции, непрерывные на отрезке, обладают рядом важных свойств, которые выражаются следующими теоремами, приводимыми здесь без доказательства.

**Теорема 10.14.** Функция, непрерывная на отрезке $[a, b]$, достигает в нем своего наименьшего значения $m$ и наибольшего значения $M$, т. е. существуют такие точки $x_1$, $x_2$ этого отрезка, что $f(x_1) = m$, $f(x_2) = M$.

Теорема имеет простой геометрический смысл (рис. 10.9).

Замечание. Функция, непрерывная на интервале, этим свойством не обладает. Например, функция $y=3x^2$ иа интервале $]0, 1[$ не достигает значений $m=0$ и $M=3$, так как эти зиачения фуикция принимает в точках $x=0$ и $x=1$, а последние данному интервалу не принадлежат.

**Теорема 10.15.** Если функция $y=f(x)$ непрерывна на отрезке $[a, b]$ и на его концах принимает неравные значения $f(a)=A$, $f(b)=B$, $A \neq B$, то каково бы ни было число

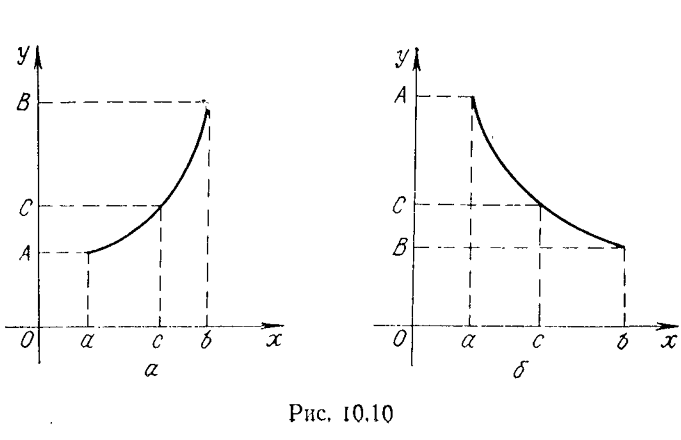



Рис. 10.10

$C$, заключенное между $A$ и $B$, найдется точка $c \in [a, b]$ такая, что $f(c) = C$.

Геометрический смысл теоремы иллюстрируется на рис. 10.10. Всякая прямая $y = C$, где $A < C < B$ (или $A > C > B$), пересекает график функции $y = f(x)$.

Следствие. Если функция непрерывна на отрезке и на его концах принимает значения разных знаков, то на этом отрезке найдется хотя бы одна точка, в которой функция обращается в нуль.

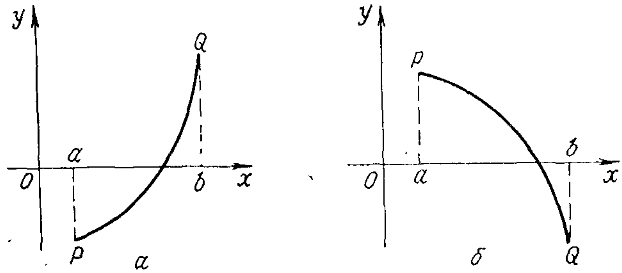



Рис. 10.11

Это частный случай теоремы: $AB < 0$, $C = 0$; геометрический смысл: график непрерывной функции $y = f(x)$, соединяющий точки $P(a, f(a))$, $Q(b, f(b))$, для которых $f(a)f(b) < 0$, пересекает ось $Ox$ (рис. 10.11).

В заключение отметим, что сумма конечного числа функций, непрерывных на некотором отрезке, непрерывна на этом отрезке.

## § 10.13. Предел последовательности

*Числовой последовательностью* (или *последовательностью*) называется функция

$$x_n = \varphi(n), \ n = 1, 2, 3, \ldots, \tag{10.33}$$

определенная на множестве натуральных чисел. Каждое значение $x_n$ ($n = 1, 2, 3, \ldots$) называется *элементом последовательности*, а число $n$ — его *номером*.

Числовую последовательность с элементом $x_n$ обозначают либо $x_n$, $n = 1, 2, 3, \ldots$, либо $(x_1, x_2, \ldots, x_n, \ldots)$, либо $(x_n)$.

Примеры числовых последовательностей:

1) $(n)=(1, 2, 3, \ldots)$; 2) $\left(\frac{1}{n}\right)=\left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right)$; 3) $\left(\frac{(-1)^{n-1}}{n^2}\right)=\left(1, -\frac{1}{4}, \frac{1}{9}, -\frac{1}{16}, \ldots\right)$; 4) $(\cos \pi n)=$ $=(-1, 1, -1, 1, \ldots)$; 5) $(1)=(1, 1, 1, \ldots)$; 6) $(c)=$ $=(c, c, c, \ldots)$.

Числовая последовательность всегда содержит бесконечное множество элементов; среди них, конечно, могут быть и равные (как в примерах 4)—6)).

Число $a$ называется *пределом последовательности* $(x_n)$, если для любого числа $\varepsilon>0$ найдется такое натуральное число $N$, что при всех $n>N$ выполняется неравенство

$$|x_n-a|<\varepsilon. \qquad (10.34)$$

Обозначения предела последовательности $(x_n)$: $\lim\limits_{n\to\infty} x_n=a$; $x_n\to a$ при $n\to\infty$ (читается: $x_n$ стремится к $a$, когда $n$ стремится к бесконечности).

Последовательность, имеющая предел, называется *сходящейся*, последовательность, у которой нет предела,— *расходящейся*.

Отметим, что неравенство (10.34) равносильно неравенствам $a-\varepsilon<x_n<a+\varepsilon$.

Определение предела имеет следующий геометрический смысл: число $a$ является пределом последовательности $(x_n)$, если в любой его $\varepsilon$-окрестности, т. е. в интервале $]a-\varepsilon, a+\varepsilon[$, содержатся почти все члены $(x_n)$ или вне этой окрестности находится лишь конечное число членов данной последовательности.

Пример 1. Последовательность $\left(3+\frac{1}{n}\right)$ сходится и имеет предел $a=3$.

Действительно, каково бы ни было число $\varepsilon>0$, найдется такое натуральное число $N$, что $|x_n-a|=\left|\left(3+\frac{1}{n}\right)-3\right|=\frac{1}{n}<\varepsilon$, $\frac{1}{n}<\varepsilon$ для $n>N$; неравенство $\frac{1}{n}<\varepsilon$ будет выполнено для всех $n>N$, если $N>\frac{1}{\varepsilon}$, т. е. в качестве $N$ можно взять большее из двух последовательных натуральных чисел, между которыми заключено число $\frac{1}{\varepsilon}$. Например, если $\varepsilon_1=0,1$, то $N_1=10$; если $\varepsilon_2=0,01$, то $N_2=100$ и т. д.

З а м е ч а н и е. Одновременно показано, что последовательность $\left(\frac{1}{n}\right)$ имеет пределом нуль, т. е.

$$\lim_{n\to\infty}\frac{1}{n}=0. \qquad (10.35)$$

П р и м е р 2. Последовательность $\left(\cos\frac{n\pi}{2}\right)$ является расходящейся.

В самом деле, каково бы ни было число $a$, вне его $\varepsilon$-окрестности, например, при $0<\varepsilon<1$, заведомо лежит бесконечное множество элементов данной последовательности (хотя среди них и много равных между собой); это означает, что $a$ не является ее пределом.

Из определения предела последовательности следует, что предел постоянной равен этой постоянной: $\lim\limits_{n\to\infty} c=c$ ($c=\mathrm{const}$), поскольку в данном случае $x_n=c$, $a=c$, $|x_n-a|=0<\varepsilon$ для любого $\varepsilon>0$.

Из определения следует также, что последовательность может иметь только один предел. Действительно, если допустить, что таких пределов два: $a$ и $b$, причем $a<b$, то каждый из интервалов $]a-\varepsilon, a+\varepsilon[$ и $]b-\varepsilon, b+\varepsilon[$, где $\varepsilon=\frac{b-a}{3}$, должен был бы содержать все точки последовательности $(x_n)$, за исключением их конечного числа, но это невозможно, потому что указанные интервалы не имеют общих точек.

Последовательность $(x_n)$ называется *ограниченной сверху (снизу)*, если существует такое число $A$, что $x_n\leqslant A$ (соответственно $x_n\geqslant A$) для всех номеров $n$.

Последовательность $(x_n)$, ограниченная сверху и снизу, называется просто *ограниченной*.

Очевидно, последовательность $(x_n)$ ограничена тогда и только тогда, когда существует такое число $c>0$, что $|x_n|\leqslant c$ для всех номеров $n$.

Например, последовательности $\left(\frac{1}{n}\right)$, $\left(\frac{1}{n^2}\right)$, $\left(\cos\frac{\pi n}{2}\right)$ ограничены, последовательность $(n^2)$ ограничена снизу, но не ограничена сверху, последовательность $\left(n\cos\frac{\pi n}{2}\right)$ не является ограниченной ни сверху, ни снизу.

**Теорема 10.15.** Если последовательность имеет предел, то она ограничена.

Д о к а з а т е л ь с т в о. Пусть последовательность $(x_n)$ сходится и $\lim\limits_{n\to\infty} x_n=a$. Это значит, что для любого

$\varepsilon > 0$ можно указать такое $N$, что $|x_n - a| < \varepsilon$ для всех $n > N$. Будем считать, что $\varepsilon < 1$. Обозначим через $d$ наибольшее из чисел 1, $|x_1 - a|$, $|x_2 - a|$, ..., $|x_N - a|$, тогда $|x_n - a| \leqslant d$, т. е. $a - d \leqslant x_n \leqslant a + d$ для всех $n$. Это означает, что последовательность $(x_n)$ ограничена.

Число $a$ называется *верхней гранью* последовательности $(x_n)$, если: 1) $x_n \leqslant a$ при всех $n$; 2) для любого $\varepsilon > 0$ существует такой номер $N$, что $x_N > a - \varepsilon$. Верхняя грань последовательности $(x_n)$ обозначается $\sup(x_n)$, или $\sup x_n$.

Аналогично определяется нижняя грань последовательности $(x_n)$ и обозначается $\inf(x_n)$ или $\inf x_n$.

В качестве примеров отметим, что $\sup\left(\frac{n-1}{n}\right) = 1$, $\inf\left(\frac{n-1}{n}\right) = 0$, $\sup(n^2) = \infty$, $\inf(n^2) = 1$.

Последовательность $(x_n)$ называется *монотонно возрастающей (монотонно убывающей)*, если $x_n \leqslant x_{n+1}$ (соответственно $x_n \geqslant x_{n+1}$) при всех $n$. Монотонно возрастающие и монотонно убывающие последовательности называются просто *монотонными*.

**Теорема 10.16.** Всякая ограниченная сверху (снизу) монотонно возрастающая (монотонно убывающая) последовательность $(x_n)$ имеет предел, причем $\lim\limits_{n\to\infty} x_n = \sup(x_n)$ (соответственно $\lim\limits_{n\to\infty} x_n = \inf(x_n)$).

Доказательство. Пусть последовательность $(x_n)$ монотонно возрастает и ограничена сверху. В силу этих условий она имеет верхнюю грань $\sup(x_n) = a$. Докажем, что $a = \lim\limits_{n\to\infty} x_n$. Зафиксируем произвольное $\varepsilon > 0$. По определению верхней грани $x_n \leqslant a$ для всех $n$ и существует такой номер $N$, что $x_N > a - \varepsilon$, тогда в силу монотонности последовательности при всех $n > N$ получаем $a - \varepsilon < x_N \leqslant x_n \leqslant a < a + \varepsilon$. Следовательно, $a - \varepsilon < x_n < a + \varepsilon$, т. е. $|x_n - a| < \varepsilon$ для $n > N$, а это означает, что $\lim\limits_{n\to\infty} x_n = a$.

Аналогично доказывается существование предела для ограниченной снизу монотонно убывающей последовательности.

В заключение отметим, что для пределов последовательностей верны теоремы, аналогичные теоремам 10.4—10.8.

## § 10.14. Число $e$

Рассмотрим последовательность

$$x_n = \left(1 + \frac{1}{n}\right)^n, \quad n = 1, 2, 3, \ldots \qquad (10.36)$$

Придавая $n$ различные натуральные значения, получаем таблицу:

| $n$ | 1 | 2 | 10 | 100 | 1000 | 10000 | 100000 |
|---|---|---|---|---|---|---|---|
| $x_n$ | 2 | 2,25 | 2,594 | 2,705 | 2,717 | 2,718 | 2,718 |

Из таблицы видно, что $x_n$ с возрастанием $n$ изменяется все медленнее и, по-видимому, стремится к некоторому пределу. Докажем, что последовательность (10.36) при $n \to \infty$ имеет предел, заключенный между числами 2 и 3.

Применяя формулу бинома Ньютона (см. § 11.14), получаем

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

$$\left.\begin{aligned} x_n = \left(1 + \frac{1}{n}\right)^n = 2 + \frac{1}{2}\left(1 - \frac{1}{n}\right) + \\ + \frac{1}{2 \cdot 3}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \ldots + \\ + \frac{1}{2 \cdot 3 \ldots n}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \ldots \left(1 - \frac{n-1}{n}\right). \end{aligned}\right\} \qquad (10.37)$$

Все слагаемые в этой формуле положительны при $n > 1$, с возрастанием $n$ увеличивается число слагаемых и каждое соответствующее слагаемое становится больше. Таким образом, последовательность $(x_n)$ возрастает вместе с $n$; наименьшее значение равно 2, $x_1 = 2$.

С другой стороны, каждое слагаемое правой части формулы (10.37) увеличится, если все множители знаменателей заменить двойкой, а каждую скобку — единицей, т. е. при всех $n$

$$x_n = \left(1+\frac{1}{n}\right)^n < 2 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}} = 2 +$$

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

Итак, последовательность $(x_n)$ ограничена и монотонно возрастает, поэтому в силу теоремы 10.16 она имеет предел. Этот предел называют числом $e$; следовательно, по определению

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n. \qquad (10.38)$$

Можно доказать, что

$$e = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x, \qquad (10.39)$$

где $x$ принимает действительные значения. Полагая $\frac{1}{x} = \alpha$ $(\alpha > -1)$, получаем еще одно выражение для числа $e$:

$$e = \lim_{\alpha\to 0}(1+\alpha)^{\frac{1}{\alpha}}. \qquad (10.40)$$

Число $e$ является иррациональным. (Доказательство этого утверждения здесь не приводится.) Укажем значение числа $e$ с первыми семью десятичными знаками: $e = 2{,}7182818\ldots$.

## § 10.15. Натуральные логарифмы

Рассмотрим логарифмическую функцию $y = \log_a x$. Число $a$ $(a \neq 1,\ a > 0)$ называется *основанием логарифмов*. Если $a = 10$, то $y$ называется *десятичным логарифмом* числа $x$ и обозначается $y = \lg x$.

Логарифмы с основанием $e = 2{,}71828\ldots$ называются натуральными логарифмами и обозначаются ln, т. е. $\log_e x = \ln x$.

Функция $y = \ln x$ определена для положительных $x$ (на рис. 10.12 изображен ее график и график функции $y = \lg x$).

Установим зависимость между десятичными и натуральными логарифмами одних и тех же чисел. Если $y =$

$=\lg x$, то $x=10^y$. Логарифмируя левую и правую часть равенства $x=10^y$ по основанию $e$, получаем $\ln x=y\ln 10$, откуда $y=\frac{1}{\ln 10}\ln x$. Следовательно, $\lg x=M\ln x$, $M=\frac{1}{\ln 10}\approx 0{,}434294$. Число $M$ называется *модулем перехода* от натуральных логарифмов к десятичным.

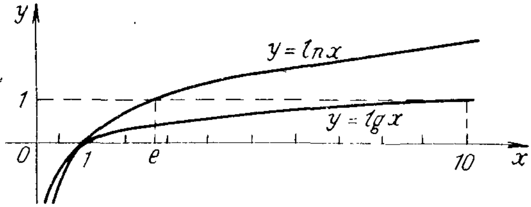



Рис. 10.12

Если положить в равенстве $\lg x=M\ln x$, $x=e$, то, поскольку $\ln e=1$, получим выражение числа $M$ через десятичные логарифмы: $M=\lg e$.

Натуральные логарифмы через десятичные выражаются формулой $\ln x=\frac{1}{M}\lg x$, $\frac{1}{M}=2{,}302585$. Функция $y=\ln|x|$ определена для всех действительных значений $x$, кроме $x=0$; ее график изображен на рис. 10.13.

## § 10.16. Показательная функция. Гиперболические функции

*Показательной* (или *экспоненциальной*) функцией называется функция $y=a^x$ $(a>0,\ a\neq 1)$. Пусть $a=e$ (см. § 10.14), в этом случае показательная (экспоненциальная) функция обозначается

$$y=e^x=\exp x. \tag{10.41}$$

Показательную функцию с другим основанием можно привести к показательной функции с основанием $e$. Действительно, по определению логарифма $a=e^{\ln a}$, поэтому $a^x=(e^{\ln a})^x=e^{kx}$, где $k=\ln a$.

*Гиперболическим синусом* называется функция

$$\operatorname{sh} x=\frac{e^x-e^{-x}}{2}, \tag{10.42}$$

*гиперболическим косинусом* — функция

$$\operatorname{ch} x = \frac{e^x + e^{-x}}{2}. \tag{10.43}$$

Аналогично определяются *гиперболический тангенс*

$$\operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{10.44}$$

и *гиперболический котангенс*

$$\operatorname{cth} x = \frac{\operatorname{ch} x}{\operatorname{sh} x} = \frac{e^x + e^{-x}}{e^x - e^{-x}}. \tag{10.45}$$

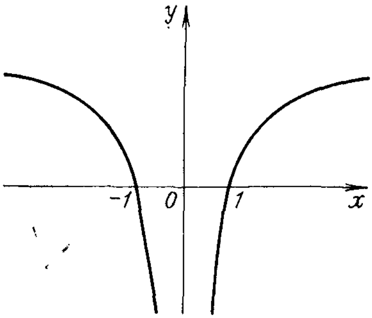



Рис. 10.13

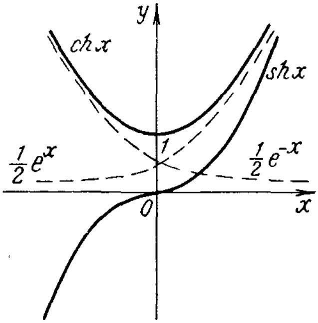



Рис. 10.14

Функции, определяемые формулами (10.42)—(10.45), называют *гиперболическими*.

Названия «синус» и «косинус» объясняются тем, что между функциями (10.42), (10.43) существуют соотношения, аналогичные соотношениям между обычными (круговыми) синусами и косинусами. Например, справедливы формулы

$$\begin{gathered} \operatorname{sh} 2x = 2 \operatorname{sh} x \operatorname{ch} x, \quad \operatorname{ch}^2 x - \operatorname{sh}^2 x = 1, \\ \operatorname{ch}^2 x + \operatorname{sh}^2 x = \operatorname{ch} 2x, \quad \operatorname{th} x \operatorname{cth} x = 1. \end{gathered} \tag{10.46}$$

*Докажем первую из них, воспользовавшись* равенствами (10.42), (10.43):

$$2 \operatorname{sh} x \operatorname{ch} x = 2 \frac{e^x - e^{-x}}{2} \frac{e^x + e^{-x}}{2} = \frac{e^{2x} - e^{-2x}}{2} = \operatorname{sh} 2x;$$

доказательство трех остальных предоставляется читателю.

Эпитет «гиперболический» в названии рассматриваемых функций связан с тем обстоятельством, что уравнения

$$x=a\operatorname{ch} t,\ y=a\operatorname{sh} t \qquad (10.47)$$

параметрически определяют гиперболу, подобно тому как уравнения (1.44) параметрически определяют окружность. Действительно, возводя в квадрат обе части каждого уравнения (10.47), почленно вычитая и используя

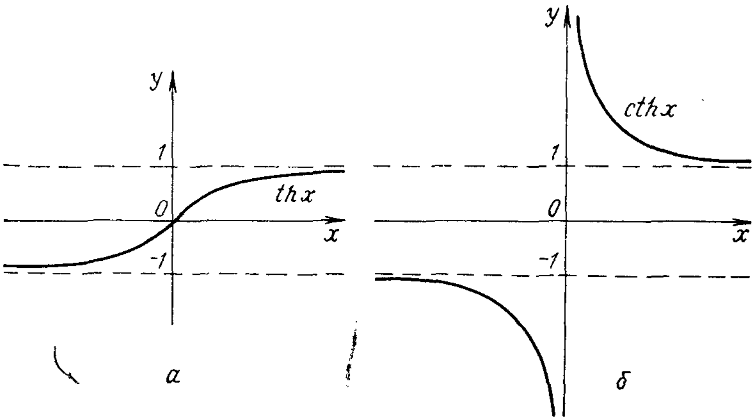



Рис. 10.15

вторую из формул (10.46), получаем уравнение $x^2-y^2=$ $=a^2$.

Функция $y=\operatorname{ch} x$ является четной, поскольку $\operatorname{ch}(-x)=$ $=\dfrac{e^{-x}+e^{x}}{2}=\operatorname{ch} x$, а функции $y=\operatorname{sh} x$, $y=\operatorname{th} x$, $y=$ $=\operatorname{cth} x$ — нечетные, так как $\operatorname{sh}(-x)=\dfrac{e^{-x}-e^{x}}{2}=$ $=-\dfrac{e^{x}-e^{-x}}{2}=-\operatorname{sh} x$, $\operatorname{th}(-x)=\dfrac{\operatorname{sh}(-x)}{\operatorname{ch}(-x)}=\dfrac{-\operatorname{sh} x}{\operatorname{ch} x}=$ $=-\operatorname{th} x$, $\operatorname{cth}(-x)=-\operatorname{cth} x$.

Гиперболические функции имеют вполне определенные значения при всех значениях $x$, кроме функции $y=$ $=\operatorname{cth} x$, которая в точке $x=0$ обращается в бесконечность. Отметим, что $\operatorname{sh} 0=0$, $\operatorname{ch} 0=1$, как и для обычных синуса и косинуса.

Гиперболические функции не обладают важнейшим свойством тригонометрических функций — свойством периодичности. Кроме того, множество значений каждой гиперболической функции существенно отличается от множества значений соответствующей тригонометрической функции. Функция $y=\operatorname{sh} x$ принимает все действительные значения, т. е. множество ее значений совпадает с бесконечным интервалом $]-\infty, +\infty[$; $y=\operatorname{ch} x$ — значения, не меньшие единицы ($1 \leqslant \operatorname{ch} x < +\infty$); значения функции $y=\operatorname{th} x$ по модулю не превышают единицы ($-1 < \operatorname{th} x < 1$); значения $y=\operatorname{cth} x$ по модулю больше единицы ($\operatorname{cth} x > 1$ при $x > 0$, $\operatorname{cth} x < -1$ при $x < 0$).

Графики гиперболических функций представлены на рис. 10.14 и 10.15. На рис. 10.14 приведены также пунктирные линии $y=\frac{1}{2}e^x$ и $y=\frac{1}{2}e^{-x}$, с помощью которых можно построить линии $y=\operatorname{sh} x$, $y=\operatorname{ch} x$. Прямые $y=+1$, $y=-1$ являются асимптотами графиков функций $y=\operatorname{th} x$, $y=\operatorname{cth} x$ (при $x$, достаточно больших по модулю, $e^{-x}$ принимает достаточно малые значения, поэтому достаточно мало отличаются числитель и знаменатель в формулах (10.44) и (10.45)). Кроме того, ось $Oy$ служит асимптотой графика функции $y=\operatorname{cth} x$.

## § 10.17. Некоторые важные пределы

Докажем, что

$$\lim_{x\to 0} \frac{\log_a (1+x)}{x} = \log_a e; \qquad (10.48)$$

$$\lim_{x\to 0} \frac{a^x - 1}{x} = \ln a; \qquad (10.49)$$

$$\lim_{x\to 0} \frac{(1+x)^\alpha - 1}{x} = \alpha. \qquad (10.50)$$

Поскольку

$$\frac{\log_a (1+x)}{x} = \frac{1}{x}\log_a (1+x) = \log_a (1+x)^{\frac{1}{x}}$$

и логарифмическая функция непрерывна, то

$$\lim_{x\to 0} \frac{\log_a (1+x)}{x} = \lim_{x\to 0} [\log_a (1+x)^{\frac{1}{x}}] =$$

$$= \log_a [\lim_{x\to 0} (1+x)^{\frac{1}{x}}] = \log_a e,$$

откуда и следует формула (10.48).

В частном случае, при $a=e$, получаем

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=1. \qquad (10.51)$$

Для доказательства формулы (10.49) положим $a^x-1=y$; в силу непрерывности функции $f(x)=a^x$, $y\to 0$ при $x\to 0$. Из равенства $a^x-1=y$ находим $a^x=1+y$, $x=\log_a(1+y)$, поэтому

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

Отметим, что при $a=e$ отсюда следует формула $\lim\limits_{x\to 0}\frac{e^x-1}{x}=1.$

Чтобы доказать формулу (10.50), положим $1+x=e^y$, тогда $x=e^y-1$; $y\to 0$ при $x\to 0$. С учетом формулы (10.49) получаем

$$\lim_{x\to 0}\frac{(1+x)^\alpha-1}{x}=\lim_{y\to 0}\frac{e^{\alpha y}-1}{e^y-1}=\alpha\cdot\lim_{y\to 0}\left(\frac{e^{\alpha y}-1}{\alpha y}\,\frac{y}{e^y-1}\right)=\alpha.$$

Пример 1. Найти $\lim\limits_{x\to 2}\frac{\ln(x^2-5x+7)}{x-2}$.

Положим $x-2=\alpha$, тогда $x^2-5x+7=(2+\alpha)^2-5(2+\alpha)+7=1-\alpha+\alpha^2$,

$$\lim_{x\to 2}\frac{\ln(x^2-5x+7)}{x-2}=\lim_{\alpha\to 0}\frac{\ln(1+\alpha^2-\alpha)}{\alpha^2-\alpha}\,\frac{\alpha^2-\alpha}{\alpha}=$$

$$=1\cdot(-1)=-1.$$

Пример 2.

$$\lim_{x\to 0}\frac{9^x-3^x}{x}=\lim_{x\to 0}3^x\cdot\frac{3^x-1}{x}=3^0\cdot\ln 3=\ln 3.$$

Пример 3.

$$\lim_{x\to 0}\frac{\sqrt[7]{1+\sin x}-1}{x}=\lim_{x\to 0}\frac{(1+\sin x)^{\frac{1}{7}}-1}{\sin x}\,\frac{\sin x}{x}=$$

$$=\frac{1}{7}\cdot 1=\frac{1}{7}.$$

## § 10.18. Сравнение бесконечно малых функций

Рассматриваемые в этом параграфе функции определены на интервале $]a, b[$ — конечном или бесконечном, кроме, быть может, точки $x_0\in]a, b[$. Если $x_0=a$ или $x_0=b$, то под пределом в точке $x_0$ понимается соответственно предел справа или слева.

В § 10.6 было показано, что сумма, разность и произведение бесконечно малых есть бесконечно малая функция. Что касается частного двух бесконечно малых, то здесь могут встретиться самые разнообразные случаи.

Бесконечно малые $\alpha(x)$ и $\beta(x)$ при $x \to x_0$ называются *величинами одного порядка,* если их отношение имеет конечный предел, отличный от нуля, т. е.

$$\lim_{x\to x_0} \frac{\alpha(x)}{\beta(x)} = c, \text{ или } \lim_{x\to x_0} \frac{\beta(x)}{\alpha(x)} = \frac{1}{c} \; (c \neq 0). \qquad (10.52)$$

В этом случае пишут

$$\alpha(x) = O(\beta(x)) \text{ при } x \to x_0 \qquad (10.53)$$

(читается: $\alpha(x)$ есть $O$ большое от $\beta(x)$ при $x \to x_0$).

Например, $\alpha(x) = \sin x$ и $\beta(x) = 5x$ при $x \to 0$ являются бесконечно малыми одного порядка, так как $\lim\limits_{x\to 0} \frac{\alpha(x)}{\beta(x)} = \frac{1}{5}$, $\sin x = O(5x)$ при $x \to 0$.

Если предел (10.52) равен нулю ($c=0$), то величина $\alpha(x)$ называется *бесконечно малой высшего порядка* по сравнению с $\beta(x)$ (а величина $\beta(x)$ — бесконечно малая низшего порядка по сравнению с $\alpha(x)$). В данном случае применяется обозначение

$$\alpha(x) = o(\beta(x)) \text{ при } x \to x_0 \qquad (10.54)$$

(читается: $\alpha(x)$ есть $o$ малое от $\beta(x)$ при $x \to x_0$). Например, $x^2 = o(\sin x)$ при $x \to 0$, поскольку

$$\lim_{x\to 0} \frac{x^2}{\sin x} = \lim_{x\to 0} x \cdot \lim_{x\to 0} \frac{x}{\sin x} = 0 \cdot 1 = 0.$$

Функция $\beta(x)$ называется бесконечно малой $k$-го порядка относительно функции $\alpha(x)$, если $\beta(x)$ и $(\alpha(x))^k$ — бесконечно малые одного порядка, т. е.

$$\lim_{x\to x_0} \frac{\beta(x)}{(\alpha(x))^k} = c \; (c \neq 0). \qquad (10.55)$$

Например, если $\alpha(x) = x$, $\beta(x) = x^4$, то при $x \to 0$ $\beta(x)$ — бесконечно малая четвертого порядка относительно бесконечно малой $\alpha(x)$ (но бесконечно малая второго порядка по сравнению с $\gamma(x) = x^2$).

Две бесконечно малые функции $\alpha(x)$ и $\beta(x)$ называются *эквивалентными* (или *равносильными*) бесконечно малыми при $x \to x_0$, если предел их отношения равен единице, т. е.

$$\lim_{x \to x_0} \frac{\alpha(x)}{\beta(x)} = 1 \quad \text{или} \quad \lim_{x \to x_0} \frac{\beta(x)}{\alpha(x)} = 1. \tag{10.56}$$

Эквивалентность бесконечно малых $\alpha(x)$ и $\beta(x)$ обозначается символом $\sim$: $\alpha(x) \sim \beta(x)$ при $x \to x_0$.

Из результатов § 10.9 и 10.17 следует, что при $x \to 0$ $x \sim \sin x \sim \operatorname{tg} x$, $\ln(1+x) \sim x$.

**Теорема 10.17.** Бесконечно малые функции $\alpha(x)$ и $\beta(x)$ эквивалентны тогда и только тогда, когда их разность есть бесконечно малая высшего порядка, чем $\alpha(x)$ и $\beta(x)$.

Доказательство. *Необходимость.* Пусть $\alpha(x)$ и $\beta(x)$ эквивалентны, т. е. выполнены равенства (10.56), тогда

$$\lim_{x \to x_0} \frac{\alpha(x) - \beta(x)}{\alpha(x)} = \lim_{x \to x_0} \left(1 - \frac{\beta(x)}{\alpha(x)}\right) = 1 - 1 = 0,$$

$$\lim_{x \to x_0} \frac{\alpha(x) - \beta(x)}{\alpha(x)} = 0.$$

Это означает, что разность $\alpha(x) - \beta(x)$ есть бесконечно малая высшего порядка, чем $\alpha(x)$ и $\beta(x)$.

*Достаточность.* Если $\lim\limits_{x \to x_0} \dfrac{\alpha(x) - \beta(x)}{\alpha(x)} = 0$, т. е.

$$\lim_{x \to x_0} \left(1 - \frac{\beta(x)}{\alpha(x)}\right) = 0, \quad \text{то} \quad \lim_{x \to x_0} \frac{\beta(x)}{\alpha(x)} = 1.$$

Аналогично, если $\lim\limits_{x \to x_0} \dfrac{\alpha(x) - \beta(x)}{\beta(x)} = 0$, т. е.

$$\lim_{x \to x_0} \left(\frac{\alpha(x)}{\beta(x)} - 1\right) = 0, \quad \text{то} \quad \lim_{x \to x_0} \frac{\alpha(x)}{\beta(x)} = 1.$$

Следовательно, бесконечно малые $\alpha(x)$ и $\beta(x)$ эквивалентны при $x \to x_0$.

**Теорема 10.18.** Если $\alpha(x) \sim \alpha_1(x)$, $\beta(x) \sim \beta_1(x)$ при $x \to x_0$ и существует $\lim\limits_{x \to x_0} \dfrac{\alpha_1(x)}{\beta_1(x)}$, то существует $\lim\limits_{x \to x_0} \dfrac{\alpha(x)}{\beta(x)}$, причем

$$\lim_{x \to x_0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to x_0} \frac{\alpha_1(x)}{\beta_1(x)}.$$

Доказательство. В силу условия

$$\lim_{x\to x_0}\frac{\alpha(x)}{\alpha_1(x)}=1,\quad \lim_{x\to x_0}\frac{\beta(x)}{\beta_1(x)}=1.$$

Следовательно,

$$\lim_{x\to x_0}\frac{\alpha(x)}{\beta(x)}=\lim_{x\to x_0}\frac{\alpha(x)}{\alpha_1(x)}\,\frac{\alpha_1(x)}{\beta_1(x)}\,\frac{\beta_1(x)}{\beta(x)}=$$

$$=\lim_{x\to x_0}\frac{\alpha(x)}{\alpha_1(x)}\cdot\lim_{x\to x_0}\frac{\alpha_1(x)}{\beta_1(x)}\cdot\lim_{x\to x_0}\frac{\beta_1(x)}{\beta(x)}=$$

$$=1\cdot\lim_{x\to x_0}\frac{\alpha_1(x)}{\beta_1(x)}\cdot 1=\lim_{x\to x_0}\frac{\alpha_1(x)}{\beta_1(x)},$$

что и требовалось доказать.

Следствие. Если $\alpha_1(x)\sim\beta(x)$, $\alpha_2(x)\sim\beta(x)$ при $x\to x_0$, то $\alpha_1(x)\sim\alpha_2(x)$ при $x\to x_0$.

Пример. Доказать, что бесконечно малые функции $\alpha(x)=x$ и $\beta(x)=x+x^2$ эквивалентны при $x\to 0$.

Поскольку разность $\beta(x)-\alpha(x)=x^2$ есть бесконечно малая высшего порядка по сравнению с каждой из них:

$$\frac{\beta(x)-\alpha(x)}{\alpha(x)}=\frac{x^2}{x}=x\to 0,\quad \frac{\beta(x)-\alpha(x)}{\beta(x)}=\frac{x^2}{x+x^2}=$$

$$=\frac{x}{1+x}\to 0,$$

то $\alpha(x)\sim\beta(x)$ при $x\to 0$.

Замечание. Если отношение $\alpha(x):\beta(x)$ двух бесконечно малых функций при $x\to x_0$ не имеет предела и не стремится к бесконечности, то бесконечно малые функции $\alpha(x)$ и $\beta(x)$ несравнимы между собой.

Например, несравнимы при $x\to 0$ бесконечно малые функции $\alpha(x)=x\sin\frac{1}{x}$, $\beta(x)=x$, так как

$$\frac{\alpha(x)}{\beta(x)}=\frac{x\sin\frac{1}{x}}{x}=\sin\frac{1}{x}$$

и $\sin\frac{1}{x}$ не имеет предела при $x\to 0$.

Глава 11

# ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ

Производная — одно из важнейших понятий современной математики. С помощью производной можно характеризовать скорость изменения функции, скорость протекания некоторого процесса, явления. С понятием производной тесно связано понятие дифференциала. Дифференциал функции — часть ее приращения, причем наиболее существенная при достаточно малых приращениях аргумента.

## § 11.1. Задачи, приводящие к понятию производной

К понятию производной приводят многие задачи геометрии, физики, химии и других наук. Рассмотрим некоторые из них.

**Задача о касательной.** *Касательной* к линии $l$ в точке $M_0$ называется прямая $M_0T$ — предельное положение секущей $M_0M$ (рис. 11.1), когда точка $M$ стремится к $M_0$

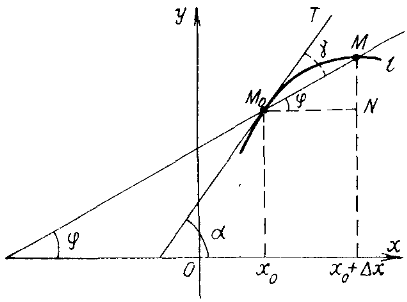



Рис. 11.1

вдоль данной линии (т. е. угол $\gamma$ стремится к нулю) произвольным образом.

Задача о касательной состоит в следующем: написать уравнение касательной к линии $y=f(x)$ в точке $M_0(x_0, y_0)$.

Введем декартову прямоугольную систему координат $Oxy$, рассмотрим график функции $y=f(x)$. Пусть $M_0(x_0, y_0)$ — фиксированная точка этого графика и $M(x, y)$ — любая другая его точка, где $x=x_0+\Delta x$, $y=y_0+\Delta y=f(x_0+\Delta x)$. Из треугольника $M_0MN$ находим $\operatorname{tg}\varphi=\frac{MN}{M_0N}=\frac{\Delta y}{\Delta x}$, где $\varphi$ — угол, образуемый секущей $M_0M$ с осью $Ox$.

Пусть $M \to M_0$ вдоль графика $y=f(x)$, $M_0T$ — касательная к нему в точке $M_0$, $\alpha$ — угол наклона касательной к оси $Ox$, тогда $\Delta x \to 0$, $\operatorname{tg}\varphi \to \operatorname{tg}\alpha$. Следовательно,

$$\operatorname{tg}\alpha = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}. \tag{11.1}$$

Таким образом, задача о касательной приводит к необходимости рассмотрения предела отношения приращения функции к приращению аргумента, когда последнее стремится к нулю.

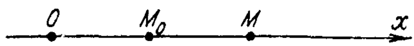



Рис. 11.2

**Задача о скорости неравномерного прямолинейного движения.** Рассмотрим неравномерное прямолинейное движение точки $M$, осуществляемое по закону $x=f(t)$, где $x=OM$ — длина пути (рис. 11.2); $f(t)$ — заданная функция времени $t$. Пусть при $t=t_0$ движущаяся точка совпадает с точкой $M_0(x_0)$, где $x_0=f(t_0)$, и в момент времени $t$ находится в точке $M$. Приращение пути $\Delta x$ за промежуток времени $\Delta t=t-t_0$ выразится формулой $\Delta x = f(t_0+\Delta t)-f(t_0)$.

Средней скоростью $v_{\text{ср}}$ движения за промежуток времени $\Delta t$ называется частное

$$v_{\text{ср}} = \frac{\Delta x}{\Delta t} = \frac{f(t_0+\Delta t)-f(t_0)}{\Delta t}. \tag{11.2}$$

Скоростью движения в момент $t_0$ называется предел отношения (11.2), когда $\Delta t \to 0$, т. е.

$$v = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(t_0+\Delta t)-f(t_0)}{\Delta t}. \tag{11.3}$$

Данная задача также сводится к нахождению аналогичного предела.

**Задача о скорости химической реакции.** Обозначим через $Q$ количество вещества, вступившего в химическую реакцию к моменту времени $t$, очевидно, $Q=Q(t)$. За промежуток времени $\Delta t$ это количество изменится на $\Delta Q$.

Средней скоростью химической реакции за промежуток времени $\Delta t$ называется отношение

$$v_{\text{ср}} = \frac{\Delta Q}{\Delta t} = \frac{Q(t+\Delta t)-Q(t)}{\Delta t},$$

а скоростью реакции в момент времени $t$ — предел указанного отношения при $\Delta t \to 0$

$$v = \lim_{\Delta t \to 0} \frac{\Delta Q}{\Delta t} = \lim_{\Delta t \to 0} \frac{Q(t+\Delta t) - Q(t)}{\Delta t}. \tag{11.4}$$

Итак, три различные задачи привели к необходимости рассмотрения предела отношения приращения функции к приращению аргумента, когда последнее стремится к нулю. Такой предел называется производной данной функции в данной точке.

## § 11.2. Понятие производной, ее геометрический и физический смысл

Рассмотрим функцию $y=f(x)$, заданную в интервале $]a, b[$; пусть $x_0 \in ]a, b[$ и $x \in ]a, b[$, тогда приращение функции в точке $x_0$ выразится формулой $\Delta y = f(x_0+\Delta x) - f(x_0)$.

*Производной функции* $y=f(x)$ *в точке* $x_0$ называется предел отношения приращения этой функции к приращению аргумента, когда последнее стремится к нулю.

Производную функции $y=f(x)$ в точке $x_0$ обозначают символом $f'(x_0)$ (читается: эф штрих от $x_0$) или $y'(x_0)$. Следовательно, по определению

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x} \tag{11.5}$$

или

$$y'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \tag{11.6}$$

Из формул (11.1) и (11.5) следует, что

$$f'(x_0) = \operatorname{tg} \alpha. \tag{11.7}$$

Формула (11.7) выражает *геометрический смысл производной:* производная от данной функции в данной точке равна тангенсу угла между осью $Ox$ и касательной к графику этой функции в соответствующей точке (рис. 11.1).

*Уравнение касательной* к линии $y=f(x)$ в точке $M_0(x_0, y_0)$ принимает вид (см. (2.8))

$$y - y_0 = f'(x_0)(x - x_0). \tag{11.8}$$

*Нормалью* к кривой в некоторой ее точке называется перпендикуляр к касательной в той же точке. Если

$f'(x_0)\neq 0$, то уравнение нормали к линии $y=f(x)$ в точке $M_0(x_0, y_0)$ записывается так:

$$y-y_0=-\frac{1}{f'(x_0)}(x-x_0) \tag{11.9}$$

(см. уравнение (2.8) и условие (2.21)).

Из формулы (11.3) и формулы, аналогичной формуле (11.6), следует, что

$$x'(t_0)=v\ \left(x'(t_0)=\lim_{\Delta t\to 0}\frac{f(t_0+\Delta t)-f(t_0)}{\Delta t}\right). \tag{11.10}$$

Формула (11.10) выражает физический смысл производной: производная от пути по времени равна скорости прямолинейного движения точки.

Если в формуле (11.5) от точки $x_0$ перейти к другой точке $x$, получим другой предел, поэтому $f'(x)$ — некоторая функция аргумента $x$:

$$f'(x)=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}. \tag{11.11}$$

Разумеется, не для всякой функции и не для любой точки $x$ предел в правой части формулы (11.11) существует. Например, для функции $y=|x|$ в точке $x=0$ указанный предел не существует, поскольку $\frac{\Delta y}{\Delta x}=1$ при $\Delta x>0$ и $\frac{\Delta y}{\Delta x}=-1$ при $\Delta x<0$.

Замечание. Употребляются и другие символы для обозначения производной функции $y=f(x)$ или $x=f(t)$:

$$\frac{dy}{dx}=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x},\quad \dot{x}=\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t}.$$

Если отношение $\Delta y:\Delta x$ при $x\to x_0$ имеет предел справа (или слева), он называется *производной справа* (соответственно *производной слева*). Такие пределы называются *односторонними производными*.

Односторонние производные функции $f(x)$ в точке $x_0$ обозначаются соответственно символами $f'_-(x_0)$, $f'_+(x_0)$:

$$f'_-(x_0)=\lim_{\Delta x\to -0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\text{ — производная слева;}$$

$$f'_+(x_0)=\lim_{\Delta x\to +0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\text{ — производная справа.}$$

Очевидно, функция $f(x)$, определенная в некоторой окрестности точки $x_0$, имеет производную $f'(x_0)$ тогда и только тогда, когда односторонние производные $f'_-(x_0)$, $f'_+(x_0)$ существуют и равны между собой, причем $f'(x_0) = f'_-(x_0) = f'_+(x_0)$.

Если для некоторого значения $x$ выполняется одно из условий

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = +\infty, \quad \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\infty,$$

то говорят, что в точке $x$ существует бесконечная производная, равная соответственно $+\infty$, $-\infty$.

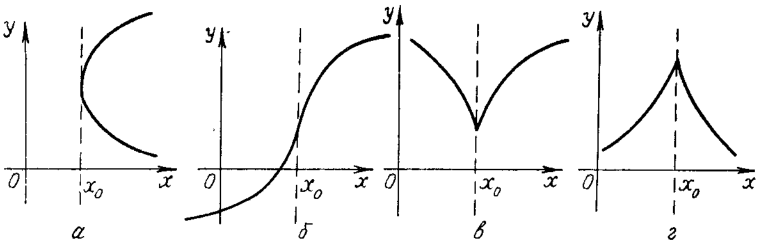



Рис. 11.3

Геометрическое истолкование производной как углового коэффициента касательной к графику распространяется и на этот случай: касательная параллельна оси $Oy$ (рис. 11.3, *а*, *б*).

Аналогично устанавливается понятие об односторонней бесконечной производной. На рис. 11.3, *в*, *г* изображены графики функций, имеющие острие, направленное вниз или вверх; односторонние бесконечные производные различны по знаку, но существует единственная вертикальная касательная.

В дальнейшем под выражением «функция имеет производную» будем понимать наличие конечной производной, если не оговорено противное.

Функция, имеющая производную в данной точке, называется *дифференцируемой* в этой точке. Функция, имеющая производную в каждой точке данного промежутка, называется дифференцируемой в этом промежутке. Если промежуток является замкнутым, на концах его имеются в виду односторонние производные.

Операция нахождения производной называется *дифференцированием*.

Зависимость между непрерывностью и дифференцируемостью функции выражается следующей теоремой.

**Теорема 11.1.** Если функция $y=f(x)$ дифференцируема в данной точке, то она и непрерывна в ней.

Доказательство. Пусть функция $y=f(x)$ дифференцируема в точке $x$, т. е. существует предел $y'(x) = \lim\limits_{\Delta x\to 0} \frac{\Delta y}{\Delta x}$. Так как $\Delta y = \frac{\Delta y}{\Delta x}\cdot\Delta x$, то $\lim\limits_{\Delta x\to 0}\Delta y = \lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x} \times \lim\limits_{\Delta x\to 0}\Delta x = y'(x)\cdot 0 = 0$, $\lim\limits_{\Delta x\to 0}\Delta y = 0$.

Замечание. Обратное утверждение не всегда верно. Например, функция $y=|x-1|$ непрерывна в точке $x=1$, но не является дифференцируемой в ней.

## § 11.3. Производные некоторых функций

Из определения производной следует схема ее нахождения: 1) фиксируется значение $x$ аргумента функции; 2) в точке $x$ аргументу придается приращение $\Delta x$; 3) вычисляется приращение $\Delta y$, соответствующее приращению $\Delta x$; 4) составляется отношение приращения функции к приращению аргумента; 5) находится предел указанного отношения, когда приращение аргумента стремится к нулю.

Рассмотрим некоторые примеры нахождения производных.

Пример 1. Найти производную функции $y=c$, где $c=\text{const}$.

Данная функция является постоянной, поэтому $y+\Delta y = c$, $\Delta y = 0$, $\frac{\Delta y}{\Delta x}=0$, $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=0$,

$$c' = 0, \tag{11.12}$$

т. е. производная постоянной равна нулю.

Пример 2. Найти производную функции $y=\sin x$.

Фиксируем некоторое значение $x$ аргумента функции, придадим ему приращение $\Delta x$, найдем выражение для соответствующего приращения функции:

$$\Delta y = \sin(x+\Delta x) - \sin x = 2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right).$$

Составим отношение приращения функции к приращению аргумента:

$$\frac{\Delta y}{\Delta x}=\frac{2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right)}{\Delta x}=$$

$$=\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\cos\left(x+\frac{\Delta x}{2}\right).$$

Переходим к пределу, принимая во внимание формулу (10.24) и непрерывность функции $f(x)=\cos x$: $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=1\cdot\cos x=\cos x$, т. е.

$$\boxed{(\sin x)'=\cos x.} \tag{11.13}$$

Пример 3. Найти производную функции $y=x^{\alpha}$, где $\alpha$ — действительное число.

Область определения степенной функции $y=x^{\alpha}$ зависит от $\alpha$.

Составим отношение приращения функции к приращению аргумента:

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

и перейдем к пределу при $\Delta x\to 0$, воспользовавшись формулой (10.50):

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=x^{\alpha-1}\lim_{\Delta x\to 0}\frac{\left(1+\frac{\Delta x}{x}\right)^{\alpha}-1}{\frac{\Delta x}{x}}=x^{\alpha-1}\cdot\alpha=\alpha x^{\alpha-1}.$$

Следовательно,

$$\boxed{(x^{\alpha})'=\alpha x^{\alpha-1}.} \tag{11.14}$$

Отметим частные случаи этой формулы:

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

## § 11.4. Основные правила дифференцирования

### Производная алгебраической суммы функций.

**Теорема 11.2.** Производная суммы (разности) двух дифференцируемых функций равна сумме (разности) производных этих функций.

Доказательство. Пусть $y=u(x)\pm v(x)$, где $u=u(x)$ и $v=v(x)$ — дифференцируемые функции. Поскольку $\Delta y=\Delta u\pm\Delta v$, то

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} \pm \frac{\Delta v}{\Delta x}, \quad \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} \pm$$

$$\pm \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta x} = u' \pm v'.$$

Следовательно,

$$\boxed{(u \pm v)' = u' \pm v'.} \qquad (11.15)$$

Следствие. Производная конечной алгебраической суммы дифференцируемых функций равна такой же алгебраической сумме производных слагаемых. Например,

$$\left(x^4 + \sqrt{x} - \frac{1}{x} + 5\right)' = 4x^3 + \frac{1}{2\sqrt{x}} + \frac{1}{x^2}.$$

**Производная произведения функций.**

**Теорема 11.3.** Производная произведения двух дифференцируемых функций равна произведению первой функции на производную второй плюс произведение второй функции на производную первой, т. е.

$$\boxed{(uv)' = u'v + uv'.} \qquad (11.16)$$

Доказательство. Пусть $y = u \cdot v$, где $u = u(x)$, $v = v(x)$ — дифференцируемые функции. Так как $\Delta y = (u + \Delta u)(v + \Delta v) - uv$, $\Delta y = u\Delta v + v\Delta u + \Delta u \Delta v$, $\frac{\Delta y}{\Delta x} = u\frac{\Delta v}{\Delta x} + v\frac{\Delta u}{\Delta x} + \frac{\Delta u}{\Delta x}\Delta v$ и $\lim\limits_{\Delta x \to 0} \Delta v = 0$ на основании теоремы 11.1, то $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = uv' + vu'$, откуда и следует формула (11.16). Например, $(x^8 \sin x)' = (x^8)' \sin x + x^8 (\sin x)' = 8x^7 \sin x + x^8 \cos x$.

Следствие 1. Постоянный множитель можно выносить за знак производной:

$$\boxed{(cv)' = cv'.} \qquad (11.17)$$

Например, $(4x^3)' = 4(x^3)' = 4 \cdot 3x^2 = 12x^2$.

Следствие 2. Производная произведения нескольких дифференцируемых функций равна сумме произведений производной каждого из них на все остальные. Например,

$$\boxed{(uvw)' = u'vw + uv'w + uvw'.} \qquad (11.18)$$

**Производная частного двух функций.**

**Теорема 11.4.** Производная частного двух дифференцируемых функций определяется формулой

$$\boxed{\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}} \quad (v \neq 0). \tag{11.19}$$

Доказательство. Если $y = \frac{u}{v}$, где $u = u(x)$, $v = v(x)$ — дифференцируемые функции, причем $v(x) \neq 0$, то

$$\Delta y = \frac{u + \Delta u}{v + \Delta v} - \frac{u}{v} = \frac{v\Delta u - u\Delta v}{v(v + \Delta v)}, \quad \frac{\Delta y}{\Delta x} = \frac{v\frac{\Delta u}{\Delta x} - u\frac{\Delta v}{\Delta x}}{v(v + \Delta v)}.$$

Следовательно,

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{u'v - uv'}{v^2},$$

так как $\lim\limits_{\Delta x \to 0} \Delta v = 0$ на основании теоремы 11.1. Например,

$$\left(\frac{x-1}{x^2+4}\right)' = \frac{(x-1)'(x^2+4) - (x^2+4)'(x-1)}{(x^2+4)^2} = \frac{1 \cdot (x^2+4) - 2x(x-1)}{(x^2+4)^2} = \frac{4 + 2x - x^2}{(x^2+4)^2}.$$

**Производная обратной функции.** Если $y = f(x)$ и $x = \varphi(y)$ — взаимно-обратные дифференцируемые функции и $y'_x \neq 0$, то

$$\boxed{x'_y = \frac{1}{y'_x}.} \tag{11.20}$$

Действительно, так как $\frac{\Delta x}{\Delta y} = \frac{1}{\frac{\Delta y}{\Delta x}}$, то

$$\lim_{\Delta y \to 0} \frac{\Delta x}{\Delta y} = \frac{1}{\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}},$$

откуда и следует равенство (11.20).

## § 11.5. Производная сложной функции

Рассмотрим сложную функцию $y=f(\varphi(x))\equiv F(x)$, где $y=f(u)$, $u=\varphi(x)$; в этом случае $u$ называют *промежуточным аргументом, $x$ — независимой переменной.*

**Теорема 11.5.** Если $y=f(u)$ и $u=\varphi(x)$ — дифференцируемые функции своих аргументов, то производная сложной функции $y=f(\varphi(x))$ существует и равна произведению производной этой функции по промежуточному аргументу на производную промежуточного аргумента по независимой переменной, т. е.

$$y'_x = y'_u \cdot u'_x, \quad \frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}. \tag{11.21}$$

Доказательство. В соответствии с условием и по определению производной

$$y'_u = \lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u}, \quad u'_x = \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x}.$$

Так как $\dfrac{\Delta y}{\Delta x} = \dfrac{\Delta y}{\Delta u}\dfrac{\Delta u}{\Delta x}$, то $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x} = \lim\limits_{\Delta u \to 0} \dfrac{\Delta y}{\Delta u} \times$ $\times \lim\limits_{\Delta x \to 0} \dfrac{\Delta u}{\Delta x}$, $y'_x = y'_u \cdot u'_x$ или $\dfrac{dy}{dx} = \dfrac{dy}{du}\dfrac{du}{dx}$.

Пример 1. Найти производную функции $y=\sin(3x-5)$.

Данную функцию можно представить так: $y=\sin u$, $u=3x-5$. Поскольку $y'_u = \cos u = \cos(3x-5)$, $u'_x = 3$, по формуле (11.21) получаем $(\sin(3x-5))'_x = \cos(3x-5)\cdot 3 = 3\cos(3x-5)$.

Пример 2. Найти производную функции $y=(3-2x)^5$.

Так как $y=u^5$, $u=3-2x$, $y'_u = 5u^4 = 5(3-2x)^4$, $u'_x = -2$, то $((3-2x)^5)'_x = -10(3-2x)^4$.

## § 11.6. Основные формулы дифференцирования

Выведем формулы для производных логарифмической, показательной, тригонометрических, обратных тригонометрических и гиперболических функций. При выводе этих формул будем пользоваться определением производной.

**Производная логарифмической функции.** Рассмотрим функцию $y=\log_a x$. Поскольку

$$\Delta y = \log_a(x+\Delta x) - \log_a x = \log_a\left(\frac{x+\Delta x}{x}\right) =$$
$$= \log_a\left(1 + \frac{\Delta x}{x}\right),$$

то

$$\frac{\Delta y}{\Delta x}=\frac{\log_a\left(1+\frac{\Delta x}{x}\right)}{\Delta x}=\frac{1}{x}\frac{\log_a\left(1+\frac{\Delta x}{x}\right)}{\frac{\Delta x}{x}}.$$

Принимая во внимание (10.48), получаем

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\frac{1}{x}\cdot\log_a e=\frac{1}{x\ln a}.$$

Следовательно,

$$(\log_a x)'=\frac{1}{x\ln a}. \tag{11.22}$$

В частном случае, когда $a=e$, получаем

$$(\ln x)'=\frac{1}{x}. \tag{11.23}$$

Найдем производную функции $y=\ln|x|$. Как уже отмечалось (см. § 10.15), эта функция определена при всех $x$, кроме $x=0$. Принимая во внимание определение модуля действительного числа, данную функцию можно записать с помощью двух равенств: $y=\ln x$ при $x>0$, $y=\ln(-x)$ при $x<0$.

На основании формул (11.21), (11.23) соответственно получаем $y'=\frac{1}{x}$ при $x>0$, $y'=\frac{1}{-x}(-x)'=-\frac{1}{x}(-1)=\frac{1}{x}$ при $x<0$.

Следовательно, $y'=\frac{1}{x}$ при $x\neq 0$, т. е.

$$(\ln|x|)'=\frac{1}{x}. \tag{11.24}$$

Если $u=f(x)$, где $f(x)$ — дифференцируемая функция, причем $f(x)>0$, то по формулам (11.21) и (11.23)

$$(\ln u)'_x=\frac{1}{u}\cdot u'_x=\frac{u'_x}{u}$$

или

$$(\ln f(x))'=\frac{f'(x)}{f(x)}. \tag{11.25}$$

Например, $(\ln(x^2+1))'=\frac{(x^2+1)'}{x^2+1}=\frac{2x}{x^2+1}$.

Производная, определяемая формулой (11.25), называется логарифмической производной функции $f(x)$.

**Производная показательной функции.** Дана функция $y=a^x$. Логарифмируя это равенство по основанию $e$, получаем $\ln y = x \ln a$, откуда по формуле (11.25) $\frac{y'}{y} = \ln a$ или $y' = y \ln a$.

Следовательно,

$$(a^x)' = a^x \ln a. \tag{11.26}$$

В частном случае, когда $a=e$, получаем

$$(e^x)' = e^x. \tag{11.27}$$

Если $u=u(x)$ — дифференцируемая функция аргумента $x$, то по формулам (11.21), (11.26), (11.27) соответственно находим

$$(a^u)'_x = a^u \ln a \cdot u'_x; \tag{11.28}$$

$$(e^u)'_x = e^u \cdot u'_x. \tag{11.29}$$

Например, $(e^{\sin x})' = e^{\sin x} \cdot \cos x$.

**Производные тригонометрических функций.** Производная функции $y=\sin x$ уже была получена в § 11.3 (см. формулу (11.13)). Найдем производные других тригонометрических функций.

Для функции $y=\cos x$

$$\Delta y = \cos(x+\Delta x) - \cos x = -2 \sin \frac{\Delta x}{2} \sin\left(x + \frac{\Delta x}{2}\right),$$

$$\frac{\Delta y}{\Delta x} = -2 \frac{\sin \frac{\Delta x}{2}}{\Delta x} \sin\left(x+\frac{\Delta x}{2}\right) =$$

$$= -\frac{\sin \frac{\Delta x}{2}}{\frac{\Delta x}{2}} \sin\left(x+\frac{\Delta x}{2}\right),$$

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\lim_{\Delta x \to 0} \frac{\sin \frac{\Delta x}{2}}{\frac{\Delta x}{2}} \cdot \lim_{\Delta x \to 0} \sin\left(x+\frac{\Delta x}{2}\right) = -1 \cdot \sin x,$$

т. е.

$$\boxed{(\cos x)' = -\sin x.} \tag{11.30}$$

Применяя формулы (11.13), (11.19), (11.30), находим

$$(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{(\sin x)' \cos x - \sin x (\cos x)'}{\cos^2 x} =$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}, \tag{11.31}$$

$$\boxed{(\operatorname{tg} x)' = \frac{1}{\cos^2 x}.}$$

Аналогично получаем

$$\boxed{(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}.} \tag{11.32}$$

**Производные обратных тригонометрических функций.** Пусть $y = \arcsin x$, т. е. $x = \sin y$, причем $-\frac{\pi}{2} \leqslant y \leqslant \leqslant \frac{\pi}{2}$, $-1 \leqslant x \leqslant 1$. Поскольку $x'_y = \cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2}$, по формуле (11.20) имеем

$$y'_x = \frac{1}{x'_y} = \frac{1}{\sqrt{1 - x^2}},$$

т. е.

$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}. \tag{11.33}$$

Аналогично находим

$$(\arccos x)' = -\frac{1}{\sqrt{1 - x^2}}. \tag{11.34}$$

Если $y = \operatorname{arctg} x$, т. е. $x = \operatorname{tg} y$, то $x'_y = \frac{1}{\cos^2 y} = 1 + \operatorname{tg}^2 y = 1 + x^2$, $y'_x = \frac{1}{x'_y} = \frac{1}{1 + x^2}$, поэтому

$$(\operatorname{arctg} x)' = \frac{1}{1 + x^2}. \tag{11.35}$$

Аналогично получаем

$$(\operatorname{arcctg} x)' = -\frac{1}{1 + x^2}. \tag{11.36}$$

**Производные гиперболических функций.** Так как $\operatorname{sh} x = \frac{e^x - e^{-x}}{2}$, $(e^{-x})' = -e^{-x}$, $(e^x)' = e^x$, то

$$(\operatorname{sh} x)' = \frac{e^x + e^{-x}}{2} = \operatorname{ch} x, \quad (\operatorname{sh} x)' = \operatorname{ch} x. \tag{11.37}$$

Аналогично находим

$$(\operatorname{ch} x)' = \operatorname{sh} x. \tag{11.38}$$

Поскольку $\operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x}$, то

$$(\operatorname{th} x)' = \frac{(\operatorname{sh} x)' \operatorname{ch} x - (\operatorname{ch} x)' \operatorname{sh} x}{\operatorname{ch}^2 x} = \frac{\operatorname{ch}^2 x - \operatorname{sh}^2 x}{\operatorname{ch}^2 x} = \frac{1}{\operatorname{ch}^2 x},$$

$$(\operatorname{th} x)' = \frac{1}{\operatorname{ch}^2 x}. \tag{11.39}$$

Аналогично получаем

$$(\operatorname{cth} x)' = -\frac{1}{\operatorname{sh}^2 x} \tag{11.40}$$

## § 11.7. Производные высших порядков

В дальнейшем производную $y' = f'(x)$ функции $y = f(x)$ будем называть первой производной этой функции.

*Второй производной* (или *производной второго порядка*) функции $y = f(x)$ называется производная от ее первой производной. Обозначения второй производной функции $y = f(x)$:

$$y'' = (y')', \ f''(x) = (f'(x))', \ \frac{d^2y}{dx^2} = \frac{df'(x)}{dx}.$$

Например, если $f(x) = x^4 - 3x^2 + 5$, то $f'(x) = 4x^3 - 6x$, $f''(x) = 12x^2 - 6 = 6(2x^2 - 1)$.

Выясним физический смысл второй производной. Пусть закон движения точки $M$ по оси $Ox$ выражается уравнением $x = f(t)$, в момент времени $t$ точка $M$ имеет скорость $v$, а в момент $t + \Delta t$ — скорость $v + \Delta v$, т. е. за промежуток времени $\Delta t$ скорость изменилась на $\Delta v$. Средним ускорением этого движения за промежуток времени $\Delta t$ называется отношение $w_{\text{ср}} = \frac{\Delta v}{\Delta t}$.

Ускорением точки $M$ в момент времени $t$ называется предел указанного отношения при $\Delta t \to 0$, т. е. $w = \lim\limits_{\Delta t \to 0} \frac{\Delta v}{\Delta t}$, $w = v'_t$. Поскольку $v = f'(t)$, то $v'_t = (f'(t))' = f''(t)$, т. е. $w = f''(t)$. Следовательно, ускорение прямолинейного движения точки равно второй производной от пути по времени (точнее, от абсциссы движущейся точки).

Аналогично определяется третья, четвертая производная. Вообще, $n$-й производной функции $y=f(x)$ называется производная от ее $(n-1)$-й производной:

$$y^{(n)}=(y^{(n-1)})', \quad f^{(n)}(x)=(f^{(n-1)}(x))'.$$

Первые три производные обозначаются штрихами, последующие — римскими цифрами. Например, если $y=\sin x$, то $y'=\cos x$, $y''=-\sin x$, $y'''=-\cos x$, $y^{\mathrm{IV}}=\sin x$, $y^{\mathrm{V}}=\cos x$ и т. д.

Замечание. Для $n$-й производной от функции $y=f(x)$ употребляется также следующее обозначение:

$$\frac{d^n y}{dx^n}=\frac{d}{dx}\left(\frac{d^{n-1}y}{dx^{n-1}}\right)$$

(левая часть читается так: дэ эн игрек по дэ икс эн). В частности, при $n=2$ получаем

$$\frac{d^2 y}{dx^2}=\frac{d}{dx}\left(\frac{dy}{dx}\right)$$

(читается: дэ два игрек по дэ икс квадрат).

Пример. Найти шестую производную функции $y=x^7+x^5+x^3-x+\cos x$.

Применяя правило дифференцирования алгебраической суммы функций, формулы (11.14), (11.30) и определение $n$-й производной ($n=1, 2, \ldots, 6$), последовательно получаем $\frac{dy}{dx}=7x^6+5x^4+3x^2-1-\sin x$, $\frac{d^2y}{dx^2}=42x^5+20x^3+6x-\cos x$, $\frac{d^5y}{dx^5}=2520x^2+120-\sin x$, $\frac{d^6y}{dx^6}=5040x-\cos x$.

## § 11.8. Понятие дифференциала, его геометрический и механический смысл

Рассмотрим функцию $y=f(x)$, имеющую производную в каждой точке ее области определения.

*Дифференциалом* функции $y=f(x)$ называется произведение производной этой функции на приращение независимой переменной $x$:

$$dy=y'\Delta x, \quad df(x)=f'(x)\Delta x. \tag{11.41}$$

Для функции $y=x$ получаем $dy=dx$, $dx=x'_x\Delta x=1\cdot\Delta x$,

$$dx=\Delta x. \tag{11.42}$$

Следовательно, дифференциал независимой переменной равен приращению этой переменной.

Из формул (11.41), (11.42) следует, что

$$\qquad (11.43)$$

т. е. дифференциал функции равен произведению ее производной на дифференциал независимой переменной.

Из формулы (11.43) находим, что

$$y' = \frac{dy}{dx} \text{ или } f'(x) = \frac{df(x)}{dx}. \qquad (11.44)$$

Это означает, что производная функции равна отношению дифференциала данной функции к дифференциалу ее аргумента; правую часть формулы (11.44) можно рассматривать и как частное (раньше это был символ для обозначения производной, см. § 11.2).

Так как по определению $y' = \lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$, то

$$\frac{\Delta y}{\Delta x} = y' + \alpha, \qquad (11.45)$$

где

$$\alpha \to 0 \text{ при } \Delta x \to 0. \qquad (11.46)$$

Умножая обе части равенства (11.45) на $\Delta x$, получаем

$$\Delta y = y'\Delta x + \alpha\Delta x \qquad (11.47)$$

или

$$\Delta y = dy + \alpha\Delta x. \qquad (11.48)$$

Поскольку $\frac{\Delta y - dy}{\Delta x} = \alpha$ и выполнено условие (11.46), то дифференциал функции есть часть приращения данной функции, отличающаяся от всего приращения на бесконечно малую величину высшего порядка по сравнению с $\Delta x$ (см. § 10.18). В соответствии с этим говорят, что дифференциал функции является главной линейной частью ее приращения (главной, так как оставшаяся часть — бесконечно малая по сравнению с $\Delta x$, линейной, поскольку она линейно выражается через $\Delta x$).

Замечание. Формулу (11.48) можно записать в виде $\Delta y = dy + o(\Delta x)$, где символ $o(\Delta x)$ означает, что данное слагаемое есть бесконечно малая высшего порядка по сравнению с $\Delta x$ (см. § 10.18).

Покажем на примере, чем отличается дифференциал от приращения функции $y=x^2$, выражающей площадь квадрата со стороной $x$ (рис. 11.4). Когда аргумент получает приращение $\Delta x$, функция получает приращение $\Delta y = (x+\Delta x)^2 - x^2 = x^2 + 2x\Delta x + \Delta x^2 - x^2$, $\Delta y = 2x\Delta x + \Delta x^2$, состоящее из двух слагаемых: первое из них линейно зависит от $\Delta x$ и является дифференциалом.

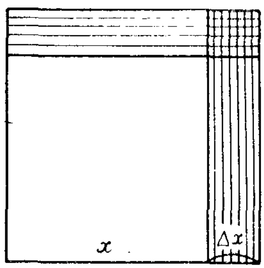



Рис. 11.4

Обратимся к графику функции $y=f(x)$. Пусть $M(x, y)$ — произвольная точка его (рис. 11.5), $M'(x+\Delta x, y+\Delta y)$ — другая точка, координаты которой равны приращенным значениям $x$ и $y$. Построим касательную к графику в точке $M$, обозначим буквой $T$ точку пересечения касательной с прямой, проведенной через точку $M'$ перпендикулярно оси $Ox$, буквой $N$ — точку пересечения указанной прямой с прямой, параллельной оси $Ox$, тогда $NM' = \Delta y$, $NT = \operatorname{tg}\alpha \cdot \Delta x = f'(x)\Delta x = dy$.

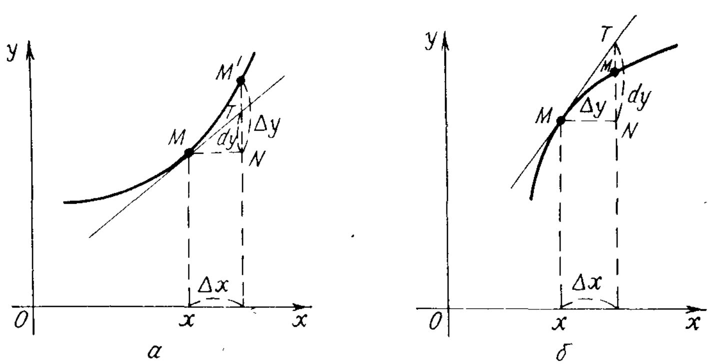



Рис. 11.5

Следовательно, дифференциал функции равен приращению ординаты касательной к графику данной функции, когда аргумент получает приращение $\Delta x$. Это утверждение выражает геометрический смысл дифференциала.

Как видно из рисунка, $dy<\Delta y$ (рис. 11.5, *а*) или $dy>$ $>\Delta y$ (рис. 11.5, *б*); если функция равна постоянной, то $dy=\Delta y=0$.

Выясним механический смысл дифференциала. Рассмотрим неравномерное прямолинейное движение точки, осуществляющееся по закону $s=f(t)$, где $s$ — длина пути, $t$ — время. Приращенному моменту времени $t+\Delta t$ соответствует приращенное значение пути $s+\Delta s=f(t+\Delta t)$, откуда $\Delta s=f(t+\Delta t)-f(t)$. Эта формула выражает истинное приращение пути за промежуток времени $\Delta t$.

Вычислим дифференциал пути. Так как $s'=f'(t)=$ $=v(t)$ — скорость в момент $t$, то $ds=s'(t)dt=f'(t)dt=$ $=v(t)dt$, $ds=v(t)dt$.

Следовательно, дифференциал пути равен приращению пути, полученному в предположении, что, начиная с данного момента времени $t$, точка движется равномерно, сохраняя приобретенную скорость.

Пример. Дана функция $y=x^3-2x+5$. Найти: 1) выражение для дифференциала, соответствующее аргументу $x$ и приращению $\Delta x$; 2) $dy$ и $\Delta y$ при переходе от точки $x_1=1$ к точке $x_2=1{,}1$.

1) По формуле (11.43) получаем $dy=(x^3-2x+5)'dx$, $dy=$ $=(3x^2-2)dx$.

2) Так как $x=1$, $\Delta x=1{,}1-1=0{,}1$, $dx=0{,}1$, то $dy=(3-2)0{,}1=0{,}1$; $\Delta y=y(1{,}1)-y(1)=4{,}131-4=0{,}131$.

Замечание 2. Дифференциал функции можно определить и по-другому. Если приращение функции $y=f(x)$ в точке $x$ представимо формулой

$$\Delta y=A\Delta x+o(\Delta x), \tag{11.49}$$

где $A$ не зависит от $\Delta x$, $o(\Delta x)$ — бесконечно малая высшего порядка по сравнению с $\Delta x$, то слагаемое $A\Delta x$ называют дифференциалом данной функции в данной точке, а функцию $y=f(x)$ в этом случае называют дифференцируемой в точке $x$. Обозначив дифференциал функции $y=f(x)$ через $dy$, по определению получим $dy=A\Delta x$.

**Теорема 11.6.** Если приращение функции $y=f(x)$ в точке $x$ представимо формулой (11.49), то $A=f'(x)$.

Доказательство. Разделим равенство (11.49) на $\Delta x\neq 0$, тогда $\dfrac{\Delta y}{\Delta x}=A+\dfrac{o(\Delta x)}{\Delta x}$. Переходя к пределу при $\Delta x\to 0$ и принимая во внимание, что $\lim\limits_{\Delta x\to 0}\dfrac{o(\Delta x)}{\Delta x}=$ $=0$, находим $\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}=A$, $f'(x)=A$. Следовательно, $dy=$ $=A\Delta x=f'(x)\Delta x$, $dy=f'(x)\Delta x$; получена формула (10.41), с помощью которой ранее определялся дифференциал функции. Поскольку $\Delta x=dx$, то $dy=f'(x)dx$.

## § 11.9. Применение дифференциала в приближенных вычислениях

**Теорема 11.7.** Бесконечно малое приращение функции эквивалентно дифференциалу этой функции при всех значениях независимой переменной, для которых производная функции конечна и отлична от нуля.

Доказательство. Поскольку функция $y=f(x)$ дифференцируема, для нее справедлива формула (11.48). Так как $y'\neq 0$, то при $\Delta x\neq 0$ $dy=y'dx\neq 0$, поэтому, разделив обе части равенства (11.48) на $dy$, получим $\frac{\Delta y}{dy}=1+\frac{\alpha}{y'}$. Переходя к пределу при $\Delta x\to 0$ и принимая во внимание (11.46), получаем $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{dy}=1$, а это означает, что $\Delta y$ и $dy$ эквивалентны при $\Delta x\to 0$.

От формулы (11.48) перейдем к приближенной формуле

$$\Delta y\approx dy,\ \Delta y\approx y'\Delta x, \tag{11.50}$$

т. е. $f(x+\Delta x)-f(x)\approx f'(x)\Delta x$, или

$$f(x+\Delta x)\approx f(x)+f'(x)\Delta x. \tag{11.51}$$

Формула (11.51) позволяет вычислить приближенное значение функции, соответствующее приращенному значению аргумента, если известно ее значение в некоторой точке и значение производной в этой точке, когда приращение аргумента является достаточно малым.

Выбирая в качестве $f(x)$ конкретные функции $f(x)=\sqrt[n]{x}$, $f(x)=\ln x$, $f(x)=\operatorname{arctg} x$ и т. д., получаем формулы

$$\sqrt[n]{x+\Delta x}\approx\sqrt[n]{x}+\frac{\Delta x}{n\sqrt[n]{x^{n-1}}}, \tag{11.51а}$$

$$\ln(x+\Delta x)\approx\ln x+\frac{\Delta x}{x}, \tag{11.51б}$$

$$\operatorname{arctg}(x+\Delta x)\approx\operatorname{arctg} x+\frac{\Delta x}{1+x^2} \tag{11.51в}$$

и т. д.

Пример. Вычислить приближенно $\sqrt[4]{1,15}$.

Пользуемся частным случаем формулы (11.51а), полагая в ней $n=4$, $x=1$, $\Delta x=0,15$:

$$\sqrt[4]{x+\Delta x}\approx\sqrt[4]{x}+\frac{\Delta x}{4\sqrt[4]{x^3}};\quad \sqrt[4]{1+0,15}\approx\sqrt[4]{1}+\frac{0,15}{4\sqrt[4]{1^3}}=$$
$$=1+\frac{0,15}{4}=1,0375.$$

Дифференциалы применяются также при оценке погрешностей. Пусть для данной функции $y=f(x)$ предельная абсолютная погрешность ее аргумента $x$ равна $\alpha_x$, т. е.

$$|\Delta x| \leqslant \alpha_x. \tag{11.52}$$

Найдем предельные погрешности функции $y$: абсолютную $\alpha_y$ и относительную $\delta_y$.

Из формулы (11.50) следует, что $|\Delta y| \approx |y'| |\Delta x|$, поэтому с учетом (11.52) $|\Delta y| \leqslant |y'| \alpha_x$. Следовательно, можно принять $\alpha_y = |y'| \alpha_x$, тогда

$$\delta_y = \frac{\alpha_y}{|y|} = |(\ln y)'| \alpha_x. \tag{11.53}$$

## § 11.10. Свойства дифференциала

Дифференциал функции обладает свойствами, аналогичными свойствам производной.

1. Дифференциал постоянной величины равен нулю:

$$dc=0, \quad c=\text{const}. \tag{11.54}$$

Если $y=c$, где $c=\text{const}$, то $y'=0$; формула (11.43) принимает вид (11.54).

2. Дифференциал алгебраической суммы нескольких дифференцируемых функций равен такой же алгебраической сумме дифференциалов слагаемых

$$d(u-v+w)=du-dv+dw. \tag{11.55}$$

В самом деле, $d(u-v+w)=(u-v+w)'dx=u'dx-v'dx+$ $+w'dx=du-dv+dw$.

Следствие. Если две дифференцируемые функции отличаются постоянным слагаемым, то их дифференциалы равны

$$d(u+c)=du. \tag{11.56}$$

Формула (11.56) следует из формул (11.55) и (11.54).

3. Дифференциал произведения двух дифференцируемых функций равен произведению первой функции на дифференциал второй плюс произведение второй на дифференциал первой, т. е.

$$d(uv)=udv+vdu. \tag{11.57}$$

Действительно, $d(uv)=(uv)'dx=(uv'+u'v)dx=$ $=u(v'dx)+v(u'dx)=udv+vdu$; $d(uv)=udv+vdu$.

Следствие. Постоянный множитель можно выносить за знак дифференциала:

$$d(cu)=cdu \quad (c=\text{const}). \tag{11.58}$$

Формула (11.58) следует из формул (11.57) и (11.54).

4. Дифференциал частного $\frac{u}{v}$ $(v\neq 0)$ двух дифференцируемых функций $u=u(x)$ и $v=v(x)$ определяется формулой

$$d\left(\frac{u}{v}\right)=\frac{vdu-udv}{v^2}. \tag{11.59}$$

В самом деле, $d\left(\frac{u}{v}\right)=\left(\frac{u}{v}\right)'dx=\frac{u'v-uv'}{v^2}dx=$ $=\frac{v(u'dx)-u(v'dx)}{v^2}=\frac{vdu-udv}{v^2}$.

5. Дифференциал сложной функции (функции от функции) равен произведению производной данной функции по промежуточному аргументу на дифференциал этого промежуточного аргумента.

Если $y=f(x)$, $x=\varphi(t)$ — дифференцируемые функции своих аргументов, то производная функции $y=f(\varphi(t))$ выражается формулой $y'_t=y'_x\cdot x'_t$. По определению $dy=$ $=y'_tdt$. С учетом предыдущей формулы получаем $dy=$ $=y'_tdt=y'_x\cdot x'_tdt=y'_x(x'_tdt)=y'_xdx$,

$$dy=y'_xdx. \tag{11.60}$$

Замечание. Формулы (11.43) и (11.60) имеют один и тот же вид, но между ними есть принципиальное различие: в формуле (11.43) $dx=\Delta x$, поскольку $x$ — независимая переменная; в формуле (11.60), вообще говоря, $dx\neq\Delta x$, так как $x$ — функция независимой переменной $t$.

Из формул (11.43) и (11.60) следует свойство независимости вида дифференциала от выбора независимой переменной: дифференциал функции равен произведению производной на дифференциал аргумента, независимо от того, является ли этот аргумент независимой переменной или функцией другой независимой переменной.

Принимая во внимание основные формулы дифференцирования, получаем таблицу для дифференциалов:

I. $dx^\alpha=\alpha x^{\alpha-1}dx$. II. $da^x=a^x\ln a dx$.

III. $d(\log_a x) = \frac{dx}{x \ln a}$, $\qquad d(\ln|x|) = \frac{dx}{x}$.

IV. $d(\sin x) = \cos x\, dx$.

V. $d(\cos x) = -\sin x dx$. $\qquad$ VI. $d(\operatorname{tg} x) = \frac{dx}{\cos^2 x}$.

VII. $d(\operatorname{ctg} x) = -\frac{dx}{\sin^2 x}$. $\qquad$ VIII. $d(\arcsin x) = \frac{dx}{\sqrt{1-x^2}}$.

IX. $d(\arccos x) = -\frac{dx}{\sqrt{1-x^2}}$. $\qquad$ X. $d(\operatorname{arctg} x) = \frac{dx}{1+x^2}$.

XI. $d(\operatorname{arcctg} x) = -\frac{dx}{1+x^2}$. $\qquad$ XII. $d(\operatorname{sh} x) = \operatorname{ch} x dx$.

XIII. $d(\operatorname{ch} x) = \operatorname{sh} x dx$. $\qquad$ XIV. $d(\operatorname{th} x) = \frac{dx}{\operatorname{ch}^2 x}$.

XV. $d(\operatorname{cth} x) = -\frac{dx}{\operatorname{sh}^2 x}$. $\qquad$ XVI. $df(u) = f'(u)du$.

## § 11.11. Дифференциалы высших порядков

Если $x$ — независимая переменная и $y=f(x)$ — дифференцируемая функция, то $df(x)=f'(x)dx$, т. е. дифференциал функции есть функция, зависящая от двух аргументов $x$ и $dx$. Этот дифференциал будем называть также *дифференциалом первого порядка* (или *первым дифференциалом*). Считая $dx$ постоянной, получаем, что $df(x)$ — функция одной переменной. Предположим, что функция $y=f(x)$ имеет не только первую производную, но и $n-1$ последовательных производных в некотором интервале $]a, b[$:

$$y''=f''(x),\quad y'''=f'''(x),\quad \ldots,\quad y^{(n-1)}=f^{(n-1)}(x),\quad y^{(n)}=f^{(n)}(x).$$

Обозначим символом $\delta$ дифференциал в некоторой точке $x_0 \in ]a, b[$ функции $dy=f'(x)dx$, зависящей только от $x$ (т. е. при фиксированном $dx$), тогда

$$\delta(dy)=\delta[f'(x)\,dx]\,|_{x=x_0}=[f'(x)\,dx]'\,|_{x=x_0}\delta x=f''(x_0)dx\delta x, \tag{11.61}$$

где $\delta x$ — новое приращение независимой переменной $x$.

Определение второго дифференциала можно сформулировать так: значение дифференциала $\delta(dy)$, т. е. дифференциала от первого дифференциала, при $\delta x = dx$ называется вторым дифференциалом функции $y=f(x)$ в точке $x_0$ и обозначается $d^2y$:

$$d^2y = f''(x_0)\,dx^2 \quad (dx^2 = (dx)^2). \tag{11.62}$$

Отметим, что второй дифференциал независимой переменной $x$ равен нулю:

$$d^2x = 0, \tag{11.63}$$

так как при вычислении дифференциалов приращение $dx = \Delta x$ мы считаем постоянным.

Дифференциал $n$-го порядка $d^n y$ функции $y = f(x)$ в точке $x_0$ определяется как дифференциал от дифференциала $(n-1)$-го порядка $d^{n-1}y$, для которого $\delta x = dx$:

$$d^n y = \delta(d^{n-1}y)\,|_{\delta x = dx}. \tag{11.64}$$

Докажем по индукции, что

$$d^n y = y^{(n)}\,dx^n,\ n = 1,\ 2,\ 3,\ \ldots . \tag{11.65}$$

Действительно, при $n=1$ и $n=2$ эта формула верна. Пусть она справедлива для дифференциала порядка $n-1$: $d^{n-1}y = y^{(n-1)}\,dx^{n-1}$. Для вычисления дифференциала $n$-го порядка согласно определению надо взять дифференциал (обозначим его символом $\delta$) от дифференциала $d^{n-1}y$:

$$\delta(d^{n-1}y) = \delta(y^{(n-1)}dx^{n-1}) = (y^{(n-1)}dx^{n-1})'\delta x = y^{(n)}dx^{n-1}\delta x$$

и положить $\delta x = dx$ : $d^n y = \delta(d^{n-1}y)\,|_{\delta x = dx} = y^{(n)}dx^n$.

Формула (11.65) доказана. Из этой формулы следует, что

$$y^{(n)} = \frac{d^n y}{dx^n}. \tag{11.66}$$

Замечание. Формулы (11.65) и (11.66) при $n > 1$ справедливы тогда и только тогда, когда $x$ является независимой переменной (в отличие от случая $n=1$, когда они верны и для $x = x(t)$, где $t$ — новая переменная).

## § 11.12. Основные теоремы дифференциального исчисления

**Теорема Лагранжа**[1]. Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема в интервале $]a, b[$, то существует такая точка $c \in ]a, b[$, что

$$f(b) - f(a) = f'(c)(b-a) \quad (a < c < b). \tag{11.67}$$

[1] *Жозеф Луи Лагранж (Joseph Louis Lagrange,* 1736—1813) — французский математик и механик.

Доказательство. Через точки $A(a, f(a))$ и $B(b, f(b))$ графика данной функции (рис. 11.6) проведем секущую $AB$. Угол, образуемый секущей $AB$ с осью $Ox$, обозначим через $\alpha$, тогда

$$\operatorname{tg}\alpha = \frac{f(b) - f(a)}{b - a}.$$

Будем перемещать эту секущую параллельно исходному положению до тех пор, пока она не превратится в

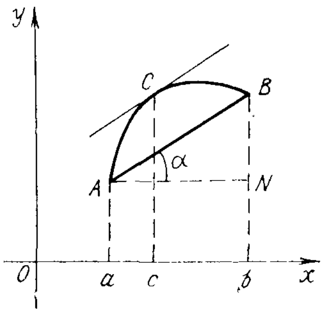



Рис. 11.6

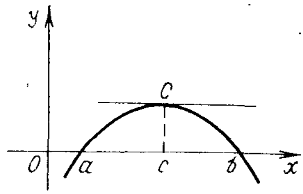



Рис. 11.7

касательную к графику функции $y=f(x)$ в некоторой точке $C(c, f(c))$, где $a<c<b$ (здесь опускается доказательство того факта, что такое предельное положение существует). Согласно построению и геометрическому смыслу производной $\operatorname{tg}\alpha=f'(c)$. Из двух полученных равенств следует равенство

$$\frac{f(b) - f(a)}{b - a} = f'(c) \quad (a<c<b)$$

и равенство (11.67), которое и требовалось доказать.

Следствие 1. Если производная функции равна нулю в каждой точке некоторого промежутка, то функция есть тождественная постоянная в этом промежутке.

Пусть $f'(x)=0$ для всех $x$ из данного промежутка. Если $x_0$ и $x$ — две точки этого промежутка, то по доказанной теореме $f(x)-f(x_0)=f'(c)(x-x_0)$, $x_0<c<x$. Поскольку $f'(c)=0$, то $f(x)-f(x_0)=0$, $f(x)=f(x_0)=\text{const}$.

Следствие 2. Если две функции имеют равные производные в некотором промежутке, то они отличаются в этом промежутке лишь постоянным слагаемым.

Действительно, если $f_1'(x)=f_2'(x)$, $x\in[a, b]$, то $[f_1(x)-f_2(x)]'=f_1'(x)-f_2'(x)=0$. В силу следствия 1 $f_1(x)-f_2(x)=C$, $x\in]a, b[$.

Замечание. Формула (11.67) называется *формулой конечных приращений.*

Прежде чем переходить к формулировке следующей теоремы, введем одно вспомогательное понятие.

*Корнем* (или *нулем*) функции $y=f(x)$ называется такое значение $x=x_0$ ее аргумента, при котором функция эта обращается в нуль. Геометрически корень функции означает абсциссу точки, в которой график функции пересекает ось $Ox$ или касается ее.

**Теорема Ролля**[1]. Между двумя различными корнями дифференцируемой функции содержится по меньшей мере один корень ее производной.

Доказательство. Пусть $a$ и $b$ — различные корни дифференцируемой функции, т. е. $f(a)=0$, $f(b)=0$. Из формулы (11.67) получаем $f'(c)(b-a)=0$ $(a<c<b)$.

Так как $b-a\neq 0$ (корни различны), то $f'(c)=0$, где $a<c<b$, что и требовалось доказать.

Замечание 1. Теорема имеет простую геометрическую интерпретацию. Между значениями $a$ и $b$ имеется по меньшей мере одно значение $c$ такое, что в точке $C(c, f(c))$ графика функции касательная к графику параллельна оси $Ox$ (рис. 11.7).

Замечание 2. Теорему можно сформулировать в более общем виде. Если $y=f(x)$ — функция, дифференцируемая иа отрезке $[a, b]$ и $f(a)=f(b)$, то между $a$ и $b$ найдется точка $c$, в которой производная равна иулю, т. е. $f'(c)=0$. Действительно, случай $f(a)=f(b)=0$ рассмотрен выше; если $f(a)\neq 0$, то введем функцию $F(x)=f(x)-f(a)$, тогда $F(x)$ дифференцируема и $F'(x)=f'(x)$, $F(a)=f(a)-f(a)=0$, $F(b)=f(b)-f(a)=0$, т. е. для функции $F(x)$ выполнены условия теоремы Ролля.

**Теорема Коши**[2]. Если $y=f(x)$ и $y=\varphi(x)$ — две функции, непрерывные на отрезке $[a, b]$ и дифференцируемые в интервале $]a, b[$, причем $\varphi'(x)\neq 0$ для любого $x\in]a, b[$, то между $a$ и $b$ найдется такая точка $c$, что

$$\frac{f(b)-f(a)}{\varphi(b)-\varphi(a)}=\frac{f'(c)}{\varphi'(c)}. \qquad (11.68)$$

Замечание. Знаменатель в левой части равенства (11.68) отличен от нуля (допустив противное, т. е. $\varphi(b)=\varphi(a)$, получили бы

[1] *Мишель Ролль (Michele Rolle,* 1652—1719) — французский математик.

[2] *Огюстен Луи Коши (Augustin Louis Cauchy,* 1789—1857) — знаменитый французский математик.

$\varphi'(c)=0$, что противоречит условию), поэтому выражение (11.68) имеет смысл.

Доказательство. Введем в рассмотрение функцию

$$F(x)=f(x)-f(a)-\frac{f(b)-f(a)}{\varphi(b)-\varphi(a)}[\varphi(x)-\varphi(a)]. \tag{11.69}$$

Эта функция удовлетворяет условиям теоремы Ролля. Она дифференцируема в интервале $]a,\ b[$, так как дифференцируемы в нем функции $f(x)$ и $\varphi(x)$:

$$F'(x)=f'(x)-\frac{f(b)-f(a)}{\varphi(b)-\varphi(a)}\varphi'(x).$$

$F(x)$ непрерывна на отрезке $[a,\ b]$, поскольку непрерывны на нем функции $f(x)$ и $\varphi(x)$. На концах отрезка $[a,\ b]$ функция $F(x)$ обращается в нуль, что проверяется непосредственной подстановкой значений $x=a$, $x=b$ в правую часть формулы (11.69).

Следовательно, между $a$ и $b$ найдется такая точка $c$, для которой $F'(c)=0$, т. е.

$$f'(c)-\frac{f(b)-f(a)}{\varphi(b)-\varphi(a)}\varphi'(c)=0.$$

Разделив это равенство почленно на $\varphi'(c)\neq 0$, получим равенство (11.68).

## § 11.13. Формула Тейлора[1]

В § 11.9 мы видели, что замена приращения функции ее дифференциалом дает возможность получить многие приближенные формулы (см. (11.51а,б,в)). Оказывается, эти формулы можно уточнить, если применить дифференциалы высших порядков.

Предположим, что функция $y=f(x)$ имеет все производные до $(n+1)$-го порядка включительно в некотором промежутке, содержащем точку $x=a$. Найдем многочлен $P_n(x)$ степени не выше $n$, такой, что

$$P_n(a)=f(a),\ P_n'(a)=f'(a),\ P_n''(a)=f''(a),\ \ldots$$
$$\ldots,\ P_n^{(n)}(a)=f^{(n)}(a), \tag{11.70}$$

[1] *Брук Тейлор (Brook Taylor,* 1685—1731) — английский математик, последователь Ньютона.

т. е. в точке $x=a$ значения многочлена и функции, а также их соответствующих производных совпадают. Многочлен $P_n(x)$ будем искать в виде

$$P_n(x)=C_0+C_1(x-a)+C_2(x-a)^2+\ldots$$
$$\ldots+C_n(x-a)^n. \tag{11.71}$$

Коэффициенты этого многочлена определим из условий (11.70), для чего предварительно найдем его производные:

$$\left.\begin{aligned}
&P_n'(x)=C_1+2C_2(x-a)+3C_3(x-a)^2+\ldots+nC_n(x-a)^{n-1},\\
&P_n''(x)=2C_2+2\cdot 3C_3(x-a)+\ldots+n(n-1)C_n(x-a)^{n-2},\\
&P_n'''(x)=2\cdot 3C_3+\ldots+n(n-1)(n-2)C_n(x-a)^{n-3},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&P_n^{(n)}(x)=n(n-1)(n-2)(n-3)\ \ldots\ 2\cdot 1C_n.
\end{aligned}\right\} \tag{11.72}$$

Подставляя значение $x=a$ в формулы (11.71) и (11.72), получаем

$$\begin{gathered}
P_n(a)=C_0,\ P_n'(a)=C_1,\ P_n''(a)=2C_2,\\
P_n'''(a)=2\cdot 3C_3,\ \ldots,\ P_n^{(n)}(a)=n(n-1)\ldots 2C_n.
\end{gathered} \tag{11.73}$$

Из равенств (11.70) и (11.73) находим

$$\left.\begin{gathered}
C_0=f(a),\ \ C_1=f'(a),\ \ C_2=\frac{1}{1\cdot 2}f''(a),\\
C_3=\frac{f'''(a)}{1\cdot 2\cdot 3},\ \ldots,\ C_n=\frac{f^{(n)}(a)}{1\cdot 2\cdot 3\ldots n},
\end{gathered}\right\} \tag{11.74}$$

или

$$C_k=\frac{f^{(k)}(a)}{k!},\ k=0,\ 1,\ 2,\ \ldots,\ n,$$
$$k!=1\cdot 2\cdot 3\ldots k,\ f^0(a)=f(a),\ 0!=1,\ 1!=1.$$

Подставляя эти значения $C_k$ ($k=0, 1, \ldots, n$) в формулу (11.71), получаем искомый многочлен:

$$P_n(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots$$
$$\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{11.75}$$

Обозначим через $R_n(x)$ разность между данной функцией $y=f(x)$ и многочленом (11.75): $f(x)-P_n(x)=R_n(x)$, откуда

$$f(x)=P_n(x)+R_n(x), \tag{11.76}$$

или

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^n(a)}{n!}(x-a)^n+R_n(x), \tag{11.77}$$

$R_n(x)$ называется *остаточным членом.*

Следовательно, при тех значениях $x$, для которых $R_n(x)$ достаточно мало, вместо функции $y=f(x)$ можно рассматривать ее многочлен (11.75).

Замечание 1. Очевидно, $R_n(x)=0$, если функция является многочленом степени $n$, т. е. $f(x)=P_n(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$. В этом случае многочлен (11.75) представляет лишь другую форму данного многочлена. Любой такой многочлен можно разложить по степеням $(x-a)$. Например, многочлен $P_2(x)=x^2+2x-5$ можно разложить по степеням $(x-3)$. Для этого нужно в правую часть подставить выражение $x=[3+(x-3)]$ и раскрыть квадратные скобки, оставив круглые: $P_2(x)=[3+(x-3)]^2+2[3+(x-3)]-5=9+6(x-3)+$ $+(x-3)^2+6+2(x-3)-5=(x-3)^2+8(x-3)+10$. Читателю предлагается убедиться в том, что это разложение получается с помощью формулы (11.75).

Оценим величину остаточного члена $R_n(x)$. Запишем его в виде

$$R_n(x)=\frac{(x-a)^{n+1}}{(n+1)!}Q(x), \tag{11.78}$$

где $Q(x)$ — функция, которую нужно определить. Формула (11.77) принимает вид

$$\left.\begin{aligned} f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots \\ \ldots+\frac{f^n(a)}{n!}(x-a)^n+\frac{(x-a)^{n+1}}{(n+1)!}Q(x). \end{aligned}\right\} \tag{11.79}$$

При фиксированных значениях $x$ и $a$ функция $Q(x)$ имеет определенное значение, обозначим его через $Q$. Введем в рассмотрение вспомогательную функцию переменной $t$ $(a<t<x)$:

$$\begin{aligned} F(t)=f(x)-f(t)-\frac{x-t}{1!}f'(t)-\frac{(x-t)^2}{2!}f''(t)-\ldots \\ \ldots-\frac{(x-t)^n}{n!}f^{(n)}(t)-\frac{(x-t)^{n+1}}{(n+1)!}Q, \end{aligned} \tag{11.80}$$

где $Q$ имеет значение, определенное формулой (11.79) при фиксированных $a$ и $x$.

Применяя правила дифференцирования алгебраической суммы и произведения двух функций, находим производную этой функции:

$$F'(t) = -\frac{(x-t)^n}{n!} f^{(n+1)}(t) + \frac{(x-t)^n}{n!} Q.$$

Из формулы (11.80) следует, что $F(x)=0$, $F(a)=0$. Следовательно, к функции $F(t)$ применима теорема Ролля, т. е. существует такое значение $t=\xi$, заключенное между $a$ и $x$, что $F'(\xi)=0$ или $-\frac{(x-\xi)^n}{n!} f^{(n+1)}(\xi) + \frac{(x-\xi)^n}{n!} Q = 0$, откуда $Q = f^{(n+1)}(\xi)$. Подставляя это выражеиие в формулу (11.78), получаем так называемую форму Лагранжа для остаточного члена:

$$R_n(x) = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi). \qquad (11.81)$$

Поскольку $\xi$ заключено между $a$ и $x$, его можно представить в виде $\xi = a + \Theta(x-a)$, где $0 < \Theta < 1$, тогда

$$R_n(x) = \frac{(x-a)^{n+1}}{n+1!} f^{(n+1)}(a + \Theta(x-a)).$$

Формула

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots$$
$$\ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(a+\Theta(x-a))}{(n+1)!}(x-a)^{n+1} \qquad (11.82)$$

называется *формулой Тейлора* с остаточным членом в *форме Лагранжа*.

Если $a=0$, то формула принимает вид

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots$$
$$\ldots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\Theta x)}{(n+1)!}x^{n+1}, \qquad (11.83)$$

где $0 < \Theta < 1$; называется она также *формулой Маклорена* [1)].

Формулу (11.82) можно записать в виде

[1)] *Колин Маклорен* (*Colin Maclaurin*, 1698—1746) — шотландский математик.

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+$$
$$+o((x-a)^n), \qquad (11.84)$$

где $o((x-a)^n)$ — бесконечно малая порядка выше $n$-го по сравнению с $x-a$; эта форма остаточного члена была указана Пеано[1].

Замечание 2. Если в формулах (11.82) и (11.84) перенести в левые части $f(a)$ и обозначить $x-a=\Delta x$, $f(x)-f(a)=\Delta f(a)$, то

$$\Delta f(a)=f'(a)\Delta x+\frac{f''(a)}{2!}\Delta x^2+\ldots+\frac{f^{(n)}(a)}{n!}\Delta x^n+$$
$$+\frac{f^{(n+1)}(\xi)}{(n+1)!}\Delta x^{n+1},$$
$$\Delta f(a)=f'(a)\Delta x+\frac{1}{2!}f''(a)\Delta x^2+\ldots+\frac{1}{n!}f^{(n)}(a)\Delta x^n+o(\Delta x^n).$$

Если в этих формулах $\Delta x$ заменить на $dx$ и принять во внимание формулы (11.62), (11.65), (11.66), получим соответственно

$$\Delta f(a)=df(a)+\frac{1}{2!}d^2f(a)+\ldots+\frac{1}{n!}d^nf(a)+$$
$$+\frac{1}{(n+1)!}d^{n+1}f(\xi),$$
$$\Delta f(a)=df(a)+\frac{1}{2!}d^2f(a)+\ldots+\frac{1}{n!}d^nf(a)+o(\Delta x^n).$$

Следовательно, если предположить, что $\Delta x\to 0$, то по этим формулам из бесконечно малого приращения функции $\Delta f(a)$ можно выделить не только его главный член — первый дифференциал, но и члены более высоких порядков малости, совпадающие (с точностью до факториалов в знаменателях) с последовательными дифференциалами $d^2f(a)$, $d^3f(a)$, ..., $d^nf(a)$.

## § 11.14. Формула Тейлора для некоторых функций

Рассмотрим конкретные разложения по формуле (11.83) некоторых важнейших элементарных функций.

**Разложение функции** $f(x)=e^x$. Поскольку

$$f'(x)=e^x,\ f''(x)=e^x,\ \ldots,\ f^{(n)}(x)=e^x,\ f^{(n+1)}(x)=e^x,$$
$$f(0)=1,\ f'(0)=1,\ f''(0)=1,\ \ldots,\ f^{(n)}(0)=1,\ f^{(n+1)}(\Theta x)=e^{\Theta x},$$

[1] *Джузеппе Пеано* (*Giuseppe Peano,* 1858—1932) — итальянский математик.

формула (11.83) для функции $f(x)=e^x$ принимает вид

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} e^{\Theta x} (0 < \Theta < 1). \tag{11.85}$$

Отметим, что при любом $x$ остаточный член формулы (11.85) стремится к нулю при неограниченном возрастании $n$, т. е.

$$\lim_{n\to\infty} R_n(x) = \lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} e^{\Theta x} = 0.$$

Действительно, при фиксированном $x$ величина $e^{\Theta x}$, где $0<\Theta<1$, ограничена (она меньше $e^x$ при $x>0$ и меньше 1 при $x<0$), а $\lim\limits_{n\to\infty} \frac{x^{n+1}}{(n+1)!} = 0$.

**Разложение функции** $f(x) = \sin x$. Так как

$$f'(x) = \cos x = \sin\left(x + \frac{\pi}{2}\right),\ f''(x) = -\sin x = \sin\left(x + 2\cdot\frac{\pi}{2}\right),$$

$$f'''(x) = -\cos x = \sin\left(x + 3\cdot\frac{\pi}{2}\right),\ f^{\mathrm{IV}}(x) = \sin x =$$

$$= \sin\left(x + 4\cdot\frac{\pi}{2}\right),\ \ldots,$$

$$f^{(n)}(x) = \sin\left(x + \frac{n\pi}{2}\right),\ f^{(n+1)}(x) = \sin\left(x + \frac{(n+1)\pi}{2}\right);$$

$$f(0) = 0,\ f'(0) = 1,\ f''(0) = 0,\ f'''(0) = -1,\ f^{\mathrm{IV}}(0) = 0,$$

$$f^{(n)}(0) = \sin\frac{n\pi}{2},\quad f^{(n+1)}(\Theta x) = \sin\left(\Theta x + \frac{(n+1)\pi}{2}\right),$$

то

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{x^n}{n!}\sin\frac{n\pi}{2} + $$
$$+ \frac{x^{n+1}}{(n+1)!}\sin\left(\Theta x + \frac{(n+1)\pi}{2}\right). \tag{11.86}$$

Остаточный член

$$R_n(x) = \frac{x^{n+1}}{(n+1)!}\sin\left(\Theta x + \frac{(n+1)\pi}{2}\right)$$

формулы (11.86) также стремится к нулю при $n\to\infty$.

**Разложение функции** $f(x) = \cos x$. Поскольку

$$f'(x) = -\sin x = \cos\left(x + \frac{\pi}{2}\right),\ f''(x) = -\cos x = \cos\left(x + 2\cdot\frac{\pi}{2}\right),$$

$$f'''(x) = \sin x = \cos\left(x + 3\cdot\frac{\pi}{2}\right),\ f^{IV}(x) = \cos x =$$

$$= \cos\left(x + 4\cdot\frac{\pi}{2}\right),\ \ldots,$$

$$f^{(n)}(x) = \cos\left(x + \frac{n\pi}{2}\right),\ f^{(n+1)}(x) = \cos\left(x + \frac{(n+1)\pi}{2}\right);$$

$$f(0) = 1,\ f'(0) = 0,\ f''(0) = -1,\ f'''(0) = 0,\ f^{IV}(0) = 1,\ \ldots,$$

$$f^{(n)}(0) = \cos\frac{n\pi}{2},\quad f^{(n+1)}(\Theta x) = \cos\left(\Theta x + \frac{(n+1)\pi}{2}\right),$$

то

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + \frac{x^n}{n!}\cos\frac{n\pi}{2} + $$
$$+ \frac{x^{n+1}}{(n+1)!}\cos\left(\Theta x + \frac{(n+1)\pi}{2}\right). \qquad (11.87)$$

Каково бы ни было $x$, остаточный член формулы (11.87) стремится к нулю при $n \to \infty$.

**Разложение функции** $f(x) = (a+x)^n$, где $a$ — действительное число, $n$ — натуральное число.

Поскольку $k$-я производная данной функции

$$f^{(k)}(x) = n(n-1)\ldots(n-k+1)(a+x)^{n-k} \quad (k=1,\ 2,\ \ldots,\ n)$$

при $x=0$ принимает значение

$$f^{(k)}(0) = n(n-1)\ldots(n-k+1)a^{n-k} \quad (k=1,\ 2,\ \ldots,\ n),$$

то

$$(a+x)^n = a^n + na^{n-1}x + \frac{n(n-1)}{2!}a^{n-2}x^2 +$$
$$+ \frac{n(n-1)(n-2)}{3!}a^{n-3}x^3 + \ldots + \frac{n(n-1)\ldots 2}{(n-1)!}ax^{n-1} + x^n. \quad (11.88)$$

Это равенство называется формулой *бинома Ньютона.*

## § 11.15. Приближенные формулы

Если в формуле (11.83) отбросить остаточный член, то получится приближенная формула

$$f(x) \approx f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n, \quad (11.89)$$

заменяющая данную функцию многочленом $n$-й степени. Качество этой формулы оценивается двояко: указываются границы погрешности $R_n(x)$ с помощью выражения (11.81) для остаточного члена, либо порядок малости этой погрешности при $x \to 0$ $R_n(x) = o(x^n)$. Для примеров обратимся к рассмотренным в § 11.14 разложениям некоторых функций по формулам Тейлора.

В случае функции $f(x) = e^x$ получаем приближенную формулу

$$e^x \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}. \qquad (11.90)$$

Поскольку $R_n(x) = \frac{e^{\theta x}}{(n+1)!} x^{n+1}$, то, например, при $x > 0$ погрешность оценивается неравенствами $0 < R_n(x) < \frac{e^x}{(n+1)!} x^{n+1}$.

В частности, при $x = 1$ получаем

$$e \approx 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}, \quad 0 < R_n(x) < \frac{3}{(n+1)!}.$$

Если взять $n=8$ и произвести вычисления с пятью десятичными знаками, то получим $e=2{,}71827$. Здесь верны первые четыре знака, так как ошибка не превосходит $\frac{3}{9!}$ или 0,00001.

Взяв $f(x) = \sin x$ и положив в равенстве (11.86) $n=2m$, получим приближенную формулу

$$\sin x \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + \frac{(-1)^{m-1} x^{2m-1}}{(2m-1)!}, \qquad (11.91)$$

остаточный член которой

$$R_{2m}(x) = \frac{\sin\left(\theta x + (2m+1)\frac{\pi}{2}\right)}{(2m+1)!} x^{2m+1} =$$

$$= (-1)^m \cos \theta x \frac{x^{2m+1}}{(2m+1)!}$$

оценивается соотношением $|R_{2m}(x)| \leqslant \frac{|x|^{2m+1}}{(2m+1)!}$.

Для функции $f(x) = \cos x$ аналогично получаем (положив в формуле (11.87) $n=2m$)

$$\cos x \approx 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^m \frac{x^{2m}}{(2m)!}. \qquad (11.92)$$

Погрешность приближенной формулы (11.92) выражается остаточным членом $R_{2m+1}(x) = (-1)^{m+1} \cos\theta x \frac{x^{2m+2}}{(2m+2)!}$ и оценивается неравенством $|R_{2m+1}(x)| \leqslant \frac{|x|^{2m+2}}{(2m+2)!}$.

Например, для формулы $\cos x \approx 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$ погрешность $|R_5(x)| \leqslant \frac{|x|^6}{6!} = \frac{x^6}{720}$.

# Глава 12

# ПРАВИЛО ЛОПИТАЛЯ—БЕРНУЛЛИ. МАКСИМУМЫ И МИНИМУМЫ

Эта глава посвящена некоторым приложениям производной. С помощью производной можно находить многие пределы (раскрывать соответствующие неопределенности), исследовать функции, строить их графики, решать задачи на отыскание наибольших и наименьших значений функций.

## § 12.1. Правило Лопиталя—Бернулли

При исследовании функций может появиться необходимость нахождения предела дроби $f(x) : \varphi(x)$, числитель и знаменатель которой при $x \to a$ стремятся к нулю или бесконечности. Нахождение таких пределов называют *раскрытием неопределенностей* соответствующего вида. Основой его является правило Лопиталя — Бернулли[1], выражаемое следующей теоремой.

**Теорема 12.1.** Если функции $f(x)$ и $\varphi(x)$ дифференцируемы в окрестности точки $x=a$, обращаются в нуль в этой точке и существует предел отношения $f'(x) : \varphi'(x)$ при $x \to a$, то существует предел отношения самих функций, равный отношению производных:

$$\lim_{x\to a} \frac{f(x)}{\varphi(x)} = \lim_{x\to a} \frac{f'(x)}{\varphi'(x)}. \tag{12.1}$$

Доказательство. Пусть точка $x \neq a$ принадле-

[1] Правило сформулировано швейцарским математиком И. Бернулли (1667—1748). Опубликовано в первом печатном учебнике анализа бесконечно малых (1696), написанном французским математиком Г. Лопиталем (1661—1704).

жит интервалу, в котором функции дифференцируемы. По теореме Коши

$$\frac{f(x)-f(a)}{\varphi(x)-\varphi(a)}=\frac{f'(c)}{\varphi'(c)},$$

где $c$ лежит между $x$ и $a$.

По условию $f(a)=\varphi(a)=0$, поэтому

$$\frac{f(x)}{\varphi(x)}=\frac{f'(c)}{\varphi'(c)}.$$

Если $x\to a$, то и $c\to a$, так как $c$ заключено между $x$ и $a$. Переходя к пределу в последнем равенстве, получаем

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

откуда и следует формула (12.1).

Пример 1. Найти $\lim\limits_{x\to 0}\dfrac{\sin 4x}{2x-\sin 3x}$.

При $x=0$ числитель и знаменатель обращаются в нуль, имеем неопределенность вида $\frac{0}{0}$. Чтобы раскрыть ее, применяем правило Лопиталя—Бернулли:

$$\lim_{x\to 0}\frac{\sin 4x}{2x-\sin 3x}=\lim_{x\to 0}\frac{4\cos 4x}{2-3\cos 3x}=\frac{4}{2-3}=-4.$$

Замечание 1. Теорема верна и в том случае, когда функции $f(x)$ и $\varphi(x)$ не определены в точке $x=a$, но $\lim\limits_{x\to a}f(x)=0$, $\lim\limits_{x\to a}\varphi(x)=0$. Действительно, этот случай можно свести к рассмотренному в теореме, если доопределить функции $f(x)$ и $\varphi(x)$ так, чтобы они были непрерывны в точке $x=a$, для чего нужно положить $f(a)=\lim\limits_{x\to a}f(x)=0$, $\varphi(a)=\lim\limits_{x\to a}\varphi(x)=0$; предел отношения $f(x):\varphi(x)$ при $x\to a$ не зависит от того, определены ли данные функции в точке $x=a$.

Замечание 2. Теорема верна и при $a=\infty$, т. е. когда $\lim\limits_{x\to\infty}f(x)=0$, $\lim\limits_{x\to\infty}\varphi(x)=0$.

В самом деле, положив $x=1:z$, видим, что $z\to 0$ при $x\to\infty$ и

$$\lim_{z\to 0}f\left(\frac{1}{z}\right)=0,\quad \lim_{z\to 0}\varphi\left(\frac{1}{z}\right)=0;$$

далее,

$$\lim_{x\to\infty}\frac{f(x)}{\varphi(x)}=\lim_{z\to 0}\frac{f\left(\frac{1}{z}\right)}{\varphi\left(\frac{1}{z}\right)}=\lim_{z\to 0}\frac{f'\left(\frac{1}{z}\right)\left(-\frac{1}{z^2}\right)}{\varphi'\left(\frac{1}{z}\right)\left(-\frac{1}{z^2}\right)}=$$

$$= \lim_{z\to 0} \frac{f'\left(\frac{1}{z}\right)}{\varphi'\left(\frac{1}{z}\right)} = \lim_{x\to\infty} \frac{f'(x)}{\varphi'(x)}.$$

Следовательно,

$$\lim_{x\to\infty} \frac{f(x)}{\varphi(x)} = \lim_{x\to\infty} \frac{f'(x)}{\varphi'(x)}, \tag{12.2}$$

если предел в правой части существует.

Замечание 3. Если $f'(a)=0$, $\varphi'(a)=0$, функции $f'(x)$, $\varphi'(x)$ дифференцируемы в окрестности точки $x=a$ и существует предел отношения $f''(x) : \varphi''(x)$ при $x\to a$, то

$$\lim_{x\to a} \frac{f'(x)}{\varphi'(x)} = \lim_{x\to a} \frac{f''(x)}{\varphi''(x)}. \tag{12.3}$$

Другими словами, правило Лопиталя — Бернулли при выполнении соответствующих условий можно применять несколько раз.

Правило Лопиталя — Бернулли применимо и при раскрытии неопределенностей вида $\frac{\infty}{\infty}$, поскольку ее можно привести к неопределенности вида $\frac{0}{0}$, представив рассматриваемую дробь так:

$$\frac{f(x)}{\varphi(x)} = \frac{1}{\varphi(x)} : \frac{1}{f(x)}.$$

Пример 2. Найти $\lim\limits_{x\to+\infty} \frac{\ln x}{x^\alpha}$, где $\alpha > 0$.

При $x\to+\infty$ числитель и знаменатель неограниченно возрастают, имеем неопределенность вида $\frac{\infty}{\infty}$. Чтобы раскрыть ее, применим правило Лопиталя—Бернулли:

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

Этот пример показывает, что при неограниченном возрастании $x$ логарифмическая функция растет медленнее степенной функции.

Пример 3. Найти $\lim\limits_{x\to+\infty} \frac{x^n}{e^x}$, где $n$ — натуральное число.

Применяя правило Лопиталя—Бернулли $n$ раз, получаем

$$\lim_{x\to+\infty} \frac{x^n}{e^x} = \lim_{x\to+\infty} \frac{nx^{n-1}}{e^x} = \lim_{x\to+\infty} \frac{n(n-1)x^{n-2}}{e^x} = \ldots$$

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

Следовательно, при неограниченном возрастании аргумента степенная функция растет медленнее показательной функции.

С помощью тождественных преобразований к основному виду $\frac{0}{0}$ или $\frac{\infty}{\infty}$ можно свести неопределенности других видов, таких как $0 \cdot \infty$, $\infty - \infty$, $1^\infty$, $0^0$, $\infty^0$.

Неопределенность вида $0 \cdot \infty$, т. е. произведение $f(x)\varphi(x)$, где $f(x) \to 0$, $\varphi(x) \to \infty$ при $x \to a$, приводится к виду $\frac{0}{0}$ или $\frac{\infty}{\infty}$ по формулам

$$f(x) \cdot \varphi(x) = f(x) : \frac{1}{\varphi(x)}, \quad f(x)\varphi(x) = \varphi(x) : \frac{1}{f(x)},$$

а затем применяется правило Лопиталя — Бернулли.

Аналогично раскрывается неопределенность вида $\infty - \infty$, т. е. находится предел

$$\lim_{x \to a} [f(x) - \varphi(x)]$$

при условии, что $\lim\limits_{x \to a} f(x) = \infty$, $\lim\limits_{x \to a} \varphi(x) = \infty$. С помощью преобразования

$$f(x) - \varphi(x) = \left[\frac{1}{\varphi(x)} - \frac{1}{f(x)}\right] : \frac{1}{f(x)\varphi(x)}$$

эта неопределенность сводится к неопределенности вида $\frac{0}{0}$.

Раскрыть неопределенность вида $1^\infty$ — это значит найти предел $\lim\limits_{x \to a} (f(x))^{\varphi(x)}$ при условии, что $\lim\limits_{x \to a} f(x) = 1$, $\lim\limits_{x \to a} \varphi(x) = \infty$. С примером неопределенности вида $1^\infty$ мы уже встречались (§ 10.14).

Раскрыть неопределенности вида $0^0$, $\infty^0$ — это значит найти предел $\lim\limits_{x \to a} (f(x))^{\varphi(x)}$ при соответствующем условии: 1) $\lim\limits_{x \to a} f(x) = 0$, $\lim\limits_{x \to a} \varphi(x) = 0$; 2) $\lim\limits_{x \to a} f(x) = \infty$, $\lim\limits_{x \to a} \varphi(x) = 0$.

Неопределенности $1^\infty$, $0^0$, $\infty^0$ раскрываются с помощью способа, в котором используется тождество $(f(x))^{\varphi(x)} = e^{\varphi(x) \ln f(x)}$.

При раскрытии неопределенностей вида $1^\infty$, $0^0$, $\infty^0$ данное выражение предварительно логарифмируют и находят предел его логарифма.

Пример 4. Найти $\lim\limits_{x\to 1} x^{\frac{1}{1-x}}$.

Это неопределенность вида $1^{\infty}$. Обозначим $y = x^{\frac{1}{1-x}}$ и прологарифмируем равенство, получим $\ln y = \dfrac{\ln x}{1-x}$.

Применяя к правой части этого равенства правило Лопиталя — Бернулли, находим

$$\lim_{x\to 1}\ln y = \lim_{x\to 1}\frac{\ln x}{1-x} = \lim_{x\to 1}\frac{\frac{1}{x}}{-1} = -1.$$

Следовательно, $\lim\limits_{x\to 1}\ln y = -1$, $\lim\limits_{x\to 1} y = e^{-1}$, $\lim\limits_{x\to 1} x^{\frac{1}{1-x}} = e^{-1}$.

## § 12.2. Признаки постоянства, возрастания и убывания функции

Рассмотрим функцию $y=c$, где $c=\text{const}$, сохраняющую одно и то же значение. Принимая во внимание следствие 1 из теоремы Лагранжа и формулу $y'=0$, заключаем, что необходимое и достаточное условие постоянства функции выражается равенством

$$y'=0. \tag{12.4}$$

Введем определения возрастающей и убывающей функции.

Функция $y=f(x)$ называется *возрастающей* в некотором промежутке $(a, b)$, если для любых двух значений $x_1$ и $x_2$, принадлежащих этому промежутку, из неравенства

$$x_1<x_2 \tag{12.5}$$

следует неравенство

$$f(x_1)<f(x_2) \tag{12.6}$$

(рис. 12.1, *а*).

Функция $y=f(x)$ называется *убывающей* в некотором промежутке, если для любых двух значений, принадлежащих этому промежутку, из неравенства (12.5) следует неравенство

$$f(x_1)>f(x_2) \tag{12.7}$$

(рис. 12.1, *б*).

Достаточное условие возрастания (убывания) функции выражается следующей теоремой.

**Теорема 12.2.** Если в данном промежутке производная функции положительна, то функция возрастает в этом промежутке; если производная отрицательна, то функция убывает в соответствующем промежутке.

Доказательство. Пусть $x_1$ и $x_2$ принадлежат промежутку, в котором $f'(x)>0$; будем считать, что $x_1< <x_2$. По теореме Лагранжа

$$f(x_2)-f(x_1)=f'(c)(x_2-x_1), \qquad (12.8)$$

где $x_1<c<x_2$.

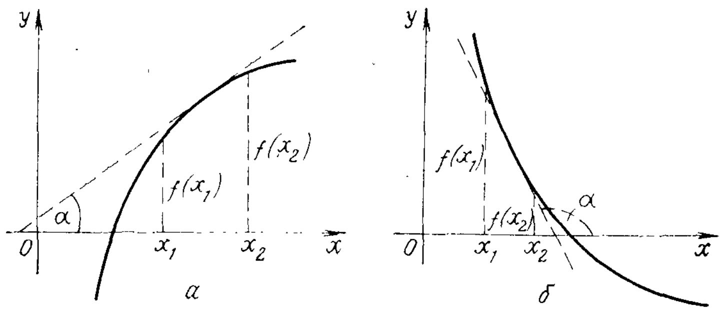



Рис. 12.1

Поскольку $f'(c)>0$, то разности $f(x_2)-f(x_1)$ и $(x_2-x_1)$ одного знака, причем $x_2-x_1>0$, поэтому $f(x_2)-f(x_1)>0$. Следовательно, из неравенства $x_1<x_2$ следует неравенство $f(x_1)<f(x_2)$, т. е. функция возрастает в промежутке, где $f'(x)>0$.

Если $f'(x)<0$ для всех $x$ из данного промежутка, то $f'(c)<0$. Из неравенства (12.8) следует, что $f(x_2)- -f(x_1)<0$ при $x_2-x_1>0$, т. е. $f(x_1)>f(x_2)$, когда $x_1<x_2$. Это означает, что функция убывает в данном промежутке.

Замечание. Теорема имеет простой геометрический смысл. Если в некотором промежутке касательная к графику функции $y=f(x)$ образует с осью $Ox$ острый угол $\alpha$ ($\operatorname{tg}\alpha>0$), то функция возрастает в этом промежутке (рис. 12.1, *а*). Если касательная к графику образует с осью $Ox$ тупой угол $\alpha$ ($\operatorname{tg}\alpha<0$), функция убывает (рис. 12.1, *б*).

Пример. Найти промежутки возрастания и убывания функции

$$f(x)=x^3-6x^2+9x-2.$$

Находим производную функции и разлагаем на множители соответствующий квадратный трехчлен: $f'(x)=3x^2-12x+9=3(x^2-4x+3)$, $f'(x)=3(x-1)(x-3)$.

Если $x<1$ и $x>3$, то $f'(x)>0$; функция возрастает в интервалах $]-\infty, 1[$, $]3, +\infty[$. Если $1<x<3$, то $f'(x)<0$; функция убывает в интервале $]1, 3[$.

## § 12.3. Экстремум функции. Необходимое условие экстремума

Рассмотрим функцию $y=f(x)$, областью определения которой является промежуток $(a, b)$.

Если можно указать такую $\delta$-окрестность точки $x_1$, принадлежащую промежутку $(a, b)$, что для всех $x\in O(x_1, \delta)$, $x\neq x_1$, выполняется неравенство 70

$$f(x_1)>f(x), \tag{12.9}$$

то $y_1=f(x_1)$ называют максимумом функции $y=f(x)$ (рис. 12.2). Максимум[1] функции $y=f(x)$ обозначим через $\max f(x)$.

Если можно указать такую $\delta$-окрестность точки $x_2$, принадлежащую промежутку $(a, b)$, что для всех $x\in O(x_2, \delta)$, $x\neq x_2$, выполняется неравенство 71

$$f(x_2)<f(x), \tag{12.10}$$

то $y_2=f(x_2)$ называют минимумом функции $y=f(x)$ (рис. 12.2).

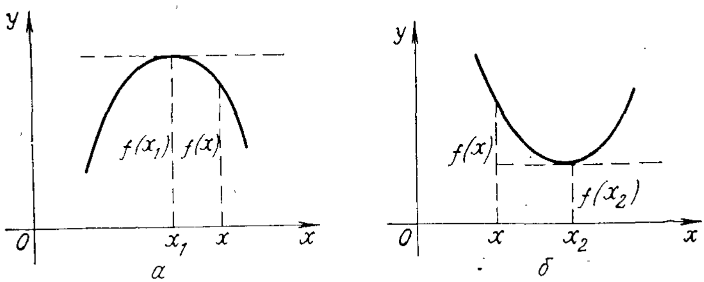



Рис. 12.2

Минимум функции $y=f(x)$ обозначим через $\min f(x)$.

Другими словами, максимумом (минимумом) функции $y=f(x)$ называют такое ее значение, которое больше

[1] Латинские слова maximum и minimum означают соответственно «наибольшее» и «наименьшее» (значение).

(меньше) всех других значений, принимаемых в точках, достаточно близких к данной и отличных от нее.

Отметим, что максимум и минимум функции имеют локальный характер (это наибольшее и наименьшее значение функции в достаточно малой окрестности соответствующей точки); отдельные минимумы некоторой функции могут оказаться больше максимумов той же функции (рис. 12.3).

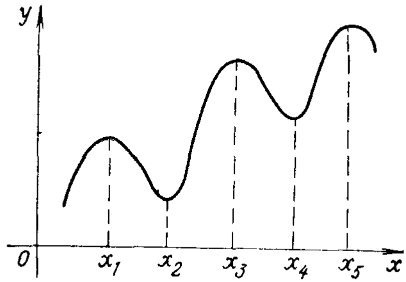



Рис. 12.3

Максимум и минимум функции называются экстремумом[1]. Значение аргумента, при котором достигается экстремум, называется точкой экстремума. Необходимое условие экстремума выражается следующей теоремой.

**Теорема 12.3.** В точке экстремума дифференцируемой функции производная ее равна нулю.

Доказательство. Пусть $x_0$ — точка экстремума дифференцируемой функции $y=f(x)$. Для определенности предположим, что $x_0$ — точка максимума, тогда при достаточно малых $\Delta x$ ($|\Delta x|<\delta$, $\delta>0$) $f(x_0)>f(x_0+\Delta x)$, поэтому $f(x_0+\Delta x)-f(x_0)<0$;

$$\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}<0 \text{ при } \Delta x>0,$$

$$\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}>0 \text{ при } \Delta x<0,$$

откуда при переходе к пределу получаем

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\leqslant 0 \quad (\Delta x>0),$$

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\geqslant 0 \quad (\Delta x<0).$$

[1] Латинское extremum означает «крайнее» (значеине).

Поскольку $f'(x_0)$ является числом, не зависящим от способа стремления $\Delta x$ к нулю, два последних соотношения совместимы лишь в том случае, когда

$$f'(x_0)=0. \qquad (12.11)$$

Аналогично доказывается теорема и для случая, когда $x_0$ — точка минимума функции.

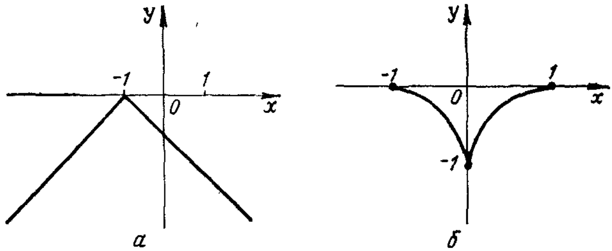



Рис. 12.4

Теорема имеет простой геометрический смысл: касательная к графику дифференцируемой функции в соответствующей точке параллельна оси $Ox$ (рис. 12.2).

Замечание 1. Если $f'(x_0)=0$, то отсюда еще не следует, что $x_0$ — точка экстремума. Например, для функции $f(x)=x^3$ $f'(x)=3x^2$, $f'(0)=0$, но $x_0=0$ не является точкой экстремума, так как $f(x)>0$ при $x>0$ и $f(x)<0$ при $x<0$ (неравенство (12.9) здесь не выполняется).

Замечание 2. Функция может достигать экстремума также в точке, в которой производная не существует. Например, функция $y=-|x+1|$ не имеет производной в точке $x=-1$, но достигает в ней максимума: $y=0$ при $x=-1$, а для всякой другой точки $y<0$ (рис. 12.4, $a$). Функция $y=-\left(1-x^{\frac{2}{3}}\right)^{\frac{3}{2}}$ не имеет конечной производной в точке $x=0$, поскольку $y'=\left(1-x^{\frac{2}{3}}\right)x^{-\frac{1}{3}}$ при $x=0$ обращается в бесконечность, но в этой точке функция имеет минимум: $f(0)=-1$, $f(x)>-1$ при $x\neq 0$ (рис. 12.4, $б$).

Точка, в которой производная равна нулю, называется *стационарной*. Стационарная точка, а также точка, в которой функция имеет бесконечную производную или в которой производная не существует, называется *критической*. Из сказанного следует, что точки экстремума следует искать среди критических точек.

15

## § 12.4. Достаточное условие экстремума

Для формулировки теоремы, выражающей достаточное условие экстремума с помощью первой производной, нам понадобится одно предварительное понятие.

Говорят, что функция $y=f(x)$ меняет знак при переходе через точку $x=x_0$, если $f(x_1)f(x_2)<0$ для любых $x_1$ и $x_2$ из некоторой окрестности этой точки, удовлетворяющих неравенствам $x_1<x_0<x_2$; знак меняется с плюса на минус, если $f(x_1)>0$, а $f(x_2)<0$; знак меняется с минуса на плюс, если $f(x_1)<0$, $f(x_2)>0$.

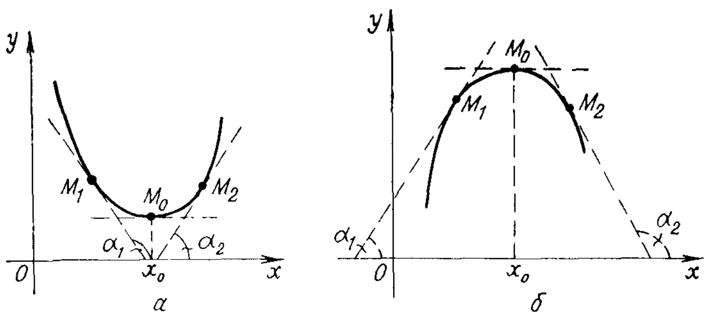



Рис. 12.5

Формулируя теоремы 12.4 и 12.5, будем предполагать, что функция $y=f(x)$ дифференцируема в некоторой окрестности точки $x_0$.

**Теорема 12.4.** Если в точке $x=x_0$ производная функции $y=f(x)$ равна нулю и меняет знак при переходе через точку, то $x_0$ является точкой экстремума, причем: 1) $x_0$ — точка максимума, если знак меняется с плюса на минус; 2) $x_0$ — точка минимума, если знак меняется с минуса на плюс.

Доказательство. Пусть в точке $x_0$ производная равна нулю и меняет знак с минуса на плюс, т. е. $f'(x_0)=0$, $f'(x)<0$ при $x_0-\varepsilon<x<x_0$, $f'(x)>0$ при $x_0<x<x_0+\varepsilon$ $(\varepsilon>0)$, тогда по теореме 12.2 функция $f(x)$ убывает в промежутке $(x_0-\varepsilon,\ x_0)$, возрастает в промежутке $(x_0,\ x_0+\varepsilon)$, т. е. $f(x_0)<f(x)$ для всех $x\in(x_0-\varepsilon,\ x_0+\varepsilon)$. Это означает, что $x_0$ — точка минимума.

Аналогично доказывается первое утверждение теоремы.

Теорема имеет следующий геометрический смысл: если в точке $M_0(x_0, f(x_0))$ графика дифференцируемой функции касательная параллельна оси $Ox$, в точках слева от $M_0$ образует острый угол с осью $Ox$, в точках справа — тупой, то $x_0$ — точка максимума (рис. 12.5, *б*).

Замечание. Теорема верна и в случае, если $x_0$ — точка непрерывности функции $f(x)$, производная в ней не существует и меняет знак при переходе через эту точку.

Достаточное условие экстремума можно выразить также с помощью второй производной.

**Теорема 12.5.** Если в точке $x=x_0$ первая производная функции $y=f(x)$ равна нулю, а вторая производная отлична от нуля, то $x_0$ является точкой экстремума, причем: 1) $x_0$ — точка минимума, если $f''(x_0)>0$; 2) $x_0$ — точка максимума, если $f''(x_0)<0$.

Доказательство. По определению второй производной

$$f''(x_0)=\lim_{x\to x_0}\frac{f'(x)-f'(x_0)}{x-x_0},$$

поэтому

$$f''(x_0)=\lim_{x\to x_0}\frac{f'(x)}{x-x_0},$$

так как $f'(x_0)=0$.

Если $f''(x_0)>0$, то в некоторой $\varepsilon$-окрестности точки $x_0$ (см. теорему 10.10) $\frac{f'(x)}{x-x_0}>0$, т. е. $f'(x)>0$, когда $x_0<x<x_0+\varepsilon$, $f'(x)<0$, когда $x_0-\varepsilon<x<x_0$.

Из последних неравенств на основании теоремы 12.2 заключаем, что $f(x)$ убывает в промежутке $(x_0-\varepsilon, x_0)$ и возрастает в промежутке $(x_0, x_0+\varepsilon)$, т. е. $f(x_0)<f(x)$ для всех $x\in(x_0-\varepsilon, x_0+\varepsilon)$. Следовательно, $x_0$ — точка минимума.

Второе утверждение теоремы доказывается аналогично.

Пример 1. Найти экстремумы функции $f(x)=x^4-10x^2+15$.

Поскольку $f'(x)=4x^3-20x=4x(x^2-5)$, то критическими точками (для которых $f'(x)=0$) являются $x_1=-\sqrt{5}$, $x_2=0$, $x_3=\sqrt{5}$.

Исследуем знак второй производной $f''(x)=12x^2-20$ в этих точках: $f''(-\sqrt{5})=12\cdot5-20>0$, $f''(\sqrt{5})=12\cdot5-20>0$, $f''(0)=-20<0$.

Следовательно, $x_1 = -\sqrt{5}$ и $x_3 = \sqrt{5}$ — точки минимума, $x_2 = 0$ — точка максимума, причем $\min f(x) = f(-\sqrt{5}) = f(\sqrt{5}) = 25 - 50 + 15 = -10$, $\max f(x) = f(0) = 15$. Данная функция достигает максимума, равного $-10$, при $x = 0$ и минимума, равного $-15$, при $x = -\sqrt{3}$, $x = \sqrt{3}$.

## § 12.5. Исследование функций с помощью формулы Тейлора

Теорема 12.5 не дает ответа на вопрос о наличии экстремума, если первые две производные функции в некоторой точке равны нулю.

Исследование этого случая проводится с помощью следующей теоремы.

**Теорема 12.6.** Пусть в точке $x_0$ первые $n$ производные равны нулю, а $(n+1)$-я отлична от нуля и непрерывна в этой точке, тогда: 1) если $(n+1)$ — четное число, то $x_0$ — точка экстремума, причем точка максимума при $f^{(n+1)}(x_0) < 0$ и точка минимума при $f^{(n+1)}(x_0) > 0$; 2) если $(n+1)$ — нечетное число, то $x_0$ не является точкой экстремума.

Доказательство. Так как в соответствии с условием $f'(x_0) = f''(x_0) = \ldots = f^{(n)}(x_0) = 0$, то разложение данной функции в окрестности $x_0$ по формуле Тейлора принимает вид

73

$$f(x) = f(x_0) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}, \qquad (12.12)$$

где $\xi$ лежит между $x$ и $x_0$.

Формулу (12.12) перепишем так:

74

$$f(x) - f(x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}. \qquad (12.13)$$

Поскольку $f^{(n+1)}(x)$ непрерывна в точке $x_0$, то найдется такая окрестность этой точки, в которой эта производная сохраняет знак. Будем считать, что в формуле (12.13) $x$ принадлежит этой окрестности. Предположим, что $(n+1)$ — четное число, тогда $(x - x_0)^{n+1} > 0$. Если $f^{(n+1)}(x_0) < 0$, то $f^{(n+1)}(x) < 0$ для $x$, достаточно близких к точке $x_0$, и $f^{(n+1)}(\xi) < 0$, поэтому из формулы (12.13) следует, что $f(x) - f(x_0) < 0$, т. е. $f(x_0) > f(x)$, $x_0$ — точка максимума. Аналогично доказывается, что $x_0$ — точка минимума, если $f^{(n+1)}(x_0) > 0$.

Если $(n+1)$ — нечетное число, то выражение

$(x-x_0)^{n+1}$ меняет знак при переходе от точки $x_1<x_0$ к точке $x_2>x_0$, поэтому меняет знак и правая часть формулы (12.13). Это означает, что $x_0$ не является точкой экстремума.

Пример 1. Найти точки экстремума функции

$$f(x)=x^5-5x^4+5x^3+4.$$

Первая производная $f'(x)=5x^4-20x^3+15x^2=5x^2(x^2-4x+3)$ обращается в нуль при $x_1=0$, $x_2=1$, $x_3=3$. Вторая производная $f''(x)=20x^3-60x^2+30x$ в этих точках принимает соответственно значения $f''(0)=0$, $f''(1)=-10<0$, $f''(3)=90>0$.

Следовательно, $x_2=1$ — точка максимума, $x_3=3$ — точка минимума. Чтобы исследовать точку $x_1=0$, обратимся к третьей производной $f'''(x)=60x^2-120x+30$. Поскольку $f'''(0)=30\neq 0$, $n+1=3$, то $x_1=0$ не является точкой экстремума.

Пример 2. Найти точки экстремума функции

$$f(x)=x^4+4x^3+6x^2+4x+3.$$

Первая производная $f'(x)=4x^3+12x^2+12x+4=4(x^3+3x^2+3x+1)$ равна нулю в единственной точке $x=-1$. Находим выражения последующих производных и их значения в критической точке $x=-1$: $f''(x)=12x^2+24x+12$, $f''(-1)=0$, $f'''(x)=24x+24$, $f'''(-1)=0$, $f^{IV}(x)=24$. Поскольку $f^{IV}(-1)>0$ и $n+1=4$ (четное число), то $x=-1$ — точка минимума, причем $\min f(x)=f(-1)=2$.

## § 12.6. Направления выпуклости, точки перегиба

График функции $y=f(x)$ называется *выпуклым вниз (вогнутым вверх)* в данном промежутке, если он целиком расположен выше касательной в его произвольной точке (рис. 12.6, *а*).

График функции $y=f(x)$ называется *выпуклым вверх (вогнутым вниз)* в данном промежутке, если он целиком

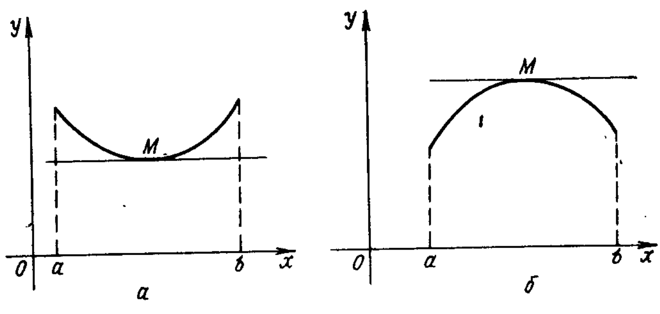



Рис. 12.6

расположен ниже касательной в его произвольной точке (рис. 12.6, *б*).

**Теорема 12.7.** Если вторая производная функции $y=f(x)$ в данном промежутке положительна, то график ее является выпуклым вниз в этом промежутке; если $f''(x)<0$, то график функции является выпуклым вверх в соответствующем промежутке.

Доказательство. Пусть функция $y=f(x)$ такова, что $f''(x)>0$ для всех $x$ из промежутка $(a, b)$. Фиксируем $x_0\in(a, b)$, запишем уравнение касательной в этой точке, обозначив текущую координату через $Y$:

$$Y-f(x_0)=f'(x_0)(x-x_0). \tag{12.14}$$

Разложим данную функцию в окрестности $x_0$ по формуле Тейлора, приняв $n=2$:

$$y=f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(\xi)}{2!}(x-x_0)^2, \tag{12.15}$$

где $\xi$ лежит между $x$ и $x_0$.

Из формул (12.14) и (12.15) получаем

$$y-Y=\frac{f''(\xi)}{2!}(x-x_0)^2. \tag{12.16}$$

Поскольку $(x-x_0)^2>0$ и по условию $f''(\xi)>0$, то $y-Y>0$, или $y>Y$.

Последнее неравенство означает, что график функции $y=f(x)$ лежит выше касательной, проведенной в его произвольной точке $M(x, f(x))$, где $x\in(a, b)$. Следовательно, в промежутке $(a, b)$ график функции $y=f(x)$ является выпуклым вниз (рис. 12.6, *а*).

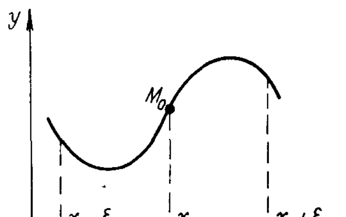


Рис. 12.7

Аналогично доказывается второе утверждение теоремы, геометрическая иллюстрация которого дана на рис. 12.6, *б*.

Точкой перегиба графика функции $y=f(x)$ называется такая его точка $M_0$ (рис. 12.7), в которой выпуклость меняется на вогнутость (по отношению к одному и тому же направлению: вверх или вниз).

**Теорема 12.8.** Если в точке $x=x_0$ вторая производная функции $y=f(x)$ обращается в нуль и меняет знак при переходе через нее, то $M_0(x_0, f(x_0))$ — точка перегиба графика этой функции.

Доказательство. Пусть $f''(x_0)=0$, $f''(x)<0$ при $x_0-\varepsilon<x<x_0 (\varepsilon>0)$ и $f''(x)>0$ при $x_0<x<x_0+\varepsilon$, тогда по теореме 12.7 график функции является выпуклым вверх в промежутке $(x_0-\varepsilon, x_0)$ и выпуклым вниз в промежутке $(x_0, x_0+\varepsilon)$. Следовательно, в точке $M_0(x_0, f(x_0))$ выпуклость вверх меняется на вогнутость вверх, т. е. $M_0$ — точка перегиба, что и требовалось доказать.

Пример. Найти интервалы выпуклости и точки перегиба графика функции $f(x)=x^3-6x^2+9x+1$.

Поскольку вторая производная $f''(x)=6x-12=6(x-2)$ обращается в нуль при $x=2$ и меняет знак при переходе через это значение, то $x=2$ — абсцисса точки перегиба; ордината этой точки $y=f(2)=3$, т. е. $M(2, 3)$ — точка перегиба.

Так как $f''(x)<0$ при $x<2$ и $f''(x)>0$ при $x>2$, то график функции является выпуклым вверх в интервале $]-\infty, 2[$ и выпуклым вниз — в интервале $]-2, +\infty[$.

## § 12.7. Асимптоты

В § 2.4 рассматривались асимптоты гиперболы. Многие другие линии также имеют асимптоты, т. е. прямые, к которым неограниченно приближается данная линия, когда ее точка неограниченно удаляется от начала координат.

По виду уравнений относительно выбранной декартовой системы координат различают асимптоты вертикальные (параллельные оси $Oy$) и наклонные (пересекающие ось $Oy$).

Переходим к определениям и рассмотрению необходимых и достаточных условий, при которых график данной функции имеет асимптоты.

Прямая $x=a$ называется *вертикальной асимптотой* графика функции $y=f(x)$, если хотя бы одно из предельных значений $\lim\limits_{x\to a-0} f(x)$, $\lim\limits_{x\to a+0} f(x)$ является бесконечным. Например, прямая $x=2$ — вертикальная асимптота графика функции $y=\frac{6}{x-2}$, так как $\lim\limits_{x\to 2-0}\frac{6}{x-2}=-\infty$, $\lim\limits_{x\to 2+0}\frac{6}{x-2}=+\infty$.

Предположим, что функция $y=f(x)$ определена при сколь угодно больших (по модулю) значениях аргумента; для определенности будем рассматривать положительные значения аргумента.

Прямая

$$y = kx + b \tag{12.17}$$

называется *наклонной асимптотой* графика функции $y = f(x)$, если эта функция представима в виде

$$f(x) = kx + b + \alpha(x), \tag{12.18}$$

где

$$\lim_{x \to +\infty} \alpha(x) = 0. \tag{12.19}$$

Покажем, что график функции $y = f(x)$ имеет при $x \to +\infty$ наклонную асимптоту (12.17) тогда и только тогда, когда существуют два конечных предела

$$\lim_{x \to +\infty} \frac{f(x)}{x} = k, \quad \lim_{x \to +\infty} [f(x) - kx] = b. \tag{12.20}$$

Пусть график функции $y = f(x)$ имеет асимптоту, определяемую уравнением (12.17), тогда справедливы равенства (12.18) и (12.19). Следовательно,

$$\lim_{x \to +\infty} \frac{f(x)}{x} = \lim_{x \to +\infty} \frac{kx + b + \alpha(x)}{x} =$$

$$= \lim_{x \to +\infty} \left[k + \frac{b}{x} + \frac{\alpha(x)}{x}\right] = k,$$

$$\lim_{x \to +\infty} [f(x) - kx] = \lim_{x \to +\infty} [b + \alpha(x)] = b,$$

т. е. существуют предельные значения (12.20).

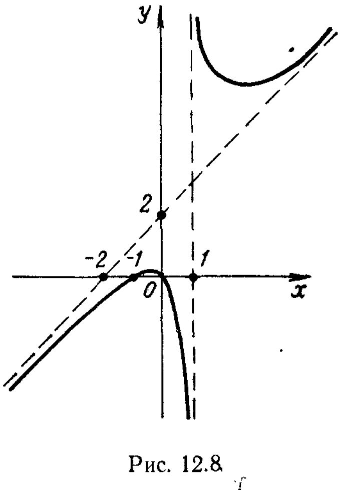


Рис. 12.8

Обратное также верно. Пусть существуют предельные значения (12.20). Второе из них означает, что разность $[f(x) - kx] - b$ является бесконечно малой при $x \to +\infty$. Обозначая эту разность через $\alpha(x)$, получаем $f(x) - kx - b = \alpha(x)$, или $f(x) = kx + b + \alpha(x)$, где $\alpha(x) \to 0$ при $x \to +\infty$. Следовательно, функция $f(x)$ представима в виде (12.18), прямая $y = kx + b$ является асимптотой графика этой функции.

Пример. Найти асимптоты графика функции $f(x) = \frac{x^2+x}{x-1}$.

Прямая $x = 1$ является вертикальной асимптотой (рис. 12.8), так как $\lim\limits_{x\to 1} f(x) = \infty$.

Поскольку $f(x) = \frac{x^2+x}{x-1} = x + 2 + \frac{2}{x-1}$ и $\lim\limits_{x\to\infty} \frac{2}{x-1} = 0$, то график функции имеет и наклонную асимптоту $y = x+2$.

## § 12.8. Исследование функций и построение их графиков

Под «исследованием функций» понимают изучение ее изменения в зависимости от изменения аргумента. На основании исследования функции строят ее график, предварительно изображая характерные точки.

Исследование функций и построение их графиков можно проводить по следующей схеме.

1. Найти область определения функции, ее точки разрыва.

2. Изучить изменение функции при стремлении аргумента к концам промежутков области определения.

3. Найти точки экстремумов, промежутки возрастания и убывания функции.

4. Вычислить значения экстремумов, построить соответствующие точки.

5. Определить промежутки выпуклости и вогнутости графика, найти точки перегиба.

6. Найти точки пересечения графика с координатными осями.

7. Найти асимптоты графика функции.

Порядок исследования иногда целесообразно выбирать, исходя из конкретных особенностей данной функции.

Если рассматриваемая функция четная или нечетная, то ее достаточно исследовать при положительных значениях аргумента из области ее определения и принять во внимание, что график четной функции симметричен относительно оси ординат, график нечетной функции симметричен относительно начала координат.

Отметим также, что графики взаимно-обратных функций симметричны относительно прямой, на которой лежит биссектриса первого координатного угла.

Пример. Исследовать функцию $f(x)=\frac{x^2+1}{x^2-1}$ и построить ее график.

1. Функция не определена лишь при $x=-1$ и $x=1$ (в точках, в которых знаменатель обращается в нуль). Следовательно, область определения состоит из трех интервалов: $]-\infty, -1[$, $]-1, 1[$, $]1, +\infty[$, два из которых являются бесконечными.

2. При стремлении аргумента к концам промежутков области определения соответственно получаем

$$\lim_{x\to-\infty}\frac{x^2+1}{x^2-1}=1,\quad \lim_{x\to-1-0}\frac{x^2+1}{x^2-1}=+\infty,\quad \lim_{x\to-1+0}\frac{x^2+1}{x^2-1}=-\infty,$$

$$\lim_{x\to1-0}\frac{x^2+1}{x^2-1}=-\infty,\quad \lim_{x\to1+0}\frac{x^2+1}{x^2-1}=+\infty,\quad \lim_{x\to+\infty}\frac{x^2+1}{x^2-1}=1.$$

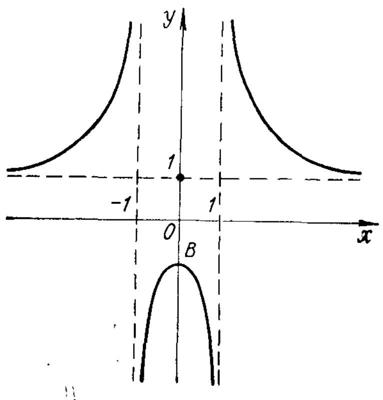



Рис. 12.9

3. Находим производные данной функции:

$$f'(x)=\frac{-4x}{(x^2-1)^2},\quad f''(x)=\frac{4(3x^2+1)}{(x^2-1)^3}.$$

Поскольку $f'(x)>0$ при $x<-1$ и $-1<x<0$, то функция возрастает в интервалах $]-\infty, -1[$ и $]-1, 0[$. Так как $f'(x)<0$ при $0<x<1$ и $x>1$, то функция убывает в интервалах $]0, 1[$ и $]1, +\infty[$.

Поскольку $f'x(0)=0$ при $x_0=0$ и $f''(x_0)=f''(0)<0$, то $x_0=0$ — точка максимума.

Других критических точек нет, ибо $f'(x)$ не определена только при $x=-1$ и $x=1$, но в этих точках не определена и сама функция.

4. Вычисляем значение максимума функции $\max f(x)=f(0)=\frac{0+1}{0-1}=-1$.

5. Поскольку $f''(x)>0$ при $x<-1$ и $x>1$, то график функции является выпуклым вниз в интервалах $]-\infty, -1[$ и $]1, +\infty[$. Так как $f''(x)<0$ при $-1<x<1$, то график функции будет выпуклым вверх в интервале $]-1, 1[$.

Точек перегиба график данной функции не имеет, ибо вторая производная в нуль нигде не обращается и не определена в тех же точках, в которых не определена и сама функция.

6. График функции не пересекает ось $Ox$, так как уравнение $\frac{x^2+1}{x^2-1}=0$ не имеет действительных корней. Если $x=0$ (уравнение оси $Oy$), то $y=-1$, в точке $B(0, -1)$ график пересекает ось $Oy$.

7. Из пункта 2 следует, что график функции имеет две вертикальные асимптоты $x=-1$ и $x=1$ и горизонтальную асимптоту $y=1$. Последнее вытекает также из того, что $\frac{x^2+1}{x^2-1}=1+\frac{2}{x^2-1}$ и $\lim\limits_{x\to\infty}\frac{2}{x^2-1}=0$.

Заметив еще, что $f(x)>0$ при $x<-1$ и $x>1$, $f(x)<0$ при $-1<x<1$, строим график функции (рис. 12.9).

## § 12.9. Задачи на наибольшие и наименьшие значения

Наибольшим значением функции $y=f(x)$ на отрезке $[a, b]$ называют такое ее значение, которое больше всех других значений, принимаемых функцией на данном отрезке. Чтобы найти наибольшее значение функции на отрезке, необходимо вычислить значения максимумов на этом отрезке и значения функции на концах отрезка, из полученных чисел выбрать самое большое. Аналогично определяется и разыскивается наименьшее значение функции.

В математике, физике, химии, технических и других науках, а также в повседневной жизни часто встречаются задачи на отыскание наибольших и наименьших значений некоторых функций.

Общая схема решения таких задач состоит в следующем. Сначала устанавливается зависимость рассматриваемой величины $y$ от некоторой независимой переменной величины $x$ (обозначения, разумеется, могут быть другими). Из условия задачи определяется промежуток, в котором может изменяться аргумент функции. Функция $y=f(x)$ исследуется с помощью теории, рассмотренной в предыдущих главах.

Пример 1. Найти наименьшее значение суммы двух положительных чисел, произведение которых постоянно и равно $a$.

Обозначим искомые числа через $x$ и $y$. По условию $xy=a$, где $a>0$, поэтому $y=\frac{a}{x}$. Сумма этих чисел $S=x+y$, $S(x)=x+$

$+\frac{a}{x}$ является функцией переменной $x$; в соответствии с условием $x > 0$. Находим производные функции $S(x): S'(x) = 1 - \frac{a}{x^2}$, $S''(x) = \frac{2a}{x^3}$. Приравнивая нулю первую производиую, получаем уравнение $1 - \frac{a}{x^2} = 0$, из которого находим критические точки $x_1 = -\sqrt{a}$, $x_2 = \sqrt{a}$; первое значение ие принадлежит области изменения аргумента данной функции. Поскольку $S''(\sqrt{a}) > 0$, то $x = \sqrt{a}$ — точка минимума, причем $\min S(x) = S(\sqrt{a}) = 2\sqrt{a}$.

Пример 2. Газовая смесь состоит из окиси азота и кислорода. Найти концентрацию кислорода, при которой содержащаяся в смеси окись азота окисляется с максимальной скоростью (см. [11]).

В условиях практической иеобратимости скорость реакции $2NO + O_2 = 2NO_2$ выражается формулой $v = cx^2y$, где $x$ — концентрация NO в любой момент времени; $y$ — концентрация $O_2$; $c$ — константа скорости реакции, не завнсящая от концентрации реагирующих компонентов и зависящая только от температуры ($c > 0$).

Выражая концентрации газов в объемных процентах, получаем $y = 100 - x$, $v = cx^2(100 - x) = c(100x^2 - x^3)$. Исследуем функцию $v = v(x)$. Находим ее производные: $v'(x) = c(200x - 3x^2)$, $v''(x) = c(200 - 6x)$. Приравнивая нулю первую производную, получаем уравнение $c(200x - 3x^2) = 0$, из которого находим критические точки $x_1 = 0$, $x_2 = \frac{200}{3}$. Поскольку $v''(0) > 0$, $v''\left(\frac{200}{3}\right) < 0$, то $x_1 = 0$ — точка минимума, $x_2 = \frac{200}{3}$ — точка максимума.

Таким образом, при $x = \frac{200}{3} = 66,7\%$ $y = 100 - x = 33,3\%$ или при $y : x = 0,5$ скорость окисления имеет максимальное значение.

# Глава 13

# КРИВИЗНА

В главе рассматривается вопрос об исследовании кривизны линии с помощью производных. Кривизна является важной характеристикой линии. Крнвизна линии — это мера изогнутости данной линии.

## § 13.1. Эквивалентность бесконечно малой дуги и стягивающей хорды. Дифференциал дуги плоской кривой

Пусть на отрезке $[a, b]$ задана дифференцируемая функция $y = f(x)$, графиком которой является дуга $\overset{\frown}{AB}$ (рис. 13.1). Отрезок $[a, b]$ разобьем на $n$ частей точками

$x_1, x_2, \ldots, x_{n-1}$. Этим точкам будут соответствовать точки $M_1, M_2, \ldots, M_{n-1}$ дуги $\smile{AB}$, соединим их ломаной линией, которую называют ломаной, *вписанной* в дугу $\smile{AB}$. Периметр данной ломаной обозначим через $s_n$, т. е.

$$s_n = \sum_{k=1}^{n} |M_{k-1} M_k| \, (M_0 = A, \ M_n = B).$$

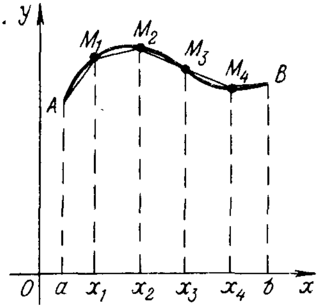



Рис. 13.1

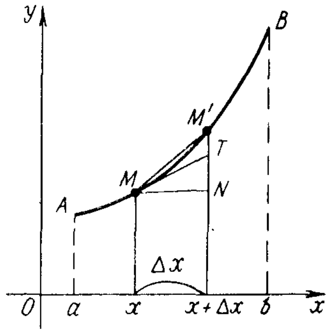



Рис. 13.2

*Длиной дуги* линии называется предел периметра вписанной в нее ломаной, когда число звеньев $M_{k-1}M_k$ неограниченно возрастает, а длина наибольшего из них стремится к нулю.

Обозначив длину дуги через $s$, получим

$$s = \lim_{n\to\infty} s_n = \lim_{\lambda\to 0} \sum_{k=1}^{n} |M_{k-1} M_k|,$$

где $\lambda$ — длина наибольшего звена.

Будем отсчитывать длину дуги от некоторой ее точки, например $A$. Пусть в точке $M(x, y)$ длина дуги $\smile{AM}$ равна $s$, а в точке $M'(x+\Delta x, y+\Delta y)$ длина дуги $\smile{AM'}$ равна $s+\Delta s$, где $\Delta s$ — длина дуги $\smile{MM'}$ (рис. 13.2). Из $\Delta MNM'$ находим длину хорды $MM'$: $MM' = \sqrt{\Delta x^2 + \Delta y^2}$.

Из геометрических соображений следует, что

$$\lim_{M'\to M} \frac{\smile{MM'}}{MM'} = 1, \ \lim_{\Delta x\to 0} \frac{\Delta s}{\sqrt{\Delta x^2 + \Delta y^2}} = 1, \qquad (13.1)$$

т. е. бесконечно малая дуга линии и стягивающая её хорда эквивалентны.

Преобразуем формулу, выражающую длину хорды $MM'$:

$$\frac{MM'}{\Delta s}\,\frac{\Delta s}{\Delta x}=\sqrt{1+\left(\frac{\Delta y}{\Delta x}\right)^2}.$$

Переходя к пределу в этом равенстве, получаем формулу для производной функции $s=s(x)$:

$$\frac{ds}{dx}=\sqrt{1+\left(\frac{dy}{dx}\right)^2},$$

из которой находим

$$ds=\sqrt{dx^2+dy^2}. \tag{13.2}$$

Формула (13.2) выражает дифференциал дуги плоской кривой и имеет простой геометрический смысл: выражает теорему Пифагора для бесконечно малого треугольника $MTN$ (рис. 13.2, $ds=MT$, $\Delta s=\smile{MM'}$). Дифференциал дуги пространственной кривой определяется формулой

$$ds=\sqrt{dx^2+dy^2+dz^2}. \tag{13.3}$$

## § 13.2. Понятие кривизны плоской линии

Рассмотрим плоскую линию, определяемую уравнением $y=f(x)$. Проведем касательную к этой линии в точке $M_0(x_0,\ y_0)$. Обозначим через $\alpha$ угол, образованный касательной с осью $Ox$ (рис. 13.3). Пусть касательная в точке $M$ образует с осью $Ox$ угол $\alpha+\Delta\alpha$.

Угол $\Delta\alpha$ между касательными в указанных точках называют *углом смежности*. Можно сказать, что при переходе из точки $M_0$ в точку $M$ данной линии касательная к ней повернулась на угол $\Delta\alpha$, которому будем приписывать соответствующий знак в зависимости от направления поворота. Величина угла поворота дает некоторое представление о степени изогнутости линии на дуге $\smile{M_0M}$. Однако сама по себе величина угла $\Delta\alpha$ еще не может служить мерой этой изогнутости; важно то, какая доля угла приходится в среднем на единицу длины дуги $\overset{\frown}{M_0M}$.

*Средней кривизной* дуги $\overset{\frown}{M_0M}$ данной линии называется абсолютное значение отношения угла смежности $\Delta\alpha$ к длине $\Delta s$ дуги $\overset{\frown}{M_0M}$:

$$K_{\text{ср}} = \left|\frac{\Delta\alpha}{\Delta s}\right|. \tag{13.4}$$

Отметим, что разные дуги одной и той же линии могут иметь разную среднюю кривизну (на рис. 13.4 сред-

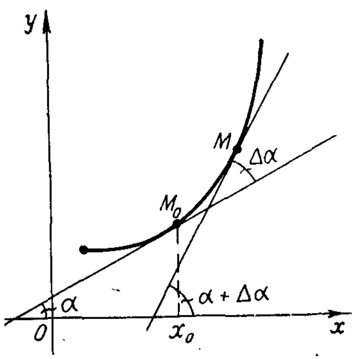



Рис. 13.3

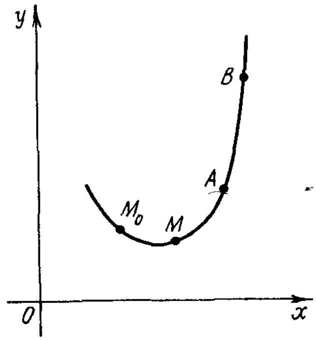



Рис. 13.4

няя кривизна дуги $\overset{\frown}{M_0M}$ значительно больше, чем дуги $\overset{\frown}{AB}$). Желая охарактеризовать степень изогнутости линии в данной ее точке, вводят понятие кривизны линии в этой точке.

*Кривизной линии* в данной точке $M_0$ называется предел средней кривизны дуги $\overset{\frown}{M_0M}$ при $M \to M_0$:

$$K = \lim_{M \to M_0}\left|\frac{\Delta\alpha}{\Delta s}\right|, \quad K = \lim_{\Delta s \to 0}\left|\frac{\Delta\alpha}{\Delta s}\right|. \tag{13.5}$$

Найдем кривизну прямой линии и окружности. В любой точке прямой касательная к ней совпадает с самой прямой, поэтому $\Delta\alpha = 0$, средняя кривизна (и кривизна) равна нулю в любой ее точке.

В случае окружности угол смежности $\Delta\alpha$ равен углу между ее радиусами $OM_0$ и $OM$ (рис. 13.5), длина дуги

$\overset{\smile}{MM_0}$ выражается формулой $\Delta s = R\Delta\alpha$, где $R$ — радиус данной окружности. Следовательно,

$$K_{\text{ср}} = \frac{\Delta\alpha}{\Delta s} = \frac{1}{R}, \quad K = \frac{1}{R},$$

т. е. кривизна окружности постоянна и равна величине, обратной радиусу этой окружности.

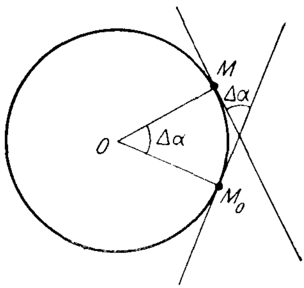



Рис. 13.5

## § 13.3. Формула кривизны

Выведем формулу для вычисления кривизны линии, заданной уравнением $y=f(x)$, в ее фиксированной точке $M_0(x_0, y_0)$. Предположим, что функция $y=f(x)$ имеет первые две производные в точке $x_0$. Обозначим через $s$ длину дуги $\overset{\smile}{AM}$ графика данной функции, ограниченной точками $A(a, b)$, $M(x, y)$. Длина дуги является функцией абсциссы $x$ конца этой дуги, т. е. $s=s(x)$. Угол наклона касательной к графику $y=f(x)$ является функцией абсциссы $x$ точки касания $M(x, y)$, т. е. $\alpha=\alpha(x)$. Будем считать, что функции $s(x)$ и $\alpha(x)$ дифференцируемы.

Преобразуем правую часть формулы (13.5). Так как

$$\frac{\Delta\alpha}{\Delta s} = \frac{\Delta\alpha}{\Delta x} : \frac{\Delta s}{\Delta x}, \quad \lim_{\Delta s\to 0} \frac{\Delta\alpha}{\Delta s} =$$

$$= \lim_{\Delta x\to 0} \frac{\Delta\alpha}{\Delta x} : \lim_{\Delta x\to 0} \frac{\Delta s}{\Delta x} = \alpha'_x : s'_x = \frac{\alpha'_x}{s'_x}, \quad \lim_{\Delta s\to 0} \frac{\Delta\alpha}{\Delta s} = \frac{\alpha'_x}{s'_x},$$

то формула (13.5) принимает вид

$$K = \left|\frac{\alpha'_x}{s'}\right|. \tag{13.6}$$

Поскольку $\operatorname{tg}\alpha = y'_x$, $\alpha = \operatorname{arctg} y'_x$, то

$$\alpha'_x = \frac{y''_x}{1 + y'^2_x}. \tag{13.7}$$

Подставляя выражения (13.2) и (13.7) в равенство (13.6), получаем искомую формулу для кривизны линии $y = f(x)$ в точке $M_0$:

$$K = \frac{|y''_x|}{(1 + y'^2_x)^{3/2}}. \tag{13.8}$$

Обращаем внимание на то, что формула (13.8) содержит значения первой и второй производной данной функции, вычисленные при $x = x_0$.

Пример. Найти кривизну косинусоиды $y = \cos x$ в точке $M_0(0, 1)$.

Поскольку $y' = -\sin x$, $y'' = -\cos x$, кривизна косинусоиды в ее произвольной точке определяется формулой

$$K = \frac{|-\cos x|}{(1 + \cos^2 x)^{3/2}}.$$

При $x = 0$ получаем $K = \dfrac{1}{(1+1)^{3/2}} = \dfrac{\sqrt{2}}{4}$.

Замечание 1. Если линия задана параметрическими уравнениями

$$x = f_1(t), \quad y = f_2(t), \tag{13.9}$$

то

$$y'_x = \frac{y'_t}{x'_t}, \quad y''_x = \frac{x'_t y''_t - x''_t y'_t}{x'^3_t}, \tag{13.10}$$

поэтому формула (13.8) принимает вид

$$K = \frac{|x'_t y''_t - x''_t y'_t|}{(x'^2_t + y'^2_t)^{3/2}}. \tag{13.11}$$

Замечание 2. Если линия задана уравнением

$$\rho = \rho(\varphi) \tag{13.12}$$

в полярных координатах, то ее параметрические уравнения примут вид

$$x = \rho(\varphi)\cos\varphi, \quad y = \rho(\varphi)\sin\varphi, \tag{13.13}$$

где роль параметра играет полярный угол $\varphi$.

Вычислив величины, входящие в формулу (13.11), получим

$$K = \frac{|\rho^2 + 2\rho'^2 - \rho\rho''|}{(\rho'^2 + \rho^2)^{3/2}}. \tag{13.14}$$

## § 13.4. Окружность кривизны. Центр и радиус кривизны

*Радиусом кривизны* данной линии в данной ее точке называется величина $R$, обратная кривизне $K$ этой линии в рассматриваемой точке:

$$R = \frac{1}{K}. \tag{13.15}$$

Принимая во внимание формулу (13.8), получаем следующее выражение для радиуса кривизны линии $y=f(x)$ в точке $M(x, y)$:

$$R = \frac{[1+(y'_x)^2]^{3/2}}{|y''_x|}. \tag{13.16}$$

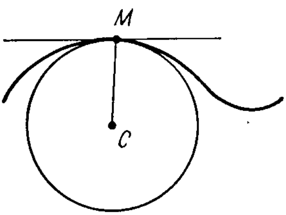


Рис. 13.6

На нормали к кривой в точке $M$ отложим отрезок $MC=R$ в сторону вогнутости кривой (рис. 13.6). Точка $C$ называется *центром кривизны* данной линии в точке $M$. Окружность радиуса $R$ с центром в точке $C$ называется *окружностью кривизны* этой линии в точке $M$.

Очевидно, в данной точке $M$ кривизна кривой и кривизна окружности кривизны равны между собой.

Выведем формулы для координат центра кривизны, обозначив их через $\alpha$ и $\beta$.

Уравнение нормали к линии $y=f(x)$ в точке $M(x, y)$ имеет вид

$$Y - y = -\frac{1}{y'}(X - x), \tag{13.17}$$

где $X$, $Y$ — текущие координаты точки нормали; $x$, $y$ — координаты точки $M$.

Поскольку точка $C(\alpha, \beta)$ лежит на нормали, ее координаты удовлетворяют уравнению (13.17)

$$\beta - y = -\frac{1}{y'}(\alpha - x). \tag{13.18}$$

Точка $C(\alpha, \beta)$ удалена от точки $M$ на расстояние, равное радиусу кривизны $R$, поэтому

$$(\alpha-x)^2+(\beta-y)^2=R^2. \tag{13.19}$$

Из уравнений (13.18) и (13.19) находим

$$(\alpha - x)^2 + \frac{1}{y'^2}(\alpha - x)^2 = R^2, \quad (\alpha - x)^2 = \frac{y'^2}{1 + y'^2} R^2,$$

откуда $\alpha = x \pm \frac{y'}{\sqrt{1 + y'^2}} R$, поэтому

$$\beta = y \mp \frac{1}{\sqrt{1 + y'^2}} R.$$

Подставив в эти формулы равенство (13.16), получим

$$\alpha = x \pm \frac{y'(1 + y'^2)}{|y''|}, \quad \beta = y \mp \frac{1 + y'^2}{|y''|}.$$

Для решения вопроса о том, какие знаки нужно взять в последних формулах, необходимо рассмотреть отдельно случаи, когда $y'' > 0$ и $y'' < 0$. Если $y'' > 0$, то в соответствующей точке кривая вогнута вверх. Следовательно, $\beta > y$ и необходимо взять нижние знаки, т. е. координаты центра кривизны определяются формулами

$$\alpha = x - \frac{y'(1 + y'^2)}{y''}, \quad \beta = y + \frac{1 + y'^2}{y''}. \tag{13.20}$$

Эти формулы остаются справедливыми и в случае $y'' < 0$; читателю предлагается убедиться в этом самостоятельно.

Если кривая задана параметрическими уравнениями (13.9), формулы (13.20) с учетом равенств (13.10) примут вид

$$\alpha = x - \frac{y'(x'^2 + y'^2)}{x'y'' - x''y'}, \quad \beta = y + \frac{x'(x'^2 + y'^2)}{x'y'' - x''y'}. \tag{13.21}$$

Множество всех центров кривизны данной линии называется ее *эволютой*. По отношению к своей эволюте исходная линия называется *эвольвентой* (или *разверткой*).

Если линия задана уравнением $y = f(x)$, уравнения (13.20) можно рассматривать как параметрические уравнения ее эволюты (с параметром $x$).

В случае параметрического задания кривой уравнения (13.21) являются параметрическими уравнениями эволюты (входящие в правые части этих уравнений величины зависят от параметра $t$).

## § 13.5. Переменная векторная величина. Вектор-функция скалярного аргумента

Рассмотрим точку $M(x, y, z)$, движущуюся по некоторой линии $\gamma$ в пространстве (рис. 13.7). Радиус-вектор $\mathbf{r}=\overline{OM}$ точки $M$ будет иметь определенное направление и длину в фиксированный момент времени $t$. С течением времени направление и длина вектора $\overline{OM}$ будут изменяться [1].

Таким образом, здесь имеем дело с переменным вектором $\overline{OM}$ или с переменной векторной величиной

$$\mathbf{r}=\mathbf{r}(t), \tag{13.22}$$

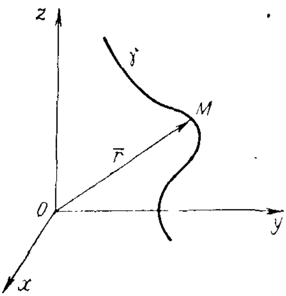



Рис. 13.7

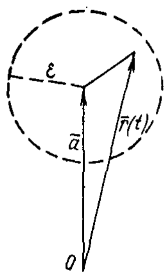



Рис. 13.8

зависящей от времени $t$. Равенство (13.22) называется *векторным уравнением* движения точки $M$.

Координаты переменного вектора $\overline{OM}=\mathbf{r}(t)=\{x(t),\ y(t),\ z(t)\}$ являются также переменными величинами (скалярными), зависящими от времени $t$:

$$x=x(t),\quad y=y(t),\quad z=z(t). \tag{13.23}$$

Уравнения (13.23) являются параметрическими уравнениями рассматриваемой линии $\gamma$.

Переменная векторная величина $\mathbf{u}$ называется *вектор-функцией* (или *векторной функцией*) скалярного аргумента $t$, если каждому значению $t_0 \in T$, где $T$ — некоторый интервал, соответствует определенный вектор $\mathbf{u}(t_0)$; в этом случае пишут

$$\mathbf{u}=\mathbf{u}(t). \tag{13.24}$$

Если $\mathbf{u}=\mathbf{u}(t)$, то и проекции $u_x$, $u_y$, $u_z$ переменного вектора $\mathbf{u}$ на оси декартовой системы координат будут (скалярными) функциями аргумента $t$:

$$u_x=u_x(t),\ u_y=u_y(t),\ u_z=u_z(t). \tag{13.25}$$

---

[1] Направление не меняется, когда $\gamma$ есть луч, исходящий из начала координат. Когда линия лежит на поверхности шара, длина не меняется. Но для двух различных значений $t_1$, $t_2$ и в этих случаях векторы $\mathbf{r}(t_1)$, $\mathbf{r}(t_2)$ будут различными.

Пример вектор-функции скалярного аргумента дает рассмотренный выше случай радиуса-вектора $\mathbf{r}=\mathbf{r}(t)$ точки, движущейся по некоторой линии в пространстве.

*Годографом* переменной векторной величины называется геометрическое место концов всех ее отдельных значений при условии, что все они отложены из одной точки. Годографом постоянного вектора является точка (конец вектора). Годограф вектор-функции $\mathbf{u}=\mathbf{u}(t)$ представляет собой некоторую линию. Если вектор $\mathbf{u}=\mathbf{u}(t)$ сохраняет постоянную длину, то ее годограф — линия, лежащая на сфере. Годографом радиуса-вектора $r=\overline{OM}$ движущейся точки $M$ является траектория этой точки.

Введем понятия предела и непрерывности для вектор-функции. Пусть $\mathbf{a}$ — постоянный вектор и $\mathbf{r}=\mathbf{r}(t)$ — вектор-функция, определенная в некотором интервале, содержащем точку $t_0$.

Вектор $\mathbf{a}$ называется *пределом вектор-функции* $\mathbf{r}=\mathbf{r}(t)$ при $t\to t_0$, если для любого $\varepsilon>0$ существует такое $\delta=\delta(\varepsilon)>0$, что $|\mathbf{r}(t)-\mathbf{a}|<\varepsilon$ для всех $t$, удовлетворяющих неравенству $|t-t_0|<\delta$, $t\neq t_0$ (рис. 13.8). Обозначения предела вектор-функции:

$$\lim_{t\to t_0}\mathbf{r}(t)=\mathbf{a}; \tag{13.26}$$

$$\mathbf{r}(t)\to\mathbf{a} \text{ при } t\to t_0. \tag{13.27}$$

Очевидно, равенство (13.26) эквивалентно равенству

$$\lim_{t\to t_0}|\mathbf{r}(t)-\mathbf{a}|=0. \tag{13.28}$$

Если $\mathbf{r}(t)=\{x(t),\ y(t),\ z(t)\}$ и $\mathbf{a}=(a_1,\ a_2,\ a_3)$, то равенство (13.26) выполняется тогда и только тогда, когда

$$\lim_{t\to t_0}x(t)=a_1,\quad \lim_{t\to t_0}y(t)=a_2,\quad \lim_{t\to t_0}z(t)=a_3. \tag{13.29}$$

Действительно, так как

$$|\mathbf{r}(t)-\mathbf{a}|=\sqrt{[x(t)-a_1]^2+[y(t)-a_2]^2+[z(t)-a_3]^2}, \tag{13.30}$$

то

$$|\mathbf{r}(t)-\mathbf{a}|\geqslant|x(t)-a_1|.$$

Из последнего соотношения следует, что условие $|\mathbf{r}(t)-\mathbf{a}|\to 0$ при $t\to t_0$ влечет за собой условие $|x(t)-a_1|\to 0$ при $t\to t_0$, т. е. первое из равенств (13.29). Аналогично доказываются два других равенства (13.29). Обратно, если выполнены равенства (13.29), то из (13.30) получаем, что $|\mathbf{r}(t)-\mathbf{a}|\to 0$ при $t\to t_0$, т. е. $\lim\limits_{t\to t_0}\mathbf{r}(t)=\mathbf{a}$.

Если $\lim\limits_{t\to t_0}\mathbf{r}(t)=\mathbf{a}$, то $\lim\limits_{t\to t_0}|\mathbf{r}(t)|=|\mathbf{a}|$; это следует из неравенства $||\mathbf{r}|-|\mathbf{a}||\leqslant|\mathbf{r}-\mathbf{a}|$ и формулы (13.28).

Если вектор-функции $\mathbf{r}_1(t)$ и $\mathbf{r}_2(t)$ определены в некоторой окрестности точки $t_0$, существуют пределы

$$\lim_{t\to t_0}\mathbf{r}_1(t)=\mathbf{a},\quad \lim_{t\to t_0}\mathbf{r}_2(t)=\mathbf{b}, \tag{13.31}$$

скалярная функция $f(t)$ имеет предел при $t\to t_0$, то существуют также пределы

$$\lim_{t\to t_0}[\mathbf{r}_1(t)+\mathbf{r}_2(t)]=\mathbf{a}+\mathbf{b}; \tag{13.32}$$

$$\lim_{t\to t_0} f(t)\mathbf{r}_1(t)=\lim_{t\to t_0} f(t) \lim_{t\to t_0}\mathbf{r}_1(t); \tag{13.33}$$

$$\lim_{t\to t_0}\mathbf{r}_1(t)\mathbf{r}_2(t)=\mathbf{ab}; \tag{13.34}$$

$$\lim_{t\to t_0}[\mathbf{r}_1(t),\ \mathbf{r}_2(t)]=[\mathbf{a},\ \mathbf{b}]. \tag{13.35}$$

Эти свойства пределов вектор-функций, в частности, можно доказать с помощью формул (13.29) и соответствующих свойств скалярных функций, если перейти от векторных соотношений к координатным. Например, если $\mathbf{r}_1(t)=\{x_1(t),\ y_1(t),\ z_1(t)\}$, $\mathbf{r}_2(t)=\{x_2(t),\ y_2(t),\ z_2(t)\}$, $\mathbf{a}=(a_1,\ a_2,\ a_3)$, $\mathbf{b}=(b_1,\ b_2,\ b_3)$, то

$$\mathbf{r}_1(t)\mathbf{r}_2(t)=x_1(t)x_2(t)+y_1(t)y_2(t)+z_1(t)z_2(t),$$
$$\lim_{t\to t_0}\mathbf{r}_1(t)\cdot\mathbf{r}_2(t)=a_1b_1+a_2b_2+a_3b_3=\mathbf{ab}. \tag{13.36}$$

Вектор-функция $\mathbf{r}=\mathbf{r}(t)$, определенная в точке $x_0$ и некоторой ее окрестности, называется *непрерывной* в этой точке, если

$$\lim_{t\to t_0}\mathbf{r}(t)=\mathbf{r}(t_0). \tag{13.37}$$

Из эквивалентности условий (13.26) и (13.28) следует, что вектор-функция $\mathbf{r}(t)=\{x(t),\ y(t),\ z(t)\}$ непрерывна в точке $t_0$ тогда и только тогда, когда непрерывны в ней функции $x(t)$, $y(t)$, $z(t)$.

## § 13.6. Дифференцирование вектор-функций

Рассмотрим вектор-функцию $\mathbf{u}=\mathbf{u}(v)$ скалярного аргумента $v$. Зафиксируем значение аргумента $v$, найдем соответствующее значение $\mathbf{u}(v)$. Придадим аргументу приращение $\Delta v$ и найдем новое значение вектор-функции $\mathbf{u}(v+\Delta v)$, а также ее приращение $\Delta\mathbf{u}=\mathbf{u}(v+\Delta v)-\mathbf{u}(v)$. Составим отношение

$$\frac{\Delta\mathbf{u}}{\Delta v}=\frac{\mathbf{u}(v+\Delta v)-\mathbf{u}(v)}{\Delta v} \tag{13.38}$$

и перейдем к пределу при $\Delta v\to 0$ (разумеется, не для всякой вектор-функции $\mathbf{u}=\mathbf{u}(v)$ и не для любого значения $v$ такой предел существует).

Предел отношения приращения вектор-функции к приращению аргумента, когда последнее стремится к нулю, называется *производной вектор-функции* $\mathbf{u}=\mathbf{u}(v)$ в точке $v$. Производная вектор-функции обозначается одним из символов

$$\mathbf{u}',\ \mathbf{u}'(v),\ \frac{d\mathbf{u}}{dv}.$$

Итак, по определению

$$\mathbf{u}'(v)=\lim_{\Delta v\to 0}\frac{\mathbf{u}(v+\Delta v)-\mathbf{u}(v)}{\Delta v}. \tag{13.39}$$

Выясним геометрический смысл производной $\mathbf{u}'(v)\neq 0$. Обратимся к годографу вектор-функции (рис. 13.9). Если $\mathbf{u}(v)=\overline{OM}$,

$\mathbf{u}(v+\Delta v)=\overline{ON}$, то $\Delta\mathbf{u}=\overline{MN}$. Отношение $\Delta\mathbf{u}/\Delta v$ есть вектор, имеющий направление вектора $\overline{MN}$, если $\Delta v>0$, или противоположное ему направление, если $\Delta v<0$ (на рис. 13.9 изображен случай $\Delta v>0$). Вектор $\overline{MP}=\Delta\mathbf{u}/\Delta v$ направлен по секущей к годографу. Если $\Delta v\to 0$, то $N\to M$, поэтому предельное положение вектора $\overline{MP}$ направлено по касательной $MQ$ к годографу в точке $M$, а это предельное положение и является производной $\mathbf{u}'(v)$. Следовательно, производная вектор-функции в данной точке есть вектор, направленный по касательной к годографу данной вектор-функции в соответствующей точке.

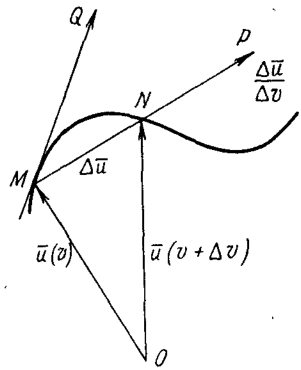


Рис. 13.9

Отметим, что при другом значении $v$ получим новое значение $\mathbf{u}'(v)$, т. е. производная вектор-функции также является вектор-функцией. Вектор-функция, имеющая производную, называется *дифференцируемой.*

Дифференциалом вектор-функции $\mathbf{u}=\mathbf{u}(v)$ называется произведение ее производной на дифференциал аргумента

$$d\mathbf{u}=\mathbf{u}'(v)\,dv, \tag{13.40}$$

где $dv=\Delta v$.

Из формулы (13.40) следует, что

$$\mathbf{u}'(v)=\frac{d\mathbf{u}}{dv}. \tag{13.41}$$

Правая часть последнего равенства представляет собой отношение двух дифференциалов.

Необходимым и достаточным условием существования производной вектор-функции

$$\mathbf{u}(v)=\{x(v),\ y(v),\ z(v)\} \tag{13.42}$$

в некоторой точке является дифференцируемость функций $x(v)$, $y(v)$, $z(v)$ в этой точке, причем в данном случае

$$\mathbf{u}'(v)=\{x'(v),\ y'(v),\ z'(v)\}. \tag{13.43}$$

Это утверждение следует из эквивалентности двух подходов (13.26) и (13.28) к определению предела вектор-функции

$$\frac{\mathbf{u}(v+\Delta v)-\mathbf{u}(v)}{\Delta v}=$$

$$=\left\{\frac{x(v+\Delta v)-x(v)}{\Delta v},\ \frac{y(v+\Delta v)-y(v)}{\Delta v},\ \frac{z(v+\Delta v)-z(v)}{\Delta v}\right\}.$$

Правила дифференцирования вектор-функции аналогичны правилам обычного дифференциального исчисления. Если $\mathbf{u}_1=\mathbf{u}_1(v)$, $\mathbf{u}_2=\mathbf{u}_2(v)$ — дифференцируемые вектор-функции скалярного аргумента $v$; $\mathbf{c}$ — постоянный вектор; $f(v)$ — дифференцируемая скалярная функция; $k$ — постоянная скалярная величина; $w$ — скалярный аргумент, связанный с $v$ формулой $w=w(v)$, где $w(v)$ — дифференцируемая функция, то правила дифференцирования выражаются следующими формулами:

$$1)\ \frac{dc}{dv}=0;\quad 2)\ \frac{d(\mathbf{u}_1\pm\mathbf{u}_2)'}{dv}=\frac{d\mathbf{u}_1}{dv}\pm\frac{d\mathbf{u}_2}{dv};$$

$$3)\ \frac{d(f\mathbf{u}_1)}{dv}=\frac{df}{dv}\mathbf{u}_1+f\frac{d\mathbf{u}_1}{dv};$$

$$3а)\ \frac{d(k\mathbf{u}_1)}{dv}=k\frac{d\mathbf{u}_1}{dv};\quad 3б)\ \frac{d(f\mathbf{c})}{dv}=\frac{df}{dv}\mathbf{c};$$

$$4)\ \frac{d(\mathbf{u}_1\mathbf{u}_2)}{dv}=\frac{d\mathbf{u}_1}{dv}\mathbf{u}_2+\mathbf{u}_1\frac{d\mathbf{u}_2}{dv};\quad 4а)\ \frac{d(\mathbf{c}\mathbf{u}_1)}{dv}=\mathbf{c}\frac{d\mathbf{u}_1}{dv};$$

$$5)\ \frac{d[\mathbf{u}_1,\mathbf{u}_2]}{dv}=\left[\frac{d\mathbf{u}_1}{dv},\ \mathbf{u}_2\right]+\left[\mathbf{u}_1,\ \frac{d\mathbf{u}_2}{dv}\right];$$

$$6)\ \frac{d\mathbf{u}_1}{dv}=\frac{d\mathbf{u}_1}{dw}\ \frac{dw}{dv}.$$

Докажем, например, формулу для дифференцирования скалярного произведения двух вектор-функций, т. е. формулу 4).

Пусть

$$\mathbf{u}_1(v)=\{x_1(v),\ y_1(v),\ z_1(v)\},\quad \mathbf{u}_2(v)=\{x_2(v),\ y_2(v),\ z_2(v)\},$$

тогда

$$\mathbf{u}_1(v)\mathbf{u}_2(v)=x_1(v)x_2(v)+y_1(v)y_2(v)+z_1(v)z_2(v).$$

Следовательно,

$$\frac{d(\mathbf{u}_1\mathbf{u}_2)}{dv}=x_1'(v)\,x_2(v)+x_1(v)\,x_2'(v)+y_1'(v)\,y_2(v)+$$
$$+y_1(v)\,y_2'(v)+z_1'(v)\,z_2(v)+z_1(v)\,z_2'(v)=(x_1'x_2+y_1'y_2+z_1'z_2)+$$
$$+(x_1x_2'+y_1y_2'+z_1z_2')=\frac{d\mathbf{u}_1}{dv}\mathbf{u}_2+\mathbf{u}_1\frac{d\mathbf{u}_2}{dv}.$$

С помощью этой формулы легко доказывается следующая лемма.

Лемма. Если модуль вектор-функции $\mathbf{u}=\mathbf{u}(v)$ сохраняет постоянное значение, то производная $\mathbf{u}'=\mathbf{u}'(v)$ перпендикулярна ей, короче, если $\mathbf{u}(v)=\mathrm{const}$, то $\mathbf{u}'\perp\mathbf{u}$.

Действительно, $\mathbf{u}\mathbf{u}=\mathbf{u}^2=|\mathbf{u}|^2=k$, $\mathbf{u}\mathbf{u}=k$, где $k$ — постоянная. Дифференцируя, получаем $\mathbf{u}'\mathbf{u}+\mathbf{u}\mathbf{u}'=0$, $2\mathbf{u}'\mathbf{u}=0$, $\mathbf{u}'\perp\mathbf{u}$.

Пусть

$$\mathbf{r}=\mathbf{r}(t),\quad \mathbf{r}(t)=x(t)\mathbf{i}+y(t)\mathbf{j}+z(t)\mathbf{k}\tag{13.44}$$

— векторное уравнение движения точки $M$ в пространстве. Приращению $\Delta t$ времени $t$ соответствует приращение $\Delta\mathbf{r}=\overline{M_0M}$ вектор-функции $\mathbf{r}=\mathbf{r}(t)$. Отношение $\Delta\mathbf{r}/\Delta t$ называется *вектором средней скорости*, этот вектор направлен по прямой $M_0M$. Предел указанного отношения при $\Delta t\to 0$ называется *вектором скорости* в момент $t_0$ (или *вектором мгновенной скорости*), обозначим его через $\mathbf{v}$, т. е.

$$\mathbf{v}=\mathbf{r}'(t)=\frac{d\mathbf{r}}{dt}.\tag{13.45}$$

Следовательно, вектор мгновенной скорости движущейся точки направлен по касательной к ее траектории. Вектор $\mathbf{r}'(t)$ характеризует направление и быстроту движения точки. В самом деле, полагая для определенности $\Delta t > 0$, получаем

$$|\mathbf{r}'(t)| = \lim_{\Delta t \to 0} \frac{|\mathbf{r}(t+\Delta t) - \mathbf{r}(t)|}{\Delta t}.$$

Поскольку $|\mathbf{r}(t+\Delta t) - \mathbf{r}(t)|$ выражает длину хорды $M_0M$, стягивающей дугу $\overline{M_0M}$ траектории точки, а указанные хорда и дуга эквивалентны (см. равенство (13.1)), то

$$|\mathbf{r}'(t)| = \lim_{\Delta t \to 0} \frac{\overset{\frown}{M_0M}}{\Delta t},$$

т. е. длина вектора скорости есть предел отношения пути, пройденного за промежуток времени $\Delta t$, к длине этого промежутка. Направление вектора $\mathbf{r}'(t)$ определяет направление движения точки $M$.

Принимая во внимание равенство (13.1), находим

$$\frac{|\Delta s|}{|\Delta \mathbf{r}|} \to 1, \quad \left|\frac{\Delta \mathbf{r}}{\Delta s}\right| \to 1, \quad \left|\frac{d\mathbf{r}}{ds}\right| = \lim_{\Delta s \to 0} \left|\frac{\Delta \mathbf{r}}{\Delta s}\right| = 1, \tag{13.46}$$

$$|d\mathbf{r}| = |ds|.$$

Следовательно, если для вектор-функции $\mathbf{r}=\mathbf{r}(t)$ в качестве параметра $t$ выбрать длину дуги $s$, отсчитываемой от некоторой точки $M_0$, то производная вектор-функции будет равна единичному вектору, направленному по касательной. Обозначив этот вектор через $\boldsymbol{\tau}$, получим

$$\frac{d\mathbf{r}}{ds} = \boldsymbol{\tau}, \quad |\boldsymbol{\tau}| = 1. \tag{13.47}$$

Отметим, что из равенства $|d\mathbf{r}| = |ds|$ следует выражение для дифференциала длины дуги в декартовых координатах:

$$ds = |d\mathbf{r}| = |d(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})| = |dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}|,$$

$$ds = \sqrt{dx^2 + dy^2 + dz^2}, \quad ds = \sqrt{x'^2 + y'^2 + z'^2}\, dt. \tag{13.48}$$

*Второй производной вектор-функции* $\mathbf{u}=\mathbf{u}(v)$ называется производная от ее производной $\mathbf{u}'(v)$:

$$\mathbf{u}''(v) = [\mathbf{u}'(v)]'. \tag{13.49}$$

Для функции (13.42) имеем

$$\mathbf{u}''(v) = \{x''(v),\ y''(v),\ z''(v)\}, \tag{13.50}$$

если существуют вторые производные функций $x(v)$, $y(v)$, $z(v)$. Аналогично определяются производные более высокого порядка для вектор-функции $\mathbf{u}=\mathbf{u}(v)$.

## § 13.7. Кривизна пространственной линии

Рассмотрим пространственную линию $\gamma$ (рис. 13.10, *а*), заданную векторно-параметрическим уравнением

$$\mathbf{r}=r(t),\ \mathbf{r}(t)=x(t)\mathbf{i}+y(t)\mathbf{j}+r(t)\mathbf{k}. \tag{13.51}$$

Фиксируем значение параметра $t$, ему соответствует некоторая точка $M$ кривой $\gamma$. Проведем касательную к линии $\gamma$ в точке $M$. Если $s$ — длина дуги, то единичный вектор касательной $\boldsymbol{\tau}$ определяется формулой (13.47). Придав аргументу $t$ приращение $\Delta t$, получим точку $N$ линии $\gamma$ и соответствующий единичный вектор касательной $\boldsymbol{\tau}+\Delta\boldsymbol{\tau}$. Степень изогнутости кривой можно характеризовать скоростью поворота вектора $\boldsymbol{\tau}$.

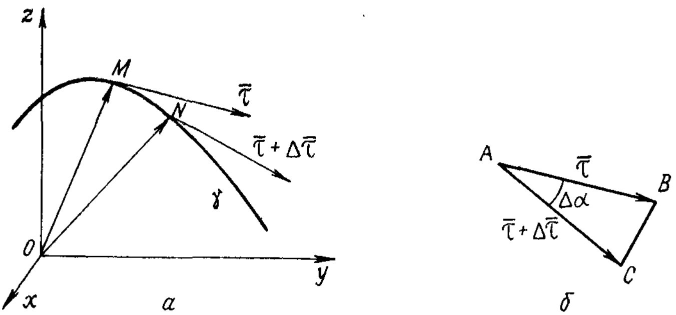



Рис. 13.10

*Кривизной линии* $\gamma$ в точке $M$ называется модуль производной вектор-функции $\boldsymbol{\tau}=\boldsymbol{\tau}(s)$ в данной точке, т. е.

$$k=\left|\frac{d\boldsymbol{\tau}}{ds}\right|. \tag{13.52}$$

Это определение кривизны равносильно определению, данному для плоской линии в § 13.2. Действительно, из рис. 13.10, *б* получаем

$$k=\left|\frac{d\boldsymbol{\tau}}{ds}\right|=\lim_{\Delta s\to 0}\left|\frac{\Delta\boldsymbol{\tau}}{\Delta s}\right|=\lim_{\Delta s\to 0}\left|\frac{\overline{BC}}{\Delta s}\right|=$$

$$=\lim_{\Delta s\to 0}\left|\frac{2\sin\dfrac{\Delta\alpha}{2}}{\Delta s}\right|=\lim_{\Delta s\to 0}\left|\frac{\Delta\alpha}{\Delta s}\right|,$$

$$k=\lim_{\Delta s\to 0}\left|\frac{\Delta\alpha}{\Delta s}\right|,$$

последнее равенство совпадает с равенством (13.5).

Выведем формулу для вычисления кривизны. Так как $|\tau| = 1$ и $d\tau/ds \perp \tau$ (согласно лемме § 13.6), то угол между векторами $\tau'_s$ и $\tau$ равен 90°, поэтому

$$k = \left|\frac{d\tau}{ds}\right| = \left|\frac{d\tau}{ds}\right| |\tau| \sin\left(\tau,\ \frac{d\tau}{ds}\right) = \left|\left[\frac{d\tau}{ds},\ \tau\right]\right|. \tag{13.53}$$

По правилу дифференцирования сложной вектор-функции

$$\tau = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt}\,\frac{dt}{ds},\quad \tau = \frac{\mathbf{r}'}{s'}. \tag{13.54}$$

Дифференцируя еще раз, находим

$$\frac{d\tau}{ds} = \frac{d\tau}{dt}\,\frac{dt}{ds} = \frac{\tau'}{s'},\quad \frac{d\tau}{ds} = \left(\frac{\mathbf{r}'}{s'}\right)'_t \frac{1}{s'} =$$

$$= \frac{\mathbf{r}''s' - s''\mathbf{r}'}{s'^2}\,\frac{1}{s'};\quad \frac{d\tau}{ds} = \frac{\mathbf{r}''s' - s''\mathbf{r}'}{s'^3}. \tag{13.55}$$

Подставляя выражения (13.54) и (13.55) в равенство (13.53), получаем формулу для кривизны пространственной линии:

$$k = \left|\left[\frac{d\tau}{ds},\ \tau\right]\right| = \left|\left[\frac{\mathbf{r}'}{s'},\ \frac{\mathbf{r}''s' - s''\mathbf{r}'}{s'^3}\right]\right| = \left|\left[\frac{\mathbf{r}'',\ \mathbf{r}'}{s'^3}\right]\right|,$$

или

$$k = \frac{|[\mathbf{r}'',\ \mathbf{r}']|}{|\mathbf{r}'|^3}, \tag{13.56}$$

так как $|\mathbf{r}'| = |s'|$ (что следует из равенства $|d\mathbf{r}| = |ds|$).

Кривизну линии можно выразить в координатах. Поскольку $\mathbf{r}(t) = \{x(t),\ y(t),\ z(t)\}$, $\mathbf{r}'(t) = \{x'(t),\ y'(t),\ z'(t)\}$, $\mathbf{r}''(t) = \{x''(t),\ y''(t),\ z''(t)\}$, на основании формулы (5.76)

$$[\mathbf{r}',\ \mathbf{r}''] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x' & y' & z' \\ x'' & y'' & z'' \end{vmatrix}. \tag{13.57}$$

Следовательно,

$$k = \frac{\sqrt{(y'z'' - y''z')^2 + (x'z'' - x''z')^2 + (x'y'' - x''y')^2}}{(x'^2 + y'^2 + z'^2)^{3/2}}. \tag{13.58}$$

Отметим, что формула (13.11) является частным случаем формулы (13.58), ибо в этом случае $z(t) = 0$ (линия лежит в плоскости $Oxy$).

Пример. Найти кривизну винтовой линии $\mathbf{r}(t) = \{a\cos t,\ a\sin t,\ bt\}$.

Поскольку $x' = -a\sin t$, $y' = a\cos t$, $z' = b$, $x'' = -a\cos t$, $y'' = -a\sin t$, $z'' = 0$, то

$$[\mathbf{r}', \mathbf{r}''] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a\sin t & a\cos t & b \\ -a\cos t & -a\sin t & 0 \end{vmatrix};$$

$$|[\mathbf{r}', \mathbf{r}'']| = \sqrt{a^2b^2 + a^4} = a\sqrt{a^2 + b^2}, \quad |\mathbf{r}'| = \sqrt{a^2 + b^2},$$

$$k = \frac{a}{a^2 + b^2}.$$

# Глава 14

# ПРИБЛИЖЕННОЕ РЕШЕНИЕ УРАВНЕНИЙ

В курсе математики средней школы рассматривались точные методы решения некоторых классов уравнений.

Существуют точные методы решения алгебраических уравнений третьей и четвертой степени. Формулы, выражающие корни этих уравнений через их коэффициенты, являются довольно громоздкими, поэтому ими редко пользуются. Что касается алгебраических уравнений степени $n$ ($n \geqslant 5$), то в общем случае для них нет таких формул; эти уравнения в радикалах не решаются.

При применении точных методов полученными результатами во многих случаях можно воспользоваться лишь приближенно, например, результатом $x = \sqrt{5}$ и т. п. По указанным причинам широко пользуются приближенными методами решения уравнений. Важнейшие из этих методов будут рассмотрены в данной главе.

## § 14.1. Графическое решение уравнений. Отделение корней уравнений

Всякое уравнение с одним неизвестным может быть записано в виде

$$f(x) = 0, \tag{14.1}$$

где $f(x)$ — некоторая функция переменной $x$.

*Корнем* уравнения (14.1) называется такое значение $x = X$, при котором левая часть уравнения обращается в нуль, т. е. $f(X) = 0$. Корень уравнения (14.1) геометрически представляет собой абсциссу точки пересечения графика функции $y = f(x)$ с осью $Ox$ (рис. 14.1, *а*), точки касания этого графика с той же осью (рис. 14.1, *б*) или другой общей точки графика и оси $Ox$ (рис. 14.1, *в*).

Из геометрической интерпретации корня следует графический способ решения уравнения с одним неизвестным: чтобы решить уравнение (14.1), необходимо построить график функции $y = f(x)$ и найти абсциссы точек пересечения графика с осью $Ox(y=0)$. Например, уравнение $2x-3=0$ имеет корень $x=1{,}5$; уравнение $x^2-4x+3=0$ — два корня: $x_1=1$, $x_2=3$; уравнение $x^3-x^2-2x=0$ — три корня: $x_1=-1$, $x_2=0$, $x_3=2$; уравнение $x^4-5x^2+4=0$ — четыре корня: $x_1=-2$, $x_2=-1$, $x_3=1$, $x_4=2$.

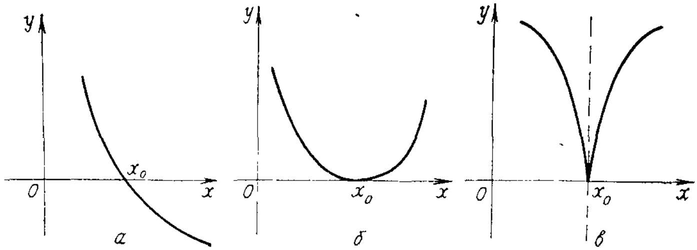



Рис. 14.1

Отделить корень уравнения (14.1) — это значит найти такой конечный промежуток, внутри которого имеется корень данного уравнения, причем корень является единственным в этом промежутке.

Для отделения корней уравнения (14.1) применяют следующий критерий: если на отрезке $[a, b]$ функция $f(x)$ непрерывна и монотонна, а ее значения на концах отрезка имеют разные знаки, то на рассматриваемом отрезке имеется единственный корень этого уравнения. Достаточным признаком монотонности функции $f(x)$ на отрезке является сохранение знака ее первой производной (если $f'(x)>0$, функция возрастает на отрезке, если $f'(x)<0$, функция убывает).

Отделение корней уравнения (14.1) можно выполнить графически, если удается построить график функции $y=f(x)$, по которому можно судить о том, в каких промежутках находятся точки пересечения его с осью $Ox$.

В некоторых случаях уравнение (14.1) целесообразно представить в эквивалентном виде

$$f_1(x)=f_2(x) \tag{14.2}$$

с таким расчетом, чтобы графики функций $y=f_1(x)$, $y=f_2(x)$ строились проще. Корень уравнения (14.2) геометрически представляет абсциссу точки пересечения графиков $y=f_1(x)$ и $y=f_2(x)$.

Таким способом легко, например, отделить корень уравнения $x^3+px+q=0$. Это будет абсцисса точек пересечения прямой $y=-px-q$ и линии $y=x^3$. Аналогично отделяются корни уравнения $x^3+ax^2+bx+c=0$.

Пример. Отделить корни уравнения $x^3+6x-5=0$.

Поскольку $f'(x)=3x^2+6>0$ для всех $x$, функция возрастает на любом промежутке. При отрицательных $x$, очевидно, функция принимает отрицательные значения; при достаточно больших положительных $x$, например при $x\geqslant 1$, — положительные значения ($f(0)<0$, $f(1)>0$). Следовательно, уравнение имеет единственный действительный корень на отрезке [0, 1].

Замечание 1. Можно указать отрезок меньшей длины, на котором находится корень. Для этого введем в рассмотрение величину $c=\frac{a+b}{2}$, т. е. середину отрезка $[a, b]$, в данном случае $c=\frac{1}{2}$. Так как $f\left(\frac{1}{2}\right)=\frac{1}{2^3}+\frac{6}{2}-5<0$, а $f(1)>0$, то корень лежит в интервале ]0, 5; 1[. Этот процесс можно продолжить: поскольку $f(0{,}75)<0$, а $f(1)>0$, корень принадлежит интервалу ]0,75; 1[.

Замечание 2. Корень данного уравнения можно отделить и графически. Придадим уравнению вид $x^3=-6x+5$, т. е. вид (14.2). Рассмотрим графики функций $y=x^3$, $y=-6x+5$ (рис. 14.2). Из рисунка видно, что указанные графики пересекаются в точке $M$, абсцисса которой находится в интервале ]0, 1[.

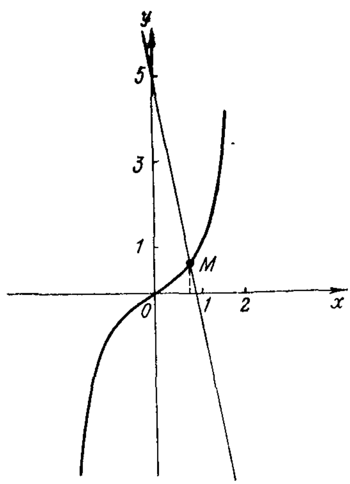


Рис. 14.2

Замечание 3. Идея деления пополам отрезка $[a, b]$, на котором находится единственный корень уравнения $f(x)=0$, лежит в основе так называемого метода «вилки» — метода приближенного вычисления корней. «Вилкой» называют всякий отрезок $[a, b]$, для концов которого $f(a)f(b)<0$. Пусть $c$ — середина отрезка $[a, b]$. Если $f(c)=0$, то $c$ — корень уравнения. В случае $f(c)\neq 0$ обозначим через $[a_1, b_1]$ один из двух отрезков $[a, c]$, $[c, b]$, на котором находится корень. Продолжая аналогичные рассуждения, имеем две возможности: 1) функция $f(x)$ равна нулю в середине некоторого отрезка (тогда корень найден); 2) указанный процесс можно продолжать неограниченно; в этом случае получаем систему вложенных отрезков $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$, длина которых стремится к нулю. Согласно теореме 10.1 эта система имеет общую точку $\xi$. Можно доказать, что $f(\xi)=0$, т. е. $\xi$ — корень уравнения.

## § 14.2. Метод хорд

Предположим, что функция $f(x)$, стоящая в левой части уравнения (14.1), на концах промежутка $[a, b]$ принимает значения различных знаков, а ее производная на указанном промежутке сохраняет знак, тогда в интервале $]a, b[$ имеется один и только один корень уравнения (14.1).

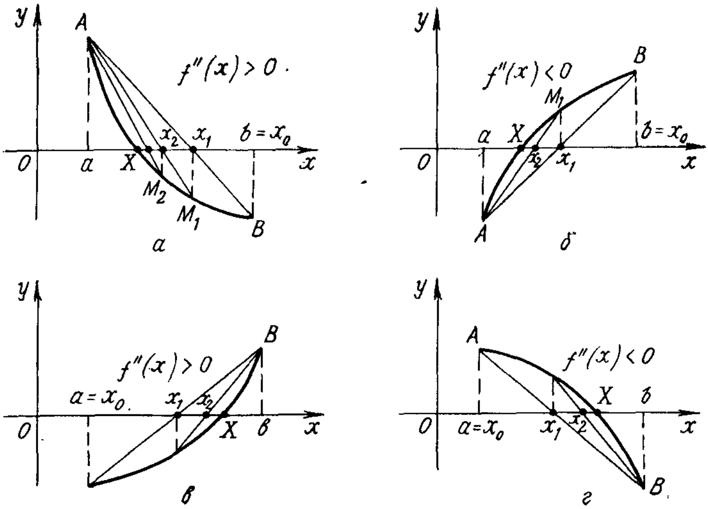



Рис. 14.3

В качестве первого приближенного значения корня данного уравнения возьмем абсциссу точки пересечения оси $Ox$ и хорды, соединяющей точки $A(a, f(a))$ и $B(b, f(b))$ — концы дуги графика функции $y=f(x)$ (рис. 14.3). На основании формулы (2.10) получаем уравнение прямой $AB$

$$\frac{y-f(a)}{f(b)-f(a)}=\frac{x-a}{b-a}. \qquad (14.3)$$

Положив в этом уравнении $y=0$, найдем первое приближение корня

$$x_1=a-\frac{f(a)(b-a)}{f(b)-f(a)} \quad \text{или} \quad x_1=\frac{af(b)-bf(a)}{f(b)-f(a)}. \qquad (14.4)$$

Чтобы получить второе приближение $x_2$ искомого корня $X$ данного уравнения, применим указанное рассуждение к тому из отрезков $[a, x_1]$, $[x_1; b]$, на концах которого функция имеет разные знаки. В случае, когда выполняется условие (см. рис. 14.3, *а, б*)

$$f(a)f''(x)>0, \tag{14.5}$$

находим

$$x_2=\frac{af(x_1)-x_1f(a)}{f(x_1)-f(a)}, \tag{14.6}$$

в случае (см. рис. 14.3, *в, г*), когда

$$f(b)f''(x)>0, \tag{14.7}$$

получаем

$$x_2=\frac{bf(x_1)-x_1f(b)}{f(x_1)-f(b)}. \tag{14.8}$$

Продолжая аналогичные рассуждения, находим последующие приближения корня $x_3$, $x_4$, $x_5$, .... Если известно $(n-1)$-е приближение $x_{n-1}$, то $n$-е приближение $x_n$ вычисляется по формуле

$$x_n=\frac{af(x_{n-1})-x_{n-1}f(a)}{f(x_{n-1})-f(a)}\ (n=1,\ 2,\ 3,\ \ldots), \tag{14.9}$$

когда выполняется условие (14.5), или по формуле

$$x_n=\frac{bf(x_{n-1})-x_{n-1}f(b)}{f(x_{n-1})-f(b)}\ (n=1,\ 2,\ 3,\ \ldots) \tag{14.10}$$

при выполнении условия (14.7).

В первом случае за начальное приближение принимается $b$, т. е. $x_0=b$, во втором—$a$, т. е. $x_0=a$ (см. рис. 14.3).

Из геометрических соображений следует, что последовательность чисел $x_n(n=1, 2, 3, \ldots)$ сходится к корню, т. е.

$$\lim_{n\to\infty} x_n = X. \tag{14.11}$$

Пользуясь формулами (14.9) и (14.10), вычисления проводят до тех пор, пока в двух последовательных приближениях не станут совпадать десятичные знаки с учетом заданной точности.

Отметим, что оценка абсолютной погрешности метода хорд определяется неравенством

$$|X-x_n|\leqslant\frac{|f(x_n)|}{\mu}, \tag{14.12}$$

где

$$\mu = \min_{a \leqslant x \leqslant b} |f'(x)|, \ f'(x) \neq 0. \qquad (14.13)$$

Чтобы доказать формулу (14.12), применим к выражению $f(x_n) = f(x_n) - f(X)$ ($f(X) = 0$, ибо $X$ — корень уравнения $f(x) = 0$) формулу Лагранжа. В результате получим $f(x_n) - f(X) = (x_n - X) f'(\xi_n)$, где $\xi_n$ находится между $x_n$ и $X$, или $f(x_n) = (x_n - X) \cdot f'(\xi_n)$, откуда и следует (14.12).

Пример. Методом хорд найти действительный корень уравнения $x^3 + x - 1 = 0$.

В данном случае $f(x) = x^3 + x - 1$. Поскольку $f(0{,}5) < 0$, $f(1) > 0$ и $f'(x) = 3x^2 + 1 > 0$ для всех $x$, то в интервале $]0{,}5;\ 1[$ находится единственный действительный корень уравнения.

Так как $f''(x) = 6x$ и $f(1) f''(x) > 0$, т. е. выполнено условие (14.7), то пользуемся формулой (14.10), положив в ней $b = 1$, $x_0 = 0{,}5$; $f(x_0) = f(0{,}5) = (0{,}5)^3 + 0{,}5 - 1 = -0{,}375$:

$$x_1 = \frac{bf(x_0) - x_0 f(b)}{f(x_0) - f(b)} = \frac{1(-0{,}375) - 0{,}5 \cdot 1}{-0{,}375 - 1} \approx 0{,}636364;$$

$$x_2 = \frac{bf(x_1) - x_1 f(b)}{f(x_1) - f(b)} = \frac{1(-0{,}105935) - 0{,}636364}{-0{,}105935 - 1} \approx 0{,}671196;$$

$$x_3 = \frac{bf(x_2) - x_2 f(b)}{f(x_2) - f(b)} = \frac{1(-0{,}026428) - 0{,}671196 \cdot 1}{-0{,}026428 - 1} \approx 0{,}679662.$$

Аналогично вычисляем последующие приближения: $x_4 = 0{,}681691$, $x_5 = 0{,}682176$, $x_6 = 0{,}682292$, $x_7 = 0{,}682319$, $x_8 = 0{,}682326$, $x_9 = 0{,}682327$. Следовательно, с точностью до 0,0001 получено значение корня $X = 0{,}6823$.

## § 14.3. Метод касательных

*Метод касательных* (или *метод Ньютона*) нахождения действительных корней состоит в следующем. Пусть в промежутке $(a, b)$ находится единственный корень уравнения (14.1). Проведем касательную к графику функции $y = f(x)$ в точке $A(a, f(a))$. Ее уравнение в соответствии с формулой (11.8) принимает вид

$$y - f(a) = f'(a)(x - a). \qquad (14.14)$$

Полагая в этом уравнении $y = 0$, находим абсциссу точки пересечения касательной с осью $Ox$:

$$x_1 = a - \frac{f(a)}{f'(a)}, \qquad (14.15)$$

если $f'(a) \neq 0$.

Формула (14.15) дает первое приближенное значение корня. Чтобы получить второе приближение, применим аналогичные рассуждения к промежутку $(x_1, b)$. Касательная к графику функции (рис. 14.4) в точке $A_1(x_1, f(x_1))$ определяется уравнением

$$y-f(x_1)=f'(x_1)(x-x_1).$$

Положив в этом уравнении $y=0$, получим второе приближение:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}. \tag{14.16}$$

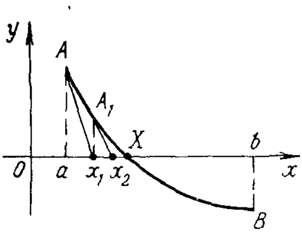



Рис. 14.4

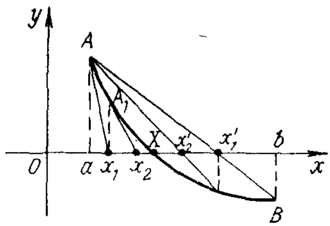



Рис. 14.5

Аналогично вычисляются последующие приближения корня. Если известно $(n-1)$-е приближение, то $n$-е приближение корня уравнения (14.1) находится по формуле

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})} \quad (n = 1,\ 2,\ 3,\ \ldots), \tag{14.17}$$

в которой за нулевое приближение $x_0$ принимается такое значение из отрезка $[a, b]$, для которого выполняется условие

$$f(x_0)f''(x) > 0, \tag{14.18}$$

т. е. начинать нужно не обязательно с точки $A$ (рис. 14.4).

Соотношения (14.11)—(14.13) остаются справедливыми и для метода касательных.

Пример. Методом касательных найти действительный корень уравнения $x^3+x-3=0$.

Придавая данному уравнению вид (14.2), получаем $x^3=-x+3$. Построив графики функций $f_1(x)=x^3$ и $f_2(x)=-x+3$, найдем, что единственный действительный корень уравнения принадлежит интервалу $]1, 2[$. Укажем промежуток меньшей длины, на котором находится корень. Поскольку $f(1{,}2)=(1{,}2)^3+1{,}2-3=-0{,}072<0$, $f(1{,}3)=(1{,}3)^3+1{,}3-3=0{,}497>0$, то корень лежит в интервале $]1{,}2;\ 1{,}3[$.

Рассмотрим еще середину отрезка [1, 2; 1, 3], т. е. точку $x=1{,}25$. Так как $f(1{,}25)=(1{,}25)^3+1{,}25-3=0{,}203125>0$ и $f(1{,}2)<0$, то искомый корень принадлежит отрезку [1,20; 1,25].

Данная функция $f(x)=x^3+x-3$ имеет производные $f'(x)=3x^2+1$, $f''(x)=6x$, принимающие положительные значения на этом отрезке.

В качестве начального приближения возьмем $x_0=1{,}25$, так как для этой точки выполнено условие (14.18).

Результаты вычислений, выполненных по формуле (14.17), заносим в табл. 3, из которой видно, что $X=1{,}21341$.

Таблица 3

| $n$ | $x_n$ | $x_n^3$ | $f(x_n)=$ $=x_n^3+$ $+x_n-3$ | $f'(x_n)=$ $=3x_n^2+1$ | $\dfrac{f(x_n)}{f'(x_n)}$ | $x_{n+1}=$ $=x_n-$ $-\dfrac{f(x_n)}{f'(x_n)}$ |
|---|---|---|---|---|---|---|
| 0 | 1,25 | 1,953125 | 0,203125 | 5,6875 | 0,035714 | 1,214286 |
| 1 | 1,214286 | 1,790452 | 0,004738 | 5,42347 | 0,000874 | 1,213412 |
| 2 | 1,213412 | 1,786590 | 0,000002 | 5,417107 | 0,0000004 | 1,2134116 |

## § 14.4. Комбинированный метод

Комбинированный метод состоит в одновременном использовании метода хорд и метода касательных. Его удобно применять, если вторая производная $f''(x)$ сохраняет знак в промежутке, на котором находится корень уравнения $f(x)=0$. В этом случае график функции не имеет точек перегиба, является выпуклым вверх или вниз (рис. 14.5), поэтому можно гарантировать приближение к корню с двух сторон: касательная пересекает ось $Ox$ со стороны выпуклости, а хорда — со стороны вогнутости графика функции $y=f(x)$. Приближения по методу касательных будут располагаться с одной стороны корня, а приближения по методу хорд — с другой. Таким образом, получаются все более суживающиеся отрезки, внутри которых заключен корень. Длина последнего из отрезков дает величину абсолютной погрешности приближенного значения корня.

Пример. Комбинированным методом найти корень уравнения $x^3-2x+7=0$.

Действительный корень данного уравнения находится на отрезке [—2,3; —2,2]. Условие (14.18) выполнено для $x_0=-2{,}3$. Приближения по методу хорд будем помечать штрихом. Пользуемся формулами (14.9) и (14.17). Замечая, что $a=x_0=-2{,}3$, $b=x_0'=-2{,}2$, $f(-2{,}3)=-0{,}567$, $f(-2{,}2)=0{,}752$, по этим формулам находим

$$x_1=x_0-\frac{f(x_0)}{f'(x_0)}=-2{,}3-\frac{-0{,}567}{13{,}87}=-2{,}3+$$
$$+0{,}040879=-2{,}259121,$$
$$x_1'=\frac{af(x_0')-x_0'f(a)}{f(x_0')-f(a)}=$$
$$=\frac{(-2{,}3)\cdot f(-2{,}2)-(-2{,}2)\cdot f(-2{,}3)}{f(-2{,}2)-f(-2{,}3)}=-2{,}257013,$$
$$x_2=-2{,}258259,\ x_2'=-2{,}258231,\ x_3=-2{,}2582589,$$
$$x_3'=-2{,}258259.$$

Следовательно, $X=-2{,}258259$ — корень данного уравнения.

## § 14.5. Метод итераций

*Метод итераций* (или *метод последовательных приближений*) решения уравнения (14.1) состоит в следующем. Каким-либо способом получают приближеиное значение $x_0$ корня данного уравнения (это может быть произвольное значение из интервала, содержащего корень; как было видно из приведенных выше примеров, такой интервал можно сделать достаточно малым).

Уравнение (14.1) представляют в виде

$$x=\varphi(x), \tag{14.19}$$

что всегда можно сделать, причем различными способами, например,

$$x=x+cf(x), \tag{14.20}$$

где $c$ — произвольная постоянная.

Обозначим через $x_1$ результат подстановки $x_0$ в правую часть уравнения (14.19):

$$x_1=\varphi(x_0), \tag{14.21}$$

далее, $x_2=\varphi(x_1)$. Продолжая указанные действия, получаем

$$x_n=\varphi(x_{n-1})\ (n=1,\ 2,\ 3,\ \ldots). \tag{14.22}$$

Процесс последовательного вычисления чисел $x_n$ $(n=1, 2, 3, \ldots)$ по формуле (14.22) называется методом итераций, или методом последовательных приближений.

Условия сходимости этого процесса выражаются следующей теоремой.

**Теорема.** Если в промежутке, содержащем корень $X$ уравнения (14.19) и его последовательные приближения $x_0, x_1, x_2, \ldots, x_n, \ldots$ ..., вычисляемые по формуле (14.22), выполнено условие

$$|\varphi'(x)|\leqslant q<1, \tag{14.23}$$

то процесс итераций сходится, т. е.

$$\lim_{n\to\infty} x_n = X. \qquad (14.24)$$

Доказательство. Поскольку $X$ — корень уравнения (14.19), то $X=\varphi(X)$. Вычитая из этого равенства равенство (14.21) и применяя теорему Лагранжа о конечном приращении, получаем

$$X-x_1=\varphi(X)-\varphi(x_0)=(X-x_0)\varphi'(\xi_0),$$

где $\xi_0$ — число, заключенное между $X$ и $x_0$.

Аналогичные равенства получаем и для последующих приближений:

$$X-x_2=\varphi(X)-\varphi(x_1)=(X-x_1)\varphi'(\xi_1),$$

$$X-x_3=\varphi(X)-\varphi(x_2)=(X-x_2)\varphi'(\xi_2),$$

$$X-x_n=\varphi(X)-\varphi(x_{n-1})=(X-x_{n-1})\varphi'(\xi_{n-1}),$$

где $\xi_i$ $(i=1, 2, \ldots, n-1)$ — некоторое число из промежутка между $x_i$ и $X$.

Перемножая почленно выписанные равенства и производя сокращения, находим

$$X-x_n=(X-x_0)\varphi'(\xi_0)\varphi'(\xi_1)\ldots\varphi'(\xi_{n-1}).$$

Принимая во внимание (14.23), получаем неравенство

$$|X-x_n|\leqslant|X-x_0|q^n.$$

Поскольку $0<q<1$ и $\lim\limits_{n\to\infty} q^n=0$, то выполняется и равенство (14.24).

Пример 1. Методом итераций найти действительный корень уравнения $x^7+x+4=0$.

Представив данное уравнение в виде $x^7=-x-4$ и построив графики функций $f_1(x)=x^7$, $f_2(x)=-x-4$, найдем, что искомый корень принадлежит отрезку $[-2, -1]$. Можно указать отрезок меньшей длины, на котором находится корень. Поскольку $f(-1{,}2)<0$ и $f(-1,1)>0$, корень лежит на отрезке $[-1, 2; -1, 1]$.

Из данного уравнения находим

$$x=\sqrt[7]{-x-4},\quad \varphi(x)=\sqrt[7]{-x-4},\quad \varphi'(x)=\frac{1}{7}\,\frac{1}{\sqrt[7]{(-x-4)^6}}.$$

Для функции $\varphi(x)$ условие (14.23) выполнено, ибо при $-1{,}2\leqslant x\leqslant -1{,}1$

$$|\varphi'(x)|=\frac{1}{7}\,\frac{1}{\sqrt[7]{(-x-4)^6}}\leqslant q=0{,}06<1.$$

В качестве начального приближения возьмем $x_0=-1{,}15$. Результаты вычислений, выполненных по формуле $x_n=\sqrt[7]{-x_{n-1}-4}$, сводим в табл. 4, из которой видно, что $X=-1{,}1607$ — корень данного уравнения.

Таблица 4

| $n$ | $x_{n-1}$ | $-x_{n-1}-4$ | $x_n=\varphi(x_{n-1})=\sqrt[7]{-x_{n-1}-4}$ |
|---|---|---|---|
| 1 | $-1,15$ | $-2,85$ | $-1,1614$ |
| 2 | $-1,1614$ | $-2,8386$ | $-1,1607$ |
| 3 | $-1,1607$ | $-2,8393$ | $-1,1607$ |

Замечание. Данное уравнение можно переписать в виде $x=\varphi_1(x)$, где $\varphi_1(x)=-4-x^7$. Поскольку для функции $\varphi_1(x)$ условие (14.23) не выполнено, то этим представлением пользоваться нельзя для вычисления корня на отрезке $[-1,2; -1,1]$.

Пример 2. Вычислить объем одного моля метана при 0° C и 50 атм (см. [11]).

В этих условиях газ заметно отклоняется от идеального состояния. Для вычисления объема воспользуемся уравнением Вандер-Ваальса, которое запишем в следующем виде:

$$V=\frac{RT}{p+\frac{a}{v^2}}+b,$$

где $V$ — объем одного моля газа; $a=2{,}253$ л$^2\cdot$атм$\cdot$моль$^{-2}$; $b=0{,}0428$ л$\cdot$моль$^{-1}$.

В качестве нулевого приближения возьмем объем, полученный из уравнения состояния идеального газа, т. е.

$$V_0=\frac{RT}{p}=\frac{0{,}082\cdot 273}{50}=0{,}448.$$

Подставляя это значение в правую часть исходного уравнения, получаем первое приближение

$$V_1=\frac{22{,}4}{50+\frac{2{,}253}{(0{,}448)^2}}+0{,}043\approx 0{,}409.$$

Аналогично вычисляем следующие приближения: $V_2\approx 0{,}395$, $V_3\approx 0{,}391$, $V_4\approx 0{,}390$, $V_5=V_6\approx 0{,}389$. Следовательно, с точностью до трех десятичных знаков находим $V=0{,}389$ (л).

## § 14.6. Метод малого параметра

*Метод малого параметра,* как и метод итераций, принадлежит к числу универсальных математических методов. Этот метод (его называют также *методом возмущений*) состоит в следующем.

Пусть требуется найти решение уравнения, содержащего кроме известных и неизвестных величин еще некоторый параметр $\alpha$, т. е. уравнения вида

$$F(x, \alpha)=0. \tag{14.25}$$

Предположим, что это уравнение при некотором значении параметра $\alpha=\alpha_0$ решается сравнительно легко; уравнение принимает вид $f(x)=0$ и имеет решение $x=x_0$. Последнее решение называется *невозмущенным*. Поскольку очень часто принимают $\alpha_0=0$, то рассматриваются достаточно малые $\alpha$, отчего и происходит название метода.

Решение уравнения (14.25) при $\alpha$, достаточно близких к $\alpha_0$, ищут в виде разложения по степеням $(\alpha-\alpha_0)$, т. е. в виде

$$x=x(\alpha)=a_0+a_1(\alpha-\alpha_0)+a_2(\alpha-\alpha_0)^2+a_3(\alpha-\alpha_0)^3+\ldots \quad (14.26)$$

Отметим, что свободный член разложения (14.26) получается при $\alpha=\alpha_0$, а именно $a_0=x_0$, т. е. совпадает с невозмущенным решением. Последующие члены дают поправки на «возмущение» решения, они имеют различный порядок малости по сравнению с $(\alpha-\alpha_0)$.

Подставляя выражение (14.26) в уравнение (14.25), получаем

$$c_0+c_1(\alpha-\alpha_0)+c_2(\alpha-\alpha_0)^2+c_3(\alpha-\alpha_0)^3+\ldots+=0, \quad (14.27)$$

где $c_i=\varphi_i(a_k)$ — некоторые функции коэффициентов разложения (14.26).

Обращая в нуль коэффициенты $c_i$, находят систему уравнений, из которой определяются коэффициенты разложения (14.26).

Этот метод определения коэффициентов разложения (14.26) называется *методом неопределенных коэффициентов*.

Метод малого параметра дает хорошие результаты лишь при $\alpha$, достаточно близких к $\alpha_0$. Чем меньше $|\alpha-\alpha_0|$, тем меньше нужно взять членов в разложении (14.26), т. е. тем меньше нужно вычислять коэффициентов искомого разложения.

Пример. Найти действительный корень уравнения $x^3+0{,}1x-1=0$.

Это уравнение никакого параметра не содержит и имеет постоянные коэффициенты, поэтому его можно решить с помощью методов, изложенных в предыдущих параграфах.

Можно найти его решение и методом малого параметра. Введем в рассмотрение уравнение, содержащее параметр $\alpha$:

$$x^3+\alpha x-1=0. \quad (1)$$

Исходное уравнение получается из уравнения при $\alpha=0{,}1$.

Найдем решение уравнения при малых $|\alpha|$ с точностью до величин порядка $\alpha^3$ включительно. Решение уравнения (1) ищем в виде

$$X=X(\alpha)=a+b\alpha+c\alpha^2+d\alpha^3+o(\alpha^3), \quad (2)$$

где $a$, $b$, $c$, $d$ — коэффициенты, которые нужно определить.

Подставляя выражение (2) в уравнение (1) и выполняя необходимые преобразования, получаем

$$[a+b\alpha+c\alpha^2+d\alpha^3+o(\alpha^3)]^3+\alpha[a+b\alpha+c\alpha^2+d\alpha^3+o(\alpha^3)]-1=0,$$
$$(a^3-1)+(a+3a^2b)\alpha+(b+3ab^2+3a^2c)\alpha^2+$$
$$+(b^3+3a^2d+6abc+c)\alpha^3+o(\alpha^3)=0.$$

Приравнивая нулю коэффициенты при степенях $\alpha$, находим систему уравнений

$$\left.\begin{aligned} a^3-1&=0;\\ a+3a^2b&=0;\\ b+3ab^2+3a^2c&=0;\\ b^3+3a^2d+6abc+c&=0,\end{aligned}\right\}$$

из которой $a=1$, $b=-\frac{1}{3}$, $c=0$, $d=\frac{1}{81}$.

Следовательно, искомое разложение с заданной точностью принимает вид

$$X=1-\frac{1}{3}\alpha+\frac{1}{81}\alpha^3. \qquad (3)$$

С помощью этой формулы, взяв $\alpha=0{,}1$, получим решение исходного уравнения $x^3+0{,}1x-1=0$, $X=0{,}967$.

Замечание. Формула (3) позволяет находить решения многих уравнений, получающихся из уравнения (1) при малых значениях параметра $\alpha$. Так, уравнение $x^3-0{,}1x-1=0$ имеет решение $X=1{,}033$; уравнение $x^3+0{,}01x-1=0$ — решение $X=0{,}996667$; уравнение $x^3+0{,}001x-1=0$ — решение $X=0{,}999667$ и т. д.

# Глава 15

# НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

В данной главе рассматривается задача, обратная задаче о дифференцировании функций. Задача состоит в следующем: дана функция $f(x)$, являющаяся производной некоторой функции $F(x)$; требуется найти функцию $F(x)$. К такой математической задаче приводят многие физические, химические и другие задачи, например, задача об отыскании закона неравномерного движения материальной точки вдоль прямой по заданной скорости, задача о нахождении закона химической реакции по известной ее скорости.

## § 15.1. Первообразная функция. Неопределенный интеграл и его свойства

Функция $F(x)$, определенная в промежутке $(a, b)$, называется *первообразной* данной функции $f(x)$ в этом промежутке, если для любого значения $x\in(a, b)$ выполняется равенство

$$F'(x)=f(x). \qquad (15.1)$$

Например, функция $F(x)=x^4$ — первообразная функции $f(x)=4x^3$ в интервале $]-\infty, +\infty[$, поскольку $(x^4)'=4x^3$ для всех $x$; функция $F(x)=\ln x$ — первообразная функции $f(x)=\frac{1}{x}$ в интервале $]0, +\infty[$, так как $(\ln x)'=\frac{1}{x}$; функция $F(x)=\arcsin x$ — первообразная функции $f(x)=\frac{1}{\sqrt{1-x^2}}$ в интервале $]-1, 1[$, ибо $(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}$.

Если $F(x)$ — первообразная функции $f(x)$, то

$$\Phi(x)=F(x)+C, \qquad (15.2)$$

где $C$ — произвольная постоянная, также является ее первообразной, поскольку $\Phi'(x)=[F(x)+C]'=F'(x)+0=f(x)$. Обратно, если $F(x)$ и $\Phi(x)$ — две первообразные функции $f(x)$, то они отличаются произвольным слагаемым, т. е.

$$\Phi(x)-F(x)=C,\ \Phi(x)=F(x)+C. \tag{15.3}$$

Это следует из теоремы Лагранжа (см. § 11.12, следствие 2).

Таким образом, выражение (15.2), в котором функция $F(x)$ удовлетворяет условию (15.1), определяет множество всех первообразных данной функции $f(x)$ в заданном промежутке $(a, b)$.

*Неопределенным* интегралом от данной функции $f(x)$ называется множество всех ее первообразных

$$\int f(x)\,dx=F(x)+C, \tag{15.4}$$

где $F'(x)=f(x)$. Знак $\int$ называется знаком неопределенного интеграла; функция $f(x)$ — подынтегральной функцией; выражение $f(x)\,dx$ — подынтегральным выражением.

Операция нахождения первообразной данной функции называется *интегрированием*.

Неопределенный интеграл обладает следующими основными свойствами.

1. Производная неопределенного интеграла равна подынтегральной функции; дифференциал неопределенного интеграла равен подынтегральному выражению:

$$\left(\int f(x)\,dx\right)'=f(x); \tag{15.5}$$

$$d\int f(x)\,dx=f(x)\,dx. \tag{15.6}$$

2. Неопределенный интеграл от дифференциала некоторой функции равен этой функции с точностью до постоянного слагаемого:

$$\int d\varphi(x)=\varphi(x)+C. \tag{15.7}$$

3. Постоянный множитель можно выносить за знак неопределенного интеграла:

$$\int kf(x)\,dx=k\int f(x)\,dx \quad (k=\text{const},\ k\neq 0). \tag{15.8}$$

4. Если функции $f_1(x)$ и $f_2(x)$ имеют первообразные, то функция $f_1(x)+f_2(x)$ также имеет первообразную, причем

$$\int [f_1(x)+f_2(x)]\,dx=\int f_1(x)\,dx+\int f_2(x)\,dx. \tag{15.9}$$

Первое свойство следует из определения неопределенного интеграла. Третье и четвертое свойства доказываются сравнением производных от обеих частей равенств (15.8) и (15.9), при этом учитывается, что неопределенный интеграл определен с точностью до постоянного слагаемого.

Чтобы доказать второе свойство, обозначим $\int d\varphi(x) = \int \varphi'(x)dx = F(x)$. На основании первого свойства получаем $\varphi'(x) = F'(x)$, откуда $F(x) = \varphi(x) + C$, т. е. $\int d\varphi(x) = \varphi(x) + C$.

## § 15.2. Таблица основных неопределенных интегралов

Таблицу простейших неопределенных интегралов нетрудно получить, воспользовавшись тем, что интегрирование является операцией, обратной дифференцированию. Будем исходить из формулы (15.7), которую запишем следующим образом: если $dF(x) = f(x)dx$, то $\int f(x)dx = F(x) + C$. Например, поскольку $d(\sin x) = \cos x dx$, то $\int \cos x dx = \sin x + C$.

Применяя аналогичное рассуждение к каждой из формул таблицы дифференциалов (см. § 11.10), получаем следующую таблицу простейших неопределенных интегралов:

1. $\int 1 \cdot dx = \int dx = x + C$.

2. $\displaystyle\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C \ (\alpha \neq -1)$.

3. $\displaystyle\int \frac{1}{x} dx = \int \frac{dx}{x} = \ln|x| + C$.

4. $\displaystyle\int a^x dx = \frac{a^x}{\ln a} + C \ (a > 0)$.

4а. $\int e^x dx = e^x + C$.

5. $\int \cos x dx = \sin x + C$.

6. $\int \sin x dx = -\cos x + C$.

7. $\displaystyle\int \frac{1}{\cos^2 x} dx = \int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$.

8. $\displaystyle\int \frac{1}{\sin^2 x} dx = \int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$.

9. $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C = -\arccos x + C.$

10. $\int \frac{dx}{1+x^2} = \operatorname{arctg} x + C = -\operatorname{arcctg} x + C.$

11. $\int \operatorname{sh} x dx = \operatorname{ch} x + C.$

12. $\int \operatorname{ch} x dx = \operatorname{sh} x + C.$

13. $\int \frac{dx}{\operatorname{ch}^2 x} = \operatorname{th} x + C.$

14. $\int \frac{dx}{\operatorname{sh}^2 x} = -\operatorname{cth} x + C.$

Отметим, что все указанные формулы справедливы в тех промежутках, в которых определены соответствующие функции. Например, формула 3 справедлива для любого промежутка, не содержащего точку $x=0$; формула 9 — для интервала $]-1, 1[$ и т. п.

Замечание. В таблице основных интегралов предполагалось, что $x$ является независимой переменной. Однако формулы этой таблицы остаются справедливыми и в случае, когда $x=\varphi(t)$, где $\varphi(t)$ — любая дифференцируемая функция новой переменной $t$. Действительно, пусть

$$F(x)+C=\int f(x)\,dx,\quad F'(x)=f(x) \tag{1}$$

и $u=\varphi(x)$ — дифференцируемая функция $x$. В силу инвариантности формы первого дифференциала $dF(u)=F'(u)\,du=f(u)\,du$, $dF(u)=f(u)\,du$, откуда

$$F(u)+C=\int f(u)\,du. \tag{2}$$

Итак, из справедливости формулы (1) следует справедливость формулы (2), которая получается из первой формулы формальной заменой $x$ на $u$.

На основании этого свойства получаем обобщенную таблицу простейших интегралов:

$$\int u^\alpha du = \frac{u^{\alpha+1}}{\alpha+1} + C\ (\alpha \neq -1),\quad \int \frac{du}{u} = \ln|u| + C,$$

$$\int e^u du = e^u + C,\quad \int \sin u\,du = -\cos u + C \text{ и т. д.},$$

где $u$ — любая дифференцируемая функция $x$.

При использовании формул этой таблицы для преобразования подынтегрального выражения к виду $f(x)dx=g(u)du$ применяются простейшие преобразования дифференциалов:

1) $dx = d(x+b)$, где $b = \text{const}$;

2) $dx = \frac{1}{a} d(ax)$, $a \neq 0$;

3) $dx = \frac{1}{a} d(ax+b)$, $a \neq 0$;

4) $xdx = \frac{1}{2} d(x^2 + b)$;

5) $\sin xdx = d(-\cos x)$;

6) $\cos xdx = d(\sin x)$;

7) $\varphi'(x)\,dx = d\varphi(x)$.

Например,

$$\int \sin 7x dx = \int \sin 7x \cdot \frac{1}{7} d(7x) = \frac{1}{7}\int \sin 7x d(7x) = -\frac{\cos 7x}{7} + C;$$

$$\int \frac{xdx}{x^2+4} = \int \frac{\frac{1}{2} d(x^2+4)}{x^2+4} = \frac{1}{2}\int \frac{d(x^2+4)}{x^2+4} = \frac{1}{2}\ln(x^2+4)+C.$$

## § 15.3. Понятие об основных методах интегрирования

К наиболее важным методам интегрирования относятся методы: непосредственного интегрирования; замены переменной; интегрирования по частям.

*Метод непосредственного интегрирования* основан на свойстве 4 неопределенного интеграла. Если функции $f_1(x), f_2(x), \ldots, f_n(x)$ имеют первообразные в некотором промежутке, то функция $f(x) = f_1(x) + f_2(x) + \ldots + f_n(x)$ также имеет первообразную в том же промежутке, причем

$$\int [f_1(x) + f_2(x) + \ldots + f_n(x)]\,dx = \int f_1(x)\,dx + \int f_2(x)\,dx + \ldots + \int f_n(x)\,dx, \qquad (15.10)$$

т. е. неопределенный интеграл от алгебраической суммы конечного числа функций равен такой же алгебраической сумме неопределенных интегралов от слагаемых. В частности,

$$\int (a_0 + a_1x + a_2x^2 + \ldots + a_nx^n)\,dx =$$
$$= \int a_0 dx + \int a_1 x dx + \int a_2x^2dx + \ldots + \int a_nx^ndx =$$
$$= a_0\int dx + a_1\int xdx + a_2\int x^2dx + \ldots + a_n\int x^ndx =$$
$$= a_0x + \frac{a_1x^2}{2} + \frac{a_2x^3}{3} + \ldots + \frac{a_nx^{n+1}}{n+1} + C.$$

Следовательно, для всякого многочлена степени $n$ существует первообразная, равная многочлену степени $n+1$.

Пример 1. Найти неопределенный интеграл $\int(\cos x-\operatorname{sh} x+e^x)\,dx$.

На основании формулы (15.10) и формул 4а, 5, 11 таблицы простейших интегралов находим

$$\int(\cos x-\operatorname{sh} x+e^x)\,dx=\int\cos x\,dx-\int\operatorname{sh} x\,dx+\int e^x dx=$$
$$=\sin x-\operatorname{ch} x+e^x+C.$$

Пример 2. Найти неопределенный интеграл

$$\int\left(\frac{1}{x}+\frac{1}{\sqrt{x}}-\frac{1}{x^2}\right)dx.$$

С помощью формул 2 и 3 $\left(\text{при } \alpha=-\frac{1}{2} \text{ и } \alpha=-2\right)$ таблицы интегралов получаем

$$\int\left(\frac{1}{x}+\frac{1}{\sqrt{x}}-\frac{1}{x^2}\right)dx=\int\left(\frac{1}{x}+x^{-\frac{1}{2}}-x^{-2}\right)dx=$$
$$=\ln|x|+2\sqrt{x}+\frac{1}{x}+C.$$

*Метод замены переменной* (или *метод подстановки*) основан на следующей теореме.

**Теорема 15.1.** Если $F(x)$ — первообразная функции $f(x)$, а $x=\varphi(t)$ — дифференцируемая функция, то функция $f[\varphi(t)]\varphi'(t)$ также имеет первообразную, причем

$$\int f[\varphi(t)]\varphi'(t)\,dt=F[\varphi(t)]+C.$$

Доказательство. По правилу дифференцирования сложной функции

$$\{F[\varphi(t)]\}'_t=F'_\varphi\cdot\varphi'_t=f[\varphi(t)]\varphi'(t),$$

т. е. функция $f[\varphi(t)]\varphi'(t)$ имеет в качестве одной из своих первообразных функцию $F[\varphi(t)]$.

Следовательно,

$$\int f[\varphi(t)]\varphi'(t)\,dt=F[\varphi(t)]+C,$$

что и требовалось доказать.

Поскольку $F[\varphi(t)]+C=F(x)+C=\int f(x)\,dx$, то

$$\int f(x)\,dx=\int f[\varphi(t)]\varphi'(t)\,dt. \qquad (15.11)$$

По формуле (15.11) осуществляется замена переменной в неопределенном интеграле.

Пример 3. Найти неопределенный интеграл $\int \sin(2-3x)\,dx$.

Введем новую переменную по формуле $2-3x=t$, тогда $x=$ $=\frac{2-t}{3}$, $dx=-\frac{dt}{3}$ и

$$\int \sin(2-3x)\,dx=\int \sin t\cdot\left(-\frac{dt}{3}\right)=-\frac{1}{3}\int \sin t dt=$$

$$=\frac{1}{3}\cos t+C=\frac{1}{3}\cos(2-3x)+C.$$

Пример 4. Найти неопределенный интеграл $\int \frac{dx}{\sqrt{x^2+\alpha}}$.

Прнменим так называемую *подстановку Эйлера* $\sqrt{x^2+\alpha}=t-x$ нли $t=x+\sqrt{x^2+\alpha}$. Поскольку

$$dt=\left(1+\frac{x}{\sqrt{x^2+\alpha}}\right)dx=\frac{x+\sqrt{x^2+\alpha}}{\sqrt{x^2+\alpha}}\,dx=$$

$$=\frac{tdx}{\sqrt{x^2+\alpha}},\ \text{или}\ \frac{dx}{\sqrt{x^2+\alpha}}=\frac{dt}{t},$$

то

$$\int \frac{dx}{\sqrt{x^2+\alpha}}=\int \frac{dt}{t}=\ln|t|+C=\ln|x+\sqrt{x^2+\alpha}|+C.$$

Следовательно,

$$\int \frac{dx}{\sqrt{x^2+\alpha}}=\ln|x+\sqrt{x^2+\alpha}|+C. \qquad (15.12)$$

Пример 5. Найти $\int \sqrt{a^2-x^2}\,dx$ $(a>0)$.

Положим $x=a\cos t$, тогда $dx=-a\sin t\,dt$. Следовательно,

$$\int \sqrt{a^2-x^2}dx=\int \sqrt{a^2-a^2\cos^2 t}\,(-a\sin t)\,dt=$$

$$=-a^2\int \sin^2 t dt=-a^2\int \frac{1-\cos 2t}{2}\,dt=$$

$$=-\frac{a^2}{2}\int dt+\frac{a^2}{2}\int \cos 2t\cdot\frac{1}{2}\,d(2t)=$$

$$=-\frac{a^2 t}{2}+\frac{a^2}{4}\sin 2t+C.$$

Возвращаемся к исходиой переменной $x$. Получаем

$$\int \sqrt{a^2-x^2}\,dx=\frac{x}{a}\sqrt{a^2-x^2}-\frac{a^2}{2}\arccos\frac{x}{a}+C. \qquad (15.13)$$

*Метод интегрирования по частям* основан на следующей формуле:

$$\int u\,dv = uv - \int v\,du. \qquad (15.14)$$

Поскольку $d(uv) = u\,dv + v\,du$ или $u\,dv = d(uv) - v\,du$, то $\int u\,dv = uv - \int v\,du$.

Пример 6. Найти $\int x \sin 2x\,dx$. Положим $x = u$, $\sin 2x\,dx = dv$ или $d\left(-\frac{1}{2}\cos 2x\right) = dv$, тогда $dx = du$, $-\frac{1}{2}\cos 2x = v$.

По формуле (15.14) находим

$$\int x \sin 2x\,dx = x\left(-\frac{1}{2}\cos 2x\right) - \int\left(-\frac{1}{2}\cos 2x\right)dx =$$

$$= -\frac{x\cos 2x}{2} + \frac{1}{2}\int \cos 2x\,dx = -\frac{x\cos 2x}{2} +$$

$$+ \frac{1}{4}\int \cos 2x\,d(2x) = -\frac{x\cos 2x}{2} + \frac{\sin 2x}{4} + C.$$

Пример 7. Найти $\int \sqrt{x^2+\alpha}\,dx$.

Положим $\sqrt{x^2+\alpha} = u$, $dx = dv$, тогда $\frac{x\,dx}{\sqrt{x^2+\alpha}} = du$, $x = v$.

По формуле (15.14) получаем

$$\int \sqrt{x^2+\alpha}\,dx = x\sqrt{x^2+\alpha} - \int x \cdot \frac{x}{\sqrt{x^2+\alpha}}\,dx =$$

$$= x\sqrt{x^2+\alpha} - \int \frac{x^2\,dx}{\sqrt{x^2+\alpha}}.$$

Преобразуем последний интеграл:

$$\int \frac{x^2\,dx}{\sqrt{x^2+\alpha}} = \int \frac{(x^2+\alpha)-\alpha}{\sqrt{x^2+\alpha}}\,dx = \int \frac{x^2+\alpha}{\sqrt{x^2+\alpha}}\,dx -$$

$$-\alpha\int \frac{dx}{\sqrt{x^2+\alpha}} = \int \sqrt{x^2+\alpha}\,dx - \alpha\int \frac{dx}{\sqrt{x^2+\alpha}}.$$

Следовательно,

$$\int \sqrt{x^2+\alpha}\,dx = x\sqrt{x^2+\alpha} - \int \sqrt{x^2+\alpha}\,dx + \alpha\int \frac{dx}{\sqrt{x^2+\alpha}},$$

откуда $2\int \sqrt{x^2+\alpha}\,dx = x\sqrt{x^2+\alpha} + \alpha\int \frac{dx}{\sqrt{x^2+\alpha}}$. Принимая во внимание формулу (15.12), получаем

$$\int \sqrt{x^2+\alpha}\,dx = \frac{1}{2}\left[x\sqrt{x^2+\alpha} + \alpha\ln\left|x+\sqrt{x^2+\alpha}\right|\right] + C. \qquad (15.15)$$

## § 15.4. Интегрирование рациональных дробей с квадратным трехчленом в знаменателе

Пусть требуется найти $\int R(x)\,dx$, где

$$R(x)=\frac{P_n(x)}{P_2(x)}=\frac{A_nx^n+A_{n-1}x^{n-1}+\ldots+A_2x^2+A_1x+A_0}{ax^2+bx+c}.$$

Произведя деление, получим

$$R(x)=Q_m(x)+\frac{kx+l}{ax^2+bx+c},$$

где $Q_m(x)$ — многочлен, степень которого ниже степени многочлена $P_n(x)$.

Первообразная от многочлена $Q_m(x)$ находится легко (см. § 15.3), остается найти первообразную функции

$$f(x)=\frac{kx+l}{ax^2+bx+c}.$$

Рассмотрим сначала некоторые основные интегралы:

$$\int\frac{dx}{x^2+a^2},\quad \int\frac{dx}{x^2-a^2},\quad \int\frac{xdx}{x^2+\alpha}.$$

Первый из этих интегралов приводится к табличному простым преобразованием подынтегрального выражения

$$\int\frac{dx}{x^2+a^2}=\int\frac{\frac{1}{a^2}\,dx}{\frac{x^2}{a^2}+1}=\frac{1}{a}\int\frac{d\left(\frac{x}{a}\right)}{1+\left(\frac{x}{a}\right)^2}=$$

$$=\frac{1}{a}\operatorname{arctg}\frac{x}{a}+c;$$

$$\int\frac{dx}{x^2+a^2}=\frac{1}{a}\operatorname{arctg}\frac{x}{a}+c. \tag{15.16}$$

Чтобы найти второй интеграл, преобразуем подынтегральную функцию

$$\frac{1}{x^2-a^2}=\frac{1}{(x-a)(x+a)}=\frac{1}{2a}\,\frac{(x+a)-(x-a)}{(x-a)(x+a)}=$$

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

Следовательно,

$$\int\frac{dx}{x^2-a^2}=\int\frac{1}{2a}\left(\frac{1}{x-a}-\frac{1}{x+a}\right)dx=$$

$$= \frac{1}{2a}\int \frac{d(x-a)}{x-a} - \frac{1}{2a}\int \frac{d(x+a)}{x+a} =$$

$$= \frac{1}{2a}\left[\ln|x-a| - \ln|x+a|\right] + C = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + C;$$

$$\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + C. \qquad (15.17)$$

Третий интеграл также легко приводится к табличному

$$\int \frac{xdx}{x^2+\alpha} = \frac{1}{2}\int \frac{d(x^2+\alpha)}{x^2+\alpha} = \frac{1}{2}\ln|x^2+\alpha| + C;$$

$$\int \frac{xdx}{x^2+\alpha} = \frac{1}{2}\ln|x^2+\alpha| + C. \qquad (15.18)$$

Интегралы (15.16)—(15.18) следует присоединить к таблице основных интегралов и запомнить их.

Неопределенный интеграл

$$\int \frac{kx+l}{ax^2+bx+c}\,dx$$

в случае $k=0$ приводится к интегралу (15.16) или (15.17); в случае $k \neq 0$ — к интегралу (15.18) и одному из интегралов (15.16), (15.17). Основной способ нахождения этого интеграла состоит в предварительном выделении полного квадрата из квадратного трехчлена и применении формул обобщенной таблицы интегралов, соответствующих формулам (15.16)—(15.18).

Выделение полного квадрата осуществляется следующим образом:

$$ax^2+bx+c = a\left[\left(x^2 + 2\cdot\frac{b}{2a}x + \frac{b^2}{4a^2}\right) - \frac{b^2}{4a^2}\right] + c =$$

$$= a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c,$$

$$ax^2+bx+c = a\left[\left(x+\frac{b}{2a}\right)^2 + \frac{1}{a}\left(c - \frac{b^2}{4a}\right)\right].$$

Пример. Найти $\int \frac{x+3}{x^2+2x+5}\,dx$.

$$\int \frac{x+3}{x^2+2x+5}\,dx = \frac{1}{2}\int \frac{2x+6}{x^2+2x+5}\,dx =$$

$$= \frac{1}{2}\int \frac{(x^2+2x+5)'\,dx}{x^2+2x+5} + 2\int \frac{d(x+1)}{(x+1)^2+4} =$$

$$= \frac{1}{2} \ln |x^2 + 2x + 5| + 2 \cdot \frac{1}{2} \operatorname{arctg} \frac{x+1}{2} + c =$$

$$= \frac{1}{2} \ln (x^2 + 2x + 5) + \operatorname{arctg} \frac{x+1}{2} + c.$$

## § 15.5. Интегрирование рациональных функций

Интегрирование рациональных дробей означает нахождение неопределенных интегралов $\int R(x)dx$, где $R(x)$ — правильная рациональная дробь, т. е.

$$R(x) = \frac{P(x)}{Q(x)} = \frac{A_0 + A_1x + A_2x^2 + \ldots + A_nx^n}{B_0 + B_1x + B_2x^2 + \ldots + B_mx^m} \, (n<m). \quad (15.19)$$

Нахождение указанных интегралов основано иа разложении рациональной дроби в сумму элементарных дробей, т. е. дробей вида

$$\frac{A}{(x-a)^\alpha}; \quad (15.20)$$

$$\frac{Bx+C}{(x^2+px+q)^\beta}, \quad (15.21)$$

где $\alpha$, $\beta$ — натуральные числа; $a$, $p$, $q$, $A$, $B$, $C$ — действительные числа; $\frac{p^2}{4} - q < 0$ (корни трехчлена являются комплексными).

Это разложение определяется теоремой, которая здесь приводится без доказательства. Доказательство можно найти в учебниках математического анализа, например [5, 6].

**Теорема 15.2.** Если дана рациональная дробь (15.19) и

$$Q(x) = (x-a_1)^{\alpha_1} \ldots (x-a_r)^{\alpha_r} (x^2 + p_1x + q_1)^{\beta_1} \ldots (x^2 + p_sx + q_s)^{\beta_s},$$

где $a_i \, (i = 1, 2, \ldots, r)$ — попарно различные вещественные корни многочлена $Q(x)$ кратности $\alpha_i$, а $x^2+p_kx+q_k = (x-z_k)(x-\bar{z}_k)$, где $z_k$ и $\bar{z}_k$ $(k=1, 2, \ldots, s)$ — попарно различные при разных $k$ комплексные корни многочлена $Q(x)$ кратности $\beta_k$, то существуют вещественные числа

$$A_i^\alpha \, (i = 1, 2, \ldots, r; \; \alpha = 1, 2, \ldots, \alpha_i), \; B_k^\beta, \; C_k^\beta$$

$$(k = 1, 2, \ldots, s; \; \beta = 1, 2, \ldots, \beta_k)$$

такие, что

$$\frac{P(x)}{Q(x)}=\frac{A_1^1}{x-a_1}+\frac{A_1^2}{(x-a_1)^2}+\ldots+\frac{A_1^{\alpha_1}}{(x-a_1)^{\alpha_1}}+\ldots$$

$$\ldots+\frac{A_r^1}{(x-a_r)}+\ldots+\frac{A_r^2}{(x-a_r)^2}+\ldots+\frac{A_r^{\alpha_r}}{(x-a_r)^{\alpha_r}}+$$

$$+\frac{B_1^1x+C_1^1}{x^2+p_1x+q_1}+\frac{B_1^2x+C_1^2}{(x^2+p_1x+q_1)^2}+\ldots$$

$$\ldots+\frac{B_1^{\beta_1}x+C_1^{\beta_1}}{(x^2+p_1x+q_1)^{\beta_1}}+\ldots+\frac{B_s^1x+C_s^1}{x^2+p_sx+q_s}+$$

$$+\frac{B_s^2x+C_s^2}{(x^2+p_sx+q_s)^2}+\ldots+\frac{B_s^{\beta}x+C_s^{\beta_s}}{(x^2+p_sx+q_s)^{\beta_s}}.$$

Отметим, что каждому действительному корню $a$ кратности $\alpha$ многочлена $Q(x)$ в этом разложении соответствует сумма $\alpha$ элементарных дробей вида (15.20), т. е. сумма

$$\frac{A_1}{x-a}+\frac{A_2}{(x-a)^2}+\ldots+\frac{A_\alpha}{(x-a)^\alpha},$$

а каждой паре комплексно-сопряженных корней $z$ и $\bar{z}$ (таких, что $(x-z)(x-\bar{z})=x^2+px+q$) кратности $\beta$ — сумма элементарных дробей вида (15.21), т. е. сумма

$$\frac{B_1x+C_1}{x^2+px+q}+\frac{B_2x+C_2}{(x^2+px+q)^2}+\ldots+\frac{B_\beta x+C_\beta}{(x^2+px+q)^\beta}.$$

Приведем примеры разложения рациональных функций на элементарные дроби:

$$1)\ \frac{7x^2-x+1}{x^3+1}=\frac{7x^2-x+1}{(x+1)(x^2-x+1)}=$$

$$=\frac{A}{x+1}+\frac{Bx+C}{x^2-x+1},$$

$$\frac{7x^2-x+1}{x^3+1}=\frac{(A+B)x^2+(B+C-A)x+(A+C)}{x^3+1};$$

$$7x^2-x+1=(A+B)x^2+(B+C-A)x+(A+C);$$

$$A+B=7,\ B+C-A=-1,\ A+C=1;$$

$$A=3,\ B=4,\ C=-2;$$

$$\frac{7x^2-x+1}{x^3+1}=\frac{3}{x+1}+\frac{4x-2}{x^2-x+1}.$$

В следующих примерах приводятся вид разложения и окончательные результаты; промежуточные выкладки предоставляются читателю.

$$2)\ \frac{5x^2-14x+11}{x^3-3x^2+3x-1}=\frac{5x^2-14x+11}{(x-1)^3}=$$

$$=\frac{A_1}{x-1}+\frac{A_2}{(x-1)^2}+\frac{A_3}{(x-1)^3}=\frac{5}{x-1}-$$

$$-\frac{4}{(x-1)^2}+\frac{2}{(x-1)^3};$$

$$3)\ \frac{x^3-x^2+3x-1}{x^4+2x^2+1}=\frac{x^3-x^2+3x-1}{(x^2+1)^2}=$$

$$=\frac{B_1x+C_1}{x^2+1}+\frac{B_2x+C_2}{(x^2+1)^2}=\frac{x-1}{x^2+1}+\frac{2x}{(x^2+1)^2}.$$

Пример. Найти $\int\frac{5x^2-14x+11}{x^3-3x^2+3x-1}\,dx$.

Воспользовавшись разложением этой рациональной функции на элементарные дроби и проинтегрировав их, получим

$$\int\frac{5x^2-14x+11}{x^3-3x^2+3x-1}=\int\left[\frac{5}{x-1}-\frac{4}{(x-1)^2}+\right.$$

$$\left.+\frac{2}{(x-1)^3}\right]dx=5\ln|x-1|+\frac{4}{x-1}-\frac{1}{(x-1)^2}+C.$$

## § 15.6. Интегрирование простейших иррациональных функций

Неопределенный интеграл $\int\frac{dx}{\sqrt{Ax^2+Bx+C}}$ выделением полного квадрата в подкоренном выражении и введением новой переменной $u=x+b$, в зависимости от знака $A$, приводится к одному из интегралов:

$$\int\frac{du}{\sqrt{a^2-u^2}}=\arcsin\frac{u}{a}+C; \qquad (15.22)$$

$$\int\frac{du}{\sqrt{u^2+\alpha}}=\ln|u+\sqrt{u^2+\alpha}|+C. \qquad (15.23)$$

Интеграл (15.23) в других обозначениях получен в § 15.3 (см. пример 4). Интеграл (15.22) приводится к табличному

$$\int \frac{dx}{\sqrt{a^2-x^2}} = \int \frac{\frac{1}{a}\,dx}{\frac{1}{a}\sqrt{a^2-x^2}} =$$

$$= \int \frac{d\left(\frac{x}{a}\right)}{\sqrt{1-\left(\frac{x}{a}\right)^2}} = \arcsin\frac{x}{a} + C.$$

Пример 1. Найти $\int \frac{dx}{\sqrt{3x^2+6x+4}}$.

Так как

$$3x^2+6x+4 = 3\left[(x+1)^2 + \frac{1}{3}\right],$$

то, положив $x+1=u$, по формуле (15.23) получим

$$\int \frac{dx}{\sqrt{3x^2+6x+4}} = \frac{1}{\sqrt{3}}\int \frac{d(x+1)}{\sqrt{(x+1)^2+\frac{1}{3}}} =$$

$$= \frac{1}{\sqrt{3}} \ln\left|(x+1) + \sqrt{(x+1)^2+\frac{1}{3}}\right| + C.$$

Неопределенный интеграл $\int\sqrt{Ax^2+Bx+C}$, в зависимости от знака $A$, приводится к одному из интегралов:

$$\int \sqrt{a^2-u^2}\,du = \frac{u}{2}\sqrt{a^2-u^2} - \frac{a^2}{2}\arccos\frac{u}{a} + C; \qquad (15.24)$$

$$\int \sqrt{u^2+\alpha}\,du = \frac{u}{2}\sqrt{u^2+\alpha} + \frac{\alpha}{2}\ln|u+\sqrt{u^2+\alpha}| + C. \qquad (15.25)$$

Эти интегралы также были получены в § 15.3.

Пример 2. Найти $\int\sqrt{x^2+6x+13}\,dx$.

Поскольку $x^2+6x+13=(x^2+6x+9)+4=(x+3)^2+4$, полагая $u=x+3$, по формуле (15.25) находим

$$\int \sqrt{x^2+6x+13}\,dx = \frac{x+3}{2}\sqrt{(x+3)^2+4} +$$

$$+ \frac{4}{2}\ln|(x+3)+\sqrt{(x+3)^2+4}| + C =$$

$$= \frac{x+3}{2}\sqrt{x^2+6x+13} + 2\ln|(x+3)+\sqrt{x^2+6x+13}| + C.$$

Отметим, что

$$\int \frac{f'(x)}{f(x)}\,dx = \int \frac{df(x)}{f(x)} = \ln|f(x)| + C; \qquad (15.26)$$

$$\int \frac{f'(x)}{\sqrt{f(x)}}\,dx = \int f[(x)]^{-\frac{1}{2}}\,df(x) =$$

$$= \frac{[f(x)]^{\frac{1}{2}}}{\frac{1}{2}} + C = 2\sqrt{f(x)} + C. \qquad (15.27)$$

К рассмотренным здесь интегралам приводится также интеграл

$$\int \frac{ax+b}{\sqrt{Ax^2+Bx+C}}\,dx.$$

Пример 3. Найти $\int \frac{9-4x}{\sqrt{5+8x-4x^2}}\,dx.$

Поскольку

$$(5+8x-4x^2)' = 8-8x = -8(x-1),$$

$$9-4x = -4x+4+5 = -4(x-1)+5,$$

$$5+8x-4x^2 = -4[(x^2-2x+1)-1]+5 =$$

$$= -4(x-1)^2+9 = 4\left[\frac{9}{4}-(x-1)^2\right],$$

на основании формул (15.22) и (15.26) получаем

$$\int \frac{9-4x}{\sqrt{5+8x-4x^2}}\,dx = \int \frac{-4(x-1)+5}{\sqrt{5+8x-4x^2}}\,dx =$$

$$= \frac{1}{2}\int \frac{-8(x-1)\,dx}{\sqrt{5+8x-4x^2}} + 5\int \frac{d(x-1)}{2\sqrt{\frac{9}{4}-(x-1)^2}} =$$

$$= \sqrt{5+8x-4x^2} + \frac{5}{2}\arcsin\frac{2(x-1)}{3} + C.$$

## § 15.7. Интегрирование некоторых тригонометрических выражений

Неопределенные интегралы вида

$$\int \sin ax \sin bx\,dx,\quad \int \cos ax \cos bx\,dx,\quad \int \sin ax \cos bx\,dx$$

с помощью тригонометрических формул

$$\sin\alpha\sin\beta = \frac{1}{2}[\cos(\alpha-\beta)-\cos(\alpha+\beta)];$$

$$\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha-\beta)+\cos(\alpha+\beta)];$$

$$\sin\alpha\cos\beta = \frac{1}{2}[\sin(\alpha-\beta)+\sin(\alpha+\beta)]$$

приводятся к интегралам

$$\int \cos kx\,dx = \int \cos kx \frac{1}{k} d(kx) = \frac{1}{k} \sin kx + C;$$

$$\int \sin kx\,dx = -\frac{\cos kx}{k} + C.$$

Пример 1. Найти $\int \cos 8x \cos 6\,xdx$.

Поскольку

$$\cos 8x \cos 6x = \frac{1}{2} (\cos 2x + \cos 14x),$$

то

$$\int \cos 8x \cos 6\,x\,dx = \frac{1}{2} \int (\cos 2x + \cos 14x)\,dx =$$

$$= \frac{\sin 2x}{4} + \frac{\sin 14x}{28} + C.$$

Неопределенные интегралы вида $I_{n,m} = \int \sin^n x \cos^m x dx$, где $n$ и $m$ — натуральные числа, находятся с помощью тригонометрических формул $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$, $\cos^2 x = \frac{1}{2}(1 + \cos 2x)$, $\cos x \sin x = \frac{1}{2} \sin 2x$, если $n$ и $m$ являются четными.

Если хотя бы одно из чисел $n$ или $m$ — нечетное, то от нечетной степени отделяется множитель и вводится новая переменная. В частности, если $m=2k+1$, то

$$I_{n,m} = \int \sin^n x \cos^{2k+1} x dx = \int \sin^n x \cos^{2k} x \cos x dx =$$

$$= \int \sin^n x (1-\sin^2 x)^k d(\sin x) = \int u^n (1-u^2)^k du.$$

Последний интеграл находится непосредственно (как интеграл от алгебраического многочлена).

Пример 2. Найти $\sin^8 x \cos^5 x dx$.

Поскольку одна из степеней является нечетной ($m=5$), интеграл можно найти следующим образом:

$$\int \sin^8 x \cos^5 x dx = \int \sin^8 x \cos^4 x \cos x dx =$$

$$= \int \sin^8 x (1 - \sin^2 x)^2 d(\sin x) =$$

$$= \frac{\sin^9 x}{9} - \frac{2 \sin^{11} x}{11} + \frac{\sin^{13} x}{13} + C.$$

Рассмотрим неопределенный интеграл

$$\int R(\sin x, \cos x) dx,$$

где $R(\sin x, \cos x)$ — рациональная функция от $\sin x$ и $\cos x$.

Введем новую переменную по формуле

$$\operatorname{tg}\frac{x}{2} = t, \tag{15.28}$$

тогда

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = \frac{2\sin\frac{x}{2}\cos\frac{x}{2}}{\cos^2\frac{x}{2} + \sin^2\frac{x}{2}} =$$

$$= \frac{2\operatorname{tg}\frac{x}{2}}{1 + \operatorname{tg}^2\frac{x}{2}} = \frac{2t}{1+t^2}, \quad \sin x = \frac{2t}{1+t^2};$$

$$\cos x = \cos^2\frac{x}{2} - \sin^2\frac{x}{2} = \frac{\cos^2\frac{x}{2} - \sin^2\frac{x}{2}}{\cos^2\frac{x}{2} + \sin^2\frac{x}{2}} =$$

$$= \frac{1 - \operatorname{tg}^2\frac{x}{2}}{1 + \operatorname{tg}^2\frac{x}{2}} = \frac{1-t^2}{1+t^2}, \quad \cos x = \frac{1-t^2}{1+t^2}.$$

Из равенства (15.28) следует, что $x = 2\operatorname{arctg} t$ и $dx = \frac{2dt}{1+t^2}$.

В результате указанной замены переменной получаем

$$\int R(\sin x, \cos x)\,dx =$$

$$= \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2} = \int R_1(t)\,dt,$$

где $R_1(t)$ — рациональная функция переменной $t$.

Пример 3. Найти $\int \frac{1+\sin x - \cos x}{1 - \sin x + \cos x}\,dx$.

Преобразуя подынтегральное выражение, получаем

$$\int \frac{1+\sin x - \cos x}{1 - \sin x + \cos x}\,dx = \int \frac{2 - (1 - \sin x + \cos x)}{1 - \sin x + \cos x}\,dx =$$

$$= 2\int \frac{dx}{1 - \sin x + \cos x} - \int dx.$$

Чтобы найти первый интеграл, применим подстановку (15.28):

$$\int \frac{dx}{1-\sin x+\cos x} = \int \frac{\frac{2}{1+t^2}}{1-\frac{2t}{1+t^2}+\frac{1-t^2}{1+t^2}}\,dt =$$

$$= -\int \frac{d(t-1)}{t-1} = -\ln|t-1|+C =$$

$$= -\ln\left|\operatorname{tg}\frac{x}{2}-1\right|+C.$$

Следовательно,

$$\int \frac{1+\sin x-\cos x}{1-\sin x+\cos x}\,dx = -x-2\ln\left|\operatorname{tg}\frac{x}{2}-1\right|+C.$$

## § 15.8. Некоторые замечания об интегрировании функций

В предыдущих параграфах рассмотрены вопросы интегрирования простейших классов функций. В частности, было показано, что для всякого многочлена существует первообразная, являющаяся многочленом, т. е. элементарной функцией. Неопределенный интеграл от рациональной функции также выражается через элементарные функции (рациональные функции, логарифмы, арктангенсы).

Ограничиваясь рассмотренными классами функций, обращаем внимание на то, что имеются специальные справочники, в которых приведены наиболее распространенные неопределенные интегралы; укажем, например, [14], [15].

Поставим следующие вопросы. Для всякой ли непрерывной функции существует первообразная? Всякая ли первообразная является элементарной функцией? Оказывается, для всякой непрерывной функции существует первообразная (см. § 16.6). Но эта первообразная для некоторых функций не является элементарной функцией. Соответствующие неопределенные интегралы называют «неберущимися», примерами их могут служить следующие:

$$\left.\begin{array}{l} \displaystyle\int \frac{\sin x}{x}\,dx,\quad \int \frac{\cos x}{x}\,dx,\quad \int \frac{dx}{\sqrt{1-k^2\sin^2 x}}\quad (0<k<1), \\ \displaystyle\int e^{-x^2}dx,\ \int \sin x^2\,dx,\ \int \cos x^2\,dx,\ \int \sqrt{1-k^2\sin^2 x}\,dx. \end{array}\right\} \tag{15.29}$$

Каждый из этих интегралов представляет собой функцию, не выражающуюся в элементарных функциях.

Интеграл от дифференциального бинома, т. е. интеграл вида

$$\int x^{\alpha}(a+bx^{\beta})^{\gamma}\,dx, \tag{15.30}$$

где $\alpha$, $\beta$, $\gamma$ — рациональные числа, выражается через элементарные функции в трех случаях: 1) $\gamma$ — целое число; 2) $\frac{\alpha+1}{\beta}$ — целое число; 3) $\frac{\alpha+1}{\beta}+\gamma$ — целое число. Эти случаи интегрируемости указаны Л. Эйлером в 1768 г. В 1853 г. П. Л. Чебышев[1] доказал, что интеграл (15.30), за исключением этих трех случаев, не выражается через элементарные функции.

## § 15.9. Примеры простейших дифференциальных уравнений

Дифференциальным уравнениям будет посвящен специальный (третий) раздел данной книги.

В этом параграфе ограничимся начальными сведениями об обыкновенных дифференциальных уравнениях первого порядка и простейшими примерами.

*Дифференциальным уравнением* называется уравнение, связывающее независимую переменную, искомую функцию и ее производную. Например, уравнения $y'+y-x=0$, $y'-y=0$, $y'-x=0$, $y'=0$ являются дифференциальными уравнениями, так как все они содержат производную $y'$ неизвестной функции $y=y(x)$. Отметим, что некоторые из них могут и не содержать аргумента, функции или того и другого одновременно, но обязательно должны содержать производную или соответствующие дифференциалы. Приведенные уравнения можно записать в виде

$$\frac{dy}{dx}+y-x=0,\quad \frac{dy}{dx}-y=0,\quad \frac{dy}{dx}-x=0,\quad \frac{dy}{dx}=0.$$

*Решением* дифференциального уравнения называется функция $y=\varphi(x)$, обращающая это уравнение в тождество. Так, функция $y=\frac{1}{2}x^2$ является решением уравнения $y'-x=0$, ибо $y'=\left(\frac{1}{2}x^2\right)'=x$ и подстановка производной в это уравнение дает $x-x\equiv 0$. Его решением будет также функция $y=\frac{1}{2}x^2+C$, где

[1] *Пафнутий Львович Чебышев* (1821—1894) — знаменитый русский математик.

$C$ — произвольная постоянная. Эта функция называется *общим решением* данного уравнения. Общее решение уравнения $y'-x=0$ можно получить следующим образом. Представим это уравнение в виде $dy=xdx$, возьмем интегралы от обеих частей, воспользовавшись вторым свойством неопределенного интеграла: $y=\frac{1}{2}x^2+C$.

Читателю предлагается найти общие решения уравнений $y'=0$ и $y'-y=0$.

Пример 1. *Химические реакции первого порядка* (см. [11]).

Обозначим через $a$ начальную концентрацию вещества $A$, через $x$ — количество молей на литр, прореагировавших за время $t$ от начала реакции, тогда действующая масса к этому времени равна $a-x$, а скорость реакции выразится производной $x'_t$.

Согласно закону действующих масс (скорость, с которой вещество вступает в реакцию, пропорциональна концентрации этого вещества, т. е. количеству молей в единице объема) для мономолекулярных реакций, или реакций первого порядка, справедливо уравнение

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

где $k$ — коэффициент пропорциональности.

Чтобы найти неизвестную функцию $x=x(t)$, необходимо решить это дифференциальное уравнение. Переписав уравнение в виде $\frac{dx}{a-x}=kdt$ и взяв неопределенные интегралы от обеих частей, получим $-\ln(a-x)=kt-\ln C_1$, где $\ln C_1=C$ — произвольная постоянная.

Из последнего равенства следует, что

$$x(t)=a-C_1e^{-kt}.$$

Поскольку $x=0$ при $t=0$ (в начальный момент времени нет продукта реакции), то $a-C_1=0$, $C_1=a$.

Следовательно, $x(t)=a(1-e^{-kt})$.

Пример 2. *Химические реакции второго порядка* (см. [11]).

Имеется два вещества $A$ и $B$, начальные концентрации которых, выраженные в молях на единицу объема, равны соответственно $a$ и $b$ (предположим, что $a<b$). За время $t$ образуется $x(t)$ молей продукта химической реакции.

Согласно закону действующих масс для бимолекулярной реакции, или реакции второго порядка, скорость пропорциональна произведению действующих масс $(a-x)$ и $(b-x)$, т. е. $x'_t=k(a-x)(b-x)$, где $k$ — коэффициент пропорциональности.

Найдем функцию $x=x(t)$, удовлетворяющую этому дифференциальному уравнению. Перепишем уравнение в виде

$$\frac{dx}{(a-x)(b-x)}=kdt$$

и возьмем интегралы от обеих частей.

Предварительно функцию $\frac{1}{(a-x)(b-x)}$ разложим на элементарные дроби:

$$\frac{1}{(a-x)(b-x)}=\frac{1}{a-b}\left(\frac{1}{x-a}-\frac{1}{x-b}\right).$$

Интегрируя, находим

$$kt = \frac{1}{a-b} \ln \left| \frac{x-a}{x-b} \right| + C = \frac{1}{a-b} \ln \frac{a-x}{b-x} + C,$$

поскольку $x<a<b$.

Значение постоянной $C$ определяется из начального условия: $x=0$ при $t=0$, т. е. $\frac{1}{a-b} \ln \frac{a}{b} + C = 0$, $C = \frac{1}{b-a} \ln \frac{a}{b}$.

Следовательно,

$$kt = \frac{1}{a-b} \ln \frac{1-\frac{x}{a}}{1-\frac{x}{b}},$$

откуда

$$x(t) = \frac{ab\left[1-e^{(a-b)kt}\right]}{b-ae^{(a-b)kt}}.$$

# Глава 16

# ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

Определенный интеграл — одно из основных понятий современной математики. К этому понятию приводят задачи на нахождение предела интегральной суммы (см. § 16.1).

Определенный интеграл тесно связан с неопределенным интегралом. Эта связь выражается формулой Ньютона — Лейбница, о которой будет идти речь в § 16.7.

## § 16.1. Задачи, приводящие к интегральным суммам и их пределам

**Задача о площади криволинейной трапеции.** Рассмотрим криволинейную трапецию $aABb$ (рис. 16.1), т. е. плоскую фигуру, ограниченную сверху графиком функции $y=f(x)$ $(f(x)\geqslant 0)$, слева и справа — отрезками $aA$ и $bB$ прямых $x=a$, $x=b$, снизу — осью $Ox$. Отрезок $[a, b]$ точками $a=x_0<x_1<x_2\ldots<x_{n-1}<x_n=b$ разобьем на $n$ элементарных отрезков $[a, x_1]$, $[x_1, x_2]$, $\ldots$, $[x_{n-1}, b]$, длины которых обозначим через $\Delta x_k=x_k-x_{k-1}$ $(k=1, 2, 3, \ldots, n)$. В каждом из элементарных отрезков $[x_{k-1}, x_k]$ выберем произвольно точку $\xi_k$ и вычислим в ней значение данной функции $f(\xi_k)$. Произведение $f(\xi_k)\Delta x_k$ выражает площадь прямоугольника с основанием $\Delta x_k$ и высотой $f(\xi_k)$. Со-

ставим сумму всех таких произведений

$$S_n = \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k. \qquad (16.1)$$

Эта сумма (ее называют *интегральной суммой* для функции $y=f(x)$ на отрезке $[a, b]$) выражает площадь ступенчатой фигуры, состоящей из прямоугольников и

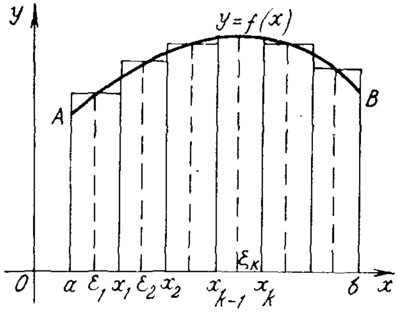



Рис. 16.1

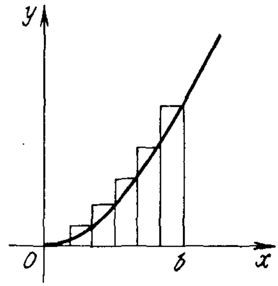



Рис. 16.2

приближенно заменяющей данную трапецию. Очевидно, сумма (16.1) зависит от способа разбиения и выбора точек $\xi_k$.

Обозначим через $\lambda$ длину наибольшего из элементарных отрезков $[x_{k-1}, x_k]$ $(k=1, 2, \ldots, n)$, т. е. $\lambda = \max \Delta x_k$ $(k=1, 2, \ldots, n)$.

Число $S$ называется *пределом интегральной суммы* (16.1), если для любого числа $\varepsilon > 0$ можно указать такое число $\delta > 0$, что при $\lambda < \delta$ выполняется неравенство $|S_n - S| < \varepsilon$ независимо от выбора точек $\xi_k$ на отрезках $[x_{k-1}, x_k]$.

Предположим, что рассматриваемая сумма имеет предел, когда число элементарных отрезков неограниченно возрастает, а длина наибольшего из них стремится к нулю; этот предел и называется *площадью криволинейной трапеции*, т. е.

$$S = \lim_{\lambda \to 0} \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k. \qquad (16.2)$$

Покажем, например, как вычислить площадь $S$ криволинейной трапеции, ограниченной дугой параболы $y=x^2$, прямой $x=b$ и осью $Ox$ (рис. 16.2).

Отрезок $[0,\ b]$ разобьем на $n$ равных частей длиной $\Delta x_k = \frac{b}{n}$ точками $x_k = k\frac{b}{n}$. Возьмем $\xi_k = x_k$, тогда

$$f(\xi_k) = \frac{k^2 b^2}{n^2}$$

и

$$S_n = \sum_{k=1}^{n} \frac{k^2 b^2}{n^2} \frac{b}{n} = \sum_{k=1}^{n} \frac{k^2 b^3}{n^3}.$$

Поскольку (см., например, [15], стр. 15)

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

то

$$S_n = \frac{b^3}{6} \frac{n(n+1)(2n+1)}{n^3} =$$

$$= \frac{b^3}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right), \ \lim_{n\to\infty} S_n = \frac{b^3}{3}; \ \ S = \frac{b^3}{3}.$$

**Задача о вычислении длины пути по заданной скорости.** Пусть точка $M$ движется прямолинейно с переменной скоростью $v = f(t)$. Вычислим длину пути, пройденного точкой $M$ за промежуток времени от $t_0$ до $T$. Промежуток $[t_0,\ T]$ разобьем на $n$ элементарных промежутков $[t_0,\ t_1]$, $[t_1,\ t_2]$, $\ldots$, $[t_{n-1},\ T]$ длиной $\Delta t_k = t_k - t_{k-1}$ $(t_n = T)$. В течение малого промежутка времени $\Delta t_k$ скорость движения можно приближенно считать постоянной и равной $f(t_k')$, где $t_k'$ — некоторое значение $t$ из промежутка $[t_{k-1},\ t_k]$, поэтому длина пути, пройденного за этот промежуток, приближенно равна $f(t_k')\,\Delta t_k$. Складывая все частичные длины $f(t_k')\,\Delta t_k$, получаем приближенное значение длины пути, пройденного точкой за промежуток от $t_0$ до $T$:

$$s_n \approx \sum_{k=1}^{n} f(t_k')\,\Delta t_k. \qquad (16.3)$$

Переходя к пределу, находим точное значение длины пути

$$s = \lim_{\lambda\to 0} \sum_{k=1}^{n} f(t_k')\,\Delta t_k, \qquad (16.4)$$

где $\lambda = \max \Delta t_k$ $(k = 1, 2, \ldots, n)$.

Итак, пройденный путь равен пределу интегральной суммы (16.3) функции $v=f(t)$ на отрезке $[t_0, T]$.

Поскольку к необходимости рассмотрения пределов интегральных сумм приводят многие другие задачи, то эти пределы стали предметом особого исследования. Называют их определенными интегралами и обозначают

$$\int_a^b f(x)\,dx = \lim_{\lambda \to 0} \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k.$$

## § 16.2. Понятие определенного интеграла

Пусть дана функция $y=f(x)$, определенная на отрезке $[a, b]$, где $a<b$. Отрезок $[a, b]$ точками $a=x_0<x_1<$ $<x_2<\ldots<x_{n-1}<x_n=b$ разобьем на $n$ элементарных отрезков $[a, x_1]$, $[x_1, x_2]$, ..., $[x_{k-1}, x_k]$, ..., $[x_{n-1}, b]$, длины которых обозначим через $\Delta x_k$, т. е. $\Delta x_k = x_k - x_{k-1}$ $(k=1, 2, \ldots$ $\ldots, n;\ x_0=a,\ x_n=b)$. В каждом из элементарных отрезков $[x_{k-1}, x_k]$ выберем произвольно одну точку $\xi_k$, значение функции в этой точке $f(\xi_k)$ умножим на длину отрезка $\Delta x_k$, получим произведение $f(\xi_k)\Delta x_k$. Составим сумму всех таких произведений

$$S_n = \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k. \tag{16.5}$$

Сумма (16.5) называется *интегральной суммой* для функции $y=f(x)$ на отрезке $[a, b]$.

Обозначим через $\lambda$ длину наибольшего из элементарных отрезков $[x_{k-1}, x_k]$ при данном $n$, т. е. $\lambda = \max \Delta x_k$ $(k=1, 2, \ldots, n)$.

Определенным интегралом от функции $y=f(x)$ на отрезке $[a, b]$ называется конечный предел ее интегральной суммы, когда число элементарных отрезков неограниченно возрастает, а длина наибольшего из них стремится к нулю. Определенный интеграл обозначается символом $\int_a^b f(x)\,dx$ (читается: определенный интеграл от $a$ до $b$); $f(x)$ называется *подынтегральной функцией, $x$ — переменной интегрирования; $a$ — нижним пределом интегрирования, $b$ — верхним пределом интегрирования.*

Следовательно, по определению

$$\int_a^b f(x)\,dx = \lim_{\lambda\to 0}\sum_{k=1}^{n} f(\xi_k)\,\Delta x_k. \qquad (16.6)$$

Из определения следует, что величина определенного интеграла не зависит от обозначения переменной интегрирования, т. е.

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \ldots = \int_a^b f(u)\,du. \qquad (16.7)$$

Функция, для которой существует предел (16.6), называется *интегрируемой* на отрезке $[a, b]$.

Очевидно, если функция $f(x)$ интегрируема на отрезке $[a, b]$, то она и ограничена на этом отрезке. В самом деле, если $f(x)$ не ограничена на отрезке $[a, b]$, то она не ограничена на некотором элементарном отрезке $[x_{k-1}, x_k]$. За счет выбора точки $\xi_k$ интегральную сумму можно сделать сколь угодно большой, а такая интегральная сумма не имеет конечного предела.

На примерах можно показать, что обратное утверждение неверно: существуют ограниченные функции, не являющиеся интегрируемыми. К ним относится функция Дирихле, равная единице в рациональных точках и нулю — в иррациональных. На любом отрезке $[a, b]$ эта функция ограничена, но не является интегрируемой на нем. Действительно, если в каждом элементарном отрезке $[x_{k-1}, x_k]$ выбрать рациональную точку $\xi_k$, то интегральная сумма

$$S_n = \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k = \sum_{k=1}^{n}\Delta x_k = \sum_{k=1}^{n}(x_k - x_{k-1}) =$$
$$= (x_1 - x_0) + (x_2 - x_1) + \ldots + (x_{n-1} - x_{n-2}) +$$
$$+ (x_n - x_{n-1}) = x_n - x_0 = b - a.$$

Если выбрать иррациональную точку $\xi_k$, то $f(\xi_k) = 0$ и $S_n = 0$. Предел интегральных сумм для функции Дирихле не существует, поэтому функция не является интегрируемой.

Отметим без доказательств, что справедливы следующие утверждения:

1. Если функция $f(x)$ интегрируема на отрезке $[a, b]$,

то она интегрируема на любом отрезке $[c, d]$, содержащемся в $[a, b]$.

2. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она и интегрируема на этом отрезке.

3. Если функция $f(x)$ имеет на отрезке $[a, b]$ конечное число точек разрыва первого рода, то она интегрируема на $[a, b]$.

## § 16.3. Геометрический смысл определенного интеграла

Возвратимся к задаче о площади криволинейной трапеции и определению определенного интеграла.

Если $a<b$ и $f(x) \geqslant 0$, то из формул (16.2) и (16.6) следует, что

$$\int_a^b f(x)\,dx = S, \qquad (16.8)$$

т. е. определенный интеграл от функции $y=f(x)$ по отрезку $[a, b]$ равен площади криволинейной трапеции,

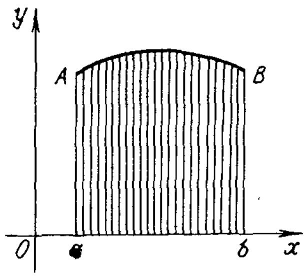



Рис. 16.3

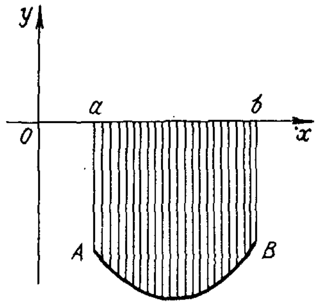



Рис. 16.4

ограниченной сверху графиком функции $y=f(x)$, слева и справа — отрезками прямых $x=a$, $x=b$, снизу — отрезком оси $Ox$ (рис. 16.3).

Если $a<b$ и $f(x) \leqslant 0$, то

$$\int_a^b f(x)\,dx = -S, \qquad (16.9)$$

т. е. определенный интеграл от функции, принимающей неположительные значения, равен площади соответствующей криволинейной трапеции, взятой со знаком минус (рис. 16.4).

Если $a<b$ и $f(x)$ меняет знак на отрезке $[a, b]$, то определенный интеграл равен алгебраической сумме

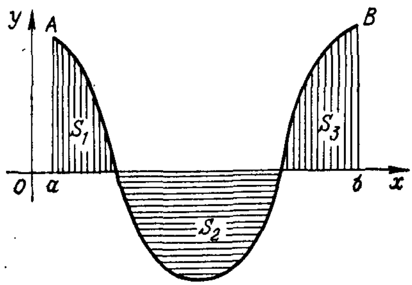



Рис. 16.5

площадей соответствующих криволинейных трапеций (рис. 16.5):

$$\int_a^b f(x)\,dx = S_1 - S_2 + S_3. \qquad (16.10)$$

## § 16.4. Основные свойства определенного интеграла

При введении понятия определенного интеграла предполагалось, что $a<b$. Рассмотрим случаи, когда $a=b$ и $a>b$. Полагаем по определению

$$\int_a^a f(x)\,dx = 0, \qquad (16.11)$$

где $f(x)$ — любая функция;

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx, \qquad (16.12)$$

где $f(x)$ — функция, интегрируемая на отрезке $[b, a]$ $(b<a)$.

Замечание. Равенство (16.11) имеет следующий геометрический смысл: соответствующая криволинейная трапеция вырождается в отрезок прямой, площадь такой «криволинейной трапеции» равна нулю.

Равенство (16.12) можно интерпретировать следующим образом: эти определенные интегралы являются пределами интегральных сумм, отличающнхся лишь знаком, так как все $\Delta x_k = x_k - x_{k-1}$ в случае $b < a$ (рис. 16.6) будут отрицательными (при $a < b$ все $\Delta x_k > 0$), поэтому интегралы отличаются только знаком.

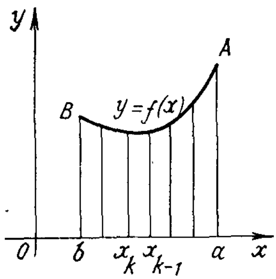



Рис. 16.6

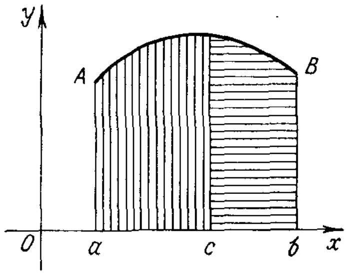



Рис. 16.7

Определенный интеграл обладает следующими свойствами:

1. Если функция $f(x)$ интегрируема на наибольшем из отрезков $[a, b]$, $[a, c]$, $[c, b]$, то она интегрируема на двух других отрезках, причем

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

при любом расположении точек $a$, $b$, $c$.

При $a < c < b$ это равенство имеет простой геометрический смысл (рис. 16.7). Другие случаи сводятся к данному. Например, если $a < b < c$, то

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

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

2. Если функция $f(x)$ интегрируема на отрезке $[a, b]$, то функция $kf(x)$, где $k=\text{const}$, также интегрируема на этом отрезке, причем $\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx$.

3. Если $f(x)$ и $\varphi(x)$ — функции, интегрируемые на отрезке $[a, b]$, то их сумма и разность также интегрируемы на этом отрезке, причем

$$\int_a^b [f(x) \pm \varphi(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b \varphi(x)\,dx.$$

Эти свойства доказываются путем построения соответствующих интегральных сумм и перехода к пределу при $\lambda \to 0$.

Докажем, например, свойство 3. Отрезок $[a, b]$ разобьем на элементарные отрезки $[x_{k-1}, x_k]$. Составим интегральные суммы для функций $f(x)$, $\varphi(x)$, $f(x) \pm \varphi(x)$. Точки $\xi_k \in [x_{k-1}, x_k]$ выберем произвольно, но одни и те же для всех сумм, тогда получим

$$\sum_{k=1}^{n} [f(\xi_k) \pm \varphi(\xi_k)]\,\Delta x_k = \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k \pm \sum_{k=1}^{n} \varphi(\xi_k)\,\Delta x_k.$$

Каждая из интегральных сумм в правой части имеет предел при $\lambda \to 0$, равный соответствующему определенному интегралу, поэтому интегральная сумма в левой части также имеет предел, т. е. функции $f(x) \pm \varphi(x)$ интегрируемы на отрезке $[a, b]$. Переходя к пределу при $\lambda \to 0$, получаем равенство, которое требовалось доказать.

Следующие свойства определенного интеграла выражаются с помощью неравенств.

4. Если функция $f(x)$ интегрируема на отрезке $[a, b]$, где $a < b$, и $f(x) \geqslant 0$ для всех $x \in [a, b]$, то

$$\int_a^b f(x)\,dx \geqslant 0.$$

Действительно, в этом случае неотрицательны также интегральные суммы и их предел.

5. Если функции $f(x)$, $\varphi(x)$ интегрируемы на отрезке $[a, b]$, где $a < b$, и $f(x) \leqslant \varphi(x)$ для всех $x \in [a, b]$, то $\int_a^b f(x)\,dx \leqslant \int_a^b \varphi(x)\,dx$.

Это свойство следует из предыдущего, так как в данном случае $\varphi(x)-f(x)\geqslant 0$ на отрезке $[a, b]$.

6. Если функция $f(x)$ интегрируема на отрезке $[a, b]$, где $a<b$, то функция $|f(x)|$ также интегрируема на $[a, b]$, причем

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx.$$

Это свойство примем без доказательства. Последнее неравенство легко интерпретировать с помощью геометрического смысла определенного интеграла (см. § 16.3).

## § 16.5. Оценка определенного интеграла. Теорема о среднем

Докажем сначала равенство

$$\int_a^b dx = b-a, \tag{16.13}$$

которое понадобится в последующем.

Подынтегральная функция равна единице, $f(x)=1$; интегральная сумма для нее выразится формулой

$$\sum_{k=1}^{n} f(\xi_k)\,\Delta x_k = \sum_{k=1}^{n} \Delta x_k = \sum_{k=1}^{n} (x_k - x_{k-1}) = b-a.$$

Итак, любая интегральная сумма для данной функции равна $b-a$, поэтому и предел ее равен $b-a$, т. е. справедливо равенство (16.13).

**Теорема 16.1.** Если функция $f(x)$ интегрируема на отрезке $[a, b]$, где $a<b$, и для всех $x\in[a, b]$ выполняется неравенство

$$m\leqslant f(x)\leqslant M, \tag{16.14}$$

то

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a). \tag{16.15}$$

Доказательство. На основании свойства 5 из неравенств (16.14) находим, что

$$\int_a^b m\,dx \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b M\,dx,$$

или

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

Принимая во внимание формулу (16.13), получаем неравенства (16.15), что и требовалось доказать.

С помощью неравенств (16.15) можно оценить определенный интеграл, т. е. указать границы, между которыми заключено его значение. Неравенства (16.15) выражают оценку определенного интеграла.

Пример. Оценить определенный интеграл $\int_0^\pi (3+\sin^6 x)\,dx$.

Поскольку в данном случае $a=0$, $b=\pi$, $b-a=\pi$, для функции $f(x)=3+\sin^6 x$ $m=3$, $M=4$, т. е. $3\leqslant 3+\sin^6 x\leqslant 4$, то $3\pi \leqslant \int_0^\pi (3+\sin^6 x)\,dx \leqslant 4\pi$.

**Теорема 16.2.** Если функция $f(x)$ интегрируема на отрезке $[a, b]$ и для всех $x\in[a, b]$ выполняются неравенства $m\leqslant f(x)\leqslant M$, то

$$\int_a^b f(x)\,dx = \mu(b-a), \tag{16.16}$$

где $m\leqslant\mu\leqslant M$.

Доказательство. Если $a<b$, то на основании теоремы 16.1 выполняются неравенства (16.15), из которых находим

$$m \leqslant \frac{1}{b-a}\int_a^b f(x)\,dx \leqslant M.$$

Полагая

$$\frac{1}{b-a}\int_a^b f(x)\,dx = \mu, \tag{16.17}$$

получаем равенство, которое требовалось доказать.

В случае $a>b$ имеем $\int_b^a f(x)\,dx=\mu(a-b)$, $\int_a^b f(x)dx=$ $=\mu(b-a)$. Доказанная теорема называется *теоремой о среднем.*

Замечание. В случае, когда функция $f(x)$ непрерывна на отрезке $[a, b]$, равенство (16.16) принимает вид

$$\int_a^b f(x)\,dx = f(c)(b-a), \qquad (16.18)$$

где $c \in [a, b]$.

Число $\mu = f(c)$, определяемое формулой (16.18), называется *средним значением* функции $f(x)$ на отрезке $[a, b]$.

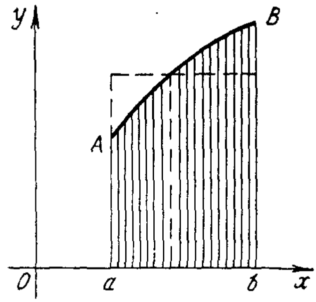



Рис. 16.8

Равенство (16.17) имеет следующий геометрический смысл: площадь криволинейной трапеции, ограниченной непрерывной линией $y=f(x)$ $(f(x) \geqslant 0)$, равна площади прямоугольника с тем же основанием и высотой, равной ординате некоторой точки этой линии (рис. 16.8).

## § 16.6. Существование первообразной для непрерывной функции

Введем сначала понятие определенного интеграла с переменным верхним пределом.

Рассмотрим функцию $y=f(x)$, интегрируемую на отрезке $[a, b]$. Если $x \in [a, b]$, то функция $f(x)$ интегрируема также на любом отрезке $[a, x]$. Предположим, что $x$ меняется на отрезке $[a, b]$, тогда на этом отрезке определена функция

$$\Phi(x) = \int_a^x f(t)\,dt^{1)}. \qquad (16.19)$$

Докажем, что функция (16.19) непрерывна на отрезке $[a, b]$. Аргументу $x$ придадим приращение $\Delta x$ такое, что $(x_0+\Delta x) \in [a, b]$, тогда по свойству 1 определенного интеграла получим

$$\Phi(x+\Delta x) = \int_a^{x+\Delta x} f(t)\,dt =$$

$$= \int_a^x f(t)\,dt + \int_x^{x+\Delta x} f(t)\,dt = \Phi(x) + \int_x^{x+\Delta x} f(t)\,dt,$$

1) Переменную интегрирования обозначили буквой $t$, переменный верхний предел — буквой $x$.

$$\Phi(x+\Delta x)-\Phi(x)=\int\limits_x^{x+\Delta x} f(t)\,dt,\quad \Delta\Phi=\int\limits_x^{x+\Delta x} f(t)\,dt.$$

Применяя теорему о среднем, находим

$$\int\limits_x^{x+\Delta x} f(t)\,dt=\mu\Delta x,$$

где $m\leqslant\mu\leqslant M$, $m$ — наименьшее, $M$ — наибольшее значение функции $f(x)$ на отрезке $[x,\ x+\Delta x]$; эти значения существуют, так как функция интегрируема, следовательно, и ограничена.

Из двух последних равенств следует, что $\Delta\Phi=\mu\Delta x$, откуда $\Delta\Phi\to 0$ при $\Delta x\to 0$, т. е. $\Phi(x)$ — непрерывная функция. Функция (16.19) будет и дифференцируемой, если $f(x)$ — непрерывная функция, о чем свидетельствует следующая теорема.

**Теорема 16.3.** Если подынтегральная функция непрерывна, то производная определенного интеграла с переменным верхним пределом существует и равна значению подынтегральной функции для этого предела, т. е.

$$\Phi'(x)=f(x). \tag{16.20}$$

Доказательство. Аргументу $x$ функции (16.19) придадим приращение $\Delta x$ такое, что $(x+\Delta x)\in[a,\ b]$, ему соответствует приращение функции

$$\Delta\Phi=\Phi(x+\Delta x)-\Phi(x)=\int\limits_x^{x+\Delta x} f(t)\,dt.$$

Применяя формулу (16.18), получаем $\int\limits_x^{x+\Delta x} f(t)\,dt=f(\xi)\,\Delta x$, $\xi\in[x,\ x+\Delta x]$, $\xi=x+\theta\Delta x$, $0<\theta<1$.

Итак, $\Delta\Phi=f(\xi)\Delta x$, откуда

$$\frac{\Delta\Phi}{\Delta x}=f(\xi),\quad \lim_{\Delta x\to 0}\frac{\Delta\Phi}{\Delta x}=\lim_{\Delta x\to 0} f(\xi)=\lim_{\Delta x\to 0} f(x+\theta\Delta x)=f(x),$$

т. е.

$$\Phi'(x)=f(x),\ \text{или}\ \left(\int\limits_a^x f(t)\,dt\right)'=f(x),$$

что и требовалось доказать.

Следствие. Определенный интеграл с переменным верхним пределом является одной из первообразных для непрерывной подынтегральной функции. Другими словами, для любой непрерывной функции существует первообразная.

Замечание. Интеграл с переменным верхним пределом интегрирования используется при определении многих новых функций, например,

$$\int_0^x e^{-t^2}\,dt, \quad \int_0^x \frac{\sin t}{t}\,dt, \quad \int_1^x \frac{\cos t}{t}\,dt, \quad \int_0^x \sin t^2\,dt, \quad \int_0^x \cos t^2\,dt.$$

Эти функции не являются элементарными; как уже отмечалось (см. § 15.8), первообразные указанных подынтегральных функций не выражаются через элементарные функции.

## § 16.7. Формула Ньютона—Лейбница

Связь между определенным и неопределенным интегралом выражает следующая теорема.

**Теорема 16.4.** Определенный интеграл от непрерывной функции равен разности значений любой ее первообразной для верхнего и нижнего предела интегрирования.

Доказательство. Рассмотрим интеграл с переменным верхним пределом

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

где $x$ принадлежит отрезку $[a, b]$, в котором $f(x)$ непрерывна. Согласно теореме 16.3 $\Phi'(x) = f(x)$.

Пусть $F(x)$ — первообразная функции $f(x)$, т. е. $F'(x) = f(x)$. Итак, функции $\Phi(x)$ и $F(x)$ имеют одинаковые производные; на основании следствия 2 из теоремы Лагранжа заключаем, что

$$\Phi(x) = F(x) + C.$$

Поскольку $\Phi(a) = 0$ (см. равенство (16.11)), при $x = a$ из последнего равенства получаем $0 = F(a) + C$, откуда $C = -F(a)$.

Следовательно,

$$\Phi(x) = F(x) - F(a), \quad \int_a^x f(t)\,dt = F(x) - F(a),$$

где $F'(x) = f(x)$.

В частности, при $x=b$ получаем

$$\int_a^b f(x)\,dx = F(b) - F(a), \tag{16.21}$$

что и требовалось доказать.

Формула (16.21) называется *формулой Ньютона — Лейбница;* ее можно переписать в виде

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b,$$

где

$$F(x)\Big|_a^b = F(b) - F(a), \tag{16.22}$$

левая часть читается так: двойная подстановка от $a$ до $b$ для функции $F(x)$.

Приведем пример применения формулы Ньютона — Лейбница:

$$\int_0^b x^2\,dx = \frac{x^3}{3}\Big|_0^b = \frac{b^3}{3}$$

(выясните геометрический смысл этого определенного интеграла и сравните решение с решением соответствующего примера из § 16.1).

## § 16.8. Замена переменной в определенном интеграле. Интегрирование по частям

**Теорема 16.5.** Если выполнены условия: 1) функция $f(x)$ непрерывна на отрезке $[a, b]$; 2) отрезок $[a, b]$ является множеством значений функции $x=\varphi(t)$, определенной на отрезке $\alpha \leqslant t \leqslant \beta$ и имеющей на нем непрерывную производную; 3) $\varphi(\alpha)=a$, $\varphi(\beta)=b$, то справедлива формула

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f[\varphi(t)]\,\varphi'(t)\,dt. \tag{16.23}$$

Доказательство. Пусть $F(x)$ — первообразная для функции $f(x)$ на отрезке $[a, b]$, т. е. $F'(x)=f(x)$ для всех $x \in [a, b]$, тогда $\int_a^b f(x)\,dx = F(b) - F(a)$.

Поскольку функции $F(x)$, $x=\varphi(t)$ являются дифференцируемыми на соответствующих отрезках, сложная функция $F[\varphi(t)]$ дифференцируема на отрезке $[a, b]$:

$$F'_t = F'_\varphi \varphi'_t = F'_x \varphi'_t = f(x)\cdot\varphi'_t = f[\varphi(t)]\,\varphi'_t,$$

$$F'_t = f[\varphi(t)]\,\varphi'(t).$$

Следовательно, функция $F[\varphi(t)]$ — первообразная для функции $f[\varphi(t)]\varphi'(t)$ на отрезке $[\alpha, \beta]$. По формуле Ньютона — Лейбница получаем

$$\int_\alpha^\beta f[\varphi(t)]\,\varphi'(t)\,dt = F[\varphi(t)]\Big|_\alpha^\beta = F[\varphi(\beta)] - F[\varphi(\alpha)].$$

Принимая во внимание условие 3) данной теоремы, находим

$$F[\varphi(\beta)] - F[\varphi(\alpha)] = F(b) - F(a) = \int_a^b f(x)\,dx.$$

Из двух последних равенств следует формула (16.23), которую называют формулой замены переменной под знаком определенного интеграла.

Пример 1. Вычислить $\int_0^3 \frac{x}{\sqrt{x+1}}\,dx$.

Введем новую переменную $t$ по формуле $t=\sqrt{x+1}$, из которой $x=t^2-1$, $dx=2t dt$. Определяем новые пределы интегрирования: из равенства $t=\sqrt{x+1}$ при $x_1=0$ получаем $t_1=1$, при $x_2=3$ $t_2=2$, т. е. $\alpha=1$, $\beta=2$.

На основании формулы (16.23) находим

$$\int_0^3 \frac{x}{\sqrt{x+1}}\,dx = \int_1^2 \frac{t^2-1}{t}\cdot 2t\,dt = 2\int_1^2 (t^2-1)\,dt.$$

Последний интеграл сводится к двум табличным:

$$\int_1^2 (t^2-1)\,dt = \int_1^2 t^2\,dt - \int_1^2 dt =$$

$$= \frac{t^3}{3}\Big|_1^2 - t\Big|_1^2 = \frac{1}{3}(2^3-1^3) - (2-1) = \frac{4}{3}.$$

Следовательно, $\int_0^3 \frac{x}{\sqrt{x+1}}\,dx = \frac{8}{3}$.

**Теорема 16.6.** Если функции $u=u(x)$, $v=v(x)$ имеют непрерывные производные на отрезке $[a, b]$, то справедлива формула

$$\int_a^b u(x)v'(x)\,dx = u(x)v(x)\Big|_a^b - \int_a^b v(x)u'(x)\,dx. \quad (16.24)$$

Доказательство. Поскольку функция $u(x)v(x)$ — первообразная для функции $u'(x)v(x)+u(x)v'(x)$, то

$$\int_a^b [u(x)v'(x) + u'(x)v(x)]\,dx = u(x)v(x)\Big|_a^b,$$

откуда и следует формула (16.24), которую можно записать в виде

$$\int_a^b u\,dv = uv\Big|_a^b - \int_a^b v\,du. \quad (16.25)$$

Пример 2. Вычислить $\int_0^1 xe^{2x}dx$.

С помощью формулы (16.25) получаем

$$\int_0^1 xe^{2x}\,dx = \frac{1}{2}\int_0^1 x\,d(e^{2x}) = \frac{1}{2}xe^{2x}\Big|_0^1 - $$

$$-\frac{1}{2}\int_0^1 e^{2x}\,dx = \frac{1}{2}xe^{2x}\Big|_0^1 - \frac{1}{4}\int_0^1 e^{2x}\,d(2x) =$$

$$= \frac{1}{2}xe^{2x}\Big|_0^1 - \frac{1}{4}e^{2x}\Big|_0^1 = \frac{1}{4}(e^2+1) \approx 2{,}097.$$

## § 16.9. Приближенное вычисление определенных интегралов

Точное вычисление определенного интеграла по формуле Ньютона — Лейбница не всегда оказывается возможным (так как первообразная подынтегральной функции иногда не выражается в элементарных функциях) или не представляется целесообразным (поскольку нахождение первообразной часто сопряжено с громоздкими выкладками). В этих случаях, а также в случае, когда подынтегральная функция задана табличным способом, целесообразно определенные интегралы вычислять приближенно. Существуют различные методы численного интегрирования функций. Рассмотрим применение простейших из них.

**Формула трапеций.** Пусть требуется вычислить определенный интеграл

$$I = \int_a^b f(x)\,dx, \tag{16.26}$$

который, как известно, равен площади соответствующей криволинейной трапеции (рис. 16.9).

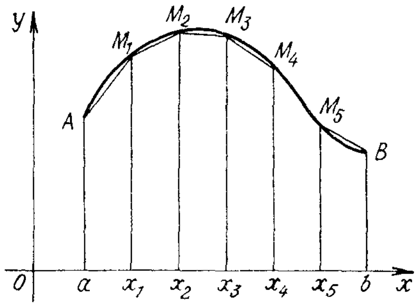



Рис. 16.9

Отрезок $[a, b]$ разобьем на $n$ равных элементарных отрезков длиной

$$h = \frac{b-a}{n}. \tag{16.27}$$

Пусть $x_0, x_1, x_2, \ldots, x_n$ — точки деления ($x_0 = a$, $x_n = b$); $y_0, y_1, y_2, \ldots, y_n$ — ординаты соответствующих точек $M_0, M_1, M_2, \ldots, M_n$ ($M_0 = A$, $M_n = B$) графика функции $y = f(x)$, тогда

$$x_k = a + kh \quad (k = 0, 1, 2, \ldots, n), \tag{16.28}$$

где $h$ — шаг разбиения, определяемый формулой (16.27),

$$y_k = f(x_k) \quad (k = 0, 1, 2, \ldots, n). \tag{16.29}$$

Соединим каждые две последовательные точки $M_{k-1}$, $M_k$ ($k = 1, 2, \ldots, n$) отрезком прямой, в результате получим $n$ прямолинейных трапеций одной и той же высоты. Они составляют фигуру, «вписанную» в данную криволинейную трапецию. Площадь этой фигуры выражается формулой

$$\frac{y_0+y_1}{2}h+\frac{y_1+y_2}{2}h+\ldots+\frac{y_{n-1}+y_n}{2}h=$$

$$=h\left(\frac{y_0}{2}+y_1+y_2+\ \ldots\ +y_{n-1}+\frac{y_n}{2}\right),$$

поэтому

$$\int_a^b f(x)\,dx\simeq\frac{h}{2}\,[(y_0+y_n)+2\,(y_1+y_2+\ \ldots\ +y_{n-1})]. \qquad (16.30)$$

Формула (16.30) и называется *формулой трапеций.*

Отметим без доказательства, что если $R_n$ — остаточный член приближенной формулы (16.30), т. е.

$$\int_a^b f(x)\,dx=\frac{h}{2}\,[y_0+2\,(y_1+y_2+\ \ldots\ +y_{n-1})+y_n]+R_n, \qquad (16.31)$$

то

$$|R_n|\leqslant\frac{(b-a)^3M}{12n^2}, \qquad (16.32)$$

где

$$M=\max_{a<x<b}|f''(x)|. \qquad (16.33)$$

**Формула парабол.** Определенный интеграл (16.26) можно вычислить приближенно и другими способами.

Отрезок $[a,\ b]$ разобьем на четное число $m=2n$ равных элементарных отрезков, длина каждого из которых

$$h=\frac{b-a}{2n}. \qquad (16.34)$$

Криволинейную трапецию $aAM_2x_2$ (рис. 16.10) заменим соответствующей параболической трапецией, ограниченной сверху дугой параболы, проходящей через три точки $A$, $M_1$, $M_2$, ось которой параллельна оси $Oy$. Уравнение такой параболы имеет вид $y=ax^2+$ $+bx+c$, коэффициенты этого уравнения определяются из условия, что парабола проходит через три точки $A$, $M_1$, $M_2$. Аналогичную замену осуществляем для других точек $M_2$, $M_3$, $M_4$; $M_4$, $M_5$, $M_6$ и т. д.

Вычислим площадь одной параболической трапеции. Покажем, что если криволинейная трапеция ограничена параболой $y=ax^2+$

$+bx+c$, осью $Ox$ и прямыми $x=-h$, $x=h$, то ее площадь определяется формулой

$$S=\frac{h}{3}(y_0+4y_1+y_2), \qquad (16.35)$$

где $y_0=y(x_0)=y(-h)$, $y_1=y(x_1)=y(0)$, $y_2=y(x_2)=y(h)$.

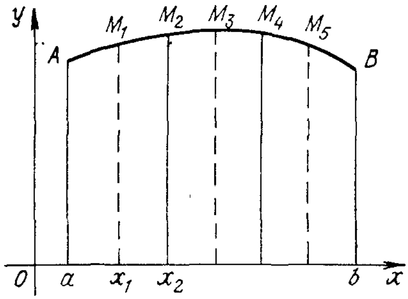



Рис. 16.10

Коэффициенты уравнения параболы $y=ax^2+bx+c$ определяются из системы уравнений

$$y_0=ah^2-bh+c,\ y_1=c,\ y_2=ah^2+bh+c. \qquad (16.36)$$

Считая коэффициенты $a$, $b$, $c$ известными, находим площадь рассматриваемой параболической трапеции

$$S=\int_{-h}^{h}(ax^2+bx+c)\,dx=\left(\frac{ax^3}{3}+\frac{bx^2}{2}+cx\right)\Bigg|_{-h}^{h},$$

$$S=\frac{h}{3}(2ah^2+6c). \qquad (16.37)$$

С другой стороны, из системы (16.36) следует

$$y_0+4y_1+y_2=2ah^2+6c. \qquad (16.38)$$

Из двух последних формул получаем формулу (16.35).

Поскольку

$$\int_a^b f(x)\,dx=\int_a^{x_2} f(x)\,dx+\int_{2}^{x_4} f(x)\,dx+\ldots+\int_{x_{2n-2}}^{x_{2n}} f(x)\,dx,$$

то с помощью формулы (16.35) находим

$$\int_a^b f(x)\,dx\approx\frac{h}{3}(y_0+4y_1+y_2)+\frac{h}{3}(y_2+4y_3+y_4)+\ldots$$

$$\ldots+\frac{h}{3}(y_{2n-2}+4y_{2n-1}+y_{2n})$$

или

$$\int_a^b f(x)\,dx \approx \frac{h}{3}\,[y_0 + y_{2n} + 4\,(y_1 + y_3 + \ldots + y_{2n-1}) +$$
$$+ 2\,(y_2 + y_4 + \ldots + y_{2n-2})]. \qquad (16.39)$$

Это и есть *формула парабол*; она называется также *формулой Симпсона.*

Для остаточного члена формулы (16.39) выполняется неравенство

$$|R_n| \leqslant \frac{(b-a)^5 M}{180\,(2n)^4}, \qquad (16.40)$$

где

$$M = \max_{a \leqslant x \leqslant b} |f^{\mathrm{IV}}(x)|. \qquad (16.41)$$

Пример. По формуле трапеций, приняв $n=10$, вычислить $\int\limits_0^1 \frac{\sin x}{x}\,dx$.

Составим таблицу значений функции $f(x) = \frac{\sin x}{x}$, необходимых для приближенного вычисления данного определенного интеграла (табл. 5).

Таблица 5

| $k$ | $x_k$ | $\sin x_k$ | $y_k = f(x_k) = \frac{\sin x_k}{x_k}$ | |
|---|---|---|---|---|
| | | | $y_0, y_{10}$ | $y_k\ (k=1, 2, \ldots, 9)$ |
| 0 | 0 | 0 | 1 | |
| 1 | 0,1 | 0,09985 | | 0,99850 |
| 2 | 0,2 | 0,19867 | | 0,99335 |
| 3 | 0,3 | 0,29552 | | 0,98507 |
| 4 | 0,4 | 0,38942 | | 0,97355 |
| 5 | 0,5 | 0,47943 | | 0,95886 |
| 6 | 0,6 | 0,56464 | | 0,94107 |
| 7 | 0,7 | 0,64422 | | 0,92031 |
| 8 | 0,8 | 0,71736 | | 0,89670 |
| 9 | 0,9 | 0,78333 | | 0,87037 |
| 10 | 1,0 | 0,84147 | 0,84147 | |
| $\Sigma$ | | | 1,84147 | 8,53778 |

Поскольку $y_0+y_{10}=1{,}84147$, $2(y_1+y_2+\ldots+y_9)=2\times 8{,}53778=$ $=17{,}07556$, $h=0{,}1$, то по формуле (16.30) получаем

$$\int_0^1 \frac{\sin x}{x}\,dx \approx \frac{0{,}1}{2}(1{,}84147+17{,}07556)=0{,}94585.$$

Замечание. Соответствующий неопределенный интеграл $\int \frac{\sin x}{x}\,dx$ не выражается в элементарных функциях (см. § 15.8), поэтому формулу Ньютона — Лейбница здесь применить нельзя.

## § 16.10. Несобственные интегралы

При введении понятия определенного интеграла предполагалось, что выполняются условия: 1) пределы интегрирования $a$ и $b$ являются конечными; 2) подынтегральная функция $f(x)$ ограничена на отрезке $[a, b]$. В этом случае определенный интеграл называют *собственным*. Если хотя бы одно из двух указанных условий не выполняется, интеграл называют *несобственным*.

**Интегралы с бесконечными пределами.** Пусть функция $y=f(x)$ непрерывна при любом $x \geqslant a$. Рассмотрим интеграл с переменным верхним пределом

$$I(b)=\int_a^b f(x)\,dx. \tag{16.42}$$

Как было показано в § 16.4, интеграл (16.42) является дифференцируемой функцией верхнего предела. Предположим, что при $b \to +\infty$ функция (16.42) имеет конечный предел; этот предел называется *сходящимся несобственным интегралом* от функции $f(x)$ по промежутку $[a, +\infty[$ и обозначается

$$\int_a^\infty f(x)\,dx = \lim_{b\to+\infty}\int_a^b f(x)\,dx. \tag{16.43}$$

Если предел (16.43) не существует или равен бесконечности, то несобственный интеграл называется *расходящимся*.

Геометрически несобственный интеграл от неотрицательной функции выражает площадь бесконечной криволинейной трапеции, ограниченной сверху графиком функции $y=f(x)$, слева — отрезком прямой $x=a$, снизу —

осью $Ox$ (рис. 16.11) (в случае сходящегося интеграла эта площадь является конечной, в случае расходящегося — бесконечной).

Пример 1. Исследовать, при каких $\alpha$ сходится несобственный интеграл $\int\limits_1^{\infty} \frac{dx}{x^{\alpha}}$.

Если $\alpha \neq 1$, то

$$\int\limits_1^{\infty} \frac{dx}{x^{\alpha}} = \lim_{b \to +\infty} \int\limits_1^{b} \frac{dx}{x^{\alpha}} = \lim_{b \to +\infty} \left. \frac{x^{-\alpha+1}}{-\alpha+1} \right|_1^b =$$

$$= \lim_{b \to +\infty} \frac{1}{1-\alpha} (b^{1-\alpha} - 1).$$

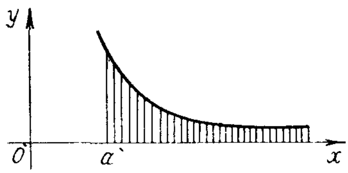



Рис. 16.11

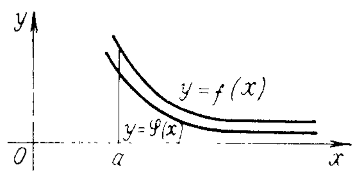



Рис. 16.12

Следовательно,

$$\int\limits_1^{\infty} \frac{dx}{x^{\alpha}} = \frac{1}{\alpha - 1} \text{ при } \alpha > 1, \quad \int\limits_1^{\infty} \frac{dx}{x^{\alpha}} = +\infty \text{ при } \alpha < 1.$$

В случае $\alpha = 1$ $\int\limits_1^{\infty} \frac{dx}{x} = \ln x \Big|_1^{\infty} = +\infty$. Итак, несобственный интеграл сходится при $\alpha > 1$.

Если $F(x)$ — первообразная для $f(x)$, то

$$\int\limits_a^{\infty} f(x)\, dx = \lim_{b \to +\infty} \int\limits_a^{b} f(x) =$$

$$= \lim_{b \to +\infty} [F(b) - F(a)] = F(+\infty) - F(a), \quad (16.44)$$

где $F(+\infty) = \lim\limits_{b \to +\infty} F(b)$.

Аналогично определяется несобственный интеграл с бесконечным нижним пределом

$$\int_{-\infty}^{b} f(x)\,dx = \lim_{a\to-\infty} \int_{a}^{b} f(x)\,dx \tag{16.45}$$

и несобственный интеграл с обоими бесконечными пределами

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{+\infty} f(x)\,dx, \tag{16.46}$$

где $c$ — любая точка из интервала $]-\infty, +\infty[$.

Приведем без доказательства две теоремы, с их помощью можно исследовать вопрос о сходимости некоторых несобственных интегралов.

**Теорема 16.7.** Если при $x \geqslant a$ выполнены неравенства $0 \leqslant \varphi(x) \leqslant f(x)$ и $\int_a^\infty f(x)\,dx$ сходится, то сходится и $\int_a^\infty \varphi(x)\,dx$, причем $\int_a^\infty \varphi(x)\,dx \leqslant \int_a^\infty f(x)\,dx$; если $\int_a^\infty \varphi(x)\,dx$ расходится, то расходится и $\int_a^\infty f(x)\,dx$.

Геометрическое значение этой теоремы иллюстрируется на рис. 16.12.

**Теорема 16.8.** Если в промежутке $[a, +\infty[$ функция $y = f(x)$ меняет знак и $\int_a^\infty |f(x)|\,dx$ сходится, то сходится также $\int_a^\infty f(x)\,dx$.

**Интегралы от неограниченных функций.** Если функция $y=f(x)$ не ограничена в окрестности точки $c$ отрезка $[a, b]$ и непрерывна при $a \leqslant x < c$ и $c < x \leqslant b$, то несобственный интеграл от этой функции определяется формулой

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0} \int_a^{c-\varepsilon} f(x)\,dx + \lim_{\eta\to 0} \int_{c+\eta}^{b} f(x)\,dx, \tag{16.47}$$

где $\varepsilon > 0$, $\eta > 0$. В случае $c=b$ или $c=a$ получаем

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0} \int_a^{b-\varepsilon} f(x)\,dx; \tag{16.48}$$

$$\int_a^b f(x)\,dx = \lim_{\eta\to 0}\int_{a+\eta}^b f(x)\,dx. \qquad (16.49)$$

Несобственный интеграл (16.48) или (16.49) называется *сходящимся*, если существует конечный предел соответствующего определенного интеграла; в противном случае интеграл называется *расходящимся.*

Несобственный интеграл (16.47) называется сходящимся, если существуют и конечны оба предела в правой части.

Для интегралов от неограниченных функций справедливы теоремы, аналогичные теоремам 16.7 и 16.8. Они применяются для исследования вопроса о сходимости несобственных интегралов и оценки их значений. В качестве функции, с которой сравнивают подынтегральную функцию, часто выбирают $\varphi(x) = \frac{1}{(x-a)^\alpha}$. Легко видеть, что $\int_a^b \frac{dx}{(x-a)^\alpha}$ сходится при $\alpha < 1$, расходится при $\alpha \geqslant 1$.

Пример 2. Исследовать, сходится ли интеграл

$$\int_0^2 \frac{dx}{\sqrt[3]{x}+3x^2}.$$

Подынтегральная функция не ограничена в окрестности точки $x=0$. Поскольку

$$\frac{1}{\sqrt[3]{x}+3x^2} < \frac{1}{\sqrt[3]{x}} \quad \text{и} \quad \int_0^2 \frac{dx}{\sqrt[3]{x}} = \int_0^2 \frac{dx}{x^{\frac{1}{3}}}$$

сходится, то сходится и данный интеграл.

## § 16.11. Понятие об эйлеровых интегралах

*Гамма-функция,* или *эйлеров интеграл второго рода,* определяется формулой

$$\Gamma(p) = \int_0^\infty e^{-x}x^{p-1}dx. \qquad (16.50)$$

Этот интеграл является несобственным, так как верхний предел бесконечен; кроме того, если $p-1<0$, то подынтегральная функция не ограничена при $x \to 0$.

Разобьем этот интеграл на два интеграла

$$\Gamma(p)=\int_0^1 e^{-x}x^{p-1}\,dx+\int_1^\infty e^{-x}x^{p-1}\,dx$$

и исследуем сходимость каждого из них.

Поскольку при $0<x<1$ $e^{-x}x^{p-1}=\frac{e^{-x}}{x^{1-p}}<\frac{1}{x^{1-p}}$ и $\int_0^1\frac{dx}{x^{1-p}}$ сходится в случае $1-p<1$ или $p>0$, то первый интеграл сходится, когда $p>0$.

Так как $e^{-x}x^{p-1}<x^{-2}$ при достаточно больших $x$ $\left(\frac{e^{-x}\,x^{p-1}}{x^{-2}}=\frac{x^{p+1}}{e^x}<1,\right.$ ибо $\lim\limits_{x\to+\infty}\frac{x^\alpha}{e^x}=0$ (см. пример 3 § 12.1)$\left.\right)$ и

$$\int_1^\infty\frac{dx}{x^2}=-\frac{1}{x}\Big|_1^\infty=1,$$

то при любом $p$ сходится и второй интеграл.

Следовательно, интеграл (16.50) сходится при $p>0$. Каждому положительному значению $p$ соответствует вполне определенное значение $\Gamma(p)$. Функция $\Gamma(p)$ не является элементарной.

Чтобы получить одно важное свойство этой функции, к интегралу (16.50) применим метод интегрирования по частям:

$$\Gamma(p+1)=\int_0^\infty e^{-x}x^{(p+1)-1}dx=\int_0^\infty e^{-x}x^p\,dx=\int_0^\infty -x^p d(e^{-x})=$$

$$=-e^{-x}x^p\Big|_0^\infty+p\int_0^\infty e^{-x}x^{p-1}\,dx=-e^{-x}x^p\Big|_0^\infty+p\Gamma(p).$$

Поскольку $-e^{-x}x^p\Big|_0^\infty=\lim\limits_{x\to+\infty}\left(0-\frac{x^p}{e^x}\right)=0$, то

$$\Gamma(p+1)=p\Gamma(p). \qquad (16.51)$$

При $p=1$ интеграл находится непосредственно:

$$\Gamma(1)=\int_0^\infty e^{-x}\,dx=-e^{-x}\Big|_0^\infty=1.$$

Подставляя в формулу (16.51) значения $p=1, 2, \ldots, n$, получаем

$$\Gamma(2)=1\cdot\Gamma(1)=1=1!,\ \Gamma(3)=2\Gamma(2)=2\cdot 1=2!,$$
$$\Gamma(4)=3\Gamma(3)=3\cdot 2\cdot 1=3!,$$
$$\Gamma(n+1)=n(n-1)\ldots 2\cdot 1=n! \qquad (16.52)$$

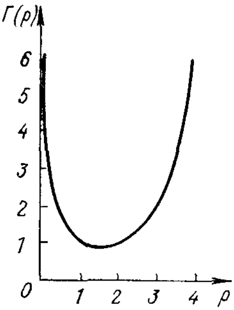



Рис. 16.13

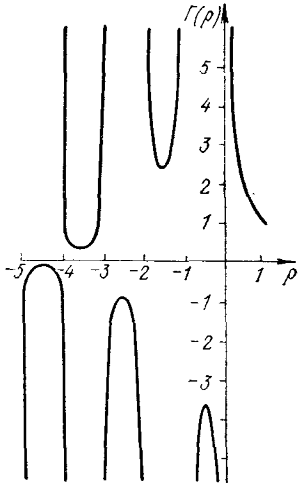



Рис. 16.14

Итак, при натуральных значениях аргумента гамма-функция совпадает с факториалом, т. е. с функцией $f(n)=n!$ Но гамма-функция определена не только при натуральных $n$, но и при любых положительных значениях аргумента.

Из формулы (16.52) следует, что $0!=\Gamma(1)=1$. График гамма-функции изображен на рис. 16.13.

Замечание. Гамма-функция определяется и при отрицательных значениях $p$. В этом случае необходимо применить формулу (16.51), переписав ее в виде

$$\Gamma(p)=\frac{\Gamma(p+1)}{p} \qquad (16.53)$$

(формулой (16.50) пользоваться уже нельзя, поскольку при $p\leqslant 0$ интеграл расходится).

Если $-1<p<0$, то $0<p+1<1$, поэтому правая часть формулы (16.53) имеет смысл, ею и определяется $\Gamma(p)$ при этих значениях $p$, отметим, что в этом случае $\Gamma(p)<0$. Продолжая аналогичные рассуждения, убеждаемся в том, что гамма-функция определена для всех отрицательных значений $p$, кроме $p=-k$, где $k=1, 2, 3, \ldots, n$ и кроме $p=0$.

График гамма-функции при отрицательных значениях $p$ изображен на рис. 16.14.

*Бета-функция,* или *эйлеров интеграл первого рода,* определяется формулой

$$B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx. \qquad (16.54)$$

Подынтегральная функция не ограничена в окрестности точки $x=0$ при $p-1<0$ и в окрестности точки $x=1$ при $q-1<0$.

Разобьем интеграл (16.54) на два интеграла

$$B(p, q) = \int_0^{0,5} x^{p-1}(1-x)^{q-1}\,dx + \int_{0,5}^{1} x^{p-1}(1-x)^{q-1}\,dx.$$

Так как $x^{p-1}(1-x)^{q-1} < cx^{p-1}$ при $0 < x < 0{,}5\, x^{p-1}$ и $\int\limits_0^{0,5} x^{p-1}\,dx = \int\limits_0^{0,5} \frac{dx}{x^{1-p}}$ сходится при $1-p<1$, т. е. при $p>0$, то первый интеграл также сходится при $p>0$. Аналогично можно показать, что второй интеграл сходится при $q>0$.

Следовательно, интеграл (16.54) сходится при $p>0$, $q>0$.

Отметим, что между бета- и гамма-функциями существует связь, выражаемая формулой

$$B(p, q) = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}. \qquad (16.55)$$

Доказательство равенства (16.55) можно найти, например, в [6].

Полагая в формуле (16.55) $p=\frac{1}{2}$, $q=\frac{1}{2}$, получаем

$$\Gamma^2\left(\frac{1}{2}\right) = B\left(\frac{1}{2}, \frac{1}{2}\right) = \int_0^1 x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}}\,dx =$$

$$= \int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \int_0^1 \frac{dx}{\sqrt{x-x^2}} = \int_0^1 \frac{d\left(x-\frac{1}{2}\right)}{\sqrt{\frac{1}{4}-\left(x-\frac{1}{2}\right)^2}} =$$

$$= \arcsin(2x-1)\Big|_0^1 = \pi, \quad \Gamma^2\left(\frac{1}{2}\right) = \pi.$$

Так как $\Gamma(p)>0$ при $p>0$, то

$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}\approx 1{,}772. \qquad (16.56)$$

С помощью формулы (16.56) можно вычислить $\int\limits_0^\infty e^{-x^2}\,dx$.

Действительно, полагая $x=\sqrt{t}$, находим

$$\int\limits_0^\infty e^{-x^2}\,dx=\frac{1}{2}\int\limits_0^\infty e^{-t}t^{-\frac{1}{2}}\,dt=\frac{1}{2}\int\limits_0^\infty e^{-t}t^{\frac{1}{2}-1}\,dt=$$

$$=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2},$$

$$\int\limits_0^\infty e^{-x^2}\,dx=\frac{\sqrt{\pi}}{2}. \qquad (16.57)$$

# Глава 17

# ПРИЛОЖЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

Понятие определенного интеграла широко применяется при решении многих задач. С помощью определенного интеграла можно вычислить площадь криволинейной фигуры, длину дуги кривой, длину прямолинейного пути по заданной переменной скорости, определить объем и поверхность тела вращения, массу и центр тяжести неоднородного стержня, работу, производимую переменной силой. и др.

## § 17.1. Площадь криволинейной фигуры в прямоугольных декартовых координатах

Как было показано в § 16.3, площадь криволинейной трапеции $aABb$, ограниченной сверху графиком функции $y=f(x)$, слева и справа — прямыми $x=a$ и $x=b$ соответственно, снизу — осью $Ox$ (см. рис. 16.1), вычисляется по формуле

$$S=\int\limits_a^b f(x)\,dx. \qquad (17.1)$$

Площадь криволинейной трапеции $cCDd$ (рис. 17.1), ограниченной справа графиком функции $x=\varphi(y)$, сверху и снизу — соответственно прямыми $y=d$, $y=c$, слева — осью $Oy$, определяется формулой

$$S=\int_c^d \varphi(y)\,dy. \qquad (17.2)$$

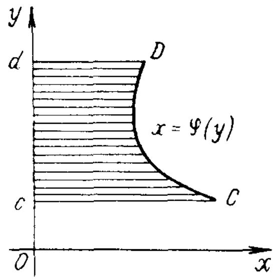


Рис. 17.1

Площадь криволинейной фигуры $A_1A_2B_1B_2$, ограниченной сверху графиком функции $y_2=f_2(x)$, снизу — графиком функции $y_1=f_1(x)$, слева и справа — прямыми $x=a$, $x=b$ (рис. 17.2, *а*), вычисляется по формуле

$$S=\int_a^b [f_2(x)-f_1(x)]\,dx. \qquad (17.3)$$

Эта формула получена с помощью формулы (17.1), так как указанная фигура представляет разность двух криволинейных трапеций.

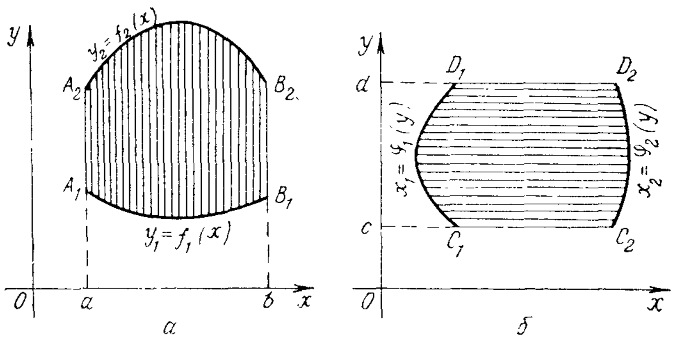


Рис. 17.2

Площадь фигуры $C_1D_1D_2C_2$, ограниченной слева и справа соответственно графиками функций $x_1=\varphi_1(y)$, $x_2=\varphi_2(y)$, снизу и сверху — прямыми $y=c$, $y=d$ (рис. 17.2, *б*), определяется формулой

$$S=\int_c^d [\varphi_2(y)-\varphi_1(y)]\,dy. \qquad (17.4)$$

В более общем случае криволинейную фигуру разбивают на части, площади которых вычисляются по приведенным формулам или определяются непосредственно.

Пример 1. Вычислить площадь криволинейной фигуры, ограниченной параболой $y=x^2$ и прямой $y=x$ (рис. 17.3). Данные линии пересекаются в точках $O(0, 0)$ и $N(1, 1)$. Воспользуемся формулой

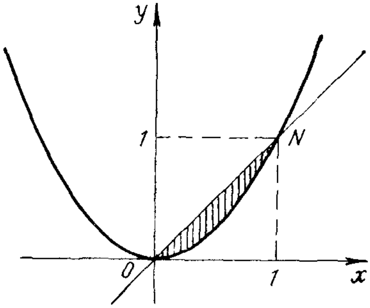



Рис. 17.3

(17.3), поскольку данная фигура — частный случай фигуры $A_1A_2B_2B_1$ (рис. 17.2, $a$), точки $A_1$ и $A_2$ совпадают с точкой $O$, точки $B_1$ и $B_2$ — с точкой $N$. Полагая в этой формуле $y_1=f_1(x)=x^2$, $y_2=f_2(x)=x$, $a=0$, $b=1$, получаем

$$S=\int_0^1 (x-x^2)\,dx=\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\Big|_0^1=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}.$$

Рассмотрим случай, когда линия, ограничивающая криволинейную трапецию сверху (рис. 16.3), задана параметрическими уравнениями

$$x=\varphi_1(t),\quad y=\varphi_2(t), \tag{17.5}$$

где $\alpha\leqslant t\leqslant\beta$, $\varphi_1(\alpha)=a$, $\varphi_1(\beta)=b$.

Уравнения (17.5) определяют некоторую функцию $y=f(t)$ на отрезке $[a, b]$. Площадь криволинейной трапеции вычисляется по формуле

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

Перейдем к новой переменной по формуле $x=\varphi_1(t)$, тогда $dx=\varphi_1'(t)\,dt$. На основании уравнений (17.5) получаем $y=f(x)=f[\varphi_1(t)]=\varphi_2(t)$.

Следовательно,

$$S=\int_\alpha^\beta \varphi_2(t)\,\varphi_1'(t)\,dt. \tag{17.6}$$

Пример 2. Вычислить площадь области, ограниченной эллипсом $x=a\cos t$, $y=b\sin t$.

В силу симметрии эллипса относительно координатных осей достаточно вычислить площадь области, лежащей в первой четверти, и результат умножить на 4. Замечаем, что в этом случае $x$ меняется от 0 до $a$, поэтому $t$ будет меняться от $\frac{\pi}{2}$ до 0, по формуле (17.6) находим

$$S=4\int_{\frac{\pi}{2}}^{0}(b\sin t)(-a\sin t)\,dt=4ab\int_0^{\frac{\pi}{2}}\sin^2 t\,dt=$$

$$=4ab\int_0^{\frac{\pi}{2}}\left(\frac{1-\cos 2t}{2}\right)dt=4ab\left(\frac{t}{2}-\frac{\sin 2t}{4}\right)\Bigg|_0^{\frac{\pi}{2}}=\pi ab,\quad S=\pi ab.$$

Замечание. В частном случае, когда $a=b=R$, получаем $S=\pi R^2$ — площадь круга радиуса $R$.

## § 17.2. Площадь в полярных координатах

Рассмотрим криволинейный сектор $OAB$ (рис. 17.4), ограниченный линией, заданной уравнением $\rho=\rho(\varphi)$ в полярных координатах, двумя лучами $OA$ и $OB$, для которых соответственно $\varphi=\alpha$, $\varphi=\beta$. Сектор $OAB$ разобьем на элементарные секторы $OM_{k-1}M_k$ $(k=1, 2, \ldots, n$, $M_0=A$, $M_n=B)$. Обозначим через $\Delta\varphi_k$ угол между лучами $OM_{k-1}$ и $OM_k$, образующими с полярной осью углы $\varphi_{k-1}$ и $\varphi_k$ соответственно. Каждый из элементарных секторов $OM_{k-1}M_k$ заменим круговым сектором с радиусом $\rho_k=\rho(\varphi_k')$, где $\varphi_k'$ — значение угла $\varphi$ из промежутка $[\varphi_{k-1}, \varphi_k]$, и центральным углом $\Delta\varphi_k$. Площадь последнего сектора выражается формулой $\Delta s_k=\frac{1}{2}\rho_k^2\Delta\varphi_k$.

Сумма

$$S_n=\sum_{k=1}^{n}\frac{1}{2}\rho_k^2\Delta\varphi_k=\frac{1}{2}\sum_{k=1}^{n}[\rho(\varphi_k')]^2\Delta\varphi_k$$

выражает площадь «ступенчатого» сектора, аппроксимирующего (приближенно заменяющего) данный сектор $OAB$.

Площадью сектора $OAB$ называется предел площади «ступенчатого» сектора при $n \to \infty$ и $\lambda = \max \Delta\varphi_k \to 0$:

$$S = \lim_{\lambda \to 0} \sum_{k=1}^{n} \frac{1}{2} [\rho(\varphi_k')]^2 \Delta\varphi_k = \lim_{\lambda \to 0} \sum_{k=1}^{n} \frac{1}{2} \rho_k^2 \Delta\varphi_k.$$

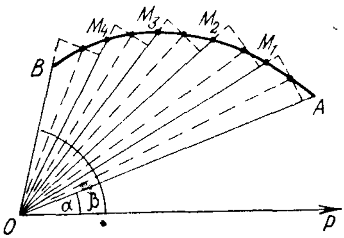



Рис. 17.4

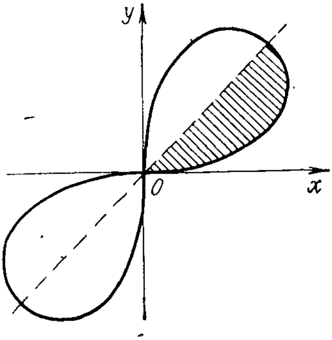



Рис. 17.5

Так как

$$\lim_{\lambda \to 0} \sum_{k=1}^{n} \frac{1}{2} \rho_k^2 \Delta\varphi_k = \frac{1}{2} \int_{\alpha}^{\beta} \rho^2(\varphi)\, d\varphi,$$

то

$$S = \frac{1}{2} \int_{\alpha}^{\beta} \rho^2(\varphi)\, d\varphi. \tag{17.7}$$

Пример. Вычислить площадь области, ограниченной лемнискатой $\rho^2 = a^2 \sin 2\varphi$.

Принимая во внимание симметрию линии относительно ее оси (рис. 17.5), по формуле (17.7) получаем

$$\frac{1}{4} S = \frac{1}{2} \int_0^{\frac{\pi}{4}} a^2 \sin 2\varphi d\varphi = \frac{1}{4} a^2 \int_0^{\frac{\pi}{4}} \sin 2\varphi d(2\varphi) =$$

$$= -\frac{a^2}{4} \cos 2\varphi \Big|_0^{\frac{\pi}{4}} = \frac{1}{4} a^2, \quad S = a^2.$$

## § 17.3. Длина дуги кривой

В § 13.1 дано определение длины дуги кривой, в § 13.6 получена формула для дифференциала длины дуги пространственной линии. Если эта линия задана параметрическими уравнениями

$$x=\varphi_1(t),\ y=\varphi_2(t),\ z=\varphi_3(t)\ (\alpha\leqslant t\leqslant\beta), \qquad (17.8)$$

где $\varphi_i(t)$ $(i=1, 2, 3)$ — дифференцируемые функции аргумента $t$, то

$$ds=\sqrt{x_t'^2+y_t'^2+z_t'^2}\,dt. \qquad (17.9)$$

Интегрируя равенство (17.9) по промежутку $[\alpha, \beta]$, получаем формулу для вычисления длины дуги линии (17.8)

$$s=\int_\alpha^\beta\sqrt{x_t'^2+y_t'^2+z_t'^2}\,dt. \qquad (17.10)$$

Если линия (17.8) лежит в плоскости $Oxy$, то $z=0$ при всех $t\in[\alpha, \beta]$, поэтому

$$s=\int_\alpha^\beta\sqrt{x_t'^2+y_t'^2}\,dt. \qquad (17.11)$$

В случае, когда плоская линия задана уравнением $y=f(x)$ $(a\leqslant x\leqslant b)$, где $f(x)$ — дифференцируемая функция, последняя формула принимает вид

$$s=\int_a^b\sqrt{1+y_x'^2}\,dx. \qquad (17.12)$$

Пусть плоская линия задана уравнением $\rho=\rho(\varphi)$ $(\alpha\leqslant\varphi\leqslant\beta)$ в полярных координатах. Принимая во внимание формулы (1.17), получаем параметрические уравнения рассматриваемой линии $x=\rho(\varphi)\cos\varphi$, $y=\rho(\varphi)\sin\varphi$, в которых роль параметра играет полярный угол $\varphi$. Поскольку $x_\varphi'=\rho'(\varphi)\cos\varphi-\rho(\varphi)\sin\varphi$, $y_\varphi'=\rho'(\varphi)\sin\varphi+\rho(\varphi)\cos\varphi$, $x_\varphi'^2+y_\varphi'^2=\rho^2+\rho'^2$, формула (17.11) принимает вид

$$s=\int_\alpha^\beta\sqrt{\rho^2+\rho'^2}\,d\varphi. \qquad (17.13)$$

Формула (17.13) выражает длину дуги линии $\rho=\rho(\varphi)$ $(\alpha\leqslant\varphi\leqslant\beta)$ в полярных координатах.

Пример 1. Вычислить длину дуги винтовой линии $x=a\cos t$, $y=a\sin t$, $z=bt$ между точками, для которых $t=0$, $t=\beta$.

Поскольку $x'=-a\sin t$, $y'=a\cos t$, $z'=b$, $\sqrt{x'^2+y'^2+z'^2}=\sqrt{a^2+b^2}$, по формуле (17.10) находим

$$s=\int_0^\beta \sqrt{a^2+b^2}\,dt=\sqrt{a^2+b^2}\,\beta,\quad s=\sqrt{a^2+b^2}\,\beta.$$

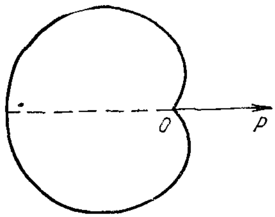



Рис. 17.6

Пример 2. Найти длину кардиоиды $\rho=2a(1-\cos\varphi)$ (рис. 17.6).

Так как $\rho'=2a\sin\varphi$, $\rho^2+\rho'^2=4a^2\sin^2\varphi+4a^2(1-\cos\varphi)^2=8a^2(1-\cos\varphi)=16a^2\sin^2\frac{\varphi}{2}$, то по формуле (17.13) получаем

$$s=2\int_0^\pi 4a\sin\frac{\varphi}{2}\,d\varphi=16a\int_0^\pi \sin\frac{\varphi}{2}\,d\left(\frac{\varphi}{2}\right)=-16a\cos\frac{\varphi}{2}\Big|_0^\pi=16a.$$

## § 17.4. Объем тела

Покажем, что объем тела можно вычислить с помощью определенного интеграла, если известна площадь любого поперечного сечения этого тела, перпендикулярного некоторому направлению.

Разобьем данное тело на элементарные слои плоскостями, перпендикулярными оси $Ox$ и определяемыми уравнениями $x=\mathrm{const}$ (рис. 17.7). Для любого фиксированного $x\in[a, b]$ известна площадь $S=S(x)$ поперечного сечения данного тела.

Элементарный слой, отсеченный плоскостями $x=x_{k-1}$, $x=x_k$ $(k=1, 2, \ldots, n,\ x_0=a,\ x_n=b)$, заменим цилиндром с высотой $\Delta x_k=x_k-x_{k-1}$ и площадью основания $S(\xi_k)$, $\xi_k\in[x_{k-1}, x_k]$.

Объем указанного элементарного цилиндра выражается формулой $\Delta v_k=S(\xi_k)\Delta x_k$.

Составим сумму всех таких произведений

$$V_n = \sum_{k=1}^{n} v_k = \sum_{k=1}^{n} S(\xi_k)\,\Delta x_k, \tag{17.14}$$

являющуюся интегральной суммой для данной функции $S = S(x)$ на отрезке $[a, b]$. Она выражает объем ступенчатого тела, состоящего из элементарных цилиндров и приближенно заменяющего данное тело.

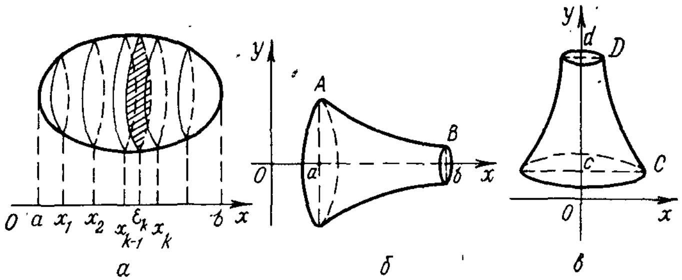



Рис. 17.7

*Объемом* данного тела называют предел объема указанного ступенчатого тела при $\lambda \to 0$, где $\lambda$ — длина наибольшего из элементарных отрезков $\Delta x_k$. Обозначим через $V$ объем данного тела, тогда по определению

$$V = \lim_{\lambda \to 0} \sum_{k=1}^{n} S(\xi_k)\,\Delta x_k. \tag{17.15}$$

С другой стороны, $\lim\limits_{\lambda \to 0} \sum\limits_{k=1}^{n} S(\xi_k)\,\Delta x_k = \int\limits_a^b S(x)\,dx.$

Из двух последних равенств следует искомая формула для вычисления объема тела по заданным поперечным сечениям

$$V = \int_a^b S(x)\,dx. \tag{17.16}$$

Если тело образовано вращением вокруг оси $Ox$ криволинейной трапеции $aABb$, ограниченной сверху дугой непрерывной линии $y = f(x)$, где $a \leqslant x \leqslant b$ (рис. 17.7, *б*), то $S(x)=\pi r^2=\pi y^2=\pi f^2(x)$ и формула (17.16) принимает вид

$$V_x=\pi\int_a^b y^2dx \quad \text{или} \quad V_x=\pi\int_a^b f^2(x)\,dx. \qquad (17.17)$$

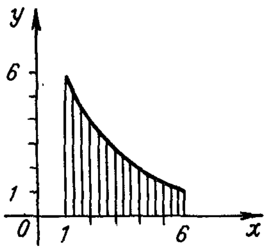


Рис. 17.8

Замечание. Объем тела, полученного вращением вокруг оси $Oy$ криволинейной трапеции $CcdD$ (рис. 17.7, *в*), где $CD$— дуга кривой $x=\varphi(y)$ $(c\leqslant y\leqslant d)$, определяется формулой

$$V_y=\pi\int_c^d x^2dy$$

или

$$V_y=\pi\int_c^d \varphi^2(y)\,dy. \qquad (17.18)$$

Пример 1. Вычислить объем тела, полученного вращением вокруг оси $Ox$ криволинейной трапеции, ограниченной линиями $xy=6$, $x=1$, $x=6$ (рис. 17.8).

*Из уравнения гиперболы* $xy=6$, *ограничивающей* сверху данную трапецию, находим выражение для $y$ и применяем формулу (17.17):

$$V_x=\pi\int_1^6 \frac{36}{x^2}\,dx=-36\pi\left.\frac{1}{x}\right|_1^6=-36\pi\left(\frac{1}{6}-1\right)=30\pi.$$

Пример 2. Вычислить объем тела, полученного вращением вокруг оси $Oy$ криволинейной трапеции, ограниченной линиями $y=x^2$, $x=0$, $y=2$.

Пользуемся формулой (17.18). В рассматриваемом случае $c=0$, $d=2$, $x^2=y$, поэтому

$$V_y=\pi\int_0^2 y\,dy=\pi\left.\frac{y^2}{2}\right|_0^2=2\pi.$$

## § 17.5. Площадь поверхности вращения

Рассмотрим поверхность, полученную вращением дуги линии $y=f(x)$ $(a\leqslant x\leqslant b)$ вокруг оси $Ox$ (предположим, что функция $y=f(x)$ имеет непрерывную производную). Фиксируем значение $x\in[a, b]$, аргументу функции придадим приращение $dx$, которому соответствует «элементарное кольцо» (рис. 17.9), полученное вращением элементарной дуги $\Delta l$. Это «кольцо» заменим цилиндрическим кольцом — боковой поверхностью тела, образованного вращением прямоугольника с основанием, равным дифференциалу дуги $dl$, и высотой $h=f(x)$. Разрезав последнее кольцо и развернув его, получим полоску шириной $dl$ и длиной $2\pi y$, где $y=f(x)$.

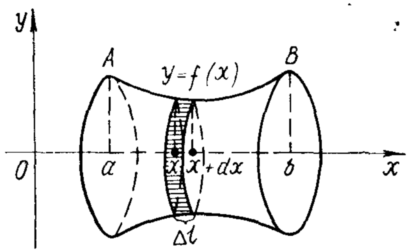



Рис. 17.9

Следовательно, дифференциал площади поверхности выразится формулой

$$dS = 2\pi y dl = 2\pi y \sqrt{1+y'^2}\,dx, \qquad (17.19)$$

откуда

$$S = 2\pi \int_a^b y \sqrt{1+y'^2}\,dx. \qquad (17.20)$$

Формула (17.20) выражает площадь поверхности, полученной вращением дуги линии $y=f(x)$ $(a\leqslant x\leqslant b)$ вокруг оси $Ox$.

Пример. Вычислить площадь поверхности, полученной вращением отрезка прямой $y=x$ $(0\leqslant x\leqslant h)$ вокруг оси $Ox$.

Поскольку $y'=1$, $a=0$, $b=h$, по формуле (17.20) получаем

$$S = 2\pi \int_0^h x\sqrt{1+1}\,dx = 2\sqrt{2}\pi \int_0^h x dx = \sqrt{2}\,\pi x^2 \Big|_0^h = \sqrt{2}\pi h^2.$$

## § 17.6. Работа переменной силы

Под действием переменной силы $F=F(x)$ материальная точка $M$ движется по прямой $Ox$; направление силы совпадает с направлением движения. Требуется найти работу, производимую силой $F=F(x)$, где $F(x)$ — функция, непрерывная на отрезке $[a, b]$, при перемещении точки $M$ из положения $x=a$ в положение $x=b$.

Отрезок $[a, b]$ разобьем на $n$ элементарных отрезков $[x_{k-1}, x_k]$ длиной $\Delta x_k = x_k - x_{k-1}$ $(k=1, 2, \ldots, n, x_0=a, x_n=b)$. В каждом элементарном отрезке $[x_{k-1}, x_k]$ произвольно выберем точку $\xi_k$ и образуем произведение $F(\xi_k)\Delta x_k$, выражающее приближенно работу силы на пути $\Delta x_k$ (здесь предполагается, что сила на элементарном пути постоянна и равна $F(\xi_k)$). Составим сумму всех таких произведений

$$A_n = \sum_{k=1}^{n} F(\xi_k)\,\Delta x_k, \tag{17.21}$$

которая выражает приближенно работу силы $F=F(x)$ на отрезке $[a, b]$.

Сумма (17.21) является интегральной суммой для непрерывной функции $F=F(x)$ на отрезке $[a, b]$. Предел этой интегральной суммы при $n \to \infty$ и $\lambda \to 0$, где $\lambda$ — длина наибольшего из отрезков $\Delta x_k$, существует и выражает работу переменной силы $F=F(x)$ на прямолинейном пути от $x=a$ до $x=b$:

$$A = \lim_{\lambda \to 0} \sum_{k=1}^{n} F(\xi_k)\,\Delta x_k = \int_a^b F(x)\,dx,$$

$$A = \int_a^b F(x)\,dx. \tag{17.22}$$

## § 17.7. Понятие об эллиптических интегралах

Интегралы вида

$$\int R\left(x, \sqrt{ax^3+bx^2+cx+d}\right)dx, \quad \int R\left(x, \sqrt{ax^4+bx^3+cx^2+dx+e}\right)dx, \tag{17.23}$$

где $R$ — рациональная функция своих аргументов, называются *эллиптическими интегралами*. В общем случае эти интегралы не выражаются через элементарные функции.

Эллиптические интегралы с помощью элементарных подстановок (с точностью до слагаемых, выражающихся в конечном виде) приводятся к следующим трем стандартным интегралам:

$$\left.\begin{array}{c} \displaystyle\int \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}}, \quad \int \frac{z^2dz}{\sqrt{(1-z^2)(1-k^2z^2)}}, \\ \displaystyle\int \frac{dz}{(1+hz^2)\sqrt{(1-z^2)(1-k^2z^2)}} \quad (0<k<1). \end{array}\right\} \tag{17.24}$$

Интегралы (17.24) называются эллиптическими интегралами соответственно первого, второго, третьего рода.

Подстановкой $z=\sin\varphi$ $\left(0\leqslant\varphi\leqslant\frac{\pi}{2}\right)$ эти интегралы приводятся к интегралам

$$\int\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}},\quad \int\sqrt{1-k^2\sin^2\varphi}\,d\varphi,$$

$$\int\frac{d\varphi}{(1+h\sin^2\varphi)\sqrt{1-k^2\sin^2\varphi}} \tag{17.25}$$

и их комбинациям.

Интегралы (17.25) называются эллиптическими интегралами первого, второго и третьего рода в форме Лежандра[1].

Особое значение и многочисленные приложения имеют первые два интеграла. Лежандр ввел в рассмотрение функции

$$F(k,\ \varphi)=\int_0^{\varphi}\frac{dt}{\sqrt{1-k^2\sin^2 t}}; \tag{17.26}$$

$$E(k,\ \varphi)=\int_0^{\varphi}\sqrt{1-k^2\sin^2 t}\,dt. \tag{17.27}$$

Для функций $F(k,\ \varphi)$, $E(k,\ \varphi)$ составлены таблицы их значений при различных $\varphi$ и $k$, построены графики. Изучены свойства этих функций; установлен ряд формул, относящихся к ним.

В случае, когда $\varphi=\frac{\pi}{2}$, интегралы (17.26) и (17.27) называются полными эллиптическими интегралами первого и второго рода:

$$F(k)=\int_0^{\frac{\pi}{2}}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}=F\left(k,\ \frac{\pi}{2}\right); \tag{17.28}$$

$$E(k)=\int_0^{\frac{\pi}{2}}\sqrt{1-k^2\sin^2\varphi}\,d\varphi=E\left(k,\ \frac{\pi}{2}\right). \tag{17.29}$$

Замечание 1. К эллиптическому интегралу второго рода в форме Лежандра приводит задача о вычислении длины дуги эллипса. Действительно, пусть эллипс задан параметрическими уравнениями $x=a\sin t$, $y=b\cos t$ $(0\leqslant t<2\pi)$. Предполагая, что $a>b$, по формуле (17.11) получаем

[1] *Адриан Мари Лежандр* (*Adrien Marie Legendre*, 1752—1833) — французский математик.

$$s=\int_0^{\varphi}\sqrt{a^2\cos^2 t+в^2\sin^2 t}\,dt=\int_0^{\varphi}\sqrt{a^2(1-\sin^2 t)+b^2\sin^2 t}\,dt=$$

$$=\int_0^{\varphi}\sqrt{a^2-(a^2-b^2)\sin^2 t}\,dt=a\int_0^{\varphi}\sqrt{1-\frac{a^2-b^2}{a^2}\sin^2 t}\,dt=$$

$$=a\int_0^{\varphi}\sqrt{1-\varepsilon^2\sin^2 t}\,dt,$$

где $\varepsilon$ — эксцентриситет эллипса, т. е.

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

Следовательно,

$$s=a\int_0^{\varphi}\sqrt{1-\varepsilon^2\sin^2 t}\,dt=aE(\varepsilon,\ \varphi),$$

т. е. длина дуги эллипса выражается через эллиптический интеграл второго рода; этот факт послужил основанием для названия «эллиптический».

Отметим, что длина всего эллипса выражается через полный эллиптический интеграл второго рода

$$s=4a\int_0^{\frac{\pi}{2}}\sqrt{1-\varepsilon^2\sin^2 t}\,dt=4aE(\varepsilon).$$

Замечание 2. Эллиптические интегралы $F(k,\ \varphi)$ и $E(k,\ \varphi)$ являются примерами функций, которые изучаются по их интегральным выражениям и широко применяются, хотя и не могут быть выражены через элементарные функции. К таким функциям относятся многие другие, например, функции:

1) $\operatorname{Si}(x)=\int_0^x\frac{\sin t}{t}\,dt$ *(интегральный синус)*;

2) $\operatorname{Ci}(x)=-\int_x^{\infty}\frac{\cos t}{t}\,dt$ *(интегральный косинус)*;

3) $\operatorname{li}(x)=\int_0^x\frac{dt}{\ln t}\ (x>0)$ *(интегральный логарифм)*;

4) $\operatorname{Ei}(x)=\int_{-\infty}^{x}\frac{e^t}{t}\,dt$ *(интегральная показательная функция)*;

5) $S(x)=\int\limits_0^x \sin t^2\,dt, \quad C(x)=\int\limits_0^x \cos t^2 dt$ *(интегралы Френеля)*;

6) $\Phi(x)=\frac{2}{\sqrt{\pi}}\int\limits_0^x e^{-t^2}\,dt$ *(интеграл вероятностей)*.

Все приведенные функции являются неэлементарными, они хорошо изучены, для них составлены таблицы значений; эти функции находят широкое применение.

# ЛИТЕРАТУРА


1. *Головина Л. И.* Линейная алгебра и некоторые ее приложения.— М., 1979.—408 с.
2. *Ефимов Н. В.* Краткий курс аналитической геометрии.— М., 1975.—228 с.
3. *Ильин В. А., Позняк Э. Г.* Аналитическая геометрия.— М., 1971.—232 с.
4. *Ильин В. А., Позняк Э. Г.* Линейная алгебра.— М., 1978.— 304 с.
5. *Ильин В. А., Позняк Э. Г.* Основы математического анализа.— М., 1971, ч. 1.— 572 с.
6. *Кудрявцев Л. Д.* Математический анализ.— М., 1973, т. 1.— 588 с.
7. *Курош А. Г.* Курс высшей алгебры.— М., 1975.— 432 с.
8. *Никольский С. М.* Курс математического анализа.— М., 1975, т. 1.—432 с.
9. *Смирнов В. И.* Курс высшей математики.— М., 1974, т. 1.— 478 с.
10. *Тышкевич Р. И., Феденко А. С.* Линейная алгебра и аналитическая геометрия.— Минск, 1976.—544 с.
11. *Батунер Л. М., Позин М. Е.* Математические методы в химической технике.— 6-е изд.— Л., 1971.—824 с.
12. *Гусак А. А.* Задачи и упражнения по высшей математике.— Минск, 1972, ч. 1.—328 с.; 1973, ч. 2.—384 с.
13. *Гусак А. А.* Элементы методов вычислений.— 2-е изд.— Минск, 1982.—168 с.
14. *Выгодский М. Я.* Справочник по высшей математике.— М. 1979.—871 с.
15. *Градштейн И. С., Рыжик И. М.* Таблицы интегралов, сумм, рядов и произведений.— М., 1971.—1100 с.
16. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров.— М., 1977.—720 с.
17. *Янке Е., Эмде Ф.* Специальные функции. Формулы, графики, таблицы.— М., 1977.—342 с.
18. История отечественной математики.— Киев, 1966, т. 1.—492 с.; 1967, т. 2.—616 с.; 1968, т. 3.—726 с.; 1970, т. 4, кн. 1.—824 с.; 1970, т. 4, кн. 2.—668 с.

# АЛФАВИТНЫЙ УКАЗАТЕЛЬ

# ОГЛАВЛЕНИЕ

## Раздел II

## МАТЕМАТИЧЕСКИЙ АНАЛИЗ

*Алексей Адамович Гусак*

**ВЫСШАЯ МАТЕМАТИКА**

**Том 1**

Редактор Л. Г. Лепило
Младший редактор М. Г. Хаютина
Художественный редактор С. В. Баленок
Технический редактор А. Я. Максимова
Корректор Л. В. Лебедева

ИБ № 668

Сдано в набор 14.07.82. Подписано в печать 09.11.83. Формат 84×108 1/32. Бумага газетная. Гарнитура литературная. Высокая печать. Усл. печ. л. 24,36. Усл. кр.-отт. 24,56. Уч.-изд. л. 23,19. Тираж 10 600 экз. Заказ 1051. Цена 90 к.

Издательство БГУ им. В. И. Ленина Минвуза БССР и Госкомиздата БССР. Минск, проспект Машерова, 11.

Ордена Трудового Красного Знамени типография издательства ЦК КП Белоруссии. Минск, Ленинский проспект, 79.