М. К. Потапов,
В. В. Александров,
П. И. Пасиченко

ЛЕКЦИИ ПО АЛГЕБРЕ И ЭЛЕМЕНТАРНЫМ ФУНКЦИЯМ

М. К. Потапов, В. В. Александров, П. И. Пасиченко

М. К. Потапов,
В. В. Александров,
П. И. Пасиченко

ЛЕКЦИИ ПО АЛГЕБРЕ И ЭЛЕМЕНТАРНЫМ ФУНКЦИЯМ

Издательство
Московского университета
1978

УДК 510

Печатается по постановлению
Редакционно-издательского совета
Московского университета

Рецензенты:
акад. АН УССР Б. В. Гнеденко,
проф. А. Б. Шидловский

Потапов М. К., Александров В. В., Пасиченко П. И.
Лекции по алгебре и элементарным функциям. М.: Изд-во. Моск. ун-та, 1978.
384 с., 139 ил.

Методическое пособие написано на основе лекций, которые читались авторами в течение ряда лет на подготовительном отделении Московского университета. Большое внимание уделено тем разделам школьной программы, которые особенно важны при изучении высшей математики. Так, действительные и комплексные числа и операции над ними описаны достаточно подробно. Изложение элементарных функций включает понятия предела и непрерывности. Материал изложен доходчивым языком, причем строгость изложения нарастает постепенно, что дает возможность читателю активно включиться в повторение забытых разделов элементарной математики Книга будет полезна слушателям подготовительных отделений вузов, а также всем тем, кто готовится к вступительным экзаменам в высшие учебные заведения.

П $\frac{20202-080}{077\ (02)-78}$ БЗ№29-16-77

© *Издательство Московского университета, 1978 г.*

Оглавление

Глава I

ДЕЙСТВИТЕЛЬНЫЕ ЧИСЛА

В разумной деятельности человека постоянно возникает необходимость не только качественной, но и количественной характеристики тех или иных явлений. Так, измерение на местности или в лабораторных условиях, любое научное исследование или производственный эксперимент приводят к тому, что приходится иметь дело с разными числами и производить действия над ними. Поэтому в I главе рассмотрим действительные числа и основные операции над ними.

§ 1
НАТУРАЛЬНЫЕ ЧИСЛА

Понятие натурального числа — одно из основных понятий математики. Как всякое основное понятие оно не определяется. При строгом введении натуральных чисел их свойства вытекают из системы аксиом. Поскольку строгое (аксиоматическое) введение натуральных чисел выходит за рамки этого курса лекций, то в этом параграфе свойства натуральных чисел будут лишь описаны.

Можно сказать, что каждое натуральное число характеризует количество элементов некоторого множества. Любые множества, имеющие одинаковое количество элементов, характеризуются одним и тем же натуральным числом. Всем множествам, состоящим из одного элемента, можно поставить в соответствие натуральное число, которое принято называть числом «единица». Всем множествам, состоящим из двух элементов, можно поставить в соответствие другое натуральное число, которое принято называть числом «два». Продолжая устанавливать такое соответствие, можно получить все натуральные числа, которые обладают следующими свойствами: а) их бесконечно много; б) все натуральные числа располагаются в порядке одно за другим, начиная от единицы.

Множество натуральных чисел образует *ряд натуральных чисел*: первое число — единица, второе — два, третье — три и т. д. При этом у каждого натурального числа есть свое место в этом ряду, т. е. можно говорить о месте натурального числа n в ряду натуральных чисел. Имея ряд натуральных чисел, можно определить, какое из двух натуральных чисел *больше*: больше то, которое стоит в ряду натуральных чисел дальше от начала; *меньше* то, которое стоит в ряду натуральных чисел ближе к началу.

Чтобы обозначить, что число m больше числа n, употребляется запись $m > n$. Для обозначения того, что число m меньше числа n, употребляется запись $m < n$. Называют эти записи неравенствами натуральных чисел. Чтобы обозначить, что число m и число n — одно и то же число, употребляют запись $m = n$ и называют ее равенством натуральных чисел.

Действия сложения и умножения натуральных чисел можно определить следующим образом.

Сложить два натуральных числа m и n — значит найти натуральное число p $(p > m)$, находящееся на n-м месте от числа m. Это число p называется *суммой* чисел m и n и обозначается $m + n$. Чтобы сложить несколько натуральных чисел, надо сложить сначала первые два, затем к полученной сумме прибавить следующее натуральное число и т. д.

Умножить натуральное число m на натуральное число n — значит найти натуральное число q, равное: а) m, если $n = 1$; б) сумме n чисел, каждое из которых есть m (т. е. взять число m суммой n раз). Это число q называется *произведением* чисел m и n и обозначается mn. Числа m и n называются *сомножителями*. Чтобы перемножить несколько натуральных чисел, надо сначала перемножить первые два, затем полученное натуральное число умножить на следующее натуральное число и т. д.

Приведем *основные законы* сложения и умножения натуральных чисел:

а) $m + n = n + m$ (коммутативность сложения);

б) $(l + m) + n = l + (m + n)$ (ассоциативность сложения);

в) $mn = nm$ (коммутативность умножения);

г) $(lm)\,n = l\,(mn)$ (ассоциативность умножения);

д) $(l + m)\,n = ln + mn$ (дистрибутивность сложения относительно умножения).

Если число m взято сомножителем k раз (k — натуральное число большее единицы), то произведение

$$\underbrace{mm\ldots m}_{k \text{ раз}}$$

называют *k-й степенью числа m* и обозначают m^k, т. е. по определению $m^k = \underbrace{mm\ldots m}_{k \text{ раз}}$. Кроме того, по определению $m^1 = m$.

Справедливы следующие свойства степеней:

а) $m^k m^n = m^{k+n}$;

б) $(m^k)^n = m^{kn}$;

в) $m^k l^k = (ml)^k$.

Действительно,

а) $m^k m^n = \underbrace{(mm\ldots m)}_{k \text{ раз}}\underbrace{(mm\ldots m)}_{n \text{ раз}} = \underbrace{(mm\ldots m)}_{(k+n) \text{ раз}} = m^{k+n}$;

б) $(m^k)^n = \underbrace{(m^k)(m^k)\ \ldots\ (m^k)}_{n \text{ раз}} = \underbrace{\underbrace{(mm\ \ldots\ m)}_{k \text{ раз}}\ \ldots\ \underbrace{(mm\ \ldots\ m)}_{k \text{ раз}}}_{n \text{ раз}} =$

$= \underbrace{(mm\ldots m)}_{kn \text{ раз}} = m^{kn}$;

в) $m^k l^k = \underbrace{(mm\ldots m)}_{k \text{ раз}}\underbrace{(ll\ \ldots\ l)}_{k \text{ раз}} = \underbrace{(ml)(ml)\ldots(ml)}_{k \text{ раз}} = (ml)^k$.

Определим действия, обратные сложению и умножению, действия вычитания и деления. *Вычесть* из натурального числа n натуральное число m — значит найти натуральное число p, такое, что

$$m+p=n. \tag{1}$$

Не для любых натуральных чисел n и m существует такое натуральное число p, что выполняется равенство (1). Если такое p существует, то оно называется *разностью* чисел n и m и обозначается $n-m$.

Разделить натуральное число n на натуральное число m — значит найти натуральное число q, такое, что

$$mq=n. \tag{2}$$

Не для любых натуральных чисел n и m существует такое натуральное число, что выполняется равенство (2). Если такое q существует, то числа m и q называются *делителями* числа n и обозначаются $q=n:m$, $m=n:q$. Справедливы следующие утверждения.

Если число m есть делитель чисел n_1 и n_2, то m есть делитель суммы n_1+n_2. Действительно, поскольку m есть делитель числа n_1, то $n_1=mq_1$, аналогично $n_2=mq_2$. Применяя закон дистрибутивности сложения относительно умножения натуральных чисел, имеем $n_1+n_2=mq_1+mq_2=m(q_1+q_2)$, т. е. число n_1+n_2 делится на число m.

Если число m есть делитель чисел n_1 и n_2 и $n_1>n_2$, то число m есть делитель разности n_1-n_2. Справедливость этого утверждения доказывается аналогично.

Отметим еще несколько очевидных свойств равенств натуральных чисел:

а) если $m=n$, то $m+k=n+k$ для любого натурального числа k;

б) если $m=n$, то $m-l=n-l$ для любого натурального числа l, такого, что $m>l$;

в) если $m=n$, то $mp=np$ для любого натурального числа p;

г) если $m=n$, то $m:q=n:q$ для любого натурального числа q, являющегося делителем числа m.

Рассмотрим *число нуль*, которое не является натуральным. Нуль вводится как число, характеризующее количество элементов в пустом множестве, т. е. в множестве, не содержащем ни одного элемента. Нуль считается числом, предшествующим всем натуральным числам и обозначается символом 0. Ряд натуральных чисел с числом нуль называется *расширенным натуральным рядом.* В расширенном натуральном ряду к правилам сложения и умножения натуральных чисел добавляются правила сложения и умножения, в которых участвует число нуль:

а) $0+n=n+0=n$;

б) $0+0=0$;

в) $0\cdot n=n\cdot 0=0$;

г) $0\cdot 0=0$.

По определению нулевая степень любого натурального числа m есть единица, т. е. $m^0=1$. Возведение нуля в нулевую степень и деление на нуль являются запрещенными действиями.

Чтобы производить действия над натуральными числами, надо уметь их записывать. Запись одного и того же натурального числа зависит от системы счисления.

В основе всякой системы счисления лежит следующий принцип: некоторое определенное число единиц составляет новую единицу следующего высшего разряда. Это число называется основанием системы счисления. Если за основание системы принято число два, то система счисления называется двоичной, если за основание принято число двенадцать — система называется двенадцатиричной и т. д. Пусть за основание системы принято некоторое число k. Тогда всякое число N в этой системе счисления может быть записано в виде суммы

$$N=a_nk^n+a_{n-1}k^{n-1}+\ldots+a_3k^3+a_2k^2+a_1k^1+a_0k^0,$$

где a_0, a_1, a_2, ..., a_{n-1} есть некоторые знаки для изображения чисел от нуля до числа $(k-1)$ включительно, a_n — некоторый знак для изображения числа от единицы до $(k-1)$ включительно, n — натуральное число, зависящее от числа N.

Рассмотрим десятичную систему счисления. В этой системе вводится десять знаков, называемых цифрами; для обозначения первых девяти натуральных чисел — знаки 1, 2, 3, 4, 5, 6, 7, 8, 9, а для числа нуль — знак 0. В этой системе счисления число десять обозначается символом 10, а каждое число N представляется в виде

$$N=a_n10^n+a_{n-1}10^{n-1}+\ldots+a_210^2+a_110+a_0, \qquad (3)$$

где a_0, a_1, a_2, ..., a_{n-1} вполне определенные числа из множества 0, 1, 2, ..., 9, a_n — из множества 1, 2, ..., 9. Для записи числа N обычно употребляется другая форма записи, основанная на принципе позиционного значения цифр. Суть этого принципа заключена в том, что каждая цифра кроме своего значения в зависимости от начертания получает еще и так называемое позиционное значение. Например, цифра 5 может иметь значения: пять единиц, если стоит в изображении (3) числа N на первом месте справа; пять десятков, если стоит в изображении (3) числа N на втором месте справа, и т. д. На этом принципе и основана обычная запись натуральных чисел. Запись 2705 означает, что число состоит из двух тысяч, семи сотен, нуля десятков и пяти единиц, т. е.

$$2705 = 2 \cdot 10^3 + 7 \cdot 10^2 + 0 \cdot 10 + 5.$$

Запись (3) числа N, основанная на позиционном принципе, будет такая:

$$N = \overline{a_n a_{n-1} \ldots a_2 a_1 a_0}$$

(черта сверху ставится для того, чтобы отличать это число от произведения $a_n a_{n-1} \ldots a_2 a_1 a_0$). В дальнейшем будут употребляться две формы записи натурального числа N:

а) $N = \overline{a_n a_{n-1} \ldots a_2 a_1 a_0}$;
б) $N = a_n 10^n + a_{n-1} 10^{n-1} + \ldots + a_2 10^2 + a_2 10^1 + a_0$.

Признаки делимости. Ранее уже говорилось, что не всегда одно натуральное число делится на другое. Ниже приводится несколько признаков делимости (отметим, что нуль делится на любое натуральное число).

Теорема 1. Чтобы натуральное число $N = \overline{a_n a_{n-1} \ldots a_2 a_1 a_0}$ делилось на 2, необходимо и достаточно, чтобы последняя цифра a_0 этого числа делилась на 2.

Докажем, что если число a_0 делится на 2, то число N также делится на 2. Запишем число N в виде

$$N = (a_n 10^n + a_{n-1} 10^{n-1} + \ldots + a_2 10^2 + a_1 10) + a_0. \qquad (4)$$

Каждое слагаемое в правой части равенства (4) делится на 2, следовательно, и вся сумма делится на 2, и число N делится на 2.

Докажем обратное. Если число N делится на 2, то число a_0 также делится на 2. По свойству б) равенств из равенства (4) вытекает

$$a_0 = N - (a_n 10^n + a_{n-1} 10^{n-1} + \ldots + a_2 10^2 + a_1 10).$$

Каждый член разности, стоящей в правой части равенства, делится на 2, следовательно, вся алгебраическая сумма делится на 2, и число a_0 делится на 2. Теорема доказана.

Рассмотрим характерные черты доказательства теоремы 1. В самой теореме сформулированы, а затем и доказаны два утверждения: а) из делимости числа a_0 на 2 следует делимость на 2 числа N (достаточное условие делимости числа N на 2); б) из делимости числа N на 2 следует делимость на 2 числа a_0 (необходимое условие делимости числа N на 2). Если свойство делимости числа a_0 на 2 обозначим буквой A, а свойство делимости числа N на 2 обозначим буквой B, то первое утверждение можно кратко сформулировать так: из A следует B $(A \Rightarrow B)$, а второе утверждение — из B следует A $(B \Rightarrow A)$. Теорему с помощью введенных символов можно записать так: $A \Leftrightarrow B$, т. е. свойство A, более простое и легко проверяемое, является необходимым и достаточным условием для выполнения более сложного свойства B.

Так как свойство A является достаточным условием для свойства B $(A \Rightarrow B)$, то на практике, убедившись в том, что последняя цифра является четной, можно быть уверенным, что и все число делится на два. Так как свойство A является необходимым условием для свойства B $(B \Rightarrow A)$, то, установив, что число a_0 не делится на 2, можно утверждать, что и число N не делится на два, т. е. теорему 1 можно сформулировать так: если последняя цифра числа N делится на 2, то число N делится на 2, если она не делится на 2, то число N не делится на 2.

Если доказано, что некоторое свойство C является достаточным условием для свойства D, то еще не следует, что нет чисел, не обладающих свойством C, но в то же время обладающих свойством D. Например, достаточным условием делимости числа N на 4 есть $a_1 = a_0 = 0$. Справедливость последнего утверждения следует из представления числа N в виде $N = (a_n 10^{n-2} + \ldots + a_2) 10^2$ и делимости числа 100 на 4. В то же время, например, число 52 также делится на 4.

Если доказано, что некоторое свойство E является необходимым условием для свойства D, то еще не следует, что нет чисел, не обладающих свойством D, но в то же время обладающих свойством E. Например, необходимым условием делимости числа N на 4 является четность числа N. Справедливость последнего утверждения очевидна, ибо если число N делится на 4, то тем более оно делится на два. В то же время, например, число 22 не делится на 4, хотя оно четно.

Если же доказано, что свойство P является необходимым и достаточным условием для свойства Q, а свойство P легко проверяется для любого числа N, то, найдя все числа, обладающие свойством P, можно сказать, что найдены все числа, обладающие более сложным свойством Q.

Сформулируем и докажем признаки делимости натуральных чисел на 4 и на 9.

Теорема 2. Чтобы натуральное число $N=\overline{a_n a_{n-1} \ldots a_2 a_1 a_0}$ делилось на 4, необходимо и достаточно, чтобы число $\overline{a_1 a_0}$ делилось на 4.

Достаточность. Пусть число $\overline{a_1 a_0}$ делится на 4. Запишем число N в виде

$$N=(a_n 10^n+a_{n-1}10^{n-1}+\ldots+a_2 10^2)+\overline{a_1 a_0}. \qquad (5)$$

Каждое слагаемое в правой части равенства (5) делится на 4, следовательно, и вся сумма делится на 4, т. е. число N делится на 4. Необходимость. Пусть число N делится на 4. По свойству б) равенств из равенства (5) вытекает

$$\overline{a_1 a_0}=N-(a_n 10^n+a_{n-1}10^{n-1}+\ldots+a_2 10^2).$$

Каждый член разности, стоящий в правой части равенства, делится на 4, следовательно, вся алгебраическая сумма делится на 4, и число $\overline{a_1 a_0}$ делится на 4. Теорема доказана.

Теорема 3. Чтобы натуральное число $N=\overline{a_n a_{n-1} \ldots a_2 a_1 a_0}$ делилось на 9, необходимо и достаточно, чтобы сумма всех цифр данного числа делилась на 9.

Достаточность. Пусть сумма цифр данного числа делится на 9. Запишем число N в виде $N=a_n 10^n+a_{n-1}10^{n-1}+$ $+\ldots+a_2 10^2+a_1 10+a_0$. Легко видеть, что справедливо равенство

$$10^k=\underbrace{\overline{99\ldots 9}}_{k \text{ раз}}+1.$$

Пользуясь этим равенством, перепишем N в виде

$$N=(a_n\underbrace{\overline{99\ldots 9}}_{n \text{ раз}}+a_{n-1}\underbrace{\overline{99\ldots 9}}_{(n-1) \text{ раз}}+\ldots+a_1 9)+(a_0+a_1+\ldots+a_n). \qquad (6)$$

Каждое слагаемое в правой части равенства (6) делится на 9, следовательно, и вся сумма делится на 9, и число N делится на 9. Необходимость. Пусть число N делится на 9. По свойству б) равенств из равенства (6) вытекает

$$a_0+a_1+\ldots+a_n=N-(a_n\underbrace{\overline{9\ldots 9}}_{n \text{ раз}}+a_{n-1}\underbrace{\overline{9\ldots 9}}_{(n-1) \text{ раз}}+\ldots+a_2\overline{99}+a_1 9).$$

Каждый член разности, стоящей в правой части равенства, делится на 9, следовательно, вся алгебраическая сумма делится на 9 и число $(a_0+a_1+\ldots+a_n)$ делится на 9. Теорема доказана.

Простые и составные числа. Множество натуральных чисел состоит из единицы, простых и составных чисел. Натуральное число, большее единицы, называется *простым*, если оно не имеет делителей, кроме единицы и самого себя. Натуральное число, большее единицы, называется *составным*, если оно имеет хотя бы один делитель, отличный от единицы и самого себя.

Всякое составное число N можно разложить на простые множители, записав это число N в виде произведения простых сомножителей.

Пример. $221 = 13 \cdot 17$.

При разложении числа на простые множители некоторые из них могут встретиться в разложении не один раз. Принято писать этот простой множитель в степени, показывающей, сколько раз он является сомножителем.

Пример. $360 = 2^3 \cdot 3^2 \cdot 5^1$.

Любое натуральное число N можно записать в виде

$$N = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}, \tag{7}$$

где $p_1, \ldots, p_k$ — различные простые делители числа N, а $\alpha_1, \ldots, \alpha_k$ — число их повторений в разложении числа N. Разложение (7) натурального числа N на простые множители единственно, т. е. не существует других простых чисел, являющихся делителями числа N, и степени $\alpha_1, \ldots, \alpha_k$ не могут заменены другими степенями.

Если натуральные числа N_1 и N_2 делятся на одно и то же натуральное число n, то число n называется *общим делителем* чисел N_1 и N_2. Наибольшее натуральное число, на которое делятся N_1 и N_2, называется *наибольшим общим делителем* этих чисел (НОД). Если НОД двух чисел равен 1, то они называются *взаимно-простыми*.

Теорема 4. Если натуральные числа N_1 и N_2 взаимно-просты, а натуральное число N делится и на N_1 и на N_2, то N делится на произведение $N_1 N_2$.

Доказательство теоремы опустим.

Заметим, что если числа N_1 и N_2 не взаимно-просты, то утверждение теоремы иногда неверно. Например, натуральное число 180 делится на 4 и на 6, но не делится на их произведение — на 24.

Теорема 5. Простых чисел бесконечно много.

Доказательство этой теоремы проведем *методом от противного*. Допустим, что существует лишь конечное число простых чисел $p_1, p_2, \ldots, p_n$. Тогда каждое натуральное число, большее 1 и не совпадающее ни с одним из этих чисел, будет составным. Число $N = p_1 \cdot p_2 \cdot \ldots \cdot p_n + 1$ не совпадает ни с одним из чисел $p_1, p_2, \ldots, p_n$, так как оно больше каждого из них. По нашему предположению, простых чисел, кроме $p_1, p_2, \ldots, p_n$, нет. Следовательно, число N составное и поэтому делится на одно из чисел $p_1, p_2, \ldots, p_n$.

С другой стороны, число N не делится ни на одно из чисел $p_1, p_2, \ldots, p_n$, поскольку произведение $p_1 \cdot p_2 \cdot \ldots \cdot p_n$ делится на каждое из этих чисел, а число 1 ни на одно из них не делится.

Таким образом, предположив, что существует лишь конечное число простых чисел, приходим к противоречию. Следовательно, множество простых чисел бесконечно.

Метод доказательства от противного заключается в следующем: строится отрицание утверждения, которое следует доказать. На основании полученного таким образом утверждения путем логических рассуждений приходим к другому утверждению, которое либо заведомо неверно, либо противоречит сделанному отрицанию, а это значит, что отрицание неверно. Тем самым из двух логически возможных ситуаций: либо верно данное утверждение, либо его отрицание — остается только одна — верно данное утверждение.

§ 2
ДРОБИ

Выше отмечалось, что операция деления не всегда выполнима в множестве натуральных чисел. Например, в множестве натуральных чисел нельзя 5 разделить на 4. Чтобы всегда была выполнима операция деления, приходится рассматривать новые числа — части натуральных чисел или дроби.

Обыкновенные дроби. Число, равное k-й части числа единица (k — натуральное число, большее единицы), обозначается $1/k$. Если эта часть берется m раз (m — натуральное число), то получаемое в результате этого новое число обозначают m/k. Число, определяемое по этому правилу при помощи двух натуральных чисел p и q ($q>1$) и записываемое как p/q, называют *дробью*, или *частным*, натуральных чисел p и q, при этом p называют *числителем* этой дроби, а число q — *знаменателем*. Всякое натуральное число можно считать дробью со знаменателем единица, т. е. любое натуральное число n можно записать как дробь $n/1$. Поэтому дальше ограничение $q>1$ на знаменатель дроби снимается и говорят, что частное двух любых натуральных чисел p и q есть дробь p/q, и при этом множество всех дробей содержит в себе множество всех натуральных чисел.

Две дроби p/q и m/k считаются *равными*, если произведение числителя первой дроби на знаменатель второй равно произведению числителя второй на знаменатель первой, т. е. $p/q=m/k$, если $pk=mq$. Аналогично $p/q>m/k$, если $pk>mq$; $p/q<m/k$, если $pk<mq$.

Суммой двух дробей называется дробь, числитель которой равен сумме произведений числителя первой дроби на знаменатель второй и числителя второй дроби на знаменатель первой, а знаменатель равен произведению знаменателей этих дробей, т. е.

$$\frac{p}{q}+\frac{m}{k}=\frac{pk+qm}{qk}.$$

Произведением двух дробей называется дробь, числитель которой равен произведению числителей этих дробей, а знаменатель — произведению знаменателей, т. е. $\frac{p}{q}\frac{m}{k}=\frac{pm}{qk}$. Справедливы следующие *основные законы* сложения и умножения дробей:

а) $\frac{p}{q}+\frac{m}{k}=\frac{m}{k}+\frac{p}{q}$ (коммутативность сложения);

б) $\left(\frac{p}{q}+\frac{m}{k}\right)+\frac{l}{n}=\frac{p}{q}+\left(\frac{m}{k}+\frac{l}{n}\right)$ (ассоциативность сложения);

в) $\frac{p}{q}\frac{m}{k}=\frac{m}{k}\frac{p}{q}$ (коммутативность умножения);

г) $\left(\frac{p}{q}\frac{m}{k}\right)\frac{l}{n}=\frac{p}{q}\left(\frac{m}{k}\frac{l}{n}\right)$ (ассоциативность умножения);

д) $\left(\frac{p}{q}+\frac{m}{k}\right)\frac{l}{n}=\frac{p}{q}\frac{l}{n}+\frac{m}{k}\frac{l}{n}$ (дистрибутивность сложения относительно умножения).

Для действий сложения и умножения дробей есть обратные действия — вычитание и деление, которые определяются так же, как обратные действия для сложения и умножения натуральных чисел. Действие деление всегда выполнимо для дробей, а действие вычитание — не всегда.

Из определения равенства двух дробей вытекает *основное свойство дробей*: если числитель и знаменатель данной дроби умножить или разделить на одно и то же натуральное число k, то получится дробь, равная данной:

$$\frac{p}{q}=\frac{pk}{qk}.$$

Дробь p/q называется *несократимой*, если числа p и q взаимно-просты.

Теорема 1. Если p/q несократимая дробь, то дробь m/n равна ей тогда и только тогда, когда $m=pk$ и $n=qk$, где k — некоторое натуральное число.

Доказательство. Пусть $m=pk$ и $n=qk$. Тогда дроби p/q и m/n равны по основному свойству дробей.

Пусть $p/q=m/n$. По определению равенства дробей $pn=mq$. Левая часть этого равенства делится на число p, следовательно, правая часть делится на p. Так как числа p и q взаимно-просты, а произведение mq делится на p, то на p делится m (т. е. существует натуральное число k, такое, что $m=pk$). Подставляя значение m в равенство $pn=mq$, получаем $pn=pkq$, откуда $n=qk$. Теорема доказана.

Десятичные дроби. Дробь p/q, у которой знаменатель есть $\underbrace{100\ldots0}_{k \text{ раз}}=10^k$, называется *конечной десятичной дробью*.

Например, дроби 3721/100 и 21/1000 — конечные десятичные дроби. Для конечной десятичной дроби принята специальная форма записи, а именно пишут числитель дроби и, отсчитав с правой стороны столько цифр, сколько нулей в знаменателе, отделяют их запятой. Если в числителе меньше цифр, чем нулей в знаменателе, то пишут числитель, и перед цифрами числителя дописывают столько нулей, сколько недостает цифр в числителе, затем ставят запятую и перед ней еще один нуль. Например, вышеприведенные десятичные дроби могут быть записаны так: 37,21 и 0,021. Десятичные дроби можно складывать так же, как складывают натуральные числа, только для сложения надо записать эти дроби так, чтобы у них было одинаковое число цифр после запятой (т. е. у той дроби, у которой их меньше, надо дописать столько нулей, чтобы число цифр после запятой стало таким же, как у другой дроби) и после сложения у полученного числа отсчитать справа столько цифр, сколько их было справа от запятой у любого из слагаемых, затем отсчитанные цифры отделить запятой. Десятичные дроби умножают как натуральные числа, только в полученном числе отсчитывают справа и затем отделяют запятой столько цифр, сколько их стояло после запятой в обоих множителях вместе.

Любая конечная десятичная дробь легко переводится в обыкновенную. Для этого надо записать в числитель целое число, которое получается, если отбросить запятую у десятичной дроби, а в знаменатель написать число 10 в такой степени, сколько цифр стоит у десятичной дроби после запятой, после чего дробь можно сократить на общий множитель, если он есть.

Записать обыкновенную дробь в виде конечной десятичной — значит найти конечную десятичную дробь, равную данной. Естественно поставить вопрос: любую ли обыкновенную дробь можно записать в виде конечной десятичной дроби? Оказывается, что здесь дело обстоит намного сложнее, чем с переводом конечной десятичной дроби в обыкновенную.

Теорема 2. Всякая дробь p/q, где натуральное число q не имеет простых делителей, отличных от 2 и 5, может быть записана в виде конечной десятичной дроби.

Доказательство. Пусть дана дробь p/q, где $q=2^m5^n$. По основному свойству дроби любая обыкновенная дробь не изменится, если числитель и знаменатель ее умножить на одно и то же число. Умножая числитель и знаменатель дроби p/q на 2^n5^m, получим

$$\frac{p}{q}=\frac{p}{2^m5^n}=\frac{2^n5^mp}{2^m5^n2^n5^m}=\frac{2^n5^mp}{2^{m+n}5^{m+n}}=\frac{2^n5^mp}{10^{m+n}}.$$

Так как произведение 2^n5^mp — натуральное число, то, обозначая его через N, запишем дробь в виде $p/q=N/10^{m+n}$, откуда видно, что дробь p/q записана конечной десятичной дробью. Теорема доказана.

Теорема 3. Если данная несократимая дробь p/q может быть записана конечной десятичной дробью, то ее знаменатель не содержит простых множителей, кроме 2 и 5.

Доказательство. Если $q=1$, то теорема очевидна. Рассмотрим случай, когда $q \neq 1$. Дробь p/q по условию представлена в виде конечной десятичной дроби, т. е. справедливо равенство $p/q=A/10^k$, где A и k — натуральные числа. Так как p/q несократимая дробь, то из теоремы вытекает, что $A=pm$ и $10^k=qm$. Число 10^k имеет простые множители только 2 и 5. Значит, и число qm не имеет других простых множителей, кроме 2 и 5, что вытекает из единственности разложения числа на простые множители. Следовательно, число q не содержит других простых множителей, кроме 2 и 5. Теорема доказана.

Теорема 4. Чтобы несократимая дробь p/q могла быть записана конечной десятичной дробью, необходимо и достаточно, чтобы ее знаменатель не содержал никаких других простых множителей, кроме 2 и 5.

Справедливость теоремы 4 вытекает из теорем 2 и 3.

Теперь рассмотрим дроби p/q, где p и q взаимно-простые числа, и q содержит простые множители, отличные от 2 и 5. Как вытекает из теоремы 3, эта дробь не может быть записана конечной десятичной дробью.

Бесконечной десятичной дробью называют десятичную дробь, у которой для любого натурального числа n на n-м месте после запятой стоит одна из цифр: 0, 1, 2, 3, ..., 9. Бесконечная дробь называется *периодической*, если одна цифра или упорядоченная совокупность цифр, начиная с некоторого места после запятой, повторяется. Такая повторяющаяся цифра или упорядоченная совокупность цифр называется *периодом бесконечной дроби*.

Всякая дробь p/q, где p и q взаимно-просты и q содержит хотя бы один простой множитель, отличный от 2 и 5, может быть единственным образом записана в виде бесконечной периодической дроби. Иногда вместо того, чтобы писать период несколько раз, его пишут один раз, заключая в скобки: $5/11 = 0{,}454545\ldots = 0,(45)$; $270/7 = 38{,}571428571428\ldots = 38,(571428)$.

Итак, любую обыкновенную дробь можно записать в виде либо конечной десятичной дроби, либо бесконечной периодической десятичной дроби. Здесь без доказательства приводится правило перевода бесконечной периодической дроби в обыкновенную (это правило доказано в гл. VIII). Чтобы обратить периодическую дробь в обыкновенную, надо из числа, стоящего до второго периода, вычесть число, стоящее до первого периода и сделать эту разность числителем, а в знаменателе написать цифру 9 столько раз, сколько цифр в периоде, и после девяток дописать столько нулей, сколько цифр между запятой и первым периодом.

Пример. $0,(45)=\frac{45-0}{99}=\frac{5}{11}$; $3,1(73)=\frac{3173-31}{990}=\frac{3142}{990}=$ $=\frac{1571}{495}$.

Пользуясь этим правилом, можно показать, что любая конечная десятичная дробь представима в виде бесконечной периодической десятичной дроби, причем двумя способами.

Пример. $0,172=0,172(0)=\frac{1720-172}{9000}=\frac{1548}{9000}=\frac{172\cdot 9}{9\cdot 1000}=0,172$;

$$0,172=0,171(9)=\frac{1719-171}{9000}=\frac{1548}{9000}=\frac{172\cdot 9}{9\cdot 1000}=0,172.$$

Чтобы не было двух разных представлений одной и той же конечной десятичной дроби, принято не иметь число 9 в периоде. Тогда каждая десятичная конечная дробь может быть единственным образом записана в виде бесконечной периодической десятичной дроби с периодом 0 и, наоборот, каждая такая дробь есть конечная десятичная дробь. Итак, каждая обыкновенная дробь p/q может быть единственным образом представлена в виде бесконечной десятичной периодической дроби и, наоборот, каждая бесконечная десятичная периодическая дробь может быть единственным образом представлена в виде обыкновенной дроби p/q.

§ 3
ЦЕЛЫЕ, РАЦИОНАЛЬНЫЕ И ИРРАЦИОНАЛЬНЫЕ ЧИСЛА

Выше отмечалось, что операция вычитания не всегда выполнима в множестве натуральных чисел. Например, в множестве натуральных чисел нельзя вычесть из числа 3 число 5. Аналогично обстоит дело и с дробями. Поэтому возникает необходимость в рассмотрении новых чисел, называемых отрицательными.

Целое отрицательное число $(-n)$ вводится как число, противоположное натуральному числу n.

Множество, состоящее из всех натуральных чисел, всех целых отрицательных чисел и нуля, называется *множеством целых чисел*, а сами эти числа называются *целыми*. Часто натуральное число называют положительным целым числом.

Определим теперь операции сложения и умножения для целых чисел. Так как правила сложения и умножения чисел расширенного натурального ряда приведены выше, то рассмотрим правила сложения и умножения целых чисел, когда одно из них или оба отрицательные числа:

а) $(-m)+(-n)=-(m+n)$;
б) $(-m)+0=0+(-m)=-m$;
в) $(-m)+n=\begin{cases} -(m-n), & \text{если } m>n; \\ n-m, & \text{если } m<n; \\ 0, & \text{если } m=n; \end{cases}$
г) $(-m)n=-(nm)$;
д) $(-m)(-n)=mn$;
е) $(-m)\cdot 0=0\cdot(-m)=0$.

Основные законы сложения и умножения целых чисел аналогичны основным законам сложения и умножения натуральных чисел.

Для действий сложения и умножения целых чисел вводятся обратные действия вычитания и деления (кроме деления на нуль). При этом действие вычитания теперь всегда выполнимо.

Отрицательная дробь $(-p/q)$ также вводится как число, противоположное дроби p/q.
Множество, состоящее из всех дробей (их теперь будем называть положительными дробями), всех отрицательных дробей и нуля, называется *множеством рациональных чисел*, а сами эти числа называются *рациональными*.

Поскольку каждое натуральное число можно считать дробью со знаменателем единица, то целые отрицательные числа можно считать отрицательными дробями со знаменателем единица. Поэтому множество целых чисел — часть множества рациональных чисел. Иногда говорят, что рациональное число r есть число, которое можно представить как отношение целого числа к натуральному числу, т. е. $r=p/q$, где q — натуральное, а p — целое число.

Два рациональных числа p/q и m/n *равны*, если справедливо равенство $pn=qm$. Сложение и умножение рациональных чисел производится по следующим правилам:

а) $\frac{p}{q}+\frac{m}{n}=\frac{pn+mq}{qn}$;

б) $\frac{p}{q}\frac{m}{n}=\frac{pm}{qn}$.

Для действий сложения и умножения рациональных чисел вводятся обратные действия вычитания и деления, при этом оба эти действия, за исключением запрещенного деления на нуль, всегда выполнимы.

Основные законы сложения и умножения рациональных чисел аналогичны основным законам сложения и умножения целых чисел. Если рациональное число r взято сомножителем k раз ($k>1$), то произведение $\underbrace{rr\ldots r}_{k \text{ раз}}$ называют *k-й степенью числа* r и обозначают r^k. По определению $r^1=r$.

Как и для натуральных чисел, справедливы следующие свойства степеней рациональных чисел:

а) $r^m r^k = r^{m+k}$;

б) $r_1^m r_2^m = (r_1 r_2)^m$;

в) $(r^k)^m = r^{km}$;

г) $\left(\frac{r_1}{r_2}\right)^k = \frac{r_1^k}{r_2^k}$, если $r_2 \neq 0$;

д) $\frac{r^k}{r^m} = r^{k-m}$, если $k>m$, $r\neq 0$.

По определению $r^0=1$ для любого рационального числа r, кроме числа нуль.

В связи с понятием степени рациональных чисел часто возникает такая задача: для данного натурального числа k и для данного положительного рационального числа r_1 найти другое положительное рациональное число r_2, такое, что $r_2^k=r_1$. Эта задача не всегда имеет решение.

Теорема 1. Не существует рационального числа, квадрат которого равен 3.

Доказательство. Предположим, что существует рациональное число p/q, такое, что $(p/q)^2=3$. Не ограничивая общности, будем считать p и q взаимно-простыми (если числитель и знаменатель данного рационального числа имеют общие делители, то число p/q, полученное после сокращения, равно данному). Пользуясь определением произведения рациональных чисел, запишем наше предположение в виде $\frac{p^2}{q^2}=\frac{3}{1}$.

Из определения равенства рациональных чисел вытекает, что $p^2=3q^2$. Поскольку правая часть этого равенства делится на 3, то и левая должна делиться на 3. Но число p^2 делится на 3 только в случае, если число p делится на 3 (если p не делится на 3, то p^2 также не делится на 3). Раз p делится на 3, то существует целое число k, такое, что $p=3k$. Подставляя это значение p в равенство $p^2=3q^2$, получаем $q^2=3k^2$. Опять правая часть делится на 3, значит, и левая делится на 3, значит, число q делится на 3, т. е. $q=3m$. Итак, получили, что числа p и q имеют общий делитель, а по предположению в равенстве $(p/q)^2=3$ числа p и q взаимно-просты. Это противоречие и означает, что сделанное предположение неверно, а верно утверждение теоремы.

Таким образом, возникает необходимость ввести новые числа, отличные от рациональных, такие, например, как число, квадрат которого равен 3. К необходимости введения новых чисел можно подойти и следующим образом.

В § 2 каждая положительная дробь записывалась единственным образом в виде бесконечной периодической десятичной дроби. Отрицательные дроби также могут быть записаны единственным образом в виде отрицательной бесконечной периодической дроби. Число нуль также можно считать бесконечной периодической дробью с периодом нуль: $0 = 0,(0)$. Поэтому можно сказать, что *рациональным* называется число, представимое в виде бесконечной десятичной периодической дроби.

Теорема 2. Число 0,101001000100001..., образуемое, по правилу: за каждой единицей идет группа нулей, содержащая на один нуль больше, чем предыдущая группа, не является бесконечной периодической десятичной дробью.

Доказательство. Предположим, что это периодическая дробь. Пусть ее период состоит из n цифр. Возьмем в этой дроби разряды, в которых подряд стоят $2n+1$ нулей и рассмотрим нуль, стоящий посредине этой группы. Этот нуль находится либо в начале, либо в конце, либо внутри некоторого периода длины n, но во всех перечисленных случаях этот период целиком лежит на взятом отрезке из $(2n+1)$ нулей. Значит, период состоит из одних нулей, а этого быть не может. Теорема доказана.

Итак, существуют бесконечные непериодические десятичные дроби. Всякая бесконечная непериодическая десятичная дробь называется *иррациональным* числом. Множество бесконечных непериодических десятичных дробей образует *множество иррациональных чисел*.

§ 4
ДЕЙСТВИТЕЛЬНЫЕ ЧИСЛА

Действительным (или *вещественным*) числом называется любая бесконечная десятичная дробь. Множество всех бесконечных десятичных дробей называется *множеством действительных* (или *вещественных*) *чисел*. Поскольку бесконечные десятичные дроби есть периодические и непериодические, то множество всех действительных чисел есть множество всех рациональных чисел и множество всех иррациональных чисел. Два действительных числа *равны*, если они выражаются одной и той же десятичной дробью. Из двух положительных действительных чисел *больше* то, у которого либо больше целая часть, либо при равных целых частях и равных n первых цифрах после запятой $(n+1)$-я цифра больше. Из двух отрицательных чисел больше то, у которого противоположное число (положительное число) меньше. Положительное число больше нуля и любого отрицательного числа. Нуль больше любого отрицательного числа.

Рассмотрим приближенные значения бесконечных дробей. Оборвем на каком-то месте бесконечную положительную дробь, изображающую данное положительное действительное число. Получим конечную десятичную дробь (которую можно записать в виде бесконечной с периодом нуль). Эта дробь будет меньше (или равна) данного числа. Такая дробь называется *приближенным значением данного положительного действительного числа с недостатком*.

Если положительную бесконечную дробь оборвать на каком-то месте и к последней цифре прибавить единицу, то получим конечную десятичную дробь, которая больше данного действительного числа. Такая дробь называется *приближенным значением данного положительного действительного числа с избытком*.

Если отрицательную бесконечную дробь оборвать на каком-то месте, то получим конечную десятичную дробь, которая больше (или равна) данного действительного числа. Такая дробь называется *приближенным значением данного отрицательного действительного числа с избытком*.

Если отрицательную бесконечную дробь оборвать на каком-то месте и к последней цифре добавить единицу, то получим конечную десятичную дробь, которая меньше данного действительного числа. Такая дробь называется *приближенным значением данного отрицательного числа с недостатком*.

Примеры. Приближенным значением числа 0,4 (31) с недостатком будут следующие конечные дроби: 0,4; 0,43; 0,431; 0,4313; 0,43131 ...; приближенным значением этого же числа с избытком будут дроби 0,5; 0,44; 0,432; 0,4314; 0,43132; ....

Приближенным значением числа —3,2 (17) с недостатком будут следующие конечные дроби: —3,3; —3,22; —3,218; —3,2172; ...; приближенным значением того же числа с избытком будут дроби —3,2; —3,21; —3,217; —3,2171; ....

Определим действия сложения и умножения для действительных чисел. *Суммой* двух действительных чисел называется число, которое больше (или равно) суммы двух любых приближенных их значений с недостатком, но меньше (или равно) суммы двух любых приближенных их значений с избытком. Без доказательства примем, что такое число всегда существует и притом только одно.

Произведением двух действительных чисел называется число, которое больше (или равно) произведения двух любых приближенных значений с недостатком, но меньше (или равно) произведения двух любых приближенных их значений с избытком. Без доказательства примем, что такое число всегда существует и притом только одно.

Справедливы следующие *основные законы* сложения и умножения действительных чисел:

а) $a+b=b+a$ (коммутативность сложения);

б) $(a+b)+c=a+(b+c)$ (ассоциативность сложения);

в) $ab=ba$, (коммутативность умножения);

г) $(ab)c=a(bc)$ (ассоциативность умножения);

д) $(a+b)c=ac+bc$ (дистрибутивность сложения относительно умножения).

На множестве действительных чисел всегда выполнимы вычитание и деление (за исключением запрещенного деления на нуль).

Если действительное число a взято сомножителем k раз ($k>1$), то произведение

$$\underbrace{aa\ldots a}_{k \text{ раз}}$$

называют *k-степенью числа a* и обозначают a^k. По определению $a^1=a$. Степень числа a обладает свойствами, аналогичными свойствам степени рациональных чисел.

Положительное число b, такое, что его n-я степень есть данное положительное число a, т. е. такое положительное число b, что $b^n=a$, называется *арифметическим корнем* степени n из положительного числа a и обозначается $b=\sqrt[n]{a}$.

Теорема. Для любого натурального числа n и любого положительного числа a существует и притом единственный в множестве действительных чисел арифметический корень степени n из числа a.

Доказательство теоремы опустим.

В случае $n=2$ в обозначении корня цифру 2 не пишут; в случае $n=1$ корень 1-й степени из числа a есть само число a. Арифметические корни могут быть рациональными и иррациональными числами.

Пример. $\sqrt[3]{\frac{8}{27}}$ есть рациональное число $\frac{2}{3}$; $\sqrt{3}$ есть иррациональное число.

Заметим, что корень любой натуральной степени из числа нуль есть нуль.

Абсолютной величиной $|a|$ действительного числа a называется: само это число, если a — положительное число; нуль, если a — нуль; противоположное число $-a$, если a — отрицательное число, т. е.

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

Основные свойства абсолютной величины действительного числа будут приведены позже (гл. II).

Заметим, что в силу определения арифметического корня из неотрицательного числа для любого действительного числа справедливо равенство $\sqrt{a^2}=|a|$.

Перейдем теперь к *геометрической интерпретации* действительных чисел. Пусть дана горизонтальная прямая. Она имеет два взаимно противоположных направления. Одно из этих направлений принимается за положительное, а противоположное — за отрицательное. Обычно за положительное берется направление вправо (рис. 1). Прямую, на которой выбрано положительное направление, называют *осью*. Возьмем на оси произвольную фиксированную точку O (начало отсчета). Точка O разбивает прямую на две части, называемые *лучами*. Луч положительного направления называется *положительной полуосью*, а луч отрицательного направления — *отрицательной полуосью*. Пусть задан отрезок, принятый за единицу длины, т. е. введен масштаб.

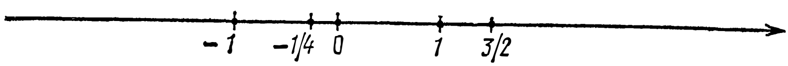


Рис. 1

Сформулируем 2 утверждения, которые здесь принимаются без доказательства.

1. Пусть дан отрезок, длина которого принята за единицу длины. Тогда длина любого отрезка (измеренного этой единицей длины) есть некоторое положительное число.

2. Если дано действительное положительное число a и дан отрезок длины 1, то существует отрезок длины a.

На основании этих утверждений каждой точке оси можно поставить в соответствие одно и только одно вещественное число по следующему правилу.

Выбранной точке O поставим в соответствие число нуль. Для каждой точки N на прямой, отличной от точки O, рассмотрим отрезок ON. Тогда его длина есть положительное рациональное или иррациональное число a.

В случае, если точка N лежит справа от точки O, то поставим ей в соответствие положительное число a, если же точка N лежит слева от точки 0, то поставим ей в соответствие отрицательное число $(-a)$.

Таким образом, каждой точке при выбранном масштабе поставлено в соответствие действительное число. Покажем, что этим процессом перебраны все действительные числа. Предположим противное, т. е. пусть некоторое действительное число b не поставлено в соответствие некоторой точке на оси. Если число b положительное, то найдется (по утверждению 2) отрезок, длина которого равна b. Отложив этот отрезок вправо от точки 0 на оси, получим точку M, которой число b должно соответствовать, т. е. получим противоречие. Если же число b отрицательное, то найдется (по утверждению 2) отрезок, длина которого равна $|b|$. Отложив этот отрезок влево от точки O на оси, получим точку, которой должно соответствовать число b, т. е. опять получим противоречие.

Итак, каждой точке на оси можно поставить в соответствие одно и только одно действительное число и при том нет действительных чисел, которые не соответствовали бы какой-либо точке; кроме того, разным точкам соответствуют разные числа. В таких случаях говорят, что между множеством всех точек числовой оси и множеством всех действительных чисел установлено *взаимно-однозначное соответствие*. Ось, точки которой изображают все действительные числа, называется *числовой осью*.

§ 5
ЧИСЛОВЫЕ РАВЕНСТВА И НЕРАВЕНСТВА

В параграфе 4 упоминалось о сравнении чисел и были приведены определения, по которым можно выяснить, равны ли два данных действительных числа или одно больше другого. Все эти определения можно записать иначе, используя сравнение действительных чисел с числом нуль, а именно так: два действительных числа a и b равны тогда и только тогда, когда их разность равна нулю, т. е. $a=b \Leftrightarrow a-b=0$; число a больше числа b тогда и только тогда, когда разность $(a-b)$ положительна, т. е. $a>b \Leftrightarrow a-b>0$; число a меньше числа b тогда и только тогда, когда разность $(b-a)$ положительна или когда разность $(a-b)$ отрицательна, т. е. $a<b \Leftrightarrow a-b<0 \Leftrightarrow b-a>0$.

Используя эти определения и правила сложения и умножения действительных чисел, легко доказать основные свойства равенств и неравенств чисел.

1. Если числа a, b и c таковы, что $a=b$ и $b=c$, то $a=c$ (свойство транзитивности равенств).

Доказательство. $a-c=(a-c)+(b-b)=(a-b)+(b-c)$. Поскольку $a=b$, то $a-b=0$; аналогично, поскольку $b=c$, то $b-c=0$, поэтому $a-c=(a-b)+(b-c)=0+0=0$, т. е. $a=c$.

2. Если числа a, b, c, d таковы, что $a=b$ и $c=d$, то $a+c=b+d$.

Доказательство. $(a+c)-(b+d)=(a-b)+(c-d)=$ $=0+0=0$, т. е. $a+c=b+d$.

3. Если числа a, b, c, d таковы, что $a=b$, $c=d$, то $ac=bd$.

Доказательство. $ac-bd=(ac-bd)+(ad-ad)=$ $=(ac-ad)+(ad-bd)=a(c-d)+d(a-b)=a\cdot 0+d\cdot 0=0+0=0$, т. е. $ac=bd$.

4. Для любых действительных чисел a, b и c равенства $a=b$ и $a+c=b+c$ равносильны, т. е. из справедливости равенства $a=b$ следует справедливость равенства $a+c=b+c$, и наоборот, из справедливости равенства $a+c=b+c$ следует справедливость равенства $a=b$, т. е. $a=b \Leftrightarrow a+c=b+c$.

Доказательство. Пусть $a=b$. Тогда $(a+c)-(b+c)=$ $=(a-b)+(c-c)=(a-b)+0=a-b=0$, т. е. $a+c=b+c$. Пусть $a+c=b+c$. Тогда $a-b-(a-b)+(c-c)=(a+c)-$ $-(b+c)=0$, т. е. $a=b$.

5. Для любых действительных чисел a и b и для любого действительного отличного от нуля числа c равенства $a=b$ и $ac=bc$ равносильны, т. е. если $c\neq 0$, то $a=b \Leftrightarrow ac=bc$.

Доказательство. Пусть $a=b$. Тогда $ac-bc=c(a-b)=$ $=c\cdot 0=0$, т. е. $ac=bc$. Пусть $ac=bc$. Тогда $c(a-b)=ca-cb=0$. Произведение двух чисел равно нулю, но число $c\neq 0$, значит, число $a-b=0$; т. е. $a=b$.

Докажем теперь аналогичные свойства для числовых неравенств.

1. Если числа a, b и c таковы, что $a>b$ и $b>c$, то $a>c$ (свойство транзитивности неравенств).

Доказательство. $a-c=(a-c)+(b-c)=(a-b)+(b-c)$. Так как $a>b$, то $a-b>0$, так как $b>c$, то $b-c>0$, но сумма двух положительных чисел положительна, поэтому $a-c>0$, т. е. $a>c$.

2. Если числа a, b, c, d таковы, что $a>b$, $c>d$, то $a+c>b+d$.

Доказательство. $(a+c)-(b+d)=(a-b)+(c-d)$. Так как $a>b$, то $(a-b)$ положительное число; так как $c>d$, то $(c-d)$ также положительное число; сумма двух положительных чисел положительна, поэтому $(a+c)-(b+d)>0$, т. е. $a+c>b+d$.

2а. Если числа a, b, c, d таковы, что $a>b$ и $c<d$, то $a-c>b-d$.

Доказательство. $(a-c)-(b-d)=(a-b)+(d-c)$. Так как $a>b$, то $(a-b)$ положительное число, так как $c<d$, то $(d-c)$ также положительное число; сумма двух положительных чисел положительна, поэтому $(a-c)-(b-d)>0$, т. е. $a-c>b-d$.

3. Если a, b, c, d положительные числа и $a>b$ и $c>d$, то $ac>bd$.

Доказательство. $ac-bd=(ac-bd)+(bc-bc)=(ac-bc)+$ $+(bc-bd)=c(a-b)+b(c-d)$. Так как $a>b$, то $(a-b)$ положительное число, так как c положительное число и так как произведение положительных чисел положительно, то $c(a-b)$ положительное число; аналогично показывается, что $b(c-d)$ положительное число; сумма двух положительных чисел положительная, поэтому $ac-bd>0$, т. е. $ac>bd$.

4. Для любых действительных чисел a, b и c неравенства $a>b$ и $a+c>b+c$ равносильны, т. е. из справедливости неравенства $a>b$ следует справедливость неравенства $a+c>$ $>b+c$ и, наоборот, из справедливости неравенства $a+c>b+c$ следует справедливость неравенства $a>b$, т. е. $a>b \Leftrightarrow a+c>$ $>b+c$.

Доказательство. Пусть $a>b$, тогда $(a+c)-(b+c)=(a-b)+(c-c)=a-b>0$, т. е. $a+c>b+c$. Пусть $a+c>b+c$. Тогда $a-b=(a-b)+(c-c)=(a+c)-(b+c)>0$, т. е. $a>b$.

5. Для любых действительных чисел a и b и любого положительного числа c неравенства $a>b$ и $ac>bc$ равносильны, т. е. если $c>0$, то $a>b \Leftrightarrow ac>bc$.

Доказательство. Пусть $a>b$, тогда $ac-bc=c(a-b)$. Так как c и $(a-b)$ положительные числа, то их произведение положительное число, т. е. $ac-bc>0$, или $ac>bc$. Пусть $ac>bc$, тогда $c(a-b)=ac-bc>0$. Если произведение двух чисел положительно и один из сомножителей положителен, то положителен и другой, т. е. так как $c>0$, то $a-b>0$, т. е. $a>b$.

5а. Для любых действительных чисел a и b и любого отрицательного числа c неравенства $a>b$ и $ac<bc$ равносильны, т. е. если $c<0$, то $a>b \Leftrightarrow ac<bc$. Доказательство этого факта аналогично доказательству свойства 5.

Итак, имеют место следующие основные свойства равенств и неравенств:

1) $a=b,\ b=c \Rightarrow a=c$;	1) $a>b,\ b>c \Rightarrow a>c$;
2) $a=b,\ c=d \Rightarrow a+c=b+d$;	2) $a>b,\ c>d \Rightarrow a+c>b+d$;
	2а) $a>b,\ c<d \Rightarrow a-c>b-d$;
3) $a=b,\ c=d \Rightarrow ac=bd$;	3) $a>b>0,\ c>d>0 \Rightarrow ac>bd$;
4) $a=b \Leftrightarrow a+c=b+c$;	4) $a>b \Leftrightarrow a+c>b+c$;
5) $a=b \Leftrightarrow ac=bc$ при $c\neq 0$,	5) $a>b \Leftrightarrow ac>bc$ при $c>0$;
	5а) $a>b \Leftrightarrow ac<bc$ при $c<0$.

Выше употреблялись знаки равенства ($=$) и строгого неравенства ($<$ или $>$). Иногда этих знаков не хватает. Есть задачи, где необходимы нестрогие неравенства.

Пример. Сегодня в Москве 0°, а в Ленинграде температура не выше.

Если температуру в Ленинграде обозначить буквой t, тогда либо $t=0$, либо $t<0$. В таких случаях принято писать $t\leqslant 0$, понимая под этой записью, что либо $t=0$, либо $t<0$.

Приведем определения нестрогих неравенств $a\geqslant b$ и $a\leqslant b$. Числовое неравенство $a\leqslant b$ считается верным и при $a<b$ и при $a=b$ и неверным лишь в случае $a>b$. Например, неравенства $6\leqslant 9$ и $3\leqslant 2+1$ оба верные неравенства, а неравенство $7\leqslant 6$ неверное (знак $\leqslant$ читается как «не больше»).

Числовое неравенство $a\geqslant b$ считается верным и при $a=b$ и при $a>b$; оно считается неверным лишь в случае $a<b$ (знак $\geqslant$ читается как «не меньше»).

Для нестрогих неравенств справедливы свойства 1 — 5а, если в них знак строгого неравенства заменить на знак нестрогого неравенства.

Если два числа соединены знаком равенства, то принято говорить, что задано числовое равенство. Однако это равенство может быть и верным и неверным, так, например, $2=5-3$, $\frac{7}{4}=7\frac{1}{\sqrt{16}}$ верные, а $3=5-1$, $6=7/3$ неверные равенства. Аналогично, например, $110{,}1<11^2$ и $\sqrt{10}>3$ верные, $-5>\sqrt{2}$, $\frac{7}{3}>5$ неверные неравенства.

Справедливость или несправедливость некоторых числовых равенств и неравенств не всегда очевидна. Например, справедливость неравенства

$$(100)^{50}<1\cdot 2\cdot 3\cdot 4\cdot\ldots\cdot 5\ldots 99\cdot 100$$

не очевидна. В этих случаях числовые равенства и неравенства надо доказывать. В гл. II будет приведено доказательство справедливости некоторых числовых равенств и неравенств.

§ 6
ЧИСЛОВЫЕ МНОЖЕСТВА И ЧИСЛОВЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Понятие множества относится к основным понятиям, т. е. к понятиям, которые не определяются.

В этом параграфе рассматриваются множества, элементами которых являются действительные числа. Приведем некоторые примеры числовых множеств:

M_1 — множество всех натуральных чисел;
M_2 — множество всех целых чисел;
M_3 — множество всех целых отрицательных чисел;
M_4 — множество рациональных чисел;
M_5 — множество иррациональных чисел;
M_6 — множество действительных чисел;
M_7 — множество целых чисел, делящихся на 2;
M_8 — множество целых чисел, делящихся на 3;
M_9 — множество чисел, расположенных на числовой оси между числами -1 и 2;
M_{10} — множество чисел, расположенных на числовой оси между числами 0 и 3, включая число 3;
M_{11} — множество чисел, расположенных на числовой оси между числами -2 и 1, включая числа -2 и 1;
M_{12} — множество, состоящее из одного числа нуль;
M_{13} — множество, не имеющее элементов, т. е. пустое множество (оно обычно обозначается символом $\varnothing$);
M_{14} — множество целых чисел, делящихся на 6.

Для множеств M_9, M_{10} и M_{11} приняты специальные обозначения, которые используют строгие и нестрогие неравенства для чисел. Множество M_9 может быть записано так: это все числа a, такие, что $-1<a<2$, для них принята запись $(-1; 2)$. Множество M_{10} может быть записано так: это все числа a, такие, что $0<a\leqslant 3$, для них принята запись $(0; 3]$. Множество M_{11} может быть записано так: это все числа a, такие, что $-2\leqslant a\leqslant 1$, для них принята запись $[-2; 1]$.

Суммой (или *объединением*) множеств α и β называется множество γ, элементами которого являются все элементы множеств α и β. Для суммы, или объединения, множеств употребляется знак $\cup$.

Пример. $M_1 \cup M_3 \cup M_{12} = M_2$; $M_4 \cup M_5 = M_6$; $M_9 \cup M_{10}$ — множество чисел, расположенных на числовой оси между числами -1 и 3, включая число 3. Иная форма записи этого примера: $(-1; 2) \cup (0; 3] = (-1; 3]$; $(-1; 2) \cup [-2; 1] = [-2; 2)$; $[-2; 1] \cup \cup (0; 3] = [-2; 3]$.

Пересечением множеств α и β называется множество γ, элементами которого будут только те элементы, которые одновременно являются элементами и первого и второго множества, т. е. пересечение двух множеств есть общая часть этих множеств. Для пересечения множеств употребляется знак $\cap$.

Пример. $M_1 \cap M_3 \neq \emptyset$; $M_1 \cap M_2 = M_1$; $M_7 \cap M_8 = M_{14}$; $(-1; 2) \cap \cap (0; 3] = (0; 2)$; $[-2; 1] \cap (0; 3] = (0; 1]$; $(-1; 3) \cap (3; 5) = \emptyset$.

Понятие объединения и пересечения множеств часто используются при решении уравнений и неравенств и систем уравнений и неравенств (гл. III).

Если каждому натуральному числу n поставлено в соответствие некоторое число, то говорят, что дана числовая последовательность.

Примеры числовых последовательностей:

а) множество натуральных чисел

$$1,\ 2,\ 3,\ 4,\ \ldots,\ n,\ \ldots;$$

б) множество отрицательных целых чисел

$$-1,\ -2,\ -3,\ -4,\ \ldots,\ -n,\ \ldots;$$

в) множество чисел вида

$$-1,\ +\frac{1}{2},\ -\frac{1}{3},\ +\frac{1}{4},\ \ldots,\ \frac{(-1)^n}{n},\ \ldots;$$

г) множество приближенных значений числа $\sqrt{2}$ с недостатком: 1; 1,4; 1,41; 1,414; 1, 4142; ....

Обычно числовая последовательность обозначается $a_1, a_2, \ldots, a_n, \ldots$, или $\{a_n\}$. Из всевозможных числовых последовательностей здесь будут рассмотрены лишь последовательности, называемые арифметической и геометрической прогрессиями.

Арифметической прогрессией называется последовательность чисел, каждый член которой, начиная со второго, равен предыдущему, сложенному с одним и тем же постоянным для данной последовательности числом, т. е. такая числовая последовательность $\{a_n\}$, что для любого натурального $n: a_{n+1} = a_n + d$, где

d — некоторое постоянное для данной последовательности число, называемое разностью прогрессии.

Например, последовательности

$$1,\ 2,\ 3,\ 4,\ \ldots, \tag{1}$$

$$3,\ 1,\ -1,\ -3,\ -5\ \ldots \tag{2}$$

арифметические прогрессии. У арифметической прогрессии (1) разность $d=1$, а у прогрессии (2) разность $d=-2$. Для любой арифметической прогрессии член, стоящий на n-м месте, всегда можно выразить через первый член и разность данной прогрессии:

$$a_n = a_1 + (n-1)d. \tag{3}$$

Эта формула называется формулой общего члена арифметической прогрессии. Доказательство ее проводится методом математической индукции (см. гл. II).

Используя формулу (3) и свойства действий над числами, легко проверить справедливость следующего утверждения: для любой арифметической прогрессии $\{a_n\}$

$$a_m + a_n = a_k + a_l \tag{4}$$

при $m+n=k+l$.

Действительно,

$$a_m + a_n = a_1 + d(m-1) + a_1 + d(n-1) = 2a_1 + d(m+n-2) =$$
$$= 2a_1 + d(k+l-2) = a_1 + d(k-1) + a_1 + d(l-1) = a_k + a_l.$$

Справедливо более общее утверждение. Если сумма индексов членов арифметической прогрессии постоянна, то и сумма членов с этими индексами постоянна.

Число, равное сумме первых n членов арифметической прогрессии, обозначается S_n, т. е. $S_n = a_1 + a_2 + a_3 + \ldots + a_{n-2} + a_{n-1} + a_n$; члены a_1 и a_n называются крайними. Используя свойство (4) и свойства действий над числами, получим следующую формулу для S_n:

$$S_n = \frac{a_1 + a_n}{2} n. \tag{5}$$

Действительно,

$$2S_n = S_n + S_n = (a_1 + a_2 + a_3 + \ldots + a_{n-2} + a_{n-1} + a_n) +$$
$$+ (a_1 + a_2 + a_3 + \ldots + a_{n-2} + a_{n-1} + a_n) =$$
$$= (a_1 + a_n) + (a_2 + a_{n-1}) + (a_3 + a_{n-2}) + \ldots +$$
$$+ (a_{n-2} + a_3) + (a_{n-1} + a_2) + (a_n + a_1) = (a_1 + b_n)\, n,$$

т. е. сумма первых n членов арифметической прогрессии равна произведению полусуммы крайних членов на число суммируемых членов.

Сумму S_n арифметической прогрессии $\{a_n\}$ можно выразить через первый член и разность данной прогрессии:

$$S_n = \frac{[2a_1 + d(n-1)]\, n}{2}.$$

Пример. Найти сумму двузначных натуральных чисел, не делящихся ни на 2, ни на 3.

Очевидно, что все двузначные натуральные числа образуют арифметическую прогрессию с первым членом $a_1 = 10$ и разностью $d = 1$: 10, 11, 12, ..., 99. Используя формулу (5), легко найти сумму всех этих чисел $S_1 = \frac{10+99}{2} 90$. Двузначные числа, делящиеся на 2, составляют арифметическую прогрессию 10, 12, 14, ..., 98 с суммой $S_2 = \frac{10+98}{2} 45$.

Аналогично числа, делящиеся на 3, образуют арифметическую прогрессию 12, 15, 18, ..., 99 с суммой $S_3 = \frac{12+99}{2} 30$.

Легко заметить, что последние две арифметические прогрессии имеют общие члены 12, 18, ... — числа, делящиеся одновременно и на 2 и на 3, т. е. сумма S всех двузначных натуральных чисел, не делящихся ни на 2, ни на 3, находится следующим образом: $S = S_1 - S_2 - S_3 + S_4$, где S_4 — сумма всех двузначных чисел, делящихся и на 2 и на 3, т. е. делящихся на 6. Используя (3), найдем число членов арифметической прогрессии, составленной из таких чисел (очевидно, что разность этой прогрессии равна 6): $96 = 12 + (n-1)6$, откуда $n = 15$. Следовательно, $S_4 = \frac{12+96}{2} 15$ и $S = \frac{15}{2}(109 \cdot 6 - 108 \cdot 3 - 111 \cdot 2 + 108) = 216$.

Геометрической прогрессией называется последовательность чисел, каждый член которой, начиная со второго, равен предыдущему, умноженному на некоторое постоянное для данной последовательности число, отличное от нуля, т. е. числовая последовательность $\{a_n\}$ такая, что для любого n $a_{n+1} = a_n q$, где q — некоторое постоянное для данной последовательности и отличное от нуля число, называемое знаменателем прогрессии.

Например, последовательности 2, 4, 8, 16, 32, 64, ...; +1, —1, +1, —1, +1, —1, ...; 1/7, —1/21, 1/63, —1/189, ... — геометрические прогрессии соответственно со знаменателями 2, (—1), (—1/3).

Как и в случае арифметической прогрессии, имеет место формула общего члена геометрической прогрессии: $a_n = a_1 q^{n-1}$ (доказательство ее справедливости дано в гл. II).

С помощью формулы общего члена найдем формулу для суммы первых n членов геометрической прогрессии: $S_n = a_1 + a_2 + a_3 + \ldots + a_n = a_1 + a_1 q + a_1 q^2 + \ldots + a_1 q^{n-1}$.

Если $q = 1$, то $S_n = n a_1$. Если $q \neq 1$, то рассмотрим выражение $S_n - qS_n = a_1 + a_1 q + a_1 q^2 + \ldots + a_1 q^{n-1} - a_1 q - a_1 q^2 - \ldots - a_1 q^{n-1} - a_1 q^n = a_1 - a_1 q^n = a_1(1 - q^n)$. Следовательно, $S_n(1 - q) = a_1(1 - q^n)$. Так как $q \neq 1$, то

$$S_n = \frac{a_1(1-q^n)}{1-q}.$$

Рассмотрим конечную числовую последовательность $\{a_1, a_2, \ldots, a_n\}$, состоящую из n элементов. Возьмем m элементов из этого множества $(1 \leqslant m \leqslant n)$ и каким-нибудь образом заново перенумеруем их. Получим новую последовательность, состоящую из m элементов. Такая последовательность называется *размещением из n элементов по m элементов*. Например, из последовательности $\{3, 7, 11\}$ можно организовать новые последовательности — размещения по два элемента: $\{3, 7\}$, $\{3, 11\}$, $\{7, 3\}$, $\{11, 3\}$, $\{7, 11\}$, $\{11, 7\}$. Уже в этом простом примере удалось организовать шесть размещений из 3 элементов по 2 элемента.

Ответим на более общий вопрос, сколько можно организовать размещений из n элементов по два элемента. На первом месте такого размещения может быть любой элемент из n элементов, на втором месте может быть любой из $(n-1)$ оставшихся. Пусть на первом месте стоит элемент a_1, тогда на втором месте может стоять любой из элементов $a_2, a_3, \ldots, a_n$. При этом получим $(n-1)$ размещений. Если на первом месте стоит элемент a_2, то на втором месте может стоять любой из элементов $a_1, a_3, a_4, \ldots, a_n$, т. е. будет еще $(n-1)$ размещений. Перебрав все элементы $a_1, a_2, \ldots, a_n$, получим n групп, в каждой из которых содержится $(n-1)$ размещений. Следовательно, всего размещений из n элементов по два будет $n(n-1)$ штук.

Если нужно узнать число размещений из n элементов по три, то следует к каждому размещению из n элементов по два добавить по очереди один элемент из $(n-2)$ оставшихся. Тогда получится $n(n-1)$ групп, в каждой из которых будет по $(n-2)$ размещений. Следовательно, всего размещений из n элементов по три будет $n(n-1)(n-2)$. Если число размещений из n элементов по m обозначать A_n^m, то можно записать следующие формулы:

$$A_n^2 = n(n-1), \quad A_n^3 = n(n-1)(n-2). \tag{6}$$

Перепишем формулы (6) в ином виде

$$A_n^2 = n[n-(2-1)], \quad A_n^3 = n(n-1)[n-(3-1)]. \tag{7}$$

Можно подметить определенную закономерность в формулах (7): число размещений равно произведению последовательных натуральных чисел начиная с n и кончая $[n-(k-1)]$, где $k=2, 3$. Рассуждая аналогично предыдущему, получим

$$A_n^m = n(n-1)\ldots[n-(m-1)], \quad 1 \leqslant m \leqslant n. \tag{8}$$

Пример. В 1-м классе 6 учебных предметов и 4 урока в день. Сколькими способами можно составить расписание одного учебного дня?

Для того чтобы решить эту задачу, надо найти число размещений из 6 элементов по 4: $A_6^4 = 6 \cdot 5 \cdot 4 \cdot 3 = 360$. Итак, возможно 360 способами составить расписание на день.

Если $m = n$, то любое размещение — это последовательность, составленная из всех элементов данной последовательности, но в другом порядке, т. е. произведена определенная перестановка элементов. Размещения из n элементов по n называются *перестановками* и обозначаются P_n. По формуле (8)

$$P_n = n(n-1)(n-2)\ldots 2 \cdot 1. \tag{9}$$

Если воспользоваться символом $n!$ (читается: n факториал), который обозначает произведение n первых чисел натурального ряда ($n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$), то формулу (9) можно записать:

$$P_n = n!. \tag{10}$$

Формулу (8) тоже можно записать, используя этот символ:

$$A_n^m = \frac{n(n-1)(n-2)\ldots[n-(m-1)](n-m)\ldots 2 \cdot 1}{(n-m)\ldots 2 \cdot 1} = \frac{n!}{(n-m)!}. \tag{8'}$$

Чтобы формула (8′) совпадала с формулой (10) при $m = n$ принято считать, что $0! = 1$.

Часто возникают задачи, когда из данного конечного множества из n элементов надо организовать множество из m элементов ($m \leqslant n$), но не обязательно последовательность, т. е. во вновь построенном множестве порядок размещения элементов не важен, а важно лишь их наличие.

Например, требуется выбрать трех учеников из десяти для уборки класса. В этом случае упорядоченность в группе из трех человек необязательна. Такие множества из n элементов по m, которые отличаются друг от друга только элементами, но не порядком их расположения, называются *сочетаниями* и обозначаются C_n^m. Очевидно, что число сочетаний из n элементов по n равно единице: $C_n^n = 1$. Расссмотрим общий случай, когда $1 \leqslant m \leqslant n$.

Пусть составлены все сочетания C_n^m из n элементов по m. Возьмем теперь любое из этих сочетаний и переставим в нем элементы всевозможными способами. Тогда число полученных всевозможных множеств из n элементов по m равно $C_n^m \cdot P_m$. Покажем, что это число совпадает с числом всех размещений из n элементов по m. Действительно, возьмем всевозможные размещения и запишем их по группам. В каждую группу включим размещения, составленные из одинаковых элементов, отличающиеся порядком их расположения. Таким образом, в каждую группу войдет столько размещений, сколько можно организовать перестановок из данных m элементов, т. е. $m!$ размещений. Все размещения, расположенные в одной группе, как сочетания одинаковы, так как содержат одинаковые элементы. Следовательно, число групп — это число различных сочетаний из n элементов по m, т. е.

$$A_n^m = C_n^m m! \tag{11}$$

Из равенства (11) получим формулу для подсчета числа сочетаний:

$$C_n^m = \frac{A_n^m}{m!}$$

или, используя формулу (8′):

$$C_n^m = \frac{n!}{(n-m)!\, m!}. \tag{12}$$

Решим с помощью формулы (12) задачу о выборе трех учеников:

$$C_{10}^3 = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10}{(1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7)\cdot(1\cdot 2\cdot 3)} = 120.$$

Если подсчитаем C_{10}^7, то получим тот же результат

$$C_{10}^7 = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10}{(1\cdot 2\cdot 3)\cdot(1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7)} = 120.$$

Покажем в общем случае, что $C_n^m = C_n^{n-m}$ $(1 \leqslant m \leqslant n-1)$. Действительно,

$$C_n^{n-m} = \frac{n!}{[n-(n+m)]!\,(n-m)!} = \frac{n!}{m!\,(n-m)!} = C_n^m. \tag{13}$$

Формула (13) позволяет легко подсчитать число сочетаний из n по m, когда m близко к n. Например

$$C_9^8 = C_9^1 = \frac{9!}{8!} = 9.$$

Покажем справедливость еще одного свойства сочетаний C_n^m:

$$C_n^{m+1} + C_n^m = C_{n+1}^{m+1}. \tag{14}$$

Действительно, по определению сочетаний можно написать

$$C_n^{m+1} + C_n^m = \frac{n!}{(m+1)!\,(n-m-1)!} + \frac{n!}{m!\,(n-m)!} =$$
$$= \frac{n!}{m!\,(m+1)\,(n-m-1)!} + \frac{n!}{m!\,(n-m-1)!\,(n-m)} =$$
$$= \frac{n!}{m!\,(n-m-1)!}\left(\frac{1}{m+1} + \frac{1}{n-m}\right) =$$
$$= \frac{n!\,(n+1)}{m!\,(n-m-1)!\,(m+1)\,(n-m)} = \frac{(n+1)!}{(m+1)!\,[(n+1)-(m+1)]!}.$$

Используя определение C_{n+1}^{m+1}, получаем справедливость формулы (14).

Глава II

АЛГЕБРАИЧЕСКИЕ ВЫРАЖЕНИЯ

В предыдущей главе были рассмотрены действительные числа и некоторые действия над ними. С помощью чисел, знаков действий и скобок составлялись различные *числовые выражения*. Приведем примеры некоторых числовых выражений:

$$27:9;\ \sqrt{8+1};\ \frac{\left(\frac{3}{11}-\frac{7}{2}\right)\frac{11}{71}-\left(\frac{6}{5}+\frac{1}{7}\right)\frac{35}{47}}{\left(\frac{1}{19}-\frac{1}{2}\right)\frac{19}{17}};\ 2\cdot 10^3+7\cdot 10^2+10-5.$$

Если в числовом выражении можно выполнить все указанные в нем действия, то полученное действительное число называют числовым значением данного числового выражения, а о числовом выражении говорят, что оно имеет смысл. В приведенных примерах каждое из первых трех числовых выражений имеет числовое значение 3, а четвертое — 2705.

Если числовое выражение состоит из одного действительного числа, то его числовым значением является само это число.

Иногда числовое выражение не имеет числового значения, так как не все указанные в нем действия выполнимы; о таком числовом выражении говорят, что оно не имеет (лишено) смысла. Например, числовые выражения $7/(3\cdot 2-6)$; $\sqrt{10-18}$; $(2-2)^0$ лишены смысла.

Таким образом, любое числовое выражение либо имеет одно числовое значение, либо лишено смысла.

Числовое выражение часто употребляют для описания какого-либо свойства числа, являющегося числовым значением этого выражения. Так, например, свойство числа -17 давать при делении на 2 остаток 1 записывают числовым выражением $2(-9)+1$. Чтобы описать свойство каждого нечетного числа из промежутка $[-2, 14]$ давать при делении на 2 остаток 1, надо написать соответствующее числовое выражение для каждого из чисел -1, 1, 3, 5, 7, 9, 11 и 13, т. е. восемь следующих числовых выражений: $2(-1)+1$, $2\cdot 0+1$, $2\cdot 1+1$, $2\cdot 2+1$, $2\cdot 3+1$, $2\cdot 4+1$, $2\cdot 5+1$ и $2\cdot 6+1$. Выписать соответствующие числовые выражения для всех тех целых нечетных чисел, каждое из которых обладает указанным свойством, практически нельзя. Замечая общность составления таких числовых выражений и используя буквенную символику, можно сокращенно записать всю бесконечную совокупность таких числовых выражений:

$$n=2l+1, \text{ где } l=0, \pm 1, \pm 2, \pm \dots \qquad (1)$$

При каждом l получаем числовое выражение, числовое значение которого есть целое число n, дающее при делении на 2 остаток 1. Для любого целого числа n, обладающего указанным свойством, можно указать число l, такое, при котором (1) превращается в числовое выражение, имеющее числовым значением число n. Запись $4l+3$, где $l=0, \pm 1, \pm 2, \pm \dots$ представляет собой бесконечную совокупность числовых выражений таких, что при каждом указанном l она превращается в числовое выражение, числовым значением которого является число n, дающее при делении на 4 остаток 3.

Приведенные выше примеры говорят о том, что часто вместо числовых выражений удобнее рассматривать выражения, в которых на некоторых местах вместо чисел стоят буквы. Всякое такое выражение называют *математическим выражением*. Примеры математических выражений:

$$\frac{7}{a}+2^{b+3}; \ \sin\frac{b-a}{c}; \ \sqrt{3+\lambda}; \ \operatorname{arctg}\frac{\alpha}{\beta}; \ \log_c\frac{3+\sqrt{m}}{n}.$$

Отметим, что понятие «математическое выражение» является простейшим и потому оно не определяется, а лишь описывается, что и было сделано выше. Математическое выражение, в котором участвуют знаки действий сложения, умножения, вычитания, деления, извлечения корня и возведения в степень, называют *алгебраическим выражением*. Примеры алгебраических выражений:

$$\frac{\alpha}{2m}+3a^b; \ \frac{a-b}{c-a}; \ \frac{3b}{\sqrt{3+a}}.$$

§ 1
ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Пусть даны два алгебраических выражения, которые обозначены буквами A и B. Определим для них арифметические операции.

Сложить два алгебраических выражения A и B — значит формально написать алгебраическое выражение $A+B$, которое называется суммой выражений A и B. Например, суммой алгебраических выражений $(a-b)/(c-a)$ и $\alpha/2p$ будет алгебраическое выражение $(a-b)/(c-a)+\alpha/2p$.

Умножить два алгебраических выражения A и B — значит формально написать алгебраическое выражение AB, которое называется произведением выражений A и B. Например, произведением алгебраических выражений $\frac{3b}{\sqrt{x+y}}$ и a^2-b^2 будет алгебраическое выражение $\frac{3b}{\sqrt{x+y}}(a^2-b^2)$.

Вычесть из алгебраического выражения A алгебраическое выражение B — это значит формально написать алгебраическое выражение $A-B$, которое называется разностью выражений A и B. Например, разностью алгебраических выражений abc^3 и a^{2mn}/pq будет алгебраическое выражение abc^3-a^{2mn}/pq.

Разделить алгебраическое выражение A на алгебраическое выражение B — это значит формально написать алгебраическое выражение $A:B$, которое называется частным от деления выражения A на выражение B. Например, частным от деления алгебраического выражения $a-b^2$ на алгебраическое выражение $p/5l$ будет алгебраическое выражение $(a-b^2):(p/5l)$. Отметим, что частное от деления алгебраического выражения A на алгебраическое выражение B часто записывается в виде A/B.

Если надо сложить несколько алгебраических выражений, то сначала складывают два первых выражения, затем к полученной сумме прибавляют третье выражение и т. д. Вот, например, как выглядит сумма пяти алгебраических выражений $\{[(A+B)+C]+D\}+E$.

Аналогично определяется и произведение нескольких алгебраических выражений. Если в произведении одно и то же алгебраическое выражение A является множителем n раз $(n>1)$, то пишут A^n вместо произведения $\underbrace{AA\ \ldots\ A}_{n\ \text{раз}}$. Кроме того, иногда пишут A^1 вместо A.

Некоторые частные случаи алгебраических выражений имеют специальные названия.

Алгебраическое выражение называется *рациональным*, если в нем участвуют относительно входящих в него букв лишь четыре арифметических действия: сложение, умножение, вычитание и деление (рациональное алгебраическое выражение может содержать любые числа, в том числе и иррациональные). Примеры рациональных алгебраических выражений:

$$a+b;\ a-c/a-b;\ 2-\gamma/\beta;\ \sqrt{2}\alpha\beta-1;\ \sqrt{3}(a^2-b^2)/(m-n)+xy.$$

Рациональное выражение называется *целым* относительно данной буквы, если оно не содержит деления на данную букву или на выражение, содержащее эту букву. *Дробное рациональное* выражение относительно данной буквы — это выражение, содержащее деление на некоторое выражение, содержащее эту букву, или на саму букву.

Например, рациональное выражение $(a+b+c)/(3a+4b)$ — целое относительно буквы c, но дробное относительно букв a и b; рациональное выражение $3a/7+\sqrt{2}/b$ — целое относительно a, но дробное относительно b.

Рациональное выражение, содержащее относительно входящих в него букв только одно действие — умножение, называется *одночленом*.

Примеры одночленов: $3a$, $2ab/3$, $2abc\frac{23ab}{7}abc$.

Рациональное выражение называется *многочленом*, если оно является целым относительно каждой буквы, входящей в это выражение.

Например, рациональное выражение $\sqrt{35}\,abc-\frac{16ad}{7}+0{,}3dc$ является многочленом, ибо это выражение является целым относительно букв a, b, c и d.

Из определения многочлена и правил действий над алгебраическими выражениями следует, что сумма, разность и произведение двух многочленов будут многочленами. Частное от деления одного многочлена на другой будет дробным рациональным выражением, которое имеет специальное название.

Алгебраической дробью называется дробное рациональное выражение, являющееся частным от деления одного многочлена на другой.

Алгебраическая дробь, которая есть частное от деления многочлена A на многочлен B, обычно записывается в виде A/B, причем многочлен A называется числителем алгебраической дроби, а многочлен B — ее знаменателем.

Примеры алгебраических дробей:

$$\frac{2a+b}{a^3+1};\quad \frac{ab-b}{d+\alpha};\quad \frac{a^2-b^2}{a-b};\quad \frac{3xy+6x}{7x+8y}.$$

Часто алгебраические выражения приходится преобразовывать. Приведем необходимые для этого определения.

Пусть дано некоторое алгебраическое выражение. Множество всех букв, входящих в это выражение, называется *буквенным набором* данного алгебраического выражения.

Если в алгебраическое выражение входит n букв a_1, a_2, ..., a_n, то буквенный набор этого алгебраического выражения записывают в виде $(a_1, a_2, \ldots, a_n)$. Каждая буква, сколько бы раз она ни встречалась в алгебраическом выражении, пишется в буквенном наборе только один раз. При составлении буквенного набора данного алгебраического выражения порядок следования букв может быть любым возможным, но раз навсегда зафиксированным.

Например, для алгебраического выражения $(2a-7b)/\sqrt{19ac}$ буквенным набором может служить набор (a, b, c), для алгебраического выражения $\sqrt[3]{\alpha/2+a^3b^k}$ — набор (k, α, a, b). Если в буквенном наборе (α, β, γ) вместо буквы α написать, например, число $-7/16$, вместо буквы β — число $\sqrt{2}$, вместо буквы γ — число $0{,}3$, то совокупность чисел $-7/16$, $\sqrt{2}$ и $0{,}3$ называют *числовым набором*, соответствующим данному буквенному набору (α, β, γ) и записывают в виде $(-7/16, \sqrt{2}, 0{,}3)$. При этом говорят, что числовой набор $(-7/16, \sqrt{2}, 0{,}3)$ соответствует буквенному набору (α, β, γ) при $\alpha=-7/16$, $\beta=\sqrt{2}$, $\gamma=0{,}3$.

Аналогично определяется числовой набор, соответствующий буквенному набору $(a_1, a_2, \ldots, a_n)$ и для любого набора из n букв (n — любое натуральное число). Одному и тому же буквенному набору можно поставить в соответствие бесконечно много разных числовых наборов. Два числовых набора считаются разными, если хотя бы на одном, но на одном и том же в каждом наборе, например i-м, месте этих числовых наборов стоят неравные числа (т. е. вместо одной и той же буквы, стоящей на i-м месте буквенного набора в этих двух числовых наборах взяты неравные числа). Например, числовые наборы $(1, -3{,}5, -\sqrt{2}, 7/3)$ и $(1, -3, -4, -\sqrt{2}, 7/3)$, соответствующие буквенному набору (a, b, c, d, e), разные, так как у них на одном и том же третьем месте стоят неравные числа 5 и -4 (т. е. в первом наборе $c=5$, а во втором — $c=-4$). Числовые наборы $(-3, -9/7)$ и $(-9/7, -3)$, соответствующие буквенному набору (x, y), разные, так как у них на первом месте стоят неравные числа (т. е. в первом наборе $x=-3$, во втором — $x=-9/7$); кроме того, они разные, так как у них на втором месте стоят неравные числа (т. е. в первом наборе $y=-9/7$, а во втором — $y=-3$).

Пусть дано некоторое алгебраическое выражение и его буквенный набор. Рассмотрим некоторый числовой набор, соответствующий этому буквенному набору. Этот числовой набор называется числовым набором для букв алгебраического выражения. Если в это алгебраическое выражение подставить вместо каждой буквы, где бы она в нем ни стояла, соответствующее ей число из данного числового набора, то получим числовое выражение, которое либо имеет смысл, либо лишено смысла. Например, рассмотрим алгебраическое выражение $(b-3a)/\sqrt{5+a}$. Запишем его буквенный набор в виде (a, b). Для числового набора $(4, 5)$, соответствующего буквенному набору (a, b) (т. е. для $a=4$, $b=5$), это алгебраическое выражение записывается в виде числового выражения $(5-3\cdot 4)/\sqrt{5+4}$ и имеет числовое значение $-7/3$. Для числового набора $(-6, 5)$ (т. е. для $a=-6$, $b=5$) это алгебраическое выражение запишется в виде числового выражения $[5-3(-6)]/\sqrt{5-6}$, которое лишено смысла.

Числовой набор, соответствующий буквенному набору данного алгебраического выражения, называется *допустимым* для этого выражения, если имеет смысл числовое выражение, которое получается из данного алгебраического выражения, если вместо каждой буквы, где бы она в нем ни стояла, подставить соответствующее ей число из данного числового набора. Совокупность всех допустимых числовых наборов, соответствующих буквенному набору данного алгебраического выражения, называется *областью допустимых значений* (ОДЗ) данного алгебраического выражения.

Отметим, что существуют алгебраические выражения, ОДЗ которых пуста. Например, пуста ОДЗ алгебраического выражения $1/[2a-(a+a)]$, ибо для любого числового значения буквы a соответствующее числовое выражение лишено смысла. Такие выражения называются выражениями, не имеющими смысла, и в дальнейшем рассматриваться не будут. Обычно ОДЗ алгебраического выражения записывают в виде набора множеств, причем указывают, какой букве соответствует какое множество. Так, например, ОДЗ алгебраического выражения $(b-3)/\sqrt{5+a}$ записывается в виде $\{a \in (-5, +\infty);\ b \in (-\infty, +\infty)\}$, или $\{-5 < a < +\infty;\ -\infty < b < +\infty\}$, или $\{a > -5;$ b — л. д. ч.$\}$ (л. д. ч. — любое действительное число).

Числовым значением, или числовой величиной, алгебраического выражения для данного числового набора из ОДЗ называют числовое значение того числового выражения, которое получится, если в данное алгебраическое выражение вместо каждой буквы, где бы она в нем ни стояла, подставить соответствующее ей число из данного числового набора.

Например, числовым значением алгебраического выражения $(b-3)/\sqrt{5+a}$ при $a=4$ и $b=5$ будет число $2/3$, а при $a=0$ и $b=4$ — число $1/\sqrt{5}$.

Часто алгебраические выражения рассматриваются не на всей своей ОДЗ, а лишь на ее части — некоторой области M. Например, рассмотрим алгебраическое выражение vt. ОДЗ этого выражения: $\{v \in (-\infty, +\infty);\ t \in (-\infty, +\infty)\}$. Пусть это алгебраическое выражение vt определяет путь, пройденный за время t со скоростью v. Тогда по физическому смыслу задачи следует наложить на v и t ограничения: $v > 0$ и $t \geqslant 0$. Другими словами, надо рассмотреть алгебраическое выражение vt на следующей области M — части ОДЗ этого выражения: $M = \{v \in (0, +\infty);\ t \in [0, +\infty)\}$. Алгебраическое выражение обычно дается вместе с областью M, на которой оно рассматривается. Если область M не указана, то алгебраическое выражение следует рассматривать на всей ОДЗ, которую предварительно надо найти.

Пусть даны два алгебраических выражения A и B. Множество всех букв этих двух выражений называют буквенным набором двух выражений A и B. Числовой набор, соответствующий буквенному набору двух алгебраических выражений, называют допустимым, если одновременно имеют смысл оба числовых выражения, которые получаются из данных алгебраических выражений, если в них вместо каждой буквы, где бы она в них ни стояла, подставить соответствующее ей число из этого числового набора. Совокупность всех допустимых числовых наборов, соответствующих буквенному набору двух алгебраических выражений, называется областью допустимых значений (ОДЗ) этих алгебраических выражений.

Аналогично определяется ОДЗ n алгебраических выражений. Два алгебраических выражения можно рассматривать не на всей ОДЗ, а лишь на некоторой ее части — некоторой области M. Поэтому дальше под областью M, принадлежащей ОДЗ алгебраических выражений, будет пониматься либо вся эта ОДЗ, либо какая-нибудь явно указываемая ее часть.

Два алгебраических выражения называются *тождественно равными на области* M, если они имеют равные числовые значения для любого данного числового набора из области M.

Например, два алгебраических выражения $\sqrt{a^2}$ и $|a|$ тождественно равны как на всей ОДЗ выражений $\sqrt{a^2}$ и $|a|$, т. е. на области $\{a \in (-\infty, +\infty)\}$, так и на любой ее части. Два алгебраических выражения $d+(a-b)/(3+c)$ и $a/(3+c)-b/(3+c)$ тождественно равны не на всей ОДЗ этих двух выражений, которая есть область $\{a \in (-\infty, \infty);\ b \in (-\infty, +\infty);\ c \in (-\infty, -3) \cup (-3, +\infty),\ d \in (-\infty, +\infty)\}$, а лишь на ее части — области M, где $M=\{a \in (-\infty, +\infty);\ b \in (-\infty, +\infty);\ c \in (-\infty, -3) \cup (-3, +\infty);\ d=0\}$. Для записи тождественного равенства на области M двух алгебраических выражений иногда употребляется знак равенства, над которым сверху написана буква M, т. е. если буквами A и B обозначены некоторые алгебраические выражения, то запись $A \overset{\text{м}}{=} B$ означает, что алгебраические выражения A и B тождественно равны на области M, а область M входит в ОДЗ двух выражений A и B.

Например, запись $(a-b)/(3+c) \overset{\text{ОДЗ}}{=\!=} a/(3+c)-b/(3+c)$ означает, что алгебраические выражения $(a-b)/(3+c)$ и $a/(3+c)-b/(3+c)$ тождественно равны на ОДЗ этих выражений, т. е. на области $\{a \in (-\infty, +\infty);\ b \in (-\infty, +\infty);\ c \in (-\infty, -3) \cup (-3, +\infty)\}$, а запись $\sqrt{a^2} \overset{\text{м}}{=} a$, где $M=\{a \in [0, 1]\}$ означает, что утверждается тождественное равенство алгебраических выражений $\sqrt{a^2}$ и a лишь на области $M=\{a \in [0, 1]\}$.

Замена алгебраического выражения A алгебраическим выражением B, тождественно равным ему на области M, принадлежащей ОДЗ выражений A и B, называется *тождественным преобразованием* на области M алгебраического выражения A. Если не указана область M, на которой происходит тождественное преобразование, то принято считать, что это преобразование происходит на ОДЗ двух выражений: данного и преобразованного.

Например, замена алгебраического выражения $\sqrt{a^2}$ алгебраическим выражением $|a|$ является тождественным преобразованием на ОДЗ этих выражений, т. е. на области M, где $M = \{a \in (-\infty, +\infty)\}$.

§ 2
РАВЕНСТВА И НЕРАВЕНСТВА ДЛЯ АЛГЕБРАИЧЕСКИХ ВЫРАЖЕНИЙ

Пусть даны два алгебраических выражения, которые обозначены буквами A и B. В § 1 запись $A \overset{M}{=} B$ употреблялась для более короткой записи следующего утверждения: известно, что два алгебраических выражения A и B тождественно равны на области M и область M принадлежит ОДЗ двух выражений A и B. Законен вопрос, а возможна ли запись $A = B$ без буквы M над знаком равенства и что эта запись означает?

Конечно, формально можно сделать запись $A = B$, но если рядом нет слов, поясняющих, как следует понимать такую запись, то такая запись не несет никакой смысловой нагрузки. Следовательно, такая запись должна употребляться только с некоторыми сопровождающими эту запись пояснениями, которые и разъяснят, как следует понимать эту запись.

Приведем теперь наиболее часто встречающиеся случаи употребления записи $A = B$ с соответствующими пояснениями, как следует понимать такую запись.

а) Пусть известно, что на некоторой области M, принадлежащей ОДЗ двух алгебраических выражений A и B, эти два выражения тождественно равны, тогда это утверждение записывают так: «Известно (или дано), что $A = B$ на области M». В этом случае говорят также, что на области M дано тождественное равенство $A = B$.

б) Пусть требуется доказать справедливость утверждения: алгебраические выражения A и B тождественно равны на области M, принадлежащей ОДЗ этих выражений, тогда пишут: «Доказать, что $A = B$ на области M». В этом случае говорят также, что требуется доказать справедливость на области M тождественного равенства $A = B$.

в) Пусть требуется найти область M, принадлежащую ОДЗ двух алгебраических выражений A и B, такую, что для любого числового набора из области M соответствующее числовое значение выражения A равно соответствующему числовому значению выражения B, а для любого числового набора, не входящего в область M, но входящего в ОДЗ этих выражений, соответствующие числовые значения данных выражений не равны. В таких случаях говорят: «Решить уравнение $A = B$».

В этой главе рассматриваются лишь тождественные равенства. Решение уравнений будет рассмотрено в гл. III, VII и IX.

Прежде всего отметим, что действия сложения и умножения алгебраических выражений производятся на основании следующих утверждений.

1. На ОДЗ двух выражений A и B справедливо тождественное равенство $A+B=B+A$.

2. На ОДЗ трех выражений A, B и C справедливо тождественное равенство $(A+B)+C=A+(B+C)$.

3. На ОДЗ двух выражений A и B справедливо тождественное равенство $AB=BA$.

4. На ОДЗ трех выражений A, B и C справедливо тождественное равенство $(AB)C=A(BC)$.

5. На ОДЗ трех выражений A, B и C справедливо тождественное равенство $A(B+C)=AB+AC$.

6. На ОДЗ выражения A справедливы тождественные равенства $A+0=A$, $0+A=A$.

7. На ОДЗ выражения A справедливы тождественные равенства

$$A\cdot 1=A, \ 1\cdot A=A.$$

8. На ОДЗ выражения A справедливо тождественное равенство $A+(-A)=0$.

9. На области M части ОДЗ выражения A, на которой ни для одного числового набора соответствующее числовое значение выражения A не обращается в нуль, справедливо тождественное равенство $A\cdot\frac{1}{A}=1$.

Поскольку метод доказательства справедливости этих утверждений один и тот же, то приведем здесь доказательство лишь утверждения 1.

Возьмем некоторый числовой набор из ОДЗ двух выражений A и B и обозначим соответствующие числовые значения этих выражений соответственно через A_0 и B_0. Тогда для чисел A_0 и B_0 по свойству коммутативности сложения чисел справедливо числовое равенство $A_0+B_0=B_0+A_0$. Значит, показано, что для данного числового набора из ОДЗ двух выражений A и B соответствующие числовые значения выражений A_0+B_0 и B_0+A_0 равны. Так как это рассуждение можно провести для любого числового набора из ОДЗ двух выражений A и B, то справедливость на ОДЗ этих выражений тождественного равенства $A+B=B+A$ доказана.

Аналогично доказываются и следующие утверждения.

10. На ОДЗ двух выражений A и B справедливо тождественное равенство $A-B=A+(-B)$.

11. На области M части ОДЗ двух выражений A и B, на которой ни для одного числового набора соответствующее

числовое значение выражения B не обращается в нуль, справедливо тождественное равенство $\frac{A}{B} = A\frac{1}{B}$.

12. На ОДЗ двух выражений A и B справедливо тождественное равенство $B+(A-B)=A$.

13. На области M части ОДЗ двух выражений A и B, на которой ни для одного числового набора соответствующее числовое значение выражения B не обращается в нуль, справедливо тождественное равенство $B\left(\frac{A}{B}\right)=A$.

Утверждения 12 и 13 показывают, что действия вычитания и деления алгебраических выражений, определенные в § 1, являются соответственно обратными к действиям сложения и умножения алгебраических выражений. При доказательстве тождественных равенств часто приходится пользоваться следующими утверждениями.

14. Если на некоторой области M, принадлежащей ОДЗ трех алгебраических выражений A, B и C, одновременно справедливы тождественные равенства $A=B$ и $B=C$, то на области M справедливо и тождественное равенство $A=C$ (транзитивность равенств).

15. Если на некоторой области M, принадлежащей ОДЗ четырех алгебраических выражений A, B, C и D, одновременно справедливы тождественные равенства $A=B$ и $C=D$, то на области M справедливо и тождественное равенство $A+C = B+D$.

16. Если на некоторой области M, принадлежащей ОДЗ четырех алгебраических выражений A, B, C и D, одновременно справедливы тождественные равенства $A=B$ и $C=D$, то на области M справедливо и тождественное равенство $AC=BD$.

17. а) Если на ОДЗ трех алгебраических выражений A, B и C справедливо тождественное равенство $A=B$, то на ОДЗ справедливо и тождественное равенство $A+C=B+C$.

б) Если на ОДЗ трех алгебраических выражений A, B и C справедливо тождественное равенство $A+C=B+C$, то на ОДЗ справедливо и тождественное равенство $A=B$.

18. Пусть некоторая область M принадлежит ОДЗ трех алгебраических выражений A, B и C и обладает свойством: ни для какого числового набора из области M соответствующее числовое значение выражения C не равно нулю, тогда:

а) если на области M справедливо тождественное равенство $A=B$, то на этой области справедливо и тождественное равенство $AC=BC$;

б) если на области M справедливо тождественное равенство $AC=BC$, то на этой области справедливо и тождественное равенство $A=B$.

Поскольку метод доказательства утверждений 14—18 один и тот же, то приведем здесь лишь доказательство утверждения 14.

Возьмем некоторый числовой набор из области M. Обозначим соответствующие числовые значения выражения A, B и C соответственно через A_0, B_0 и C_0. Из справедливости на области M тождественных равенств $A=B$ и $B=C$ вытекает справедливость числовых равенств $A_0=B_0$ и $B_0=C_0$. По свойству транзитивности числовых равенств тогда справедливо и числовое равенство $A_0=C_0$. Таким образом, показано, что для данного числового набора из области M соответствующие числовые значения выражений A и C равны. Поскольку это рассуждение можно провести для любого числового набора из области M, то справедливость на области M тождественного равенства $A=C$ доказана.

Если на некоторой области M из справедливости одного тождественного равенства вытекает справедливость второго, а из справедливости второго вытекает справедливость первого, то говорят, что такие два *тождественные равенства равносильны* на *области* M, а замену одного из них другим называют *равносильным переходом* на области M от первого ко второму. Поэтому утверждения 17 и 18 можно сформулировать так:

17. На ОДЗ трех алгебраических выражений A, B и C тождественные равенства $A=B$ и $A+C=B+C$ равносильны.

18. Пусть некоторая область M принадлежит ОДЗ трех алгебраических выражений A, B и C и обладает свойством: ни для какого числового набора из области M соответствующее числовое значение выражения C не равно нулю. Тогда на области M тождественные равенства $A=B$ и $AC=BC$ равносильны.

Приведем еще одно утверждение, которое также доказывается аналогично предыдущим утверждениям.

19. На ОДЗ двух алгебраических выражений A и B равносильны тождественные равенства $A=B$ и $A-B=0$.

Равносильный переход на области M от одного тождественного равенства к другому обозначается иногда двойной стрелкой, над которой сверху написана буква M, т. е. запись $A=B \overset{M}{\Longleftrightarrow} C=D$ означает, что на области M равенства $A=B$ и $C=D$ равносильны.

Поэтому утверждения 17, 18 и 19 можно записать так:

17. Пусть M — ОДЗ трех алгебраических выражений A, B и C, тогда $A=B \overset{M}{\Longleftrightarrow} A+C=B+C$.

18. Пусть некоторая область M принадлежит ОДЗ трех алгебраических выражений A, B и C и обладает свойством:

ни для какого числового набора из области M соответствующее числовое значение выражения C не равно нулю, тогда

$$A=B \overset{M}{\Leftrightarrow} AC=BC.$$

19. Пусть M — ОДЗ двух алгебраических выражений A и B, тогда $A=B \overset{M}{\Leftrightarrow} A-B=0$.

Принято следующее соглашение: если не указана явно область M, на которой рассматривается тождественное равенство $A=B$, то оно рассматривается на ОДЗ двух выражений A и B. Поэтому слова: дано, что $A=B$, — означают, что на ОДЗ двух выражений A и B справедливо тождественное равенство $A=B$; слова: доказать, что $A=B$, — означают: доказать, что на ОДЗ двух выражений A и B справедливо тождественное равенство $A=B$ (при этом имеется в виду, что эту ОДЗ обязательно следует отыскать).

В частности, исходя из этого, утверждения 1—9, называемые обычно законами действий сложения и умножения алгебраических выражений, можно переписать и так.

Справедливы следующие *основные законы* сложения и умножения алгебраических выражений:

1. $A+B=B+A$ (коммутативность сложения);
2. $(A+B)+C=A+(B+C)$ (ассоциативность сложения);
3. $AB=BA$ (коммутативность умножения);
4. $(AB)C=A(BC)$ (ассоциативность умножения);
5. $(A+B)C=AC+BC$ (дистрибутивность сложения относительно умножения);
6. $A+0=A$; 7. $A\cdot 1=A$;
8. $A-A=0$; 9. $A\cdot\frac{1}{A}=1$.

Отметим, что если дано равенство $A=B$, то ОДЗ двух выражений A и B называют часто ОДЗ равенства $A=B$.

Перейдем теперь к употреблению знака неравенства для алгебраических выражений. Знак неравенства ($>$, $\geqslant$, $<$ или $\leqslant$), так же как и знак равенства, употребляется для алгебраических выражений только с некоторыми пояснениями, как такую запись следует понимать. Приведем наиболее часто встречающиеся случаи употребления этих знаков.

а) Пусть известно, что на некоторой области M, принадлежащей ОДЗ двух алгебраических выражений A и B, для любого числового набора из M соответствующее числовое значение выражения A больше соответствующего числового значения выражения B, тогда это утверждение записывается так: «Известно (или дано), что $A>B$ на области M». В этом случае говорят также, что на области M справедливо тождественное неравенство $A>B$.

б) Пусть требуется доказать справедливость утверждения: для любого числового набора из области M, принадлежащей ОДЗ двух выражений A и B, соответствующее числовое значение выражения A больше соответствующего числового значения выражения B, тогда пишут: «Доказать, что $A > B$ на области M». В этом случае говорят также, что требуется доказать справедливость на области M тождественного неравенства $A > B$.

в) Пусть требуется найти область M, принадлежащую ОДЗ двух алгебраических выражений A и B, такую, что для любого числового набора из области M соответствующее числовое значение выражения A больше соответствующего числового значения выражения B, а для любого числового набора из ОДЗ, не входящего в область M, соответствующее числовое значение выражения A меньше или равно соответствующему числовому значению выражения B. В таких случаях говорят: «Решить неравенство $A > B$».

В этой главе рассматриваются лишь тождественные неравенства. Решение неравенств будет рассмотрено в главах III и VII.

При доказательстве тождественных неравенств часто приходится пользоваться следующими утверждениями.

20. Если на некоторой области M, принадлежащей ОДЗ трех алгебраических выражений A, B и C, одновременно справедливы тождественные неравенства $A > B$ и $B > C$, то на области M справедливо и тождественное неравенство $A > C$.

21. Если на некоторой области M, принадлежащей ОДЗ четырех алгебраических выражений A, B, C и D, одновременно справедливы тождественные неравенства $A > B$ и $C > D$, то на области M справедливо и тождественное неравенство $A + C > B + D$.

22. Если для любого числового набора из некоторой области M, принадлежащей ОДЗ четырех алгебраических выражений A, B, C и D, соответствующие числовые значения этих выражений A, B, C и D положительны и если на этой области одновременно справедливы тождественные неравенства $A > B$ и $C > D$, то справедливо и тождественное неравенство $AC > BD$.

23. а) Если на ОДЗ трех алгебраических выражений A, B и C справедливо тождественное неравенство $A > B$, то на ОДЗ справедливо и тождественное неравенство $A + C > B + C$.

б) Если на ОДЗ трех алгебраических выражений A, B и C справедливо тождественное неравенство $A + C > B + C$, то на ОДЗ справедливо и тождественное неравенство $A > B$.

24. Пусть некоторая область M принадлежит ОДЗ трех алгебраических выражений A, B и C и обладает свойством: для любого числового набора из области M соответствующее числовое значение выражения C положительно, тогда:

а) если на области M справедливо тождественное неравенство $A > B$, то на области M справедливо и тождественное неравенство $AC > BC$;

б) если на области M справедливо тождественное неравенство $AC > BC$, то на области M справедливо тождественное неравенство $A > B$.

Поскольку метод доказательства утверждений 20—24 один и тот же, то приведем здесь лишь доказательство утверждения 24, а).

Возьмем некоторый числовой набор из области M. Обозначим соответствующие числовые значения выражений A, B и C соответственно через A_0, B_0 и C_0. По условию утверждения 24, а) число C_0 — положительное число, и из справедливости на области M тождественного неравенства $A > B$ вытекает справедливость числового неравенства $A_0 > B_0$. Поскольку C_0 — положительное число, то справедливо числовое неравенство $A_0C_0 > B_0C_0$. Таким образом, доказано, что для данного числового набора из области M соответствующее числовое значение выражения AC больше соответствующего числового значения выражения BC. Поскольку это рассуждение можно провести для любого числового набора из области M, то справедливость на области M тождественного неравенства $AC > BC$ доказана.

Если на некоторой области M из справедливости одного тождественного неравенства вытекает справедливость второго, а из справедливости второго вытекает справедливость первого, то говорят, что такие два тождественные неравенства равносильны на области M, а замену одного из них другим называют равносильным переходом от первого ко второму. Поэтому утверждения 23 и 24 можно сформулировать так.

23. На ОДЗ трех алгебраических выражений A, B и C равносильны тождественные неравенства $A > B$ и $A + C > B + C$.

24. Пусть некоторая область M принадлежит ОДЗ трех алгебраических выражений A, B и C и обладает свойством: для любого числового набора из области M соответствующее числовое значение выражения C положительно. Тогда на области M равносильны тождественные неравенства $A > B$ и $AC > BC$.

Приведем еще одно свойство тождественных неравенств, доказываемое аналогично предыдущим.

25. На ОДЗ двух алгебраических выражений A и B равносильны тождественные неравенства $A > B$ и $A - B > 0$.

Равносильный переход на области M от одного тождественного неравенства к другому обозначается иногда двойной стрелкой, над которой сверху написана буква M, т. е. запись $A > B \overset{M}{\Leftrightarrow} C > D$ означает, что на области M неравенства $A > B$ и $C > D$ равносильны. Поэтому утверждения 23, 24 и 25 можно записать так.

23. Пусть M — ОДЗ трех алгебраических выражений A, B и C, тогда $A > B \overset{M}{\Leftrightarrow} A+C > B+C$.

24. Пусть некоторая область M принадлежит ОДЗ трех алгебраических выражений A, B и C и обладает свойством: для любого числового набора из области M соответствующее числовое значение выражения C положительно. Тогда $A > B \overset{M}{\Leftrightarrow} AC > BC$.

25. Пусть M — ОДЗ двух алгебраических выражений A и B, тогда $A > B \overset{M}{\Leftrightarrow} A-B > 0$.

Докажем еще одно свойство тождественных равенств и неравенств алгебраических выражений.

26. Пусть некоторая область M принадлежит ОДЗ двух алгебраических выражений A и B и обладает свойством: для любого числового набора из области M соответствующие числовые значения выражений A и B положительны. Тогда на области M для любого натурального числа n:

а) тождественные равенства $A=B$ и $A^n=B^n$ равносильны;

б) тождественные неравенства $A>B$ и $A^n>B^n$ равносильны.

Доказательство. При $n=1$ утверждение 26 очевидно. Для доказательства при $n \geqslant 2$ воспользуемся справедливостью на ОДЗ двух выражений A и B тождественного равенства

$$A^n - B^n = (A-B)(A^{n-1} + A^{n-2}B + \ldots + AB^{n-2} + B^{n-1})$$

(доказательство этого равенства в § 5 гл. II). Обозначим алгебраическое выражение $A^{n-1} + A^{n-2}B + \ldots + AB^{n-2} + B^{n-1}$ через C. Тогда на области M справедливо тождественное равенство

$$A^n - B^n = (A-B)C. \qquad (1)$$

Очевидно, что для любого числового набора из области M соответствующее числовое значение выражения C положительно.

Докажем утверждение а). Пусть дано, что $A=B$ на области M. Тогда на основании утверждения 19 на M справедливо равенство $A-B=0$, а отсюда по утверждению 17 получим, что на M справедливо равенство $(A-B)C=0$. По свойству транзитивности равенств из этого равенства и равенства (1) вытекает, что $A^n-B^n=0$, откуда по утверждению 19 следует, что на области M справедливо равенство $A^n=B^n$.

Итак, доказано, что на области M из справедливости равенства $A=B$ следует справедливость равенства $A^n=B^n$. Пусть теперь дано, что $A^n=B^n$ на области M. Тогда на области M по утверждению 19 справедливо равенство $A^n - B^n=0$. Отсюда и из равенства (1) на области M на основании свойства транзитивности равенств имеем, что $(A-B)C = 0$, а отсюда по утверждению 17 вытекает, что $A-B=0$, и, наконец, по утверждению 19 получаем, что $A = B$. Значит, на области M из справедливости равенства $A^n = B^n$ вытекает справедливость равенства $A = B$. Утверждение а) доказано.

Докажем теперь утверждение б). Пусть дано, что на области M $A > B$. Тогда на основании утверждения 25 на M справедливо и неравенство $A - B > 0$, а отсюда по утверждению 24 получим справедливость неравенства $(A - B) C > 0$. Учитывая справедливость равенства (1), получим, что $A^n - B^n > 0$. Наконец, по утверждению 25 получим, что на M справедливо неравенство $A^n > B^n$.

Итак, доказано, что на области M из справедливости неравенства $A > B$ следует справедливость неравенства $A^n > B^n$. Пусть теперь дано, что $A^n > B^n$ на области M. Тогда на области M по утверждению 25 справедливо неравенство $A^n - B^n > 0$. Отсюда и из равенства (1) получим, что $(A - B) C > 0$ на области M. Применяя теперь утверждение 24, находим, что $A - B > 0$ на M. Наконец, по утверждению 25, имеем $A > B$ на M. Итак, на области M из справедливости неравенства $A^n > B^n$ вытекает справедливость неравенства $A > B$. Утверждение б) доказано.

Принято следующее соглашение: если не указана явно область M, на которой рассматривается тождественное неравенство $A > B$, то оно рассматривается на ОДЗ двух выражений A и B. Поэтому слова: дано, что $A > B$, — означают, что на ОДЗ двух выражений A и B справедливо тождественное неравенство $A > B$; слова: доказать, что $A > B$, — означают: доказать, что на ОДЗ двух выражений A и B справедливо тождественное неравенство $A > B$ (при этом имеется в виду, что эту ОДЗ обязательно следует отыскать). Если дано неравенство $A > B$, то ОДЗ двух выражений A и B называют часто ОДЗ неравенства $A > B$.

Все сказанное выше о неравенствах со знаком $>$ можно было бы повторить и для неравенств со знаком $\geqslant$, $\leqslant$ и $<$ (с естественными уточнениями, связанными с новыми знаками), но в связи с полной аналогией все это повторяться не будет.

Приведем теперь некоторые способы доказательства справедливости тождественных равенств и неравенств.

I. Доказательство с помощью использования законов действий над числами и алгебраическими выражениями

Докажем этим способом равенство

$$(a + b)(a^2 + b^2) = a^3 + a^2 b + ab^2 + b^3. \tag{2}$$

Отметим, что равенство (2) доказывается на ОДЗ трех выражений $(a + b)$, $(a^2 + b^2)$ и $a^3 + a^2 b + ab^2 + b^3$, т. е. на множестве всех действительных чисел a и b. На основании закона дистрибутивности действий над алгебраическими выражениями можно утверждать справедливость равенства

$$(a+b)(a^2+b^2)=a(a^2+b^2)+b(a^2+b^2). \tag{3}$$

На основании закона коммутативности умножения справедливы равенства

$$a(a^2+b^2)=(a^2+b^2)a, \tag{4}$$

$$b(a^2+b^2)=(a^2+b^2)b. \tag{5}$$

На основании закона дистрибутивности справедливы равенства

$$(a^2+b^2)a=a^2a+b^2a, \tag{6}$$

$$(a^2+b^2)b=a^2b+b^2b. \tag{7}$$

На основании законов коммутативности и ассоциативности умножения справедливы равенства

$$a^2a=a^3, \tag{8}$$

$$b^2a=ab^2, \tag{9}$$

$$a^2b=a^2b, \tag{10}$$

$$b^2b=b^3. \tag{11}$$

Вследствие того что равенства можно складывать (см. свойство 15 равенств), складывая равенства (8) и (9), а затем (10) и (11), получаем справедливость равенств

$$a^2a+b^2a=a^3+ab^2,$$

$$a^2b+b^2b=a^2b+b^3.$$

Складывая эти равенства, а затем равенства (6) и (7), получаем справедливость равенств

$$a^2a+b^2a+a^2b+b^2b=a^3+ab^2+a^2b+b^3, \tag{12}$$

$$(a^2+b^2)a+(a^2+b^2)b=a^2a+b^2a+a^2b+b^2b. \tag{13}$$

Складывая равенства (4) и (5), получаем справедливость равенства

$$a(a^2+b^2)+b(a^2+b^2)=(a^2+b^2)a+(a^2+b^2)b. \tag{14}$$

Применяя свойство транзитивности равенств, из справедливости равенств (3), (14), (13) и (12) получаем справедливость равенства (2).

Отметим, что обычно доказательство не расписывают так длинно, а все предшествующие выкладки записывают в виде следующей цепочки равенств:

$$(a+b)(a^2+b^2)=a(a^2+b^2)+b(a^2+b^2)=(a^2+b^2)a+(a^2+b^2)b=$$
$$=a^2a+b^2a+a^2b+b^2b=a^3+a^2b+ab^2+b^3.$$

Из справедливости этой цепочки равенств делается вывод о справедливости равенства (2). В дальнейшем при доказательстве способом I будем писать лишь цепочку очевидных равенств.

II. Доказательство перебором всех возможных случаев

Докажем этим методом свойства абсолютных величин действительных чисел:

1. $|a+b| \leqslant |a|+|b|$;
2. $||a|-|b|| < |a-b|$;
3. $|ab|=|a||b|$;
4. $|a/b|=|a|/|b|$.

ОДЗ каждого из утверждений 1—3 есть $\{a \in (-\infty, +\infty); b \in (-\infty, +\infty)\}$. ОДЗ утверждения 4: $\{a \in (-\infty, +\infty); b \in (-\infty, 0) \cup (0, +\infty)\}$. Каждое из этих утверждений надо доказывать на своей ОДЗ. Начнем, например, со свойства 3. ОДЗ разобьем на 4 области:

$$\alpha)\ \{a \geqslant 0;\ b \geqslant 0\};\quad \beta)\ \{a \geqslant 0;\ b \leqslant 0\};\quad \gamma)\ \{a \leqslant 0;\ b \geqslant 0\};\quad \delta)\ \{a \leqslant 0;\ b \leqslant 0\}$$

и доказательство проведем в каждом случае.

В случае α) по определению абсолютной величины $|a|=a$, $|b|=b$ и $ab \geqslant 0$, поэтому $|ab|=ab$. Значит, в случае α) равенство $|ab|=|a||b|$ может быть записано в виде $ab=ab$, после чего оно становится очевидным.

В случае β) $ab \leqslant 0$, поэтому по определению абсолютной величины $|a|=a$, $|b|=-b$, $|ab|=-ab$. Значит, в этом случае свойство 3 может быть записано в виде $-ab=a(-b)$ или $-ab=-ab$, после чего оно становится очевидным.

В случае γ) или δ) свойство 3 доказывается аналогично. Из справедливости свойства 3 во всех возможных случаях вытекает его справедливость в той формулировке, в которой оно записано.

Докажем теперь свойство 1. Рассмотрим следующие 6 случаев:

$$\begin{aligned}
&\alpha)\ a \geqslant 0; && b \geqslant 0; && a+b \geqslant 0;\\
&\beta)\ a \geqslant 0; && b \leqslant 0; && a+b \geqslant 0;\\
&\gamma)\ a \geqslant 0; && b \leqslant 0; && a+b \leqslant 0;\\
&\delta)\ a \leqslant 0; && b \geqslant 0; && a+b \geqslant 0;\\
&\lambda)\ a \leqslant 0; && b \geqslant 0; && a+b \leqslant 0;\\
&\nu)\ a \leqslant 0; && b \leqslant 0; && a+b \leqslant 0.
\end{aligned}$$

В случае α) $|a+b|=a+b=|a|+|b|$, поэтому свойство 1 в этом случае может быть записано в виде $|a+b|=|a|+|b|$, после чего оно становится очевидным.

В случае β) $|a+b|=a+b=a-(-b)=|a|-|b|$, поэтому свойство 1 в этом случае может быть записано в виде $|a|-|b| \leqslant |a|+|b|$, после чего оно становится очевидным.

В случае γ) $|a+b|=-(a+b)=(-b)-(a)=|b|-|a|$, поэтому свойство 1 в этом случае может быть записано в виде

$|b|-|a|\leqslant|a|+|b|$, после чего оно становится очевидным.

В случаях δ), λ) и ν) доказательства свойства 1 аналогичны предыдущим. Из справедливости свойства 1 во всех возможных случаях вытекает его справедливость в той формулировке, в которой оно записано. Свойства 2 и 4 абсолютных величин доказываются аналогично.

III. Доказательство методом от противного

Докажем этим методом, что для любого положительного числа a справедливо неравенство $a+1/a\geqslant 2$. Предположим противное, т. е. предположим, что существует хотя бы одно положительное число a, такое, что для него справедливо неравенство $a+1/a<2$. Так как a положительное число, то это неравенство на основании утверждения 24 равносильно неравенству $(a+1/a)\,a<2a$, т. е. неравенству $a^2+1<2a$, которое на основании утверждения 23 равносильно неравенству $(a^2+1)-2a<2a+(-2a)$, т. е. неравенству $a^2-2a+1<0$. Переписав последнее неравенство в виде $(a-1)^2<0$, приходим к противоречию с очевидным фактом, что квадрат любого действительного числа неотрицателен. Полученное противоречие говорит о том, что сделанное предположение неверно. Следовательно, неравенство $a+1/a\geqslant 2$ выполняется для любого положительного a.

IV. Прямое доказательство

Часто в процессе поиска доказательства, переходя от данного неравенства к следующим, приходят в конце к очевидному неравенству. Если при этом совершались только равносильные переходы, т. е. в результате перехода каждый раз получали неравенство, равносильное предыдущему, то тем самым получено доказательство исходного неравенства. Докажем этим способом следующее неравенство: $(a+b)/2\geqslant\sqrt{ab}$ на области $M=\{a\in(0,\ +\infty);\ b\in(0,\ +\infty)\}$. Напишем цепочку равносильных на области M переходов:

$$(a+b)/2\geqslant\sqrt{ab}\overset{M}{\Leftrightarrow}a+b-2\sqrt{ab}\geqslant 0\overset{M}{\Leftrightarrow}$$

$$\Leftrightarrow(\sqrt{a})^2-2\sqrt{a}\sqrt{b}+(\sqrt{b})^2\geqslant 0\overset{M}{\Leftrightarrow}(\sqrt{a}-\sqrt{b})^2\geqslant 0.$$

Поскольку справедливость последнего неравенства очевидна, то из равносильности первого и последнего неравенств вытекает справедливость первого неравенства.

Доказанное неравенство часто формулируют так: среднее арифметическое двух положительных чисел не меньше их среднего геометрического.

V. Доказательство с помощью свойства транзитивности неравенств

Пусть требуется доказать на области M неравенство $A<C$. Если известно или уже доказано, что на области M справедливы неравенства $A<B$, $B<C$, то по свойству транзитивности неравенств будет справедливо и исходное неравенство. Докажем этим способом следующее неравенство:

$$\frac{n(n-1)(n-2)\dots(n-k+2)(n-k+1)}{k!\,n^k}<\frac{1}{2^{k-1}}$$

на области $M=\{n$ и k — натуральные числа; $n\geqslant 3,\ 3\leqslant k\leqslant n\}$.

Для доказательства обозначим

$$A=\frac{n(n-1)(n-2)\dots(n-k+2)(n-k+1)}{k!\,n^k},$$

$C=1/2^{k-1}$ и покажем справедливость неравенства $A<C$. Разделив каждый член числителя выражения A на n, а знаменатель на n^k, получим равенство $A=\frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)\dots$ $\dots\left(1-\frac{k-2}{n}\right)\left(1-\frac{k-1}{n}\right)$. Каждая скобка в правой части этого равенства меньше единицы, поэтому справедливо неравенство $A<B$, где $B=\frac{1}{k!}$. Если в выражении $1\cdot 2\cdot 3\dots(k-1)k$ все члены, начиная с третьего, заменить на 2, то получим верное неравенство $\frac{1}{k!}<\frac{1}{2^{k-1}}$, т. е. $B<C$. Из свойства транзитивности неравенств и получаем теперь справедливость неравенства $A<C$, что и требовалось доказать.

VI. Доказательство методом математической индукции

Существует много равенств и неравенств, в которых участвует лишь одна буква n, пробегающая множество натуральных чисел. Для доказательства подобных утверждений обычно применяется метод математической индукции. В § 5 гл. II этот метод будет сформулирован и с его помощью будет доказан ряд равенств и неравенств, справедливых для всех натуральных чисел.

§ 3
МНОГОЧЛЕНЫ

В этом параграфе изучаются многочлены, в частности, для них конкретизируются законы действий, доказывается ряд тождественных равенств, называемых формулами сокращенного умножения, изучаются некоторые свойства многочленов.

Если в многочлен входит n букв, то его ОДЗ состоит из набора n множеств, каждое из которых состоит из всех действительных чисел. Поэтому обычно, рассматривая многочлен, не говорят о его ОДЗ, понимая, что этот многочлен имеет смысл для любого числового набора букв, входящих в этот многочлен. Обычно одночлены тождественно преобразуют по законам действий, приведенных в § 2, собирая вместе все числа, входящие в одночлен, и записывая их перед буквами одночлена, а также собирая вместе одинаковые буквы, входящие в одночлен, и записывая их в виде степени этой буквы. После такого преобразования одночлен считается записанным в стандартном виде, а числовой множитель, стоящий перед буквами одночлена, называется *коэффициентом данного одночлена.*

Например, одночлен $3abc2bc\frac{3}{7}ac$ преобразуется к стандартному виду $\frac{18}{7}a^2b^2c^3$ и число $\frac{18}{7}$ есть его коэффициент.

Согласно правилам действий над алгебраическими выражениями многочлен всегда можно тождественно преобразовать к виду, когда многочлен состоит из нескольких одночленов, записанных в стандартном виде и соединенных знаками сложения и вычитания; поэтому обычно говорят, что многочлен есть алгебраическая сумма одночленов.

Подобные члены многочлена — это одночлены его, записанные в стандартном виде и отличающиеся не более чем коэффициентами. Привести подобные члены многочлена — это значит заменить алгебраическую сумму подобных членов одним членом, тождественно равным этой сумме. Исходя из правил действий над алгебраическими выражениями можно следующим образом конкретизировать законы действий над многочленами.

Чтобы сложить два многочлена, следует записать подряд все члены первого многочлена, а затем все члены второго многочлена, сохраняя у каждого одночлена знак, стоящий перед его коэффициентом, после чего необходимо привести подобные члены.

Пример: $(2cd+5a)+(x+7a-4cd)=2cd+5a+x+7a-4cd=$ $=12a+x-2cd.$

Чтобы вычесть из одного многочлена другой многочлен, следует записать подряд все члены первого многочлена, сохраняя у каждого одночлена знак, стоящий перед его коэффициентом, затем все члены второго многочлена, изменив на противоположные знаки, стоящие перед коэффициентами одночленов второго многочлена, после чего необходимо привести подобные члены.

Пример: $(x^2-y^2)-(-7x^2+8y^2-5a)=x^2-y^2+7x^2-8y^2+$ $+5a=8x^2-9y^2+5a.$

Чтобы умножить одночлен на многочлен, следует умножить этот одночлен на каждый член многочлена, записать члены произведения подряд с теми знаками, что были у членов многочлена, если перед коэффициентом одночлена стоит знак плюс, и с противоположными знаками — если перед коэффициентом одночлена стоит знак минус, каждый одночлен произведения записать в стандартном виде, а затем привести подобные члены.

Чтобы умножить многочлен на многочлен, следует каждый одночлен (вместе со знаком, стоящим перед его коэффициентом) первого многочлена умножить на второй многочлен, записать подряд все произведения, каждый полученный одночлен записать в стандартной форме, а затем привести подобные члены.

Пример: $(ab-cd)(ab+cd)=(ab)(ab)+(ab)(cd)-(cd)(ab)-(cd)(cd)=a^2b^2+abcd-abcd-c^2d^2=a^2b^2-c^2d^2$.

Пользуясь этими законами, выведем формулы сокращенного умножения многочленов.

Начнем с перемножения одинаковых многочленов вида $(a+b)$. Используя законы действий над алгебраическими выражениями, можно написать следующую цепочку тождественных равенств:

$$(a+b)^2=(a+b)(a+b)=(a)(a)+(b)(a)+(a)(b)+(b)(b)=$$
$$=a^2+ab+ab+b^2=a^2+2ab+b^2.$$

Эта формула имеет следующую словесную формулировку: квадрат суммы двух чисел равен квадрату первого числа, плюс удвоенное произведение первого числа на второе, плюс квадрат второго числа. Теперь, используя предыдущую формулу, можно написать следующую цепочку тождественных равенств:

$$(a+b)^3=(a+b)(a+b)^2=(a+b)(a^2+2ab+b^2)=$$
$$=(a)(a^2)+(a)(2ab)+(a)(b^2)+(b)(a^2)+(b)(2ab)+(b)(b^2)=$$
$$=a^3+2a^2b+ab^2+a^2b+2ab^2+b^3=a^3+3a^2b+3ab^2+b^3.$$

Эта формула имеет следующую словесную формулировку: куб суммы двух чисел равен кубу первого числа, плюс утроенное произведение квадрата первого числа на второе, плюс утроенное произведение первого числа на квадрат второго числа и плюс куб второго числа.

Напишем еще одну цепочку тождественных равенств:

$$(a+b)^4=(a+b)(a+b)^3=(a+b)(a^3+3a^2b+3ab^2+b^3)=$$
$$=(a)(a^3)+(a)(3a^2b)+(a)(3ab^2)+(a)(b^3)+(b)(a^3)+$$
$$+(b)(3a^2b)+(b)(3ab^2)+(b)(b^3)=$$
$$=a^4+3a^3b+3a^2b^2+ab^3+ba^3+3a^2b^2+3ab^3+b^4=$$
$$=a^4+4a^3b+6a^2b^2+4ab^3+b^4.$$

Приведенные формулы позволяют заметить некоторую закономерность, с помощью которой можно написать формулу для $(a+b)^n$, где n — любое натуральное число. А именно легко заметить, что всех членов будет $(n+1)$, первый член есть первое число в степени n, в каждом последующем члене степень первого числа на единицу меньше его степени в предшествующем члене, а в последнем члене оно в нулевой степени; второе число находится в первом члене в нулевой степени, во втором члене в первой степени, а в каждом последующем члене степень второго числа на единицу больше его степени в предшествующем члене, а в последнем члене второе число в степени n.

Коэффициент же при каждом члене можно найти при помощи треугольника Паскаля:

$$
\begin{array}{l ccccccccccccccccccc}
0 & & & & & & & & & & 1 & & & & & & & & & \\
1 & & & & & & & & & 1 & & 1 & & & & & & & & \\
2 & & & & & & & & 1 & & 2 & & 1 & & & & & & & \\
3 & & & & & & & 1 & & 3 & & 3 & & 1 & & & & & & \\
4 & & & & & & 1 & & 4 & & 6 & & 4 & & 1 & & & & & \\
5 & & & & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & & \\
6 & & & & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & & & \\
7 & & & 1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1 & & \\
8 & & 1 & & 8 & & 28 & & 56 & & 70 & & 56 & & 28 & & 8 & & 1 & \\
9 & 1 & & 9 & & 36 & & 84 & & 126 & & 126 & & 84 & & 36 & & 9 & & 1
\end{array}
$$

Правило образования строк треугольника Паскаля простое. Каждая строка может быть получена из предыдущей верхней строки следующим образом. В промежутке между любыми соседними числами верхней строки (но под ними) пишется сумма этих чисел, а по краям пишутся единицы. Номер строки показывает, в какую степень возводится двучлен $(a+b)$, а числа этой строки являются коэффициентами соответствующих членов, записанных в рассмотренном выше порядке.

Конечно, если надо написать формулу для $(a+b)^n$, где n — большое число (например, $n=100$), то ясно, что по треугольнику Паскаля вычислять коэффициенты правой части долго. Поэтому желательно знать общую формулу для вычисления $(a+b)^n$. Эта формула носит название формулы бинома Ньютона и будет доказана в § 5 гл. II методом математической индукции. Здесь она приводится без доказательства:

$$(a+b)^n = C_1^0 a^n + C_n^1 a_1^{n-1} b + \ldots + C_n^k a^{n-k} b^k + \ldots + C_n^n b^n,$$

где $C_n^k = \frac{n!}{k!\,(n-k)!}$, $0! = 1$, $k! = 1 \cdot 2 \ldots k$ для любого $k = 1, 2, \ldots$ $\ldots n$. Из формулы бинома Ньютона легко получить формулу для $(a-b)^n$. Обозначим $d = -b$ и применим формулу бинома Ньютона:

$$(a-b)^n = (a+d)^n = C_1^0 a^n + C_n^1 a^{n-1} d + \ldots + C_n^k a^{n-k} d^k + \ldots + C_n^n d^n.$$

Подставляя $(-b)$ вместо d, получим

$$(a-b)^n = C_n^0 a^n - C_n^1 a^{n-1} b + \ldots + (-1)^k C_n^k a^{n-k} b^k + \ldots + (-1)^n C_n^n b^n.$$

Частные случаи этой формулы для $n=2$ и $n=3$:

$$(a-b)^2=a^2-2ab+b^2,$$
$$(a-b)^3=a^3-3a^2b+3ab^2-b^3.$$

Докажем некоторые формулы, в которых берется произведение разных многочленов. Очевидна следующая цепочка тождественных равенств:

$$(a-b)(a+b)=(a)(a)+(a)(b)-(b)(a)-(b)(b)=$$
$$=a^2+ab-ab-b^2=a^2-b^2.$$

Эта формула обычно запоминается в записи, где меняются местами правая и левая части: $a^2-b^2=(a-b)(a+b)$.

Приведем ее словесную формулировку: разность квадратов двух чисел равна произведению разности этих чисел на их сумму.

Выведем формулу разности кубов двух чисел a^3-b^3. Поскольку $(a-b)(a^2+ab+b^2)=(a)(a^2)+(a)(ab)+(a)(b^2)-(b)\times\times(a^2)-(b)(ab)-(b)(b^2)=a^3+a^2b+ab^2-a^2b-ab^2-b^3=a^3-b^3$, то $a^3-b^3=(a-b)(a^2+ab+b^2)$.
Приведем словесную формулировку этой формулы: разность кубов двух чисел равна произведению разности этих чисел на неполный квадрат суммы этих чисел.

Приведенные формулы позволяют заметить некоторую закономерность, с помощью которой легко записать формулу a^n-b^n для любого натурального числа n. Эта формула будет иметь вид

$$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+a^{n-3}b^2+\ldots+ab^{n-2}+b^{n-1}).$$

Доказательство этой формулы будет приведено в § 5 гл. II методом математической индукции. Наконец, выведем следующую формулу: $a^3+b^3=(a+b)(a^2-ab+b^2)$.
Действительно, $(a+b)(a^2-ab+b^2)=(a)(a^2)-(a)(ab)+(a)(b^2)+ + (b)(a^2)-(b)(ab)+(b)(b^2) = a^3 - a^2b + ab^2 + a^2b-ab^2+b^3 = =a^3+b^3$.
Словесная формулировка этой формулы следующая: сумма кубов двух чисел равна произведению суммы этих чисел на неполный квадрат разности этих чисел.
Приведем формулы, которые желательно запомнить:

$$(a+b)^2=a^2+2ab+b^2,$$
$$(a+b)^3=a^3+3a^2b+3ab^2+b^3,$$
$$a^2-b^2=(a-b)(a+b),$$
$$a^3-b^3=(a-b)(a^2+ab+b^2),$$
$$a^3+b^3=(a+b)(a^2-ab+b^2),$$
$$(a-b)^2=a^2-2ab+b^2,$$
$$(a-b)^3=a^3-3a^2b+3ab^2-b^3.$$

Формулы, доказанные в этом параграфе, справедливы для любых числовых значений букв a и b. Иногда эти формулы употребляются и тогда, когда буквами a и b обозначены некоторые алгебраические выражения, но тогда очевидно, что эти формулы будут справедливы уже на ОДЗ двух алгебраических выражений a и b.

В ряде вопросов при действиях с многочленами удобнее рассматривать их не в обычном виде, а в виде произведения.

Тождественное преобразование многочлена к виду произведения многочленов называется разложением многочлена на множители. Собственно говоря, все формулы сокращенного умножения и есть формулы разложения многочлена на множители.

Кроме применения формул сокращенного умножения есть и другие приемы для разложения многочлена на множители, например вынесение за скобки общего множителя, группировка. Для разложения многочлена на множители употребляются все приемы.

Рассмотрим пример разложения многочлена на множители. Группируя, вынося за скобки общий множитель и пользуясь формулой сокращенного умножения, получаем цепочку тождественных равенств:

$$a^2c+2abc+b^2c+(a+b)^2 d=c(a^2+2ab+b^2)+d(a+b)^2=$$
$$=c(a+b)^2+d(a+b)^2=(a+b)^2(c+d).$$

Если в некотором алгебраическом выражении можно выделить несколько букв, над которыми производятся только действия умножения, сложения и вычитания, а над остальными буквами этого выражения производятся любые действия, включая, например, деление, извлечение корней, логарифмирование, то данное алгебраическое выражение можно рассматривать как многочлен относительно выделенных букв. Более точно это можно сказать так.

Многочленом, целым относительно данных букв x, y, z, ..., t, называется алгебраическое выражение, целое относительно этих букв (относительно других букв оно может быть и нецелым), т. е. алгебраическое выражение, в котором эти буквы связаны только знаками умножения, сложения и вычитания.

Например, следующие алгебраические выражения являются целыми многочленами относительно букв x, y, z и t:

$$x^2y+3xz^2+t^2-t^6z^2, \quad \frac{a}{b^2}x^3y+t\sqrt{ab}\,z,$$
$$\frac{xy}{a}+3x^2t+2z^2t\log_3 b+\frac{2}{\sqrt{a+c}}xyz.$$

Одночлены принято записывать в виде произведения целых степеней (т. е. если некоторая буква является множителем

n раз, то вместо этого пишут эту букву в степени n): $2axyzxyzxzxxz = 2ax^5y^2z^4$.

После того как выделены буквы, относительно которых алгебраическое выражение является многочленом, для каждого одночлена такого многочлена все остальные буквы и числа этого одночлена называют коэффициентом. Например, если выражение $3x^2y \log_a b + \frac{a}{b^2} x^4y$ есть многочлен относительно x и y, то у одночлена $3x^2y \log_a b$ коэффициентом будет $3 \log_a b$, а у одночлена $\frac{a}{b^2} x^4y$ коэффициентом будет $\frac{a}{b^2}$.

Степенью одночлена относительно данных букв x, y, z, ..., t называется сумма показателей степеней всех данных букв, входящих в этот одночлен.

Пример. Одночлен $\frac{x^6z^2}{b}$ имеет 8-ю степень относительно x и z, одночлен $2ax^4y^2z^3$ имеет 9-ю степень относительно x, y, z.

Степенью многочлена целого относительно данных букв x, y, z, ..., t называется наибольшая относительно данных букв из степеней всех одночленов, входящих в этот многочлен.

Многочлен, целый относительно одной буквы, обычно относительно буквы x, имеет одночлены разных степеней, ибо в противном случае можно привести подобные члены. Одночлены разных степеней можно упорядочить относительно возрастания или убывания степени буквы x. Обычно многочлен, целый относительно одной буквы, записывают в порядке убывания степеней. Многочлен, записанный в виде $a_0x^n + a_1x^{n-1} + a_2x^{n-2} + + \ldots + a_{n-1}x + a_n$, называется расположенным многочленом. Если $a_0 \neq 0$, то говорят, что этот многочлен имеет степень n. Если неизвестно, равен или не равен нулю коэффициент a_0, то говорят, что этот многочлен степени не выше чем n. Пользуясь определением тождественного равенства двух алгебраических выражений, можно доказать теорему.

Теорема 1. Два многочлена, целые относительно x, тождественно равны тогда и только тогда, когда равны их степени и равны коэффициенты при равных степенях x.

Д о к а з а т е л ь с т в о этой теоремы опускается.

Эта теорема может быть применена для разложения многочлена на множители. Воспользуемся так называемым методом неопределенных коэффициентов. Суть этого метода состоит в следующем. Пусть дан многочлен $P_n(x)$ степени n и его надо представить в виде произведения многочленов степеней k ($k < n$) и $(n-k)$. Тогда выписываются два многочлена $P_k(x)$ и $P_{n-k}(x)$; первый — степени k и второй — степени $(n-k)$ с коэффициентами, обозначенными некоторыми буквами, скажем, у первого α_0, α_1, ..., α_k, у второго β_0, β_1, ..., β_{n-k}. Перемножая многочлены $P_k(x)$ и $P_{n-k}(x)$, получаем многочлен $T_n(x)$ степени n с коэффициентами, зависящими от α_i $(i = 0, 1, \ldots, k)$ и β_j $(j = 0, 1, \ldots, n-k)$. Из условия, что многочлены $P_n(x)$ и $T_n(x)$

тождественно равны, получаем $(n+1)$ равенство, в которых участвуют $(n+2)$ коэффициента α_i и β_j, которые надо найти. Полагая, например, коэффициент $\alpha_0=1$, приходим к $(n+1)$ равенству, из которых надо найти $(n+1)$ коэффициент α_i ($i=1, 2, \ldots, k$) и β_j ($j=0, 1, \ldots, n-k$). Найдя их, найдем и многочлены $P_k(x)$ и $P_{n-k}(x)$.

Пример. Разложить многочлен x^3+3x+4 на множители, среди которых один — многочлен первой степени, а второй многочлен — второй степени. Будем искать многочлены $(x+\alpha_1)$ и $(\beta_0x^2+\beta_1x+\beta_2)$, такие, что справедливо тождественное равенство $(x+\alpha_1)(\beta_0x^2+\beta_1x+\beta_2)=x^3+3x+4$. Применяя теорему 1, получаем 4 равенства: $\beta_0=1$, $\beta_0\alpha_1+\beta_1=0$, $\beta_1\alpha_1+\beta_2=3$, $\alpha_1\beta_2=4$. Этим равенствам удовлетворяют $\beta_0=1$, $\alpha_1=1$, $\beta_1=-1$, $\beta_2=4$. Значит, многочлен x^3+3x+4 разлагается на множители $(x+1)$ и (x^2-x+4), т. е. $x^3+3x+4=(x+1)\times(x^2-x+4)$. Заметим, что не всякий многочлен можно разложить на множители. Например, многочлен x^2+x+1 нельзя разложить на произведение двух многочленов первой степени.

§ 4
АЛГЕБРАИЧЕСКИЕ ДРОБИ

Напомним (см. § 1), что алгебраической дробью $\frac{A}{B}$ называется алгебраическое выражение — частное от деления многочлена A на многочлен B. Для нахождения ОДЗ алгебраической дроби $\frac{A}{B}$, в которую входит n букв, надо взять сначала набор из n множеств, каждое из которых состоит из множества всех действительных чисел, и исключить из этого набора те числовые наборы, для каждого из которых соответствующее числовое значение многочлена B равно нулю.

Докажем ряд свойств алгебраических дробей.

1. Если обозначить алгебраическую дробь $\frac{A}{B}$ одной буквой C, то на ОДЗ этой дроби равносильны тождественные равенства $C=\frac{A}{B}$ и $A=CB$.

Справедливость этого свойства вытекает из справедливости утверждения 18 § 2.

2. Равенства $\frac{A}{B}=\frac{C}{D}$ и $AD=BC$ равносильны на ОДЗ первого из них.

Это свойство часто формулируют так: две дроби $\frac{A}{B}$ и $\frac{C}{D}$ тождественно равны на ОДЗ тогда и только тогда, когда на этой ОДЗ справедливо равенство $AD=BC$.

Доказательство. Пусть область M — ОДЗ двух дробей $\frac{A}{B}$ и $\frac{C}{D}$. Рассмотрим случай, когда $A=0$ на M. Тогда $\frac{A}{B}=0$ и из равенства $\frac{A}{B}=\frac{C}{D}$ следует, что и $\frac{C}{D}=0$ на M. Поэтому $C=0$ на M, а это значит, что $AD=BC$ на M. Обратно, пусть $AD=BC$ и $A=0$ на M. Так как на M $D\neq 0$ и $B\neq 0$, то $C=0$ на M. Следовательно, $\frac{A}{B}=\frac{C}{D}$.

Рассмотрим теперь случай, когда ни для одного набора из области M многочлен A не обращается в нуль, т. е. рассмотрим случай, когда $A\neq 0$ на M. Пусть $\frac{A}{B}=\frac{C}{D}$, тогда отсюда следует, что и $C\neq 0$ на M. Обозначим $\frac{A}{B}$ через α и $\frac{C}{D}$ через β. По свойству 1 алгебраических дробей $A=\alpha B$ и $C=\beta D$. По утверждению 16 § 2 имеем

$$A\beta D=C\alpha B. \tag{1}$$

Так как $\alpha=\beta\neq 0$ на M, то по утверждению 18 § 2 из (1) следует, что $AD=CB$.

Обратно, пусть $AD=BC$, тогда, так как $A\neq 0$, $B\neq 0$ и $D\neq 0$ на M, то и $C\neq 0$ на M. Следовательно, $\alpha=\frac{A}{B}$ и $\beta=\frac{C}{D}$ не равны нулю на M. Тогда умножим данное равенство $AB=BC$ на $\alpha\beta$. Получим равносильное равенство

$$\alpha\beta AD=\alpha\beta BC. \tag{2}$$

Но $\alpha B=A$, $\beta D=C$, и равенство (2) примет вид

$$\alpha AC=\beta AC. \tag{3}$$

Используя утверждение 18 § 2, получим $\alpha=\beta$, что и требовалось доказать.

Таким образом, свойство 2 алгебраических дробей доказано.

3. На ОДЗ алгебраической дроби $\frac{A}{B}$ справедливы тождественные равенства $\frac{A}{B}=\frac{-A}{-B}=-\frac{A}{-B}=-\frac{-A}{B}$.

Каждое из этих равенств становится очевидным, если воспользоваться только что доказанным свойством 2.

4. Для любого многочлена K, не обращающегося в нуль на ОДЗ алгебраической дроби $\frac{A}{B}$, справедливо тождественное равенство $\frac{A}{B}=\frac{AK}{BK}$.

Поскольку это равенство по свойству 2 равносильно на ОДЗ дроби A/B равенству $A(BK)=B(AK)$, которое является очевидным, то столь же очевидна и справедливость свойства 4.

5. На ОДЗ $\frac{A}{B}$ справедливо тождественное равенство $\frac{A}{B}=A\frac{1}{B}$. Это свойство уже доказано в § 2.

6. На ОДЗ алгебраической дроби $\frac{1}{AB}$ справедливо тождественное равенство $\frac{1}{AB}=\frac{1}{A}\frac{1}{B}$.

Действительно, на ОДЗ дроби $\frac{1}{AB}$ очевидна справедливость цепочки тождественных равенств

$$\frac{1}{AB}=\frac{1}{AB}\left(A\frac{1}{A}\right)\left(B\frac{1}{B}\right)=\left(\frac{1}{AB}AB\right)\left(\frac{1}{A}\frac{1}{B}\right)=\frac{1}{A}\frac{1}{B}.$$

7. На ОДЗ двух алгебраических дробей $\frac{A}{B}$ и $\frac{B}{A}$ справедливо тождественное равенство $\frac{A}{B}=\frac{1}{\frac{B}{A}}$.

Действительно, применяя сначала свойство 5 дробей, затем свойства действий над алгебраическими выражениями, затем свойства 6 и 5 дробей, имеем цепочку тождественных равенств

$$\frac{A}{B}=A\frac{1}{B}=A\frac{1}{B}\left(\frac{1}{A}\frac{1}{\frac{1}{A}}\right)=\left(A\frac{1}{A}\right)\left(\frac{1}{B}\frac{1}{\frac{1}{A}}\right)=\frac{1}{B\frac{1}{A}}=\frac{1}{\frac{B}{A}}.$$

Напомним следующее соглашение: если не указана явно область M, на которой рассматривается некоторое тождественное равенство, то оно рассматривается на ОДЗ двух выражений, стоящих в левой и правой частях равенства. Поэтому дальше не будет явно выписываться область, на которой будет справедливо тождественное равенство, имея в виду, что оно справедливо на ОДЗ двух выражений, стоящих в левой и правой частях равенства.

Пользуясь свойствами сложения и умножения алгебраических выражений и свойствами алгебраических дробей, легко показать справедливость тождественных равенств:

$$\frac{A}{B}+\frac{C}{D}=\frac{AD+BC}{BD};\quad \frac{A}{B}\frac{C}{D}=\frac{AC}{BD}.$$

Действительно, используя свойства алгебраических дробей, получим

$$\frac{A}{B}+\frac{C}{D}=\frac{AD}{BD}+\frac{CB}{DB}=AD\frac{1}{BD}+CB\frac{1}{BD}.$$

Применяя теперь свойства сложения и умножения алгебраических выражений, а затем опять свойства алгебраических дробей, имеем

$$AD\frac{1}{BD}+CB\frac{1}{BD}=\frac{1}{BD}(AD+CB)=\frac{AD+CB}{BD},$$

что и требовалось доказать.

Аналогично доказывается второе равенство:

$$\frac{A}{B}\,\frac{C}{D}=\left(A\,\frac{1}{B}\right)\left(C\,\frac{1}{D}\right)=AC\,\frac{1}{B}\,\frac{1}{D}=AC\,\frac{1}{BD}=\frac{AC}{BD}.$$

Так же доказываются и равенства

$$\frac{A}{B}-\frac{C}{D}=\frac{AD-BC}{BD};\quad \frac{A}{B}:\frac{C}{D}=\frac{AD}{BC}.$$

Часто надо привести алгебраические дроби к общему знаменателю, т. е. надо записать их в виде, где у всех этих дробей один и тот же знаменатель. Для этого существует следующий способ: надо разложить каждый знаменатель на множители, а затем числитель и знаменатель каждой дроби умножить на произведение тех множителей знаменателей остальных дробей, которые не содержатся в данном знаменателе, что по свойству дробей их не изменит.

Пример. Привести к общему знаменателю следующие алгебраические дроби: $\frac{a}{a^3-b^3}$; $\frac{c}{a^2-b^2}$; $\frac{d}{a^2+ab+b^2}$.

Разлагая знаменатели на множители, перепишем дроби так:

$$\frac{a}{(a-b)(a^2+ab+b^2)};\quad \frac{c}{(a-b)(a+b)};\quad \frac{d}{a^2+ab+b^2}.$$

Теперь умножая числитель и знаменатель первой дроби на $(a+b)$, второй — на a^2+ab+b^2, третий — на $(a-b)(a+b)$, получаем

$$\frac{a(a+b)}{(a-b)(a+b)(a^2+ab+b^2)};\quad \frac{c(a^2+ab+b^2)}{(a-b)(a+b)(a^2+ab+b^2)};$$

$$\frac{d(a-b)(a+b)}{(a-b)(a+b)(a^2+ab+b^2)}.$$

У этих дробей одинаковые знаменатели, т. е. первоначальные дроби приведены к общему знаменателю.

В ряде случаев требуется представить дробь в виде суммы дробей с более простыми знаменателями. Это можно сделать только в том случае, когда многочлен, стоящий в знаменателе дроби, разлагается на произведение многочленов. Покажем на примере, как это делается.

Пусть надо разложить алгебраическую дробь $\frac{1}{x^2-1}$ на простейшие. Так как многочлен x^2-1 разлагается на произведение многочленов $x-1$ и $x+1$, то это можно сделать. Для этого нужно найти алгебраические дроби $\frac{A}{x-1}$ и $\frac{B}{x+1}$, такие, чтобы было выполнено тождественное равенство $\frac{1}{x^2-1}=\frac{A}{x-1}+$

$+\frac{B}{x+1}$. Рассмотрим сумму $\frac{A}{x-1}+\frac{B}{x+1}$. По только что сформулированным правилам

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

Так как эта дробь должна тождественно равняться дроби $\frac{1}{x^2-1}$ (заметим, что эти рассуждения проводятся для любого x, кроме $x=1$ и $x=-1$), то, применяя свойство 2 дробей, получим, что эти две дроби равны только тогда, когда $[(A+B)x+ +(A-B)](x-1)(x+1)=(x-1)(x+1)$. Так как это равенство должно выполняться для любого x, кроме $x=1$ и $x-1$, то, полагая, например, $x=0$, затем $x=2$, получаем, что это будет верно только тогда, когда одновременно $A-B=1$ и $3A+B=1$. А эти два равенства справедливы одновременно только для $A=1/2$ и $B=-1/2$. Значит, данная дробь разложена на простейшие, а именно справедливо следующее тождественное равенство:

$$\frac{1}{x^2-1}=\frac{1/2}{x-1}+\frac{-1/2}{x+1}.$$

Этот способ разложения дроби в сумму более простых дробей называется способом неопределенных коэффициентов. Действительно, полагая числа A и B вначале неизвестными, получаем на ОДЗ равенство двух многочленов, один из которых с известными коэффициентами, другой с неизвестными, выраженными через A и B. Это дает возможность выписать алгебраические равенства относительно неизвестных коэффициентов (в данном случае $A-B=1$, $3A+B=1$). Найдя численные значения неизвестных коэффициентов, обращающие данные алгебраические равенства в верные числовые равенства, тем самым решим поставленную задачу о представлении дроби в виде суммы более простых дробей.

§ 5
МЕТОД МАТЕМАТИЧЕСКОЙ ИНДУКЦИИ

Существует очень много утверждений, зависящих от натурального числа n. Как понимать такие утверждения?

Поскольку натуральных чисел бесконечно много, то на самом деле каждое такое утверждение содержит в себе бесконечно много утверждений. Например, утверждение — сумма n первых натуральных чисел равна $n(n+1)/2$ — содержит в себе следующие утверждения:

для $n=1$: первое натуральное число, т. е. число единица, равно $1(1+1)/2$;

для $n=2$: сумма двух первых натуральных чисел, т. е. сумма чисел единица и два, равна $2(2+1)/2$;
для $n=3$: сумма трех первых натуральных чисел, т. е. сумма чисел единица, два и три, равна $3(3+1)/2$;
. . . .
для $n=10\,000$: сумма десяти тысяч первых натуральных чисел равна $10\,000(10\,000+1)/2$ и т. д., т. е. рассматриваемое утверждение действительно содержит бесконечно много утверждений.

Аналогично и любое другое утверждение, зависящее от натурального числа n, на самом деле есть простая форма записи бесконечного числа утверждений.

Естествен вопрос, а как убедиться в справедливости утверждения, зависящего от натурального числа?

Для доказательства утверждений, зависящих от натурального числа n, применяется общий метод доказательства — метод полной математической индукции. Этот метод основан на аксиомах натуральных чисел. Но поскольку ранее эти аксиомы не приводились, то метод полной математической индукции принимается здесь без доказательства.

Для доказательства некоторого утверждения, зависящего от натурального числа n, делается следующее:

1. Проверяется справедливость этого утверждения для $n=1$.

2. Предполагается справедливость этого утверждения для $n=k$.

3. Доказывается справедливость этого утверждения для $n=k+1$, используя предполагаемую справедливость его для $n=k$. После чего делается вывод, что утверждение справедливо для любого натурального числа n.

Пользуясь этим методом, докажем справедливость ранее сформулированного утверждения, т. е. докажем, что для любого натурального числа n справедливо равенство

$$1+2+3+\ldots+n=n(n+1)/2. \tag{1}$$

Проверяем справедливость равенства (1) для $n=1$.
Для $n=1$ оно запишется так: $1=1(1+1)/2$ и очевидно, что это равенство верное. Предположим, что равенство (1) справедливо для $n=k$, т. е. предположим, что справедливо равенство

$$1+2+3+\ldots+k=k(k+1)/2. \tag{2}$$

Используя равенство (2), докажем, что равенство (1) справедливо для $n=k+1$, т. е. докажем справедливость равенства

$$1+2+3+\ldots+(k+1)=(k+1)[(k+1)+1]/2. \tag{3}$$

Действительно, рассмотрим сумму $1+2+3+\ldots+k+(k+1)$. Используя сначала свойство ассоциативности сложения, затем равенство (2) и делая простейшие преобразования, получаем

$$1+2+3+\ldots+k+(k+1)=(1+2+3+\ldots+k)+(k+1)=$$
$$=k(k+1)/2+(k+1)=(k+1)(k+1)12=(k+1)[(k+1)+1]/2,$$

т. е. получаем справедливость равенства (3). На основании метода полной математической индукции делаем вывод, что равенство (1) справедливо для любого натурального числа n.

Докажем методом полной математической индукции еще несколько утверждений.

1. Для любого натурального числа n справедливо неравенство

$$n \leqslant 2^{n-1}. \tag{4}$$

Доказательство. Для $n=1$ неравенство (4) превращается в верное числовое неравенство $1 \leqslant 2^{1-1}$. Предположим, что неравенство (4) справедливо для $n=k$, т. е. предположим справедливость неравенства

$$k \leqslant 2^{k-1}. \tag{5}$$

Используя неравенство (5), докажем справедливость неравенства (4) для $n=k+1$, т. е. докажем справедливость неравенства

$$k+1 \leqslant 2^{(k+1)-1} \tag{6}$$

Действительно, очевидно, что $k+1 \leqslant 2k$. Откуда, используя неравенство (5) и свойство транзитивности неравенств, $k+1 \leqslant \leqslant 2 \cdot 2^{k-1}$. Правая часть последнего неравенства может быть записана в виде $2^{(k+1)-1}$, откуда следует справедливость неравенства (6). На основании метода полной математической индукции делаем вывод, что неравенство (4) справедливо для любого натурального числа n.

2. Для любого натурального числа n число $N(n)=n^3+5n$ делится на 6.

Доказательство. Для $n=1$ число $N(1)=6$ делится на 6, т. е. утверждение справедливо. Предположим, что утверждение справедливо для $n=k$, т. е. предположим, что число $N(k)=(k^3+5k)$ делится на 6. Используя, что число $N(k)$ делится на 6, докажем справедливость утверждения для $n= =k+1$, т. е. докажем, что число $N(k+1)=[(k+1)^3+5(k+1)]$ делится на 6.

Действительно, используя формулы сокращенного умножения и свойство ассоциативности и коммутативности действий над числами и алгебраическими выражениями, имеем

$$N(k+1)=[(k+1)^3+5(k+1)]=k^3+3k^2+3k+1+5k+5=$$
$$=(k^3+5k)+6+3k^2+3k=N(k)+6+3k(k+1).$$

Поскольку числа k и $k+1$ есть два рядом стоящих натуральных числа, то одно из них четное, поэтому число $3k(k+1)$ делится на 6. Учитывая, что число $N(k)$ делится на 6 и число 6 делится на 6, получаем, что число $N(k+1)$ также делится на 6. На основании метода полной математической индукции делаем вывод, что число $N(n)=n^3+5n$ делится на 6 для любого натурального числа n.

3. Общий член арифметической прогрессии вычисляется по формуле

$$a_n = a_1 + (n-1)d. \qquad (7)$$

Доказательство. Для $n=1$ формула (7) запишется в виде $a_1=a_1+0\cdot d$, т. е. оказывается справедливой. Предположим, что формула (7) справедлива для $n=k$, т. е. предположим что k-й член арифметической прогрессии вычисляется по формуле

$$a_k = a_1 + (k-1)d. \qquad (8)$$

Докажем, что формула (7) справедлива для $n=k+1$, т. е. докажем справедливость формулы

$$a_{k+1} = a_1 + [(k+1)-1]d. \qquad (9)$$

Действительно, по определению арифметической прогрессии $a_{k+1}=a_k+d$. Следовательно, используя формулу (8), можно написать, что $a_{k+1}=a_1+(k-1)d+d=a_1+[(k+1)-1]d$, т. е. получить справедливость формулы (9). На основании метода полной математической индукции делаем вывод, что формула (7) справедлива для любого натурального числа n.

4. Общий член геометрической прогрессии вычисляется по формуле

$$b_n = b_1 q^{n-1}. \qquad (10)$$

Доказательство. Для $n=1$ формула (10) запишется в виде $b_1=b_1q^0$, т. е. оказывается справедливой. Предположим, что формула (10) справедлива для $n=k$, т. е. предположим, что справедлива формула

$$b_k = b_1 q^{k-1}. \qquad (11)$$

Докажем, что формула (10) справедлива для $n=k+1$, т. е. докажем справедливость формулы

$$b_{k+1} = b_1 q^{(k+1)-1}. \qquad (12)$$

Действительно, по определению геометрической прогрессии $b_{k+1}=b_kq$. Следовательно, используя формулу (11), можно написать, что

$$b_{k+1} = (b_1 q^{k-1})\, q = b_1 q^{(k+1)-1},$$

т. е. получить справедливость формулы (12). На основании метода полной математической индукции делаем вывод, что формула (10) справедлива для любого натурального числа n.

5. Для любого натурального числа n справедлива формула бинома Ньютона:

$$(a+b)^n = C_n^0 a^n + C_n^1 a^{n-1} b + C_n^2 a^{n-2} b^2 + \\ + \ldots + C_n^m a^{n-m} b^m + \ldots + C_n^n b^n, \qquad (13)$$

где числа C_n^m — коэффициенты бинома Ньютона (биномиальные коэффициенты) вычисляются по правилу $C_n^m = \dfrac{n!}{m!\,(n-m)!}$ $(0 \leqslant \leqslant m \leqslant n)$.

Доказательство. Для $n=1$ формула (13) запишется в виде $(a+b)^1 = C_1^0 a^1 + C_1^1 b^1$. Учитывая правило для вычисления биномиальных коэффициентов, перепишем эту формулу в виде $(a+b)^1 = a^1 + b^1$, т. е. убеждаемся в справедливости формулы (13) для $n=1$. Предположим, что формула (13) справедлива для $n=k$, т. е. предположим, что справедлива формула

$$(a+b)^k = C_k^0 a^k + C_k^1 a^{k-1} b + \ldots + C_k^{l-1} a^{k-(l-1)} b^{l-1} + C_k^l a^{k-l} b^l + \\ + C_k^{l+1} a^{k-(l+1)} b^{l+1} + \ldots + C_k^{k-1} a b^{k-1} + C_k^k b^k. \qquad (14)$$

Докажем, используя справедливость формулы (14), что формула (13) верна для $n=k+1$, т. е. докажем справедливость формулы

$$(a+b)^{k+1} = C_{k+1}^0 a^{k+1} + C_{k+1}^1 a^{(k+1)-1} b + \\ + \ldots + C_{k+1}^l a^{(k+1)-l} b^l + C_{k+1}^{l+1} a^{(k+1)-(l+1)} b^{l+1} + \\ + \ldots + C_{k+1}^{(k+1)-1} a b^k + C_{k+1}^{k+1} b^{k+1}. \qquad (15)$$

Действительно, используя сначала свойства степени с натуральным показателем, затем формулу (14), затем правило перемножения многочленов, получим

$$(a+b)^{k+1} = (a+b)(a+b)^k = (a+b)(C_k^0 a_1^k + C_k^1 a^{k-1} b + C_k^2 a^{k-2} b^2 + \\ + \ldots + C_k^{l-1} a^{k-(l-1)} b^{l-1} + C_k^l a^{k-l} b^l + C_k^{l+1} a^{k-(l+1)} b^{l+1} + \\ + \ldots + C_k^{k-1} a b^{k-1} + C_k^k b^k) = C_k^0 a^{k+1} + C_k^1 a^k b + C_k^2 a^{k-1} b^2 + \\ + \ldots + C_k^{l-1} a^{k-l+2} b^{l-1} + C_k^l a^{k-l+1} b^l + C_k^{l+1} a^{k-l} b^{l+1} + \\ + \ldots + C_k^{k-1} a^2 b^{k-1} + C_k^k a b^k + C_k^0 a^k b + C_k^1 a^{k-1} b^2 + C_k^2 a^{k-2} b^3 + \\ + \ldots + C_k^{l-1} a^{k-l+1} b^l + C_k^l a^{k-l} b^{l+1} + C_k^{l+1} a^{k-l-1} b^{l+2} + \\ + \ldots + C_k^{k-1} a b^k + C_k^k b^{k+1}.$$

Приводя в этой сумме подобные члены, получим, что

$$(a+b)^{k+1} = C_k^0 a^{k+1} + (C_k^1 + C_k^0) a^k b + (C_k^2 + C_k^1) a^{k-1} b^2 + \\ + \ldots + (C_k^l + C_k^{l-1}) a^{k-l+1} b^l + (C_k^{l+1} + C_k^l) a^{k-l} b^{l+1} + \\ + \ldots + (C_k^{k-1} + C_k^{k-2}) a^2 b^{k-1} + (C_k^k + C_k^{k-1}) a b^k + C_k^k b^{k+1}. \qquad (16)$$

В § 6 гл. I доказано следующее свойство:

$$C_n^{m+1} + C_n^m = C_{n+1}^{m+1} \quad \text{для} \quad m = 0,\ 1,\ 2,\ \ldots,\ n.$$

Пользуясь этим свойством, а также очевидными равенствами $C_k^0=1=C_{k+1}^0$ и $C_k^k=1=C_{k+1}^{k+1}$, из справедливости формулы (16) получим справедливость формулы (15). На основании метода полной математической индукции делаем вывод, что формула бинома Ньютона (13) справедлива для любого натурального числа n.

Заметим, что формулу для вычисления коэффициентов бинома Ньютона можно было вывести, используя теорию сочетаний (см. § 6 гл. I). Действительно, коэффициент бинома Ньютона при члене a^kb^{n-k} показывает, сколько раз встретится этот член, если перемножать последовательно многочлен $(a+b)$ на себя n раз, а для этого надо будет брать число a сомножителем k раз из n имеющихся в наличии чисел a, а это и есть число сочетаний из n по k, т. е. C_n^k.

В заключение заметим, что часто метод полной математической индукции применяется к доказательству утверждений, справедливых не для всех натуральных чисел n, а лишь для n, больших или равных некоторого натурального числа p. Тогда формулировка сути метода полной математической индукции остается почти такой же, но с заменой пункта 1 на пункт 1а): «Проверяется справедливость этого утверждения для $n=p$». В этом случае для доказательства справедливости утверждения для любого натурального $n\geqslant p$ делается следующее.

1. Проверяется справедливость утверждения для $n=p$.

2. Предполагается справедливость этого утверждения для $n=k$ (где $k\geqslant p$).

3. Доказывается справедливость этого утверждения для $n=k+1$, используя его справедливость для $n=k$. После этого делается вывод, что утверждение справедливо для любого натурального $n\geqslant p$.

Например, докажем для любого натурального $n\geqslant 2$ неравенство Бернулли

$$(1+\alpha)^n>1+\alpha n \quad (\alpha>-1,\ \alpha\neq 0). \tag{17}$$

Действительно, при $n=2$ неравенство (17) имеет вид

$$(1+\alpha)^2>1+2\alpha \quad (\alpha>-1,\ \alpha\neq 0). \tag{18}$$

Неравенство (18) равносильно неравенству

$$\alpha^2>0 \quad (\alpha>-1,\ \alpha\neq 0). \tag{19}$$

Неравенство (19) при $\alpha\neq 0$ очевидно. Следовательно, неравенство (18) справедливо. Предположим, что при $n=k$ $(k\geqslant 2)$ неравенство (17) справедливо, т. е.

$$(1+\alpha)^k>1+\alpha k \quad (\alpha>-1,\ \alpha\neq 0). \tag{20}$$

Докажем, используя неравенство (20), справедливость неравенства (17) для $n=k+1$, т. е. докажем неравенство

$$(1+\alpha)^{k+1}>1+\alpha(k+1) \quad (\alpha>-1,\ \alpha\neq 0). \tag{21}$$

Для доказательства умножим неравенство (20) на положительное число $(\alpha+1)$ $(\alpha+1>0$, так как $\alpha>-1)$. Получим равносильное неравенство

$$(1+\alpha)^{k+1}>(1+\alpha k)(1+\alpha) \quad (\alpha>-1,\ \alpha\neq 0). \tag{22}$$

Докажем теперь справедливость неравенства

$$(1+\alpha k)(1+\alpha)>1+\alpha(k+1). \tag{23}$$

Перенеся все члены неравенства (23) в одну сторону, раскрыв скобки и приведя подобные члены, получим равносильное неравенство $\alpha^2 k>0$, которое очевидно, так как $\alpha\neq 0$, $k\geqslant 2$. Следовательно, неравенство (23) справедливо, но тогда, используя справедливость неравенств (22), (23) и свойство транзитивности неравенств, получим, что справедливо и неравенство (21). Таким образом, неравенство Бернулли доказано для любого натурального $n\geqslant 2$. Это неравенство имеет смысл запомнить, так как с его помощью можно доказать много других неравенств, например неравенство

$$\left(1+\frac{1}{n}\right)^n<\left(1+\frac{1}{n+1}\right)^{n+1}. \tag{24}$$

Сделав несложные преобразования, получим равносильное неравенство:

$$\left(1+\frac{1}{n}\right)^n<\left(1+\frac{1}{n+1}\right)^{n+1}\Leftrightarrow\frac{(n+2)^{n+1}n^n}{(n+1)^{n+1}(n+1)^n}>1\Leftrightarrow$$

$$\Leftrightarrow\frac{n+2}{n+1}\left[\frac{n(n+2)}{(n+1)^2}\right]^n>1\Leftrightarrow\frac{n+2}{n+1}\left[1-\frac{1}{(n+1)^2}\right]^n>1. \tag{25}$$

Для доказательства неравенства (25) к выражению $[1-1/(n+1)^2]^n$, считая $\alpha=-1/(n+1)^2$, применим неравенство Бернулли (17). Получим $\left[1+\left(-\frac{1}{(n+1)^2}\right)\right]^n>1-\frac{n}{(n+1)^2}$. Используя свойство транзитивности неравенств, получим следующую цепочку неравенств:

$$\frac{n+2}{n+1}\left[1-\frac{1}{(n+1)^2}\right]^n>\frac{n+2}{n+1}\left[1-\frac{n}{(n+1)^2}\right]=\frac{n^3+3n^2+3n+2}{n^3+3n^2+3n+1}>1.$$

Последнее неравенство в цепочке очевидно, так что неравенство (24) доказано для $n\geqslant 2$.

Приведем еще один пример на применение обобщенного метода математической индукции.

Доказать, что для любого натурального числа $n\geqslant 2$ и любых действительных чисел A и B справедливо равенство

$$(A-B)(A^{n-1}+A^{n-2}B+A^{n-3}B^2+\ldots+AB^{n-2}+B^{n-1})=A^n-B^n. \tag{26}$$

Доказательство. Для $n=2$ равенство (26) превращается в верное равенство $(A-B)(A+B)=A^2-B^2$, доказанное в § 3.

Предположим, что при $n=k$ $(k \geqslant 2)$ равенство (26) справедливо, т. е. справедливо равенство

$$(A-B)(A^{k-1}+A^{k-2}B+A^{k-3}B^2+\ldots+AB^{k-2}+B^{k-1})= \\ =A^k-B^k. \qquad (27)$$

Докажем, используя равенство (27), справедливость равенства (26) для $n=k+1$, т. е. докажем равенство

$$(A-B)(A^k+A^{k-1}B+A^{k-2}B^2+\ldots+AB^{k-1}+B^k)= \\ =A^{k+1}-B^{k+1}. \qquad (28)$$

Действительно, используя свойства действий над алгебраическими выражениями и равенство (27), имеем цепочку тождественных равенств

$$(A-B)(A^k+A^{k-1}B+A^{k-2}B^2+\ldots+AB^{k-1}+B^k)= \\ =(A-B)(A^k+A^{k-1}B+A^{k-2}B^2+\ldots+AB^{k-1})+(A-B)B^k= \\ =(A-B)(A^{k-1}+A^{k-2}B+A^{k-3}B^2+\ldots+B^{k-1})A+(A-B)B^k= \\ =(A^k-B^k)A+(A-B)B^k=A^{k+1}-B^kA+AB^k-B^{k+1}= \\ =A^{k+1}-B^{k+1}.$$

Тем самым равенство (28) доказано, а значит, доказано и утверждение (26).

Глава III

АЛГЕБРАИЧЕСКИЕ УРАВНЕНИЯ И НЕРАВЕНСТВА

Пусть даны два многочлена A и B. Если стоит задача (см. § 2, гл. II) решить уравнение $A = B$, то говорят, что дано *алгебраическое уравнение* $A = B$. Если же стоит задача решить неравенство $A > B$ ($A < B$, $A \geqslant B$, $A \leqslant B$), то говорят, что дано *алгебраическое неравенство* $A > B$ ($A < B$, $A \geqslant B$, $A \leqslant B$).

В этой главе рассматриваются только алгебраические уравнения и неравенства. Поэтому в этой главе часто вместо слов «алгебраическое уравнение» будем писать просто «уравнение», вместо слов «алгебраическое неравенство» будем писать «неравенство».

§ 1
УРАВНЕНИЯ С ОДНИМ НЕИЗВЕСТНЫМ

Пусть стоит задача решить уравнение

$$R(x) = Q(x), \tag{1}$$

где $R(x)$ и $Q(x)$ — многочлены, целые (см. § 3 гл. II) относительно одной буквы x, тогда букву x называют неизвестной буквой, или просто *неизвестным*, а уравнение (1) — *алгебраическим уравнением с одним неизвестным*.

Поскольку ОДЗ многочленов $R(x)$ и $Q(x)$ состоит из всех действительных чисел, то задача о решении уравнения (1) может быть сформулирована так: найти **все** числовые значения неизвестного x, каждое из которых обращает уравнение (1) в верное числовое равенство. Каждое такое число называют корнем или решением уравнения (1). Поэтому *решить уравнение* (1) — это значит найти множество всех его корней. В случае, если множество всех корней уравнения (1) есть пустое множество, говорят, что уравнение (1) не имеет корней.

Например, очевидно, что уравнение $x^2 + 1 = -x^4 - 1$ корней не имеет, ибо ни для одного числового значения неизвестного x это уравнение не превращается в верное числовое равенство.

Теперь рассмотрим *простейшее алгебраическое уравнение с одним неизвестным*

$$x = \alpha, \tag{2}$$

где α — некоторое фиксированное число. Очевидно, что это простейшее уравнение имеет единственный корень — число α, поэтому множество всех корней уравнения (2) состоит из одного числа α.

Не для всякого алгебраического уравнения с одним неизвестным так же очевидно, как в двух рассмотренных примерах, множество всех его корней. Обычно для нахождения множества всех корней уравнения это уравнение равносильными переходами (определение равносильного перехода см. ниже) сводят к одному или совокупности нескольких уравнений (определение совокупности уравнений см. ниже), каждое из которых — либо простейшее уравнение типа (2), либо уравнение, для которого очевидно, что оно не имеет корней.

В этом параграфе рассматриваются примеры равносильных переходов. Пусть даны два алгебраических уравнения с одним неизвестным $R(x)=Q(x)$ и $S(x)=T(x)$. Эти уравнения называются *равносильными*, если любой корень первого уравнения является корнем второго уравнения, и, наоборот, любой корень второго уравнения является корнем первого уравнения. В силу этого определения равносильны любые два уравнения, не имеющие корней.

Замена одного уравнения равносильным ему другим уравнением называется *равносильным переходом* от одного уравнения к другому. Равносильный переход от одного уравнения к другому обозначается двойной стрелкой. Запись $R(x)=Q(x) \Leftrightarrow$ $\Leftrightarrow T(x)=S(x)$ означает, что уравнения $R(x)=Q(x)$ и $T(x)=S(x)$ равносильны.

Приведем некоторые утверждения, при помощи которых и будут совершаться равносильные переходы.

1. Уравнения $R(x)=Q(x)$ и $R(x)-Q(x)=0$ равносильны.

2. Уравнения $R(x)=Q(x)$ и $R(x)+\alpha=Q(x)+\alpha$ равносильны для любого действительного числа α.

3. Уравнения $R(x)=Q(x)$ и $\alpha R(x)=\alpha Q(x)$ равносильны для любого отличного от нуля действительного числа α.

4. Пусть известно, что для любого действительного числа x справедливо равенство $R(x)=T(x)$, тогда равносильны уравнения $R(x)=Q(x)$ и $T(x)=Q(x)$. Доказательства справедливости этих утверждений аналогичны, поэтому докажем, например, утверждение 2. Пусть число x_1 — некоторый корень уравнения $R(x)=Q(x)$. Тогда справедливо числовое равенство $R(x_1)=Q(x_1)$. Поскольку справедливость числового равенства не нарушается при прибавлении к обеим его частям любого действительного числа (см. § 2, 5 гл. I), то справедливо числовое равенство $R(x_1)+\alpha=Q(x_1)+\alpha$. Справедливость этого числового равенства означает, что число x_1 является корнем уравнения $R(x)+\alpha=$ $=Q(x)+\alpha$. Поскольку такое рассуждение можно провести для любого корня уравнения $R(x)=Q(x)$, то тем самым показано,

что любой корень уравнения $R(x)=Q(x)$ является корнем уравнения $R(x)+\alpha=Q(x)+\alpha$.

Покажем теперь обратное. Пусть число x_2 есть некоторый корень уравнения $R(x)+\alpha=Q(x)+\alpha$, тогда справедливо числовое равенство $R(x_2)+\alpha=Q(x_2)+\alpha$. Прибавив к обеим частям этого числового равенства число $-\alpha$, получим справедливость числового равенства $R(x_2)=Q(x_2)$, откуда вытекает, что число x_2 есть корень уравнения $R(x)=Q(x)$. Поскольку такое рассуждение можно провести для любого корня уравнения $R(x)+\alpha=Q(x)+\alpha$, то тем самым показано, что любой корень уравнения $R(x)+\alpha=Q(x)+\alpha$ является корнем уравнения $R(x)=Q(x)$.

Из доказанного вытекает, что если уравнение $R(x)=Q(x)$ не имеет корней, то и уравнение $R(x)+\alpha=Q(x)+\alpha$ корней не имеет. Действительно, предположим, что уравнение $R(x)=Q(x)$ не имеет корней, а уравнение $R(x)+\alpha=Q(x)+\alpha$ имеет хотя бы один корень. Из условия, что уравнение $R(x)+\alpha=Q(x)+\alpha$ имеет корень, по доказанному выше следует, что имеет корень и уравнение $R(x)=Q(x)$, что противоречит предположению. Значит, если уравнение $R(x)=Q(x)$ не имеет корней, то и уравнение $R(x)+\alpha=Q(x)+\alpha$ корней не имеет.

Аналогично показывается, что если уравнение $R(x)+\alpha=Q(x)+\alpha$ не имеет корней, то и уравнение $R(x)=Q(x)$ корней не имеет. Итак, показано, что уравнения $R(x)=Q(x)$ и $R(x)+\alpha=Q(x)+\alpha$ равносильны, т. е. доказано утверждение 2.

Пусть дан многочлен $P(x)$ степени n целый относительно буквы x

$$P(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n \qquad (a_0\neq 0), \tag{3}$$

где буквами $a_0, a_1, \ldots, a_{n-1}, a_n$ обозначены некоторые фиксированные действительные числа, называемые коэффициентами многочлена $P(x)$.

Из утверждений 1 и 4 вытекает, что каждое алгебраическое уравнение с одним неизвестным можно привести к виду $P(x)=0$, поэтому достаточно рассмотреть лишь уравнения

$$P(x)=0, \tag{4}$$

где $P(x)$ есть многочлен вида (3). Всякое такое уравнение называется *алгебраическим уравнением степени n*. Рассмотрим, как находить корни уравнения (4). Начнем со случая, когда $P(x)$ есть многочлен первой степени, т. е. рассмотрим уравнение

$$a_0x+a_1=0 \qquad (a_0\neq 0). \tag{5}$$

Такое уравнение принято называть линейным уравнением с одним неизвестным. На основании утверждения 2 уравнение (5) равносильно уравнению

$$a_0x=-a_1 \qquad (a_0\neq 0). \tag{6}$$

Поскольку число $\frac{1}{a_0} \neq 0$, то на основании утверждения 3 уравнение (6) равносильно уравнению

$$x = -\frac{a_1}{a_0} \qquad (a_0 \neq 0). \tag{7}$$

Все равносильные переходы от уравнения (5) к уравнению (7) можно записать более коротко в виде следующей цепочки равносильных переходов: $a_0x + a_1 = 0 \quad (a_0 \neq 0) \Leftrightarrow a_0x = -a_1 \times \times (a_0 \neq 0) \Leftrightarrow x = -a_1/a_0 (a_0 \neq 0)$. Простейшее уравнение $x = -\frac{a_1}{a_0}$ имеет единственный корень — число $-\frac{a_1}{a_0}$. Так как уравнение (5) равносильно простейшему уравнению (7), то уравнение (5) также имеет единственный корень — число $-\frac{a_1}{a_0}$. Таким образом, линейное уравнение (5) с одним неизвестным имеет только один корень $x_1 = -\frac{a_1}{a_0}$.

Для решения алгебраических уравнений более высоких степеней понадобится понятие совокупности уравнений. Пусть дано m многочленов $P_1(x), P_2(x), \ldots, P_m(x)$. Говорят, что дана *совокупность m алгебраических уравнений* с одним неизвестным x:

$$P_1(x) = 0, \; P_2(x) = 0, \; \ldots, \; P_m(x) = 0, \tag{8}$$

если требуется найти все числовые значения неизвестного x, каждое из которых является корнем хотя бы одного уравнения из этой совокупности (8) (уравнения совокупности обычно записываются в строчку).

Таким образом, решить совокупность уравнений (8) — значит решить каждое уравнение $P_i(x) = 0$, где $i = 1, 2, \ldots, m$, т. е. найти множества $N_1, N_2, \ldots, N_m$ всех корней каждого из уравнений и затем взять объединение этих множеств. Это объединение $N = N_1 \cup N_2 \cup \ldots \cup N_m$ и будет решением совокупности уравнений (8), а каждое число из множества N будет называться корнем или решением совокупности (8).

Например, решим совокупность уравнений $3x + 4 = 0$, $-7x + 2 = 0$, $2x - \sqrt{5} = 0$, $-12x - 16 = 0$. Первое уравнение имеет корень $(-4/3)$, второе — $(2/7)$, третье — $(\sqrt{5}/2)$, четвертое — $(-16/12)$. Совокупность этих уравнений имеет три корня: $x_1 = -4/3$, $x_2 = 2/7$, $x_3 = \sqrt{5/2}$.

Часто возникает необходимость совершать равносильный переход от уравнения к совокупности уравнений. Будем говорить, что *уравнение*

$$P(x) = 0 \tag{4}$$

равносильно совокупности уравнений

$$P_1(x) = 0, \; P_2(x) = 0, \; \ldots, \; P_m(x) = 0, \tag{8}$$

если любое решение (любой корень) уравнения (4) является решением (корнем) совокупности (8) и, наоборот, любое решение (любой корень) совокупности (8) является решением (корнем) уравнения (4). Замена уравнения (4) равносильной ему совокупностью (8) называется *равносильным переходом от уравнения* (4) *к совокупности* (8).

Например, уравнение

$$(3x+4)(-7x+2)\left(2x-\sqrt{5}\right)(-12x-16)=0 \tag{9}$$

равносильно совокупности уравнений $3x+4=0$, $-7x+2=0$, $2x-\sqrt{5}=0$, $-12x-16=0$. Действительно, любой корень уравнения (9) обращает в нуль хотя бы один из многочленов $(3x+4)$, $(-7x+2)$, $\left(2x-\sqrt{5}\right)$, $(-12x-16)$, т. е. является корнем хотя бы одного из уравнений совокупности, и, наоборот, любой корень совокупности обращает в нуль хотя бы один из этих многочленов, т. е. удовлетворяет уравнению (9). Совокупность уравнений имеет три корня: $x_1=-4/3$, $x_2=2/7$, $x_3=\sqrt{5}/2$. Следовательно, в силу равносильного перехода эти корни и только они являются корнями уравнения (9).

Рассмотрим теперь алгебраическое уравнение второй степени, т. е. уравнение

$$ax^2+bx+c=0 \qquad (a\neq 0). \tag{10}$$

Такие уравнения принято называть квадратными уравнениями. Многочлен ax^2+bx+c, где $a\neq 0$, называют обычно *квадратным трехчленом*. Коэффициенты квадратного трехчлена имеют специальные названия: числа a, стоящее при x^2, называется первым коэффициентом, число b, стоящее при x — вторым коэффициентом, число c — свободным членом. Кроме того, число $D=b^2-4ac$ называется *дискриминантом* квадратного трехчлена, а также дискриминантом квадратного уравнения (10). Проведем тождественное преобразование квадратного трехчлена. Так как $a\neq 0$, то справедливо тождественное равенство

$$ax^2+bx+c=a\left(x^2+\frac{b}{a}x+\frac{c}{a}\right).$$

Теперь применим тождественное преобразование, которое называется «выделением полного квадрата»:

$$x^2+\frac{b}{a}x+\frac{c}{a}=\left(x^2+2x\frac{b}{2a}+\frac{b^2}{4a^2}\right)-\frac{b^2}{4a^2}+\frac{c}{a}=$$
$$=\left(x+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a^2}=\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}.$$

Окончательно получаем справедливость следующего тождественного равенства:

$$ax^2+bx+c=a\left[\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}\right] \qquad (a\neq 0).$$

На основании утверждения 4 уравнение (10) равносильно уравнению

$$a\left[\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}\right]=0 \qquad (a\neq 0), \tag{11}$$

а на основании утверждения 3, учитывая, что $a\neq 0$, получаем, что уравнение (11) равносильно уравнению

$$\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}=0 \qquad (a\neq 0). \tag{12}$$

Более коротко это можно записать так:

$$ax^2+bx+c=0\ (a\neq 0)\Leftrightarrow a\left[\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}\right]=0\ \ (a\neq 0)\Leftrightarrow$$
$$\Leftrightarrow\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}=0 \qquad (a\neq 0).$$

В зависимости от дискриминанта D возможны два случая.

а) $D<0$. Так как при любом числовом значении x_0 число $\left(x_0+\frac{b}{2a}\right)^2$ — неотрицательно, а число $\left(-\frac{D}{4a^2}\right)$ — положительно, то число $\left(x_0+\frac{b}{2a}\right)^2-\frac{D}{4a^2}$ — также положительно и потому не может равняться нулю. А это означает, что уравнение (12) не имеет действительных корней. Поскольку уравнение (10) равносильно уравнению (12), то оно также не имеет действительных корней.

б) $D\geqslant 0$. Тогда $D=(\sqrt{D})^2$ и поэтому выражение, стоящее в левой части уравнения (12) можно рассматривать как разность двух квадратов $\left(x+\frac{b}{2a}\right)^2$ и $\left(\frac{\sqrt{D}}{2a}\right)^2$. Воспользовавшись формулой сокращенного умножения, получим уравнение

$$\left[\left(x+\frac{b}{2a}\right)+\frac{\sqrt{D}}{2a}\right]\left[\left(x+\frac{b}{2a}\right)-\frac{\sqrt{D}}{2a}\right]=0 \qquad (a\neq 0), \tag{13}$$

равносильное уравнению (12). Уравнение (13) в свою очередь равносильно совокупности двух уравнений

$$x+\frac{b}{2a}+\frac{\sqrt{D}}{2a}=0\ \ (a\neq 0),\ \ x+\frac{b}{2a}-\frac{\sqrt{D}}{2a}=0\ \ (a\neq 0). \tag{14}$$

Каждое уравнение в этой совокупности — линейное и, следовательно, по доказанному выше имеет один корень. Решая каждое уравнение совокупности (14), получаем, что совокупность уравнений (14) имеет два корня:

$$x_1=\frac{-b+\sqrt{D}}{2a}\ \ (a\neq 0),\ \ x_2=\frac{-b-\sqrt{D}}{2a}\ \ (a\neq 0). \tag{15}$$

В силу равносильных переходов уравнение (10) при $D\geqslant 0$ равносильно совокупности (14) и потому имеет два корня: x_1 и x_2, вычисляемые по формулам (15). Отметим, что если $D=0$, то эти корни совпадают: $x_1=x_2=-b/2a$.

Итак, квадратное уравнение (10) не имеет действительных корней, если его дискриминант отрицателен, имеет два действительных неравных корня, если дискриминант положителен, и имеет два совпавших корня, если дискриминант равен нулю.

Отметим, что если дискриминант квадратного уравнения (10) неотрицателен, то формулы (15) для нахождения корней этого уравнения часто записывают в виде одной формулы:

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \qquad (a\neq 0). \tag{16}$$

Квадратный трехчлен, у которого первый коэффициент равен единице, называется *приведенным квадратным трехчленом*. Общепринято второй коэффициент приведенного трехчлена обозначать p, а его свободный член — q, т. е. приведенный квадратный трехчлен имеет вид x^2+px+q.

Квадратное уравнение вида

$$x^2+px+q=0 \tag{17}$$

называется приведенным квадратным уравнением.

Очевидно, что квадратное уравнение (10) равносильно соответствующему приведенному уравнению, а именно

$$ax^2+bx+c=0 \ \ (a\neq 0) \Longleftrightarrow x^2+\frac{b}{a}x+\frac{c}{a}=0 \quad (a\neq 0).$$

Если дискриминант приведенного уравнения (17) неотрицателен, то формула (16) для нахождения корней этого уравнения принимает вид

$$x_{1,2}=-\frac{p}{2}\pm\sqrt{\left(\frac{p}{2}\right)^2-q}. \tag{18}$$

Теорема Виета. Если приведенное квадратное уравнение $x^2+px+q=0$ имеет неотрицательный дискриминант, то сумма корней этого уравнения равна второму его коэффициенту, взятому с противоположным знаком, а произведение корней равно свободному члену, т. е. если x_1 и x_2 корни уравнения x^2+px+ $+q=0$, то $x_1+x_2=-p$; $x_1x_2=q$.

Доказательство. Так как $D\geqslant 0$, то, используя формулы (18), получим

$$x_1+x_2=\left(-\frac{p}{2}+\sqrt{\left(\frac{p}{2}\right)^2-q}\right)+\left(-\frac{p}{2}-\sqrt{\left(\frac{p}{2}\right)^2-q}\right)=-p,$$

$$x_1x_2=\left(-\frac{p}{2}+\sqrt{\left(\frac{p}{2}\right)^2-q}\right)\left(-\frac{p}{2}-\sqrt{\left(\frac{p}{2}\right)^2-q}\right)=$$

$$=\frac{p^2}{4}-\frac{p^2}{4}+q=q.$$

Теорема доказана.

Если дискриминант квадратного уравнения (10) ax^2+bx+ $+c=0$ $(a\neq 0)$ неотрицателен, то его корни x_1 и x_2 таковы, что

$$x_1+x_2=-b/a, \ \ x_1x_2=c/a.$$

Замечание. Следует отметить, что теорема Виета имеет место и при $D<0$, но в этом случае корнями квадратного уравнения будут комплексные сопряженные числа (гл. X).

Теорема Виета часто применяется при решении различных задач. Например, дано квадратное уравнение $x^2+x+q=0$, где свободный член неизвестен, но зато известно, что данное уравнение имеет два действительных корня x_1 и x_2 и сумма квадратов этих корней равна единице, т. е. $x_1^2+x_2^2=1$. Чтобы найти q, воспользуемся теоремой Виета. Справедлива цепочка тождественных равенств $x_1^2+x_2^2=x_1^2+2x_1x_2+x_2^2-2x_1x_2=(x_1+x_2)^2-2x_1x_2=1-2q$ и потому $1-2q=1$, откуда $q=0$.

Умение находить корни линейного и квадратного уравнений дает возможность решать алгебраические уравнения (4) более высоких степеней. Рассмотрим некоторые из них. *Симметричное* уравнение 3-й степени имеет вид

$$ax^3+bx^2+bx+a=0 \quad (a\neq 0). \tag{19}$$

Преобразуем многочлен ax^3+bx^2+bx+a, воспользовавшись способом разложения многочлена на множители. Очевидна справедливость следующей цепочки тождественных равенств $ax^3+bx^2+bx+a=a(x^3+1)+bx(x+1)=a(x+1)(x^2-x+1)+bx(x+1)=(x+1)[a(x^2-x+1)+bx]=(x+1)[ax^2+(b-a)x+a]$, т. е. уравнение (19) равносильно уравнению

$$(x+1)[ax^2+(b-a)x+a]=0 \quad (a\neq 0). \tag{20}$$

Уравнение (20), в свою очередь, равносильно совокупности уравнений

$$x+1=0,\ ax^2+(b-a)x+a=0 \quad (a\neq 0). \tag{21}$$

Следовательно, и уравнение (19) равносильно этой совокупности. Решение совокупности (21) легко находится, так как она содержит только уравнения первой и второй степени.

Пример. Найти корни уравнения

$$x^3+4x^2+4x+1=0. \tag{22}$$

Преобразуем левую часть уравнения: $x^3+4x^2+4x+1=(x^3+1)+4x(x+1)=(x+1)(x^2+3x+1)$. Очевидно, что уравнение (22) равносильно совокупности уравнений

$$x+1=0,\ x^2+3x+1=0. \tag{23}$$

Первое уравнение совокупности (23) имеет один корень -1, второе — два корня: $x_1=(-3+\sqrt{5})/2$ и $x_2=(-3-\sqrt{5})/2$. Следовательно, совокупность уравнений (23), а значит, и данное уравнение (22) имеют три корня: $x_1=-1$, $x_2=(-3+\sqrt{5})/2$, $x_3=(-3-\sqrt{5})/2$.

Уравнение четвертой степени (4), в котором коэффициенты при нечетных степенях x равны нулю, называется *биквадрат-*

ным уравнением, т. е. уравнение вида

$$ax^4+bx^2+c=0 \qquad (a\neq 0) \tag{24}$$

называется биквадратным уравнением.

Для решения биквадратного уравнения (24) его левая часть преобразуется способом «выделения полного квадрата»:

$$ax^4+bx^2+c=a\left[\left(x^4+2x^2\cdot\frac{b}{2a}+\left(\frac{b}{2a}\right)^2\right)-\frac{b^2}{4a^2}+\frac{c}{a}\right]=$$
$$=a\left[\left(x^2+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a^2}\right].$$

На основании этого тождественного равенства уравнение (24) равносильно уравнению

$$a\left[\left(x^2+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a^2}\right]=0 \qquad (a\neq 0). \tag{25}$$

Очевидно, что если $b^2-4ac<0$, то уравнение (25), а значит, и равносильное ему уравнение (24) корней не имеют. Если же $b^2-4ac\geqslant 0$, то уравнение (25), а значит, и равносильное ему уравнение (24) равносильны совокупности уравнений

$$x^2+\frac{b}{2a}-\frac{\sqrt{b^2-4ac}}{2a}=0 \ (a\neq 0), \quad x^2+\frac{b}{2a}+\frac{\sqrt{b^2-4ac}}{2a}=0 \ (a\neq 0).$$

Перепишем эту совокупность в равносильном виде

$$x^2=\frac{-b+\sqrt{b^2-4ac}}{2a} \quad (a\neq 0), \quad x^2=\frac{-b-\sqrt{b^2-4ac}}{2a} \quad (a\neq 0). \tag{26}$$

Поскольку числа, стоящие в правых частях уравнений совокупности (26), есть корни квадратного уравнения

$$at^2+bt+c=0 \qquad (a\neq 0), \tag{27}$$

имеющего неотрицательный дискриминант $D=b^2-4ac$, то совокупность уравнений (26) может быть записана в виде

$$x^2=t_1, \quad x^2=t_2, \tag{28}$$

где t_1 и t_2 — корни уравнения (27).

Таким образом, показано, что для решения биквадратного уравнения (24) надо сначала решить квадратное уравнение (27), при этом, если квадратное уравнение (27) не имеет действительных корней, т. е. если его дискриминант отрицателен, то уравнение (24) также не имеет корней; если же уравнение (27) имеет два корня t_1 и t_2, т. е. если его дискриминант неотрицателен, то уравнение (24) равносильно совокупности уравнений (28). Каждое из уравнений совокупности (28) — квадратное, поэтому корни этой совокупности, а значит, и корни равносильного этой совокупности уравнения (24) легко найти.

Пример. Решить биквадратное уравнение

$$x^4-x^2-6=0. \tag{29}$$

Для решения уравнения (29) решим сначала квадратное уравнение $t^2-t-6=0$. Корни этого уравнения: $t_1=-2$, $t_2=3$. Поэтому уравнение (29) равносильно совокупности уравнений $x^2=-2$, $x^2=3$. Первое уравнение этой совокупности действительных корней не имеет, а второе имеет два корня: $x_1=\sqrt{3}$ и $x_2=-\sqrt{3}$. Значит, уравнение (29) имеет два корня: $x_1=\sqrt{3}$ и $x_2=-\sqrt{3}$.

Преобразование левой части уравнения (4) способом «выделения полного квадрата» применяется не только для решения квадратных и биквадратных уравнений, но и для целого ряда других уравнений, левую часть которых удается представить как разность квадратов некоторых двух многочленов от x. Продемонстрируем это на примере решения *симметричного уравнения четвертой степени*, т. е. уравнения вида

$$ax^4+bx^3+cx^2+bx+a=0 \qquad (a\neq 0). \tag{30}$$

Учитывая, что $a\neq 0$, запишем это уравнение в равносильном виде $(x^4+1)+\frac{b}{a}x(x^2+1)+\frac{c}{a}x^2=0$ $(a\neq 0)$. Очевидна справедливость следующей цепочки тождественных равенств:

$$(x^4+1)+\frac{b}{a}x(x^2+1)+\frac{c}{a}x^2=(x^4+2x^2+1)+\frac{b}{a}x(x^2+1)+$$
$$+\left(\frac{c}{a}-2\right)x^2=(x^2+1)^2+2(x^2+1)\frac{bx}{2a}+\left(\frac{bx}{2a}\right)^2-\left(\frac{bx}{2a}\right)^2+$$
$$+\left(\frac{c}{a}-2\right)x^2=\left(x^2+1+\frac{bx}{2a}\right)^2-\frac{b^2-4a(c-2a)}{4a^2}x^2.$$

Из справедливости этой цепочки получаем, что уравнение (30) равносильно уравнению

$$\left(x^2+\frac{b}{2a}x+1\right)^2-\frac{b^2-4a(c-2a)}{4a^2}x^2=0 \qquad (a\neq 0). \tag{31}$$

В зависимости от числа $M=b^2-4a(c-2a)$ возможны два случая:

а) $M<0$. Уравнение (31), а значит, и равносильное ему уравнение (30) действительных корней не имеют.

б) $M\geqslant 0$. Уравнение (31), а значит, и равносильное ему уравнение (30) равносильны совокупности квадратных уравнений

$$x^2+\frac{b+\sqrt{b^2-4a(c-2a)}}{2a}x+1=0 \qquad (a\neq 0),$$
$$x^2+\frac{b-\sqrt{b^2-4a(c-2a)}}{2a}x+1=0 \qquad (a\neq 0),$$

каждое из которых легко решить.

Пример. Решить уравнение

$$x^4+x^3-x^2+x+1=0. \qquad (32)$$

Проведем следующую цепочку тождественных равенств:

$$x^4+x^3-x+x^2+1=(x^4+2x^2+1)+x(x^2+1)-3x^2=$$
$$=(x^2+1)^2+2(x^2+1)\frac{x}{2}+\left(\frac{x}{2}\right)^2-\frac{x^2}{4}-3x^2=$$
$$=\left(x^2+\frac{x}{2}+1\right)^2-\frac{13}{4}x^2=$$
$$=\left(x^2+\frac{1+\sqrt{13}}{2}x+1\right)\left(x^2+\frac{1-\sqrt{13}}{2}x+1\right),$$

откуда следует, что уравнение (32) равносильно совокупности уравнений $x^2+\frac{\sqrt{13}+1}{2}x+1=0$, $x^2+\frac{\sqrt{13}-1}{2}x+1=0$. Первое уравнение этой совокупности имеет два корня:

$$x_1=\frac{-\sqrt{13}-1+\sqrt{2\sqrt{13}-2}}{4},\quad x_2=\frac{-\sqrt{13}-1-\sqrt{2\sqrt{13}-2}}{4}, \qquad (33)$$

а второе уравнение действительных корней не имеет, так как его дискриминант отрицателен. Следовательно, уравнение (32) имеет два корня (33).

В заключение рассмотрим решение двучленных уравнений вида

$$x^n-1=0, \qquad (34)$$

где n — данное натуральное число. При $n=1$ уравнение (34) есть частный случай линейного уравнения и потому имеет единственный корень $x_1=1$. При $n=2$ уравнение (34) есть частный случай квадратного уравнения с положительным дискриминантом и потому имеет два корня: $x_1=1$ и $x_2=-1$. Покажем теперь, что при $n\geqslant 3$ для любого нечетного n уравнение (34) имеет один действительный корень $x_1=1$, а для любого четного n уравнение (34) имеет два действительных корня: $x_1=1$ и $x_2=-1$. Пусть $n\geqslant 3$ и n нечетное число, т. е. пусть $n=2k+1$, где k — данное натуральное число. Пользуясь формулой сокращенного умножения, получаем справедливость тождественного равенства $x^{2k+1}-1=(x-1)(x^{2k}+x^{2k-1}+\ldots+x^2+x+1)$. Из справедливости этого тождественного равенства вытекает, что уравнение (34) при $n=2k+1$ равносильно совокупности уравнений $x-1=0$, $x^{2k}+x^{2k-1}+\ldots+x^2+x+1=0$.

Первое уравнение этой совокупности имеет один корень $x_1=1$, второе уравнение совокупности действительных корней не имеет. Покажем, что для любого действительного x справедливо следующее неравенство:

$$x^{2k}+x^{2k-1}+\ldots+x^2+x+1>0. \tag{35}$$

Действительно, для любого $x\in[0, +\infty)$ справедливость неравенства (35) очевидна. Для любого $x\in[-1, 0)$, переписав левую часть неравенства (35) в виде $x^{2k}+x^{2k-2}(x+1)+\ldots$ $\ldots+x^2(x+1)+(x+1)$, убеждаемся, что первое слагаемое этой суммы положительно, а остальные — неотрицательны. Значит, для любого $x\in[-1, 0)$ неравенство (35) справедливо. Переписав левую часть неравенства (35) в виде $x^{2k-1}(x+1)+$ $+x^{2k-2}(x+1)+\ldots+x(x+1)+1$, убеждаемся, что для любого $x\in(-\infty, -1)$ все слагаемые этой суммы положительны. Значит, для любого $x\in(-\infty, -1)$ неравенство (35) справедливо.

Итак, показана справедливость неравенства (35) для любого действительного x, а это означает, что уравнение $x^{2k}+x^{2k-1}+$ $+\ldots+x^2+x+1=0$ действительных корней не имеет. Значит, уравнение (34) при $n=2k+1$ имеет один действительный корень $x_1=1$.

Пусть теперь $n=2k$ (k — данное натуральное число, $k\geqslant 2$). Пользуясь формулой сокращенного умножения, получаем справедливость тождественного равенства

$$x^{2k}-1=(x^2-1)(x^{2(k-1)}+x^{2(k-2)}+\ldots+x^4+x^2+1).$$

Из справедливости этого тождественного равенства вытекает, что уравнение (34) при $n=2k$ ($k\geqslant 2$) равносильно совокупности уравнений $x^2-1=0$, $x^{2(k-1)}+x^{2(k-2)}+\ldots+x^4+x^2+1=0$. Первое уравнение этой совокупности имеет два корня: $x_1=1$ и $x_2=-1$, а второе уравнение действительных корней не имеет, так как для любого действительного x очевидна справедливость неравенства $x^{2(k-1)}+x^{2(k-2)}+\ldots+x^4+x^2+1>0$. Значит, уравнение (34) имеет при $n=2k$ два действительных корня: $x_1=1$ и $x_2=-1$.

Итак, уравнение (34) при любом нечетном n имеет только один действительный корень $x_1=1$, а при любом четном n — только два действительных корня: $x_1=1$ и $x_2=-1$.

Отметим, что о решении алгебраических уравнений говорится также в гл. X, XI.

Замечание. Выше было показано, как решать любое линейное и любое квадратное уравнение и были выведены формулы для нахождения их корней. Что касается уравнений, степени которых выше, чем два, то были рассмотрены лишь отдельные примеры. Связано это с тем, что хотя для уравнений третьей и четвертой степеней такие формулы есть, они очень громоздки и потому применяются редко, а для уравнений пятой степени и выше таких формул нет.

§ 2

НЕРАВЕНСТВА С ОДНИМ НЕИЗВЕСТНЫМ

Пусть стоит задача: решить неравенство

$$R(x) > Q(x) \quad (\text{или } R(x) < Q(x)), \tag{1}$$

где $R(x)$ и $Q(x)$ многочлены, целые (см. § 3, гл. II) относительно одной буквы x. Поскольку ОДЗ многочленов $R(x)$ и $Q(x)$ состоит из всех действительных чисел, то задачу о решении неравенства (1) можно сформулировать так: найти все числовые значения буквы x, каждое из которых обращает неравенство (1) в верное числовое неравенство. Каждое такое числовое значение называется решением неравенства (1). Буква x называется неизвестной буквой, или, просто, *неизвестным*, неравенство (1) — *алгебраическим неравенством с одним неизвестным*. Любые два алгебраических неравенства $R(x) > Q(x)$ и $T(x) < S(x)$ называются *равносильными*, если любое решение первого неравенства является решением второго, и, наоборот, любое решение второго неравенства является решением первого. Замена одного неравенства равносильным ему другим неравенством называется *равносильным переходом* от одного неравенства к другому. Равносильный переход принято обозначать двойной стрелкой. Запись $R(x) > Q(x) \Leftrightarrow T(x) < S(x)$ означает, что неравенства $R(x) > Q(x)$ и $T(x) < S(x)$ равносильны. Приведем некоторые утверждения, при помощи которых будут совершаться равносильные переходы.

1. Неравенства $R(x) > Q(x)$ и $R(x) - Q(x) > 0$ равносильны.

2. Неравенства $R(x) > Q(x)$ и $R(x) + \alpha > Q(x) + \alpha$ равносильны для любого действительного числа α.

3. а) Неравенства $R(x) > Q(x)$ и $\alpha R(x) > \alpha Q(x)$ равносильны для любого положительного числа α.

б) Неравенства $R(x) > Q(x)$ и $\alpha R(x) < \alpha Q(x)$ равносильны для любого отрицательного числа α.

4. Пусть известно, что для любого действительного числа x справедливо равенство $R(x) = T(x)$, тогда равносильны неравенства $R(x) > Q(x)$ и $T(x) > Q(x)$.

Доказательства справедливости этих утверждений аналогичны, поэтому докажем, например, утверждение 1. Пусть число x_1 есть некоторое решение неравенства $R(x) > Q(x)$, т. е. пусть справедливо числовое неравенство $R(x_1) > Q(x_1)$. Тогда по свойству числовых неравенств справедливо и числовое неравенство $R(x_1) - Q(x_1) > 0$. Справедливость этого числового неравенства означает, что число x_1 является решением неравенства $R(x) - Q(x) > 0$. Поскольку такое рассуждение можно провести для любого решения неравенства $R(x) > Q(x)$, то любое решение неравенства $R(x) > Q(x)$ есть решение неравенства $R(x) - Q(x) > 0$.

Покажем теперь обратное. Пусть число x_2 есть некоторое решение неравенства $R(x)-Q(x)>0$, т. е. пусть справедливо числовое неравенство $R(x_2)-Q(x_2)>0$. Из справедливости последнего неравенства следует справедливость числового неравенства $R(x_2)>Q(x_2)$, а это означает, что число x_2 — решение неравенства $R(x)>Q(x)$. Поскольку такое рассуждение можно провести для любого решения неравенства $R(x)-Q(x)>0$, то любое решение неравенства $R(x)-Q(x)>0$ есть решение неравенства $R(x)>Q(x)$. Утверждение 1 доказано.

Из утверждений 1 и 4 вытекает, что каждое алгебраическое неравенство можно привести или к виду $P(x)>0$ или к виду $P(x)<0$, поэтому достаточно рассмотреть лишь неравенства вида

$$P(x)>0 \tag{2}$$

и

$$P(x)<0, \tag{3}$$

где $P(x)$ — многочлен степени n, целый относительно буквы x, т. е. $P(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n \quad (a_0\neq 0)$. Такие неравенства называют *алгебраическими неравенствами степени* n. Покажем, как находить решения неравенств (2) и (3). Пусть $P(x)$ есть многочлен первой степени, т. е. пусть надо решить неравенство

$$a_0x+a_1>0 \quad (a_0\neq 0). \tag{4}$$

На основании утверждения 2 неравенство (4) равносильно неравенству

$$a_0x>a_1 \quad (a_0\neq 0). \tag{5}$$

Рассмотрим случаи $a_0>0$ и $a_0<0$. Пусть $a_0>0$, тогда на основании утверждения 3а) неравенство (5) равносильно неравенству

$$x>-\frac{a_1}{a_0} \quad (a_0\neq 0). \tag{6}$$

Очевидно, что любое x из промежутка $\left(-\frac{a_1}{a_0};\ +\infty\right)$ удовлетворяет неравенству (6). Следовательно, промежуток $\left(-\frac{a_1}{a_0};+\infty\right)$ (рис. 2) есть множество всех решений неравенства (6). Так как неравенство (4) при $a_0>0$ равносильно неравенству (6), то множеством всех решений неравенства (4) будет промежуток $\left(-\frac{a_1}{a_0};\ +\infty\right)$. Все найденные решения неравенства (4) записывают часто так: $x\in(-a_1/a_0;\ +\infty)$ или $x>-a_1/a_0$. Все равносильные переходы от неравенства (4) к неравенству (5), а затем к очевидному неравенству (6) записываются более коротко в виде следующей цепочки равносильных переходов:

$$a_0x+a_1>0\ (a_0>0)\Leftrightarrow a_0x>-a_1\ (a_0>0)\Leftrightarrow x>-\frac{a_1}{a_0}\ (a_0>0).$$

Аналогично справедливы следующие цепочки равносильных переходов:

$$a_0x+a_1>0\ (a_0<0)\Leftrightarrow a_0x>-a_1\ (a_0<0)\Leftrightarrow x<-\frac{a_1}{a_0}\ (a_0<0);$$

$$a_0x+a_1<0\ (a_0>0)\Leftrightarrow a_0x<-a_1\ (a_0>0)\Leftrightarrow x<-\frac{a_1}{a_0}\ (a_0>0);$$

$$a_0x+a_1<0\ (a_0<0)\Leftrightarrow a_0x<-a_1\ (a_0<0)\Leftrightarrow x>-\frac{a_1}{a_0}\ (a_0<0).$$

В каждой из этих цепочек последнее неравенство — множество всех решений первого неравенства данной цепочки (при указанном ограничении на a_0).

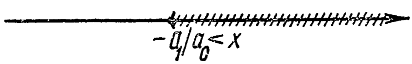


Рис. 2

Все вышенаписанное о решении линейных неравенств часто формулируют так: многочлен первой степени a_0x+a_1 $(a_0\neq 0)$

а) при $a_0>0$ положителен для любого $x\in\left(-\frac{a_1}{a_0};\ +\infty\right)$ и отрицателен для любого $x\in\left(-\infty;\ -\frac{a_1}{a_0}\right)$;

б) при $a_0<0$ положителен для любого $x\in\left(-\infty;\ -\frac{a_1}{a_0}\right)$ и отрицателен для любого $x\in\left(-\frac{a_1}{a_0};\ +\infty\right)$.

В частности, двучлен $(x-\alpha)$ положителен для всех x, находящихся на числовой оси справа от точки, изображающей число α, и отрицателен для всех x, находящихся слева от этой точки. Другими словами, точка α делит числовую ось на две части: в части, находящейся справа от точки α, двучлен $(x-\alpha)$ положителен, а в части, находящейся слева от точки α, — отрицателен. Это свойство двучлена $(x-\alpha)$ лежит в основе *метода интервалов* и часто используется для решения алгебраических неравенств более высоких степеней.

Пусть требуется решить неравенство

$$(x-c_1)(x-c_2)(x-c_3)\ \dots\ (x-c_{n-2})(x-c_{n-1})(x-c_n)>0, \quad (7)$$

где $c_1,\ c_2,\ c_3,\ \dots,\ c_{n-2},\ c_{n-1},\ c_n$ — фиксированные числа, среди которых нет равных, причем такие, что $c_1<c_2<c_3<\dots<c_{n-2}<c_{n-1}<c_n$.

Рассмотрим многочлен

$$P(x)=(x-c_1)(x-c_2)(x-c_3)\ \dots\ (x-c_{n-2})(x-c_{n-1})(x-c_n). \quad (8)$$

На основании сделанного выше замечания, очевидно, что для любого числа x_0, такого, что $x_0 > c_n$, соответствующее числовое значение любого сомножителя в произведении (8) положительно, поэтому соответствующее числовое значение многочлена $P(x_0)$ также положительно. Для любого числа x_0, взятого из промежутка (c_{n-1}, c_n), т. е. $c_{n-1} < x_0 < c_n$, соответствующее числовое значение последнего сомножителя отрицательно, а соответствующее числовое значение любого из оставшихся сомножителей положительно, поэтому число $P(x_0)$ — отрицательно; аналогично для любого числа x_0 из промежутка (c_{n-2}, c_{n-1}), т. е. $c_{n-2} < x_0 < c_{n-1}$ число $P(x_0)$ — положительно и т. д. На этом рассуждении и основан метод интервалов, состоящий в следующем. На числовую ось наносят числа $c_1, c_2, c_3, \dots, c_{n-2}, c_{n-1}, c_n$; в промежутке справа от наибольшего из них ставят знак плюс, в следующем за ним справа налево промежутке ставят знак минус, затем — знак плюс, затем — знак минус и т. д. (рис. 3). Тогда решением неравен-

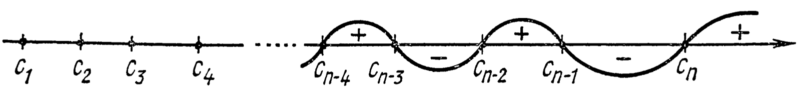


Рис. 3

ства (7) будет совокупность всех промежутков, в которых поставлен знак плюс. Методом интервалов можно решать те алгебраические неравенства, которые цепочкой равносильных переходов можно свести к неравенствам вида (7).

Пример. Решить неравенство

$$(x-3)(2+x)(4-x)>0. \quad (9)$$

Умножая неравенство (9) на —1, получим равносильное ему неравенство

$$[x-(-2)](x-3)(x-4)<0. \quad (10)$$

Для решения неравенства (10) применим метод интервалов: на числовую ось наносим числа -2, 3, 4. В промежутках справа налево расставим знаки плюс и минус (рис. 4). Множество всех x из промежутков $(-\infty, -2)$ и $(3, 4)$ — множество всех решений неравенства (10). Поскольку неравенство (9) равносильно неравенству (10), то решением неравенства (9) есть множество $(-\infty, -2) \cup (3, 4)$.

Применим метод интервалов к решению алгебраических неравенств второй степени. Отметим, что обычно их называют *квадратными неравенствами*. Рассмотрим квадратное нера-

венство
$$ax^2+bx+c>0 \quad (a\neq 0). \tag{11}$$

Применяя тождественное преобразование «выделение полного квадрата» (см. § 1 гл. III), получаем
$$ax^2+bx+c=a\left[\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}\right],$$
где $D=b^2-4ac$. Поэтому неравенство (11) равносильно неравенству
$$a\left[\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}\right]>0 \quad (a\neq 0). \tag{12}$$

Пусть $a>0$. Тогда неравенство (12) равносильно неравенству
$$\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}>0 \quad (a>0). \tag{13}$$

а) Если $D<0$, то при любом числовом значении неизвестного $x=x_0$ в левой части неравенства (13) стоит сумма неотрицательного числа $\left(x_0+\frac{b}{2a}\right)^2$ и положительного числа $\left(-\frac{D}{4a^2}\right)$, т. е. неравенство (13) превращается в верное числовое неравенство. Следовательно, неравенство (13) справедливо при

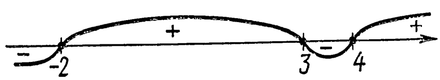


Рис. 4

любом x. Другими словами, решением неравенства (13) в этом случае будет множество всех действительных чисел.

б) Если $D=0$, то очевидно, что неравенство (13) превращается в верное числовое неравенство для любого числа x_0, кроме $x_0=-b/2a$. Следовательно, решением неравенства (13) в этом случае будет множество $(-\infty,\ -b/2a)\cup(-b/2a,\ +\infty)$.

в) Если $D>0$, то неравенство (13) равносильно неравенству
$$(x-x_1)(x-x_2)>0, \quad (a>0), \tag{14}$$
где $x_1=\frac{-b-\sqrt{D}}{2a}$, $x_2=\frac{-b+\sqrt{D}}{2a}$. Очевидно, что $x_1<x_2$, поэтому, применяя метод интервалов, получим, что множество всех решений неравенства (14) состоит из двух интервалов: $(-\infty,\ x_1)\cup(x_2,\ +\infty)$.

Пусть $a<0$. Тогда неравенство (12) равносильно неравенству
$$\left(x+\frac{b}{2a}\right)^2-\frac{D}{4a^2}<0 \quad (a<0). \tag{15}$$

а) Если $D<0$, то очевидно, что для любого числа x_0 это неравенство превращается в неверное числовое неравенство, а потому неравенство (15) не имеет решений.

б) Если $D=0$, то столь же очевидно, что неравенство (15) не имеет решений.

в) Если $D>0$, то неравенство (15) равносильно неравенству

$$(x-x_1)(x-x_2)<0 \quad (a<0), \tag{16}$$

где $x_1=\frac{-b-\sqrt{D}}{2a}$, $x_2=\frac{-b+\sqrt{D}}{2a}$. Очевидно, что $x_1>x_2$, поэтому, применяя метод интервалов, получим, что множество всех решений неравенства (16) состоит из интервала (x_2, x_1).

Аналогично проводится решение неравенства $ax^2+bx+c<0$ $(a\neq 0)$. Приведенные выше рассуждения можно собрать вместе (табл. 1).

Таблица 1.

a	D	Неравенство	Решение неравенства
$a>0$	$D>0$	$ax^2+bx+c>0$	$\left(-\infty, \frac{-b-\sqrt{D}}{2a}\right)\bigcup\left(\frac{-b+\sqrt{D}}{2a}, +\infty\right)$
$a>0$	$D>0$	$ax^2+bx+c<0$	$\left(\frac{-b-\sqrt{D}}{2a}, \frac{-b+\sqrt{D}}{2a}\right)$
$a>0$	$D=0$	$ax^2+bx+c>0$	$\left(-\infty, -\frac{b}{2a}\right)\bigcup\left(-\frac{b}{2a}, +\infty\right)$
$a>0$	$D=0$	$ax^2+bx+c<0$	нет решений
$a>0$	$D<0$	$ax^2+bx+c>0$	$(-\infty, \infty)$
$a>0$	$D<0$	$ax^2+bx+c<0$	нет решений
$a<0$	$D>0$	$ax^2+bx+c>0$	$\left(\frac{-b+\sqrt{D}}{2a}, \frac{-b-\sqrt{D}}{2a}\right)$
$a<0$	$D>0$	$ax^2+bx+c<0$	$\left(-\infty, \frac{-b+\sqrt{D}}{2a}\right)\bigcup\left(\frac{-b-\sqrt{D}}{2a}, +\infty\right)$
$a<0$	$D=0$	$ax^2+bx+c>0$	нет решений
$a<0$	$D=0$	$ax^2+bx+c<0$	$\left(-\infty, -\frac{b}{2a}\right)\bigcup\left(-\frac{b}{2a}, +\infty\right)$
$a<0$	$D<0$	$ax^2+bx+c>0$	нет решений
$a<0$	$D<0$	$ax^2+bx+c<0$	$(-\infty, +\infty)$.

Отметим, что запоминать эту таблицу не надо, для решения конкретного квадратного неравенства лучше каждый раз повторить те рассуждения, которые были сделаны выше.

Некоторые алгебраические неравенства степеней, более высоких чем два, цепочкой равносильных переходов приводятся к виду

$$(x-c_1)^{k_1}(x-c_2)^{k_2}(x-c_3)^{k_3}\dots(x-c_{n-2})^{k_{n-2}}\times$$
$$\times(x-c_{n-1})^{k_{n-1}}(x-c_n)^{k_n}>0, \qquad (17)$$

где $k_1, k_2, \dots, k_{n-1}, k_n$ — фиксированные натуральные числа, $c_1, c_2, \dots, c_{n-1}, c_n$ — некоторые фиксированные действительные числа, среди которых нет равных, и такие, что $c_1<c_2<\dots$ $\dots<c_{n-1}<c_n$ (отметим, что если хотя бы одно из чисел $k_l\geqslant 2$, то для решения неравенства (17) неприменим приведенный выше метод интервалов). Тогда неравенства вида (17) решаются так называемым *обобщенным методом интервалов*. Рассмотрим многочлен

$$P(x)=(x-c_1)^{k_1}(x-c_2)^{k_2}\dots(x-c_{n-1})^{k_{n-1}}(x-c_n)^{k_n}. \qquad (18)$$

Очевидно, что для любого числа x_0, такого, что $x_0>c_n$, соответствующее значение любого сомножителя в произведении (18) положительно, поэтому числовое значение многочлена $P(x_0)$ также положительно.

Для любого числа x_0, взятого из промежутка (c_{n-1}, c_n), т. е. $c_{n-1}<x_0<c_n$, соответствующее числовое значение любого сомножителя, кроме последнего сомножителя положительно, соответствующее числовое значение последнего сомножителя положительно, если k_n — четное число, и отрицательно, если k_n — нечетное число, поэтому число $P(x_0)$ — положительно, если k_n четное число, и число $P(x_0)$ — отрицательно, если k_n — нечетное число. Обычно в этих случаях говорят, что многочлен $P(x)$ при переходе через точку c_n меняет знак, если k_n — нечетное число и не меняет знака, если k_n — четное число. Аналогично показывается, что если известен знак многочлена $P(x)$ на промежутке (c_i, c_{i+1}), то на промежутке (c_{i-1}, c_i) знак определяется по правилу: многочлен $P(x)$ при переходе через точку c_i меняет знак, если k_i — нечетное число, и не меняет знака, если k_i — четное число. На этом рассуждении и основан обобщенный метод интервалов: на числовую ось наносятся числа $c_1, c_2, \dots, c_{n-1}, c_n$. В промежутке справа от наибольшего ставят знак плюс, в следующем за ним справа налево промежутке ставят знак плюс, если k_n — четное число, и знак минус, если k_n — нечетное число, в следующем за ним справа налево промежутке ставят знак, пользуясь правилом: многочлен $P(x)$ при переходе через точку c_{n-1} меняет знак, если k_{n-1} — нечетное число, и не меняет знак, если k_{n-1} — четное число, затем рассматривается следующий за ним справа налево промежуток, в нем ставят знак, пользуясь тем же правилом.

Таким образом рассматриваются все промежутки. Решением неравенства (17) будет совокупность всех промежутков, в которых поставлен знак плюс.

Пример. Решить неравенство

$$(x+5)(2x-3)^5(-x+7)^3(3x+8)^2<0. \qquad (19)$$

Прежде всего, умножая это неравенство на $-1/2^5 3^2$, получим равносильное ему неравенство

$$[x-(-5)][x-(-8/3)]^2(x-3/2)^5(x-7)^3>0. \qquad (20)$$

Для решения неравенства (20) применим обобщенный метод интервалов. На числовой оси отметим числа -5, $-8/3$, $3/2$, 7 (рис. 5). Справа от наибольшего, т. е. от числа 7, ставим знак плюс. При переходе через точку 7 многочлен

$$P(x)=[x-(-5)][x-(-8/3)]^2(x-3/2)^5(x-7)^3 \qquad (21)$$

меняет знак, так как двучлен $(x-7)$ содержится в произведении (21) в нечетной степени, поэтому в промежутке $(3/2, 7)$ ставим знак минус. При переходе через точку $3/2$ многочлен $P(x)$ меняет знак, так как двучлен $(x-3/2)$ содержится в произведении (21) в нечетной степени, поэтому в промежутке $(-8/3, 3/2)$

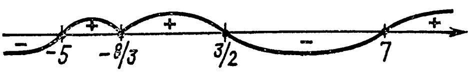


Рис. 5

ставим знак плюс. При переходе через точку $-8/3$ многочлен $P(x)$ не меняет знака, так как двучлен $[x-(-8/3)]$ содержится в произведении (21) в четной степени, поэтому в промежутке $(-5, -8/3)$ ставим знак плюс. Наконец, при переходе через точку -5 многочлен $P(x)$ меняет знак, так как двучлен $[x-(-5)]$ содержится в произведении (21) в первой степени, поэтому в промежутке $(-\infty, -5)$ ставим знак минус. Итак, решение неравенства (20) и равносильного ему неравенства (19) — совокупность всех промежутков, где поставлен знак плюс, т. е. решение неравенства (19) есть множество $(-5, -8/3) \cup (-8/3, 3/2) \cup (7, +\infty)$. Перейдем теперь к решению нестрогих неравенств

$$P(x) \geqslant 0, \qquad (22)$$

$$P(x) \leqslant 0. \qquad (23)$$

Если некоторое число x_0 есть решение неравенства (22), то справедливо числовое неравенство $P(x_0) \geqslant 0$. В силу определения нестрогого знака неравенства справедливо или числовое равенство $P(x_0)=0$, или числовое неравенство $P(x_0)>0$. Другими словами, если число x_0 — решение неравенства (22), то оно — либо решение уравнения $P(x)=0$, либо — неравенства $P(x)>0$. Такое рассуждение можно провести для любого решения неравенства $P(x) \geqslant 0$. Аналогично показывается, что любое решение неравенства $P(x) > 0$ и любое решение уравнения $P(x) = 0$ также есть решение неравенства (22).

Таким образом, решение нестрогого неравенства (22) является объединением двух множеств: множества всех решений строгого неравенства $P(x) > 0$ и множества всех решений уравнения $P(x) = 0$. Аналогично решение нестрогого неравенства (23) является объединением двух множеств: множества всех решений строгого неравенства $P(x) < 0$ и множества всех решений

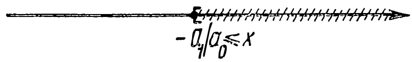


Рис. 6

уравнения $P(x) = 0$. На этом и основано правило решения нестрогих неравенств: сначала решаются соответствующее строгое неравенство и соответствующее уравнение, а затем множество решений строгого неравенства и уравнения объединяются: объединение этих множеств является множеством всех решений нестрогого неравенства.

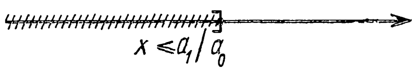


Рис. 7

Примеры. 1. Решить линейное нестрогое неравенство

$$a_0x + a_1 \geqslant 0 \quad (a_0 \neq 0). \tag{24}$$

Решаем сначала уравнение

$$a_0x + a_1 = 0 \quad (a_0 \neq 0). \tag{25}$$

Его решение — число $(-a_1/a_0)$. Затем решаем неравенство

$$a_0x + a_1 > 0 \quad (a_0 \neq 0). \tag{26}$$

При $a_0 > 0$ его решение — множество $(-a_1/a_0, +\infty)$, при $a_0 < 0$ его решение — множество $(-\infty, -a_1/a_0)$; объединяя решения уравнения (25) и неравенства (26), получаем для $a_0 > 0$ решение неравенства (24) есть любое число, большее или равное числу $(-a_1/a_0)$, т. е. множество $\left[-\frac{a_1}{a_0}, +\infty\right)$; для $a_0 < 0$ решение неравенства (24) есть любое число, меньшее или равное числу $(-a_1/a_0)$, т. е. множество $(-\infty, -a_1/a_0]$. Эти множества решений неравенства (24) принято графически обозначать так, как на рис. 6 и рис. 7 соответственно.

2. Решить неравенство

$$(x^2-3x+2)(x^3-3x^2)(4-x^2)\geqslant 0. \tag{27}$$

Поскольку справедливы следующие тождественные равенства

$$x^2-3x+2=(x-2)(x-1),$$
$$x^3-3x^2=x^2(x-3),$$
$$4-x^2=-(x-2)(x+2),$$

то согласно утверждениям 4 и 3, б этого параграфа неравенство (27) равносильно неравенству

$$(x+2)x^2(x-1)(x-2)^2(x-3)\leqslant 0. \tag{28}$$

Решим сначала уравнение

$$(x+2)x^2(x-1)(x-2)^2(x-3)=0. \tag{29}$$

Оно имеет корни: $x_1=-2$, $x_2=0$, $x_3=1$, $x_4=2$, $x_5=3$. Затем решаем строгое неравенство

$$(x+2)x^2(x-1)(x-2)^2(x-3)<0 \tag{30}$$

обобщенным методом интервалов (рис. 8). Множеством всех его решений будет множество $(-\infty, -2)\cup(1, 2)\cup(2, 3)$. Объединяя множество решений уравнения (29) и строгого неравен-

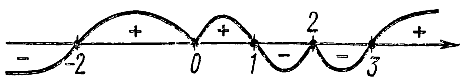


Рис. 8

ства (30), получим множество решений неравенства (28), а в силу равносильного перехода — неравенства (27). Итак, решение неравенства (27) есть множество $-\infty<x\leqslant -2$, $x=0$, $1\leqslant x\leqslant 3$.

§ 3
СИСТЕМЫ УРАВНЕНИЙ С НЕСКОЛЬКИМИ НЕИЗВЕСТНЫМИ

Пусть надо решить уравнение

$$R(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t), \tag{1}$$

где $R(x, y, z, \ldots, t)$ и $Q(x, y, z, \ldots, t)$ — многочлены, целые (см. § 3 гл. II) относительно букв $x, y, z, \ldots, t$. Тогда говорят, что дано *алгебраическое уравнение с неизвестными* x, y, $z, \ldots, t$. Заметим, что неизвестные $x, y, z, \ldots, t$ представляют собой множество всех неизвестных, содержащихся как в левой, так и в правой частях уравнения (1). Например, уравнение $4x^2=yz+5y^2$ есть уравнение относительно неизвестных x, y, z, ибо многочлены, стоящие в левой и правой частях этого уравнения, могут быть записаны в виде $R(x, y, z)=4x^2=4x^2+oy+oz$, $Q(x, y, z)=yz+5y^2=0x+yz+5y^2$, откуда видно, что эти многочлены действительно целые относительно букв x, y, z.

Упорядоченный набор $(x, y, z, \ldots, t)$ называется *набором неизвестных уравнения* (1). *ОДЗ уравнения* (1) есть множество всех числовых наборов, соответствующих набору неизвестных $(x, y, z, \ldots, t)$, у каждого из которых на месте каждого неизвестного может стоять любое действительное число.

Числовой набор $(x_0, y_0, z_0, \ldots, t_0)$, соответствующий набору неизвестных $(x, y, z, \ldots, t)$, называется *решением уравнения* (1), если равны числовые значения многочленов R и Q, соответствующие этому числовому набору, т. е. если справедливо числовое равенство $R(x_0, y_0, z_0, \ldots, t_0)=Q(x_0, y_0, z_0, \ldots, t_0)$.

Решить уравнение (1) — это значит найти **все** его решения, т. е. найти **все** числовые наборы, каждый из которых обращает уравнение (1) в верное числовое равенство.

Пусть даны два алгебраических уравнения с одними и теми же неизвестными $x, y, z, \ldots, t$:

$$R(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)$$

и

$$T(x, y, z, \ldots, t)=S(x, y, z, \ldots, t).$$

Эти уравнения называются *равносильными*, если любое решение первого уравнения является решением второго уравнения и, наоборот, любое решение второго уравнения является решением первого уравнения. Замена одного уравнения равносильным ему другим уравнением называется *равносильным переходом* от первого уравнения ко второму. Справедливы следующие утверждения.

1. Уравнения $R(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)$ и $R(x, y, z, \ldots, t)-Q(x, y, z, \ldots, t)=0$ равносильны.

2. Уравнения $R(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)$ и $R(x, y, z, \ldots, t)+S(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)+S(x, y, z, \ldots, t)$, где $S(x, y, z, \ldots, t)$ — многочлен, целый относительно букв $x, y, z, \ldots, t$, равносильны.

3. Уравнения $R(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)$ и $\alpha R(x, y, z, \ldots, t)=\alpha Q(x, y, z, \ldots, t)$ равносильны для любого, отличного от нуля действительного числа α.

4. Пусть известно, что справедливо тождественное равенство $R(x, y, z, \ldots, t)=T(x, y, z, \ldots, t)$, тогда уравнения $R(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)$ и $T(x, y, z, \ldots, t)=Q(x, y, z, \ldots, t)$ равносильны.

Справедливость этих утверждений доказывается аналогично доказательству соответствующих утверждений § 1 и потому опускается.

Из утверждений 1 и 4 вытекает, что каждое алгебраическое уравнение можно привести к виду $P(x, y, z \ldots, t)=0$, поэтому можно рассматривать лишь уравнение вида

$$P(x, y, z, \ldots, t)=0, \tag{2}$$

где $P(x, y, z, \ldots, t)$ — многочлен, целый относительно букв x, $y, z, \ldots, t$. Говорят, что дана *совокупность m алгебраических уравнений с неизвестными* $x, y, z, \ldots, t$:

$$P_1(x, y, z, \ldots, t)=0, \quad P_2(x, y, z, \ldots, t)=0, \ldots, \\ P_m(x, y, z, \ldots, t)=0, \tag{3}$$

если требуется найти все числовые наборы, соответствующие набору неизвестных $(x, y, z, \ldots, t)$, каждый из которых является решением хотя бы одного уравнения из совокупности (3). Каждый такой набор называется *решением совокупности* (3). Таким образом, решить совокупность уравнений (3) — это значит решить каждое уравнение $P_i(x, y, z, \ldots, t)=0$, где $i=1, 2, 3, \ldots, m$, а затем взять объединение этих решений.

Уравнение (2) *равносильно совокупности уравнений* (3), если любое решение уравнения (2) является решением совокупности (3), и, наоборот, любое решение совокупности (3) является решением уравнения (2). Замена уравнения (2) равносильной совокупностью (3) называется равносильным переходом от уравнения (2) к совокупности (3).

Пусть даны многочлены $P_1, P_2, P_3, \ldots, P_m$, целые относительно букв $x, y, z, \ldots, t$. Говорят, что дана *система m алгебраических уравнений с неизвестными* $x, y, z, \ldots, t$:

$$\begin{cases} P_1(x, y, z, \ldots, t)=0, \\ P_2(x, y, z, \ldots, t)=0, \\ P_3(x, y, z, \ldots, t)=0, \\ \ldots\ldots\ldots\ldots\ldots \\ P_m(x, y, z, \ldots, t)=0, \end{cases} \tag{4}$$

если требуется найти все числовые наборы, соответствующие набору неизвестных $(x, y, z, \ldots, t)$, каждый из которых является решением каждого из уравнений системы (4), т. е. если требуется найти все такие числовые наборы неизвестных $x, y, z, \ldots, t$, при подстановке каждого из которых во все уравнения системы (4) последние обращались бы в верные числовые равенства. Каждый такой числовой набор называется решением системы (4) (уравнения системы обычно записываются в столбик и объединяются фигурной скобкой).

Две системы алгебраических уравнений с одними и теми же неизвестными $x, y, z, \ldots, t$ называются *равносильными*, если любое решение первой системы является решением второй системы, и, наоборот, любое решение второй системы является решением первой. Говорят, что дана *совокупность k алгебраических систем уравнений с неизвестными* $x, y, z, \ldots, t$:

$$\begin{cases} P_{11}(x, y, z, \ldots, t)=0, \\ P_{21}(x, y, z, \ldots, t)=0, \\ P_{31}(x, y, z, \ldots, t)=0, \\ \ldots\ldots\ldots\ldots\ldots \\ P_{m1}(x, y, z, \ldots, t)=0, \end{cases} \quad \begin{cases} P_{12}(x, y, z, \ldots, t)=0, \\ P_{22}(x, y, z, \ldots, t)=0, \\ P_{32}(x, y, z, \ldots, t)=0, \ldots \\ \ldots\ldots\ldots\ldots\ldots \\ P_{m2}(x, y, z, \ldots, t)=0, \end{cases}$$

$$\begin{cases} P_{1k}(x, y, z, \ldots, t)=0, \\ P_{2k}(x, y, z, \ldots, t)=0, \\ P_{3k}(x, y, z, \ldots, t)=0, \\ \ldots\ldots\ldots\ldots\ldots \\ P_{mk}(x, y, z, \ldots, t)=0, \end{cases} \tag{5}$$

если требуется найти **все** числовые наборы, каждый из которых является решением хотя бы одной из системы уравнений совокупности (5). Каждый такой набор называется *решением совокупности систем* уравнений (5). *Система уравнений* (4) *равносильна совокупности систем уравнений* (5), если любое решение системы уравнений (4) является решением совокупности систем уравнений (5), и, наоборот, любое решение совокупности систем уравнений (5) является решением системы уравнений (4).

Приведем некоторые утверждения равносильности систем уравнений.

1. Если изменить порядок уравнений системы (4), то полученная система равносильна (4).

2. Если одно из уравнений системы (4) заменить на равносильное уравнение, полученная система равносильна (4).

3. Пусть в системе уравнений с неизвестными $x, y, z, \ldots$ $\ldots, t$ одно из уравнений записано в виде, где в левой части стоит одно из неизвестных, например x, в первой степени, а в правой части — многочлен, целый относительно других букв. Тогда говорят, что неизвестное x выражено через другие неизвестные. Если неизвестное x выражено из первого уравнения системы через другие неизвестные, то, подставив в другие уравнения системы вместо x этот многочлен от других неизвестных, получим равносильную систему уравнений, т. е. равносильны две следующие системы:

$$\begin{cases} x=Q(y, z, \ldots, t), \\ P_2(x, y, z, \ldots, t)=0, \\ P_3(x, y, z, \ldots, t)=0, \\ \ldots\ldots\ldots\ldots\ldots \\ P_m(x, y, z, \ldots, t)=0, \end{cases} \quad \text{и} \quad \begin{cases} x=Q(y, z, \ldots, t), \\ P_2[Q(y, z, \ldots, t), y, z, \ldots, t]=0, \\ P_3[Q(y, z, \ldots, t), y, z, \ldots, t]=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ P_m[Q(y, z, \ldots, t), y, z, \ldots, t]=0. \end{cases}$$

Заметим, что если во второй системе рассмотреть только уравнения $P_2=0,\ P_3=0, \dots, P_m=0$, то они образуют систему уравнений с числом неизвестных меньшим, чем в первой системе.

4. Если первое уравнение системы (4) заменить уравнением, равным сумме первого уравнения, умноженного на некоторое действительное число $\beta \neq 0$, и второго уравнения, умноженного на некоторое действительное число α, то полученная система уравнений равносильна системе уравнений (4), т. е. при любых действительных α и $\beta \neq 0$ две следующие системы уравнений равносильны:

$$\begin{cases} P_1(x,\ y,\ z, \dots,\ t)=0. \\ P_2(x,\ y,\ z, \dots,\ t)=0, \\ P_3(x,\ y,\ z, \dots,\ t)=0, \\ \dots\dots\dots\dots\dots \\ P_m(x,\ y,\ z, \dots,\ t)=0 \end{cases} \quad \text{и}$$

$$\begin{cases} \beta P_1(x,\ y,\ z, \dots,\ t)+\alpha P_2(x,\ y,\ z, \dots,\ t)=0, \\ P_2(x,\ y,\ z, \dots,\ t)=0, \\ P_3(x,\ y,\ z, \dots,\ t)=0, \\ \dots\dots\dots\dots\dots \\ P_m(x,\ y,\ z, \dots,\ t)=0. \end{cases}$$

В качестве следствия утверждения 4 имеем утверждение:

5. Если первое уравнение системы (4) заменить на сумму (или разность) первого и второго уравнений системы, то полученная система уравнений будет равносильна системе уравнений (4).

6. Если первое уравнение системы (4) равносильно совокупности уравнений: $Q_1(x, y, z, \dots, t)=0$, $Q_2(x, y, z, \dots, t)=0, \dots$, $Q_k(x,\ y,\ z, \dots,\ t)=0$, то система (4) равносильна следующей совокупности систем уравнений:

$$\begin{cases} Q_1(x,\ y,\ z, \dots,\ t)=0, \\ P_2(x,\ y,\ z, \dots,\ t)=0, \\ P_3(x,\ y,\ z, \dots,\ t)=0, \\ \dots\dots\dots\dots\dots \\ P_m(x,\ y,\ z, \dots,\ t)=0, \end{cases} \quad \begin{cases} Q_2(x,\ y,\ z, \dots,\ t)=0, \\ P_2(x,\ y,\ z, \dots,\ t)=0, \\ P_3(x,\ y,\ z, \dots,\ t)=0, \\ \dots\dots\dots\dots\dots \\ P_m(x,\ y,\ z, \dots,\ t)=0, \end{cases} \dots$$

$$\dots \begin{cases} Q_k(x,\ y,\ z, \dots,\ t)=0, \\ P_2(x,\ y,\ z, \dots,\ t)=0, \\ P_3(x,\ y,\ z, \dots,\ t)=0, \\ \dots\dots\dots\dots\dots \\ P_m(x,\ y,\ z, \dots,\ t)=0. \end{cases} \tag{6}$$

Если другие уравнения системы (4) равносильны своим совокупностям уравнений, то к каждой системе совокупности (6) применимо утверждение 6 и каждая система совокупности (6) может быть заменена своей совокупностью систем уравнений. Полная совокупность систем уравнений, равносильная системе (4), получается перебором всех логически возможных случаев (см. пример ниже).

Доказательство всех этих утверждений опускается.

Рассмотрим некоторые способы решения систем уравнений.

I. Способ подстановки или способ исключения неизвестного (способ основан на утверждении 3).

Рассмотрим применение этого способа на примере решения системы двух линейных уравнений:

$$\begin{cases} a_1x+b_1y+c_1=0, \\ a_2x+b_2y+c_2=0, \end{cases} \tag{7}$$

где $a_1^2+b_1^2\neq 0$ и $a_2^2+b_2^2\neq 0$, другими словами, где хотя бы один из двух коэффициентов a_1 и b_1, а также хотя бы один из двух коэффициентов a_2 и b_2 отличны от нуля (в противном случае по крайней мере один из многочленов $a_1x+b_1y+c_1$ или $a_2x+b_2y+c_2$ не был бы многочленом первой степени ни относительно неизвестного x, ни относительно неизвестного y). Пусть для определенности $a_1\neq 0$. Тогда первое уравнение системы (7) можно записать в следующем равносильном виде: $x=-\frac{b_1}{a_1}y-\frac{c_1}{a_1}$.

Пользуясь утверждением 3, запишем систему

$$\begin{cases} x=-\frac{b_1}{a_1}y-\frac{c_1}{a_1}, \\ a_2\left(-\frac{b_1}{a_1}y-\frac{c_1}{a_1}\right)+b_2y+c_2=0, \end{cases}$$

равносильную системе (7). Равносильно преобразовав второе уравнение этой системы, получим систему

$$\begin{cases} x=-\frac{b_1}{a_1}y-\frac{c_1}{a_1}, \\ \frac{a_1b_2-b_1a_2}{a_1}y=\frac{c_1a_2-c_2a_1}{a_1}, \end{cases} \tag{8}$$

равносильную системе (7).

Рассмотрим второе уравнение системы (8). Число $a_1b_2-a_2b_1$ обозначим буквой Δ и будем называть его *основным определителем системы* (7). Пусть $\Delta\neq 0$. Тогда из второго уравнения системы (8) находим (см. § 1), что это уравнение имеет единственное решение $y=\frac{c_1a_2-c_2a_1}{\Delta}$. Подставив это значение y в первое уравнение системы (8), получим, что $x=\frac{b_1c_2-c_1b_2}{\Delta}$. Таким образом, если $\Delta\neq 0$, то система (7) имеет единственное решение (x_1, y_1), где $x=\frac{b_1c_2-c_1b_2}{\Delta}$, $y_1=\frac{a_2c_1-c_2a_1}{\Delta}$, $\Delta=a_1b_2-a_2b_1$. Если $\Delta=0$, то возможны два случая:

а) $c_1a_2-c_2a_1=0$. Тогда любое действительное число y удовлетворяет второму уравнению системы (8), а это значит, что система (8), а значит, и равносильная ей система (7) имеет решением бесконечное множество пар вида $\left(-\frac{b_1}{a_1}y-\frac{c_1}{a_1},\ y\right)$, где y — любое действительное число;

б) $c_1a_2-c_2a_1\neq 0$. Тогда не существует действительного числа y, обращающего второе уравнение системы (8) в верное числовое равенство, а это означает, что не существует пары действительных чисел (x, y), которая обращала бы оба уравнения системы (8) в верные числовые равенства, т. е. система (8) не имеет решений. Значит, и равносильная ей система (7) решений не имеет.

II. Способ линейного преобразования (способ базируется на утверждении 4 и заключен в равносильной замене первого уравнения системы другим уравнением, равным сумме первого уравнения, умноженного на число $\beta\neq 0$, и любого другого уравнения, умноженного на число α).

Рассмотрим применение этого способа на примере решения следующей системы уравнений:

$$\begin{cases} x^2+y^2+3y+1=0, \\ x^2+y^2+y-3=0. \end{cases} \tag{9}$$

Вычитая из первого уравнения второе, получим на основании утверждения 5 равносильную систему

$$\begin{cases} 2y+4=0, \\ x^2+y^2+y-3=0. \end{cases} \tag{10}$$

Первое уравнение системы (10) имеет решение $y_1=-2$. Подставив это значение y_1 во второе уравнение системы (10), получим, что эта система, а значит, и равносильная ей система (9) имеют решения $(1, -2)$ и $(-1, -2)$. Отметим, что часто эти решения записываются в виде

$$\begin{cases} x_1=1, \\ y_1=-2; \end{cases} \quad \begin{cases} x_2=-1, \\ y_2=-2. \end{cases}$$

III. Способ замены системы уравнений совокупностью систем уравнений (способ базируется на утверждении 6 о равносильности системы уравнений совокупности систем уравнений).

Рассмотрим применение этого способа на примере решения следующей системы уравнений

$$\begin{cases} x^2-y^2-x+y=0, \\ x+y^2-2=0. \end{cases} \tag{11}$$

Поскольку первое уравнение этой системы равносильно совокупности уравнений $x-y=0$ и $x+y-1=0$, то система (11) равносильна совокупности систем:

$$\begin{cases} x-y=0, \\ x+y^2-2=0, \end{cases} \quad \begin{cases} x+y-1=0, \\ x+y^2-2=0. \end{cases} \tag{12}$$

Каждая из систем совокупности (12) легко решается способом подстановки. Первая система имеет два решения: $(1,\ 1)$; $(-2,\ -2)$; вторая система имеет два решения: $((1+\sqrt{5})/2, (1-\sqrt{5})/2)$; $((1-\sqrt{5})/2,\ (1+\sqrt{5})/2)$. Следовательно, система (11) имеет четыре решения:

$$\begin{cases} x_1=1, \\ y_1=\ 1; \end{cases} \quad \begin{cases} x_2=-2, \\ y_2=-2; \end{cases} \quad \begin{cases} x_3=\dfrac{1+\sqrt{5}}{2}, \\ y_3=\dfrac{1-\sqrt{5}}{2}; \end{cases} \quad \begin{cases} x_4=\dfrac{1-\sqrt{5}}{2}, \\ y_4=\dfrac{1+\sqrt{5}}{2}. \end{cases}$$

Рассмотрим еще применение этого способа на примере решения системы двух уравнений с двумя неизвестными, если одно из уравнений этой системы будет однородным уравнением второй степени. Уравнение $ax^2+bxy+cy^2=0$ называется *однородным уравнением второй степени*. Итак, решим систему уравнений

$$\begin{cases} ax^2+bxy+cy^2=0, \\ P(x,\ y)=0, \end{cases} \tag{13}$$

где $P(x,\ y)$ — многочлен, целый относительно x и y.

1. Пусть $a=0$. Очевидно, что система (13) равносильна совокупности систем

$$\begin{cases} y=0, \\ P(x,\ y)=0, \end{cases} \quad \begin{cases} bx+cy=0, \\ P(x,\ y)=0. \end{cases}$$

Каждую из этих систем можно решить способом подстановки.

2. Пусть $a\neq 0$. Применим к левой части первого уравнения системы (13) тождественное преобразование «выделение полного квадрата»:

$$ax^2+bxy+cy^2=a\left(x^2+\frac{b}{a}xy+\frac{c}{a}y^2\right)=a\left[\left(x^2+2xy\frac{b}{2a}+y^2\frac{b^2}{4a^2}\right)+\right.$$
$$\left.+\frac{c}{a}y^2-\frac{b^2}{4a^2}y^2\right]=a\left[\left(x+\frac{b}{2a}y\right)^2-\frac{D}{4a^2}y^2\right], \tag{14}$$

где $D=b^2-4ac$. В случае, если $D\geqslant 0$, левая часть первого уравнения системы (13) представляется в виде произведения

$$a\left(x+\frac{b}{2a}y+\frac{\sqrt{D}}{2a}y\right)\left(x+\frac{b}{2a}y-\frac{\sqrt{D}}{2a}y\right),$$

и поэтому система (13) равносильна совокупности систем

$$\begin{cases} x+\dfrac{b+\sqrt{D}}{2a}y=0, \\ P(x,\ y)=0, \end{cases} \quad \begin{cases} x+\dfrac{b-\sqrt{D}}{2a}y=0, \\ P(x,\ y)=0. \end{cases} \tag{15}$$

Каждую из этих систем можно решить способом подстановки, Отметим, что в случае, если $D=0$, совокупность систем (15) состоит из двух одинаковых систем уравнений, т. е. на самом деле есть одна система уравнений. В случае, если $D<0$. из равенства (14) вытекает, что первое уравнение системы (13) имеет единственное решение: $\begin{cases} x_1=0, \\ y_1=0, \end{cases}$ и поэтому остается проверить, удовлетворяет ли это решение второму уравнению системы (13). Решим этим способом следующую систему уравнений:

$$\begin{cases} x^2+xy+y^2=19, \\ x^2-xy+y^2=7. \end{cases} \tag{16}$$

Применим сначала способ линейного преобразования системы: умножая первое уравнение на 7, а второе на 19 и вычитая затем из второго уравнения первое, приходим к равносильной системе уравнений:

$$\begin{cases} 12x^2-26xy+12y^2=0, \\ x^2-xy+y^2=7. \end{cases} \tag{17}$$

Применим к левой части первого уравнения тождественное преобразование «выделение полного квадрата»:

$$12x^2-26xy+12y^2=12\left[x^2-2x\frac{13}{12}y+\frac{169}{144}y^2+y^2-\frac{169}{144}y^2\right]=$$
$$=12\left[\left(x-\frac{13}{2}y\right)^2-\frac{25}{144}y^2\right]=12\left(x-\frac{13}{12}y+\frac{5}{12}y\right)\left(x-\frac{13}{12}y-\frac{5}{12}y\right).$$

На основании этого тождественного преобразования и утверждения 6 можно утверждать, что система уравнений (17) равносильна совокупности систем уравнений:

$$\begin{cases} x-\dfrac{2}{3}y=0, \\ x^2-xy+y^2=7, \end{cases} \quad \begin{cases} x-\dfrac{3}{2}y=0, \\ x^2-xy+y^2=7. \end{cases}$$

Решая каждую из этих систем способом подстановки, получаем, что первая система имеет два решения: (2, 3) и (—2, —3) — и вторая система имеет два решения: (3, 2) и (—3, —2). Следовательно, система (16) имеет четыре решения:

$$\begin{cases} x_1=2; \\ y_1=3; \end{cases} \quad \begin{cases} x_2=-2, \\ y_2=-3; \end{cases} \quad \begin{cases} x_3=3, \\ y_3=2; \end{cases} \quad \begin{cases} x_4=-3, \\ y_4=-2. \end{cases}$$

Иногда для решения системы уравнений утверждение 6 надо применить не один раз, а несколько раз. Например, так надо поступить при решении системы уравнений

$$\begin{cases} x^3-y^3=19(x-y), \\ x^3+y^3=7(x+y). \end{cases} \tag{18}$$

Перепишем эту систему в следующем равносильном виде:

$$\begin{cases} (x-y)(x^2+xy+y^2-19)=0, \\ (x+y)(x^2-xy+y^2-7)=0. \end{cases}$$

На основании утверждения 6 эта система равносильна совокупности систем:

$$\begin{cases} (x-y)=0, \\ (x+y)(x^2-xy+y^2-7)=0, \end{cases} \quad \begin{cases} x^2+xy+y^2-19=0, \\ (x+y)(x^2-xy+y^2-7)=0. \end{cases}$$

Применяя к каждой системе опять утверждение 6, получим, что исходное уравнение (18) равносильно совокупности систем уравнений:

$$\begin{cases} x-y=0, \\ x+y=0, \end{cases} \quad \begin{cases} x-y=0, \\ x^2+xy+y^2-7=0, \end{cases} \quad \begin{cases} x^2+xy+y^2-19=0, \\ x+y=0, \end{cases}$$

$$\begin{cases} x^2+xy+y^2-19=0, \\ x^2-xy+y^2-7=0. \end{cases}$$

Первые три системы легко решаются способом подстановки, а четвертая система уже была решена выше. Собирая вместе решения всех этих систем, получаем, что исходная система (18) имеет девять решений:

$$\begin{cases} x_1=0, \\ y_1=0; \end{cases} \quad \begin{cases} x_2=\sqrt{7}, \\ y_2=\sqrt{7}; \end{cases} \quad \begin{cases} x_3=-\sqrt{7}, \\ y_3=-\sqrt{7}; \end{cases} \quad \begin{cases} x_4=\sqrt{19}, \\ y_4=-\sqrt{19}; \end{cases}$$

$$\begin{cases} x_5=-\sqrt{19}, \\ y_5=\sqrt{19}; \end{cases} \quad \begin{cases} x_6=2, \\ y_6=3; \end{cases} \quad \begin{cases} x_7=-2, \\ y_7=-3; \end{cases} \quad \begin{cases} y_8=3, \\ y_8=2; \end{cases} \quad \begin{cases} x_9=-3, \\ y_9=-2. \end{cases}$$

В заключение отметим, что при решении систем уравнений редко применяется только один из рассмотренных способов решения, обычно для решения системы приходится применять несколько способов.

Глава IV
СТЕПЕНИ И ЛОГАРИФМЫ

§ 1
СТЕПЕНЬ С ЦЕЛЫМ ПОКАЗАТЕЛЕМ

В зависимости от того, является ли целое число натуральным числом, целым отрицательным числом или нулем, приведем соответствующие каждому случаю определения.

Пусть a — любое действительное число, n — любое натуральное число, тогда *степенью числа a с натуральным показателем n* (или n-й степенью числа a) называется число, записываемое как a^n и определяемое по правилу

$$a^n = \begin{cases} \underbrace{a \cdot a \ldots a}_{n \text{ раз}}, & \text{если} \quad n \geqslant 2, \\ a, & \text{если} \quad n = 1. \end{cases}$$

Пусть a — любое отличное от нуля действительное число, тогда *нулевой степенью* этого числа называется число единица, т. е. по определению $a^0 = 1$ для любого отличного от нуля действительного числа a. Нулевая степень числа нуль не определяется и символ 0^0 считается лишенным смысла.

Пусть a — любое отличное от нуля действительное число, n — любое натуральное число, тогда *степенью числа a с целым отрицательным показателем* $(-n)$ называется число $\frac{1}{a^n}$, т. е. по определению $a^{-n} = \frac{1}{a^n}$ для любого отличного от нуля действительного числа a и любого целого отрицательного числа $(-n)$. Целая отрицательная степень числа нуль не определяется и символ 0^{-n} считается лишенным смысла.

Итак, натуральная (n-я) степень определяется для любого действительного числа, а нулевая или целая отрицательная степень лишь для любого отличного от нуля действительного числа.

Если a — любое отличное от нуля действительное число, то можно дать определение степени с целым показателем, которое есть объединение предыдущих определений в этом случае.

Пусть a — любое отличное от нуля действительное число, α — любое целое число, тогда под числом a^α понимают число, определяемое по правилу:

$$\begin{cases} a, & \text{если } \alpha = 1, \\ \underbrace{a \cdot a \ldots a}_{m \text{ раз}}, & \text{если } \alpha = m, \ m \text{ — натуральное число}, \\ 1, & \text{если } \alpha = 0, \\ \dfrac{1}{a^n}, & \text{если } \alpha = -n, \ -n \text{ — целое отрицательное число}, \end{cases}$$

при этом число a^α называется *степенью с целым показателем*, число a — *основанием степени*, число α — *показателем степени*.

Пусть a, b — любые действительные числа, не равные нулю, α, β — любые целые числа. Перечислим свойства степени с целым показателем:

а) $(ab)^\alpha = a^\alpha b^\alpha$;

б) $\left(\dfrac{a}{b}\right)^\alpha = \dfrac{a^\alpha}{b^\alpha}$;

в) $a^\alpha b^\beta = a^{\alpha+\beta}$;

г) $a^\alpha : a^\beta = \dfrac{a^\alpha}{a^\beta} = a^{\alpha-\beta}$;

д) $(a^\alpha)^\beta = a^{\alpha\beta}$.

Докажем их справедливость.

Пусть α натуральное число, тогда справедливость свойства 1 уже доказана ранее (см. гл. I).

Пусть $\alpha = 0$, тогда $(ab)^\alpha = (ab)^0 = 1 = 1 \cdot 1 = a^0 b^0 = a^\alpha b^\alpha$, т. е. $(ab)^\alpha = a^\alpha b^\alpha$.

Пусть $\alpha = -m$ и m — натуральное число. По определению степени с отрицательным

показателем	$(ab)^\alpha = (ab)^{-m} = \dfrac{1}{(ab)^m} =$
по свойству степени с натуральным показателем	$= \dfrac{1}{a^m b^m} =$
по свойству дробей	$= \dfrac{1}{a^m} \dfrac{1}{b^m} =$
по определению степени с отрицательным показателем	$= a^{-m} b^{-m} = a^\alpha b^\alpha$.

Следовательно, свойство а) справедливо. Свойство б) доказывается аналогично.

Докажем свойство в).

Пусть α и β — натуральные числа, тогда свойство в) уже доказано ранее (см. гл. I).

Пусть $\alpha = n$, $\beta = -m$, где n и m натуральные числа, тогда по определению степени с целым отрицательным показателем

$$a^\alpha a^\beta = a^n \frac{1}{a^m}.$$

Применяя правило умножения дробей, имеем $a^n \dfrac{1}{a^m} = \dfrac{a^n}{a^m}$.

Пусть $n > m$, тогда, применяя свойство степени с натуральным показателем, получаем

$$\frac{a^n}{a^m} = a^n : a^m = a^{n-m} = a^{n+(-m)} = a^{\alpha+\beta}.$$

Пусть $n = m$, тогда по определению степени с нулевым показателем получаем

$$\frac{a^n}{a^m} = 1 = a^0 = a^{n-m} = a^{n+(-m)} = a^{\alpha+\beta}.$$

Пусть $n < m$, тогда, применяя свойство степени с натуральным показателем и определение степени с целым отрицательным показателем, получаем

$$\frac{a^n}{a^m} = \frac{1}{a^m \frac{1}{a^n}} = \frac{1}{a^m \cdot a^{-n}} = \frac{1}{a^{m-n}} = a^{-(m-n)} = a^{-m+n} = a^{n+(-m)} = a^{\alpha+\beta}.$$

Пусть $\alpha = n$, n — натуральное число, $\beta = 0$, тогда, применяя определение степени с нулевым показателем, получаем

$$a^\alpha a^\beta = a^n \cdot 1 = a^n = a^{n+0} = a^{\alpha+\beta}.$$

Пусть $\alpha = 0$, β — любое целое число, тогда, применяя определение степени с нулевым показателем, получаем

$$a^\alpha a^\beta = 1 \cdot a^\beta = a^\beta = a^{0+\beta} = a^{\alpha+\beta}.$$

Пусть $\alpha = -n$, $\beta = m$, где n и m — натуральные числа. Этот случай аналогичен случаю $\alpha = n$, $\beta = -m$.

Пусть, наконец, $\alpha = -n$, $\beta = -m$, где n, m — натуральные числа.

Тогда	$a^\alpha a^\beta = a^{-n} a^{-m} =$
по определению степени с целым отрицательным показателем	$= \frac{1}{a^n} \frac{1}{a^m} =$
по правилу перемножения дробей	$= \frac{1}{a^n a^m} =$
по свойству степени с натуральным показателем	$= \frac{1}{a^{n+m}} =$
по определению степени с целым отрицательным показателем	$= a^{-(n+m)} =$
	$= a^{(-n)+(-m)} = a^{\alpha+\beta}.$

Следовательно, свойство в) справедливо. Свойство г) доказывается аналогично.

Докажем свойство д).

Пусть α и β — натуральные числа, тогда справедливость свойства д) уже доказана (см. гл. I).

Пусть α — любое целое число, $\beta = 0$, тогда по определению степени с нулевым показателем $(a^\alpha)^\beta = (a^\alpha)^0 = 1 = a^0 = a^{\alpha \cdot 0} = a^{\alpha\beta}$.

Пусть $\alpha = n$, $\beta = -m$, где n и m — натуральные числа.

Тогда	$(a^{\alpha})^{\beta}=(a^{n})^{-m}=$
по определению степени с целым отрицательным показателем	$=\dfrac{1}{(a^{n})^{m}}=$
по определению степени с натуральным показателем	$=\dfrac{1}{a^{nm}}=$
по определению степени с целым отрицательным показателем	$=a^{-(nm)}=$
	$=a^{n(-m)}=a^{\alpha\beta}.$

Пусть $\alpha=-n$, $\beta=m$, где n и m — натуральные числа. Справедливость этого свойства доказывается аналогично случаю $\alpha=n$, $\beta=-m$.

Пусть, наконец, $\alpha=-n$, $\beta=-m$, где n и m — натуральные числа.

Тогда	$(a^{\alpha})^{\beta}=(a^{-n})^{-m}=$
по определению степени с целым отрицательным показателем	$=\dfrac{1}{(a^{-n})^{m}}=$
по только что разобранному случаю	$=\dfrac{1}{a^{-nm}}=$
по определению степени с целым отрицательным показателем	$=\dfrac{1}{\dfrac{1}{a^{nm}}}=$
по свойству дробей	$=1:\dfrac{1}{a^{nm}}=$
по правилу деления дробей	$=a^{nm}=$
	$=a^{(\;n)(-m)}=a^{\alpha\beta}.$

Следовательно, свойство д) справедливо.

§ 2
АРИФМЕТИЧЕСКИЕ И АЛГЕБРАИЧЕСКИЕ КОРНИ

В § 4 гл. I было дано следующее определение арифметического корня из положительного числа.

Пусть n — натуральное число, a — положительное число. Тогда положительное число b, такое, что $b^{n}=a$, называется *арифметическим корнем* n-й степени из числа a и обозначается $b=\sqrt[n]{a}$.

Без доказательства принимается утверждение, что для каждого положительного числа a существует и притом единственный арифметический корень n-й степени.

По определению $\sqrt[n]{a}$ справедливо следующее утверждение:

$$\sqrt[n]{a}\Longleftrightarrow\begin{cases} a\text{ — положительное число,}\\ n\text{ — натуральное число,}\\ \sqrt[n]{a}\text{ — положительное число,}\\ \left(\sqrt[n]{a}\right)^{n}=a.\end{cases}$$

Поставим задачу: для любого действительного числа a и любого натурального числа n найти действительные числа b, такие, что $b^n = a$. Если такие числа существуют, то они называются *действительными алгебраическими корнями* n-степени из действительного числа a. Если число a положительное, то, как говорилось выше, существует одно положительное число b, такое, что $b^n = a$, т. е. для любого положительного числа a всегда существует хотя бы один алгебраический корень b, для обозначения которого есть специальный символ $\sqrt[n]{a}$ (арифметический корень). В случае существования других алгебраических корней, кроме арифметического корня, нет специального символа для их обозначения.

Рассмотрим вопрос существования алгебраического корня из действительного числа. Заметим, что утверждение о количестве действительных корней для данного действительного числа принимается без доказательства. Их справедливость будет вытекать из общей теоремы о количестве корней из комплексного числа (гл. X).

Пусть $a = 0$, тогда для любого натурального числа n существует и притом только один алгебраический корень n-й степени — число b, равное 0.

Пусть a — положительное число и n — нечетное натуральное число ($n = 2k+1$). Тогда существует и притом только один арифметический корень $b_1 = \sqrt[2k+1]{a}$ из этого числа и других вещественных алгебраических корней из этого числа нет. Таким образом, существует только один алгебраический корень нечетной степени из положительного числа, а именно арифметический корень.

Пусть a — положительное число и n — четное натуральное число ($n = 2k$). Тогда существует и притом только один арифметический корень $b_1 = \sqrt[2k]{a}$ и один вещественный алгебраический корень $b_2 = -\sqrt[2k]{a}$ из этого числа. Таким образом, существуют два вещественных алгебраических корня четной степени из положительного числа a: $b_1 = \sqrt[2k]{a}$ и $b_2 = -\sqrt[2k]{a}$.

Пусть a — отрицательное число и n — четное натуральное число ($n = 2k$). Поскольку любое не равное нулю действительное число в четной степени есть положительное число, а число 0 в любой натуральной степени есть нуль, то нет ни одного действительного числа b, такого, что b^{2k} — отрицательное число. Значит, нет действительного алгебраического корня четной степени из отрицательного числа.

Пусть a — отрицательное число и n — нечетное натуральное число ($n = 2k+1$). Покажем, что есть одно действительное отрицательное число b, такое, что $b^n = a$. Обозначим $c = -a$. Тогда $c > 0$ и потому существует единственный арифметический корень d степени $2k+1$ из числа c: $d^{2k+1}=c$ или $d=\sqrt[2k+1]{c}=$ $=\sqrt[2k+1]{-a}=\sqrt[2k+1]{|a|}$. Положим теперь $b=-d$. Тогда $b^{2k+1}=$ $=(-1)^{2k+1}d^{2k+1}=(-1)\,c=(-1)(-a)=a$. Значит, $b=-\sqrt[2k+1]{|a|}$ есть отрицательное число, такое, что $b^{2k+1}=a$, т. е. b — вещественный алгебраический корень из отрицательного числа a.

Пример. Пусть $a=-7$, $n=5$, тогда вещественный алгебраический корень 5-й степени из числа -7 есть число $b=$ $=-\sqrt[5]{|-7|}=-\sqrt[5]{7}$. Пусть $a=-8$, $n=3$, тогда вещественный алгебраический корень 3-й степени из числа -8 есть число $b=-\sqrt[3]{|-8|}=-\sqrt[3]{8}=-2$.

Замечание. Иногда корень нечетной степени из отрицательного числа a записывают в виде $b=\sqrt[2k+1]{a}$, понимая под этим число $b=-\sqrt[2k+1]{|a|}$, например, вместо $b=-\sqrt[5]{7}$ пишут $b=$ $=\sqrt[5]{-7}$, понимая под этим $b=-\sqrt[5]{7}$, но так делать не следует, т. к. выражение $\sqrt[2k+1]{a}$ определено только для положительных чисел.

Для корней справедливы следующие свойства. Пусть a и b — положительные числа, n и k — натуральные числа, тогда:

а) $\left(\sqrt[n]{a}\right)^n=a$; б) $\left(\sqrt[n]{a}\right)^k=\sqrt[n]{a^k}$;

в) $\sqrt[nk]{a^k}=\sqrt[n]{a}$; г) $\sqrt[k]{\sqrt[n]{a}}=\sqrt[nk]{a}$;

д) $\sqrt[n]{ab}=\sqrt[n]{a}\,\sqrt[n]{b}$; е) $\sqrt[n]{\dfrac{a}{b}}=\dfrac{\sqrt[n]{a}}{\sqrt[n]{b}}$.

Справедливость этих свойств будет вытекать из справедливости свойств степени с рациональными показателями, доказываемых в следующем параграфе.

§ 3
СТЕПЕНЬ С РАЦИОНАЛЬНЫМ ПОКАЗАТЕЛЕМ

Пусть a — положительное число, $r=p/q$ — рациональное число, причем q — натуральное число ($q>0$). Положительное число b, такое, что $b=\sqrt[q]{a^p}$, называется *r-й степенью* числа a и обозначается $b=a^r$, т. е. $a^{p/q}=\sqrt[q]{a^p}$. Заметим, что $\sqrt[n]{a}=a^{\frac{1}{n}}$ $(a>0)$. Если a и b — любые положительные числа, то для рациональной степени положительных чисел справедливы следующие свойства:

а) $(ab)^r=a^rb^r$;

б) $\left(\dfrac{a}{b}\right)^r=\dfrac{a^r}{b^r}$;

в) $a^{r_1}a^{r_2}=a^{r_1+r_2}$,

г) $a^{r_1}:a^{r_2}=a^{r_1-r_2}$;

д) $(a^{r_1})^{r_2}=a^{r_1r_2}$;

Докажем справедливость этих свойств.

а) Пусть $r=p/q$, причем q — натуральное число.

Рассмотрим	$[(ab)^{p/q}]^q=$
по определению рациональной степени	$=\left[\sqrt[q]{(ab)^p}\right]^q=$
по определению арифметического корня	$=(ab)^p=$
по свойству степени с целым показателем	$=a^pb^p=$
по определению арифметического корня	$=\left(\sqrt[q]{a^p}\right)^q\left(\sqrt[q]{b^p}\right)^q=$
по определению рациональной степени	$=(a^{p/q})^q(b^{p/q})^q=$
по свойству степени с натуральным показателем	$=(a^{p/q}b^{p/q})^q$.

Итак, $[(ab)^r]^q=[a^rb^r]^q$. По свойству 26 равенств (см. § 2 гл. II) алгебраических выражений это равенство равносильно равенству $(ab)^r=a^rb^r$ и свойство а) доказано. Свойство б) доказывается аналогично.

в) Пусть $r_1=p/q$, $r_2=m/n$. Тогда $a^{r_1}a^{r_2}=a^{p/q}a^{m/n}$

Рассмотрим	$[a^{p/q}a^{m/n}]^{qn}=$
по свойству степени с натуральным показателем	$=(a^{p/q})^{qn}(a^{m/n})^{qn}=$
по свойству степени с натуральным показателем	$=[(a^{p/q})^q]^n[(a^{m/n})^n]^q=$
по определению рациональной степени	$=\left[\left(\sqrt[q]{a^p}\right)^q\right]^n\left[\left(\sqrt[n]{a^m}\right)^n\right]^q=$
по определению арифметического корня	$=[a^p]^n[a^m]^q=$
по свойству степени с целым показателем	$=a^{pn}a^{mq}=$
по свойству степени с целым показателем	$=a^{pn+mq}=$
по определению арифметического корня	$=\left(\sqrt[nq]{a^{pn+mq}}\right)^{nq}=$
по определению рациональной степени	$=\left(a^{\frac{pn+mq}{nq}}\right)^{nq}$.

Итак, учитывая, что $(pn+mq)/nq=r_1+r_2$, имеем $(a^{r_1}a^{r_2})^{qn}=$ $=(a^{r_1+r_2})^{qn}$. По свойству 26 (см. § 2 гл. II) равенств алгебраических выражений это равенство равносильно равенству $a^{r_1}a^{r_2}=$ $=a^{r_1+r_2}$ и свойство в) доказано. Свойство г) доказывается аналогично.

д) Пусть $r_1=p/q$, $r_2=m/n$. Тогда $(a^{r_1})^{r_2}=(a^{p/q})^{m/n}$.

Рассмотрим	$[(a^{p/q})^{m/n}]^{nq}=$
по свойству степени с натуральным показателем	$=\{[(a^{p/q})^{m/n}]^n\}^q=$
по определению степени с рациональным показателем	$=\left\{\left[\sqrt[n]{(a^{p/q})^m}\right]^n\right\}^q=$
по определению арифметического корня	$=\{(a^{p/q})^m\}^q=$

по свойству степени с целым показателем $= (a^{p/q})^{mq} =$
по свойству степени с целым показателем $= [(a^{p/q})^q]^m =$
по определению степени с рациональным показателем $= \left[\left(\sqrt[q]{a^p}\right)^q\right]^m =$
по определению арифметического корня $= [a^p]^m =$
по свойству степени с целым показателем $= a^{pm} =$
по определению арифметического корня $= \left(\sqrt[qn]{a^{pm}}\right)^{qn} =$
по определению степени с рациональным показателем $= (a^{pn/qn})^{qn}$.

Итак, $[(a^{r_1})^{r_2}]^{nq} = (a^{r_1 r_2})^{nq}$. По свойству 26 (см. § 2 гл. II) (из справедливости этого равенства вытекает справедливость свойства д).

Докажем еще несколько свойств степени с рациональным показателем.

е) Пусть a — положительное число, $r = p/q$ — рациональное число, q и n — натуральные числа. Тогда $a^{p/q} = a^{pn/qn}$.

Рассмотрим $(a^{p/q})^{qn} =$
по свойству степени с натуральным показателем $= [(a^{p/q})^q]^n =$
по определению степени с рациональным показателем $= \left[\left(\sqrt[q]{a^p}\right)^q\right]^n =$
по определению арифметического корня $= (a^p)^n =$
по свойству степени с целым показателем $= a^{pn} =$
по определению арифметического корня $= \left(\sqrt[qn]{a^{pn}}\right)^{qn} =$
по определению степени с рациональным показателем $= (a^{pn/qn})^{qn}$.

Итак, $(a^{p/q})^{qn} = (a^{pn/qn})^{qn}$, откуда и вытекает справедливость свойства е).

ж) Пусть $a > 1$, $r = p/q$ — положительное рациональное число ($p > 0$, $q > 0$). Тогда $a^r > 1$.

Рассмотрим $(a^{p/q})^q =$
по определению степени с рациональным показателем $= \left(\sqrt[q]{a^p}\right)^q =$
по определению арифметического корня $= a^p$.

По свойству 26 (см. § 2 гл. II) условия $a > 1$ и $a^p > 1^p$ равносильны, значит, из условия $a > 1$ вытекает, что $a^p > 1$, но тогда $(a^{p/q})^q > 1$, т. е. $(a^r)^q > 1^q$, откуда по свойству 26 получаем, что $a^r > 1$. Свойство ж) доказано.

з) Пусть $0 < a < 1$, $r = p/q$ — положительное число ($p > 0$, $q > 0$). Тогда $a^r < 1$.

Доказательство этого свойства аналогично доказательству свойства ж).

и) Пусть $a > 1$ и r_1 и r_2 — рациональные числа, такие, что $r_1 > r_2$. Тогда $a^{r_1} > a^{r_2}$.

Доказательство. Поскольку $(r_1 - r_2)$ — рациональное положительное число, то по свойству ж) $a^{r_1 - r_2} > 1$. Умножая это неравенство на положительное число a^{r_2}, по свойству 24 (см. § 2 гл. II) получаем $a^{r_2}(a^{r_1-r_2})>a^{r_2}$. Применяя к левой части свойство в) степени с рациональным показателем, получаем $a^{r_1}>a^{r_2}$, т. е. свойство и) доказано.

к) Пусть $0<a<1$, r_1, r_2 — рациональные числа, такие, что $r_1>r_2$. Тогда $a^{r_1}<a^{r_2}$.

Доказательство этого свойства аналогично доказательству свойства и).

§ 4
СТЕПЕНЬ С ИРРАЦИОНАЛЬНЫМ ПОКАЗАТЕЛЕМ

Возьмем приближенные значения числа $\sqrt{2}$ с недостатком:

$$1<1{,}4<1{,}41<1{,}414<\ldots$$

и приближенные значения числа $\sqrt{2}$ с избытком:

$$2>1{,}5>1{,}42>1{,}415>\ldots.$$

По свойству и) степени с рациональным показателем

$$3^1<3^{1{,}4}<3^{1{,}41}<3^{1{,}414}<\ldots \tag{1}$$

и

$$3^2>3^{1{,}5}>3^{1{,}42}>3^{1{,}415}>\ldots. \tag{2}$$

В последовательностях (1) и (2) членов бесконечно много и любой член последовательности (1) меньше любого члена последовательности (2). Естественно под числом $3^{\sqrt{2}}$ понимать число, которое больше любого члена последовательности (1) и меньше любого члена последовательности (2), т. е. под числом $3^{\sqrt{2}}$ понимается число, которое больше, чем 3 в любой рациональной степени, приближающей $\sqrt{2}$ с недостатком, и которое меньше, чем 3 в любой рациональной степени приближающей $\sqrt{2}$ с избытком. Без доказательства принимается, что такое число существует и притом только одно.

Так же определяется и a^{α}, где $a>1$, α — иррациональное число, $\alpha>0$. Находятся рациональные числа r_i, приближающие число α с недостатком: $r_1<r_2<r_3<\ldots<\alpha$, затем — числа l_k, приближающие число α с избытком: $l_1>l_2>l_3>\ldots>\alpha$, и составляются две последовательности: $a^{r_1}<a^{r_2}<a^{r_3}<\ldots$ и $a^{l_1}>a^{l_2}>a^{l_3}>\ldots$. Под a^{α} понимается число, которое больше любого числа a^{r_i} и меньше любого числа a^{l_k}.

Пусть дано число $a>1$ и положительное иррациональное число α. Через r_i обозначим рациональные числа, приближающие α с недостатком, через l_k — с избытком. Под числом a^{α} понимается действительное число a^{α}, такое, что для любых r_i и l_k справедливо неравенство $a^{r_i}<a^{\alpha}<a^{l_k}$. Без доказательства принимается, что такое число существует и притом только одно.

Пусть дано число a, такое, что $0 < a < 1$ и положительное иррациональное число α. Через r_i обозначим рациональные числа, приближающие α с недостатком, через l_k — с избытком. Под числом a^α понимается действительное число a^α, такое, что для любых r_i и l_k справедливо неравенство $a^{r_i} > a^\alpha > a^{l_k}$. Без доказательства принимается, что такое число существует и притом только одно.

Пусть дано положительное вещественное число a, такое, что $a \neq 1$ и отрицательное иррациональное число α. Под числом a^α понимается действительное число, равное $1/a^{|\alpha|}$, т. е. если $a \neq 1$ и α — отрицательное иррациональное число, то $a^\alpha = 1/a^{-\alpha}$. Поскольку число $a^{|\alpha|}$ не равно нулю и в множестве действительных чисел действие деление всегда выполнимо, то существует и притом единственное число, равное частному $1/a^{|\alpha|}$. Это число называют числом a^α.

Замечание. 1. Если $a = 1$, то для любого вещественного числа $\alpha : a^\alpha = 1$. Поэтому в вышеприведенных определениях случай $a = 1$ не рассматривался.

2. В силу вышеприведенных определений и определения степени с рациональным показателем для $a > 0$ и любого вещественного числа α число a^α всегда положительно.

Для степеней с иррациональным показателем справедливы следующие свойства. Пусть $a > 0$, $b > 0$, α — иррациональное число, β — рациональное или иррациональное, тогда:

а) $(ab)^\alpha = a^\alpha b^\alpha$;

б) $\left(\frac{a}{b}\right)^\alpha = \frac{a^\alpha}{b^\alpha}$;

в) $a^\alpha a^\beta = a^{\alpha+\beta}$;

г) $a^\alpha : a^\beta = a^{\alpha-\beta}$;

д) $(a^\alpha)^\beta = a^{\alpha\beta}$.

Доказательство этих свойств проводится с помощью теории пределов и поэтому здесь опускается.

§ 5
СТЕПЕНЬ ПОЛОЖИТЕЛЬНОГО ЧИСЛА

Вышеизложенные § 1 — 4 позволяют определить любую действительную степень положительного числа. Заметим, что число a^α существует и притом только одно для любого действительного α.

Пусть дано положительное число a и действительное число α. Под числом a^α понимают положительное число, определяемое по следующему правилу.

I. Если $\alpha > 0$ и:

1. $\alpha = m$, m — натуральное число, то $\underbrace{a^\alpha = aa \ldots a}_{m \text{ раз}}$;

2. $\alpha = 1/q$, q — натуральное число, то $a^\alpha = \sqrt[q]{a}$ (арифметический корень q-й степени из положительного числа);

3. $\alpha=p/q$, p, q — натуральные числа, то $a^\alpha=\sqrt[q]{a^p}$;

4. α — иррациональное число, то:

а) если $a>1$, то a^α — число большее, чем a^{r_i} и меньшее, чем a^{l_k}, где r_i — любое рациональное приближение α с недостатком, l_k — любое рациональное приближение α с избытком;

б) если $0<a<1$, то a^α — число меньшее, чем a^{r_i}, и большее, чем a^{l_k} (r_i и l_k — те же, что и выше);

в) если $a=1$, то $a^\alpha=1$.

II. Если $\alpha=0$, то $a^\alpha=1$.

III. Если $\alpha<0$, то $a^\alpha=1/a^{|\alpha|}$.

Число a^α *называется степенью*, число a — *основанием степени*, число α — *показателем степени*.

Из § 1—4 вытекает, что степень положительного числа обладает следующими основными свойствами: если a и b — положительные числа, α и β — любые действительные числа, то:

а) $(ab)^\alpha=a^\alpha b^\alpha$;

б) $(a/b)^\alpha=a^\alpha/b^\alpha$;

в) $a^\alpha a^\beta=a^{\alpha+\beta}$;

г) $a^\alpha : a^\beta=a^{\alpha-\beta}$;

д) $(a^\alpha)^\beta=a^{\alpha\beta}$.

Докажем еще несколько свойств степеней.

Теорема 1. Если $a>1$ и $\alpha>0$, то $a^\alpha>1$.

Доказательство. Если $\alpha=p/q$ — рациональное число (p и q — натуральные числа), то свойство $a^\alpha>1$ уже доказано в § 3. Если α — иррациональное число, то возьмем любое положительное рациональное число r, приближающее α с недостатком, тогда $a^\alpha>a^r$ по определению иррациональной степени. В то же время по уже доказанному в § 3 свойству $a^r>1$. По свойству транзитивности неравенств из справедливости двух неравенств $a^\alpha>a^r$ и $a^r>1$ вытекает справедливость неравенства $a^\alpha>1$. Теорема 1 доказана.

Теорема 2. Если $a>1$ и $\alpha<0$, то $a^\alpha<1$.

Доказательство. Число $\beta=-\alpha$ — положительное число, поэтому по теореме 1 $\alpha^\beta>1$. Умножим обе части этого неравенства на положительное число a^α. По свойству неравенств $a^\beta a^\alpha>a^\alpha$; по свойству в) и определению степеней $a^\beta a^\alpha=a^{\alpha+\beta}=a^0=1$, т. е. $a^\alpha<1$ и теорема 2 доказана.

Теорема 3. Если $a>1$ и $a^\alpha>1$, то $\alpha>0$.

Доказательство. Предположим, что $a>1$ и $a^\alpha>1$, но $\alpha\leqslant 0$, т. е. либо $\alpha=0$, либо $\alpha<0$. Если $\alpha=0$, то $a^\alpha=1$ по определению. Если $\alpha<0$ и $a>1$, то по теореме 2 $a^\alpha<1$. Итак, если $\alpha\leqslant 0$, то $a^\alpha\leqslant 1$, что противоречит предположению $a^\alpha>1$. Теорема доказана.

Теорема 4. Если $a>1$ и $a^\alpha<1$, то $\alpha<0$.

Доказательство теоремы 4 аналогично доказательству теоремы 3. Обобщим теоремы 1—4.

Утверждение 1. Если $a>1$, то условия $a^\alpha>1$ и $\alpha>0$ равносильны; кроме того, равносильны условия $a^\alpha<1$ и $\alpha<0$, т. е., если $a>1$, то

$$a^\alpha>1 \Leftrightarrow \alpha>0,$$
$$a^\alpha<1 \Leftrightarrow \alpha<0.$$

Теорема 5. Если $0<a<1$ и $\alpha>0$, то $a^\alpha<1$.

Доказательство. Рассмотрим число $b=1/a$. Так как $b>1$, то по теореме 1 $b^\alpha>1$. Умножим обе части этого неравенства на положительное число a^α. По свойству неравенств $b^\alpha a^\alpha>a^\alpha$. По свойствам степеней $b^\alpha a^\alpha=(ab)^\alpha=1^\alpha=1$, т. е. $a^\alpha<1$.

Теорема 6. Если $0<a<1$ и $\alpha<0$, то $a^\alpha>1$.

Теорема 7. Если $0<a<1$ и $a^\alpha>1$, то $\alpha<0$.

Теорема 8. Если $0<a<1$ и $a^\alpha<1$, то $\alpha>0$.

Доказательство этих теорем аналогично доказательству теоремы 5. Обобщим теоремы 5—8.

Утверждение 2. Если $0<a<1$, то условие $a^\alpha>1$ и $\alpha<0$ равносильны, кроме того, равносильны условия $a^\alpha<1$ и $\alpha>0$, т. е. если $0<a<1$, то

$$a^\alpha>1 \Leftrightarrow \alpha<0,$$
$$a^\alpha<1 \Leftrightarrow \alpha>0.$$

Из утверждений 1 и 2 легко получить следующее следствие.

Утверждение 3. Если $a>0$ и $a\neq 1$, то условия $a^\alpha=1$ и $\alpha=0$ равносильны, т. е. если $a>0$ и $a\neq 1$, то

$$a^\alpha=1 \Leftrightarrow \alpha=0.$$

Теорема 9. Если $a>1$ и $\alpha_1>\alpha_2$, то $a^{\alpha_1}>a^{\alpha_2}$.

Доказательство. Рассмотрим число $\beta=\alpha_1-\alpha_2$. Так как $\beta>0$, то по теореме 1 $a^\beta>1$. Умножим обе части этого неравенства на положительное число a^{α_2}. По свойству неравенств $a^\beta a^{\alpha_2}>a^{\alpha_2}$; по свойству степеней $a^\beta a^{\alpha_2}=a^{\alpha_1}$, т. е. $a^{\alpha_1}>a^{\alpha_2}$ и теорема 9 доказана.

Теорема 10. Если $a>1$ и $\alpha_1<\alpha_2$, то $a^{\alpha_1}<a^{\alpha_2}$.

Теорема 11. Если $a>1$ и $a^{\alpha_1}>a^{\alpha_2}$, то $\alpha_1>\alpha_2$.

Теорема 12. Если $a>1$ и $a^{\alpha_1}<a^{\alpha_2}$, то $\alpha_1<\alpha_2$.

Доказательство этих теорем аналогично доказательству теоремы 9. Объединим теоремы 9—12.

Утверждение 4. Если $a>1$, то условия $a^{\alpha_1}>a^{\alpha_2}$ и $\alpha_1>\alpha_2$ равносильны, кроме того, равносильны условия $a^{\alpha_1}<a^{\alpha_2}$ и $\alpha_1<\alpha_2$, т. е. если $a>1$, то

$$a^{\alpha_1}>a^{\alpha_2} \Leftrightarrow \alpha_1>\alpha_2,$$
$$a^{\alpha_1}<a^{\alpha_2} \Leftrightarrow \alpha_1<\alpha_2.$$

Теорема 13. Если $0<a<1$ и $\alpha_1>\alpha_2$, то $a^{\alpha_1}<a^{\alpha_2}$.

Теорема 14. Если $0<a<1$ и $\alpha_1<\alpha_2$, то $a^{\alpha_1}>a^{\alpha_2}$.

Теорема 15. Если $0<a<1$ и $a^{\alpha_1}>a^{\alpha_2}$, то $\alpha_1<\alpha_2$.

Теорема 16. Если $0 < a < 1$ и $a^{\alpha_1} < a^{\alpha_2}$, то $\alpha_1 > \alpha_2$.

Доказательство этих теорем аналогично доказательству теоремы 9.

Объединим теоремы 13 — 16.

Утверждение 5. Если $0 < a < 1$, то условия $a^{\alpha_1} > a^{\alpha_2}$ и $\alpha_1 < \alpha_2$ равносильны, кроме того, равносильны условия $a^{\alpha_1} < a^{\alpha_2}$ и $a_1 > \alpha_2$, т. е. если $0 < a < 1$, то

$$a^{\alpha_1} > a^{\alpha_2} \Leftrightarrow \alpha_1 < \alpha_2,$$
$$a^{\alpha_1} < a^{\alpha_2} \Leftrightarrow \alpha_1 > \alpha_2.$$

Из утверждений 4 и 5 легко получить следующее следствие.

Утверждение 6. Если $a > 0$ и $a \neq 1$, то условия $a^{\alpha_1} = a^{\alpha_2}$ и $\alpha^1 = \alpha_2$ равносильны, т. е. если $a > 0$ и $a \neq 1$, то

$$a^{\alpha_1} = a^{\alpha_2} \Leftrightarrow \alpha_1 = \alpha_2.$$

Утверждения 4 и 5 словесно формулируются следующим образом.

Если основание степени больше 1, то большему показателю соответствует большая степень и наоборот, большей степени соответствует больший показатель.

Если основание степени меньше 1 $(0 < a < 1)$, то большему показателю соответствует меньшая степень и наоборот, меньшей степени соответствует больший показатель.

Замечание. Если $\alpha > 0$, то понятие операции возведения в степень можно расширить на множество всех неотрицательных чисел: по определению $0^\alpha = 0$, если $\alpha > 0$.

§ 6
ЛОГАРИФМЫ

Рассмотрим основные задачи, которые возникают при изучении степеней.

1. Даны действительные числа a и α. Найти действительное число x, такое, что $x = a^\alpha$. Это — задача о возведении положительного числа в степень. Она разрешима для любого положительного числа a и любого действительного числа α. Если $a = 0$, $\alpha > 0$, то $x = 0$ (см. § 5 гл. IV). При $a < 0$ эта задача здесь рассматриваться не будет.

2. Даны действительные числа b и α. Найти действительное число x, такое, что $x^\alpha = b$.

Если b — любое положительное число и α — любое отличное от нуля действительное число, то эта задача сводится к предыдущей, ибо ответ дает число $x = b^{1/\alpha}$. Действительно, $x^\alpha = (b^{1/\alpha})^\alpha = b^{\frac{1}{\alpha}\alpha} = b^1 = b$. Если $\alpha = 0$ и $b = 1$, то решением этой задачи является любое действительное отличное от нуля число x. Если $\alpha = 0$ и $b \neq 1$, то эта задача не имеет решения. При $b \leqslant 0$ эта задача здесь рассматриваться не будет.

3. Даны действительные числа a и b, найти действительное число x, такое, что $a^x=b$. Будем рассматривать эту задачу только для действительных положительных чисел a и b. Если $a=1$ и $b=1$, то решением этой задачи является любое действительное число x. Если $a=1$ и $b\neq 1$, то задача не имеет решения. Рассмотрим случай $a\neq 1$.

Теорема 1. Для любой пары действительных чисел a и b, таких, что $a>0$, $a\neq 1$ и $b>0$, существует и притом только одно действительное число x, такое, что $a^x=b$.

Доказательство существования такого числа x сложно и поэтому здесь опускается. Докажем единственность. Предположим, что существуют действительные числа x_1 и x_2, такие, что $a^{x_1}=b$ и $a^{x_2}=b$. По свойству транзитивности равенств $a^{x_1}=a^{x_2}$. На основании утверждения 6 (см. § 5) $x_1=x_2$, что и требовалось доказать.

Действительное число α называется *логарифмом числа b по основанию a* ($a>0$, $a\neq 1$, $b>0$) и обозначается $\alpha=\log_\alpha b$, если $a^\alpha=b$.

Заметим, что определение логарифма можно дать только после доказательства теоремы 1, так как до нее было неясно, существует ли такое число α, что $a^\alpha=b$, и единственно ли оно. Подчеркнем еще раз, что логарифм определяется только для положительного числа по положительному не равному единице основанию, т. е. для любого $a\leqslant 0$, $a=1$ и любого $b\leqslant 0$ понятие логарифма лишено смысла: утверждение, что число 3 является логарифмом числа -8 по основанию -2 лишено смысла. Итак, в определении логарифма $\log_a b$ всегда $a>0$, $a\neq 1$, $b>0$. Из определения логарифма вытекает *основное логарифмическое тождество* $a^{\log_a b}=b$. Пользуясь определением логарифма, получаем, что $\log_a a=1$ и $\log_a 1=0$. Используя единственность логарифма, имеем, что если $\mu>0$ и $\mu\neq 1$, то всегда $\log_a\mu\neq 0$. Перейдем к рассмотрению основных свойств логарифмов.

Пусть числа M, N, a, b, α и β таковы, что $M>0$, $N>0$, $a>0$, $b>0$, $a\neq 1$, $b\neq 1$, α и β — любые действительные числа ($\beta\neq 0$). Тогда:

а) $\log_a MN=\log_a M+\log_a N$;

б) $\log_a \frac{M}{N}=\log_a M-\log_a N$;

в) $\log_a M^\alpha=\alpha\log_a M$;

г) $\log_{a^\beta} M^\alpha=\frac{\alpha}{\beta}\log_a M$;

д) $\log_b M=\frac{\log_a M}{\log_a b}$ (правило перехода от логарифмов по одному основанию к логарифмам по другому основанию);

е) $\log_a M=\log_a N\Leftrightarrow M=N$;

ж) для $a>1$ $\log_a M<\log_a N\Leftrightarrow M<N$;

з) для $0<a<1$ $\log_a M<\log_a N\Leftrightarrow M>N$.

г$'$) $\log_{a^\alpha} M^\alpha = \log_a M$ для $\alpha \neq 0$:

д$'$) $\log_b a \log_a b = 1$.

Докажем эти свойства.

а) Рассмотрим $a^{\log_a MN} =$

по основному логарифмическому тождеству $= MN =$

по основному логарифмическому тождеству $= a^{\log_a M} a^{\log_a N} =$

по свойству степени положительного числа $= a^{\log_a M + \log_a N}$.

Итак, $a^{\log_a (MN)} = a^{\log_a M + \log_a N}$. Применяя к последнему равенству утверждение 6 свойств степеней, получим, что $\log_a MN = \log_a M + \log_a N$. Свойство б) доказывается аналогично.

в) Рассмотрим $a^{\log_a (M^\alpha)} =$

по основному логарифмическому тождеству $= M^\alpha = (a^{\log_a M})^\alpha =$

по свойству степени положительного числа $= a^{\alpha \log_a M}$.

Итак, $a^{\log_a (M^\alpha)} = a^{\alpha \log_a M}$. Применяя утверждение 6 свойств степеней, получим $\log_a (M^\alpha) = \alpha \log_a M$.

г) Рассмотрим $(a^\beta)^{\log_{a^\beta} (M^\alpha)} =$

по основному логарифмическому тождеству $= M^\alpha = a^{(\log_a M)\alpha} =$

по свойству степени положительного числа $= a^{\alpha \log_a M} =$

так как $\beta \neq 0$, то $1 = \beta \frac{1}{\beta}$, поэтому $= \left(a^{\beta \frac{1}{\beta}}\right)^{\alpha \log_a M} =$

по свойству степени положительного числа $= (a^\beta)^{\frac{\alpha}{\beta} \log_a M}$.

Итак, $(a^\beta)^{\log_{a^\beta} M^\alpha} = (a^\beta)^{\frac{\alpha}{\beta} \log_a M}$. Применяя к последнему равенству утверждение 6 свойств степеней, получим $\log_{a^\beta} M^\alpha = \frac{\alpha}{\beta} \log_a M$.

д) Рассмотрим $a^{\log_a M} =$

по основному логарифмическому тождеству $= M =$

по основному логарифмическому тождеству $= b^{\log_b M} =$

по основному логарифмическому тождеству $= (a^{\log_a b})^{\log_b M} =$

по свойству степени положительного числа $= a^{\log_a b \log_b M}$.

Итак, $a^{\log_a M} = a^{\log_a b \log_b M}$. Применяя к последнему равенству утверждение 6 свойств степеней, получим, что $\log_a M = \log_a b \log_b M$. По свойству равенств обе части этого равенства можно умножить на число $1/\log_a b$ (так как $b \neq 1$, то $\log_a b \neq 0$) и получить справедливость равенства

$$\log_b M = \frac{\log_a M}{\log_a b}.$$

е) По основному логарифмическому тождеству $M = a^{\log_a M}$ и $N = a^{\log_a N}$, следовательно,

$$M = N \Leftrightarrow a^{\log_a M} = a^{\log_a N}. \qquad (1)$$

По утверждению 6 свойств степеней имеем

$$a^{\log_a M} = a^{\log_a N} \Leftrightarrow \log_a M = \log_a N. \qquad (2)$$

Из (1) и (2) вытекает, что

$$M = N \Leftrightarrow \log_a M = \log_a N.$$

ж) По основному логарифмическому тождеству $M = a^{\log_a M}$ и $N = a^{\log_a N}$, следовательно,

$$M < N \Leftrightarrow a^{\log_a M} < a^{\log_a N}. \qquad (3)$$

На основании утверждения 4 свойств степеней имеем

$$a^{\log_a M} < a^{\log_a N} \Leftrightarrow \log_a M < \log_a N. \qquad (4)$$

Из (3) и (4) вытекает, что

$$M < N \Leftrightarrow \log_a M < \log_a N.$$

Свойство з) доказывается аналогично. Свойства ж) и з) имеют следующую словесную формулировку.

При основании, большем единицы, меньшему числу из двух положительных чисел соответствует меньший логарифм и меньшему логарифму соответствует меньшее число.

При основании, меньшем единицы, меньшему из двух положительных чисел соответствует больший логарифм и большему логарифму соответствует меньшее число.

Логарифмы по основанию 10 называются *десятичными логарифмами* и вместо обозначения $\log_{10} M$ употребляется обозначение $\lg M$.

Логарифмы по основанию e (e — иррациональное число, приближенное значение которого 2,7183 ...) называются *натуральными логарифмами* и вместо обозначения $\log_e N$ употребляется обозначение $\ln N$.

Свойства логарифмов используются для преобразования различных логарифмических выражений как с числами, так и с буквами.

Пример. Вычислить $A = \left(\sqrt[7]{\frac{1}{27}}\right)^{\frac{1}{5\log_5 3} + \log_{9\sqrt{3}} 125}$.

По свойству д') логарифмов $\frac{1}{5\log_5 3} = \frac{1}{5}\log_3 5$; по свойству д) логарифмов $\log_{9\sqrt{3}} 125 = \log_{3^{5/2}} 5^3 = \frac{6}{5}\log_3 5$. Используя свойства степеней и основное логарифмическое тождество, получим

$$A = \left(3^{-\frac{3}{7}}\right)^{\frac{1}{5}\log_3 5 + \frac{6}{5}\log_3 5} = 3^{-\frac{3}{5}\log_3 5} = \left(3^{\log_3 5}\right)^{-\frac{3}{5}} =$$

$$= 5^{-\frac{3}{5}} = \sqrt[5]{0{,}008}.$$

Глава V
ТРИГОНОМЕТРИЯ

§ 1
УГЛЫ И ИХ ИЗМЕРЕНИЕ

Пусть даны два луча, выходящие из точки O и занимающие одно и то же положение, т. е. два совпадающих луча OA и OB (рис. 9).

Рис. 9

Пусть луч OA поворачивается в плоскости вокруг точки O и после того, как он совершит некоторый поворот, займет фиксированное положение. Тогда для любого такого поворота луч OB считается *неподвижным* (начальным) лучом поворота, а луч OA — *подвижным*, совершившим данный поворот.

Любой поворот подвижного луча OA от неподвижного луча OB может быть совершен в двух противоположных направлениях (по часовой и против часовой стрелки). Если на подвижном луче OA приспособить пишущее устройство, равномерно удаляющееся от точки O вдоль луча OA, то при вращении луча OA оно оставит след на плоскости. После того как луч OA совершит некоторый поворот, этот след будет представлять раскручивающуюся вокруг точки поворота O кривую, которая начинается у неподвижного луча OB и оканчивается у подвижного луча OA. С помощью такой кривой на чертежах показывают повороты. При этом возле подвиж-

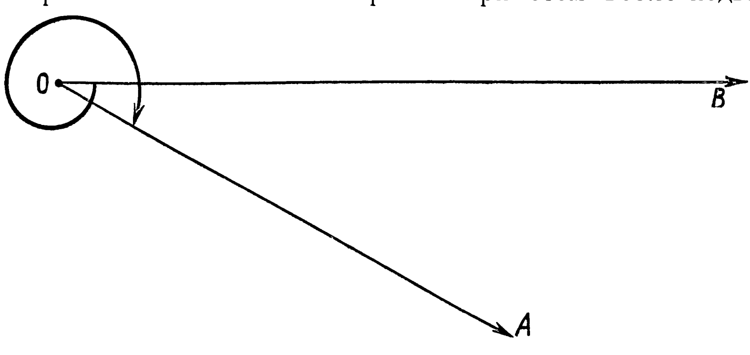


Рис. 10

ного луча кривая обязательно оканчивается стрелкой, указывающей направление совершенного поворота. На рис. 10 показан один из поворотов по часовой стрелке. На рис. 11 показан один из поворотов против часовой стрелки.

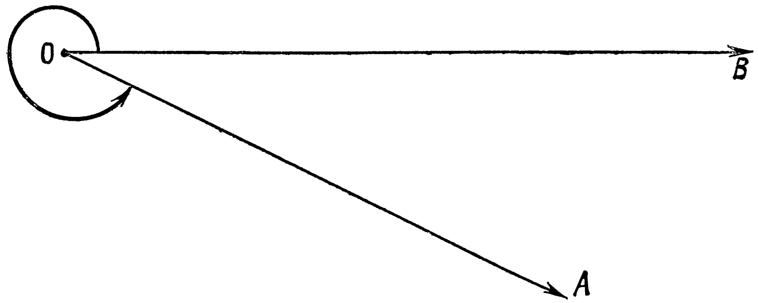


Рис. 11

Пусть подвижный луч OA совпадает с неподвижным лучом OB, не совершив поворота; тогда принято говорить, что подвижный луч OA совершил нулевой поворот (см. рис. 9).

Пусть подвижный луч OA совершил такой поворот по часовой стрелке, что луч OA впервые совпал с начальным лучом OB. Этот поворот принято называть *полным оборотом по часовой стрелке*. (рис. 12).

Рис. 12

Пусть подвижный луч OA совершил такой поворот против часовой стрелки, что луч OA впервые совпал с начальным лучом OB. Этот поворот принято называть *полным оборотом против часовой стрелки* (рис. 13).

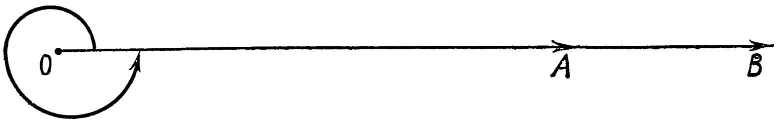


Рис. 13

На рис. 14 показан поворот, равный трем полным оборотам против часовой стрелки. На рис. 15 показан поворот, равный двум полным оборотам по часовой стрелке.

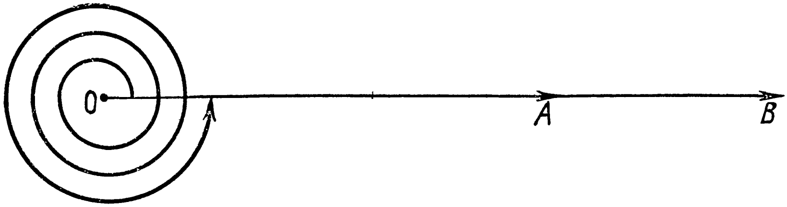


Рис. 14

Пусть подвижный луч OA совершил некоторый поворот в плоскости вокруг точки O от неподвижного луча OB. Тогда говорят, что подвижный луч OA задает угол α, соответствующий этому повороту. Точку O называют вершиной угла α,

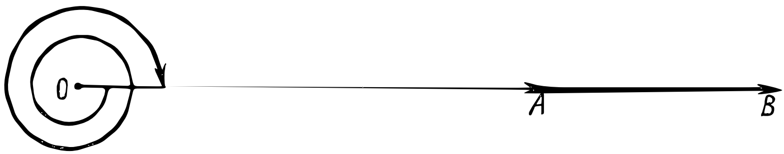


Рис. 15

неподвижный луч OB — началом отсчета угла α, подвижный луч OA — подвижным лучом, задающим угол α. Неподвижный луч OB (начало отсчета для любого угла) на чертежах принято располагать горизонтально вправо (правый горизонтальный луч). Принято считать: если подвижный луч совершит некоторый поворот против часовой стрелки, то тем самым образуется соответствующий *положительный угол*; если подвижный луч совершит некоторый поворот по часовой стрелке, то тем самым образуется соответствующий *отрицательный угол*; если подвижный луч совершил нулевой поворот, то считается, что он задает *нулевой угол*.

Например, если подвижный луч OA совершил полный оборот против часовой стрелки, то образуется *положительный полный угол*; если подвижный луч OA совершил полный оборот по часовой стрелке, то образуется *отрицательный полный угол*.

Пусть подвижный луч OA совершил поворот, равный $^1/_{360}$ части полного оборота против часовой стрелки. Тогда говорят, что подвижный луч OA задает угол в *один градус* (1°). Следовательно, положительный полный угол и угол в 360° один и тот же угол — угол, который задает подвижный луч OA, совершивший полный оборот против часовой стрелки (см.

рис. 13). Для частей угла в один градус приняты специальные наименования — *минута* и *секунда*.

Пусть подвижный луч OA совершил поворот, равный $^1/_{60}$ части поворота, соответствующего углу в один градус, против часовой стрелки. Тогда говорят, что подвижный луч OA задает угол в *одну минуту* $(1')$. Следовательно, угол в $60'$ и угол в 1° один и тот же угол.

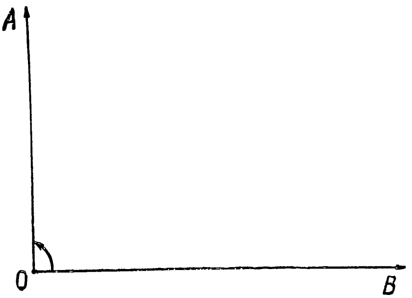


Рис. 16

Пусть подвижный луч OA совершил поворот, равный $^1/_{60}$ части поворота, соответствующего углу в одну минуту, против часовой стрелки. Тогда говорят, что подвижный луч OA задает угол в *одну секунду* $(1'')$. Следовательно, угол в $60''$ и угол в $1'$ один и тот же угол.

Пусть подвижный луч OA совершил поворот, равный $^1/_4$ части полного оборота против часовой стрелки. Тогда говорят, что подвижный луч OA задает *положительный прямой угол*, или угол в 90° (рис. 16). Луч OA, задающий угол в 90°, принято называть *верхним вертикальным лучом*.

Пусть подвижный луч OA совершил поворот, равный $^1/_2$ полного оборота против часовой стрелки. Тогда говорят, что подвижный луч OA задает *развернутый положительный* угол, или угол в 180° (рис. 17). Луч OA, задающий угол в 180°, принято называть *левым горизонтальным лучом*.

Рис. 17

Пусть подвижный луч OA совершил поворот, равный $^1/_4$ части полного оборота по часовой стрелке. Тогда подвижный луч OA задает *отрицательный прямой угол*, или угол в -90° (рис. 18). Луч OA, задающий угол в -90°, принято называть *нижним вертикальным лучом*.

Рассмотрим *радианное измерение углов*. Пусть подвижный луч OA совпадает с неподвижным лучом OB, не совершив поворота. Возьмем произвольную точку M на неподвижном луче OB и точку N подвижного луча OA, которая совпадает с точкой M. Проведем окружность с центром в точке O радиуса R, равным длине отрезка ON. Если подвижный луч OA будет вращаться вокруг точки O, то точка N будет двигаться по этой окружности (рис. 19).

Пусть подвижный луч OA совершил такой поворот против часовой стрелки, что точка N, двигаясь по окружности, прошла расстояние, равное радиусу этой окружности. Тогда говорят, что подвижный луч OA задает *угол в один радиан* (рис. 20).

При радианном измерении углов радианную меру угла определяют следующим образом. Пусть дан некоторый угол α, задаваемый подвижным лучом OA. *Радианной мерой угла* α называют число, равное отношению расстояния S, пройденного по окружности радиуса R точкой N подвижного луча OA, совершившего этот поворот, к радиусу. Отношение берется со знаком плюс, если поворот совершен против часовой стрелки (угол α положительный) и со знаком минус, если поворот совершен по часовой стрелке (угол α отрицательный). Итак, радианной мерой угла α называют число α, равное положительному числу S/R, если поворот совершен против часовой стрелки, равное отрицательному числу $(-S/R)$, если

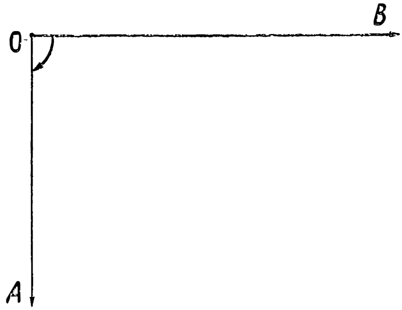


Рис. 18

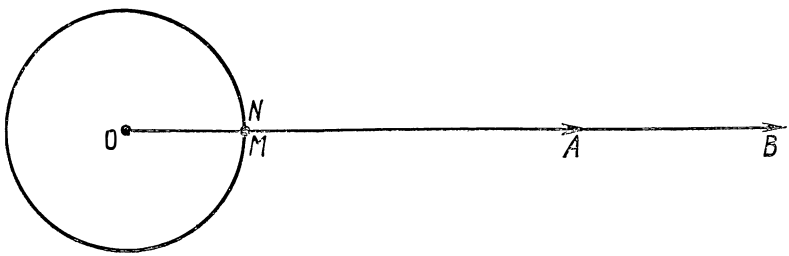


Рис. 19

поворот совершен по часовой стрелке. Если угол задается подвижным лучом OA, совпадающим с неподвижным лучом OB и подвижный луч OA не совершил поворота, то угол α будет нулевым и радианную меру этого угла полагают равной нулю.

Пусть подвижный луч OA совершит полный оборот против часовой стрелки, тогда точка N луча OA, двигаясь по окружности радиуса R, пройдет расстояние $2\pi R$. Следовательно, положительный полный угол α и положительный угол в 2π радиан один и тот же угол (рис. 21), т. е. угол в $360°$ и угол, равный 2π радиан, — один и тот же угол.

Если подвижный луч OA совершит полный оборот по часовой стрелке, то он образует угол, равный -2π радиан

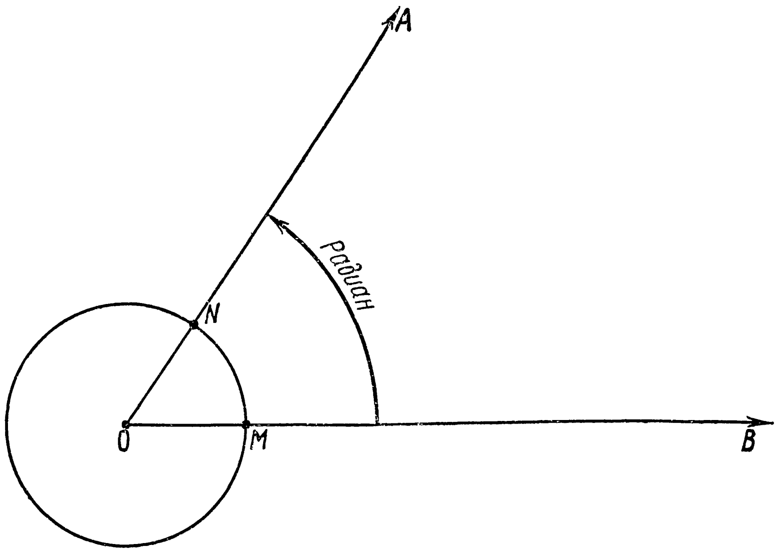


Рис. 20

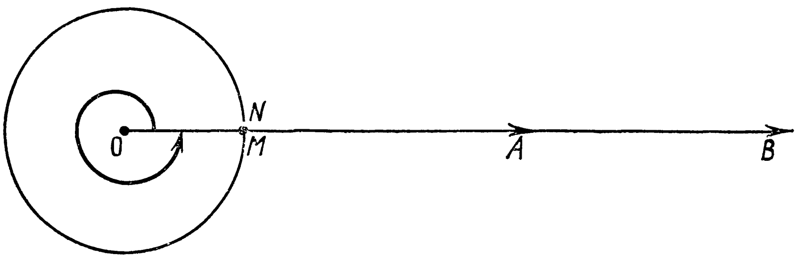


Рис. 21

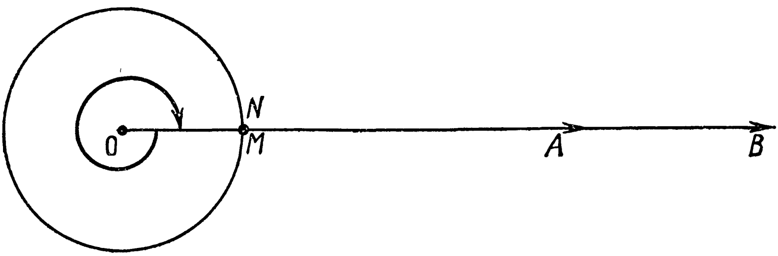


Рис. 22

(рис. 22), т. е. угол в $-360°$ и угол, равный -2π радиан, — один и тот же угол.

Пусть подвижный луч OA совершит $^1/_4$ полного оборота против часовой стрелки, заняв верхнее вертикальное положение. Тогда точка N подвижного луча, двигаясь по окружности радиуса R, пройдет расстояние $\pi R/2$. Следовательно, $^1/_4$ полного оборота есть угол равный $\pi/2$ радиан (рис. 23), т. е. угол в $90°$ и угол, равный $\pi/2$ радиан, — один и тот же угол.

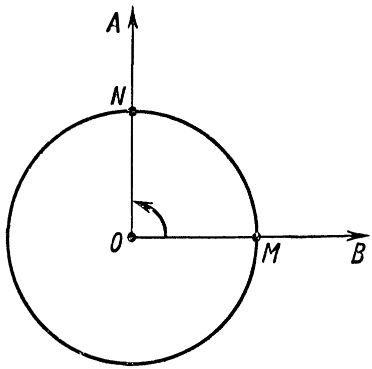


Рис. 23

Если подвижный луч OA совершил $^1/_4$ полного оборота по часовой стрелке, то он образует угол, равный $-\pi/2$ радиан (рис. 24), т. е. угол в $-90°$ и угол, равный $-\pi/2$ радиан, — один и тот же угол.

Пусть подвижный луч OA совершит $^1/_2$ полного оборота против часовой стрелки, заняв левое горизонтальное положение. Тогда точка N подвижного луча OA, двигаясь по окружности радиуса R, пройдет расстояние πR, и угол, образованный подвижным лучом OA, будет равен π радиан (рис. 25), т. е. угол в $180°$ и угол, равный π радиан, — один и тот же угол.

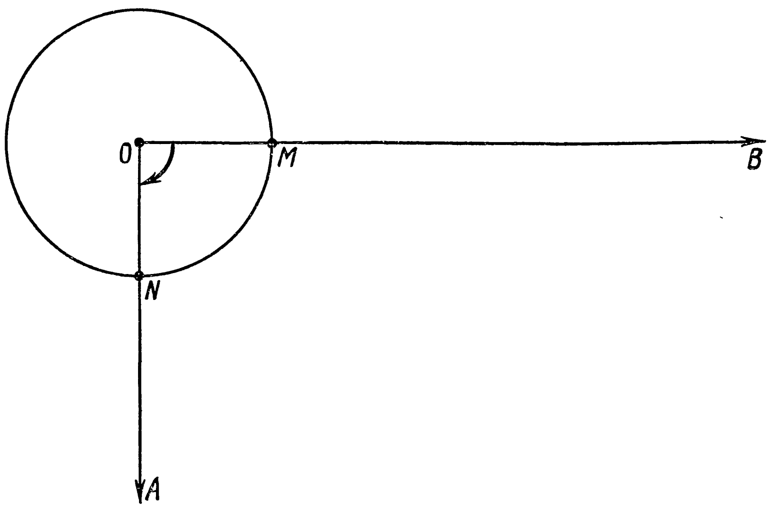


Рис. 24

Аналогично, угол в $-180°$ и угол, равный $-\pi$ радиан, — один и тот же угол, задаваемый подвижным лучом OA, совершившим $1/2$ полного оборота по часовой стрелке (см. рис. 25).

Если радианная мера некоторого угла есть β радиан, а градусная мера — α градусов, то эти числа связаны пропорцией

$$\alpha° : 360° = \beta : 2\pi.$$

Пользуясь этой пропорцией, легко переводить радианную меру угла в градусную и, наоборот — градусную меру в радианную. Рассмотренные выше примеры — частные случаи этой пропорции. Приведем еще примеры.

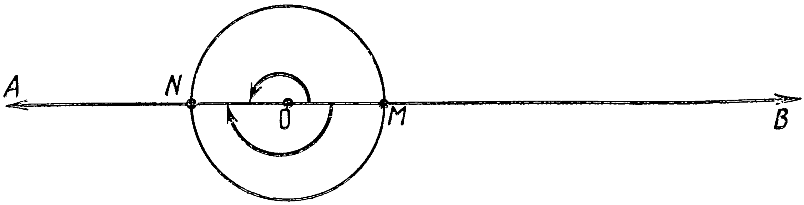


Рис. 25

Угол в $30°$ и угол $\pi/6$ радиан есть один и тот же угол. Это вытекает из справедливости пропорции $30° : 360° = \pi/6 : 2\pi$. Угол в $45°$ и угол $\pi/4$ радиан есть один и тот же угол. Это вытекает из справедливости пропорции $45° : 360° = \pi/4 : 2\pi$. Угол в $60°$ и угол $\pi/3$ радиан есть один и тот же угол. Это вытекает из справедливости пропорции $60° : 360° = \pi/3 : 2\pi$.

Замечание. В обозначениях меры угла в радианах часто опускают слово радиан, т. е. под углом $\alpha = 12$ понимается угол в 12 радиан, под углом $\beta = 7/9$ понимается угол в $7/9$ радиана, под углом $(x+y)$ понимается угол в $(x+y)$ радиан и т. д.

Пусть подвижный луч OA совпадает с неподвижным лучом OB. Но не известно, какой поворот совершил луч OA от луча OB, прежде чем совпасть с ним (см. рис. 19). Ответить на вопрос, какой поворот совершил луч OA, однозначно нельзя, ибо такое положение луча OA возможно после любого поворота, кратного полному обороту как по часовой, так и против часовой стрелки, а также при нулевом повороте. Следовательно, одно и то же положение подвижного луча OA при совпадении с неподвижным лучом OB задает бесконечно много как положительных, так и отрицательных углов, кратных $360°$ (или 2π радиан), или угол, равный нулю.

Пусть подвижный луч OA совершил некоторый поворот, после чего лучи OA и OB не совпадают. Если из этого положения как из начального положения подвижный луч OA совершит дополнительно любой поворот, кратный полному обороту по часовой или против часовой стрелки, то он окажется в том же положении OA. Следовательно, одно и то же положение подвижного луча OA задает бесконечно много поворотов как по часовой, так и против часовой стрелки, при этом, любые два из них отличаются друг от друга на поворот, кратный полному обороту по часовой или против часовой стрелки, совершаемый из положения OA, в котором находится подвижный луч, задающий эти повороты. Следовательно, подвижный луч в некотором фиксированном положении задает бесконечно много положительных и отрицательных углов, отличающихся друг от друга на угол, кратный углу в $360°$ (или 2π радиан). Часто для простоты эти углы записываются одной формулой. Приведем примеры такой записи.

Все углы, задаваемые подвижным лучом OA, находящимся в правом горизонтальном положении, можно записать формулами $360°n$, или $2\pi n$, где $n=0, \pm 1, \pm 2, \pm 3, \pm \ldots$. Действительно, из первой формулы при $n=0$ получается угол в $0°$, при $n=-1$ — угол в $-360°$, при $n=2$ — угол в $720°$ и т. д. Из второй формулы при $n=0$ получается угол в 0 радиан, при $n=1$ — угол в 2π радиан и т. д.

Все углы, задаваемые подвижным лучом OA, находящимся в верхнем вертикальном положении, можно записать формулами $90°+360°k$, или $\pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$. Действительно, из первой формулы при $k=0$ получается угол в $90°$, при $k=1$ — угол в $450°$, при $k=-1$ — угол в $-270°$, при $k=2$ — угол в $810°$, при $k=-2$ — угол в $-630°$ и т. д. Из второй формулы при $k=0$ получается угол в $\pi/2$ радиан, при $k=1$ — угол в $5\pi/2$ радиан, и т. д.

Все углы, задаваемые подвижным лучом OA, находящимся в левом горизонтальном положении, можно записать формулами $180°+360°m$, или $\pi+2\pi m$, где $m=0, \pm 1, \pm 2, \pm 3, \pm \ldots$. Действительно, из первой формулы при $m=0$ получается угол в $180°$, при $m=-1$ — угол в $-180°$, при $m=1$ — угол в $540°$ и т. д. Из второй формулы при $m=0$ получается угол в π радиан, при $m=-1$ — угол в $-\pi$ радиан и т. д.

Все углы, задаваемые подвижным лучом OA, находящимся в нижнем вертикальном положении, можно записать формулами $-90°+360°l$ или $-\pi/2+2\pi l$, где $l=0, \pm 1, \pm 2, \pm 3, \pm \ldots$. Действительно, из первой формулы при $l=0$ получается угол в $-90°$, при $l=1$ — угол в $270°$, при $l=-1$ — угол в $-450°$ и т. д. Из второй формулы при $l=0$ получается угол в $-\pi/2$ радиан, при $l=1$ — угол в $3\pi/2$ радиан, при $l=-1$ — угол в $-5\pi/2$ радиан и т. д.

В общем случае, если подвижный луч OA не совпадает с неподвижным лучом OB, т. е., если луч OA задает угол $\beta \in (0;\ 2\pi)$, то он задает бесконечное множество углов, которые можно записать с помощью формул $\beta+2\pi k$, где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$.

§ 2
СИСТЕМЫ КООРДИНАТ

Система координат на прямой. В § 5 гл. I для геометрической интерпретации действительных чисел было введено понятие числовой оси и показано, что между множеством точек, лежащих на оси, и множеством действительных чисел существует взаимно-однозначное соответствие.

Если на прямой выбрано начало отсчета — некоторая точка O, положительная и отрицательная полуоси и введен масштаб, то говорят, что на прямой задана система координат. При этом сама прямая называется координатной осью, а точка O — началом координат. Каждой точке M прямой ставят в соответствии число, называемое координатой точки в заданной системе координат, равное длине отрезка OM и взятое со знаком плюс, если точка M находится на положительной полуоси, и со знаком минус, если точка M находится на отрицательной полуоси.

Пусть данная координатная ось расположена горизонтально, причем так, что положительная полуось направлена вправо. Тогда любая точка, лежащая справа от начала координат O, имеет положительную координату, а любая точка, лежащая слева от начала координат O, — отрицательную координату. Координата точки O, начала координат, равна нулю.

Если рассматривается несколько разных фиксированных точек оси t, то часто их обозначают некоторой заглавной буквой с разными номерами, например, $M_1, M_2, M_3, \ldots, M_n, \ldots$; координаты этих точек обозначают прописной буквой оси с соответствующими номерами, т. е. $t_1, t_2, t_3, \ldots, t_n, \ldots$. Чтобы указать, что данная точка имеет данную координату, записывают эту координату в круглых скобках рядом с обозначением самой точки, например $M_1(t_1)$, $M_2(t_2)$, $M_3(t_3)$, ..., $M_n(t_n)$, ... Говоря, что дана точка, понимают, что дана ее координата; говоря, что надо найти точку, ищут ее координату.

Теорема 1. При любом расположении на координатной оси двух разных точек $M_1(t_1)$ и $M_2(t_2)$ расстояние d между этими точками равно модулю разности их координат, т. е.

$$d = |t_1 - t_2|.$$

Доказательство. Если точки M_1 и M_2 совпали, то утверждение теоремы очевидно. Пусть точки M_1 и M_2 не совпадают и пусть для определенности точка $M_2(t_2)$ лежит правее точки $M_1(t_1)$, (если точка $M_1(t_1)$ лежит правее точки $M_2(t_2)$, то доказательство повторяется с заменой t_2 на t_1).

Пусть $M_1(t_1)$ и $M_2(t_2)$ любые несовпадающие точки, лежащие правее начала координат O (рис. 26). Тогда длина отрезка M_1M_2 равна длине отрезка OM_2, которая равна t_2

минус длина отрезка OM_1, которая равна t_1, т. е. $d = t_2 - t_1 = |t_2 - t_1|$.

Пусть $M_1(t_1)$ совпадает с началом координат $O(0)$, т. е. $t_1 = 0$, а $M_2(t_2)$ — любая точка, лежащая правее начала координат O (рис. 27). Тогда длина d отрезка M_1M_2 равна длине отрезка OM_2, которая равна t_2, т. е. $d = t_2 = t_2 - 0 = t_2 - t_1 = |t_2 - t_1|$.

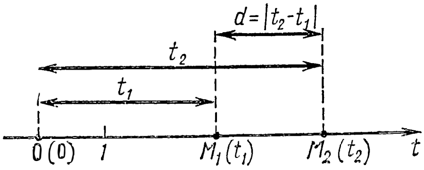


Рис. 26

Пусть $M_1(t_1)$ — любая точка, лежащая левее начала координат, а точка $M_2(t_2)$ — любая точка, лежащая правее начала координат O (рис. 28). Тогда длина d отрезка M_1M_2 равна длине отрезка OM_2, которая равна t_2 плюс длина отрезка OM_1, которая равна $-t_1$, т. е. $d = t_2 + (-t_1) = t_2 - t_1 = |t_2 - t_1|$.

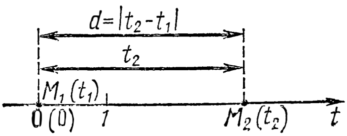


Рис. 27

Пусть $M_1(t_1)$ — любая точка, лежащая левее начала координат, а точка $M_2(t_2)$ совпадает с началом координат O, т. е. $t_2 = 0$ (рис. 29). Тогда длина d отрезка M_1M_2 совпадает с длиной отрезка M_1O, которая равна $-t_1$, т. е. $d = -t_1 = 0 - t_1 = t_2 - t_1 = |t_2 - t_1|$.

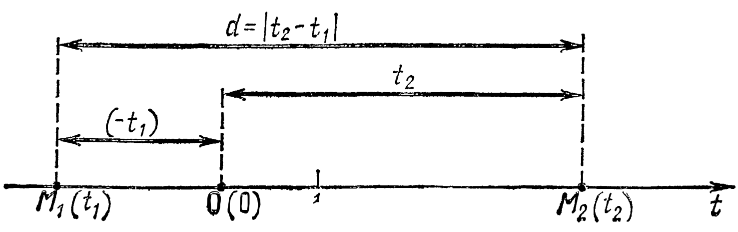


Рис. 28

Пусть $M_1(t_1)$ и $M_2(t_2)$ — любые несовпадающие точки, лежащие левее начала координат O (рис. 30). Тогда длина d отрезка M_1M_2 равна длине отрезка M_1O, которая равна $-t_1$, минус длина отрезка M_2O, которая равна $-t_2$, т. е. $d = (-t_1) - (-t_2) = t_2 - t_1 = |t_2 - t_1|$.

Итак, во всех случаях $d = |t_2 - t_1|$. Теорема доказана.

Пример 1. Найти расстояние между точками $M(3)$ и $P(-2)$.
$d=|3-(-2)|=|3+2|=5$, или $d=|-2-3|=|-5|=5$.

2. Найти расстояние между точками $M(-4)$ и $P(-10)$.

$$d=|-4-(-10)|=|-4+10|=6 \text{ или}$$
$$d=|(-10)-(-4)|=|(-10)+4|=6.$$

Прямоугольная система координат на плоскости. Если указан способ, позволяющий устанавливать положение точек на плоскости заданием пар чисел, то говорят, что на плоскости

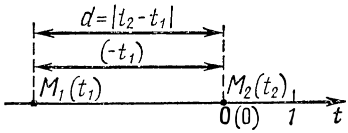


Рис. 29

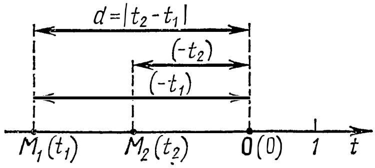


Рис. 30

задана система координат. Рассмотрим простейшую и чаще всего употребляемую систему координат, которая называется прямоугольной.

Две оси, пересекающиеся в некоторой точке, называются *взаимно перпендикулярными*, если положительная полуось одной из этих двух осей при повороте на угол $\pi/2$ по часовой стрелке или против совпадает с положительной полуосью другой оси.

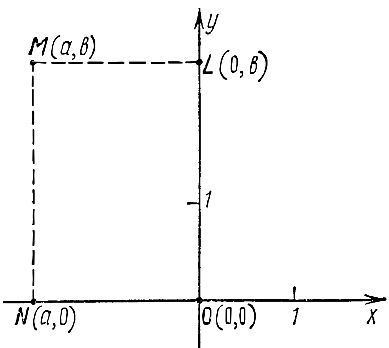


Рис. 31

Прямоугольная система координат на плоскости определяется заданием единицы для измерения длин и заданием двух взаимно перпендикулярных осей с указанием, какая из осей является первой и какая второй. Обычно принято первой осью считать ту из двух, положительная полуось которой при повороте на угол $\pi/2$ против часовой стрелки совпадает с положительной полуосью другой. Точка пересечения осей называется *началом координат*. Сами оси — *координатными осями*, причем первую из них принято называть *осью абсцисс*, а вторую — *осью ординат*.

Обозначим начало координат буквой O, ось абсцисс — буквами OX и ось ординат — буквами OY. На чертежах координатные оси обычно располагаются так, чтобы ось абсцисс была горизонтальной и ее положительная полуось направлена вправо, а положительная полуось ординат — вверх (рис. 31).

Введем на каждой из координатных осей OX и OY систему координат, сохранив данную единицу измерения и данные направления осей OX и OY, а в качестве начала координат на каждой оси выберем точку пересечения осей O.

Пусть M — любая точка плоскости. Проведем через точку M прямые, параллельные координатным осям, до пересечения с ними. Пусть прямая, проходящая через точку M и параллельная оси OY, пересечет ось абсцисс в точке N, а прямая, проходящая через точку M, параллельная оси OX, пересечет ось ординат в точке L. Так как на осях заданы системы координат, то точка N имеет в своей системе координат на оси абсцисс координату a, точка L имеет в своей системе координат на оси ординат координату b. Тогда координатами точки M в выбранной системе координат с осями OX и OY называют пару чисел (a, b). Число a называется первой координатой, или *абсциссой точки* M, число b называется второй координатой, или *ординатой точки* M. Тот факт, что точка M имеет абсциссу a и ординату b записывается $M(a, b)$. При этом сначала пишется абсцисса, затем ордината точки M (см. рис. 31).

Часто, когда рассматривается несколько разных фиксированных точек плоскости, их обозначают некоторой заглавной буквой с разными номерами, например $M_1, M_2, M_3, \ldots, M_n, \ldots$. Координаты этих точек помечаются соответствующими номерами: $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$, $\ldots$, $M_n(x_n, y_n)$, $\ldots$.

Так как через любую точку плоскости можно провести только одну прямую, параллельно данной координатной оси, а каждая такая прямая пересечет соответствующую ось только в одной точке, то каждой точке плоскости соответствует только одна упорядоченная пара чисел — координаты этой точки.

Между точками, лежащими на любой оси и множеством действительных чисел, имеется взаимно-однозначное соответствие (см. гл. I), следовательно, разным точкам плоскости XOY будут соответствовать разные упорядоченные пары действительных чисел. Итак, между множеством точек на плоскости и множеством упорядоченных пар действительных чисел существует следующее соответствие:

1. Каждой точке плоскости соответствует одна упорядоченная пара действительных чисел.

2. Двум разным точкам плоскости соответствуют разные упорядоченные пары действительных чисел.

3. Нет ни одной упорядоченной пары действительных чисел, которая бы не соответствовала какой-нибудь точке плоскости.

Такое соответствие называется взаимно-однозначным соответствием. Таким образом, введение на плоскости прямоугольной системы координат позволяет установить взаимно-однозначное соответствие между множеством всех точек плоскости и множеством упорядоченных пар действительных чисел. Это соответствие дает возможность сводить изучение множества точек плоскости к изучению множества пар действительных чисел, т. е. применять к изучению вопросов геометрии алгебраические методы.

Сделаем несколько замечаний.

1. Абсцисса точки M равна нулю тогда и только тогда, когда точка M лежит на оси OY.

2. Ордината точки M равна нулю тогда и только тогда, когда точка M лежит на оси OX.

3. Точка O — начало координат (и только она) имеет обе координаты, равные нулю.

4. Ось ординат делит плоскость на две полуплоскости: правую, каждая точка которой имеет положительную абсциссу ($x > 0$), и левую, каждая точка которой имеет отрицательную абсциссу ($x < 0$).

5. Ось абсцисс делит плоскость на две полуплоскости: верхнюю, каждая точка которой имеет положительную ординату ($y > 0$), и нижнюю, каждая точка которой имеет отрицательную ординату ($y < 0$).

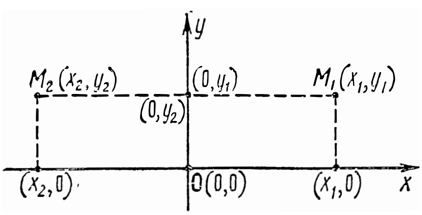


Рис. 32

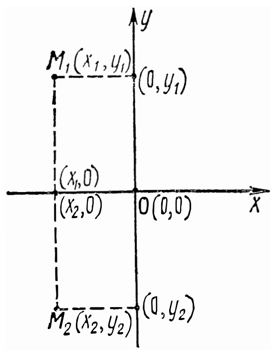


Рис. 33

6. Оси координат делят всю плоскость на четыре части, называемые *четвертями* (или *квадрантами*). I четверть та, в которой $x > 0$, $y > 0$; II — та, в которой $x < 0$, $y > 0$; III — та, в которой $x < 0$, $y < 0$; IV — та, в которой $x > 0$, $y < 0$. Точки, лежащие на самих осях, не относятся ни к какой четверти.

7. Положительная полуось абсцисс ($y = 0$, $x > 0$) называется правой горизонтальной полуосью; отрицательная полуось абсцисс ($y = 0$, $x < 0$) — левой горизонтальной полуосью; положительная полуось ординат ($x = 0$, $y > 0$) — верхней вертикальной полуосью; отрицательная полуось ординат ($x = 0$, $y > 0$) — нижней вертикальной полуосью.

8. Две точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ называются: *симметричными относительно оси ординат*, если $x_1 = -x_2$ и $y_1 = y_2$ (рис. 32); *симметричными относительно оси абсцисс*, если $x_1 =$

$= x_2$ и $y_1 = -y_2$ (рис. 33); *симметричными относительно начала координат*, если $x_1 = -x_2$ и $y_1 = -y_2$ (рис. 34).

Теорема 2. При любом расположении двух точек $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ на плоскости расстояние d между ними (длина отрезка M_1M_2) определяется формулой $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$, т. е. расстояние между двумя любыми точками плоскости, заданными своими координатами, равно корню квадратному из суммы квадратов разностей одноименных координат.

Доказательство. Пусть даны две несовпадающие точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$. Прямая M_1M_2 может быть: параллельна оси OY (или совпадать с ней) (а); параллельна оси OX (или совпадать с ней) (б); не параллельна ни оси OY, ни оси OX (в). Доказательство теоремы проведем для каждого из этих случаев отдельно.

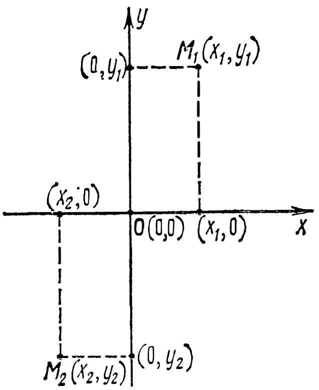


Рис. 34

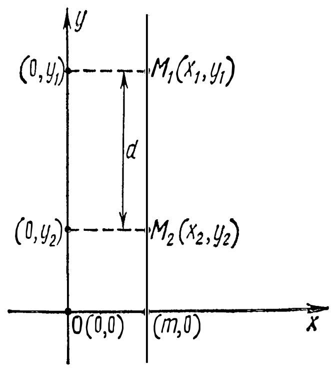


Рис. 35

а) Пусть прямая, на которой лежат точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ параллельна оси OY (или совпадает с ней). Тогда у любой точки, лежащей на этой прямой, одна и та же абсцисса, т. е. у точек M_1 и M_2 одинаковые абсциссы $x_1 = x_2 = m$ (рис. 35).

Эта прямая может быть рассмотрена как ось, с положительным направлением вверх, с тем же самым единичным отрезком, что и для системы координат XOY, и началом в точке $(m, 0)$. Координата любой точки этой оси будет совпадать с ординатой той же точки, как точки плоскости. Согласно теореме 1 расстояние между двумя точками M_1 и M_2, как точками этой оси, равно $d = |y_2 - y_1| = \sqrt{|y_2 - y_1|^2} = \sqrt{0 + (y_2 - y_1)^2} = \sqrt{(m - m)^2 + (y_2 - y_1)^2}$, т. е. $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

б) Пусть прямая, на которой лежат точки M_1 и M_2, параллельна оси OX (или совпадает с ней). Тогда у любой точки, лежащей на этой прямой, одна и та же ордината, т. е. у точек M_1 и M_2 одинаковые ординаты, $y_1 = y_2 = n$ (рис. 36).

Эта прямая может быть рассмотрена как ось, с положительным направлением вправо, с тем же самым единичным отрезком, что и для системы координат XOY, и началом в точке $(0, n)$. Координата любой точки этой оси будет совпадать с абсциссой той же точки, как точки плоскости.

Согласно теореме 1 расстояние между двумя точками M_1 и M_2, как точками этой оси, равно $d = |x_2 - x_1| = \sqrt{|x_2 - x_1|^2} = \sqrt{(x_2 - x_1)^2 + 0} = \sqrt{(x_2 - x_1)^2 + (n - n)^2}$, т. е. $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

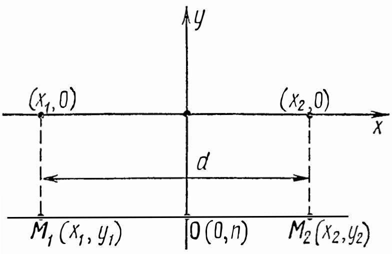


Рис. 36

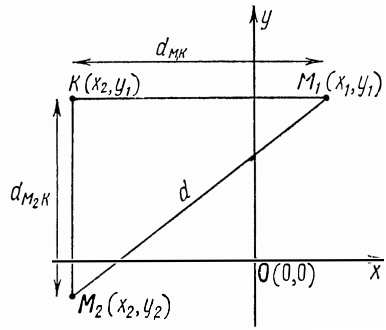


Рис. 37

в) Пусть теперь точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ не лежат ни на прямой, параллельной оси ординат, ни на прямой, параллельной оси абсцисс. Тогда одноименные координаты этих точек будут разные числа, т. е. $x_1 \neq x_2$ и $x_1 \neq y_2$ (рис. 37). Проведем через точку M_1 прямую, параллельную оси абсцисс, а через точку M_2 — прямую, параллельную оси ординат. Эти прямые пересекутся в точке $K(x_2, y_1)$. Точка M_1 и точка K лежат на прямой, параллельной оси абсцисс, следовательно, как было установлено в случае б), расстояние между этими точками (длина отрезка M_1K) равно $d_{M_1K} = |x_2 - x_1|$.

Точка M_2 и точка K лежат на прямой, параллельной оси ординат, следовательно, как было установлено в случае а), расстояние между этими точками (длина отрезка M_2K) равно $d_{M_2K} = |y_2 - y_1|$. Так как треугольник M_1KM_2 — прямоугольный, то по теореме Пифагора $d = \sqrt{d_{M_1K}^2 + d_{M_2K}^2} = \sqrt{|x_2 - x_1|^2 + |y_2 - y_1|^2}$. На основании свойств абсолютной величины получаем $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

Пример. Найти расстояние d между точками $M_1(-2, 1)$ и $M_2(-3, -2)$.

$$d = \sqrt{[-2 - (-3)]^2 + [+1 - (-2)]^2} = \sqrt{10}.$$

§ 3
ТРИГОНОМЕТРИЧЕСКИЕ ОПЕРАЦИИ НАД УГЛАМИ

Введем на плоскости прямоугольную систему координат XOY с положительной полуосью абсцисс OX, направленной вправо, а положительной полуосью ординат OY, направленной вверх (рис. 38). Пусть дана окружность радиуса, равного единице измерения длин, и с центром в начале координат. Такую окружность принято называть единичной окружностью, а радиус этой окружности — единичным радиусом. Единичная окружность пересекает положительную полуось абсцисс в точке $M(1, 0)$. Примем за вершину любого угла начало координат (точку $O(0, 0)$). Положительную полуось абсцисс примем за неподвижный луч, т. е. за начало отсчета для любого угла α. Единичный отрезок OM положительной полуоси абсцисс будем считать за отрезок, задающий неподвижный луч OB, и будем называть его неподвижным единичным радиусом.

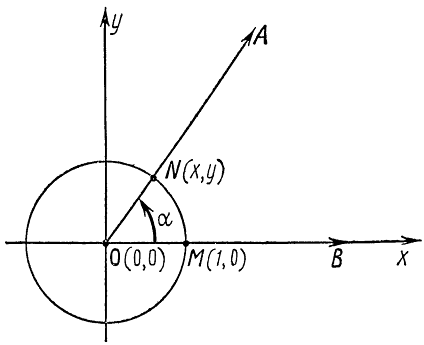


Рис. 38

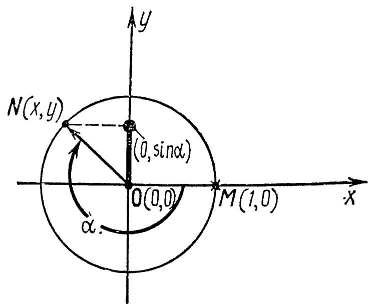


Рис. 39

Пусть некоторый угол α, образованный поворотом подвижного луча, пересекает единичную окружность в некоторой точке $N(x, y)$. Будем говорить, что отрезок ON определяет подвижный луч OA угла α и будем называть отрезок ON подвижным единичным радиусом, задающим угол α.

Пусть дан любой угол α. Число, равное ординате конца подвижного единичного радиуса, задающего этот угол α, называется *синусом* угла α и обозначается $\sin\alpha$ (рис. 39). Из определения следует, что для любого угла α существует $\sin\alpha$ и притом единственный.

Так как при любом угле α ордината конца подвижного радиуса, задающего угол α, не может быть меньше -1 и больше 1 и заключена между ними, включая -1 и 1, то для любого угла α: $-1 \leqslant \sin\alpha \leqslant 1$.

Пример. Найти синус угла α, равного 0, π, $\pi/2$, $-\pi/2$, $\pi/6$, $\pi/3$, $\pi/4$.

При $\alpha_1 = 0$ и $\alpha_2 = \pi$ ордината конца подвижного единичного радиуса равна нулю (рис. 40) и, следовательно, $\sin 0 = 0$, $\sin \pi = 0$.

При $\alpha_3 = \pi/2$ и $\alpha_4 = -\pi/2$ абсолютная величина ординаты конца подвижного единичного радиуса равна единице (рис. 41) и, следовательно, $\sin(\pi/2) = 1$, $\sin(-\pi/2) = -1$.

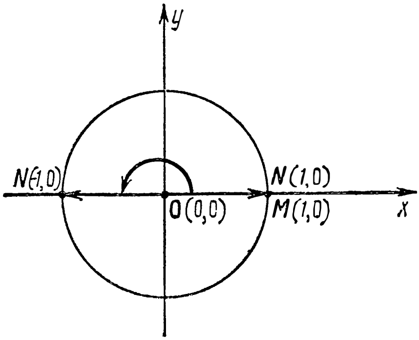


Рис. 40

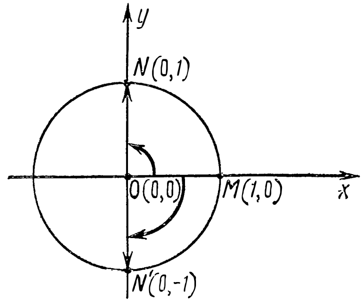


Рис. 41

При $\alpha_5 = \pi/6$ ордината конца подвижного единичного радиуса равна половине этого радиуса как катет прямоугольного треугольника, лежащий против угла в $\pi/6$. Следовательно, $\sin {}^{\pi}/_{6} = 1/2$ (рис. 42). Если $\alpha_6 = \pi/3$, то, воспользовавшись этим же свойством и теоремой Пифагора, получим $\sin {}^{\pi}/_{3} = \sqrt{1-(1/2)^2} = \sqrt{3}/2$ (рис. 43).

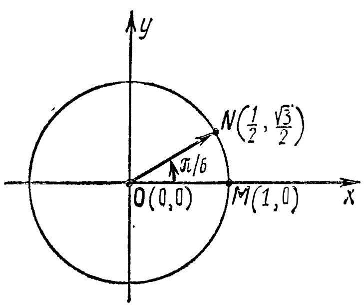


Рис. 42

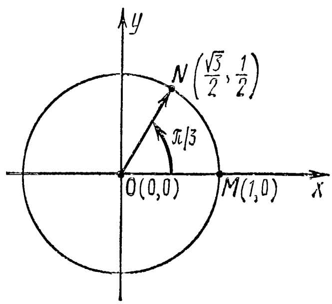


Рис. 43

При $\alpha_7 = \pi/4$ ордината конца подвижного единичного радиуса положительна и образует с абсциссой равнобедренный прямоугольный треугольник. По теореме Пифагора $\sin^2 {}^{\pi}/_{4} + \sin^2 {}^{\pi}/_{4} = 1 \Rightarrow \sin {}^{\pi}/_{4} = \sqrt{2}/2$ (рис. 44).

Возьмем любой угол α, задаваемый подвижным единичным радиусом ON (см. рис. 39). Если подвижный единичный радиус из положения ON совершит дополнительный полный оборот против часовой стрелки или по часовой стрелке, то он окажется в том же положении ON. Ордината конца N подвижного единичного радиуса ON останется прежней, хотя угол α изменится и станет соответственно равным $(\alpha+2\pi)$ или $(\alpha-2\pi)$, следовательно, $\sin\alpha=\sin(\alpha+2\pi)=\sin(\alpha-2\pi)$.

Итак, синус любого угла α повторяется через полный оборот по часовой стрелке или против часовой стрелки, т. е. через 2π. Число 2π называется *периодом синуса угла.*

При любом $k=1, 2, 3, \ldots$ после совершения подвижным единичным радиусом ON дополнительного поворота на k полных оборотов против часовой стрелки или по часовой стрелке угол α изменится и станет соответственно равным $(\alpha+2\pi k)$ или $(\alpha-2\pi k)$, но ордината его конца N останется прежней, следовательно, $\sin\alpha=\sin(\alpha+2\pi k)=\sin(\alpha-2\pi k)$ при любом $k=1, 2, 3, \ldots$. Итак, синус любого угла α повторяется при изменении угла на $2\pi n$, где $n=\pm1, \pm2, \pm3, \pm\ldots$, т. е. число $2\pi n$ тоже является периодом синуса угла. Дополнительный нулевой поворот не меняет ни положения подвижного единичного радиуса ON, ни угла, т. е. $\sin\alpha=\sin(\alpha+0\cdot2\pi)$.

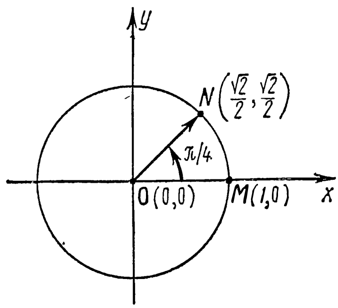


Рис. 44

Тот факт, что число $2\pi n$ является периодом синуса угла, записывается формулой $\sin\alpha=\sin(\alpha+2\pi n)$, где $n=0, \pm1, \pm2, \pm3, \ldots$. Знак числа n показывает направление поворота, а число $|n|$ — число полных оборотов, дополнительно совершенных после поворота на угол α.

Если подвижный единичный радиус находится в любом положении в I или II четвертях, или в верхнем вертикальном положении, то ордината его конца N положительна. Следовательно, $\sin x$ положителен для любого угла x из всех тех, которые задаются подвижным единичным радиусом ON, находящимся в I или II четвертях, или в верхнем вертикальном положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств: $2\pi k< <x<\pi+2\pi k$, где $k=0, \pm1, \pm2, \pm3, \ldots$). На числовой оси (рис. 45) показаны числа x, для каждого из которых $\sin x$ положителен.

Если подвижный единичный радиус ON находится в правом или левом горизонтальном положении, то ордината его конца N равна нулю. Следовательно, $\sin x=0$ для любого угла x из всех тех, которые задаются подвижным единичным радиусом ON, находящимся в правом или левом горизонтальном положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором m одному из следующих числовых равенств: $x = \pi m$, где $m = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 46) показаны числа x, для каждого из которых $\sin x = 0$.

Если подвижный единичный радиус ON находится в любом положении в III или IV четвертях, или в нижнем вертикаль-

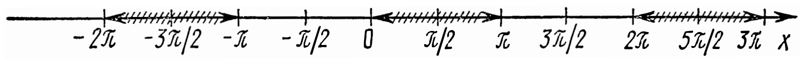


Рис. 45

ном положении, то ордината его конца N отрицательна. Следовательно, $\sin x$ отрицателен для любого угла x из всех тех, которые задаются подвижным единичным радиусом ON, находя-

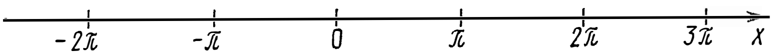


Рис. 46

щимся в III или IV четвертях или в нижнем вертикальном положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором n одному из следующих двойных неравенств: $2\pi n - \pi < x < 2\pi n$, где $n = 0$,

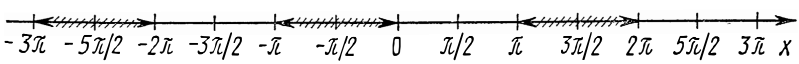


Рис. 47

$\pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 47) показаны числа x, для каждого из которых $\sin x$ отрицателен.

Возьмем произвольный угол α, заданный подвижным единичным радиусом в положении ON. Тогда угол $(-\alpha)$ задается

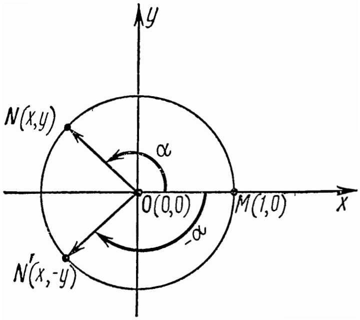


Рис 48

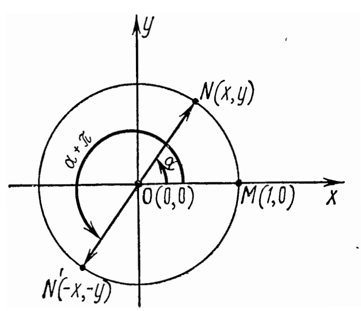


Рис. 49

подвижным единичным радиусом в положении ON', полученным после его вращения в направлении, противоположном направлению вращения для угла α. Подвижные единичные радиусы ON и ON' находятся во взаимно симметричном, относительно оси OX, положении (рис. 48).

Концы радиусов N и N', как точки взаимно симметричные относительно оси OX, имеют ординаты, равные друг другу по абсолютной величине и противоположные по знаку. Согласно определению синуса угла ордината конца радиуса ON есть $\sin\alpha$, а ордината конца радиуса ON' $\sin(-\alpha)$, поэтому можно написать равенства: $\sin(-\alpha) = -\sin\alpha$. Это свойство синуса угла называется *нечетностью* и его можно выразить словами так: для любого угла α синус угла минус α равен минус синусу угла α.

Итак, нечетность синуса угла заключается в том, что знак минус можно выносить за знак синуса или вносить под знак синуса, т. е. $\sin(-\alpha) = -\sin\alpha = \sin(-\alpha)$.

Пример. $\sin(-\pi/6) = -\sin(\pi/6) = -1/2$; $\sin(-\pi/3) = -\sin(\pi/3) = -\sqrt{3}/2$.

Отметим еще некоторые свойства синуса угла.

1. $\sin(\alpha+\pi) = -\sin\alpha$ (из симметрии относительно точки $O(0, 0)$ (рис. 49)).

2. $\sin(\pi - a) = \sin\alpha$ (из симметрии относительно вертикальной оси (рис. 50))

Пример. $\sin(5\pi/4) = \sin(\pi+\pi/4) = -\sin(\pi/4) = -\sqrt{2}/2$; $\sin(7\pi/6) = \sin(\pi+\pi/6) = -\sin(\pi/6) = -1/2$; $\sin(3\pi/4) = \sin(\pi-\pi/4) = \sin(\pi/4) = \sqrt{2}/2$; $\sin(5\pi/6) = \sin(\pi-\pi/6) = \sin(\pi/6) = 1/2$.

Пусть дан любой угол α. Число, равное абсциссе конца подвижного единичного радиуса, задающего этот угол α, называется *косинусом* угла α и обозначается $\cos\alpha$ (рис. 51). Из определения следует, что для любого угла α существует $\cos\alpha$ и притом единственный.

Так как при любом угле α абсцисса конца подвижного единичного радиуса, задающего угол α, не может быть меньше -1 и больше 1 и заключена между ними, включая -1 и 1, то для любого угла α: $-1 \leqslant \cos\alpha \leqslant 1$.

Пример. Найти косинус угла α, равного 0, π, $\pi/2$, $-\pi/2$, $\pi/6$, $\pi/3$, $\pi/4$.

Так же как и в случае нахождения синуса угла, легко видеть, что $\cos 0 = 1$, $\cos\pi = -1$, $\cos(\pi/2) = 0$, $\cos(-\pi/2) = 0$ (см. рис. 40, 41).

При $\alpha_5 = \pi/6$ и $\alpha_6 = \pi/3$, используя соотношения в прямоугольном треугольнике с углами $\pi/6$ и $\pi/3$, получим, что $\cos(\pi/6) = \sqrt{3}/2$, $\cos(\pi/3) = 1/2$ (см. рис. 42, 43).

При $\alpha_7 = \pi/4$ абсцисса и ордината конца подвижного единичного радиуса образуют равнобедренный прямоугольный треугольник. Следовательно, $\cos {}^{\pi}/_{4} = \sin {}^{\pi}/_{4} = \sqrt{2}/2$ (см. рис. 44).

Возьмем любой угол α, задаваемый подвижным единичным радиусом ON. Если подвижный единичный радиус из положе-

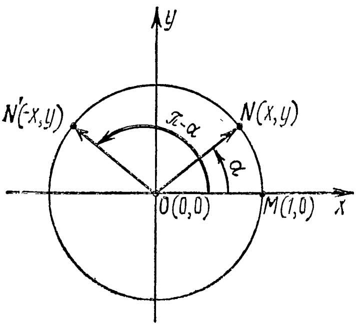


Рис. 50

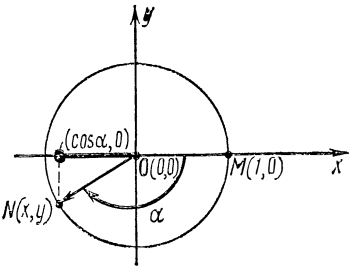


Рис. 51

ния ON совершит дополнительный полный оборот против часовой стрелки или по часовой стрелке, то он окажется в том же положении ON. Абсцисса конца N подвижного единичного радиуса останется прежней, хотя угол α изменится и станет соответственно равным $(\alpha + 2\pi)$ или $(\alpha - 2\pi)$ (см. рис. 51), следовательно, $\cos\alpha = \cos(\alpha + 2\pi) = \cos(\alpha - 2\pi)$.

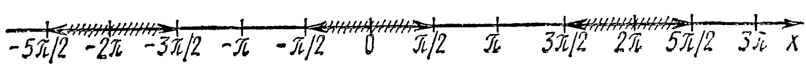


Рис. 52

Итак, косинус любого угла α повторяется через полный оборот по часовой стрелке или против часовой стрелки, т. е. через 2π. Число 2π называется *периодом косинуса угла*. Аналогично показывается справедливость равенств: $\cos\alpha = \cos(\alpha + 2\pi k) = \cos(\alpha - 2\pi k)$, при любом $k = 1, 2, 3, \ldots$.

Итак, косинус любого угла α повторяется при изменении угла на $2\pi n$, где $n = \pm 1, \pm 2, \pm 3, \pm \ldots$, т. е. число $2\pi n$ тоже является периодом косинуса угла. Дополнительный нулевой поворот не меняет ни положения подвижного единичного радиуса ON, ни угла α, т. е. $\cos\alpha = \cos(\alpha + 0 \cdot 2\pi)$.

Тот факт, что число $2\pi n$ является периодом косинуса угла записывается следующей формулой: $\cos\alpha = \cos(\alpha + 2\pi k)$, где $n = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$. Знак числа n показывает направление поворота, а число $|n|$ — число полных оборотов, дополнительно совершенных после поворота на угол α.

Если подвижный единичный радиус находится в любом положении в I или IV четвертях, или правом горизонтальном положении, то абсцисса конца подвижного единичного радиуса положительна. Следовательно, $\cos x$ положителен для любого угла x, из всех тех, которые задаются подвижным единичным радиусом ON, находящимся в I или IV четверти, или в правом горизонтальном положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором l одному из следующих двойных числовых неравенств: $2\pi l - \pi/2 < x < \pi/2 + 2\pi l$, где $l = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 52) показаны числа x, для каждого из которых $\cos x$ положителен.

Если подвижный единичный радиус находится в любом положении во II или III четвертях, или в левом горизонтальном положении, то абсцисса конца подвижного единичного радиуса отрицательна. Следовательно, $\cos x$ отрицателен для любого угла x, из всех тех, которые задаются подвижным единичным радиусом ON, находящимся во II или III четвертях, или в левом горизонтальном положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором p одному из следующих двойных числовых неравенств $2\pi p + \pi/2 < x < 3\pi/2 + 2\pi p$, где $p = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 53) пакозаны числа x, для каждого из которых $\cos x$ отрицателен.

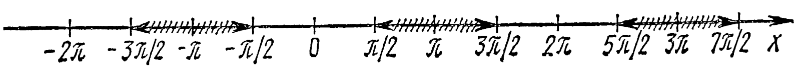


Рис. 53

Если подвижный единичный радиус находится в верхнем или нижнем вертикальном положении, то абсцисса конца этого подвижного единичного радиуса равна нулю. Следовательно, $\cos x = 0$ для любого угла x из всех тех, которые задаются подвижным единичным радиусом ON, находящимся в верхнем и нижнем вертикальном положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором n одному из следующих числовых равенств: $x = \pi/2 + \pi n$, где $n = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 54) показаны числа x, для каждого из которых $\cos x = 0$.

Возьмем произвольный угол α, заданный подвижным единичным радиусом в положении ON. Тогда угол $(-\alpha)$ задается подвижным единичным радиусом в положении ON', полученным после его вращения в направлении, противоположном направлению вращения для угла α. Подвижные единичные радиусы ON и ON' находятся во взаимно симметричном относительно оси OX положении (см. рис. 48).

Концы радиусов N и N', как точки взаимно симметричные относительно оси OX, имеют одну и ту же абсциссу, которая

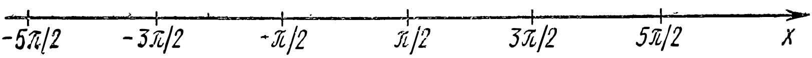


Рис. 54

для конца подвижного радиуса ON является $\cos\alpha$, а для ON' является $\cos(-\alpha)$, поэтому можно написать равенства: $\cos(-\alpha)=\cos\alpha$. Это свойство косинуса угла называется *четностью* и его можно выразить словами так: для любого угла α косинус угла α равен косинусу угла минус α.

Итак, четность косинуса угла заключается в том, что знак перед углом, стоящим под косинусом, можно менять, не меняя значения косинуса угла, т. е. $\cos(-\alpha)=\cos\alpha=\cos(-\alpha)$.

Пример. $\cos(-\pi/6)=\cos\pi/6=\sqrt{3}/2$; $\cos(-\pi/4)=\cos\pi/4=$ $=\sqrt{2}/2$; $\cos(-\pi/3)=\cos\pi/3=1/2$.

Отметим еще некоторые свойства косинуса угла.

1. $\cos(\pi+\alpha)=-\cos\alpha$ (из симметрии относительно точки $O(0,\ 0)$, (см. рис. 49)).

2. $\cos(\pi-\alpha)=-\cos\alpha$ (из симметрии относительных вертикальной оси (см. рис. 50)).

Пример. $\cos 5\pi/4=\cos(\pi+\pi/4)=-\cos\pi/4=-\sqrt{2}/2$; $\cos 7\pi/6=\cos(\pi+\pi/6)=-\cos\pi/6=-\sqrt{3}/2$; $\cos 3\pi/4=$ $=\cos(\pi-\pi/4)=-\cos\pi/4=-\sqrt{2}/2$; $\cos 5\pi/6=\cos(\pi-\pi/6)=$ $=-\cos\pi/6=-\sqrt{3}/2$.

Пусть дан любой угол α, такой, что $\alpha\neq\pi/2+\pi k$, где $k=0,\ \pm1,\ \pm2,\ \pm3,\ \pm\ldots$. *Тангенсом* угла α называется число, равное отношению синуса этого угла к косинусу того же угла α и обозначаемое $\operatorname{tg}\alpha$, т. е. $\operatorname{tg}\alpha=\frac{\sin\alpha}{\cos\alpha}$. Из определения следует, что для любого угла α, задаваемого любым положением подвижного единичного радиуса, кроме верхного и нижнего вертикальных, тангенс угла α существует и притом единственный.

Пример. $\operatorname{tg}0=\frac{\sin 0}{\cos 0}=0$; $\operatorname{tg}\pi=\frac{\sin\pi}{\cos\pi}=0$; $\operatorname{tg}\frac{\pi}{4}=\frac{\sin(\pi/4)}{\cos(\pi/4)}=$

$$=1;\quad \operatorname{tg}\frac{\pi}{6}=\frac{\sin(\pi/6)}{\cos(\pi/6)}=\frac{\sqrt{3}}{3};\ \operatorname{tg}\frac{\pi}{3}=\frac{\sin(\pi/3)}{\cos(\pi/3)}=\sqrt{3}.$$

Для любого угла x, синус и косинус которого одного знака, тангенс угла x положителен, т. е. $\operatorname{tg}x$ положителен для любого угла x, задаваемого подвижным единичным радиусом, находящимся в I или III четвертях (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором к одному из следующих двойных числовых неравенств: $\pi k < x < \pi/2 + \pi k$, где $k = 0, \pm 1 \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 55) показаны числа x, для каждого из которых $\operatorname{tg} x$ положителен.

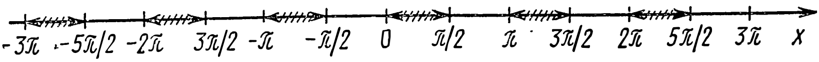


Рис. 55

Для любого угла x, синус и косинус которого разных знаков, тангенс угла x отрицателен, т. е. $\operatorname{tg} x$ отрицателен для любого угла x, задаваемого подвижным единичным радиусом, находящимся во II или IV четвертях (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором l одному из следующих двойных числовых неравенств: $\pi l - \pi/2 < x < \pi l$, где $l = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (рис. 56) показаны числа x, для каждого из которых $\operatorname{tg} x$ отрицателен.

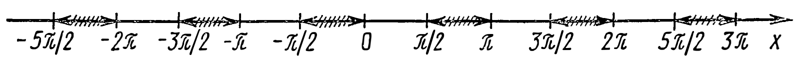


Рис. 56

Для любого угла x, синус которого равен нулю, тангенс угла x тоже равен нулю, т. е. $\operatorname{tg} x = 0$ для любого угла x, задаваемого подвижным единичным радиусом, находящимся в горизонтальном правом или левом положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором m одному из следующих числовых равенств: $x = \pi m$, где $m = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (см. рис. 46) показаны числа x, для каждого из которых $\operatorname{tg} x = 0$.

Так для любого угла α $\sin(-\alpha) = -\sin\alpha$ и $\cos(-\alpha) = \cos\alpha$, то для любого угла α ($\alpha \neq \pi/2 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$) по определению тангенса $\operatorname{tg}(-\alpha) = \frac{\sin(-\alpha)}{\cos(-\alpha)} = \frac{-\sin\alpha}{\cos\alpha} = -\frac{\sin\alpha}{\cos\alpha} = -\operatorname{tg}\alpha$. Это свойство тангенса угла называется *нечетностью* и его можно выразить словами так: для любого допустимого угла α (т. е. для любого угла α, для которого определен тангенс) тангенс угла, взятого со знаком минус, равен минус тангенсу угла.

Пример. $\operatorname{tg}(-\pi/4) = -\operatorname{tg}\pi/4 = -1$; $\operatorname{tg}(-\pi/6) = -\operatorname{tg}(\pi/6) = -\sqrt{3}/3$; $\operatorname{tg}(-\pi/3) = -\operatorname{tg}(\pi/3) = -\sqrt{3}$.

Используя свойства синуса и косинуса угла, можно написать, что для любого допустимого угла α

$$\operatorname{tg}(\alpha+\pi)=\frac{\sin(\alpha+\pi)}{\cos(\alpha+\pi)}=\frac{-\sin\alpha}{-\cos\alpha}=\frac{\sin\alpha}{\cos\alpha}=\operatorname{tg}\alpha,$$
$$\operatorname{tg}(\alpha-\pi)=\frac{\sin(\alpha-\pi)}{\cos(\alpha-\pi)}=\frac{-\sin\alpha}{-\cos\alpha}=\frac{\sin\alpha}{\cos\alpha}=\operatorname{tg}\alpha.$$

Пример. $\operatorname{gt} {}^{5\pi}/_4=\operatorname{tg}(\pi+\pi/4)=\operatorname{tg} {}^{\pi}/_4=1$; $\operatorname{tg} {}^{7\pi}/_6=\operatorname{tg}(\pi+\pi/6)=\operatorname{tg} {}^{\pi}/_6=\sqrt{3}/3$; $\operatorname{tg} {}^{3\pi}/_4=\operatorname{tg}(\pi-\pi/4)=\operatorname{tg}(-\pi/4)=-1$; $\operatorname{tg} 5\pi/6=\operatorname{tg}(\pi-\pi/6)=\operatorname{tg}(-\pi/6)=-\sqrt{3}/3$.

Итак, тангенс любого допустимого угла α повторяется при изменении угла на половину полного оборота против часовой стрелки или по часовой стрелке, т. е. через π. Число π называется *периодом тангенса угла.* Учитывая это, получим

$$\operatorname{tg}\alpha=\operatorname{tg}(\alpha+\pi)=\operatorname{tg}[(\alpha+\pi)+\pi]=\operatorname{tg}(\alpha+2\pi)=\operatorname{tg}[(\alpha+2\pi)+\pi]=$$
$$=\operatorname{tg}(\alpha+3\pi)=\ldots,$$
$$\operatorname{tg}\alpha=\operatorname{tg}(\alpha-\pi)=\operatorname{tg}[(\alpha-\pi)-\pi]=\operatorname{tg}(\alpha-2\pi)=\operatorname{tg}[(\alpha-2\pi)-\pi]=$$
$$=\operatorname{tg}(\alpha-3\pi)=\ldots,$$

т. е. для любого $k=1, 2, 3, \ldots$ $\operatorname{tg}\alpha=\operatorname{tg}(\alpha+k\pi)=\operatorname{tg}(\alpha-k\pi)$.

Итак, тангенс любого допустимого угла α повторяется при изменении угла на πn, где $n=\pm 1, \pm 2, \pm 3, \ldots$, т. е. число πn тоже является периодом тангенса угла. Тот факт, что πn является периодом тангенса угла, записывается следующей формулой: $\operatorname{tg}\alpha=\operatorname{tg}(\alpha+\pi n)$, где $n=0, \pm 1, \pm 2, \pm 3, \pm \ldots$. При этом знак числа n показывает направление поворота, а число $|n|$ — число половин полного оборота, дополнительно совершенных после поворота на угол α.

Пусть дан любой угол α, такой, что $\alpha\neq\pi m$, где $m=0, \pm 1, \pm 2, \pm 3, \pm \ldots$. *Котангенсом* этого угла α называется число, равное отношению косинуса этого угла α к синусу того же угла α и обозначаемое $\operatorname{ctg}\alpha$, т. е. $\operatorname{ctg}\alpha=\frac{\cos\alpha}{\sin\alpha}$. Из определения следует, что для любого угла α, задаваемого любым положением подвижного единичного радиуса, кроме правого и левого горизонтальных, котангенс угла α существует и притом единственный.

Пример. $\operatorname{ctg}\frac{\pi}{2}=\frac{\cos(\pi/2)}{\sin(\pi/2)}=0$; $\operatorname{ctg}\left(-\frac{\pi}{2}\right)=\frac{\cos(\pi/2)}{\sin(-\pi/2)}=0$; $\operatorname{ctg}\frac{\pi}{4}=\frac{(\cos\pi/4)}{(\sin\pi/4)}=1$; $\operatorname{ctg}\frac{\pi}{6}=\frac{\cos(\pi/6)}{\sin(\pi/6)}=\sqrt{3}$; $\operatorname{ctg}\frac{\pi}{3}=\frac{\cos(\pi/3)}{\sin(\pi/3)}=\frac{\sqrt{3}}{3}$.

Для любого угла x, косинус и синус которого одного знака, котангенс угла x положителен, т. е. $\operatorname{ctg} x$ положителен для любого угла x, задаваемого подвижным единичным радиусом, находящимся в I или III четвертях (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором p одному из следующих двойных числовых неравенств: $\pi p < x < \pi/2 + \pi p$, где $p = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (см. рис. 55) показаны числа x, для каждого из которых $\operatorname{ctg} x$ положителен.

Для любого угла x, косинус и синус которого разных знаков, котангенс угла x отрицателен, т. е. $\operatorname{ctg} x$ отрицателен для любого угла x, задаваемого подвижным единичным радиусом, находящимся во II или IV четвертях (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором n одному из следующих двойных числовых неравенств: $\pi n + \pi/2 < x < \pi + \pi n$, где $n = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (см. рис. 56) показаны числа x, для каждого из которых $\operatorname{ctg} x$ отрицателен.

Для любого угла x, косинус которого равен нулю, котангенс угла x тоже равен нулю, т. е. $\operatorname{ctg} x = 0$ для любого угла x, задаваемого подвижным единичным радиусом, находящимся в вертикальном верхнем или вертикальном нижнем положении (т. е. для любого числа x, как радианной меры соответствующего угла x, удовлетворяющего при некотором m одному из следующих числовых равенств: $x = \pi/2 + \pi m$, где $m = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$). На числовой оси (см. рис. 54) показаны такие числа x, для каждого из которых $\operatorname{ctg} x = 0$.

Так как для любого угла α $\sin(-\alpha) = -\sin\alpha$ и $\cos(-\alpha) = \cos\alpha$, то для любого угла α ($\alpha \neq \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$) по определению котангенса

$$\operatorname{ctg}(-\alpha) = \frac{\cos\alpha}{\sin(-\alpha)} = \frac{\cos\alpha}{-\sin\alpha} = -\frac{\cos\alpha}{\sin\alpha} = -\operatorname{ctg}\alpha.$$

Это свойство котангенса угла называется *нечетностью* котангенса и его можно выразить словами так: для любого допустимого угла (т. е. для любого угла α, для которого определен котангенс) котангенс угла, взятого со знаком минус, равен минус котангенсу угла α.

Пример. $\operatorname{ctg}(-\pi/4) = -\operatorname{ctg}(\pi/4) = -1$; $\operatorname{ctg}(-\pi/6) = -\operatorname{ctg}(\pi/6) = -\sqrt{3}$; $\operatorname{ctg}(-\pi/3) = -\operatorname{ctg}(\pi/3) = -\sqrt{3}/3$.

Используя свойства синуса и косинуса угла, можно написать, что для любого допустимого угла α

$$\operatorname{ctg}(\alpha + \pi) = \frac{\cos(\alpha + \pi)}{\sin(\alpha + \pi)} = \frac{-\cos\alpha}{-\sin\alpha} = \frac{\cos\alpha}{\sin\alpha} = \operatorname{ctg}\alpha,$$

$$\operatorname{ctg}(\alpha - \pi) = \frac{\cos(\alpha - \pi)}{\sin(\alpha - \pi)} = \frac{-\cos\alpha}{-\sin\alpha} = \frac{\cos\alpha}{\sin\alpha} = \operatorname{ctg}\alpha.$$

Итак, котангенс любого допустимого угла α повторяется при изменении угла α на половину полного оборота против часовой стрелки или по часовой стрелке, т. е. через π. Число π называется *периодом котангенса угла*. Учитывая это, получим

$$\operatorname{ctg} = \operatorname{ctg}(\alpha+\pi) = \operatorname{ctg}[(\alpha+\pi)+\pi] = \operatorname{ctg}(\alpha+2\pi) =$$
$$= \operatorname{ctg}[(\alpha+2\pi)+\pi] = \operatorname{ctg}(\alpha+3\pi) = \ldots,$$
$$\operatorname{ctg}\alpha = \operatorname{ctg}(\alpha-\pi) = \operatorname{ctg}[(\alpha-\pi)-\pi] = \operatorname{ctg}(\alpha-2\pi) =$$
$$= \operatorname{ctg}[\alpha-2\pi)-\pi] = \operatorname{ctg}(\alpha-3\pi) = \ldots,$$

т. е. для любого $k=1,\ 2,\ 3,\ \ldots.\ \operatorname{ctg}\alpha = \operatorname{ctg}(\alpha+k\pi) = \operatorname{ctg}(\alpha-k\pi)$.

Итак, котангенс любого допустимого угла повторяется при изменении угла на πn, где $n=\pm 1,\ \pm 2,\ \pm 3\ \pm\ldots$, т. е. число πn тоже является периодом котангенса угла. Тот факт, что πn является периодом котангенса угла записывается следующей формулой: $\operatorname{ctg}\alpha = \operatorname{ctg}(\alpha+\pi n)$, где $n=0,\ \pm 1,\ \pm 2,\ \pm 3,\ \pm\ldots.$ При этом знак числа n показывает направление поворота, а число $|n|$ — число половин полного оборота, дополнительно совершенных после поворота на угол α.

§ 4
ОСНОВНОЕ ТРИГОНОМЕТРИЧЕСКОЕ ТОЖДЕСТВО

Теорема 1. Для любого угла α справедливо основное тригонометрическое тождество:

$$\sin^2\alpha + \cos^2\alpha = 1, \tag{1}$$

т. е. квадрат синуса любого угла плюс квадрат косинуса того же угла равен единице.

Доказательство. Пусть дан некоторый угол α. Тогда координаты конца подвижного единичного радиуса, задающего угол α, будут $(\cos\alpha,\ \sin\alpha)$ (рис. 57). Так как квадрат расстояния между любыми двумя точками плоскости, заданными своими координатами, равен сумме квадратов разности одноименных координат, то для точки $(\cos\alpha,\ \sin\alpha)$ и точки $(0,\ 0)$ имеем $(\cos\alpha-0)^2+(\sin\alpha-0)^2=1$, или $\sin^2\alpha+\cos^2\alpha=1$ и теорема 1 доказана.

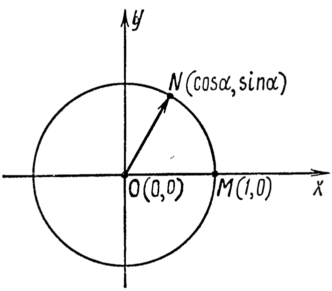


Рис. 57

Основное тригонометрическое тождество показывает, в какой зависимости находится синус и косинус одного и того же угла. Зная одну из величин, входящих в основное тригонометрическое тождество для некоторого угла α, можно найти другую величину того же угла α. Действительно, по основному тригонометрическому тождеству $\cos^2\alpha = 1 - \sin^2\alpha$ или $\sin^2\alpha = 1 - \cos^2\alpha$. Полученные тождества равносильны следующим:

$$|\cos\alpha| = \sqrt{1-\sin^2\alpha}, \tag{2}$$

$$|\sin\alpha| = \sqrt{1-\cos^2\alpha}. \tag{3}$$

Из тождества (2) имеем

$$\cos\alpha = \sqrt{1-\sin^2\alpha} \quad (2а)$$

для любого угла α, для которого $\cos\alpha$ не отрицателен (т. е. для любого α, удовлетворяющего при некотором k одному из двойных числовых неравенств:

$$2\pi k - \frac{\pi}{2} \leqslant \alpha \leqslant \frac{\pi}{2} + 2\pi k,$$

где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$).

$$\cos\alpha = -\sqrt{1-\sin^2\alpha} \quad (2б)$$

для любого угла α, для которого $\cos\alpha$ не положителен (т. е. для любого α, удовлетворяющего при некотором k одному из двойных числовых неравенств:

$$2\pi k + \frac{\pi}{2} \leqslant \alpha \leqslant \frac{3\pi}{2} + 2\pi k,$$

где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$).

Из тождества (3) имеем

$$\sin\alpha = \sqrt{1-\cos^2\alpha} \quad (3а)$$

для любого угла α, для которого $\sin\alpha$ не отрицателен (т. е. для любого α, удовлетворяющего при некотором m одному из двойных числовых неравенств:

$$2\pi m \leqslant \alpha \leqslant \pi + 2\pi m,$$

где $m = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$).

$$\sin\alpha = -\sqrt{1-\cos^2\alpha} \quad (3б)$$

для любого угла α, для которого $\sin\alpha$ не положителен (т. е. для любого α, удовлетворяющего при некотором m одному из двойных числовых неравенств:

$$\pi + 2\pi m \leqslant \alpha \leqslant 2\pi + 2\pi m,$$

где $m = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$).

Замечание. Формулы (2*а*) и (2б) при граничных значениях угла α ($\alpha = \pi/2 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$) дают одно и то же значение $\cos\alpha = 0$; формулы (3а) и (3б) при граничных значениях угла α ($\alpha = \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$) дают одно и то же значение $\sin\alpha = 0$.

Пример. 1. Пусть $\cos\alpha = -9/11$ и $\pi < \alpha < 3\pi/2$. Вычислить $\sin\alpha$.

В указанном промежутке $\sin\alpha$ отрицателен и поэтому $\sin\alpha = -\sqrt{1-\cos^2\alpha} = -2\sqrt{10}/11$.

2. Вычислить $\cos\alpha$, если $-3\pi/2 < \alpha < -\pi$ и $\sin\alpha = -1/3$.

В указанном промежутке $\cos\alpha$ отрицателен и поэтому $\cos\alpha = -\sqrt{1-\sin^2\alpha} = -2\sqrt{2}/3$.

Следствие 1. Для любого угла α, такого, что $\alpha \neq \pi/2 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$, справедливо тождество:

$$1 + \operatorname{tg}^2\alpha = \frac{1}{\cos^2\alpha}. \tag{4}$$

Доказательство. Так как $\alpha \neq \pi/2 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$, то $\cos\alpha \neq 0$ и поэтому $\operatorname{tg}\alpha$ определен и основное тригонометрическое тождество (1) можно почленно разделить на $\cos^2\alpha$. Тогда для любого такого α имеем $\frac{(\cos^2\alpha+\sin^2\alpha)}{\cos^2\alpha}=$ $=\frac{1}{\cos^2\alpha}\Leftrightarrow\frac{\cos^2\alpha}{\cos^2\alpha}+\frac{\sin^2\alpha}{\cos^2\alpha}=\frac{1}{\cos^2\alpha}\Leftrightarrow 1+\left(\frac{\sin\alpha}{\cos\alpha}\right)^2=\frac{1}{\cos^2\alpha}\Leftrightarrow 1+\mathrm{tg}^2\alpha=$ $=\frac{1}{\cos^2\alpha}$.

Тождество (4) доказано.

Тождество (4) показывает, в какой зависимости находятся тангенс и косинус одного и того же допустимого угла α. Зная одну из величин, входящих в тождество (4) для некоторого допустимого угла α, можно найти другую величину того же угла α. Действительно, по тождеству (4) $\cos^2\alpha=\frac{1}{1+\mathrm{tg}^2\alpha}$ или $\mathrm{tg}^2\alpha=\frac{1-\cos^2\alpha}{\cos^2\alpha}$. Полученные тождества равносильны следующим:

$$|\cos\alpha|=\frac{1}{\sqrt{1+\mathrm{tg}^2\alpha}}, \tag{5}$$

$$|\mathrm{tg}\,\alpha|=\frac{\sqrt{1-\cos^2\alpha}}{|\cos\alpha|}. \tag{6}$$

Из тождества (5) имеем

$$\cos\alpha=\frac{1}{\sqrt{1+\mathrm{tg}^2\alpha}} \qquad \text{(5а)}$$

для любого α, удовлетворяющего при некотором k одному из двойных числовых неравенств:

$2\pi k-\pi/2<\alpha<\pi/2+2\pi k$,
где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$.

$$\cos\alpha=\frac{1}{-\sqrt{1+\mathrm{tg}^2\alpha}} \qquad \text{(5б)}$$

для любого α, удовлетворяющего при некотором k одному из двойных числовых неравенств:

$2\pi k+\pi/2<\alpha<3\pi/2+2\pi k$,
где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$.

Из тождества (6) имеем

$$\mathrm{tg}\,\alpha=\frac{\sqrt{1-\cos^2\alpha}}{\cos\alpha} \qquad \text{(6а)}$$

для любого угла α, для которого $\mathrm{tg}\,\alpha$ и $\cos\alpha$ одного знака (т. е. для любого угла α, удовлетворяющего при некотором m одному из двойных числовых неравенств:

$2\pi m\leqslant\alpha<\pi/2+2\pi m$,
$\pi/2+2\pi m<\alpha\leqslant\pi+2\pi m$,
где $m=0, \pm 1, \pm 2, \pm 3, \pm\ldots$).

$$\mathrm{tg}\,\alpha=\frac{-\sqrt{1-\cos^2\alpha}}{\cos\alpha} \qquad \text{(6б)}$$

для любого угла α, для которого $\mathrm{tg}\,\alpha$ и $\cos\alpha$ разных знаков (т. е. для любого угла α, удовлетворяющего при некотором m одному из двойных числовых неравенств:

$2\pi m-\pi\leqslant\alpha<-\pi/2+2\pi m$,
$-\pi/2+2\pi m<\alpha\leqslant 2\pi m$,
где $m=0, \pm 1, \pm 2, \pm 3, \pm\ldots$).

Замечание. Формулы (6а) и (6б) при граничных значениях угла α ($\alpha=\pi m$, где $m=0, \pm 1, \pm 2, \pm 3, \ldots$) дают одно и то же значение $\mathrm{tg}\,\alpha=0$.

Примеры. 1. Вычислить $\operatorname{tg}\alpha$, если $\cos\alpha = -\sqrt{5}/5$ и $\pi < \alpha < 3\pi/2$.

В указанном промежутке $\operatorname{tg}\alpha$ положителен, а $\cos\alpha$ отрицателен, и поэтому $\operatorname{tg}\alpha = \frac{-\sqrt{1-\cos^2\alpha}}{\cos\alpha} = 2$.

2. Вычислить $\cos\alpha$, если $\operatorname{tg}\alpha = -3$ и $-\pi/2 < \alpha < 0$.

В указанном промежутке $\cos\alpha$ положителен и поэтому $\cos\alpha = \frac{1}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{\sqrt{10}}{10}$.

Следствие 2. Для любого угла α, такого, что $\alpha \neq \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$, справедливо тождество:

$$1 + \operatorname{ctg}^2\alpha = \frac{1}{\sin^2\alpha}. \qquad (7)$$

Доказательство. Так как $\alpha \neq \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$, то $\sin\alpha \neq 0$ и поэтому $\operatorname{ctg}\alpha$ определен и основное тригонометрическое тождество (1) можно почленно разделить на $\sin^2\alpha$. Тогда для любого такого α имеем $\frac{\cos^2\alpha + \sin^2\alpha}{\sin^2\alpha} = \frac{1}{\sin^2\alpha} \Leftrightarrow \frac{\cos^2\alpha}{\sin^2\alpha} + \frac{\sin^2\alpha}{\sin^2\alpha} = \frac{1}{\sin^2\alpha} \Leftrightarrow \left(\frac{\cos\alpha}{\sin\alpha}\right)^2 + 1 = \frac{1}{\sin^2\alpha} \Leftrightarrow \operatorname{ctg}^2\alpha + 1 = \frac{1}{\sin^2\alpha}$. Тождество (7) доказано.

Тождество (7) показывает, в какой зависимости находятся котангенс и синус одного и того же допустимого угла α. Зная одну из величин, входящих в тождество (7) для некоторого допустимого угла α, можно найти другую величину того же угла α. Действительно, по тождеству (7) $\sin^2\alpha = \frac{1}{(1+\operatorname{ctg}^2\alpha)}$, или $\operatorname{ctg}^2\alpha = \frac{(1-\sin^2\alpha)}{\sin^2\alpha}$. Полученные тождества равносильны следующим:

$$|\sin\alpha| = \frac{1}{\sqrt{1+\operatorname{ctg}^2\alpha}}, \qquad (8)$$

$$|\operatorname{ctg}\alpha| = \frac{\sqrt{1-\sin^2\alpha}}{|\sin\alpha|}. \qquad (9)$$

Из тождества (8) имеем

$$\sin\alpha = \frac{1}{\sqrt{1+\operatorname{ctg}^2\alpha}} \qquad (8а)$$

для любого α, удовлетворяющего при некотором n одному из следующих двойных числовых неравенств:

$$2\pi n < \alpha < \pi + 2\pi n,$$

где $n = 0, \pm 1, \pm 2, \pm 3, \pm \ldots$.

$$\sin\alpha = \frac{1}{-\sqrt{1-\operatorname{ctg}^2\alpha}} \qquad (8б)$$

для любого α, удовлетворяющего его при некотором n одному из следующих двойных числовых неравенств:

$$2\pi n + \pi < \alpha < 2\pi + 2\pi n,$$

где $n = 0, \pm 1, \pm 2, \pm 3, \ldots$.

Из тождества (9) имеем

$$\operatorname{ctg}\alpha=\frac{\sqrt{1-\sin^2\alpha}}{\sin\alpha} \quad (9а)$$

для любого угла α, для которого $\operatorname{ctg}\alpha$ и $\sin\alpha$ одного знака (т. е. для любого α, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств:

$$2\pi k-\pi/2\leqslant\alpha<2\pi k,$$
$$2\pi k<\alpha\leqslant\pi/2+2\pi k,$$

где $k=0, \pm1, \pm2, \pm3, \pm\ldots$).

$$\operatorname{ctg}\alpha=\frac{-\sqrt{1-\sin^2\alpha}}{\sin\alpha} \quad (9б)$$

для любого угла α, для которого $\operatorname{ctg}\alpha$ и $\sin\alpha$ разных знаков (т. е. для любого α, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств:

$$2\pi k+\pi/2\leqslant\alpha<\pi+2\pi k,$$
$$\pi+2\pi k<\alpha\leqslant 3\pi/2+2\pi k,$$

$k=0, \pm1, \pm2, \pm3, \pm\ldots$).

Замечание. Формулы (9а) и (9б) при граничных значениях угла α ($\alpha=\pi/2+\pi k$), где $k=0, \pm1, \pm2, \pm3, \pm\ldots$) дают одно и то же значение $\operatorname{ctg}\alpha=0$.

Примеры. 1. Пусть $\sin\alpha=-24/25$ и $\alpha\in(-\pi, -\pi/2)$. Найти $\operatorname{ctg}\alpha$.

Для указанного угла $\operatorname{ctg}\alpha$ положителен, а $\sin\alpha$ отрицателен, и поэтому

$$\operatorname{ctg}\alpha=\frac{-\sqrt{1-\sin^2\alpha}}{\sin\alpha}=7/24.$$

2. Вычислить $\sin\alpha$, если $\operatorname{ctg}\alpha=4$ и $\alpha\in(\pi, 3\pi/2)$.

Для указанного угла $\sin\alpha$ отрицателен и поэтому

$$\sin\alpha=\frac{1}{-\sqrt{1+\operatorname{ctg}^2\alpha}}=-\frac{\sqrt{17}}{17}.$$

Из определения тангенса и котангенса одного и того же угла следует справедливость следующего утверждения.

Для любого угла α, такого, что $\alpha\neq\frac{\pi}{2}n$, где $n=0, \pm1, \pm2, \pm3, \pm\ldots$, справедливо тождество:

$$\operatorname{tg}\alpha\operatorname{ctg}\alpha=1. \quad (10)$$

Тождество (10) равносильно тождествам:

$$\operatorname{tg}\alpha=\frac{1}{\operatorname{ctg}\alpha}, \quad (11)$$

$$\operatorname{ctg}\alpha=\frac{1}{\operatorname{tg}\alpha}. \quad (12)$$

Тождества (11) и (12) показывают, в какой зависимости находятся тангенс и котангенс одного и того же допустимого угла α. Зная одну из величин, входящих в тождества (11) и (12), можно найти другую величину того же угла.

Примеры. 1. Пусть $\sin\alpha=-3/5$ и $\alpha\in(-\pi, -\pi/2)$. Вычислить $\cos\alpha$, $\operatorname{tg}\alpha$, $\operatorname{ctg}\alpha$.

Для указанного угла $\cos\alpha$ отрицателен и поэтому

$$\cos\alpha=-\sqrt{1-\sin^2\alpha}=-4/5.$$

Так как $\cos\alpha \neq 0$, то $\operatorname{tg}\alpha = \sin\alpha/\cos\alpha = 3/4$. Аналогично, так как $\sin\alpha \neq 0$, то $\operatorname{ctg}\alpha = \cos\alpha/\sin\alpha = 4/3$.

2. Пусть $\operatorname{ctg}\alpha = -7$ и $\alpha \in (\pi/2, \pi)$. Вычислить $\cos\alpha$, $\sin\alpha$, $\operatorname{tg}\alpha$.

Так как $\operatorname{ctg}\alpha \neq 0$, то $\operatorname{tg}\alpha = 1/\operatorname{ctg}\alpha = -1/7$. Для указанного угла $\sin\alpha$ положителен, а $\cos\alpha$ отрицателен, следовательно, $\sin\alpha = 1/\sqrt{1+\operatorname{ctg}^2\alpha} = \sqrt{2}/10$, $\cos\alpha = -\sqrt{1-\sin^2\alpha} = -7\sqrt{2}/10$.

§ 5
ФОРМУЛЫ СЛОЖЕНИЯ

Теорема 1. Для любых углов α, β справедливо тождество

$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta, \tag{1}$$

т. е. косинус разности двух любых углов равен произведению косинуса первого угла на косинус второго угла плюс произведение синуса первого угла на синус второго угла.

Доказательство. Построим единичную окружность с центром в точке O. Для построения любого угла будем считать началом отсчета неподвижный единичный радиус OM этой окружности (рис. 58).

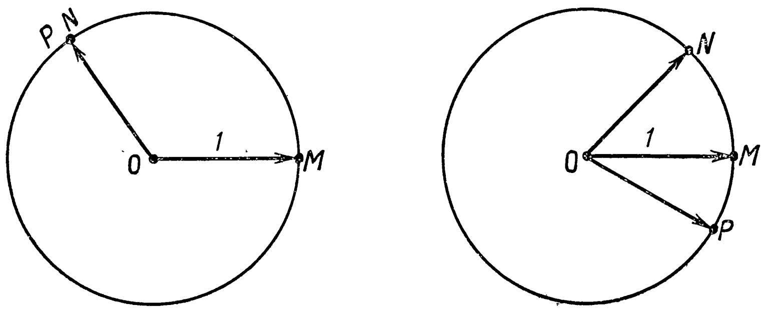


Рис. 58

Пусть угол α задается подвижным единичным радиусом ON, угол β — подвижным единичным радиусом OP. Возможны два случая расположения радиусов ON и OP. Они либо совпадают, либо не совпадают. Доказательство теоремы проведем отдельно в каждом из этих случаев.

1. Пусть радиусы ON и OP совпадают. В этом случае углы α и β таковы, что $\alpha = \beta + 2\pi k$ для некоторого фиксированного целого числа k и тождество (1) можно переписать в виде $\cos 2\pi k = \cos\beta\cos(\beta+2\pi k) + \sin\beta\sin(\beta+2\pi k)$ или, поскольку $\cos 2\pi k = 1$, $\cos(\beta+2\pi k) = \cos\beta$, $\sin(\beta+2\pi k) = \sin\beta$, в виде $1 = \cos^2\beta + \sin^2\beta$, т. е. в рассматриваемом случае тождество (1) есть равносильная форма записи основного тригонометрического тождества, следовательно, тождество (1) справедливо.

2. Пусть углы α и β таковы, что $\alpha \neq \beta + 2\pi k$ ни для какого целого числа k. Вычислим двумя способами длину отрезка PN. Сначала введем прямоугольную систему координат XOY так (рис. 59), чтобы единица масштаба совпала с уже выбранной ранее единицей длины; неподвижный единичный радиус OM оказался лежащим на положительной полуоси абсцисс (OX); положительная полуось ординат (OY) образовала бы положительный угол $\pi/2$ с положительной полуосью абсцисс (OX).

В выбранной системе координат точки N и P, как концы подвижных единичных радиусов, задающих углы α и β, имеют координаты $N(\cos\alpha, \sin\alpha)$; $P(\cos\beta, \sin\beta)$. По теореме о длине отрезка

$$d_{NP}^2 = (\cos\alpha - \cos\beta)^2 + (\sin\alpha - \sin\beta)^2. \tag{2}$$

Введем теперь другую прямоугольную систему координат $X'OY'$ так, чтобы единица масштаба совпала с уже выбранной ранее единицей длины; положительная полуось абсцисс (OX') была бы продолжением единичного радиуса OP; положительная полуось ординат (OY') образовывала бы положительный угол $\pi/2$ с положительной полуосью абсцисс (OX').

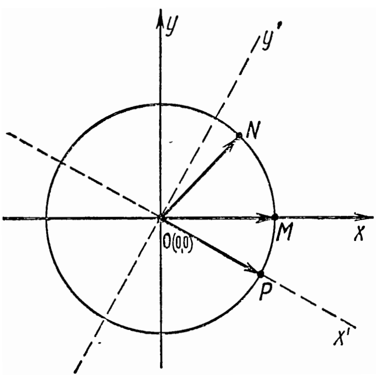


Рис. 59

В выбранной системе координат точка P будет иметь координаты $P(1, 0)$. Примем теперь радиус OP за неподвижный единичный радиус, т. е. за новое начало отсчета углов. Тогда относительно нового начала отсчета углов, радиус ON будет задавать угол $(\alpha - \beta)$ и в системе координат $X'OY'$ точка N будет иметь координаты $N[\cos(\alpha - \beta), \sin(\alpha - \beta)]$. По теореме о длине отрезка

$$d_{NP}^2 = [1 - \cos(\alpha - \beta)]^2 + [\sin(\alpha - \beta)]^2. \tag{3}$$

Используя свойство транзитивности равенств, из равенств (2) и (3) имеем

$$(\cos\alpha - \cos\beta)^2 + (\sin\alpha + \sin\beta)^2 = [1 - \cos(\alpha - \beta)]^2 + [\sin(\alpha - \beta)]^2.$$

Раскрывая скобки и группируя, имеем

$$(\cos^2\alpha + \sin^2\alpha) + (\cos^2\beta + \sin^2\beta) - 2(\cos\alpha\cos\beta + \sin\alpha\sin\beta) =$$
$$= 1 + [\cos^2(\alpha - \beta) + \sin^2(\alpha - \beta)] - 2\cos(\alpha - \beta).$$

Применяя трижды основное тригонометрическое тождество, перепишем это равенство в виде

$$2 - 2[\cos\alpha\cos\beta + \sin\alpha\sin\beta] = 2 - 2\cos(\alpha - \beta),$$

Откуда следует справедливость теоремы во втором случае. Теорема доказана.

Пример. Вычислить $\cos {}^{\pi}/_{12}$.

Поскольку $^{\pi}/_{12} = {}^{\pi}/_3 - {}^{\pi}/_4$, то $\cos {}^{\pi}/_{12} = \cos({}^{\pi}/_3 - {}^{\pi}/_4) = \cos {}^{\pi}/_3 \cos {}^{\pi}/_4 + \sin {}^{\pi}/_3 \sin {}^{\pi}/_4 = {}^1/_2 \cdot {}^{\sqrt{2}}/_2 + {}^{\sqrt{3}}/_2 \cdot {}^{\sqrt{2}}/_2 = {}^{\sqrt{2}(1+\sqrt{3})}/_4$.

Приведем несколько следствий из доказанной теоремы.

Следствие 1. Для любого угла α справедливы формулы $\cos({}^{\pi}/_2 - \alpha) = \sin\alpha$ и $\sin({}^{\pi}/_2 - \alpha) = \cos\alpha$.

Действительно, применяя тождество (1), имеем

$$\cos({}^{\pi}/_2 - \alpha) = \cos {}^{\pi}/_2 \cos\alpha + \sin {}^{\pi}/_2 \sin\alpha = \\ = 0 \cdot \sin\alpha + 1 \cdot \sin\alpha = \sin\alpha,$$

т. е. получаем справедливость первой формулы. Обозначая $({}^{\pi}/_2 - \alpha) = \beta$, получаем из уже доказанной первой формулы, что $\sin\beta = \cos({}^{\pi}/_2 - \beta)$. Так как $({}^{\pi}/_2 - \beta) = \alpha$, то $\sin({}^{\pi}/_2 - \alpha) = \cos\alpha$.

Два угла α и β, в сумме составляющие $^{\pi}/_2$ (т. е. $\alpha + \beta = {}^{\pi}/_2$), называются *дополнительными* один к другому. Так, угол α дополнительный к углу $({}^{\pi}/_2 - \alpha)$ и наоборот. Учитывая это определение, следствие 1 может быть сформулировано так: синус любого угла α равен косинусу дополнительного угла $({}^{\pi}/_2 - \alpha)$, а косинус любого угла β равен синусу дополнительного угла $({}^{\pi}/_2 - \beta)$.

Следствие 2. Для любых углов α и β справедливо тождество $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$, т. е. косинус суммы двух углов равен произведению косинуса первого угла на косинус второго угла минус произведение синуса первого угла на синус второго угла.

Доказательство. Представим $(\alpha + \beta)$ в виде $[\alpha - (-\beta)]$ и применим теорему 1. Используя четность косинуса и нечетность синуса, имеем $\cos(\alpha + \beta) = \cos[\alpha - (-\beta)] = \cos\alpha\cos(-\beta) + \sin\alpha\sin(-\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$, что и требовалось доказать.

Пример. Вычислить $\cos {}^{5\pi}/_{12}$.

Поскольку $^{5\pi}/_{12} = {}^{\pi}/_6 + {}^{\pi}/_4$, то $\cos {}^{5\pi}/_{12} = \cos({}^{\pi}/_6 + {}^{\pi}/_4) = \cos {}^{\pi}/_6 \cos {}^{\pi}/_4 - \sin {}^{\pi}/_6 \sin {}^{\pi}/_4 = {}^{\sqrt{3}}/_2 \cdot {}^{\sqrt{2}}/_2 - {}^1/_2 \cdot {}^{\sqrt{2}}/_2 = {}^{\sqrt{2}(\sqrt{3}-1)}/_4$.

Следствие 3. Для любых углов α и β справедливо тождество $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$, т. е. синус суммы любых двух углов равен произведению синуса первого угла на косинус второго плюс произведение косинуса первого угла на синус второго.

Доказательство. Используя следствие 1, затем тождество 1 и еще раз следствие 1, имеем

$$\sin(\alpha + \beta) = \cos({}^{\pi}/_2 - \alpha - \beta) = \cos[({}^{\pi}/_2 - \alpha) - \beta] =$$

$= \cos({}^{\pi}/_2 - \alpha)\cos\beta + \sin({}^{\pi}/_2 - \alpha)\sin\beta = \sin\alpha\cos\beta + \cos\alpha\sin\beta$, что и требовалось доказать.

Пример. Вычислить $\sin({}^{5\pi}/_{12})$

Поскольку $^{5\pi}/_{12}={}^{\pi}/_6+{}^{\pi}/_4$, то $\sin {}^{5\pi}/_{12}=\sin({}^{\pi}/_6+{}^{\pi}/_4)=$ $=\sin{}^{\pi}/_6\cos{}^{\pi}/_4+\sin{}^{\pi}/_4\cos{}^{\pi}/_6={}^1/_2\cdot{}^{\sqrt{2}}/_2+{}^{\sqrt{2}}/_2\cdot{}^{\sqrt{3}}/_2={}^{\sqrt{2}(1+\sqrt{3})}/_4$.

Следствие 4. Для любых углов α и β справедливо тождество $\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta$, т. е. синус разности любых двух углов равен произведению синуса первого угла на косинус второго угла минус произведение косинуса первого угла на синус второго угла.

Доказательство. Представим $(\alpha-\beta)$ в виде $[\alpha+(-\beta)]$ и применим следствие 3. Используя четность косинуса и нечетность синуса, имеем $\sin(\alpha-\beta)=\sin[\alpha+(-\beta)]=$ $=\sin\alpha\cos(-\beta)+\cos\alpha\sin(-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta$, что и требовалось доказать.

Пример. Вычислить $\sin({}^{\pi}/_{12})$.

Поскольку ${}^{\pi}/_{12}={}^{\pi}/_3-{}^{\pi}/_4$, то $\sin{}^{\pi}/_{12}=\sin({}^{\pi}/_3-{}^{\pi}/_4)=$ $=\sin{}^{\pi}/_3\cos{}^{\pi}/_4-\cos{}^{\pi}/_3\sin{}^{\pi}/_4={}^{\sqrt{3}}/_2\cdot{}^{\sqrt{2}}/_2-{}^1/_2\cdot{}^{\sqrt{2}}/_2={}^{\sqrt{2}(\sqrt{3}-1)}/_4$.

Для любых углов α и β справедливы следующие формулы:

$$\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta,$$
$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta.$$

Складывая и вычитая эти равенства, получаем формулы для вычисления произведения косинусов и произведения синусов:

$$\cos\alpha\cos\beta={}^1/_2[\cos(\alpha-\beta)+\cos(\alpha+\beta)], \quad (4)$$
$$\sin\alpha\sin\beta={}^1/_2[\cos(\alpha-\beta)-\cos(\alpha+\beta)]. \quad (5)$$

Эти формулы имеют следующие словесные формулировки.

Следствие 5. Произведение косинуса любого угла α на косинус любого угла β равно полусумме косинуса разности этих углов и косинуса суммы этих углов.

Следствие 6. Произведение синуса любого угла α на синус любого угла β равно полуразности косинуса разности этих углов и косинуса суммы этих углов.

Замечание. В формулах (4) и (5) в силу четности косинуса угла при взятии косинуса разности двух углов можно брать косинус и $(\alpha-\beta)$ и $(\beta-\alpha)$.

Примеры. 1. $\cos({}^{\pi}/_6+{}^{\alpha}/_2)\cos({}^{\pi}/_6-{}^{\alpha}/_2)={}^1/_2[\cos({}^{\pi}/_6+{}^{\alpha}/_2-$ $-{}^{\pi}/_6+{}^{\alpha}/_2)+\cos({}^{\pi}/_6+{}^{\alpha}/_2+{}^{\pi}/_6-{}^{\alpha}/_2)={}^1/_2(\cos\alpha+\cos{}^{\pi}/_3)=$ $={}^1/_2\cos\alpha+{}^1/_4$; **2.** $2\sin({}^{\pi}/_4-x)\sin({}^{\pi}/_4+x)=\cos({}^{\pi}/_4-x-{}^{\pi}/_4-$ $-x)-\cos({}^{\pi}/_4-x+{}^{\pi}/_4+x)=\cos(-2x)-\cos{}^{\pi}/_2=\cos 2x$.

Для любых углов α и β справедливы следующие формулы:

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta,$$
$$\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta.$$

Складывая эти равенства, получаем формулу для вычисления произведения синуса угла на косинус другого угла:

$$\sin\alpha\cos\beta={}^1/_2[\sin(\alpha+\beta)+\sin(\alpha-\beta)]. \quad (6)$$

Эта формула имеет следующую словесную формулировку.

Следствие 7. Произведение синуса любого угла α на косинус любого угла β равно полусумме синуса суммы углов α и β и синуса разности углов α и β, причем разность берется так, что от угла, стоящего под знаком синуса, вычитается угол, стоящий под знаком косинуса.

Пример. Вычислить $4\sin {}^{\pi}/_{6} \cos {}^{\pi}/_{3}$.

Применяя формулу (6), имеем $4\sin {}^{\pi}/_{6} \cos {}^{\pi}/_{3} = 2[\sin({}^{\pi}/_{6} + {}^{\pi}/_{3}) + \sin({}^{\pi}/_{6} - {}^{\pi}/_{3})] = 2[\sin {}^{\pi}/_{2} + \sin(-{}^{\pi}/_{6})] = 2[1 - {}^{1}/_{2}] = 1$.

Формулы (4), (5) и (6) часто читаются не слева направо, а наоборот, справа налево. Для более удобного запоминания формул (4), (5) и (6) в этих ситуациях обозначим

$$\begin{cases} \alpha + \beta = x, \\ \alpha - \beta = y. \end{cases} \tag{7}$$

Складывая эти равенства, получим

$$\begin{cases} \alpha = \frac{x+y}{2}, \\ \beta = \frac{x-y}{2}. \end{cases} \tag{8}$$

Из равенства (8) следует, что для любой пары x и y всегда найдется пара α и β, такая, что справедливы равенства (7). Если в формулах (4), (5) и (6) произвести замену α и β на x и y по формулам (7) и (8), то получим справедливость следующих формул:

$$\cos x + \cos y = 2\cos\frac{x+y}{2}\cos\frac{x-y}{2}, \tag{9}$$

$$\cos x - \cos y = 2\sin\frac{x+y}{2}\sin\frac{y-x}{2}, \tag{10}$$

$$\sin x + \sin y = 2\sin\frac{x+y}{2}\cos\frac{x-y}{2}. \tag{11}$$

Пользуясь нечетностью синуса угла, из формулы (11) получим

$$\sin x - \sin y = \sin x + \sin(-y) = 2\sin\frac{x-y}{2}\cos\frac{x+y}{2},$$

т. е. справедлива следующая формула:

$$\sin x - \sin y = 2\sin\frac{x-y}{2}\cos\frac{x+y}{2}. \tag{12}$$

Формулы (9), (10), (11) и (12) имеют следующие словесные формулировки.

Следствие 8. Сумма косинусов любых двух углов равна удвоенному произведению косинуса полусуммы этих углов на косинус полуразности этих углов.

Следствие 9. Разность косинусов любых двух углов равна удвоенному произведению синуса полусуммы этих углов на синус обратной полуразности этих углов (под обратной разностью углов понимается такая разность, когда из угла, стоящего под знаком вычитаемого косинуса, вычитается угол, стоящий под знаком уменьшаемого косинуса).

Следствие 10. Сумма синусов любых двух углов равна удвоенному произведению синуса полусуммы этих углов на косинус полуразности этих углов.

Следствие 11. Разность синусов любых двух углов равна удвоенному произведению синуса полуразности этих углов на косинус полусуммы этих углов, при этом синус полуразности берется так, что из угла, стоящего под знаком уменьшаемого синуса, вычитается угол, стоящий под знаком вычитаемого синуса.

Примеры. 1. $\cos 4x + \cos 6x = 2\cos\frac{(4x+6x)}{2}\cdot\cos\frac{4x-6x}{2} = 2\cos 5x\cos(-x) = 2\cos 5x\cos x.$

2. $\cos(\pi/3 - x) - \cos(\pi/6 + x) = 2\sin\frac{\pi/3 - x + \pi/6 + x}{2}\times\sin\frac{(\pi/6 + x - \pi/3 + x)}{2} = 2\sin(\pi/4)\sin(x - \pi/12) = \sqrt{2}\sin(x - \pi/12).$

3. $\sin x + 1/2 = \sin x + \sin \pi/6 = 2\sin(x/2 + \pi/12)\cdot\cos(x/2 - \pi/12).$

4. $\sin\alpha - \sin 5\alpha = 2\sin\frac{(\alpha - 5\alpha)}{2}\cdot\cos\frac{(\alpha+5\alpha)}{2} = 2\sin(-2\alpha)\times\cos 3\alpha = -2\sin 2\alpha\cos 3\alpha.$

Следствие 12. Тангенс суммы двух любых допустимых углов равен дроби, числитель которой есть сумма тангенсов этих углов, а знаменатель есть разность между единицей и произведением тангенсов тех же углов, т. е. для любых двух углов α и β, таких, что $\alpha \neq \pi/2 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm 3, \pm\ldots$, $\beta \neq \pi/2 + \pi m$, где $m = 0, \pm 1, \pm 2, \pm 3, \pm\ldots$, $(\alpha+\beta) \neq \pi/2 + \pi n$, где $n = 0, \pm 1, \pm 2, \pm 3, \ldots$, справедливо тождество

$$\operatorname{tg}(\alpha+\beta) = \frac{\operatorname{tg}\alpha + \operatorname{tg}\beta}{1 - \operatorname{tg}\alpha\operatorname{tg}\beta}.$$

Действительно, для любых допустимых углов α и β имеем:

$$\operatorname{tg}(\alpha+\beta) = \frac{\sin(\alpha+\beta)}{\cos(\alpha+\beta)} = \frac{\sin\alpha\cos\beta + \cos\alpha\sin\beta}{\cos\alpha\cos\beta - \sin\alpha\sin\beta} = \frac{\dfrac{\sin\alpha\cos\beta + \cos\alpha\sin\beta}{\cos\alpha\cos\beta}}{\dfrac{\cos\alpha\cos\beta - \sin\alpha\sin\beta}{\cos\alpha\cos\beta}} =$$

$$= \frac{\dfrac{\sin\alpha\cos\beta}{\sin\alpha\cos\beta} + \dfrac{\cos\alpha\sin\beta}{\cos\alpha\cos\beta}}{\dfrac{\cos\alpha\cos\beta}{\cos\alpha\cos\beta} - \dfrac{\sin\alpha\sin\beta}{\cos\alpha\cos\beta}} = \frac{\dfrac{\sin\alpha}{\cos\alpha} + \dfrac{\sin\beta}{\cos\beta}}{1 - \dfrac{\sin\alpha}{\cos\alpha}\dfrac{\sin\beta}{\cos\beta}} = \frac{\operatorname{tg}\alpha + \operatorname{tg}\beta}{1 - \operatorname{tg}\alpha\operatorname{tg}\beta}.$$

Пример. Вычислить $\operatorname{tg} 5\pi/12$.

Поскольку $5\pi/12 = \pi/4 + \pi/6$, то $\operatorname{tg} 5\pi/12 = \operatorname{tg}(\pi/4 + \pi/6) = \frac{\operatorname{tg}(\pi/4) + \operatorname{tg}(\pi/6)}{1 - \operatorname{tg}(\pi/4)\operatorname{tg}(\pi/6)} = \frac{1 + 1/\sqrt{3}}{1 - 1\cdot 1/\sqrt{3}} = \frac{\sqrt{3}+1}{\sqrt{3}-1} = \frac{(\sqrt{3}+1)(\sqrt{3}+1)}{(\sqrt{3}-1)(\sqrt{3}+1)} = 2 + \sqrt{3}.$

Следствие 13. Тангенс разности двух любых допустимых углов α и β равен дроби, числитель которой есть разность тангенса α и тангенса β, а знаменатель — сумма единицы и произведения тангенсов тех же углов, т. е. для любых двух углов α и β, таких, что $\alpha \neq {}^{\pi}/_2 + \pi k$, где $k=0, \pm1, \pm2, \pm3, \ldots$, $\beta \neq {}^{\pi}/_2 + \pi m$, где $m=0, \pm1, \pm2, \pm3, \pm\ldots$, $(\alpha-\beta) \neq {}^{\pi}/_2 + \pi n$, где $n=0, \pm1, \pm2, \pm3, \pm\ldots$, справедливо тождество

$$\operatorname{tg}(\alpha-\beta) = \frac{\operatorname{tg}\alpha - \operatorname{tg}\beta}{1+\operatorname{tg}\alpha\operatorname{tg}\beta}.$$

Действительно, для любых допустимых углов α и β имеем

$$\operatorname{tg}(\alpha-\beta) = \operatorname{tg}[\alpha+(-\beta)] = \frac{\operatorname{tg}\alpha+\operatorname{tg}(-\beta)}{1-\operatorname{tg}\alpha\operatorname{tg}(-\beta)} = \frac{\operatorname{tg}\alpha-\operatorname{tg}\beta}{1+\operatorname{tg}\alpha\operatorname{tg}\beta},$$

что и требовалось доказать.

Пример. Вычислить $\operatorname{tg}({}^{\pi}/_{12})$.

Поскольку ${}^{\pi}/_{12} = {}^{\pi}/_3 - {}^{\pi}/_4$, то $\operatorname{tg}{}^{\pi}/_{12} = \operatorname{tg}({}^{\pi}/_3 - {}^{\pi}/_4) =$

$$= \frac{\operatorname{tg}{}^{\pi}/_3 - \operatorname{tg}{}^{\pi}/_4}{1+\operatorname{tg}{}^{\pi}/_3\operatorname{tg}{}^{\pi}/_4} = \frac{\sqrt{3}-1}{\sqrt{3}+1} = \frac{(\sqrt{3}-1)(\sqrt{3}-1)}{(\sqrt{3}+1)(\sqrt{3}-1)} = 2-\sqrt{3}.$$

Аналогично получаются соответствующие формулы котангенса суммы и котангенса разности двух любых допустимых углов.

Следствие 14. Для любых углов α и β, таких, что $\alpha \neq \pi k$, где $k=0, \pm1, \pm2, \pm3, \pm, \ldots$, $\beta \neq \pi m$, где $m=0, \pm1, \pm2, \pm3, \pm, \ldots$, $(\alpha+\beta) \neq \pi n$, где $n=0, \pm1, \pm2, \pm3, \ldots$, справедливо тождество

$$\operatorname{ctg}(\alpha+\beta) = \frac{\operatorname{ctg}\alpha\operatorname{ctg}\beta - 1}{\operatorname{ctg}\alpha+\operatorname{ctg}\beta}.$$

Следствие 15. Для любых двух углов α и β, таких, что $\alpha \neq \pi k$, где $k=0, \pm1, \pm2, \pm3, \ldots$, $\beta \neq \pi m$, где $m=0, \pm1, \pm2, \pm3, \pm, \ldots$, $(\alpha-\beta) \neq \pi n$, где $n=0, \pm1, \pm2, \pm3, \pm\ldots$, справедливо тождество

$$\operatorname{ctg}(\alpha-\beta) = \frac{\operatorname{ctg}\alpha\operatorname{ctg}\beta+1}{\operatorname{ctg}\beta-\operatorname{ctg}\alpha}.$$

Примеры. 1. Вычислить $\operatorname{ctg}{}^{5\pi}/_{12}$.

Поскольку ${}^{5\pi}/_{12} = {}^{\pi}/_4 + {}^{\pi}/_6$, то $\operatorname{ctg}({}^{5\pi}/_{12}) = \operatorname{ctg}({}^{\pi}/_4 + {}^{\pi}/_6) =$

$$= \frac{\operatorname{ctg}{}^{\pi}/_4\operatorname{ctg}{}^{\pi}/_6 - 1}{\operatorname{ctg}{}^{\pi}/_4 + \operatorname{ctg}{}^{\pi}/_6} = \frac{1\cdot\sqrt{3}-1}{1+\sqrt{3}} = \frac{(\sqrt{3}-1)(\sqrt{3}-1)}{(\sqrt{3}+1)(\sqrt{3}-1)} = 2-\sqrt{3}.$$

2. Вычислить $\operatorname{ctg}{}^{\pi}/_{12}$.

Поскольку ${}^{\pi}/_{12} = {}^{\pi}/_3 - {}^{\pi}/_4$, то $\operatorname{ctg}{}^{\pi}/_{12} = \operatorname{ctg}({}^{\pi}/_3 - {}^{\pi}/_4) =$

$$= \frac{\operatorname{ctg}{}^{\pi}/_3\operatorname{ctg}{}^{\pi}/_4 + 1}{\operatorname{ctg}{}^{\pi}/_4 - \operatorname{ctg}{}^{\pi}/_3} = \frac{{}^{1}/_{\sqrt{3}}\cdot 1+1}{1-{}^{1}/_{\sqrt{3}}} = \frac{(\sqrt{3}+1)(\sqrt{3}+1)}{(\sqrt{3}-1)(\sqrt{3}+1)} = 2+\sqrt{3}.$$

§ 6
ТРИГОНОМЕТРИЧЕСКИЕ ОПЕРАЦИИ ДЛЯ ДВОЙНЫХ И ПОЛОВИННЫХ УГЛОВ

Пусть дан некоторый угол α (некоторое число α, представляющее радианную меру этого угла). Тогда под углом 2α понимается угол, радианная мера которого есть число 2α; угол 2α часто называется *двойным углом.*

Приведем некоторые формулы, выражающие тригонометрические операции для двойного угла 2α через тригонометрические операции для угла α.

1. Синус двойного угла 2α равен удвоенному произведению синуса угла α на косинус угла α, т. е.

$$\sin 2\alpha = 2\cos\alpha\sin\alpha. \tag{1}$$

Доказательство. Полагая $\alpha=\beta$ в формуле для синуса суммы двух углов $\sin(\alpha+\beta)=\sin\alpha\cos\beta+\sin\beta\cos\alpha$, получаем справедливость утверждения 1.

2. Косинус двойного угла 2α равен квадрату косинуса угла α минус квадрат синуса угла α, т. е.

$$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha. \tag{2}$$

Доказательство. Полагая $\alpha=\beta$ в формуле для косинуса суммы двух углов $\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta$, получаем справедливость утверждения 2.

3. Пусть дан угол α, такой, что $\alpha\neq{}^{\pi}/_2+\pi n$ и $\alpha\neq{}^{\pi}/_4+{}^{\pi n}/_2$ ни для какого целого n. Тогда для любого допустимого угла справедливо следующее утверждение: тангенс двойного угла 2α равен удвоенному тангенсу угла α, деленному на разность между единицей и квадратом тангенса угла α, т. е.

$$\operatorname{tg} 2\alpha = \frac{2\operatorname{tg}\alpha}{1-\operatorname{tg}^2\alpha}. \tag{3}$$

Доказательство. Представляя 2α как $\alpha+\alpha$ и применяя формулу тангенса суммы двух углов, имеем $\operatorname{tg}2\alpha = \operatorname{tg}(\alpha+\alpha)=\frac{\operatorname{tg}\alpha+\operatorname{tg}\alpha}{1-\operatorname{tg}\alpha\operatorname{tg}\alpha}=\frac{2\operatorname{tg}\alpha}{1-\operatorname{tg}^2\alpha}$, что и требовалось доказать.

Примеры. 1. Вычислить $\sin 2\alpha$ и $\cos 2\alpha$, зная, что $\pi/2<\alpha<\pi n$ и $\sin\alpha={}^3/_5$.

Для указанного угла $\cos\alpha=-\sqrt{1-\sin^2\alpha}=-{}^4/_5$. Применяя формулы 1 и 2, получаем $\sin 2\alpha=2\sin\alpha\cos\alpha=-{}^{24}/_{25}$, $\cos 2\alpha=\cos^2\alpha-\sin^2\alpha={}^7/_{25}$.

2. Пусть $\operatorname{tg}2\alpha=3$. Вычислить $\operatorname{tg}({}^{\pi}/_4+\alpha)-\operatorname{tg}({}^{\pi}/_4-\alpha)$.

Применяя формулы для тангенса суммы двух углов, затем для тангенса двойного угла, получаем

$$\operatorname{tg}({}^{\pi}/_4+\alpha)-\operatorname{tg}({}^{\pi}/_4-\alpha)=\frac{\operatorname{tg}({}^{\pi}/_4)+\operatorname{tg}\alpha}{1-\operatorname{tg}({}^{\pi}/_4)\operatorname{tg}\alpha}-\frac{\operatorname{tg}({}^{\pi}/_4)-\operatorname{tg}\alpha}{1+\operatorname{tg}({}^{\pi}/_4)\operatorname{tg}\alpha}=$$

$$=\frac{1+\operatorname{tg}\alpha}{1-\operatorname{tg}\alpha}-\frac{1-\operatorname{tg}\alpha}{1+\operatorname{tg}\alpha}=\frac{4\operatorname{tg}\alpha}{1-\operatorname{tg}^2\alpha}=2\operatorname{tg}2\alpha=6.$$

Рассмотрим угол $n\alpha$, где n — любое натуральное число. Под углом $n\alpha$ понимается угол, радианная мера которого есть число $n\alpha$. Можно вывести и соответствующие формулы, выражающие тригонометрические операции для угла $n\alpha$ через тригонометрические операции угла α. В качестве примера получим формулы для $\sin 3\alpha$ и $\cos 3\alpha$.

$$\begin{aligned}\sin 3\alpha &= \sin(2\alpha+\alpha) = \sin 2\alpha\cos\alpha + \cos 2\alpha\sin\alpha = \\ &= 2\sin\alpha\cos^2\alpha + (\cos^2\alpha - \sin^2\alpha)\sin\alpha = \\ &= 3\sin\alpha\cos^2\alpha - \sin^3\alpha = 3\sin\alpha - 4\sin^3\alpha; \\ \cos 3\alpha &= \cos(2\alpha+\alpha) = \cos 2\alpha\cos\alpha - \sin 2\alpha\sin\alpha = \\ &= (\cos^2\alpha - \sin^2\alpha)\cos\alpha - 2\sin\alpha\cos\alpha\sin\alpha = \\ &= \cos^3\alpha - 3\cos\alpha\sin^2\alpha = 4\cos^3\alpha - 3\cos\alpha.\end{aligned}$$

Пусть дан некоторый угол α (некоторое число α, представляющее радианную меру этого угла). Тогда под углом $\alpha/2$ понимается угол, радианная мера которого есть число $\alpha/2$; угол $\alpha/2$ часто называется *половинным углом*. Приведем некоторые формулы, выражающие тригонометрические операции половинного угла $\alpha/2$ через тригонометрические операции угла α. Выведем некоторые следствия из формулы для косинуса двойного угла. Очевидно, что угол α можно считать двойным углом по отношению к углу $\alpha/2$. Поэтому для любого угла α справедливо следующее тождество: $\cos^2\alpha = \cos^2(\alpha/2) - \sin^2(\alpha/2)$.

Применяя основное тригонометрическое тождество, можно заменить $\cos^2(\alpha/2)$ на $1-\sin^2(\alpha/2)$ или $\sin^2(\alpha/2)$ на $1-\cos^2(\alpha/2)$ и получить следующие два тождества:

$$\sin^2(\alpha/2) = \frac{1-\cos\alpha}{2}, \tag{4}$$

$$\cos^2(\alpha/2) = \frac{1+\cos\alpha}{2}. \tag{5}$$

Покажем, как отыскивать тригонометрические операции половинного угла $\alpha/2$, зная лишь косинус угла α.

Из тождества (4) вытекают следующие формулы:

$$\sin(\alpha/2) = \sqrt{\frac{1-\cos\alpha}{2}} \tag{4а}$$

для любого угла, для которого $\sin\alpha/2$ неотрицателен (т. е. для любого α, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств: $2\pi k \leqslant \alpha/2 \leqslant \pi + 2\pi k$, где $k=0, \pm1, \pm2, \pm3, \pm\ldots$).

$$\sin(\alpha/2) = -\sqrt{\frac{1-\cos\alpha}{2}} \tag{4б}$$

для любого угла, для которого $\sin\alpha/2$ неположителен (т. е. для любого α, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств: $-\pi + 2\pi k \leqslant \alpha/2 \leqslant 2\pi k$, где $k=0, \pm1, \pm2, \pm3, \pm\ldots$).

Замечание. Формулы (4а) и (4б) для граничных значений угла α ($\alpha = 2\pi k$, где $k=0, \pm1, \pm2, \pm3, \pm\ldots$) дают одно и то же значение $\sin(\alpha/2) = 0$.

Пример. Пусть $\sin\alpha = -\sqrt{32}/9$ и $\pi < \alpha < 3\pi/2$. Вычислить $\sin(\alpha/2)$.

Поскольку для рассматриваемого угла $\cos\alpha$ отрицателен, то $\cos\alpha = -\sqrt{1-\sin^2\alpha} = -\sqrt{1-(-32/9)^2} = -7/9$. Так как

$\pi < \alpha < {}^{3\pi}/_2$, то ${}^{\pi}/_2 < {}^{\alpha}/_2 < {}^{3\pi}/_4$, т. е. угол ${}^{\alpha}/_2$ задается подвижным единичным радиусом, находящимся во II четверти, следовательно, $\sin({}^{\alpha}/_2)$ положителен, поэтому $\sin({}^{\alpha}/_2) = \sqrt{(1-\cos\alpha)/_2} = {}^{2\sqrt{2}}/_3$.

Из тождества (5) вытекают следующие формулы:

$\cos({}^{\alpha}/_2) = \sqrt{(1+\cos\alpha)/_2}$ (5а)

для любого угла α, такого что $\cos({}^{\alpha}/_2)$ неотрицателен (т. е. для любого α, удовлетворяющего при некотором n одному из следующих двойных числовых неравенств:

$2\pi k - {}^{\pi}/_2 \leqslant {}^{\alpha}/_2 \leqslant {}^{\pi}/_2 + 2\pi k$,

где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$).

$\cos({}^{\alpha}/_2) = -\sqrt{(1-\cos\alpha)/_2}$ (5б)

для любого угла α, такого что $\cos({}^{\alpha}/_2)$ неположителен (т. е. для любого α, удовлетворяющего при некотором n одному из следующих двойных числовых неравенств:

${}^{\pi}/_2 + 2\pi k \leqslant {}^{\alpha}/_2 \leqslant {}^{3\pi}/_2 + 2\pi k$,

где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$).

Замечание. Формулы (5а) и (5б) для граничных значений угла α ($\alpha = \pi + 2\pi k$, где $k=0, \pm 1, \pm 2, \pm 3, \pm \ldots$) дают одно и то же значение $\cos({}^{\alpha}/_2) = 0$.

Пример. Пусть $\sin\alpha = {}^4/_5$ и $-{}^{3\pi}/_2 < \alpha < -\pi$. Вычислить $\cos({}^{\alpha}/_2)$.

Прежде всего найдем $\cos\alpha$. Поскольку для данного угла $\cos\alpha$ отрицателен, то $\cos\alpha = -\sqrt{1-\sin^2\alpha} = -3/5$. Так как $-3\pi/2 < \alpha < -\pi$, то $-3\pi/4 < \alpha/2 < -\pi/2$, т. е. угол $\alpha/2$ задается подвижным единичным радиусом, находящимся в третьей четверти, следовательно, $\cos(\alpha/2)$ также отрицателен и поэтому $\cos(\alpha/2) = -\sqrt{(1+\cos\alpha)/2} = -\sqrt{5}/5$.

Пусть теперь угол α таков, что $\alpha \neq \pi + 2\pi n$ ни для какого целого числа n, тогда для любого такого угла $\cos(\alpha/2) \neq 0$, а поэтому из тождеств (4) и (5) получаем равенство

$$\operatorname{tg}^2(\alpha/2) = \frac{1-\cos\alpha}{1+\cos\alpha},$$

из которого следует, что

$$\operatorname{tg}(\alpha/2) = \sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}}$$

для любого угла α, для которого $\operatorname{tg}(\alpha/2)$ неотрицателен (т. е. для любого α, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств

$k\pi \leqslant \alpha/2 < \pi/2 + k\pi$,

где $k=0, \pm 1, \pm 2, \pm \ldots$).

$$\operatorname{tg}(\alpha/2) = -\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}}$$

для любого угла α, для которого $\operatorname{tg}(\alpha/2)$ неположителен (т. е. для любого α, удовлетворяющего при некотором k одному из следующих двойных числовых неравенств $-\pi/2 + k\pi < \alpha/2 \leqslant \pi k$; где $k=0, \pm 1, \pm 2, \pm \ldots$).

Замечание. В случае граничных углов $\alpha = 2\pi k$ обе формулы дают одно и то же значение $\operatorname{tg}(\alpha/2) = 0$.

Пример. Пусть угол α таков, что $\cos 2\alpha = 7/32$ и $-2\pi < 2\alpha < -3\pi/2$. Вычислить $\operatorname{tg}(\alpha/2)$.

Прежде всего найдем $\cos\alpha$. Так как $-2\pi < 2\alpha < -3\pi/2$, то $-\pi < \alpha < -3\pi/4$ и, следовательно, $\cos\alpha$ отрицателен, поэтому

$$\cos\alpha = -\sqrt{(1-\cos 2\alpha)/2} = -{}^5/_8.$$

Так как $-\pi < \alpha < -3\pi/4$, то $-\pi/2 < \alpha/2 < -3\pi/8$ и, следовательно, $\operatorname{tg}{}^{\alpha}/_2$ также отрицателен, поэтому

$$\operatorname{tg}{}^{\alpha}/_2 = -\sqrt{(1-\cos\alpha)/(1+\cos\alpha)} = -\sqrt{39}/3.$$

Для тангенса половинного угла можно вывести и другие формулы. Пусть угол α таков, что $\alpha \neq \pi k$ ни для какого целого числа k. Тогда для любого такого угла α справедливо равенство

$$\operatorname{tg}\alpha/2 = (1-\cos\alpha)/\sin\alpha. \tag{6}$$

Действительно, так как $\alpha \neq \pi k$, то $\sin\alpha/2 \neq 0$, $\cos\alpha/2 \neq 0$ и справедливы следующие выкладки:

$$\operatorname{tg}{}^{\alpha}/_2 = \frac{\sin(\alpha/2)}{\cos(\alpha/2)} = \frac{\sin(\alpha/2)\,2\sin(\alpha/2)}{\cos(\alpha/2)\,2\sin(\alpha/2)} = \frac{2\sin^2(\alpha/2)}{2\sin(\alpha/2)\cos(\alpha/2)}.$$

Применяя формулы (1) и (4), получаем справедливость равенства (6).

Аналогично, если угол α таков, что $\alpha \neq \pi + 2\pi k$ ни для какого целого числа k, то для любого такого угла α справедливо равенство

$$\operatorname{tg}\alpha/2 = \sin\alpha/(1+\cos\alpha). \tag{7}$$

Пример. Пусть угол α таков, что $\alpha \neq \pi + 2\pi k$ и $\alpha \neq \pi/2 + \pi k$ ни для какого целого числа k. Упростить выражение $A = \frac{\sin 2\alpha}{1+\cos\alpha}\,\frac{\cos\alpha}{1+\cos 2\alpha}$.

Поскольку для рассматриваемого угла α: $\cos\alpha \neq 0$, $\cos\alpha \neq -1$, $\cos 2\alpha \neq -1$, то

$$A = \frac{\sin 2\alpha}{1+\cos\alpha}\,\frac{\cos\alpha}{1+\cos 2\alpha} = \frac{2\sin\alpha\cos\alpha}{1+\cos\alpha}\,\frac{\cos\alpha}{2\sqrt{\cos^2\alpha}} = \frac{\sin\alpha}{1+\cos\alpha}.$$

Так как в данном случае применима формула (7), то $A = \operatorname{tg}\alpha/2$.

Глава VI

ФУНКЦИИ И ИХ ГРАФИКИ

При рассмотрении количественных отношений явлений реального мира приходится иметь дело с численными значениями различных величин, например времени, пути, скорости, ускорения, объема и т. д. В зависимости от рассматриваемых условий одни из величин имеют постоянные числовые значения, у других — эти значения переменные. Такие величины называются соответственно постоянными и переменными. Например, при равномерном движении скорость v постоянна, время t и путь S переменные, причем $S=vt$. При свободном падении тела ускорение силы тяжести g постоянное, время движения t и пройденный путь S переменные, причем $S=gt^2/2$.

Изучение окружающих явлений показывает, что переменные величины изменяются не независимо друг от друга, а изменение численных значений одних из них влечет за собой изменение значений других. Здесь будут рассматриваться лишь *пары переменных,* значения одной (*зависимой*) из которых изменяются в зависимости от значений другой (*независимой*). В примерах $S=vt$ и $S=gt^2/2$ естественно считать t независимой переменной, а S — зависимой переменной. Приведем другие примеры таких пар переменных величин: площадь круга $S=\pi R^2$ (R — радиус, π — константа), где S меняется в зависимости от R; закон Бойля — Мариотта $p=c/V$ (V — объем некоторого газа, p — давление этого газа, c — некоторая константа), где p меняется в зависимости от V.

Независимая переменная может принимать не любые числовые значения, а лишь значения, продиктованные условиями рассматриваемой задачи. Например, при свободном падении тела время движения t берется из промежутка $0\leqslant t\leqslant\sqrt{2H/g}$, где H — высота, с которой падает тело. Можно привести и другие примеры, из которых вытекает, что по физическим соображениям или в силу иных условий **независимая переменная пробегает некоторое множество числовых значений.** Каждому такому значению независимой переменной соответствует некоторое значение зависимой переменной.

Подчеркнем, что в приведенных примерах каждому значению независимой переменной соответствует **только одно** значение зависимой переменной. Здесь будет рассматриваться только такая зависимость пар переменных. Приведем соответствующие определения.

§ 1
ОПРЕДЕЛЕНИЕ И ПРИМЕРЫ ФУНКЦИЙ

Функциональной зависимостью называется закон (правило), по которому каждому значению величины x (называемой независимой переменной, или *аргументом*) из некоторого множества чисел (называемого *областью определения функции*) ставится в соответствие одно вполне определенное значение величины y (называемой зависимой переменной, или *функцией*). Совокупность значений, которые принимает зависимая переменная y, называется *областью изменения функции*.

То же самое можно сказать другими словами.

Пусть задано некоторое множество чисел и известно, что независимая переменная x может принимать любое из этих значений, тогда переменная y называется *функцией* от переменной x на этом множестве, если по некоторому правилу или закону каждому значению x из этого множества ставится в соответствие одно определенное значение y. При этом независимая переменная x называется *аргументом* функции. Множество, на котором определена функция, называется областью *определения* функции. Множество значений, которые принимает при этом переменная y, называется областью *изменения* функции.

Замечание. Область определения функции может задаваться либо условиями решаемой задачи, либо физическим смыслом изучаемого явления, либо математическими соглашениями, а область изменения функции получается после задания области определения.

Приведем примеры функций.

1. Пусть n — некоторое натуральное число, тогда каждому действительному числу a можно поставить в соответствие (см. гл. IV) одно действительное число a^n. Значит, можно сказать, что указан закон, по которому каждому значению x из множества всех действительных чисел ставится в соответствие одно числовое значение x^n. Другими словами, задана функция $y=x^n$ с областью определения — множеством всех действительных чисел.

2. Пусть $(-n)$ — некоторое целое отрицательное число, тогда каждому отличному от нуля действительному числу a можно поставить в соответствие (см. гл. IV) одно действительное число a^{-n}. Значит, этим соответствием задана функция $y=x^{-n}$ с областью определения — множеством всех отличных от нуля действительных чисел.

3. Пусть α — некоторое положительное нецелое число, тогда каждому неотрицательному числу a можно поставить в соответствие (см. гл. IV) одно действительное число a^α. Значит, этим соответствием задана функция $y=x^\alpha$ с областью определения — множеством всех неотрицательных чисел.

4. Пусть $(-\alpha)$ — некоторое отрицательное нецелое число, тогда каждому положительному числу a можно поставить в соответствие (см. гл. IV) одно число $a^{-\alpha}$. Значит, этим соответствием задана функция $y=x^{-\alpha}$ с областью определения — множеством всех положительных чисел.

5. Пусть a — некоторое положительное, не равное единице число, тогда каждому действительному числу β можно поставить в соответствие (см. гл. IV) одно число a^{β}. Значит, этим соответствием задана функция $y=a^x$ с областью определения — множеством всех действительных чисел.

6. Пусть a — некоторое положительное, не равное единице число, тогда каждому положительному числу b можно поставить в соответствие (см. гл. IV) одно число $\log_a x$. Значит, этим соответствием задана функция $y=\log_\alpha b$ с областью определения — множеством всех положительных чисел.

7. Каждому действительному числу α (как радианной мере угла α) можно поставить в соответствие (см. гл. V) одно число $\sin\alpha$. Значит, этим соответствием задана функция $y=\sin x$ с областью определения — множеством всех действительных чисел.

8. Каждому действительному числу α (как радианной мере угла α) можно поставить в соответствие (см. гл. V) одно число $\cos\alpha$. Значит, этим соответствием задана функция $y=\cos x$ с областью определения — множеством всех действительных чисел.

9. Каждому действительному числу α, кроме $\alpha=\pi/2+k\pi$, где k — любое целое число (как радианной мере угла α), можно поставить в соответствие (см. гл. V) одно число $\operatorname{tg}\alpha$. Значит, этим соответствием задана функция $y=\operatorname{tg} x$ с областью определения — множеством всех действительных чисел, кроме $x=\pi/2+k\pi$, где k — любое целое число.

10. Каждому действительному числу α, кроме $\alpha=\pi k$, где k — любое целое число (как радианной мере угла α), можно поставить в соответствие (см. гл. V) одно число $\operatorname{ctg}\alpha$. Значит, этим соответствием задана функция $y=\operatorname{ctg} x$ с областью определения — множеством всех действительных чисел, кроме $x=\pi k$, где k — любое целое число.

Замечание. Функции, рассмотренные в примерах 1 — 10, называются *основными элементарными функциями.*

11. Рассмотрим функцию, определенную для любого действительного x по правилу: $y=1$, если x — рациональное число, $y=0$, если x — иррациональное число. Эта функция называется функцией Дирихле. Коротко эту функцию записывают так:

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

12. Рассмотрим функцию, определенную для любого действительного x по правилу: $y=1$, если x — положительное число,

$y = -1$, если x — отрицательное число, $y = 0$, если $x = 0$. Эта функция называется знаком x и обозначается так: $y = \operatorname{sign} x$. Определение этой функции коротко записывают так:

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

13. Рассмотрим функцию, определенную для любого действительного x по правилу: $y = n$, если x — положительное число, причем $x = n + \alpha$, где n — натуральное число и $0 \leqslant \alpha < 1$; $y = -m$, если x — отрицательное число, причем $x = -m + \beta$, где m — натуральное число и $0 \leqslant \beta < 1$; $y = 0$, если $0 \leqslant x < 1$. Эта функция называется целой частью x и обозначается $y = [x]$. Коротко функцию $y = [x]$ можно определить так: $[x]$ — есть наибольшее целое число, которое меньше или равно x.

14. Если всем действительным числам поставлено в соответствие одно и то же действительное число c, то говорят, что задана функция $y = c$, с областью определения — множеством всех действительных чисел.

§ 2
ИССЛЕДОВАНИЕ ФУНКЦИЙ

Область существования и область изменения функции. Первый вопрос, на который надо ответить при исследовании функции, — это вопрос об области определения и области изменения этой функции. Начнем с примера. Формула $S = \pi R^2$ определяет площадь круга S как функцию радиуса R. По смыслу задачи эта функция определена в области всех положительных чисел ($R > 0$). Отвлекаясь от содержания задачи, рассмотрим формулу $S = \pi R^2$. По этой формуле каждому значению R, где R — любое действительное число, можно поставить в соответствие определенное значение S, т. е. определить S как функцию R в области всех действительных чисел R. Таким образом, область определения R в формуле для вычисления площади круга оказалась лишь частью той естественной области, в которой формула $S = \pi R^2$ имеет смысл. С таким явлением приходится часто сталкиваться при решении конкретных задач. Если функция задана и не указана ее область определения, то в математике принято рассматривать естественную область определения этой функции.

Естественной областью определения, или областью существования, функции $y = f(x)$ называют совокупность всех действительных значений независимой переменной x, для каждого из которых функция принимает действительные значения. Итак, область существования функции определяется самим существом функции, а область *определения* ее задается условиями или смыслом решаемой задачи, т. е. областью определения функции может быть любая часть области существования функции или они могут полностью совпадать. Например, для функции $y=gx^2/2$ область существования $(-\infty, +\infty)$, а область ее определения при падении тела с высоты H есть $[0; \sqrt{2H/g}]$.

Область определения X функции $y=f(x)$ задается вместе с заданием самой функции, а область изменения функции вычисляется по данной области определения. Например, для функции $y=\sqrt{\sin x - 3}$ область существования — пустое множество, следовательно, и область изменения — пустое множество. Приведем примеры других функций, указав их области существования и области изменения (табл. 2).

Таблица 2

Функция	Область существования	Область изменения
$y=\sqrt{\log_2 \sin x}$	$x=\pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$	$y=0$
$y=\sqrt{1-x^2}$	$-1 \leqslant x \leqslant 1$	$0 \leqslant y \leqslant 1$
$y=1/\sqrt{1-x^2}$	$-1 < x < 1$	$1 \leqslant y < +\infty$

Ограниченность функций. Функция $y=f(x)$, определенная на множестве X, называется *ограниченной снизу* на данном множестве, если существует число A, такое, что $A \leqslant f(x)$ для любого $x \in X$.

Пример. Функция $y=a^x$ ограничена снизу на всей области существования, так как (см. гл. IV) $0 < a^x$ для любого действительного x.

Функция $y=f(x)$, определенная на множестве X, называется *ограниченной сверху* на данном множестве, если существует число B, такое, что $f(x) \leqslant B$ для любого $x \in X$.

Пример. Функция $y=\sqrt{1-x^2}$ ограничена сверху на всей области существования, так как $\sqrt{1-x^2} \leqslant 1$ для любого x, такого, что $-1 \leqslant x \leqslant 1$.

Функция $y=f(x)$, определенная на множестве X, называется *ограниченной* на данном множестве, если существует число $M > 0$, такое, что $|f(x)| \leqslant M$ для любого $x \in X$.

Пример. Функция $y=\sin x$ ограничена на всей области существования, так как $|\sin x| \leqslant 1$ для любого действительного x.

Можно доказать, что функция $y=f(x)$ ограничена тогда и только тогда, когда она одновременно ограничена и снизу и сверху.

Пример. Функция $y=\sqrt{1-x^2}$ ограничена на множестве существования $-1 \leqslant x \leqslant 1$, так как она снизу ограничена нулем, а сверху — единицей.

Четность и нечетность функций. Функция $y=f(x)$ называется *четной*, если область ее определения X есть множество, симметричное относительно начала координат, и если $f(-x)=f(x)$ при любом $x \in X$. Примеры четных функций: $y=x^2$, $y=\cos x$, $y=\sqrt{1-x^2}$.

Функция $y=f(x)$ называется *нечетной*, если область ее определения X есть множество, симметричное относительно начала координат, и если $f(-x)=-f(x)$ при любом $x \in X$. Примеры нечетных функций: $y=x$, $y=1/x$, $y=\sin x$.

Наряду с четными и нечетными функциями есть функции, не являющиеся ни теми, ни другими, например $y=2^x$, $y=\lg x$.

Теорема 1. Всякую функцию, определенную на множестве, симметричном относительно начала координат, можно представить в виде суммы двух функций, одна из которых четная, а другая нечетная.

Доказательство. Пусть функция $y=f(x)$ имеет область определения, симметричную относительно начала координат. Покажем, что существуют функции $\varphi(x)$ и $\psi(x)$, такие, что $f(x)=\varphi(x)+\psi(x)$, где $\varphi(x)$ — четная, а $\psi(x)$ — нечетная функции. Положим $\varphi(x)=\frac{f(x)+f(-x)}{2}$, $\psi(x)=\frac{f(x)-f(-x)}{2}$. Ясно, что $\varphi(-x)=\varphi(x)$, а $\psi(-x)=-\psi(x)$ и $\varphi(x)+\psi(x)=f(x)$. Теорема доказана.

Пример. Функцию $y=2^x$ можно представить в виде суммы двух функций $\varphi(x)=\frac{2^x+2^{-x}}{2}$ и $\psi(x)\frac{-2^{-x}+2^x}{2}$, функция $\varphi(x)$ — четная, $\psi(x)$ — нечетная.

Замечание. Наряду с понятием четной функции, т. е. функции, симметричной относительно оси ординат, можно ввести более общее понятие функции, симметричной относительно вертикальной прямой, проходящей через точку $(a, 0)$. Функция $y=f(x)$ называется *симметричной относительно вертикальной прямой*, проходящей через точку с координатами $(a, 0)$, если область ее определения есть множество, симметричное относительно точки $(a, 0)$ и при любом x из области определения $f(2a-x)=f(x)$.

Пример. Функция $y=\sin x$ симметрична относительно вертикальной прямой, проходящей через точку $(\pi/2, 0)$. Действительно, областью определения является вся числовая прямая, следовательно, это симметричное множество относительно любой точки, в том числе и относительно $\pi/2$; равенство $\sin(2\pi/2-x)=\sin x$ для любого действительного x очевидна.

Возрастание и убывание функций. Пусть функция $y=f(x)$ определена на промежутке (a, b).

Функция $y=f(x)$ называется *возрастающей* в промежутке $(a;\ b)$, если для любой пары чисел x_1 и x_2 этого промежутка из неравенства $x_1<x_2$ следует, что $f(x_1)<f(x_2)$.

Пример. Функция $y=x$ возрастает в промежутке $-\infty<x<\infty$; $y=x^2$ — в промежутке $0\leqslant x<\infty$; $y=\sin x$ — в промежутке $-\pi/2\leqslant x\leqslant \pi/2$.

Функция $y=f(x)$ называется *убывающей* в промежутке $(a;\ b)$, если для любой пары чисел x_1 и x_2 этого промежутка из неравенства $x_1<x_2$ следует, что $f(x_1)>f(x_2)$.

Пример. Функция $y=(1/2)^x$ убывает в промежутке $-\infty<x<\infty$; $y=\sin x$ — в промежутке ${}^{\pi}/_2\leqslant x\leqslant {}^{3\pi}/_2$.

Функция $y=f(x)$ называется *неубывающей* в промежутке $(a;\ b)$, если для любой пары чисел x_1 и x_2 этого промежутка из неравенства $x_1<x_2$ следует, что $f(x_1)\leqslant f(x_2)$.

Пример. Функция $y=\sqrt{x+|x|}$ является неубывающей в промежутке $-\infty<x<\infty$.

Функция $y=f(x)$ называется *невозрастающей* в промежутке $(a;\ b)$, если для любой пары чисел x_1 и x_2 этого промежутка из неравенства $x_1<x_2$ следует, что $f(x_1)\geqslant f(x_2)$.

Пример. Функция $y=\begin{cases} x^2 \text{ для } x<0, \\ 0 \text{ для } x\geqslant 0 \end{cases}$ является невозрастающей в промежутке $-\infty<x<+\infty$.

Функции возрастающие, убывающие, невозрастающие и неубывающие называются *монотонными* функциями. Функции возрастающие и убывающие называются *строго монотонными* функциями.

Периодичность функций. Функция $y=f(x)$ называется *периодической*, если существует число $T\neq 0$, такое, что для любого x из области определения функции $y=f(x)$ числа $(x+T)$ и $(x-T)$ также входят в область определения и для любого x из области определения $f(x+T)=f(x)$.

Замечание. Для периодической функции имеет место равенство $f(x-T)=f(x)$. Действительно, функция $y=f(x)$ в точке $(x-T)$ определена и $f(x)=f(x-T+T)=f(x-T)$.

Теорема 2. Если число T есть период функции $y=f(x)$, то и число $Q=mT\ (m=\pm 1,\ \pm 2,\ \pm 3,\ \pm\ldots)$ будет периодом этой функции.

Доказательство. Докажем, что для любого x из области определения функции $y=f(x)$ и любого натурального n:

а) точки $(x+nT)$ и $(x-nT)$ принадлежат области определения функции $y=f(x)$;

б) $f(x)=f(x+nT)$ и $f(x)=f(x-nT)$.

Пусть $n=1$, тогда согласно определению и замечанию:

а) точки $(x+T)$ и $(x-T)$ принадлежат области определения функции $y=f(x)$;

б) $f(x)=f(x+T)$ и $f(x)=f(x-T)$.

Предположим, что для $n=k$ справедливо утверждение а).

Докажем справедливость утверждения для $n=k+1$. Действительно, по предположению точки $(x+kT)$ и $(x-kT)$ принадлежат области определения функции $y=f(x)$ и T есть ее период. Следовательно, точки $[(x+kT)+T]$ и $[(x-kT)-T]$, т.е. точки $[x+(k+1)T]$ и $[x-(k+1)T]$ принадлежат области ее определения.

Итак, для любого x из области определения функции $y=f(x)$ при любом $m=\pm 1, \pm 2, \pm 3, \pm \ldots$ точка $(x+mT)$ принадлежит области ее определения.

Предположим теперь, что утверждение б) справедливо для любого $n=k$, т. е. $f(x)=f(x+kT)$ и $f(x)=f(x-kT)$. Докажем справедливость утверждения б) для $n=k+1$. Действительно, так как T является периодом функции $y=f(x)$, то для точки $(x+kT)$ имеем $f(x+kT)=f([x+kT]+T)=f(x+[k+1]T)$, но по предположению $f(x)=f(x+kT)$, следовательно, $f(x)=$ $=f(x+[k+1]T)$. Аналогично для точки $(x-kT)$ доказывается, что $f(x)=f(x-[k+1]T)$, т. е. для любого $m=\pm 1, \pm 2, \pm 3, \pm \ldots$ утверждение б) доказано.

Итак, теорема доказана.

Примеры. 1. Функция $y=\sin x$ имеет периодом число $T=2\pi$, так как для любого x числа $(x+2\pi)$ и $(x-2\pi)$ входят в область определения этой функции и $\sin(x+2\pi)=\sin x$.

2. Функция $y=x-[x]$ имеет период $T=1$, так как она определена для любого x и $(x+1)-[x+1]=x-[x]$.

3. Функция, определенная следующим образом:

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

имеет периодом любое рациональное число,

4. Функция $y=\sin(\sqrt{x})$ не является периодической, так как, например, для числа $x=0$, число $x-T$ $(T>0)$ и число $x+T$ $(T<0)$ не принадлежат области определения этой функции.

Число T называется *главным периодом*, если оно положительно и является наименьшим среди всех положительных периодов. Заметим, что функция, приведенная в примере 3, не имеет главного периода.

Экстремальные значения. Наименьшее или наибольшее значение функции, принимаемое ею на данном промежутке $(a; b)$, называется экстремальным значением функции на этом промежутке.

Примеры. 1. Функция $y=\sin x$ на промежутке $0 \leqslant x \leqslant 2\pi$ принимает экстремальные значения: наибольшее $y=+1$ (при $x=\pi/2$) и наименьшее $y=-1$ (при $x=-\pi/2$).

2. Функция $y=x^2$ на промежутке $-\infty<x<+\infty$ принимает наименьшее значение $y=0$ (при $x=0$), но нет такого x из области существования функции, где она принимает наибольшее значение.

3. Функция $y=2^x$ на промежутке $-\infty<x<+\infty$ не принимает экстремальных значений, но на промежутке $0\leqslant x<+\infty$ принимает наименьшее значение $y=1$ (при $x=0$).

График функции. Введем на плоскости прямоугольную систему координат XOY и рассмотрим функцию $y=f(x)$. Придавая x последовательно значения $x_1, x_2, \ldots, x_n, \ldots$, получим соответствующие значения $y_1, y_2, \ldots, y_n, \ldots$. Отметим на плоскости точки с координатами $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n), \ldots$.

Графиком функции $y=f(x)$ называется множество точек на плоскости, удовлетворяющее следующим условиям:

а) всякая точка с координатами (x_0, y_0), где $y_0=f(x_0)$, принадлежит этому множеству;

б) всякая точка, принадлежащая геометрическому месту точек, имеет координаты (x_1, y_1), такие, что $y_1=f(x_1)$.

Другими словами, график функции $y=f(x)$ — это множество всех точек, координаты которых удовлетворяют условию $y=f(x)$ и не содержащее других точек.

Например, графиком функции $y=\sqrt{\log_2 \sin x}$ будет множество точек плоскости $(\pi/2+2\pi k, 0)$, где k — любое целое число, и только эти точки.

Итак, при исследовании функции необходимо ответить на следующие вопросы:

а) Какова область существования функции?

б) Какова область изменения функции?

в) Ограниченная ли это функция?

г) Есть ли у нее экстремальные значения?

д) Периодическая ли она?

е) Является ли эта функция четной или нечетной, или ни той и ни другой?

ж) Есть ли у нее промежутки, где она монотонна?

з) Есть ли точки пересечения графика с осями координат?

и) Каков график этой функции?

Замечание. Наглядность графика является незаменимым вспомогательным средством исследования функции, но он только иллюстрирует свойства функции, а не доказывает их.

Исследуем сначала основные элементарные функции, а затем, зная их свойства, перейдем к более сложным функциям.

§ 3
ОСНОВНЫЕ ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ

Линейная функция $y=x$. Зависимость $y=x$ называется прямой пропорциональной зависимостью. Легко проверяются следующие свойства этой функции:

а) область существования: $-\infty<x<+\infty$;

б) область изменения: $-\infty<x<+\infty$;

в) функция не ограничена ни снизу, ни сверху;

г) функция не имеет экстремальных значений;

д) функция не периодическая;

е) функция нечетная;

ж) функция возрастает на всем промежутке $-\infty < x < +\infty$;

з) точка (0, 0) — единственная точка пересечения с осями координат.

Теорема 1. График функции $y=x$ есть прямая, проходящая через начало координат и являющаяся биссектрисой первого и третьего координатных углов (рис. 60).

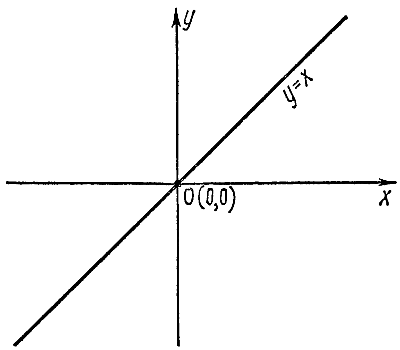


Рис. 60

Доказательство. Пусть точка $M_1(x_1\ y_1)$ принадлежит графику функции $y=x$. Тогда ее координаты удовлетворяют условию $y_1 = x_1$. Если $x_1 = y_1 = 0$, то точка M_1 совпадает с началом координат. Если $x_1 = y_1 \neq 0$, то числа x_1 и y_1 либо оба положительны, либо оба отрицательны, т. е. точка M_1 лежит либо в первом, либо в третьем координатном угле. Поскольку из условий $y_1 = x_1$ следует $|y_1| = |x_1|$, то точка M_1 равноудалена от осей координат. Следовательно, она лежит либо на биссектрисе первого координатного угла (если ее координаты положительны), либо на биссектрисе третьего координатного угла (если ее координаты отрицательны). Итак, любая точка графика функции $y=x$ либо совпадает с началом координат, либо лежит на одной из биссектрис первого или третьего координатных углов.

Наоборот. Так как координаты начала координат удовлетворяют условию $0=0$, то начало координат принадлежит графику функции $y=x$. Если точка $M_2(x_2,\ y_2)$ лежит на одной из биссектрис либо первого, либо третьего координатного угла, то $|x_2| = |y_2|$ (расстояния до осей координат равны). Так как числа x_2 и y_2 или оба положительны (если M_2 лежит в первом координатном угле) или оба отрицательны (если точка M_2 лежит в третьем координатном угле), то из условий $|y_2| = |x_2|$ следует $y_2 = x_2$, т. е. точка M_2 принадлежит графику функции $y=x$. Итак, начало координат и любая точка, лежащая на одной из биссектрисс либо первого, либо третьего координатных углов, принадлежит графику функции $y=x$.

По определению графика функции прямая, проходящая через начало координат и являющаяся биссектрисой первого и третьего координатных углов — график функции $y=x$ (рис. 60). Теорема доказана.

Гипербола $y = 1/x$. Зависимость $y = 1/x$ называется *обратной пропорциональной зависимостью*. Легко проверяются следующие свойства этой функции:

а) область существования: $-\infty < x < 0$, $0 < x < +\infty$;

б) область изменения: $-\infty < y < 0$, $0 < y < +\infty$;

в) функция не является ограниченной на всей области существования, но ограничена снизу ($y > 0$) на промежутке $0 < x < +\infty$ и ограничена сверху ($y < 0$) на промежутке $-\infty < x < 0$;

г) функция не принимает экстремальных значений;

д) функция не является периодической;

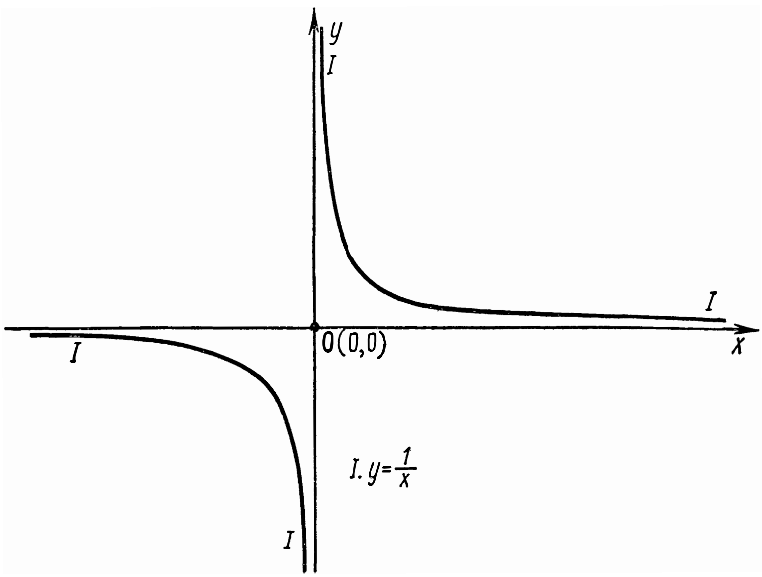


Рис. 61

е) функция нечетная;

ж) функция не является монотонной на всей области существования, но убывает на промежутке $-\infty < x < 0$, кроме того, она убывает и на промежутке $0 < x < +\infty$;

з) точек пересечения с осями нет.

Графиком функции $y = 1/x$ является линия, называемая *гиперболой* (рис. 61).

Теорема 2. График нечетной функции симметричен относительно начала координат.

Доказательство. Пусть дана нечетная функция $y = f(x)$ и пусть точка (x_0, y_0) лежит на ее графике. Тогда $y_0 = f(x_0)$ и в силу нечетности функции $-y_0 = f(-x_0)$, т. е. точка $(-x_0, -y_0)$, симметричная точке (x_0, y_0) относительно начала координат, тоже лежит на графике. Теорема доказана.

Замечание. Для построения графика нечетной функции достаточно построить его для $x \geqslant 0$. Для $x < 0$ он получается симметричным отображением построенной части графика относительно начала координат.

Парабола $y = x^2$. Зависимость $y = x^2$ называется *квадратичной зависимостью*. Легко проверяются следующие свойства этой функции:

а) область существования: $-\infty < x < +\infty$;

б) область изменения: $0 \leqslant y < +\infty$;

в) функция ограничена снизу: $y \geqslant 0$;

г) функция принимает наименьшее значение $y = 0$ при $x = 0$;

д) функция не является периодической;

е) функция четная;

ж) функция не является монотонной на всей области существования, но убывает на промежутке $-\infty < x \leqslant 0$ и возрастает на промежутке $0 \leqslant x < < +\infty$;

з) точка (0, 0) — единственная точка пересечения с осями координат.

Графиком функции $y = x^2$ является линия, называемая *параболой* (рис. 62).

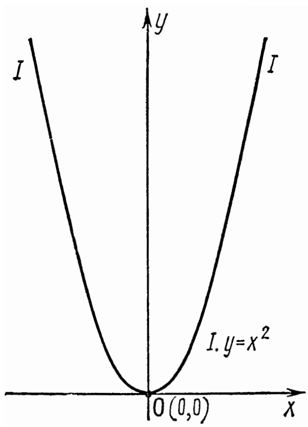


Рис. 62

Теорема 3. График четной функции симметричен относительно оси OY.

Доказательство. Пусть точка (x_0, y_0) лежит на графике четной функции $y = f(x)$, т. е. пусть $y_0 = f(x_0)$. В силу четности функции $y_0 = f(-x_0)$, т. е. точка $(-x_0, y_0)$, симметричная точке (x_0, y_0) относительно оси OY, тоже лежит на графике функции $y = f(x)$. Теорема доказана.

Замечание. Для построения графика четной функции достаточно построить его для $x \geqslant 0$. Для $x < 0$ он получается симметричным отображением построенной части графика относительно оси OY.

Степенная функция $y = x^\alpha$. Рассмотренные выше функции $y = x$, $y = x^2$, $y = 1/x$ являются частными случаями этой функции. Рассмотрим другие случаи.

1. $y = x^{2m}$ (m — некоторое натуральное число).

а) область существования: $-\infty < x < +\infty$;

б) область изменения: $0 \leqslant x < +\infty$;

в) функция ограничена снизу ($y \geqslant 0$);

г) функция принимает наименьшее значение $y = 0$ при $x = 0$;

д) функция не является периодической;

е) функция четная;

ж) функция не является монотонной на всей области существования, но убывает на промежутке $-\infty < x \leqslant 0$ и возрастает на промежутке $0 \leqslant x < +\infty$;

з) точка $(0, 0)$ — единственная точка пересечения с осями координат.

Графики функций $y=x^2$, $y=x^4$, $y=x^6$ изображены на рис. 63.

2. $y=x^{2m-1}$ (m — некоторое натуральное число).

а) область существования: $-\infty < x < +\infty$;

б) область изменения: $-\infty < y < +\infty$;

в) функция не является ограниченной ни сверху, ни снизу;

г) функция не принимает экстремальных значений;

д) функция не является периодической;

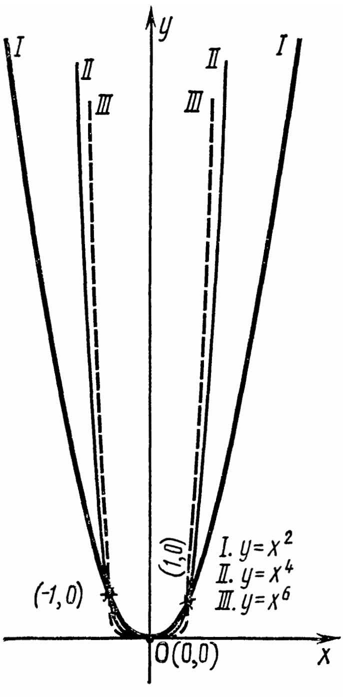


Рис. 63

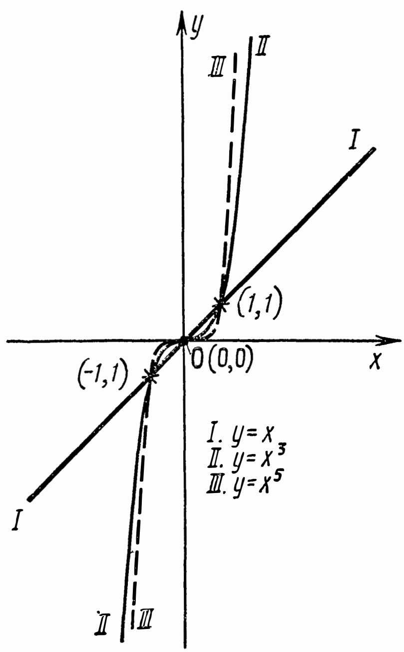


Рис. 64

е) функция нечетная;

ж) функция возрастает на всей области существования;

з) точка $(0, 0)$ — единственная точка пересечения с осями координат.

Графики функций $y=x$, $y=x^3$, $y=x^5$ изображены на рис. 64.

3. $y=x^{-2m}$ (m — некоторое натуральное число).

а) область существования: $-\infty < x < 0$, $0 < x < +\infty$;

б) область изменения: $0 < y < +\infty$;

в) функция ограничена снизу ($y > 0$);

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция четная;

ж) функция не является монотонной на всей области существования, но возрастает на промежутке $-\infty < x < 0$ и убывает на промежутке $0 < x < +\infty$;

з) точек пересечения с осями координат нет.

Графики функций $y = x^{-2}$, $y = x^{-4}$ изображены на рис. 65.

4. $y = x^{-2m+1}$ (m — некоторое натуральное число).

а) область существования: $-\infty < x < 0$, $0 < x + \infty$;

б) область изменения: $-\infty < y < 0$, $0 < y < +\infty$;

в) функция не является ограниченной ни сверху, ни снизу;

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция нечетная;

ж) функция не является монотонной на всей области существования, но убывает на промежутке

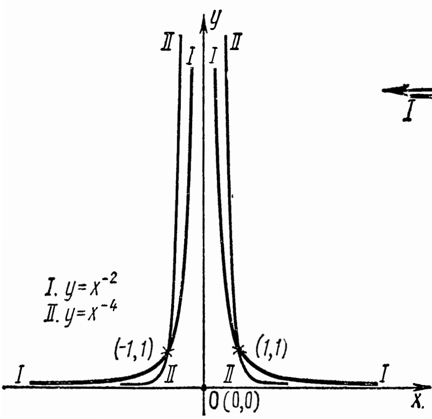


Рис. 65

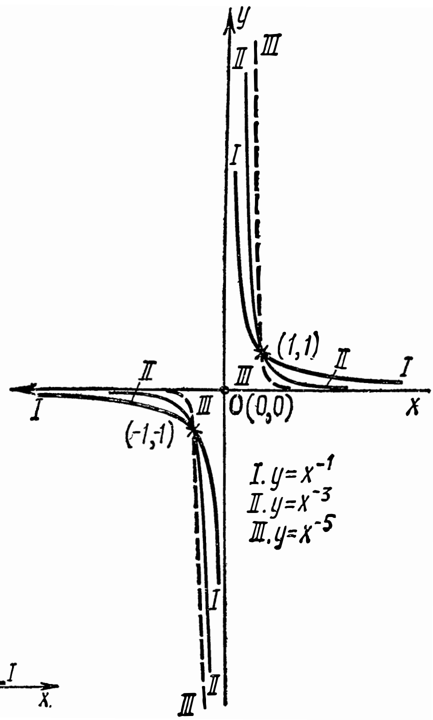


Рис. 66

$-\infty < x < 0$, кроме того, она убывает на промежутке $0 < x < +\infty$;

з) точек пересечения с осями координат нет.

Графики функций $y = x^{-1}$, $y = x^{-3}$, $y = x^{-5}$ изображены на рис. 66.

5. $y = x^{\alpha}$ (α — положительное нецелое число).

а) область существования: $0 \leqslant x < +\infty$;

б) область изменения: $0 \leqslant y < +\infty$;

в) функция ограничена снизу ($0 \leqslant y$);

г) функция принимает наименьшее значение $y = 0$ при $x = 0$;

д) функция не является периодической;

е) функция не является ни четной, ни нечетной;

ж) функция возрастает на всей области существования;

з) точка (0, 0) — единственная точка пересечения с осями координат.

График функции изображен на рис. 67.

6. $y = x^{-\alpha}$ (α — положительное нецелое число).

а) область существования: $0 < x < +\infty$;

б) область изменения: $0 < y < +\infty$;

в) функция ограничена снизу ($y > 0$);

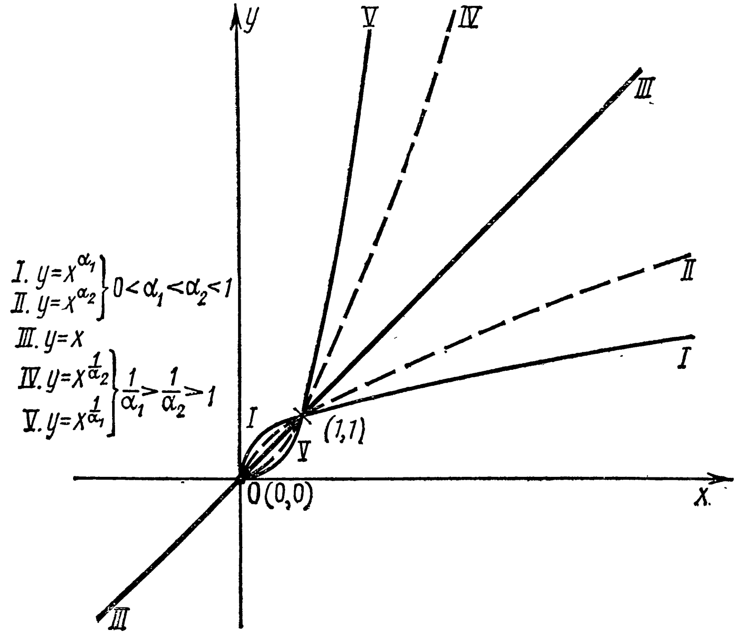


Рис. 67

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция не является ни четной, ни нечетной;

ж) функция убывает на всей области существования;

з) точек пересечения с осями координат нет.

График функции изображен на рис. 68.

Показательная функция $y = a^x$. Функция $y = a^x$, где $a > 0$ и $a \neq 1$ называется *показательной* функцией. Показательная функция обладает следующими свойствами:

а) область существования: $-\infty < x + \infty$;

б) область изменения: $0 < x < +\infty$;

в) функция ограничена снизу ($y > 0$);

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция не является ни четной, ни нечетной;

ж) если $a>1$, то функция $y=a^x$ возрастает на всей области существования; если $0<a<1$, то функция $y=a^x$ убывает на всей области существования;

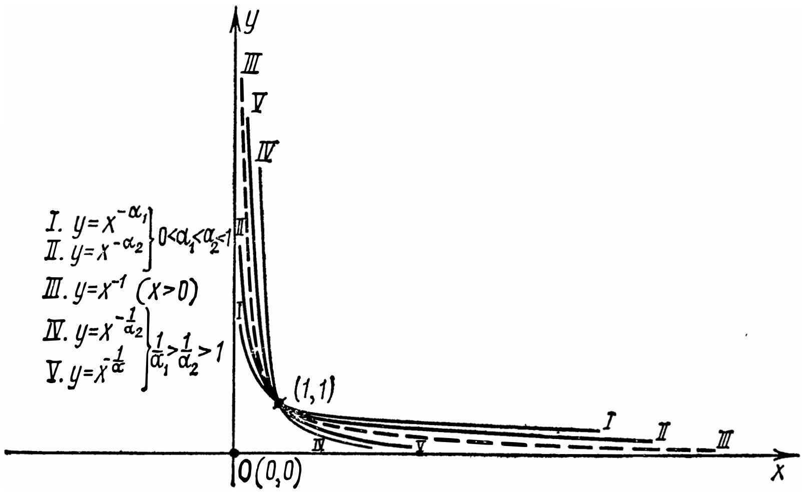


Рис. 68

з) точка (0, 1) — единственная точка пересечения с осями координат.

График функции $y=a^x$ при $a>1$ изображен на рис. 69, при $0<a<1$ — на рис. 70.

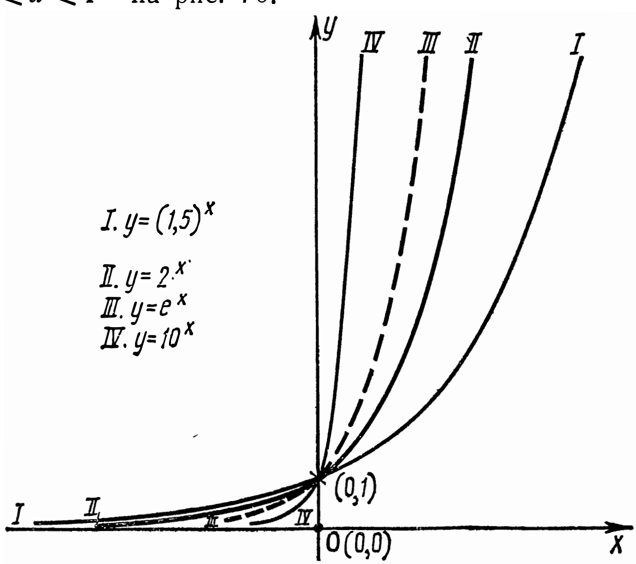


Рис. 69

Логарифмическая функция $y = \log_a x$. Функция $y = \log_a x$, где $a > 0$, $a \neq 1$, называется *логарифмической* функцией. Логарифмическая функция обладает следующими свойствами:

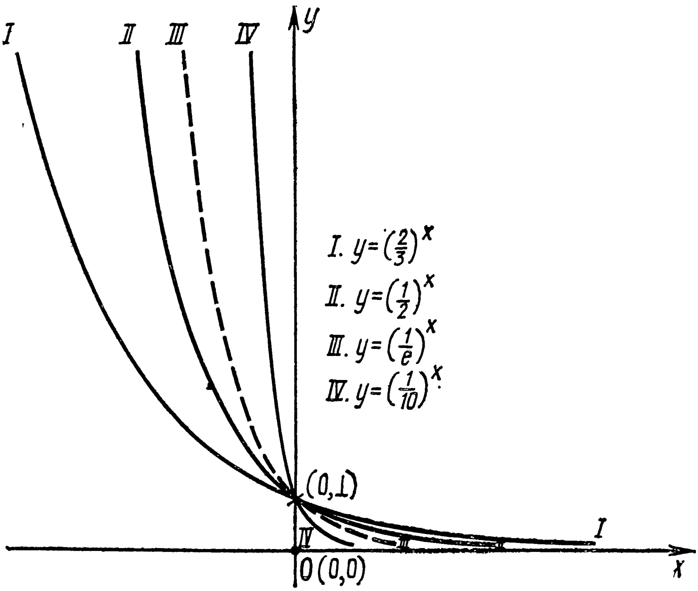


Рис. 70

а) область существования: $0 < x < +\infty$;

б) область изменения: $-\infty < y < +\infty$;

в) функция не является ограниченной ни сверху, ни снизу;

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция не является ни четной, ни нечетной;

ж) если $a > 1$, то функция $y = \log_a x$ возрастает на всей области существования; если $0 < a < 1$, то функция $y = \log_a x$ убывает на всей области существования;

з) точка $(1, 0)$ — единственная точка пересечения с осями координат.

График функции $y = \log_a x$ при $a > 1$ изображен на рис. 71, при $0 < a < 1$ — на рис. 72.

Основные тригонометрические функции. Прежде чем переходить к исследованию тригонометрических функций, докажем теорему.

Теорема 4. Для построения графика функции, имеющей главный период T, достаточно построить его на отрезке длины T, а затем продолжить периодически.

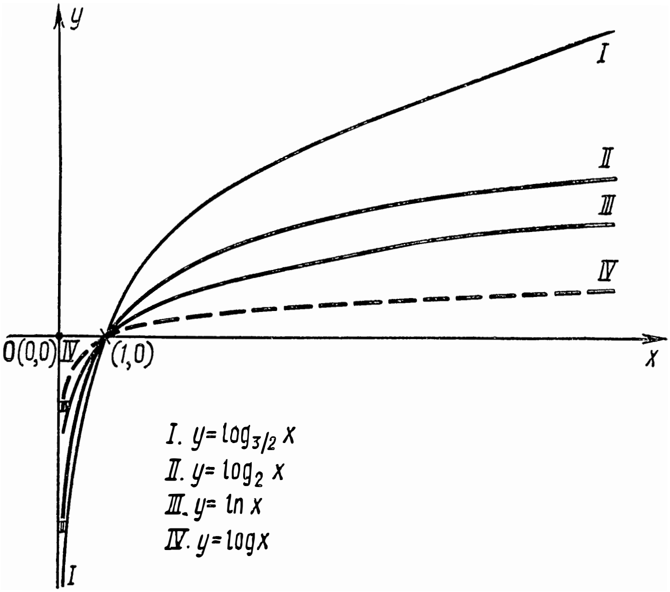

y
I
II
III
IV
O(0,0)
IV
(1,0)
x
I
I. $y=\log_{3/2} x$
II. $y=\log_2 x$
III. $y=\ln x$
IV. $y=\lg x$

Рис. 71

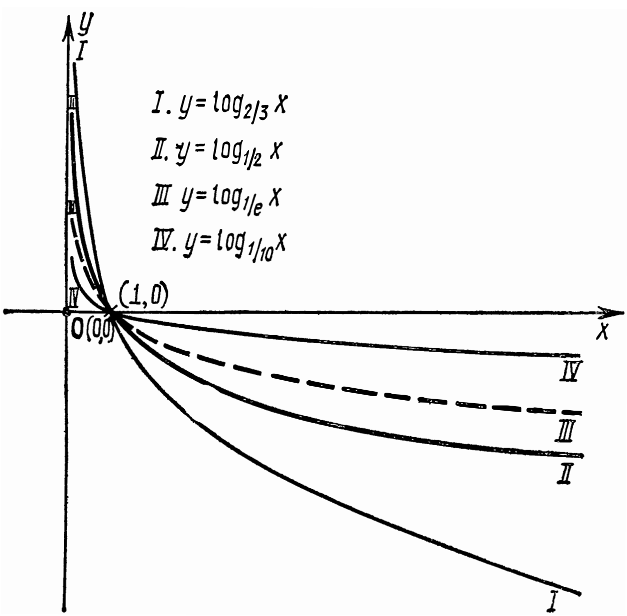

y
I
I. $y=\log_{2/3} x$
II. $y=\log_{1/2} x$
III $y=\log_{1/e} x$
IV. $y=\log_{1/10} x$
IV
(1,0)
O(0,0)
x
IV
III
II
I

Рис. 72

Доказательство вытекает из определения графика функции и определения периодичности функции. Так, если $f(x+nT)=f(x)$, то вместе с точкой (x_0, y_0) графику принадлежат точки (x_0+Tn, y_0).

Синусоида $y=\sin x$. Используя свойства операции нахождения синуса, получим следующие свойства функции $y=\sin x$:

а) область существования: $-\infty<x<+\infty$;

б) область изменения: $-1\leqslant y\leqslant 1$;

в) функция ограничена и снизу и сверху;

г) функция принимает наименьшее значение $y=-1$ при каждом $x=-\pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$, и наибольшее значение $y=1$ при каждом $x=\pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$;

д) функция периодическая, главный период $T=2\pi$;

е) функция нечетная;

ж) функция не является монотонной на всей области существования, но функция возрастает на каждом (при некотором k) промежутке $2\pi k-\pi/2\leqslant x\leqslant \pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$, и функция убывает на каждом (при некотором k) промежутке $2\pi k+\pi/2\leqslant x\leqslant 3\pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$.

Покажем, например, что на отрезке $[-\pi/2, \pi/2]$ функция $y=\sin x$ возрастает, т. е. что для любой пары чисел x_1 и x_2, такой, что $-\pi/2\leqslant x_1<x_2\leqslant \pi/2$, будет справедливо неравенство $\sin x_1<\sin x_2$. Для любой пары чисел x_1 и x_2 по формуле разности синусов

$$\sin x_1-\sin x_2=2\sin\frac{x_1-x_2}{2}\cos\frac{x_1+x_2}{2}. \qquad (1)$$

Докажем, что правая часть равенства (1) отрицательна, если $-\pi/2\leqslant x_1<x_2\leqslant \pi/2$. Условие $x_2\leqslant \pi/2$ равносильно условию $-\pi/2\leqslant -x_2$. Сложив это неравенство с неравенством $-\pi/2\leqslant x_1$, получим $-\pi\leqslant x_1-x_2$. Учитывая, что $x_1<x_2$ равносильно $x_1-x_2<0$, имеем $-\pi\leqslant x_1-x_2<0$, или $-\pi/2\leqslant \frac{x_1-x_2}{2}<0$. Следовательно, $\sin\frac{x_1-x_2}{2}<0$. Складывая неравенства $-\pi/2\leqslant x_1<\pi/2$ и $-\pi/2<x_2\leqslant \pi/2$, получаем $-\pi<x_1+x_2<\pi$, или $-\pi/2<(x_1+x_2)/2<\pi/2$. Следовательно, $\cos\frac{x_1+x_2}{2}>0$.

Итак, левая часть равенства (1) меньше нуля, следовательно, $\sin x_1<\sin x_2$.

Покажем, что на отрезке $[\pi/2, 3\pi/2]$ функция $y=\sin x$ убывает, т. е. что для любой пары x_1 и x_2, такой, что $\pi/2\leqslant x_1<x_2\leqslant 3\pi/2$ справедливо неравенство $\sin x_1>\sin x_2$. Прибавляя $-\pi$ к неравенствам $\pi/2\leqslant x_1<x_2\leqslant 3\pi/2$, имеем $-\pi/2\leqslant x_1-\pi<x_2-\pi\leqslant \pi/2$. В силу того, что на отрезке $[-\pi/2, \pi/2]$ функция $y=\sin x$ монотонно возрастает, то для $x_1-\pi$ и $x_2-\pi$

имеем

$$\sin(x_1-\pi)<\sin(x_2-\pi)\Leftrightarrow\sin[-(\pi-x_1)]<\sin[-(\pi-x_2)]\Leftrightarrow$$
$$\Leftrightarrow -\sin(\pi-x_1)<-\sin(\pi-x_2)\Leftrightarrow\sin(\pi-x_1)>\sin(\pi-x_2)\Leftrightarrow$$
$$\Leftrightarrow\sin x_1>\sin x_2,$$

что и требовалось доказать.

Аналогично доказывается, что функция $y=\sin x$ является возрастающей на каждом (при некотором k) промежутке $[2\pi k-\pi/2;\ 2\pi k+\pi/2]$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$, и убывающей на каждом (при некотором k) промежутке $[2\pi k+\pi/2,\ 2\pi k+3\pi/2]$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$.

з) точки пересечения с осями координат — точки $(\pi k,\ 0)$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$.

Учитывая периодичность, можно построить график функции $y=\sin x$, называемой *синусоидой* (рис. 73).

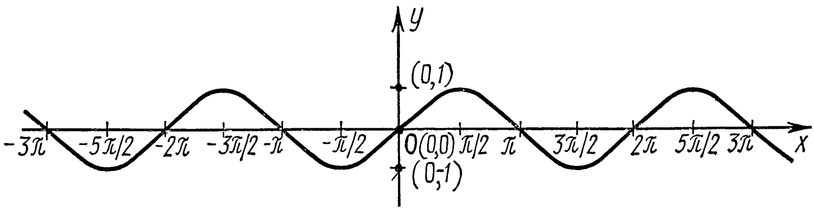


Рис. 73

Косинусоида $y=\cos x$. Используя свойства операции нахождения косинуса, получим следующие свойства функции $y=\cos x$:

а) область существования: $-\infty<x<+\infty$;

б) область изменения: $-1\leqslant y\leqslant 1$;

в) функция ограничена и снизу и сверху;

г) функция имеет экстремальные значения: наименьшее значение $y=-1$ достигается при $x=2\pi k+\pi$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$, наибольшее значение $y=1$ достигается при $x=2\pi k$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$;

д) функция периодическая, главный период $T=2\pi$;

е) функция четная;

ж) функция не является монотонной на всей области существования, но функция возрастает на каждом (при некотором k) промежутке $2\pi k-\pi\leqslant x\leqslant 2\pi k$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$, и функция убывает на каждом (при некотором k) промежутке $2\pi k\leqslant x\leqslant\pi+2\pi k$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$;

з) точка пересечения с осью OY — $(0,1)$; с осью OX точек пересечения бесконечно много, т. е. каждая (при некотором k) из точек $(\pi/2+\pi k,\ 0)$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$.

Учитывая периодичность, можно построить график функции $y=\cos x$, называемый *косинусоидой* (рис. 74).

Тангенсоида $y=\operatorname{tg} x$. Используя свойства операции нахождения тангенса, получим следующие свойства функции $y=\operatorname{tg} x$:

а) область существования: любое x, кроме $x=\pi/2+\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$;

б) область изменения: $-\infty < y < +\infty$;

в) функция не является ограниченной;

г) функция не принимает экстремальных значений;

д) функция периодическая, главный период $T=\pi$;

е) функция нечетная;

ж) функция не является монотонной на всей области существования, но функция возрастает на каждом (при некотором k)

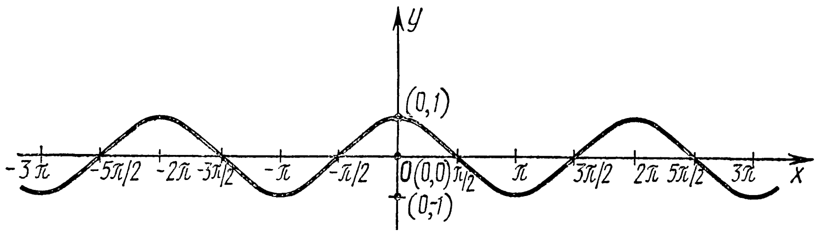


Рис. 74

из следующих промежутков $\pi k-\pi/2 < x < \pi/2+\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$;

з) точки пересечения с осями координат — точки $(\pi k, 0)$, где $k=0, \pm 1, \pm 2, \pm \ldots$;

Учитывая периодичность, можно построить график функции $y=\operatorname{tg} x$, называемый *тангенсоидой* (рис. 75).

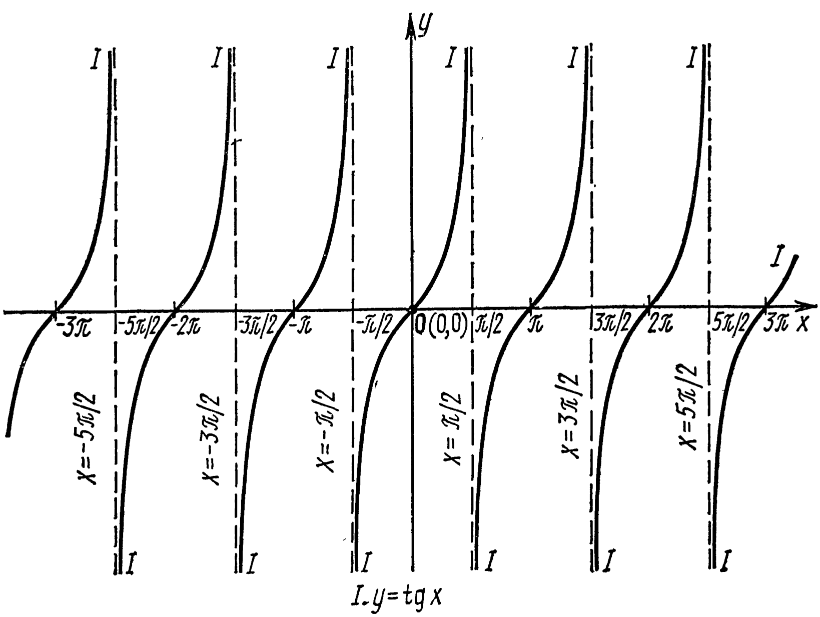


Рис. 75

Котангенсоида $y = \operatorname{ctg} x$. Используя свойства операции нахождения котангенса, получим следующие свойства функции $y = \operatorname{ctg} x$:

а) область существования: любое x, кроме $x = \pi k$, где $k = 0$, ± 1, ± 2, $\pm \ldots$;

б) область изменения: $-\infty < y < +\infty$;

в) функция не является ограниченной;

г) функция не принимает экстремальных значений;

д) функция периодическая, главный период $T = \pi$;

е) функция нечетная;

ж) функция не является монотонной на всей области существования, но функция убывает на каждом (при некотором k) из следующих промежутков: $\pi k < x < \pi + \pi k$, где $k = 0$, ± 1, ± 2, $\pm \ldots$;

з) точки пересечения с осями координат — точки $(\pi/2 + \pi k, 0)$, где $k = 0$, ± 1, ± 2, $\pm \ldots$.

Учитывая периодичность, можно построить график функции $y = \operatorname{ctg} x$, называемый *котангенсоидой* (рис. 76).

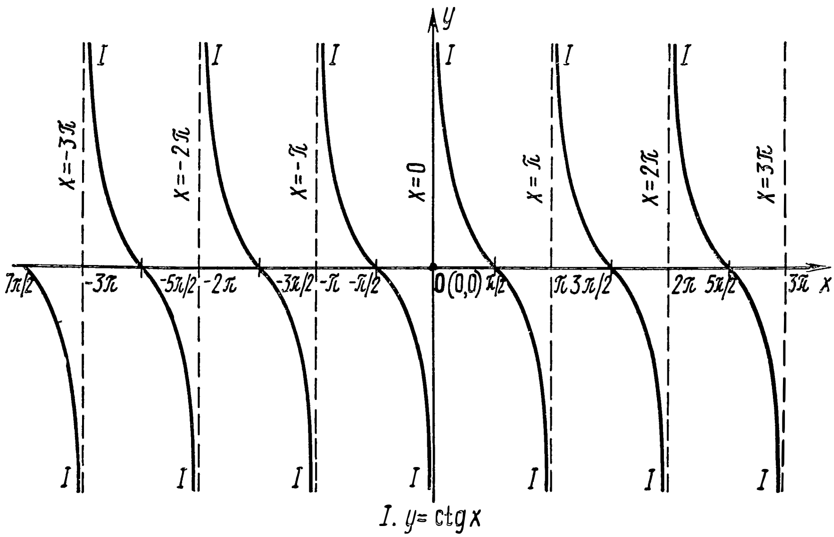


Рис. 76

§ 4
ОБРАТНЫЕ ФУНКЦИИ

Каждая функция $y = f(x)$ производит отображение области существования функции на область ее изменения так, что каждому x из области существования соответствует единственное значение y из области изменения.

Рассмотрим несколько функций вместе с их областями существования и изменения (табл. 3).

Таблица 3

функция	область существования	область изменения
$y=x^2$	$-\infty<x<\infty$	$0\leqslant y<+\infty$
$y=\sqrt{1-x^2}$	$-1\leqslant x\leqslant 1$	$0\leqslant y\leqslant 1$
$y=\dfrac{1}{1+x^2}$	$-\infty<x<+\infty$	$0<y<1$
$y=2^x$	$-\infty<x<+\infty$	$0<y<+\infty$

Некоторые из приведенных функций разным x из области существования ставят в соответствие разные y. Например, для функции $y=2^x$ каждое значение y из области изменения получается лишь при одном значении x из области существования. В таких случаях говорят, что функция осуществляет *взаимно-однозначное отображение* своей области существования на область изменения. Заметим, что все остальные функции, приведенные в таблице, таким свойством не обладают:

функция $y=x^2$ при $x=1$ и при $x=-1$ принимает одно и то же значение $y=1$;

функция $y=\sqrt{1-x^2}$ при $x=1/2$ и $x=-1/2$ принимает одно и то же значение $y=\sqrt{3}/2$;

функция $x=1/(1+x)^2$ при $x=2$ и $x=-2$ принимает одно и то же значение $y=1/5$.

Итак, функции можно разделить на два класса:

1) функции, осуществляющие взаимно-однозначное отображение области существования на область изменения;

2) функции, не обладающие этим свойством.

Если функции второго класса рассматривать не на всей области существования, то часто удается выбрать такую область определения (часть области существования), что функция будет отображать эту область определения на соответствующую область изменения уже взаимно-однозначно. Заметим, что любая функция $y=f(x)$ на той части области определения X (из области существования функции), где она является строго монотонной (т. е. возрастающей или убывающей), принадлежит к первому классу. Например, для функции $y=x^2$ такой областью будет промежуток $0\leqslant x<\infty$ или промежуток $-\infty<x\leqslant 0$.

Рассмотрим функцию $y=x^2$ на области определения $-\infty<x\leqslant 0$ (соответствующая ей область изменения $0\leqslant y<+\infty$). Эта функция осуществляет взаимно-однозначное отображение области $-\infty<x\leqslant 0$ на область $0\leqslant y<\infty$. При этом по каждому y из области $0\leqslant y<\infty$ можно однозначно восстановить x из области $-\infty<x\leqslant 0$. Действительно, если

$0 \leqslant y_0 < \infty$, то соответствующее значение x_0, такое, что $y_0 = x_0^2$, ищется по правилу $x_0 = -\sqrt{y_0}$. При этом если $y_0 \neq \bar{y}_0$, то $x_0 \neq \bar{x}_0$, где $x_0 = -\sqrt{y_0}$, а $\bar{x}_0 = -\sqrt{\bar{y}_0}$. Другими словами, можно установить взаимно-однозначное отображение области $0 \leqslant y < \infty$ на область $-\infty < x \leqslant 0$ по правилу $x = -\sqrt{y}$.

Итак, функция $y = x^2$ осуществляет взаимно-однозначное отображение области $-\infty < x \leqslant 0$ на область $0 \leqslant y < \infty$, а правило $x = -\sqrt{y}$ осуществляет взаимно-однозначное отображение области $0 \leqslant y < \infty$ на область $-\infty < x \leqslant 0$. Правило $x = -\sqrt{y}$ назовем *обратным правилом* для функции $y = x^2$ на области определения $-\infty < x \leqslant 0$ и области изменения $0 \leqslant y < +\infty$.

Функция $y = x^2$ и правило $x = -\sqrt{y}$, если $-\infty < x \leqslant 0$, а $0 \leqslant y < +\infty$, выражают одну и ту же связь между переменными x и y, только для любой пары (x_0, y_0) из рассматриваемых переменных величин функция $y = x^2$ дает возможность, зная x_0, найти y_0, а правило $x = -\sqrt{y}$ — зная y_0, найти x_0. Если в обратном правиле заменить x на y, а y на x с одновременной заменой области определения на область изменения и области изменения на область определения, то получим новую функцию $y = -\sqrt{x}$, у которой область определения $0 \leqslant x < +\infty$ и область изменения $-\infty < y \leqslant 0$. Новую функцию $y = -\sqrt{x}$ с областью определения $0 \leqslant x < +\infty$ и областью изменения $-\infty < y \leqslant 0$ называют обратной функцией к функции $y = x^2$ с областью определения $-\infty < x \leqslant 0$ и областью изменения $0 \leqslant y < +\infty$.

Для функции $y = x^2$ с областью определения $0 \leqslant x < +\infty$ и областью изменения $0 \leqslant y < +\infty$ обратной функцией будет функция $y = \sqrt{x}$ с областью определения $0 \leqslant x < +\infty$ и областью изменения $0 \leqslant y < +\infty$.

Приведем определение обратной функции в общем случае.

Пусть область определения функции $y = f(x)$ такова, что функция осуществляет взаимно-однозначное отображение области ее определения $a \leqslant x \leqslant b$ на область изменения $c \leqslant y \leqslant d$. Тогда, по каждому y из области $c \leqslant y \leqslant d$ можно однозначно восстановить x из области $a \leqslant x \leqslant b$ следующим образом: в равенстве $f(x) - y = 0$ считают y фиксированным и отыскивают x, удовлетворяющее этому равенству. Найденное x обозначается $f^{-1}(y)$, равенство $x = f^{(-1)}(y)$ называется обратным правилом. *Обратной функцией* к функции $y = f(x)$ $(a \leqslant x \leqslant b,\ c \leqslant y \leqslant d)$ называется функция, которая получается из обратного правила $x = f^{-1}(y)$ заменой x на y, а y на x с одновременной заменой области определения на область изменения, а области изменения на область определения. После такой замены область изменения функции $y = f(x)$ становится областью определения обратной функции $y=f^{-1}(x)$, а область определения функции $y=f(x)$ становится областью изменения обратной функции $y=f^{-1}(x)$.

Покажем на нескольких примерах, как отыскивается обратное правило и обратная функция. Во всех нижеприведенных примерах области определения выбраны так, что соответствующие функции осуществляют взаимно-однозначное отображение области определения на область изменения (табл. 4).

Таблица 4

Функция $y=f(x)$	Область определения $f(x)$	Область изменения $y=f(x)$	Обратное правило $x=f^{-1}(y)$	Обратная функция $y=f^{-1}(x)$	Область определения $y=f^{-1}(x)$	Область изменения $y=f^{-1}(x)$
$y=x^2$	$0\leqslant x<\infty$	$0\leqslant y<\infty$	$x=\sqrt{y}$	$y=\sqrt{x}$	$0\leqslant x<\infty$	$0\leqslant y<\infty$
$y=x^2$	$-\infty<x\leqslant 0$	$0\leqslant y<\infty$	$x=-\sqrt{y}$	$y=-\sqrt{x}$	$0\leqslant x<\infty$	$-\infty<y\leqslant 0$
$y=2^x$	$-\infty<x<\infty$	$0<y<\infty$	$x=\log_2 y$	$y=\log_2 x$	$0<x<\infty$	$-\infty<y<\infty$
$y=\sqrt{1-x^2}$	$0\leqslant x\leqslant 1$	$0\leqslant y\leqslant 1$	$x=\sqrt{1-y^2}$	$y=\sqrt{1-x^2}$	$0\leqslant x\leqslant 1$	$0\leqslant y\leqslant 1$
$y=\sqrt{1-x^2}$	$-1\leqslant x\leqslant 0$	$0\leqslant y\leqslant 1$	$x=-\sqrt{1-y^2}$	$y=-\sqrt{1-x^2}$	$0\leqslant x\leqslant 1$	$-1\leqslant y\leqslant 0$
$y=\frac{1}{1+x^2}$	$0\leqslant x<\infty$	$0<y\leqslant 1$	$x=\sqrt{1/y-1}$	$y=\sqrt{1/x-1}$	$0<x\leqslant 1$	$0\leqslant y<\infty$
$y=2x+1$	$-\infty<x<\infty$	$-\infty<y<+\infty$	$x=(y-1)/2$	$y=(x-1)/2$	$-\infty<x<\infty$	$-\infty<y<\infty$
$y=(x+1)^2$	$-1\leqslant x<\infty$	$0\leqslant y<\infty$	$x=\sqrt{y}-1$	$y=\sqrt{x}-1$	$0\leqslant x<\infty$	$-1\leqslant y<\infty$

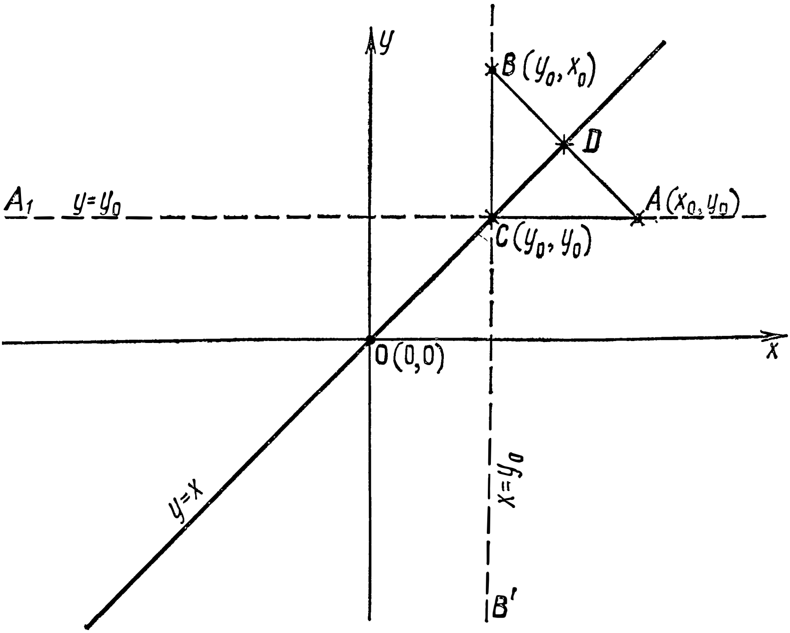
Замечание. Не для всякой функции удается найти такую область определения, на которой она взаимно-однозначна. Например, функция $y=1$ не является взаимно-однозначной ни на каком промежутке. В качестве другого примера можно привести функцию Дирихле.

График обратной функции. Прежде всего выясним, как расположены точки, координаты одной из которых получаются из координат другой заменой x на y, а y на x.

При замене x на y, а y на x любая точка, лежащая на прямой $y=x$, остается на месте, ибо абсцисса и ордината такой точки равны, и после перестановки их местами координаты этой точки не изменятся.

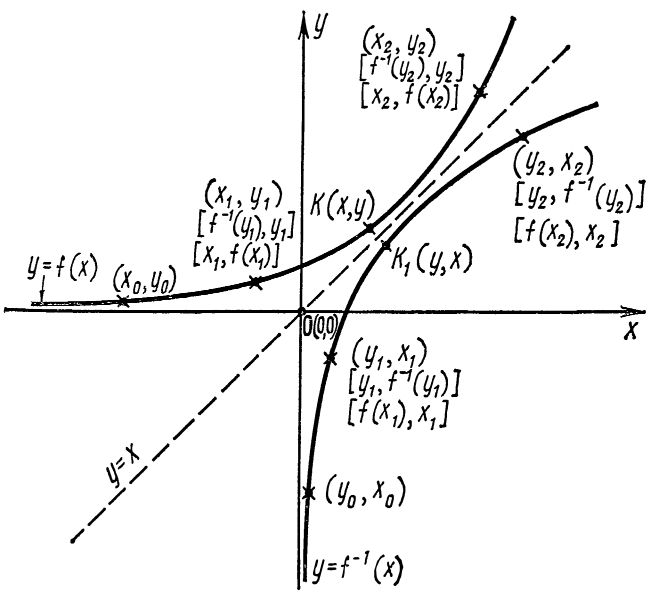
Теорема. *Любые точки $A(x_0, y_0)$ и $B(y_0, x_0)$ симметричны относительно прямой $y=x$.*

Доказательство. Если $x_0=y_0$, тогда точки A и B (рис. 77) совпадают и лежат на прямой $y=x$, т. е. в этом случае утверждение теоремы очевидно. Пусть $x_0\neq y_0$ и пусть точка $A(x_0, y_0)$ лежит в первой четверти и $x_0>y_0$. Проведем из точки A параллельно оси OX прямую AA_1, т. е. прямую $y=y_0$; из точки B — прямую параллельно оси OY, т. е. прямую $x=y_0$. Прямые AA_1 и BB_1 пересекутся в точке $C(y_0, y_0)$, т. е. в точке, лежащей на прямой $y=x$. Рассмотрим треугольник BCA, он прямоугольный (прямой угол BCA) и равнобедренный ($AC=BC=|x_0-y_0|$). Биссектриса CD угла BCA совпадает с прямой



y
B (y₀, x₀)
D
A₁ y=y₀
A (x₀, y₀)
C (y₀, y₀)
O (0,0)
x
y=x
x=y₀
B′

Рис. 77



y
(x₂, y₂)
[f⁻¹(y₂), y₂]
[x₂, f(x₂)]
(y₂, x₂)
[y₂, f⁻¹(y₂)]
[f(x₂), x₂]
(x₁, y₁)
[f⁻¹(y₁), y₁]
[x₁, f(x₁)]
K(x,y)
K₁(y,x)
y=f(x)
(x₀, y₀)
O(0,0)
x
(y₁, x₁)
[y₁, f⁻¹(y₁)]
[f(x₁), x₁]
y=x
(y₀, x₀)
y=f⁻¹(x)

Рис. 78

$y=x$. Поскольку треугольник BCA равнобедренный, то его биссектриса CD является и высотой и медианой, следовательно, $CD \perp AB$ и $AD=BD$. Это означает, что точки A и B симметричны относительно прямой $y=x$. Аналогично проводится доказательство теоремы в случае, когда точка $A(x_0, y_0)$ лежит в первой четверти и $x_0<y_0$, а также в случае, когда точка $A(x_0, y_0)$ лежит не в первой четверти. Теорема доказана.

Пусть функция $y=f(x)$ взаимно-однозначно отображает область определения $a \leqslant x \leqslant b$ на область изменения $c \leqslant y \leqslant d$. Тогда график этой функции таков, что по любому x_1 однозначно находится $y_1=f(x_1)$ и, наоборот, по любому y_2 однозначно находится $x_2=f^{-1}(y_2)$ (рис. 78).

Если точка $M(x_0, y_0)$ лежит на графике функции $y=f(x)$, то ее координаты удовлетворяют условию $y_0=f(x_0)$, а следовательно, и условию $x_0=f^{-1}(y_0)$. Другими словами, все точки (и только они) графика $y=f(x)$ удовлетворяют условию $x=$ $=f^{-1}(y)$. Так как для получения обратной функции надо в обратном правиле $x=f^{-1}(y)$ заменить x на y, а y на x, то каждая точка графика $y=f^{-1}(x)$ получается из точки графика функции $y=f(x)$ заменой x на y, а y на x, т. е. если точка $k(x, y)$ — точка графика $y=f(x)$, то точка $K_1(y, x)$ — точка графика $y=f^{-1}(x)$. Из этого рассуждения и теоремы вытекает, что график обратной функции $y=f^{-1}(x)$ получается из графика функции $y=f(x)$ симметричным отображением относительно прямой $y=x$ (рис. 78).

Замечание. Часто возникает ситуация, когда вид обратной функции получается непростым. Например, функция $y=x^{2m+1}$, где m — фиксированное натуральное число, взаимно-однозначна на всей числовой прямой. Для того чтобы найти обратное правило, с помощью которого строится обратная функция, необходимо решить уравнение $x^{2m+1}-y_0=0$. Как известно (см. гл. II), в области неотрицательных значений и x и y это уравнение имеет решение $x=\sqrt[2m+1]{y_0}$, а в области отрицательных значений и x и y уравнение имеет решение $x=-\sqrt[2m+1]{|y_0|}$. Поэтому обратное правило задается следующим образом:

$$x=\begin{cases} \sqrt[2m+1]{y}, & \text{если } y \geqslant 0, \\ -\sqrt[2m+1]{|y|}, & \text{если } y<0. \end{cases}$$

Следовательно, и обратная функция будет иметь вид

$$y=f^{-1}(x)=\begin{cases} \sqrt[2m+1]{x}, & \text{если } x \geqslant 0, \\ -\sqrt[2m+1]{|x|}, & \text{если } x<0. \end{cases}$$

Иногда обратную функцию к степенной функции $y=x^{2m+1}$ записывают с помощью одной формулы $y=\sqrt[2m+1]{x}$, но так делать не следует, так как операция взятия арифметического корня определена только для неотрицательных чисел (см. гл. IV).

Основные обратные тригонометрические функции. Для любой из основных тригонометрических функций $y=\sin x$, $y=\cos x$, $y=\operatorname{tg} x$, $y=\operatorname{ctg} x$ можно выделить много областей определения, каждую из которых соответствующая тригонометрическая функция отображает взаимно-однозначно на соответствующую область изменения. Поэтому для определения обратных тригонометрических функций условились каждую основную тригонометрическую функцию рассматривать на своей специально выбранной области определения (табл. 5).

Таблица 5

Основная тригонометрическая функция $y=f(x)$	Область определения $y=f(x)$	Область изменения $y=f(x)$	Обратное правило $x=f^{-1}(y)$	Обратная функция $y=f^{-1}(x)$	Область определения $y=f^{-1}(x)$	Область изменения $y=f^{-1}(x)$
$y=\sin x$	$-\pi/2\leqslant x\leqslant\pi/2$	$-1\leqslant y\leqslant 1$	$x=\arcsin y$	$y=\arcsin x$	$-1\leqslant x\leqslant 1$	$-\pi/2\leqslant y\leqslant\pi/2$
$y=\cos x$	$0\leqslant x\leqslant\pi$	$-1\leqslant y\leqslant 1$	$x=\arccos y$	$y=\arccos x$	$-1\leqslant x\leqslant 1$	$0\leqslant y\leqslant\pi$
$y=\operatorname{tg} x$	$-\pi/2<x<\pi/2$	$-\infty<y<+\infty$	$x=\operatorname{arctg} y$	$y=\operatorname{arctg} x$	$-\infty<x<+\infty$	$-\pi/2<y<\pi/2$
$y=\operatorname{ctg} x$	$0<x<\pi$	$-\infty<y<+\infty$	$x=\operatorname{arcctg} y$	$y=\operatorname{arcctg} x$	$-\infty<x<+\infty$	$0<y<\pi$

Приведем свойства и построим графики обратных тригонометрических функций.

1. $y=\arcsin x$.

а) Функция определена только на орезке $[-1;\ 1]$, т. е. $-1\leqslant x\leqslant 1$;

б) по определению множество всех значений функции составляет отрезок $[-\pi/2;\ \pi/2]$, т. е. $-\pi/2\leqslant\arcsin x\leqslant\pi/2$;

в) функция ограничена и снизу ($y\leqslant-\pi/2$) и сверху ($y\leqslant\pi/2$);

г) функция принимает наименьшее значение $y=-\pi/2$ при $x=-1$ и наибольшее значение $y=\pi/2$ при $x=1$;

д) функция не является периодической;

е) функция нечетная, т. е. $\arcsin(-x)=-\arcsin x$ для любого x, принадлежащего области существования.

Докажем это свойство. Обозначим $\alpha_1=\arcsin(-x)$, $\alpha_2=\arcsin x$. Ясно, что $\sin\alpha_1=-x$ и $\sin\alpha_2=x$, поэтому $\sin\alpha_1=-\sin\alpha_2$. Используя нечетность функции $y=\sin x$, получаем, что $\sin\alpha_1=\sin(-\alpha_2)$. Поскольку α_1 и α_2 лежат на сегменте $[-\pi/2;\ \pi/2]$, а на этом сегменте функция $y=\sin x$ принимает каждое значение y лишь для одного x, то отсюда следует, что $\alpha_1=-\alpha_2$, что и требовалось доказать.

ж) Функция возрастает на всей области определения.

Докажем это свойство. Для этого докажем, что если $-1\leqslant x_1<x_2\leqslant 1$, то $\arcsin x_1<\arcsin x_2$. Обозначим $\alpha_1=\arcsin x_1$ и $\alpha_2=\arcsin x_2$. Ясно, что $\sin\alpha_1=x_1$ и $\sin\alpha_2=x_2$, т. е. надо доказать, что если $\sin\alpha_1 < \sin\alpha_2$, то $\alpha_1 < \alpha_2$. Доказательство проведем от противного: пусть $\alpha_1 \geqslant \alpha_2$. Так как на сегменте $[-\pi/2;\ \pi/2]$ функция $y=\sin x$ возрастает, то из условия $\alpha_1 \geqslant \alpha_2$ следует, что $\sin\alpha_1 \geqslant \sin\alpha_2$, что противоречит условию $\sin\alpha_1 < \sin\alpha_2$. Значит, предположение неверно, т. е. функция $y=\arcsin x$ возрастает.

з) Точка (0, 0) — единственная точка пересечения с осями координат.

Построим график (рис. 79) функции $y=\arcsin x$, воспользовавшись общим правилом построения графика обратной функции.

Аналогично доказываются свойства и для других обратных тригонометрических функций.

2. $y=\arccos x$.

а) Функция определена только на отрезке $[-1;\ 1]$, т. е. $-1 \leqslant x \leqslant 1$;

б) по определению множество всех значений функции составляет отрезок $[0;\ \pi]$, т. е. $0 \leqslant y \leqslant \pi$;

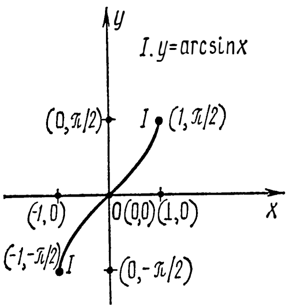


Рис. 79

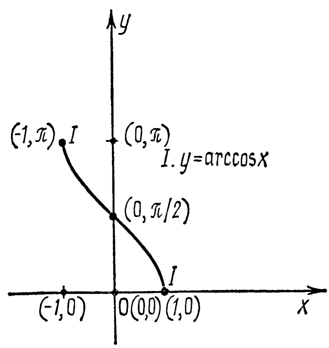


Рис. 80

в) функция ограничена и снизу ($y \geqslant 0$) и сверху ($y \leqslant \pi$);

г) функция принимает наибольшее значение $y=\pi$ при $x=-1$ и наименьшее значение $y=0$ при $x=1$;

д) функция не является периодической;

е) функция не является ни четной, ни нечетной;

ж) функция убывает на всей области определения;

з) точки (0, $\pi/2$) и (1,0) — точки пересечения с осями координат.

График этой функции изображен на рис. 80.

3. $y=\operatorname{arctg} x$.

а) Функция определена для любого действительного x;

б) по определению множество всех значений функции составляет промежуток $(-\pi/2,\ \pi/2)$, т. е. $-\pi/2 < y < \pi/2$;

в) функция ограничена и снизу ($y > -\pi/2$) и сверху ($y < \pi/2$);

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция нечетная;

ж) функция возрастает на всей области определения;

з) точка (0, 0) — единственная точка пересечения с осями координат.

График этой функции изображен на рис. 81.

4. $y = \operatorname{arcctg} x$.

а) Функция определена для любого действительного x;

б) по определению множество всех значений функции составляет промежуток $(0; \pi)$, т. е. $0 < \operatorname{arcctg} x < \pi$;

в) функция ограничена и снизу ($y > 0$) и сверху ($y < \pi$);

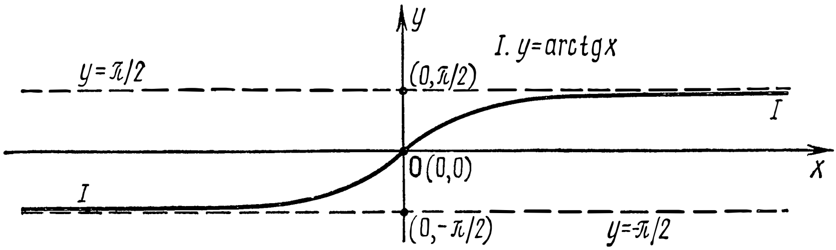


Рис. 81

г) функция не принимает экстремальных значений;

д) функция не является периодической;

е) функция не является ни четной, ни нечетной;

ж) функция убывает на всей области определения;

з) точка $(0, \pi/2)$ — единственная точка пересечения с осями координат.

График этой функции изображен на рис. 82.

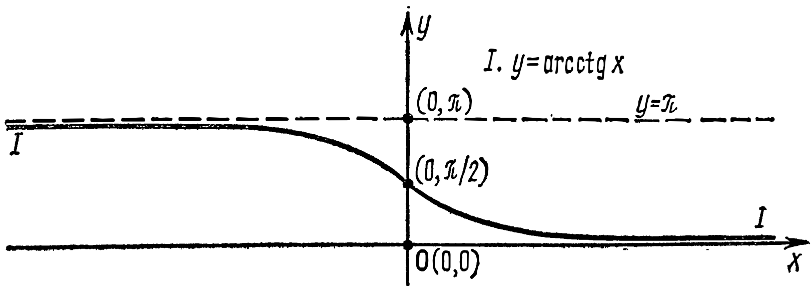


Рис. 82

§ 5
СУПЕРПОЗИЦИИ ФУНКЦИЙ И ИХ ГРАФИКИ

Сложная функция. Пусть функция $u = \varphi(x)$ определена на множестве X и множество значений этой функции входит в область существования функции $y = F(u)$. Тогда любому x из области определения X функции $u = \varphi(x)$ соответствует определенное значение переменной u, а этому значению u функция $y = F(u)$

ставит в соответствие определенное значение переменной y, т. е. переменная y является функцией от x на множестве X: $y = F(\varphi(x))$. Полученная функция от функции называется *сложной функцией* переменной x. Функцию $u=\varphi(x)$ называют внутренней функцией, а $y=F(u)$ — внешней. Сложную функцию $y=F[\varphi(x)]$ называют часто *суперпозицией* двух функций: внутренней $u=\varphi(x)$ и внешней $y=F(u)$. Например, если $u=\cos x$ и $y=2^u$, то для любого действительного x определена сложная функция $y=2^{\cos x}$. Сложными функциями будут, например, функции $y=\sin(2x+4)$, $y=\log_2 \operatorname{tg} x$, $y=\operatorname{tg}\log_2 x$.

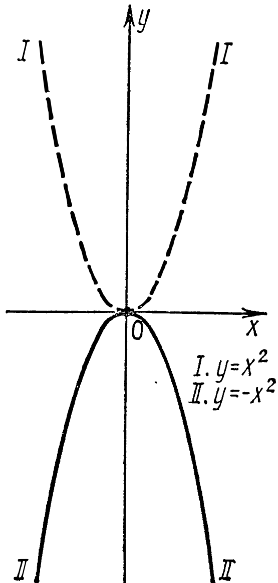


Рис. 83

Функции, полученные из основных элементарных функций с помощью конечного числа алгебраических операций и применения конечного числа суперпозиций, принято называть *элементарными функциями*. Элементарными функциями будут, например, функции $y=\operatorname{tg}\sqrt{1-x^2}$, $y=\log_2\sin 3^{x-4}$, $y=\sqrt{\frac{1}{x^2+4x+1}}$ и т. д.

Рассмотрим примеры, показывающие, как построить график сложной функции $y=F[\varphi(x)]$, зная график внутренней функции $u=\varphi(x)$ и свойства внешней функции $y=F(u)$ (из рассмотренных примеров будет ясно, как построить график любой элементарной функции, зная свойства и графики основных элементарных функций).

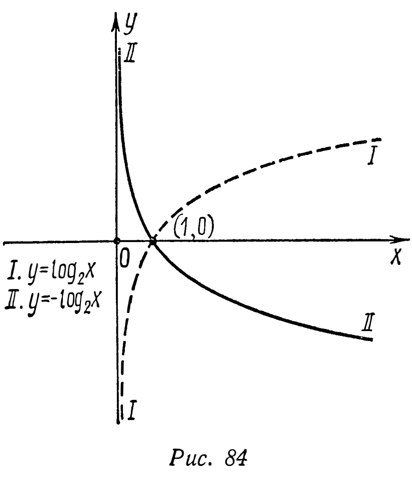


Рис. 84

Построение графика функции $y=-f(x)$ по графику функции $y=f(x)$. Пусть некоторая точка $M_0(x_0, y_0)$ принадлежит графику функции $y=f(x)$, т. е. $y_0=f(x_0)$. Возьмем точку $M_1(x_0, -y_0)$, симметричную точке $M(x_0, y_0)$ относительно оси OX. Координаты точки $M_1(x_0, -y_0)$ удовлетворяют условию $-y_0=-f(x_0)$, поэтому точка $M_1(x_0, -y_0)$ принадлежит графику функции $y=-f(x)$. Так как точка $M_0(x_0, y_0)$, принадлежащая графику функции $y=f(x)$, бралась любая, и функции $y=f(x)$ и $y=-f(x)$ имеют одну и ту же область определения, то график функции $y=-f(x)$ получается из графика функции $y=f(x)$ симметричным отображением последнего относительно оси OX. Построим этим способом графики функций $y=-x^2$ (рис. 83), $y=-1/x$ (рис. 85), $y=-\log_2 x$ (рис. 84).

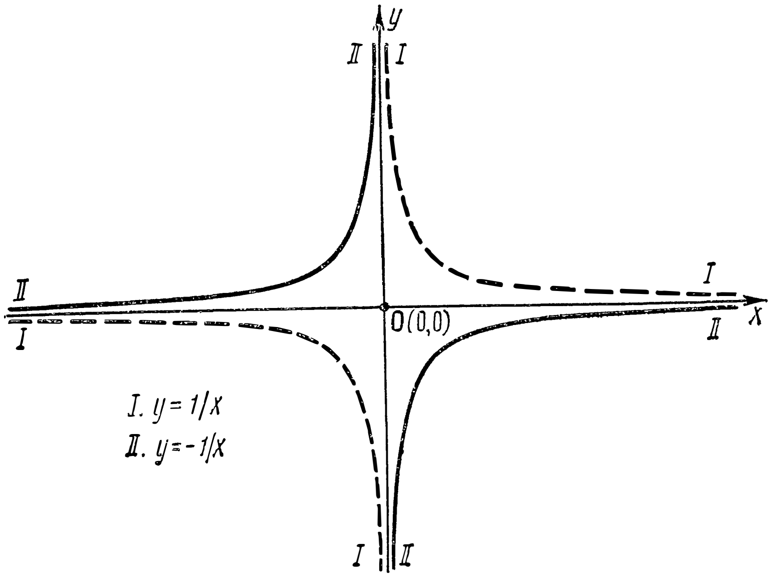


Рис. 85

Построение графика функции $y=f(-x)$ по графику функции $y=f(x)$. Пусть некоторая точка $M_0(x_0, y_0)$ принадлежит графику функции $y=f(x)$, т. е. пусть $y_0=f(x_0)$. Возьмем точку $M_1(-x_0, y_0)$, симметричную точке $M_0(x_0, y_0)$ относительно оси OY. Координаты точки $M_1(-x_0, y_0)$ удовлетворяют условию $y_0=f[-(-x_0)]$, поэтому точка $M_1(-x_0, y_0)$ принадлежит графику функции $y=f(-x)$. Так как точка $M_0(x_0, y_0)$, принадлежащая графику функции $y=f(x)$, бралась любая и функции $y=f(-x)$ и $y=f(x)$ имеют области определения, симметричные относительно начала координат, то график функции $y=f(-x)$ получается из графика функции $y=f(x)$ симметричным отображением последнего относительно оси OY.

Построим этим способом графики функций $y=2^{-x}$ (рис. 86), $y=\log_2(-x)$ (рис. 87).

Построение графика функции $y=Bf(x)$, где $B\neq 0$ по графику функции $y=f(x)$. Функции $y=f(x)$ и $y=Bf(x)$ имеют одну и ту же область определения. Следовательно, зная как для любого x по ординате функции $y=f(x)$ найти ординату функции $y=Bf(x)$, можно по графику функции $y=f(x)$ построить график функции $y=Bf(x)$. Пусть некоторая точка

$M_0(x_0, y_0)$ принадлежит графику функции $y = f(x)$, т. е. пусть $y_0 = f(x_0)$. Возьмем точку $M_1(x_0, By_0)$. Координаты ее удовлетворяют условию $By_0 = Bf(x_0)$, поэтому точка $M_1(x_0, By_0)$ при-

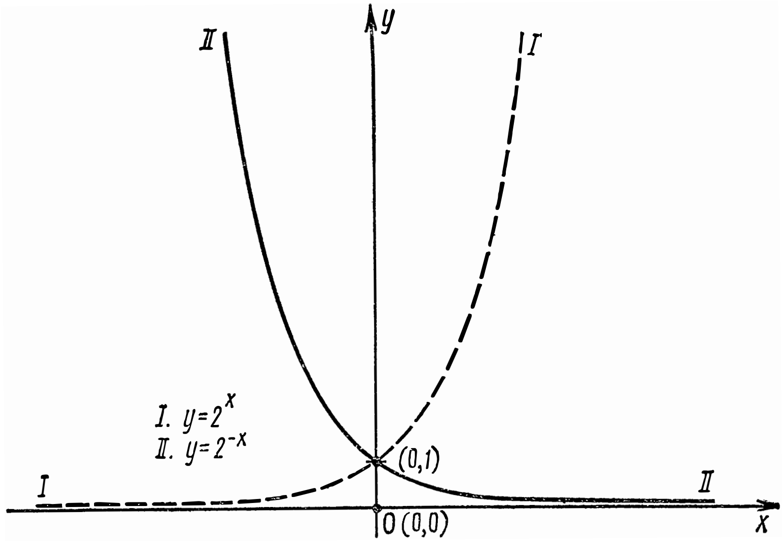


Рис. 86

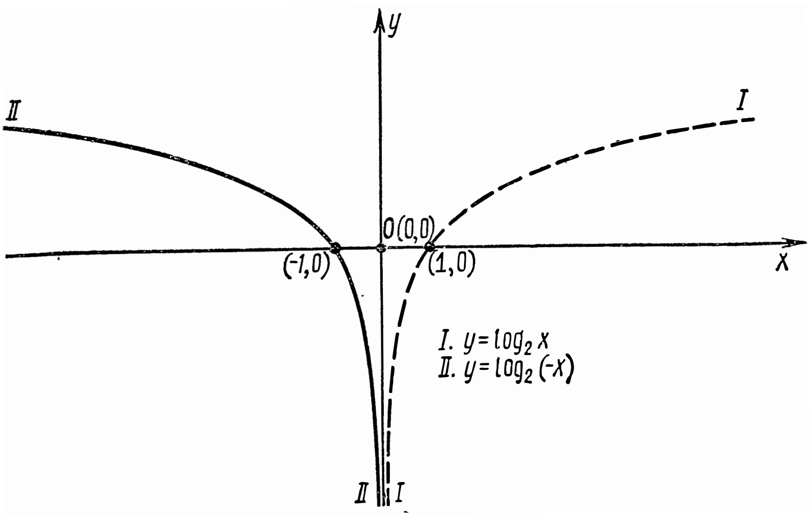


Рис. 87

надлежит графику функции $y = Bf(x)$. Рассмотрим возможные случаи в зависимости от числа B.

1. $B > 1$. Точка $M_1(x_0, By_0)$ получается из точки $M_0(x_0, y_0)$ растяжением ординаты точки M_0 в B раз и график функции

$y=Bf(x)$ получается из графика функции $y=f(x)$ растяжением вдоль оси OY графика функции $y=f(x)$, как упругого тела, в B раз.

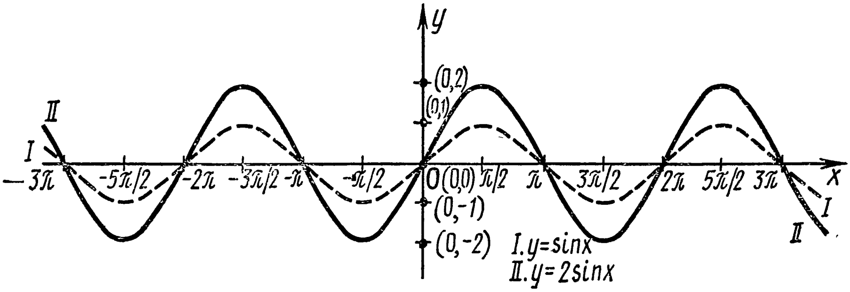


Рис. 88

2. $B=1$. Все точки графика $y=f(x)$ остаются на месте.

3. $0<B<1$. Поскольку $B=\frac{1}{(1/B)}$, то точка $M_1(x_0, By_0)$ получается из точки $M_0(x_0, y_0)$ сжатием ординаты точки M_0 в $1/B$ раз и график функции $y=Bf(x)$ получается из графика функции $y=f(x)$ сжатием ординат всех точек в $1/B$ раз, т. е. сжатием вдоль оси OY графика функции $y=f(x)$ как упругого тела в $1/B$ раз.

4. $B<0$. Тогда $B=-|B|$ и построение графика функции $y=Bf(x)$ разбивается на 2 этапа:

а) построение графика функции $y=|B|f(x)$ по графику функции $y=f(x)$;

б) построение графика функции $y=-[|B|f(x)]$ по графику функции $y=|B|f(x)$. Построим этим способом графики функций $y=2\sin x$ (рис. 88), $y=-2x^2$ (рис. 89), $y=\frac{1}{2}\cos x$(рис. 90).

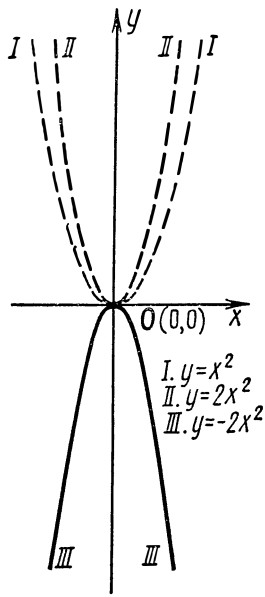


Рис. 89

Построение графика функции $y=f(kx)$, где $k\neq 0$, по графику функции $y=f(x)$. Функция $y=f(kx)$ определена для всех тех x, для которых число kx принадлежит области определения функции $y=f(x)$. Пусть некоторая точка $M_0(x_0, y_0)$ принадлежит графику функции $y=f(x)$, т. е. пусть $y_0=f(x_0)$. Точка $M_1\left(\frac{1}{k}x_0, y_0\right)$ принадлежит графику функции $y=f(kx)$, так как ее координаты удовлетворяют условию $y_0=f\left(k\frac{x_0}{k}\right)$. Рассмотрим различные случаи в зависимости от числа k.

1. $k>1$. Точка $M_1\left(\frac{1}{k}x_0,\ y_0\right)$ получается из точки $M_0(x_0, y_0)$ сжатием абсциссы точки M_0 в k раз и график функции $y=f(kx)$ получается из графика функции $y=f(x)$ сжатием абсцисс всех точек в k раз, т. е. сжатием вдоль оси OX графика функции $y=f(x)$ как упругого тела в k раз.

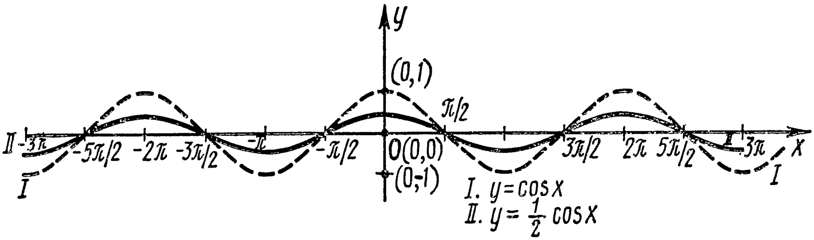


Рис. 90

2. $k=1$. Все точки графика $y=f(x)$ остаются на месте.

3. $0<k<1$. Точка $M_1\left(\frac{1}{k}x_0,\ y_0\right)$ получается из точки $M_0(x_0, y_0)$ растяжением абсциссы точки M_0 в $1/k$ раз и график

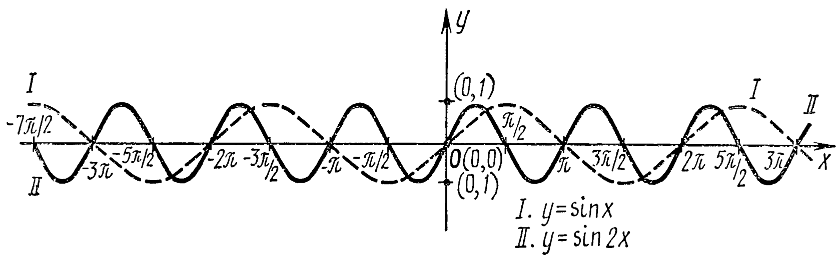


Рис. 91

функции $y=f(kx)$ получается из графика функции $y=f(x)$ растяжением абсцисс всех точек в $1/k$ раз, т. е. растяжением вдоль оси OX графика функции $y=f(x)$ как упругого тела в $1/k$ раз.

4. $k<0$. Тогда $k=-|k|$ и построение графика функции $y=f(kx)$ разбивается на 2 этапа:

а) построение графика функции $y=f(|k|x)$ по графику функции $y=f(x)$;

б) построение графика $y=f(-|k|x)$ по графику $y=f(|k|x)$. Построим этим способом графики функций $y=\sin 2x$ (рис. 91), $y=2^{-2x}$ (рис. 92), $y=\log_2\left(-\frac{1}{3}x\right)$ (рис. 93).

Построение графика функции $y=f(x-\alpha)$, где $\alpha\neq 0$, по графику функции $y=f(x)$. Функция $y=f(x-\alpha)$ определена для всех тех x, для которых $(x-\alpha)$ принадлежит области определения функции $y=f(x)$. Пусть точка $M_0(x_0, y_0)$ принадлежит графику функции $y=f(x)$, т. е. пусть $y_0=f(x_0)$. Точка $M_1(x_0+\alpha, y_0)$ принадлежит графику функции $y=f(x-\alpha)$, так как ее координаты удовлетворяют условию $y_0=f[(x_0-\alpha)+\alpha]=f(x_0)$. Следовательно, каждая точка M_1 графика функции $y=f(x-\alpha)$ получается из соответствующей точки M_0 графика функции $y=f(x)$ сдвигом этой точки вдоль оси OX на величину α. При этом если $\alpha>0$, то сдвиг производится вправо на величину α, а если $\alpha<0$ — сдвиг влево на величину $|\alpha|$. Итак, график функции $y=f(x-\alpha)$ получается из графика функции $y=f(x)$ сдвигом последнего как жесткого тела вдоль оси OX на величину α. Построим этим способом графики функций $y=\log_2(x+3)$ (рис. 94), $y=(x-2)^2$ (рис. 95), $y=\cos(x+\pi/3)$ (рис. 96).

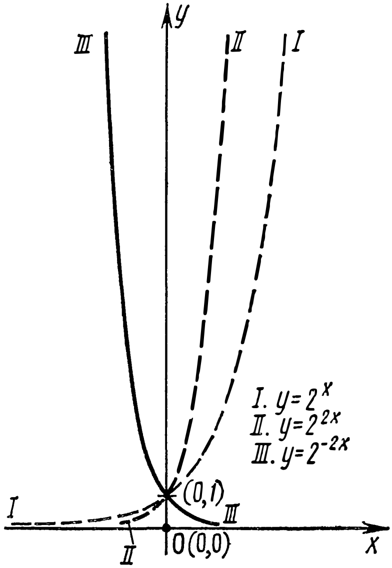


Рис. 92

Построение графика функции $y=f(x)+A$, где $A\neq 0$, по графику функции $y=f(x)$. Функции $y=f(x)$ и $y=f(x)+A$

Рис. 93

имеют одну и ту же область определения. Следовательно, зная как для любого x по ординате функции $y=f(x)$ найти ординату функции $y=f(x)+A$, можно по графику функции $y=f(x)$

построить график функции $y=f(x)+A$. Пусть некоторая точка $M_0(x_0, y_0)$ принадлежит графику функции $y=f(x)$, т. е. пусть $y_0=f(x_0)$. Возьмем точку $M_1(x_0, y_0+A)$. Координаты ее удов-

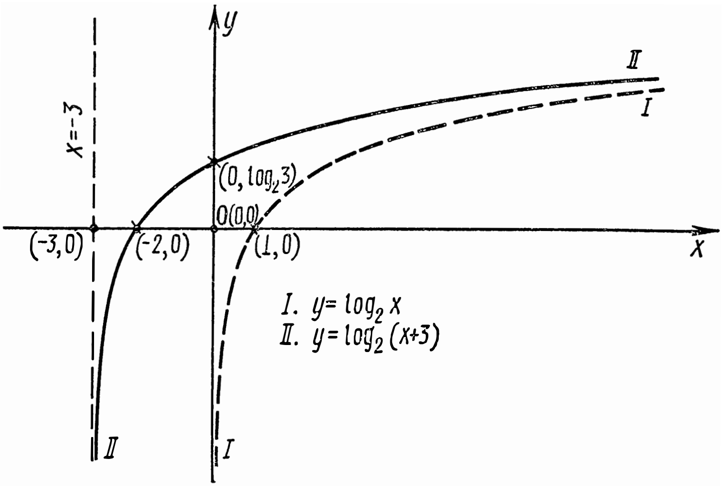


Рис. 94

летворяют условию $y_0+A=f(x_0)+A$. Следовательно, чтобы получить точку M_1, нужно точку M_0 сдвинуть вдоль оси OY на величину A. При этом если $A>0$, то сдвиг производится вверх на величину A, если $A<0$ — вниз на величину $|A|$. Построим этим способом графики функций $y=2^x-3$ (рис. 97), $y=\sin x+1$ (рис. 98), $y=x^2-1$ (рис. 99).

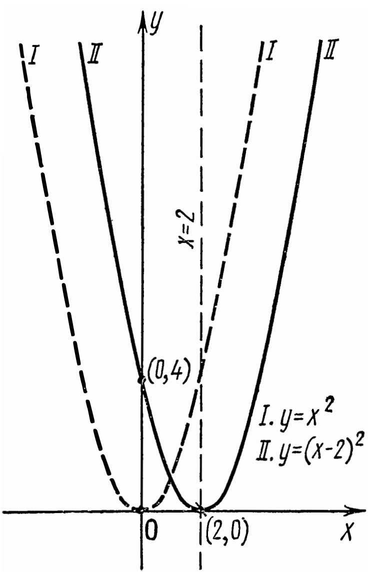


Рис. 95.

Построение графика функции $y=Bf[k(x-\alpha)]+A$ по графику функции $y=f(x)$. График функции $y=Bf[k(x-\alpha)]+A$ строится по графику функции $y=f(x)$ последовательным применением предыдущих способов:

$$y=f(x)\to y=f(kx)\to y=B(kx)\to y=Bf[k(x-\alpha)]\to y=Bf[k(x-\alpha)]+A.$$

Покажем это на примерах. Построить график $y=ax^2+bx+c$, где $a\neq 0$.

Преобразуем квадратный трехчлен ax^2+bx+c, выделив полный квадрат:

$$ax^2+bx+c=a\left(x^2+\frac{b}{a}x\right)+c=a\left(x^2+2x\frac{b}{2a}+\frac{b^2}{4a^2}-\frac{b^2}{4a^2}\right)+c=$$
$$=a\left(x^2+2x\frac{b}{2a}+\frac{b^2}{4a^2}\right)+c-\frac{b^2}{4a}=a\left[x-\left(-\frac{b}{2a}\right)\right]^2+\frac{4ac-b^2}{4a}.$$

Итак, следует построить график функции

$$y=a\left[x-\left(-\frac{b}{2a}\right)\right]^2+\frac{4ac-b^2}{4a}.$$

Сначала построим график функции $y=x^2$. Затем растяжением его вдоль оси OY в $|a|$ раз — график функции $y=|a|x^2$. Если

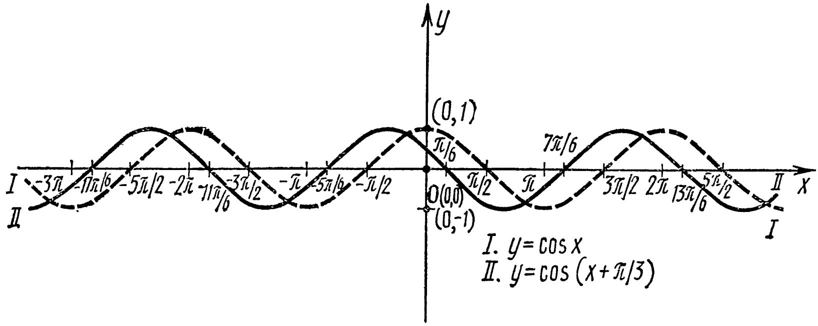


Рис. 96

$a>0$, то возьмем построенный график функции $y=ax^2$; если $a<0$, то отобразим график функции $y=|a|x^2$ относительно оси OX и получим график функции $y=ax^2$. Наконец, сдвигом графика функции $y=ax^2$ вдоль оси OX на $-b/2a$, а затем сдвигом полученного графика функции $y=a(x+b/2a)^2$ вдоль оси OY на $(4ac-b^2)/4a$ получим график функции $y=ax^2+bx+c$. Покажем все эти этапы построения графика квадратного трехчлена на примере $y=-2x^2+3x+1$. Преобразуем квадратный трехчлен $-2x^2+3x+1=-2\,(x-{}^3/_4)^2+{}^{17}/_8$ и построим по изложенной схеме график функции $y=-2\,(x-{}^3/_4)^2+{}^{17}/_8$ (рис. 100).

Теорема 1. График функции $y=kx+b$ есть прямая линия, пересекающая ось OY в точке $M\,(0,\ b)$ и образующая с положительным направлением оси OX угол, тангенс которого равен k.

Доказательство. Докажем, что график функции $y=kx$ есть прямая, проходящая через начало координат и образующая с положительным направлением оси OX угол, тангенс которого равен k.

Рассмотрим несколько случаев.

а) $k=0$. Тогда график функции $y=0$ есть ось OX — прямая линия, и $\operatorname{tg}\alpha=0$;

б) $k>0$. Тогда все точки графика функции $y=kx$ лежат в I или III четвертях. Начало координат принадлежит графику функции $y = kx$. Возьмем некоторую точку M_0 (x_0, y_0), отличную от начала координат, принадлежащую графику функции $y = kx$, т. е. такую, что $y_0 = kx_0$. Проведем через точки $M_0(x_0, y_0)$ и $O(0, 0)$ прямую и покажем, что эта прямая и будет графиком функции $y = kx$.

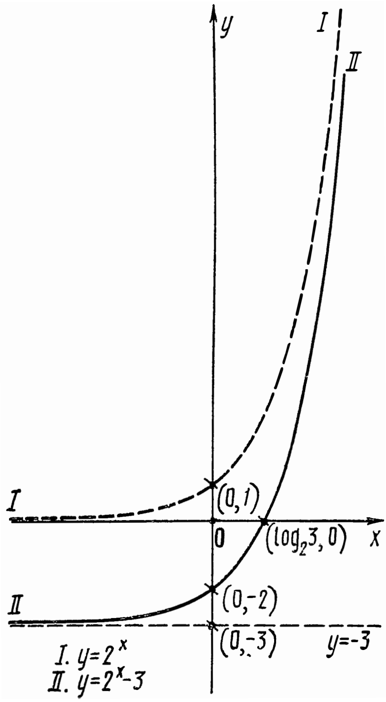


Рис. 97

Пусть построенная прямая образует с положительным направлением оси OX угол α, тогда $\operatorname{tg}\alpha = |y_0|/|x_0| = y_0/x_0 = kx_0/x_0 = k$. Возьмем некоторую точку $M_1(x_1, y_1)$, отличную от точек O и M_0, на этой прямой (пусть для определенности точка M_1 лежит в I четверти). Из прямоугольного треугольника OAM_1 (рис. 101) находим, что $AM_1 = OA \operatorname{tg}\alpha$. Так как $AM_1 = y_1$, $AO = x_1$, $tg\,\alpha = k$, то $y_1 = kx_1$. Это означает, что координаты любой точки построенной прямой удовлетворяют условию $y = kx$.

Пусть теперь x_2 и y_2 таковы, что $y_2 = kx_2$ (пусть для определенности $x_2 > 0$ и, следовательно, $y_2 > 0$). Построим точку M_2 (x_2, y_2). Точка M_2 должна оказаться на построенной прямой, ибо если точка M_2 не попадает на эту прямую, то через начало

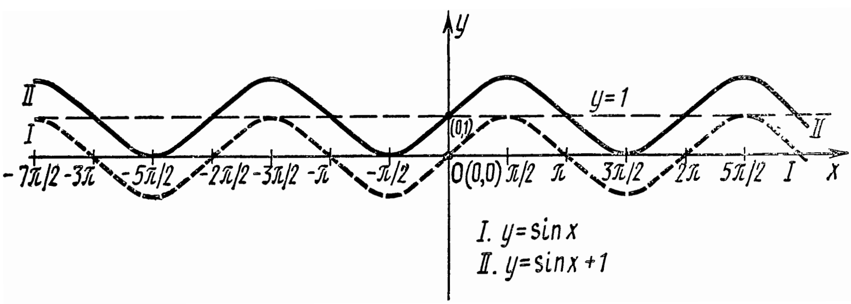


Рис. 98

координат будут проведены 2 прямые, образующие один и тот же угол α с положительным направлением оси OX, что невозможно.

Итак, точки построенной прямой и только они удовлетворяют условию $y = kx$, т. е. график функции $y = kx$ есть построенная прямая.

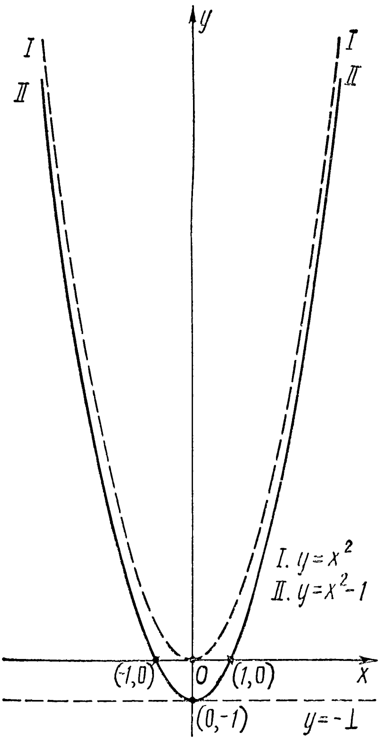


Рис. 99

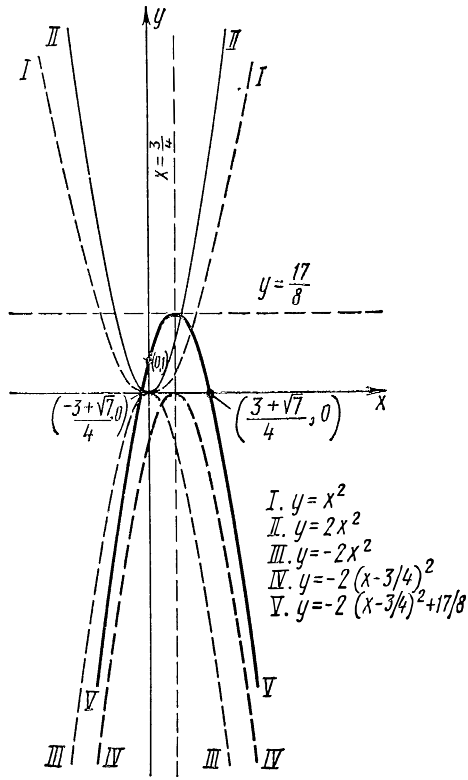


Рис. 100

в) $k < 0$. Тогда график функции $y = |k|x$ есть прямая, проходящая через начало координат и образующая с положительным направлением оси OX угол, тангенс которого равен $|k|$. График функции $y = -|k|x$ получается из этого графика симметрией относительно оси OX, поэтому этот график есть прямая, образующая с положительным направлением оси OX угол, тангенс которого равен $-|k| = k$.

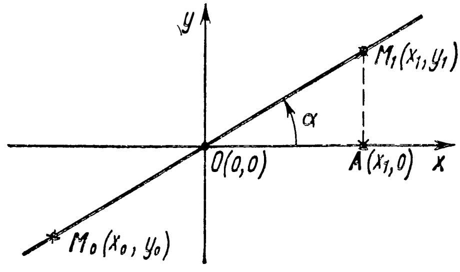


Рис. 101

Для завершения доказательства остается сказать, что график функции $y=kx+b$ получается из прямой $y=kx$ сдвигом вдоль оси OY как жесткого тела на величину b. При этом точка $O(0, 0)$ графика функции $y=kx$ перейдет в точку $A(0, b)$ графика функции $y=kx+b$. Теорема доказана.

Построим этим способом графики функций $y=\frac{3}{4}x+3$ (рис. 102, I), $y=-2x+2$ (рис. 102, II).

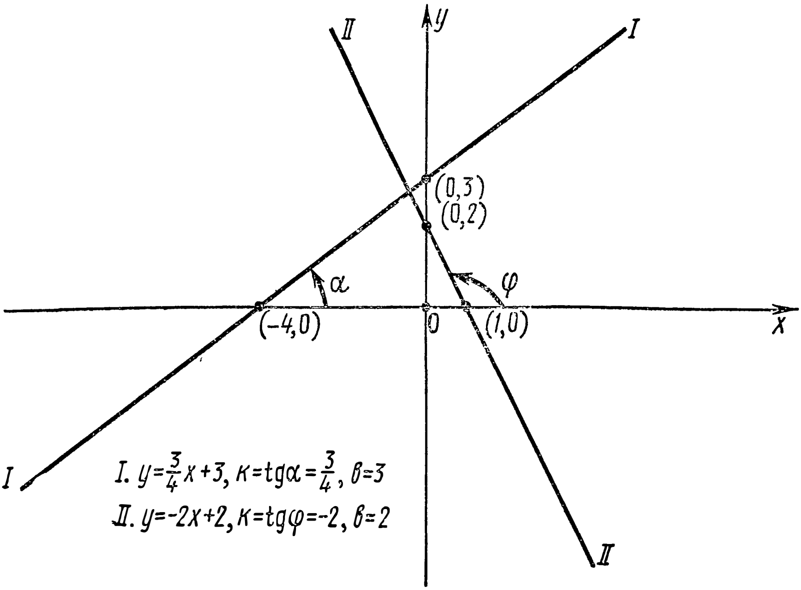


Рис. 102

Построение графика функции $y=|f(x)|$ по графику функции $y=f(x)$. Прежде всего напомним определение:

$$|f(x)| = \begin{cases} f(x) & \text{для тех } x, \text{ где } f(x) \geqslant 0, \\ -f(x) & \text{для тех } x, \text{ где } f(x) < 0. \end{cases}$$

Пусть некоторая точка $M_0(x_0, y_0)$ принадлежит графику функции $y=f(x)$, т. е. пусть $y_0=f(x_0)$. Рассмотрим два случая:

а) $y_0 \geqslant 0$. Тогда, поскольку $|f(x_0)|=f(x_0)=y_0$, точка $M_0(x_0, y_0)$ принадлежит графику функции $y=|f(x)|$;

б) $y_0 < 0$. Тогда, поскольку $|f(x_0)|=-f(x_0)=-y_0$, точка $M_0(x_0, -y_0)$ принадлежит графику функции $y=|f(x)|$. Следовательно, график функции $y=|f(x)|$ получается из графика функции $y=f(x)$ следующим образом:

все точки графика $y=f(x)$, лежащие на оси OX и выше ее остаются на месте;

все точки графика $y=f(x)$, лежащие ниже оси OX, симметрично отображаются относительно оси OX, Заметим, что график функции $y=|f(x)|$ не имеет точек ниже оси OX.

Построим этим способом графики функций $y=|x^2-1|$ (рис. 103), $y=|2^x|$ (рис. 69, II), $y=|\log_2 x|$ (рис. 104), $y=|\sin x|$ (рис. 105).

Построение графика функции $y=f(|x|)$ по графику функции $y=f(x)$. Заметим, что функция $y=f(|x|)$ четная функция, так как $f(|-x|)=f(|x|)$. График четной функции строится так: строится график этой функции **для всех** $x \geqslant 0$; для построения графика этой функции для $x<0$ построенная часть отображается симметрично относительно оси OY. Поскольку $|x|=x$ для $x \geqslant 0$, то для $x \geqslant 0$ график функции $y=f(|x|)$ совпадает с графиком функции $y=f(x)$. Для построения графика функции $y=f(|x|)$ для $x<0$ надо часть графика функции $y=f(|x|)$ для $x \geqslant 0$ симметрично отобразить относительно оси OY. Для построения графика функции $y=f(|x|)$ существенную роль играют точки графика функции $y=f(x)$, лежащие на оси OY или справа от нее; точки графика, лежащие слева от оси OY, никакой роли не играют, следовательно, для построения графика функции $y=f(|x|)$ надо:

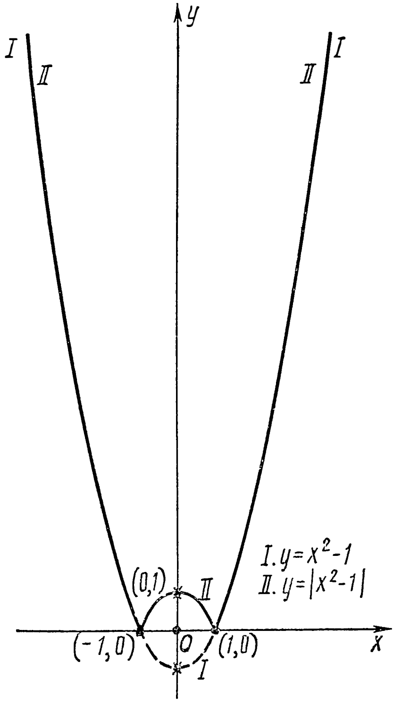


Рис. 103

а) стереть все точки графика функции $y=f(x)$, лежащие слева от оси OY;

б) оставить на месте все точки графика функции $y=f(x)$, лежащие на оси OY и справа от нее;

в) отобразить правую часть графика симметрично относительно оси OY.

Построим этим способом графики функций $y=2^{|x|}$ (рис. 106), $y=\log_2|x|$ (рис. 107), $y=\sin|x|$ (рис. 108).

Построение графика функции $y=F(f(x))$ по графику функции $y=f(x)$. В случаях, более сложных, чем рассмотренные выше случаи, график функции $y=F(f(x))$ строится, используя график $y=f(x)$ и свойства функции $y=F(x)$. Не давая общих рекомендаций, покажем это на нескольких примерах.

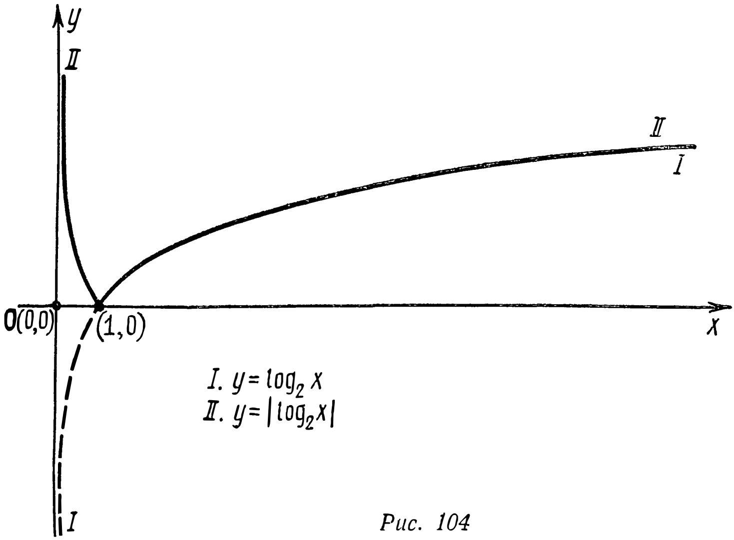


Рис. 104

Используя график функции $y=\sin x$ (см. рис. 73), построить графики:

$$y=2^{\sin x}$$

а) область определения функции $y=2^{\sin x}$ — все действительные x;

б) поскольку функция $y=$ $=\sin x$ периодическая с главным периодом $T=2\pi$, то функция $y=2^{\sin x}$ тоже периодическая с периодом 2π;

$$y=\log_2 \sin x$$

а) область определения функции $y=\log_2 \sin x$ — все те x, для которых $\sin x>0$, т. е. все те x, где график функции $y=\sin x$ выше оси OX;

б) поскольку функция $y=$ $=\sin x$ периодическая с главным периодом $T=2\pi$, то функция $y=\log_2 \sin x$ тоже периодическая с периодом 2π;

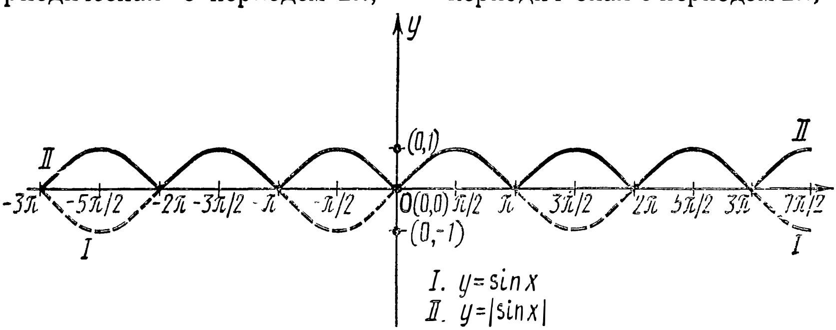


Рис. 105

Поэтому будем строить графики обеих функций только на отрезке $[0;\ 2\pi]$, а потом продолжим графики периодически.

в) поскольку на промежутке $0 \leqslant x \leqslant \pi/2$ функция $y = \sin x$ возрастает от 0 до 1, то функция $y = 2^{\sin x}$ на этом промежутке возрастает от 1 до 2; на промежутке от $\pi/2$ до $3\pi/2$ функция $y = \sin x$ убывает от 1 до -1, а функция $y = 2^{\sin x}$ убывает от 2 до $^1/_2$;

на промежутке от $3\pi/2$ до 2π функция $y = \sin x$ возрастает от -1 до 0, а функция $y = 2^{\sin x}$ возрастает от $^1/_2$ до 1.

в) поскольку на промежутке $0 < x \leqslant \pi/2$ функция $y = \sin x$ возрастает от 0 до 1, то функция $y = \log_2 \sin x$ на этом промежутке возрастает от $-\infty$ до 0;

на промежутке от $\pi/2$ до π функция $y = \sin x$ убывает от 1 до 0, а функция $y = \log_2 \sin x$ убывает от 0 до $-\infty$;

на промежутке $\pi \leqslant x \leqslant 2\pi$ функция $y = \sin x$ неположительна, поэтому на этом промежутке функция $y = \log_2 \sin x$ не определена (и точек графика этой функции здесь нет)

Перечисленные свойства позволяют нам построить требуемые графики на промежутке $[0; 2\pi]$ и продолжить их периодически (рис. 109, 110). Предыдущие рассуждения показывают, как график функции помогает выбрать необходимые промежутки для исследования свойств сложных функций и тем самым помогает построить график сложной функции.

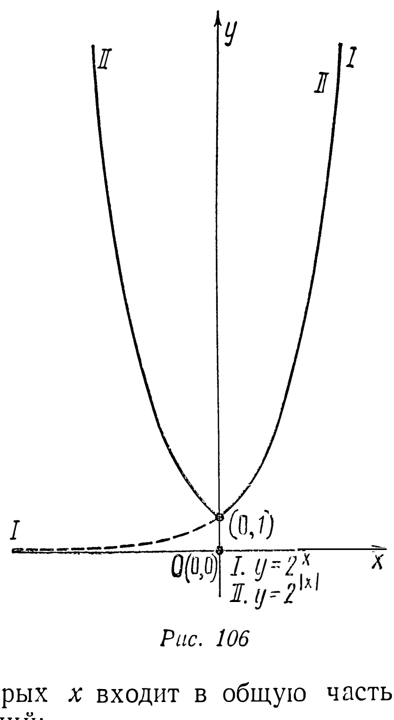


Рис. 106

Сложение графиков. Пусть даны функции $y = f(x)$ и $y = g(x)$. Тогда на общей части их областей определения определена функция $y = f(x) + g(x)$. Пусть точка $M_1(x_0, y_1)$ принадлежит графику функции $y = f(x)$, а точка $M_2(x_0, y_2)$ принадлежит графику функции $y = g(x)$, причем число x_0 принадлежит общей части областей определения функций $y = f(x)$ и $y = g(x)$. Тогда точка $M_3(x_0, y_1 + y_2)$ принадлежит графику функции $y = f(x) + g(x)$. Значит, для построения графика функции $y = f(x) + g(x)$ надо:

а) оставить те точки графиков $y = f(x)$ и $y = g(x)$, у которых x входит в общую часть областей определения этих функций;

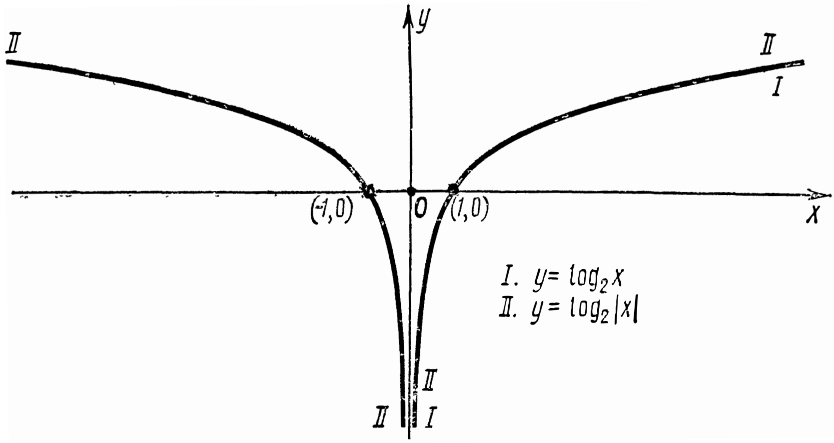


Рис. 107

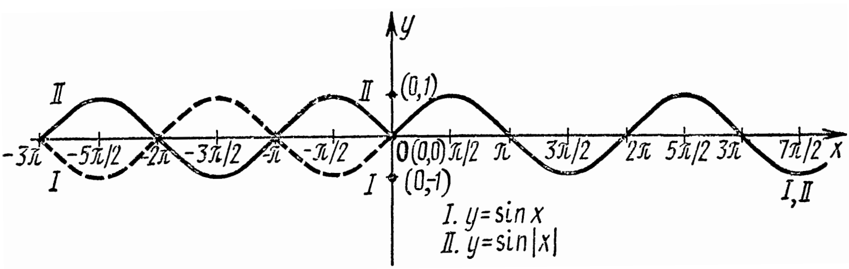


Рис. 108

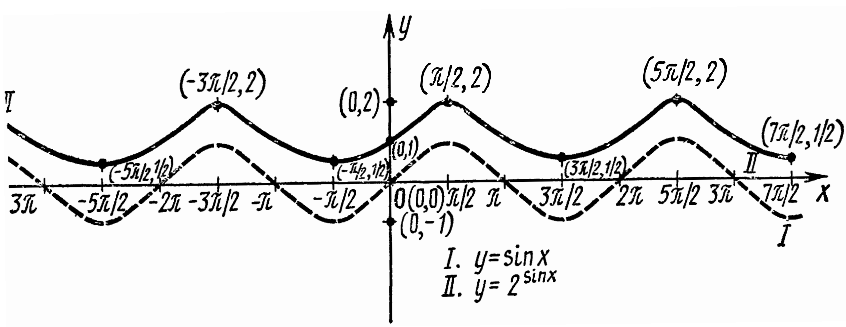


Рис. 109

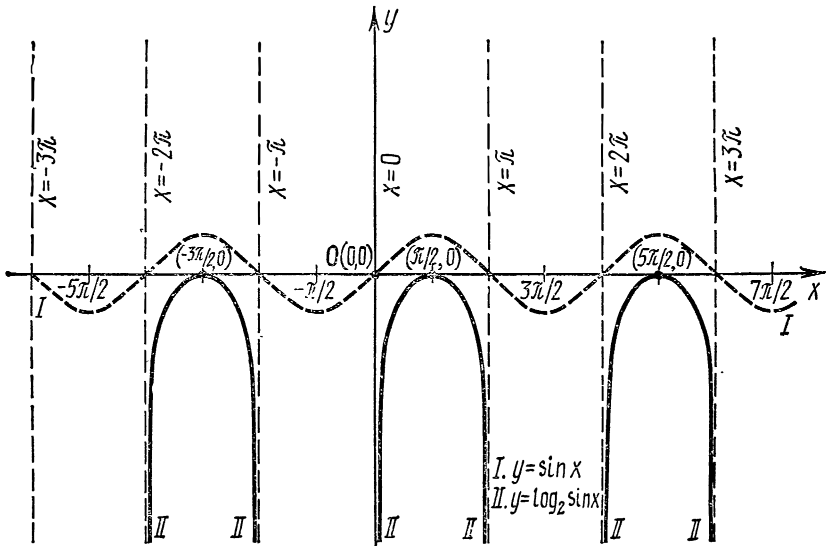


Рис. 110

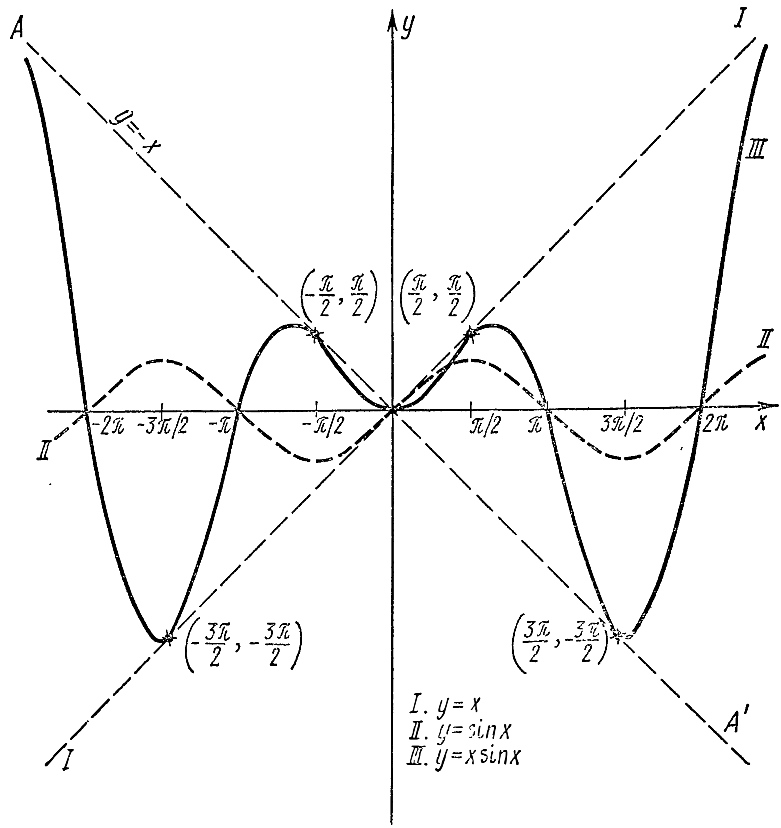


Рис. 111

б) для каждого такого x произвести алгебраическое сложение ординат (соответствующих данному x) этих двух графиков.

Построим этим методом график функции $y=x+\sin x$ (рис. 112).

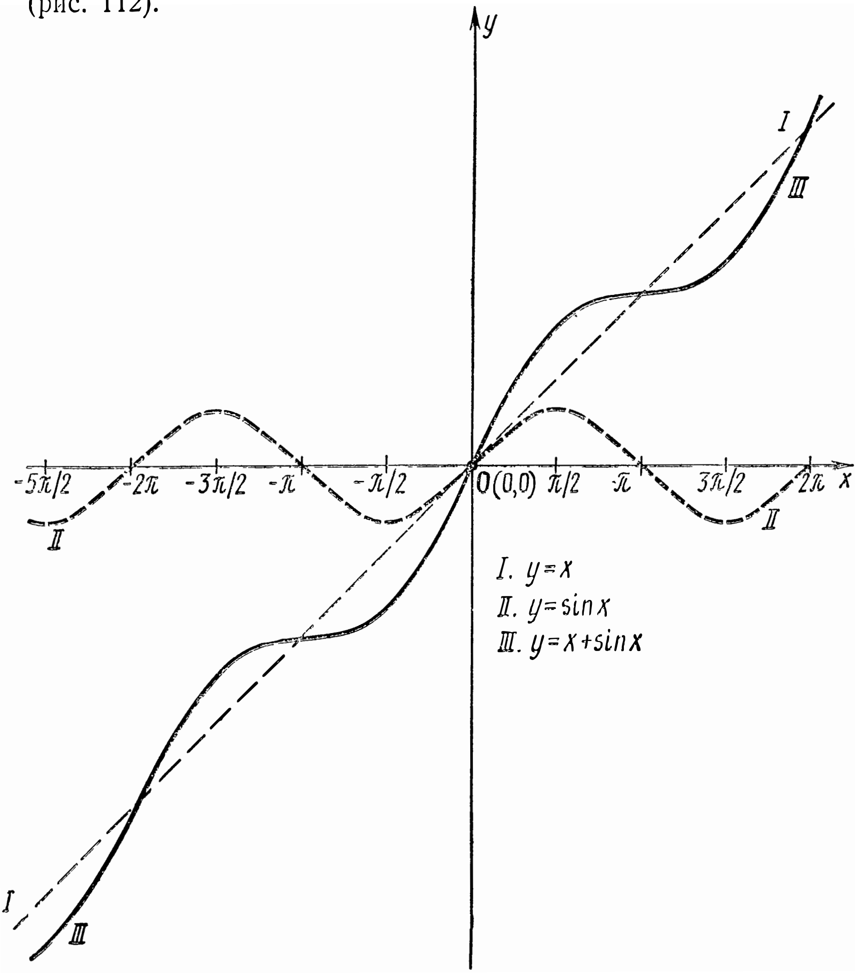


Рис. 112

Умножение графиков. Пусть даны функции $y=f(x)$ и $y=g(x)$. Тогда на общей части их областей определения определена функция $y=f(x)g(x)$. Пусть точка $M_1(x_0, y_1)$ принадлежит графику функции $y=f(x)$, а точка $M_2(x_0, y_2)$ принадлежит графику функции $y=g(x)$, причем число x_0 принадлежит общей части областей определения функции $y=f(x)$ и $y=g(x)$. Тогда точка $M_3(x_0, y_1y_2)$ принадлежит графику функции $y=f(x)g(x)$. Значит, для построения графика функции $y=f(x)g(x)$ надо:

а) оставить те точки графиков $y=f(x)$ и $y=g(x)$, у которых x входит в общую часть областей определения этих функций;

б) для каждого такого x произвести умножение ординат (соответствующих данному x) этих двух графиков.

Построим этим методом график функции $y=x\sin x$ (рис. 111).

Глава VII

УРАВНЕНИЯ И НЕРАВЕНСТВА

Пусть даны две функции: $y=f(x)$ и $y=g(x)$. У каждой из них есть своя область определения и потому есть некоторая область M — пересечение областей определения этих функций (в частности, область M может быть пустым множеством).

Пусть стоит задача: найти все числа α из области M, для каждого из которых справедливо числовое равенство $f(\alpha)=g(\alpha)$. В таких случаях говорят, что стоит задача: *решить уравнение* $f(x)=g(x)$ или что *дано уравнение* $f(x)=g(x)$.

Пусть стоит задача: найти все числа α из области M, для каждого из которых справедливо числовое неравенство $f(\alpha)>g(\alpha)$. В таких случаях говорят, что стоит задача: *решить неравенство* $f(x)>g(x)$ или что *дано неравенство* $f(x)>g(x)$. Аналогично формулируются задачи: решить неравенство $f(x)<g(x)$; решить неравенство $f(x)\geqslant g(x)$; решить неравенство $f(x)\leqslant g(x)$.

В этой главе рассматриваются некоторые способы решения таких уравнений и неравенств. Отметим, что если область M — пустое множество, то задачи: решить уравнение $f(x)=g(x)$ и решить неравенство $f(x)>g(x)$ [$f(x)<g(x)$; $f(x)\geqslant g(x)$; $f(x)\leqslant g(x)$] уже решены. Нет ни одного числа α, для которого справедливо либо числовое равенство $f(\alpha)=g(\alpha)$, либо числовое неравенство $f(\alpha)>g(\alpha)$ [$f(\alpha)<g(\alpha)$; $f(\alpha)\geqslant g(\alpha)$; $f(\alpha)\leqslant g(\alpha)$].

§ 1
ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И УТВЕРЖДЕНИЯ РАВНОСИЛЬНОСТИ УРАВНЕНИЙ

Областью допустимых значений (ОДЗ) уравнения $f(x)=g(x)$ называется общая часть (пересечение) областей существования функций $y=f(x)$ и $y=q(x)$, т. е. множество всех числовых значений неизвестного x, при каждом из которых имеют смысл (определены) левая и правая части уравнения. Всякое число x из ОДЗ уравнения называется *допустимым значением* для данного уравнения.

Число α из ОДЗ уравнения называется *решением* (*или корнем*) *уравнения* $f(x)=g(x)$, если при подстановке его вместо неизвестного x уравнение превращается в верное числовое равенство $f(\alpha)=g(\alpha)$.

Решить уравнение — это значит найти множество всех его корней. Отметим, что это множество может быть пустым в двух случаях: а) если ОДЗ уравнения — пустое множество; б) если ОДЗ уравнения — непустое множество x, но ни для одного числа $\alpha \in X$ не выполняется числовое равенство $f(\alpha)=g(\alpha)$. Если множество всех корней уравнения $f(x)=g(x)$ — пустое множество, то обычно говорят, что уравнение $f(x)=g(x)$ не имеет корней, поэтому иногда говорят так: решить уравнение $f(x)=$ $=g(x)$ — это значит найти все его корни или доказать, что корней нет.

Пусть даны два уравнения: $f(x)=g(x)$ и $p(x)=\varphi(x)$. Если любой корень первого уравнения является корнем второго уравнения, то *второе уравнение называется следствием первого.*

Пусть даны два уравнения: $f(x)=q(x)$ и $p(x)=\varphi(x)$. Если любой корень первого уравнения является корнем второго уравнения, а любой корень второго уравнения является корнем первого уравнения, то такие *два уравнения называются равносильными* (*или эквивалентными*). Другими словами, два уравнения равносильны, если каждое из них является следствием другого. Замена одного уравнения другим уравнением, ему равносильным, называется *равносильным переходом* от одного уравнения к другому.

Пусть даны два уравнения: $f(x)=g(x)$ и $p(x)=\varphi(x)$ — и пусть дано некоторое множество M значений неизвестного x. Если любой корень первого уравнения, принадлежащий множеству M, является корнем второго уравнения, а любой корень второго уравнения, принадлежащий множеству M, является корнем первого уравнения, то такие два уравнения называются *равносильными на множестве* M. Замена одного уравнения другим уравнением, равносильным ему на множестве M, называется *равносильным переходом на множестве* M от одного уравнения к другому.

Приведем несколько примеров, иллюстрирующих введенные понятия. Область допустимых значений уравнения $\sqrt{1-x}=$ $=\log_2(x-1)$ есть пустое множество. Действительно, область существования функции $y=\sqrt{1-x}$ есть множество: $-\infty<x\leqslant 1$, а область существования функции $y=\log_2(x-1)$ — множество $1<x<+\infty$. Общая часть (пересечение) этих областей — пустое множество. В данном примере, после того как найдена ОДЗ уравнения, оно уже решено, ибо установлено, что корней нет.

Область допустимых значений уравнения $\sqrt{x^2-4}=\sqrt{4-x^2}$ есть два числа 2 и -2, ибо область существования функции $y=\sqrt{x^2-4}$ есть множество $|x|\geqslant 2$, а область существования функции $y=\sqrt{4-x^2}$ — множество $|x|\leqslant 2$. Общая часть (пересечение) этих областей — множество $|x|=2$, т. е. $x=2$ и $x=-2$. Подстановкой убеждаемся, что оба эти числа являются корнями уравнения. Значит, и в этом примере, после того как найдена ОДЗ уравнения, оно решено.

Приведенные примеры показывают, что при решении уравнений бывает полезно знать ОДЗ этого уравнения. Однако можно привести примеры уравнений, для решения которых не обязательно знать ОДЗ этого уравнения. Например, уравнение $\sqrt{\log_2(x+\sin x)}=-1$ не имеет корней, так как при любом значении x из ОДЗ уравнения имеем неверное числовое равенство. В то же время вычисление ОДЗ этого уравнения было бы непростой задачей.

Два уравнения: $x+4=0$ и $(x^2+1)(x+4)=0$ — равносильны на множестве всех действительных чисел, ибо каждое из этих уравнений имеет только один корень $x=-4$.

Рассмотрим два уравнения: $\sqrt{x}=1$ и $x^2=1$. Первое уравнение имеет только один корень $x_1=1$, который является корнем второго, поэтому второе уравнение — следствие первого. Но второе уравнение имеет еще корень $x_2=-1$, который не удовлетворяет первому уравнению и даже не входит в его ОДЗ. Таким образом, данные уравнения не являются равносильными на множестве всех действительных чисел, но они равносильны на ОДЗ первого уравнения (т. е. на множестве неотрицательных чисел), ибо на этом множестве каждое из них имеет только один корень $x_1=1$.

Приведем некоторые утверждения равносильности уравнений.

1. Уравнения $f(x)=g(x)$ и $f(x)-g(x)=0$ равносильны.

2. Уравнения $f(x)=g(x)$ и $f(x)+\alpha=g(x)+\alpha$ равносильны для любого действительного числа α.

3. Уравнения $f(x)=q(x)$ и $\alpha f(x)=\alpha g(x)$ равносильны для любого действительного отличного от нуля числа α.

4. Уравнения $a^{f(x)}=a^{g(x)}$ и $f(x)=g(x)$ равносильны для любого фиксированного положительного и не равного единице числа a.

Доказательства справедливости этих утверждений аналогичны, поэтому докажем, например, утверждение 4.

Пусть число x_1 является некоторым корнем уравнения $a^{f(x)}=a^{g(x)}$, т. е. пусть существуют числа $f(x_1)$ и $g(x_1)$, для которых справедливо числовое равенство $a^{f(x_1)}=a^{g(x_1)}$. Поскольку фиксированное число a удовлетворяет условиям: $a>0$ и $a\neq 1$, то из справедливости числового равенства $a^{f(x_1)}=a^{g(x_1)}$ вытекает справедливость числового равенства $f(x_1)=g(x_1)$. Следовательно, число x_1 является корнем уравнения $f(x)=g(x)$. Такое рассуждение можно провести для любого корня уравнения $a^{f(x)}=a^{g(x)}$. Значит, любой корень уравнения $a^{f(x)}=a^{g(x)}$ является корнем уравнения $f(x)=g(x)$.

Покажем теперь обратное. Пусть число x_2 является некоторым решением уравнения $f(x)=g(x)$, т. е. пусть существуют числа $f(x_2)$ и $g(x_2)$, для которых справедливо числовое равенство $f(x_2)=g(x_2)$. Тогда на основании свойства числовых равенств для $a>1$ и $a\neq 1$ справедливо равенство $a^{f(x_2)}=a^{g(x_2)}$. Следовательно, число x_2 является корнем уравнения $a^{f(x)}=a^{g(x)}$. Такое рассуждение можно провести для любого корня уравнения $f(x)=g(x)$. Значит, любой корень уравнения $f(x)=g(x)$ является корнем уравнения $a^{f(x)}=a^{q(x)}$. Утверждение 4 доказано. Приведем некоторые утверждения, когда одно уравнение является следствием другого.

5. Пусть n — натуральное число, тогда уравнение $[f(x)]^n=$ $=[g(x)]^n$ — следствие уравнения $f(x)=g(x)$.

Доказательство. По утверждению 1 уравнение $[f(x)]^n=$ $=[g(x)]^n$ равносильно уравнению $[f(x)^n]-[g(x)]^n=0$, которое на основании формул сокращенного умножения может быть записано в виде

$$[f(x)-g(x)]\{[f(x)]^{n-1}+[f(x)]^{n-2}g(x)+\ldots+[g(x)]^{n-1}\}=0. \quad (1)$$

Пусть число x_0 является некоторым корнем уравнения $f(x)=g(x)$, т. е. пусть существуют числа $f(x_0)$ и $g(x_0)$, для которых справедливо равенство $f(x_0)=g(x_0)$. Но тогда справедливо и числовое равенство

$$[f(x_0)-g(x_0)]\{[f(x_0)]^{n-1}+[f(x_0]^{n-2}g(x_0)+\ldots+[g(x_0]^{n-1}\}=0.$$

Следовательно, число x_0 является корнем уравнения (1), которое равносильно первоначальному уравнению, а потому число x_0 является корнем уравнения $[f(x)]^n=[g(x)]^n$. Такое рассуждение можно провести для любого корня первоначального уравнения. Значит, любой корень уравнения $f(x)=g(x)$ есть корень уравнения $[f(x)]^n=[g(x)]^n$.

6. Уравнение $f(x)=g(x)$ — следствие уравнения $\log_a f(x)=$ $=\log_a g(x)$, где $a>0$ и $a\neq 1$.

Доказательство. Пусть число x_0 — некоторый корень уравнения $\log_a f(x)=\log_a g(x)$, т. е. существуют числа $\log_a f(x_0)$ и $\log_a g(x_0)$, для которых справедливо числовое равенство $\log_a f(x_0)=\log_a g(x_0)$. Из равенства логарифмов двух чисел по одному и тому же основанию следует равенство самих этих чисел (см. § 5 гл. IV), т. е. $f(x_0)=g(x_0)$, следовательно, число x_0 является корнем уравнения $f(x)=g(x)$. Такое рассуждение можно провести для любого корня уравнения $\log_a f(x)=$ $=\log_a g(x)$. Значит, любой корень уравнения $\log_a f(x)=\log_a g(x)$ является корнем уравнения $f(x)=g(x)$. Приведем несколько утверждений о равносильности уравнений на множестве.

7. Пусть n — натуральное число и пусть на некотором множестве M функции $y=f(x)$ и $y=g(x)$ неотрицательны. Тогда на этом множестве уравнения $f(x)=g(x)$ и $[f(x)]^n=[g(x)]^n$ равносильны.

Доказательство. Доказанное выше утверждение 5 говорит, что всякий корень уравнения $f(x)=g(x)$ (в том числе и из множества M) есть корень уравнения $[f(x)]^n=[g(x)]^n$.

Докажем обратное. Пусть число $x_0 \in M$ есть некоторый корень уравнения $[f(x)]^n = [g(x)]^n$, т. е. пусть существуют неотрицательные числа $f(x_0)$ и $g(x_0)$, для которых справедливо числовое равенство

$$[f(x_0)]^n = [g(x_0)]^n. \tag{2}$$

Предположим, что число x_0 таково, что одно из чисел $f(x_0)$ или $g(x_0)$ равно нулю. Тогда из равенства (2) следует, что и другое из этих чисел равно нулю, т. е. в этом случае $f(x_0) = g(x_0)$. Следовательно, в этом случае число x_0 является корнем уравнения $f(x) = g(x)$. Предположим теперь, что число x_0 таково, что одно из чисел $f(x_0)$ или $g(x_0)$ не равно нулю. Тогда из числового равенства (2) следует, что и другое из этих чисел также не равно нулю, следовательно, согласно условию утверждения 7 оба числа $f(x_0)$ и $g(x_0)$ положительны. Числовое равенство (2) равносильно числовому равенству

$$[f(x_0) - g(x_0)]\{[f(x_0)]^{n-1} + [f(x_0)]^{n-2} g(x_0) + \ldots + [g(x_0)]^{n-1}\} = 0. \tag{3}$$

Поскольку любая натуральная степень некоторого положительного числа есть положительное число, произведение и сумма положительных чисел есть положительное число, то число $[f(x_0)]^{n-1} + [f(x_0)]^{n-2} g(x_0) + \ldots + [g(x_0)]^{n-1}$ положительно, следовательно, числовое равенство (3) равносильно числовому равенству $f(x_0) - g(x_0) = 0$ или равенству $f(x_0) = g(x_0)$. Последнее числовое равенство означает, что число x_0 является корнем уравнения $f(x) = g(x)$. Такое рассуждение можно провести для любого корня, принадлежащего множеству M уравнения $[f(x)]^n = [g(x)]^n$. Значит, любой корень, принадлежащий множеству M уравнения $[f(x)]^n = [g(x)]^n$, является корнем уравнения $f(x) = g(x)$.

8. Пусть фиксированное число a таково, что $a > 0$ и $a \neq 1$, и пусть на некотором множестве M функции $y = f(x)$ и $y = g(x)$ положительны. Тогда на множестве M уравнения $f(x) = g(x)$ и $\log_a f(x) = \log_a g(x)$ равносильны.

Д о к а з а т е л ь с т в о. Доказанное выше утверждение 6 говорит о том, что всякий корень уравнения $\log_a f(x) = \log_a g(x)$ (в том числе и из множества M) есть корень уравнения $f(x) = g(x)$.

Докажем обратное. Пусть число $x_0 \in M$ является некоторым корнем уравнения $f(x) = g(x)$, т. е. $f(x_0) > 0$, $g(x_0) > 0$ и $f(x_0) = g(x_0)$. Но тогда равны и логарифмы этих чисел по одному и тому же основанию, т. е. тогда справедливо числовое равенство $\log_a f(x_0) = \log_a g(x_0)$. Следовательно, число x_0 является корнем уравнения $\log_a f(x) = \log_a g(x)$. Такое рассуждение можно провести для любого корня из множества M уравнения $f(x) = g(x)$. Итак, любой корень из множества M уравнения $f(x) = g(x)$ является корнем уравнения $\log_a f(x) = \log_a g(x)$.

9. Пусть на некотором множестве M, принадлежащем ОДЗ уравнения $f(x)=g(x)$, функция $y=\varphi(x)$ определена и отлична от нуля. Тогда на множестве M уравнения $f(x)=g(x)$ и $f(x)\varphi(x)=g(x)\varphi(x)$ равносильны.

Доказательство. Пусть число $x_1 \in M$ является некоторым корнем уравнения $f(x)=g(x)$, т. е. существуют числа $f(x_1)$ и $g(x_1)$, для которых справедливо числовое равенство $f(x_1)-g(x_1)=0$. Так как по условию существует число $\varphi(x_1)$ и $\varphi(x_1)\neq 0$, то справедливо и числовое равенство $\varphi(x_1)f(x_1)-\varphi(x_1)g(x_1)=0$, которое равносильно числовому равенству $\varphi(x_1)f(x_1)=\varphi(x_1)g(x_1)$. Последнее числовое равенство означает, что число x_1 является корнем уравнения $\varphi(x)f(x)=\varphi(x)g(x)$. Такое рассуждение можно провести для любого корня из множества M уравнения $f(x)=g(x)$. Значит, любой корень из множества M уравнения $f(x)=g(x)$ является корнем уравнения $\varphi(x)f(x)=\varphi(x)g(x)$.

Докажем обратное. Пусть число $x_2 \in M$ и является некоторым корнем уравнения $\varphi(x)f(x)=\varphi(x)g(x)$, т. е. пусть существуют числа $f(x_2)$, $g(x_2)$ и $\varphi(x_2)$, для которых справедливо числовое равенство $\varphi(x_2)f(x_2)=\varphi(x_2)g(x_2)$, равносильное числовому равенству $\varphi(x_2)[f(x_2)-g(x_2)]=0$. По условию утверждения 9 число $\varphi(x_2)\neq 0$, следовательно, последнее числовое равенство равносильно числовому равенству $f(x_2)-g(x_2)=0$, которое равносильно равенству $f(x_2)=g(x_2)$. Последнее числовое равенство означает, что число x_2 — корень уравнения $f(x)=g(x)$. Такое рассуждение можно провести для любого корня из множества M уравнения $\varphi(x)f(x)=\varphi(x)g(x)$. Значит, любой корень из множества M уравнения $\varphi(x)f(x)=\varphi(x)g(x)$ является корнем уравнения $f(x)=g(x)$.

Приведем определения совокупности и системы уравнений. Говорят, что дана *совокупность n уравнений*:

$$f_1(x)=g_1(x),\ f_2(x)=g_2(x),\ \ldots,\ f_n(x)=g_n(x), \tag{4}$$

если требуется найти **все** такие числа α, каждое из которых является корнем хотя бы одного уравнения совокупности (4) (уравнения совокупности обычно записываются в строчку).

Таким образом, решить совокупность уравнений (4) — значит сначала решить каждое уравнение $f_i(x)=g_i(x)$, $i=1, 2, \ldots, n$, т. е. найти множества $M_1, M_2, \ldots, M_n$ всех корней каждого из этих уравнений, и затем взять объединение этих множеств $M=M_1\cup M_2\cup\ldots\cup M_n$. Это множество и будет решением совокупности уравнений (4), а каждое число α из множества M будет называться корнем совокупности (4).

Часто возникает необходимость совершить *равносильный переход от уравнения к совокупности уравнений*. Говорят, что уравнение

$$P(x)=\varphi(x) \tag{5}$$

равносильно на множестве M совокупности уравнений

$$f_1(x)=g_1(x),\ f_2(x)=g_2(x),\ \ldots,\ f_n(x)=g_n(x), \qquad (6)$$

если любой корень уравнения (5), принадлежащий множеству M, является корнем совокупности (6) и, наоборот, если любой корень совокупности (6), принадлежащий множеству M, является корнем уравнения (5). Такая замена уравнения (5) совокупностью (6) называется равносильным переходом на множестве M от уравнения (5) к совокупности (6).

Отметим, что если в совокупности (4) уравнений бесконечно много, то говорят, что дана бесконечная совокупность уравнений. Говорят, что дана *система n уравнений*:

$$\begin{cases} f_1(x)=g_1(x), \\ f_2(x)=g_2(x), \\ \ldots\ldots\ldots \\ f_n(x)=g_n(x), \end{cases} \qquad (7)$$

если требуется найти **все** такие числа α, каждое из которых является корнем каждого уравнения системы (7) (уравнения системы обычно записывают в столбик и объединяют фигурной скобкой).

Таким образом, решить систему уравнений (7) — значит сначала решить каждое уравнение $f_i(x)=g_i(x)$, $i=1, 2, \ldots, n$, т. е. найти множества $M_1, M_2, \ldots, M_n$ всех корней каждого из этих уравнений, и затем взять пересечение всех этих множеств $L=M_1 \cap M_2 \cap \ldots \cap M_n$. Это множество, и будет решением системы уравнений (7), а каждое число α из множества L будет называться корнем системы (7).

Иногда возникает необходимость совершить *равносильный переход от уравнения к системе уравнений*.
Говорят, что уравнение

$$P(x)=\varphi(x) \qquad (8)$$

равносильно на множестве E системе уравнений

$$\begin{cases} P_1(x)=g_1(x), \\ P_2(x)=g_2(x), \\ \ldots\ldots\ldots \\ P_n(x)=g_n(x), \end{cases} \qquad (9)$$

если любой корень уравнения (8), принадлежащий множеству E, является корнем системы (9) и, наоборот, если любой корень системы (9), принадлежащей множеству E, является корнем уравнения (8). Такая замена уравнения (8) системой (9) называется равносильным переходом на множестве M от уравнения (8) к системе (9).

§ 2
ПРОСТЕЙШИЕ УРАВНЕНИЯ

Простейшим уравнением называется уравнение вида

$$f(x)=b,$$

где $f(x)$ — основная элементарная функция (см. гл. VI), b — данное действительное число. Очевидно, что ОДЗ простейшего уравнения совпадает с областью существования основной элементарной функции, которая стоит в левой части уравнения.

Рассмотрим решение простейшего уравнения на некоторой области $X \subset$ ОДЗ, причем в качестве области X будем брать следующие промежутки: либо отрезок $[x_1;\ x_2]$, либо интервал $(x_1;\ x_2)$, либо полуинтервалы $(x_1;\ x_2]$; $[x_1;\ x_2)$ либо лучи $[x_1;\ +\infty)$; $(x_1+\infty)$; $(-\infty;\ x_1]$, $(-\infty;\ x_1)$, либо всю числовую прямую $(-\infty;\ +\infty)$. Через Y обозначим ту часть области значений функции $y=f(x)$, которая получается только для $x \in X$. Тогда, если основная элементарная функция $y=f(x)$ строго монотонна на X и $b \in Y$, то уравнение $f(x)=b$ имеет единственное решение на области X; если $b \notin Y$ (т. е. если число b не принадлежит области Y), то уравнение $f(x)=b$ не имеет решения на множестве X, так как $f(x_1) \in Y$ при любом $x_1 \in \in X$ и, следовательно, $f(x_1) \neq b$. Рассмотрим применение этого утверждения для решения простейших уравнений.

Алгебраическое уравнение $x^n=b$ (n — данное натуральное число) (1).

Элементарная функция $y=x^n$ определена на всей числовой прямой, значит, ОДЗ уравнения (1): $-\infty<x<+\infty$.

При $n=2m+1$ ($m=0,\ 1,\ 2,\ldots$) уравнение (1) принимает вид $x^{2m+1}=b$. Функция $y=x^{2m+1}$ является возрастающей функцией и областью значений Y этой функции является вся числовая прямая, т. е. $-\infty<y<+\infty$. Следовательно, уравнение $x^{2m+1}=b$ при любом b имеет единственное решение, которое определяется с помощью операции арифметического корня (см. гл. II), а именно, если $b \geqslant 0$, то решение имеет вид $x_1=\sqrt[2m+1]{b}$, если $b<0$ то $x_1=-\sqrt[2m+1]{|b|}$.

В случае $n=2m$ ($m=1,\ 2,\ldots$) уравнение (1) принимает вид $x^{2m}=b$. Областью значений функции $y=x^{2m}$ является луч $Y=[0;\ +\infty)$. Поэтому при $b<0$ уравнение $x^{2m}=b$ решений не имеет. Так как на луче $X_1=[0,\ +\infty)$ функция $y=x^{2m}$ строго возрастает, то на луче X_1 при каждом $b \geqslant 0$ уравнение $x^{2m}=b$ имеет единственное решение $x_1=\sqrt[2m]{b}$.

Так как на луче $X_2=(-\infty,\ 0]$ функция $y=x^{2m}$ строго убывает, то на луче X_2 при каждом $b \geqslant 0$ уравнение $x^{2m}=b$ также имеет единственное решение x_2, которое в силу четности функции $y=x^{2m}$ таково, что $x_2=-x_1$, т. е. имеет вид $x_2 = -\sqrt[2m]{b}$. Все возникшие случаи решения уравнения (1) можно объединить (табл. 6).

Таблица 6

	$b<0$	$b=0$	$b>0$
$n=2m+1$ $(m=0, 1, 2, \ldots)$	$x_1=-\sqrt[n]{\lvert b\rvert}$	$x_1=0$	$x_1=\sqrt[n]{b}$
$n=2m$ $(m=1, 2, \ldots)$	нет решений	$x_1=0$	$x_1=\sqrt[n]{b}$; $x_2=-\sqrt[n]{b}$

Дробное уравнение $x^{-n}=b$ (n — данное натуральное число) (2).

Элементарная функция $y=x^{-n}$ определена на множестве всех отличных от нуля действительных чисел. Значит, ОДЗ уравнения (2): $0<x<+\infty$, $-\infty<x<0$.

Если $n=2m+1$ $(m=0, 1, 2, \ldots)$, то уравнение (2) принимает вид $x^{-(2m+1)}=b$. Областью значений функции $y=x^{-(2m+1)}$ являются два луча: $Y_1=(-\infty, 0)$, если $x\in X_1=(-\infty, 0)$, и $Y_2=(0, +\infty)$, если $x\in X_2=(0, +\infty)$. Следовательно, при $b=0$ уравнение $x^{-(2m+1)}=b$ не имеет решения. Если $b>0$, то уравнение $x^{-(2m+1)}=b$ на луче X_1 решений не имеет, так как при $x\in X_1$ значения функции $y=x^{-(2m+1)}$ отрицательны. На луче X_2 функция $y=x^{-(2m+1)}$ строго убывает и, значит, при каждом $b>0$ уравнение $x^{-(2m+1)}=b$ имеет единственное решение x_1. Поскольку $x_1\in(0, +\infty)$ и $b>0$, то справедлива следующая цепочка равносильных числовых равенств

$$x_1^{-(2m+1)}=b \Leftrightarrow x_1^{2m+1}=1/b \Leftrightarrow x_1=\sqrt[2m+1]{1/b}.$$

Следовательно, при каждом $b>0$ уравнение $x^{-(2m+1)}=b$ имеет единственное решение $x_1=\sqrt[2m+1]{1/b}$. Рассуждая аналогично, получим, что при каждом $b<0$ уравнение $x^{-(2m+1)}=b$ имеет единственное решение $x_1=-\sqrt[2m+1]{1/\lvert b\rvert}$.

Если $n=2m$ $(m=1, 2\ldots)$, то уравнение (2) принимает вид $x^{-2m}=b$ и областью значений функции $y=x^{-2m}$ является луч $Y=(0; +\infty)$. Поэтому при $b\leqslant 0$ уравнение $x^{-2m}=b$ не имеет решений. При $b>0$ рассмотрим сначала решение на луче $X_2=(0; +\infty)$. Так как на луче X_2 функция $y=x^{-2m}$ убывает, то при каждом $b>0$ на луче X_2 уравнение $x^{-2m}=b$ имеет единственное решение x_1. Поскольку $x_1\in(0; +\infty)$ и $b>0$, то справедлива следующая цепочка равносильных числовых равенств:

$$x_1^{-2m}=b \Leftrightarrow x_1^{2m}=1/b \Leftrightarrow x_1=\sqrt[2m]{1/b}.$$

Следовательно, при каждом $b>0$ уравнение $x^{-2m}=b$ имеет на луче $X_2=(0; +\infty)$ единственное решение $x_1=\sqrt[2m]{1/b}$.

На луче $X_1=(-\infty;\ 0)$ функция $y=x^{-2m}$ возрастает. Поэтому на луче X_1 при каждом $b>0$ уравнение $x^{-2m}=b$ также имеет единственное решение x_2, которое в силу четности функции $y=x^{-2m}$ таково, что $x_2=-x_1$, т. е. $x_2=-\sqrt[2m]{1/b}$. Решения простейшего уравнения (2) можно объединить (табл. 7).

Таблица 7

	$b<0$	$b=0$	$b>0$
$n=2m+1$ $(m=0,\ 1,\ 2,\ \ldots)$	$x_1=-\sqrt[n]{1/\|b\|}$	нет решений	$x_1=\sqrt[n]{1/b}$
$n=2m$ $(m=1,\ 2,\ \ldots)$	нет решений	нет решений	$x_1=\sqrt[n]{1/b}$; $x_2=-\sqrt[n]{1/b}$

Степенные уравнения $x^{\alpha}=b$ (3), **$x^{-\alpha}=b$** (4) (a — данное положительное нецелое число).

Функция $y=x^{\alpha}$ при данном положительном нецелом числе α на всей области допустимых значений строго возрастает от нуля до бесконечности, т. е. областью значений этой функции является луч $Y=[0;\ +\infty)$. Поэтому уравнение (3) при $b<0$ решений не имеет, а при каждом $b\geqslant 0$ имеет единственное решение x_1. Если $b=0$, то очевидно, что $x_1=0$. Если же $b>0$, то $x_1\in(0;\ +\infty)$ и поэтому равносильны числовые равенства $x_1^{\alpha}=b$ и $x_1=b^{1/\alpha}$. Следовательно, при каждом $b>0$ уравнение $x^{\alpha}=b$ имеет решение $x_1=b^{1/\alpha}$.

Функция $y=x^{-\alpha}$ при данном положительном нецелом числе α является строго убывающей на всей ОДЗ, а областью значений этой функции будет луч $Y=(0;\ +\infty)$. Поэтому уравнение (4) при $b\leqslant 0$ решений не имеет, а при каждом $b>0$ имеет единственное решение x_1. Поскольку $x_1\in(0;\ +\infty)$ и $b>0$, то справедлива следующая цепочка равносильных числовых равенств: $x_1^{-\alpha}=b\Leftrightarrow x_1^{\alpha}=1/b\Leftrightarrow x_1=(1/b)^{1/\alpha}$. Следовательно, при каждом $b>0$ уравнение $x^{-\alpha}=b$ имеет решение $x_1=(1/b)^{1/\alpha}$.

Итак, имеем следующую таблицу (табл. 8) решений простейших степенных уравнений (3) и (4), где α — данное положительное нецелое число.

Таблица 8

	$b<0$	$b=0$	$b>0$
$x^{\alpha}=b$	нет решений	$x_1=0$	$x_1=b^{1/a}$
$x^{-\alpha}=b$	нет решений	нет решений	$x_1=\left(\frac{1}{b}\right)^{1/\alpha}$

Показательное уравнение $a^x = b$ (a — данное положительное не равное единице число) (5).

Функция $y = a^x$ ($a > 0$, $a \neq 1$) строго монотонна на всей области допустимых значений. Областью значений этой функции является луч $Y = (0; +\infty)$. Следовательно, если $b \leq 0$, то уравнение (5) не имеет решений, а при каждом $b > 0$ уравнение (5) имеет единственное решение x_1. Поскольку $a^{x_1} > 0$ и $b > 0$, то равносильны числовые равенства $a^{x_1} = b$ и $\log_a a^{x_1} = \log_a b$. Согласно основному логарифмическому тождеству $\log_a a^{x_1} = x_1$. Следовательно, при каждом $b > 0$ уравнение $a^x = b$ имеет единственное решение $x_1 = \log_a b$.

Логарифмическое уравнение $\log_a x = b$ (a — данное положительное не равное единице число) (6).

Функция $y = \log_a x$ ($a > 0$, $a \neq 1$) является строго монотонной функцией на всей области допустимых значений. Областью значений этой функции является вся числовая прямая $Y = (-\infty; +\infty)$. Следовательно, при каждом b уравнение (6) имеет единственное решение x_1. Поскольку $x_1 \in (0; +\infty)$, то равносильны числовые равенства $\log_a x_1 = b$ и $a^{\log_a x_1} = a^b$. Согласно основному логарифмическому тождеству $a^{\log_a x_1} = x_1$. Следовательно, при каждом действительном b уравнение $\log_a x = b$ имеет единственное решение $x_1 = a^b$.

Тригонометрические уравнения. Сделаем несколько общих замечаний. Пусть надо решить простейшее уравнение $f(x) = b$, где $f(x)$ — основная элементарная тригонометрическая функция. Будем говорить, что простейшее уравнение $f(x) = b$ имеет главный период T, если функция $y = f(x)$ имеет главный период T. Поэтому очевидно, что если для некоторого простейшего тригонометрического уравнения с главным периодом T найдено некоторое решение x_0, то любое число $x_k = x_0 + kT$ при любом целом k также является решением этого уравнения. При этом множество всех решений вида $x_k = x_0 + kT$, где k пробегает все целые числа, т. е. $k = 0, \pm 1, \pm 2 \pm \ldots$, называется серией решений этого уравнения. Для нахождения множества всех решений данного простейшего тригонометрического уравнения с главным периодом T надо найти **все** его решения на промежутке длиной в T, затем для каждого найденного решения выписать соответствующую серию решений; объединение всех таких серий и будет множеством всех решений данного уравнения. Промежуток длиной в главный период T следует выбирать таким, чтобы он содержал промежуток, на котором для функции $y = f(x)$ определена обратная тригонометрическая функция (см. § 4 гл. VI), и таким, чтобы все решения уравнения на этом промежутке, по возможности, просто отыскивались.

Тригонометрическое уравнение $\cos x = b$ (7).

Функция $y = \cos x$ определена на всей числовой оси и имеет главный период 2π. Областью значений этой функции является отрезок $Y=[-1;\ 1]$. Следовательно, если $|b|>1$, то уравнение $\cos x=b$ не имеет решений.

Рассмотрим теперь решение уравнения $\cos x=b$ при $|b|\leqslant 1$ на промежутке $(-\pi;\ \pi]$ длиной в главный период. Сначала найдем решение уравнения (7) на отрезке $[0;\ \pi]$. На этом отрезке функция $y=\cos x$ строго убывает, областью ее изменения для $x\in[0;\ \pi]$ является отрезок $Y=[-1;\ 1]$, поэтому уравнение $\cos x=b$ для каждого b, такого, что $|b|\leqslant 1$, имеет единственное решение x_1. Тогда справедливо числовое равенство

$$\cos x_1=b. \tag{11}$$

В силу определения функции, обратной к функции $y=\cos x$, числовое равенство (11) равносильно числовому равенству $\arccos(\cos x_1)=\arccos b$. По определению $\arccos(\cos x_1)=x_1$ для любого числа $x_1\in[0;\ \pi]$. Следовательно, уравнение (7) на отрезке имеет единственное решение $x_1=\arccos b$. Теперь найдем решение уравнения (7) на промежутке $(-\pi;\ 0)$. На этом промежутке функция $y=\cos x$ строго возрастает; областью ее изменения для $x\in(-\pi;\ 0)$ является промежуток $Y_1=(-1;\ 1)$. Следовательно, уравнение $\cos x=b$ для $b=1$ и $b=-1$ не имеет решений на рассматриваемом промежутке $(-\pi,\ 0)$, а для каждого b, такого, что $|b|<1$, имеет единственное решение x_2. Учитывая, что функция $y=\cos x$ четная функция, получаем, что $x_2=-x_1$.

Итак, уравнение $\cos x=b$ на промежутке $(-\pi;\ \pi]$ имеет:

для $b=1$ единственное решение $x_1=\arccos 1$, т. е. $x_1=0$;

для $b=-1$ единственное решение $x_1=\arccos(-1)$, т. е. $x_1=\pi$;

для каждого b, такого, что $|b|<1$, имеет два решения: $x_1=\arccos b$ и $x_2=-\arccos b$.

Каждое из этих решений дает серию решений уравнения (7). Объединение этих серий и будет множеством всех решений уравнения (7). Значит, множеством всех решений уравнения (7) будет:

для $b=1$ серия $x=2\pi k$, где $k=0,\ \pm 1,\ \pm 2,\ \pm\ldots$;

для $b=-1$ серия $x=\pi+2\pi m$, где $m=0,\ \pm 1,\ \pm 2,\ \pm\ldots$;

для каждого b, такого, что $|b|<1$, две серии: $x=\arccos b+2\pi p$, где $p=0,\ \pm 1,\ \pm 2,\ \pm\ldots$; $x=-\arccos b+2\pi q$, где $q=0,\ \pm 1,\ \pm 2,\ \pm\ldots$. Заметим, что иногда две серии решений уравнения (7) записываются с помощью одной формулы: $x=\pm\arccos b+2\pi n$, где $n=0,\ \pm 1,\ \pm 2,\ \pm\ldots$.

Тригонометрическое уравнение $\sin x=b$ (8). Функция $y=\sin x$ определена на всей числовой прямой и имеет главный период 2π. Областью значений этой функции является отрезок $Y=[-1;\ 1]$. Следовательно, если $|b|>1$, то уравнение $\sin x=b$ не имеет решений.

Рассмотрим теперь решение уравнения $\sin x = b$ при $|b| \leqslant 1$ на промежутке $[-\pi/2; 3\pi/2)$ длиной в главный период. Сначала найдем решение уравнения (8) на отрезке $[-\pi/2; \pi/2]$. На этом отрезке функция $y = \sin x$ строго возрастает, а областью ее изменения для $x \in [-\pi/2; \pi/2]$ является отрезок $Y = [-1; 1]$, поэтому уравнение $\sin x = b$ для каждого b, такого, что $|b| \leqslant 1$, имеет единственное решение x_1. Но тогда справедливо числовое равенство

$$\sin x_1 = b. \qquad (12)$$

В силу определения функции, обратной к функции $y = \sin x$, числовое равенство (12) равносильно числовому равенству $\arcsin(\sin x_1) = \arcsin b$. Но по определению $\arcsin(\sin x_1) = x_1$ для любого числа $x_1 \in [-\pi/2; \pi/2]$. Следовательно, уравнение (8) на отрезке $[-\pi/2; \pi/2]$ имеет единственное решение $x_1 = \arcsin b$. Теперь найдем решение уравнения (8) на интервале $(\pi/2; 3\pi/2)$. На этом интервале функция $y = \sin x$ строго убывает, областью ее изменения для $x \in (\pi/2; 3\pi/2)$ является интервал $Y_1 = (-1; 1)$. Следовательно, уравнение $\sin x = b$ для $b = 1$ и $b = -1$ не имеет решений на рассматриваемом интервале $(\pi/2; 3\pi/2)$, а для каждого b, такого, что $|b| < 1$, имеет единственное решение x_2.

Учитывая, что функция $y = \sin x$ симметрична относительно вертикальной прямой, проходящей через точку $(\pi/2, 0)$, получаем, что $x_2 = \pi - x_1$, т. е. $x_2 = \pi - \arcsin b$.

Итак, уравнение $\sin x = b$ на промежутке $[-\pi/2; 3\pi/2)$ имеет:

для $b = 1$ единственное решение $x_1 = \arcsin 1$, т. е. $x_1 = \pi/2$;

для $b = -1$ единственное решение $x_1 = \arcsin(-1)$, т. е. $x_1 = -\pi/2$;

для каждого b, такого, что $|b| < 1$, имеет два решения: $x_1 = \arcsin b$ и $x_2 = \pi - \arcsin b$. Каждое из этих решений дает серию решений уравнения (8), объединение этих серий и будет множеством всех решений уравнения (8). Значит, множеством всех решений уравнения (8) будет:

для $b = 1$ серия $x = \pi/2 + 2\pi k$, где $k = 0, \pm 1, \pm 2, \pm \ldots$;

для $b = -1$ серия $x = -\pi/2 + 2\pi m$, где $m = 0, \pm 1, \pm 2, \pm \ldots$;

для каждого b, такого, что $|b| < 1$, две серии: $x = \arcsin b + 2\pi p$, где $p = 0, \pm 1, \pm 2, \pm \ldots$; $x = \pi - \arcsin b + 2\pi q$, где $q = 0, \pm 1, \pm 2, \pm \ldots$. Заметим, что иногда две серии решений уравнения (8) записываются с помощью одной формулы: $x = (-1)^n \arcsin b + \pi n$, где $n = 0, \pm 1, \pm 2, \pm \ldots$.

Тригонометрическое уравнение $\operatorname{tg} x = b$ (9). Функция $y = \operatorname{tg} x$ определена на всей числовой прямой, кроме точек $x = \pi/2 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm \ldots$, и имеет главный период π. Областью значений этой функции является вся числовая прямая $Y = (-\infty; +\infty)$.

Рассмотрим решение уравнения $\operatorname{tg} x = b$ на интервале $(-\pi/2, \pi/2)$. На этом интервале функция $y = \operatorname{tg} x$ строго возрастает, областью ее изменения для $x \in (-\pi/2, \pi/2)$ является вся числовая прямая $Y = (-\infty; +\infty)$, поэтому для каждого действительного b уравнение $\operatorname{tg} x = b$ имеет единственное решение x_1. Но тогда справедливо числовое равенство

$$\operatorname{tg} x_1 = b. \tag{13}$$

Согласно определению функции, обратной к функции $y = \operatorname{tg} x$, числовое равенство (13) равносильно числовому равенству

$$\operatorname{arctg} (\operatorname{tg} x_1) = \operatorname{arctg} b.$$

Но по определению $\operatorname{arctg} (\operatorname{tg} x_1) = x_1$ для любого числа $x_1 \in (-\pi/2; \pi/2)$. Следовательно, уравнение (9) на интервале $(-\pi/2; \pi/2)$ имеет единственное решение $x_1 = \operatorname{arctg} b$.
Это решение дает серию решений, которая и будет множеством всех решений уравнения (9). Значит множеством всех решений уравнения (9) является серия $x = \operatorname{arctg} b + \pi n$, где $n = 0, \pm 1, \pm 2, \pm \ldots$.

Тригонометрическое уравнение $\operatorname{ctg} x = b$ (10). Функция $y = \operatorname{ctg} x$ определена на всей числовой оси, кроме точек $x = \pi k$, где $k = 0, \pm 1, \pm 2, \pm \ldots$, и имеет главный период π. Областью значений этой функции является вся числовая прямая $Y = (-\infty; +\infty)$.

Рассмотрим решение уравнения $\operatorname{ctg} x = b$ на интервале $(0; \pi)$. На этом интервале функция $y = \operatorname{ctg} x$ строго убывает, областью ее изменения для $x \in (0; \pi)$ является вся числовая прямая $Y = (-\infty; +\infty)$, поэтому для каждого действительного b уравнение $\operatorname{ctg} x = b$ имеет единственное решение x_1. Но тогда справедливо числовое равенство

$$\operatorname{ctg} x_1 = b. \tag{14}$$

Согласно определению функции, обратной к функции $y = \operatorname{ctg} x$, числовое равенство (14) равносильно числовому равенству

$$\operatorname{arcctg} (\operatorname{ctg} x_1) = \operatorname{arcctg} b.$$

Но по определению $\operatorname{arcctg} (\operatorname{ctg} x_1) = x_1$ для любого числа $x_1 \in (0; \pi)$. Следовательно, уравнение (10) на интервале $(0; \pi)$ имеет единственное решение $x_1 = \operatorname{arcctg} b$. Это решение дает серию решений, которая и будет множеством всех решений уравнения (10). Значит множеством всех решений уравнения (10) является серия $x = \operatorname{arcctg} b + \pi n$, где $n = 0, \pm 1, \pm 2, \pm \ldots$.

§ 3
РАВНОСИЛЬНЫЕ ПРЕОБРАЗОВАНИЯ УРАВНЕНИЙ

В этом и следующем параграфах рассматриваются уравнения, которые некоторыми преобразованиями могут быть сведены к одному или совокупности нескольких простейших уравнений.

При решении таких уравнений (не являющихся простейшими) обычно приходится проводить довольно много преобразований, т. е. переходов от одного уравнения к другому. При этом каждый раз уравнение заменяется на какое-то новое, а у нового уравнения, естественно, могут быть и другие корни. Задача правильного решения уравнения заключена в том, чтобы при этих преобразованиях уравнение заменить новым, имеющим все корни предыдущего и только их, т. е. чтобы не произошло потери или приобретения корней. Если каждый раз заменять уравнение на равносильное, корни последнего уравнения и будут корнями исходного.

В этом параграфе рассматриваются только равносильные преобразования уравнений. Неравносильные преобразования будут рассмотрены в следующем параграфе.

Как уже отмечалось в § 1, утверждения 1—3 дают примеры равносильных преобразований. Например, с их помощью в главе III решено линейное уравнение $ax+b=0$ $(a\neq 0)$. Приведем еще несколько примеров равносильных преобразований.

Пусть дано уравнение

$$f(x)=\varphi(x) \tag{1}$$

и пусть для любого действительного x справедливо тождественное равенство $\varphi(x)=g(x)$, тогда уравнение (1) равносильно уравнению

$$f(x)=g(x). \tag{2}$$

Это равносильное преобразование дает возможность применять для решения уравнений различные формулы, справедливые при всех действительных значениях входящих в него букв. Примеры таких преобразований дают формулы сокращенного умножения многочленов, основное тригонометрическое тождество, формулы для синусов и косинусов сумм и разностей углов и некоторые другие формулы. Отметим, что с помощью такого равносильного преобразования, используя формулы сокращенного умножения многочленов, в главе III решены квадратные и некоторые другие алгебраические уравнения. Приведем еще примеры применения этого преобразования. Пусть надо решить уравнение

$$\cos^3 2x=1-2\cos^2 x. \tag{3}$$

Используя основное тригонометрическое тождество и формулу косинуса двойного угла, можно написать тождественное равенство $1+\cos 2x=2\cos^2 x$ или $1-2\cos^2 x=-\cos 2x$, справедливое для любого действительного x. Значит, уравнение (3) равносильно уравнению $\cos^3 2x=-\cos 2x$. Применяя утверждение 1 § 1 и группируя, получим, что уравнение (3) равносильно уравнению $\cos 2x(\cos^2 2x+1)=0$. Последнее уравнение равносильно совокупности двух простейших тригонометрических уравнений: $\cos 2x=0$, $\cos^2 2x+1=0$. Первое уравнение совокупности имеет корни $x=\pi/4+\pi k/2$, где $k=0, \pm 1, \pm 2, \pm\ldots$, а второе уравнение корней не имеет. Значит, уравнение (3) имеет корни $x=\pi/4+k\pi/2$, где $k=0, \pm 1, \pm 2, \pm\ldots$.

Рассмотрим теперь уравнение

$$a\sin x+b\cos x=c. \tag{4}$$

В случае, когда либо $a=0$, либо $b=0$, это уравнение при помощи утверждений 2 и 3 § 1 сводится к простейшему уравнению $\cos x=c/b$ ($b\neq 0$; $a=0$) либо $\sin x=c/a$ ($a\neq 0$; $b=0$).

Пусть теперь $a\neq 0$ и $b\neq 0$. Значит, $a^2+b^2\neq 0$. Применяя утверждение 3, получаем, что уравнение (4) равносильно уравнению

$$\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{b}{\sqrt{a^2+b^2}}\cos x=\frac{c}{\sqrt{a^2+b^2}}. \tag{5}$$

Рассмотрим случай $a>0$. Построим прямоугольный треугольник с катетами длиной a и $|b|$. Угол, лежащий против катета длиной $|b|$, обозначим φ. Тогда имеем числовые равенства $\sin\varphi=|b|/\sqrt{a^2+b^2}$, $\cos\varphi=a/\sqrt{a^2+b^2}$ и уравнение (5) примет вид $\cos\varphi\sin x+\operatorname{sign} b\sin\varphi\cos x=c/\sqrt{a^2+b^2}$, так как $b=|b|\operatorname{sign} b$. По формулам синуса суммы или синуса разности имеем тождественные равенства

$$\cos\varphi\sin x+\operatorname{sign} b\sin\varphi\cos x=$$
$$=\begin{cases}\cos\varphi\sin x+\sin\varphi\cos x=\sin(x+\varphi), \text{ если } b>0;\\ \cos\varphi\sin x-\sin\varphi\cos x=\sin(x-\varphi), \text{ если } b<0,\end{cases}$$

или, используя определение функции $y=\operatorname{sign} t$ (см. гл. VI), имеем $\cos\varphi\sin x+\operatorname{sign} b\sin\varphi\cos x=\sin(x+\varphi\cdot\operatorname{sign} b)$. Так как $\varphi=\operatorname{arctg}\frac{|b|}{a}$, то $\varphi\cdot\operatorname{sign} b=\operatorname{sign} b\cdot\operatorname{arctg}\frac{|b|}{a}=\operatorname{arctg}\frac{b}{a}$. Действительно, если $b>0$, $\operatorname{sign} b=1$, $|b|=b$, и равенство очевидно; если $b<0$, то $\operatorname{sign} b=-1$, $|b|=-b$ и, воспользовавшись нечетностью функции $y=\operatorname{arctg} t$ (см. гл. VI), имеем

$$\operatorname{sign} b\operatorname{arctg}\frac{|b|}{a}=(-1)\operatorname{arctg}\frac{-b}{a}=\operatorname{arctg}\frac{b}{a}.$$

Итак,

$$\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{b}{\sqrt{a^2+b^2}}\cos x=\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{|b|\operatorname{sign} b}{\sqrt{a^2+b^2}}\cos x=$$

$$=\cos\varphi\sin x+\operatorname{sign} b\sin\varphi\cos x=\sin(x+\varphi\operatorname{sign} b)=\sin\left(x+\operatorname{arctg}\frac{b}{a}\right),$$

т. е. уравнение (5) при $a>0$ равносильно уравнению

$$\sin\left(x+\operatorname{arctg}\frac{b}{a}\right)=\frac{c}{\sqrt{a^2+b^2}},$$

которое является простейшим уравнением.

Замечания. 1. Часто это простейшее уравнение записывают в виде:

а) при $b>0,\ a>0\quad \sin(x+\varphi)=\frac{c}{\sqrt{a^2+b^2}}$, где $\varphi=$ $=\arcsin\frac{b}{\sqrt{a^2+b^2}}=\arccos\frac{a}{\sqrt{a^2+b^2}}=\operatorname{arctg}\frac{b}{a}$;

б) при $b<0, a>0$ $\sin(x-\varphi),=\frac{c}{\sqrt{a^2+b^2}}$, где $\varphi=\arcsin\frac{-b}{\sqrt{a^2+b^2}}=$ $=\arccos\frac{a}{\sqrt{a^2+b^2}}=\operatorname{arctg}\left(\frac{-b}{a}\right)$.

2. Случай $a<0$ сводится к рассмотренному умножением левой и правой частей уравнения (4) на -1.

Пример. Решить уравнение

$$\cos x+\sqrt{3}\sin x+1=0. \tag{6}$$

Поскольку уравнение (6) таково, что $a=\sqrt{3}$, $b=1$, $\sqrt{a^2+b^2}=2$, то оно равносильно уравнению

$$\frac{\sqrt{3}}{2}\sin x+\frac{1}{2}\cos x=-\frac{1}{2}.$$

Найдем угол φ: $\varphi=\operatorname{arctg}{}^1/_{\sqrt{3}}={}^{\pi}/_6$. Значит, уравнение (6) равносильно уравнению $\sin(x+{}^{\pi}/_6)=-{}^1/_2$. Решая это простейшее уравнение, получаем две серии решений: $x_1=-{}^{\pi}/_3+$ $+2\pi k$, где $k=0,\ \pm1,\ \pm2,\ \pm\ldots$; $x_2=\pi+2\pi m$, где $m=0$, $\pm1,\ \pm2,\ \pm\ldots$.

Довольно часто встречаются уравнения

$$f(x)=0, \tag{7}$$

где $f(x)=P[g(x)]$ — сложная функция, являющаяся суперпозицией двух функций: основной элементарной $y=g(x)$ и квадратного трехчлена $P(g)=ag^2+bg+c$. Для решения такого уравнения решают сначала квадратное уравнение

$$at^2+bt+c=0. \tag{8}$$

В случае, если дискриминант $D=b^2-4ac$ этого уравнения неотрицателен, то уравнение (8) имеет два корня t_1 и t_2, и в этом случае уравнение (7) равносильно совокупности уравнений

$$g(x)=t_1,\quad g(x)=t_2. \tag{9}$$

Заметим, что если $D=0$, то $t_1=t_2$. Для получения решения уравнения (7) остается решить совокупность уравнений (9). В случае, если дискриминант уравнения (8) отрицателен, то уравнение (8), а значит, и уравнение (7) не имеют действительных корней.

Пример. Решить уравнение $4^x-32^x+2=0$.

Решив квадратное уравнение $t^2-3t+2=0$, получим его корни $t_1=2$, $t_2=1$. Следовательно, исходное уравнение равносильно совокупности уравнений $2^x=2$, $2^x=1$, корни которой и будут корнями исходного уравнения, а именно: $x_1=1$, $x_2=0$.

Не всегда данное уравнение удается сразу преобразовать к виду (7). Часто для этого надо проделать дополнительные равносильные преобразования. Например, для решения уравнения

$$\sin 2x-12(\sin x-\cos x)+12=0 \tag{10}$$

в качестве функции $g(x)$ можно взять функцию $(\sin x-\cos x)$. Чтобы представить уравнение (10) в виде (7) воспользуемся тригонометрическими формулами $\sin 2x=2\cos x\sin x$ и $1=\cos^2 x+\sin^2 x$, т. е. равенствами, справедливыми для любого действительного x. Тогда уравнение (10) равносильно уравнению

$$-(\sin x-\cos x)^2-12(\sin x-\cos x)+13=0. \tag{11}$$

Для решения уравнения (11) решим квадратное уравнение $-t^2-12t+13=0$. Оно имеет два корня: $t_1=1$, $t_2=-13$. Следовательно, уравнение (11) равносильно совокупности уравнений

$$\sin x-\cos x=1,\quad \sin x-\cos x=-13. \tag{12}$$

Решим каждое из этих уравнений способом, изложенным выше, для чего перепишем совокупность уравнений (12) в равносильном виде

$$\sin(x-\pi/4)=\sqrt{2}/2,\quad \sin(x-\pi/4)=-13/\sqrt{2}. \tag{13}$$

Первое уравнение в совокупности (13) имеет две серии решений: $x_1=\pi+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$; $x_2=\pi/2+2\pi n$, где $n=0, \pm 1, \pm 2, \pm \ldots$. Второе уравнение корней не имеет, так как $-13/\sqrt{2}<-1$. Следовательно, уравнение (10) имеет две серии решений: $x_1=\pi+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$; $x_2=\pi/2+2\pi n$, где $n=0, \pm 1, \pm 2, \pm \ldots$.

Замечание. Следует отметить, что этот способ называется способом «замены неизвестного» и решение выглядит следующим образом. В уравнении (11) делаем замену неизвестного

$t=\sin x-\cos x$, получаем квадратное уравнение $-t^2-12t+$ $+13=0$, которое имеет корни $t_1=1$ и $t_2=-13$. Значит, уравнение (11) равносильно совокупности уравнений $\sin x-$ $-\cos x=1$, $\sin x-\cos x=-13$. Решив эти уравнения, получаем ответ: уравнение (10) имеет две серии решений: $x_1=\pi+$ $+2\pi k$, где $k=0, \pm 1, \pm 2, \pm \ldots$; $x_2=\pi/2+2\pi n$, где $n=$ $=0, \pm 1, \pm 2, \pm \ldots$.

Рассмотрим уравнение

$$f(x)=0, \tag{14}$$

где $f(x)$ есть сложная функция, являющаяся суперпозицией нескольких основных элементарных функций. Для решения такого уравнения применяют равносильный переход, рассмотренный выше.

Пусть, например, $f(x)$ есть суперпозиция трех элементарных функций $f(x)=g\{\varphi[u(x)]\}$. Тогда для решения уравнения (14) сначала ищут решения простейшего уравнения $g(t)=0$. Это уравнение может иметь как конечное число корней, так и бесконечное. Пусть это будут числа $t_1, t_2, \ldots$ $\ldots, t_n, \ldots$. Значит, получена совокупность уравнений $\varphi(v)=t_1$, $\varphi(v)=t_2, \ldots, \varphi(v)=t_n, \ldots$, равносильная искомому уравнению. Теперь к каждому из уравнений этой совокупности применяют тот же прием: ищут решение каждого из простейших уравнений $\varphi(v)=t_i$. Пусть это будут числа $v_{i1}, v_{i2}, \ldots, v_{in}, \ldots$; затем записывают совокупность уравнений $u(x)=v_{1k}$, где $k=1, 2, \ldots, u(x)=v_{2k}$, где $k=1, 2, \ldots, u(x)=v_{3k}$, где $k=1, 2, \ldots, u(x)=v_{nk}$, где $k=1, 2, \ldots, \ldots$. Решения этой совокупности уравнений и будут решениями исходного уравнения.

Рассмотрим, например, уравнение

$$2^{\sin\sqrt{x}}=1. \tag{15}$$

Решая простейшее уравнение $2^t=1$, получаем его решение $t_1=0$. Значит, уравнение (15) равносильно уравнению

$$\sin\sqrt{x}=0. \tag{16}$$

Простейшее уравнение $\sin v=0$ имеет корни $v=\pi k$, где $k=$ $=0, \pm 1, \pm 2, \pm \ldots$. Значит, уравнение (15) равносильно бесконечной совокупности уравнений

$$\sqrt{x}=k\pi, \quad \text{где} \quad k=0, \pm 1, \pm 2, \pm \ldots. \tag{17}$$

Ясно, что для любого $k\geqslant 0$ каждое такое простейшее уравнение имеет корень $x=(k\pi)^2$, а для $k<0$ ни одно из этих уравнений не имеет корней. Следовательно, решениями исходного уравнения (15) будут все $x=(k\pi)^2$, где $k=0, 1, 2, 3, \ldots$.

Рассмотрим теперь решение уравнений, в которых некоторые участвующие в уравнении функции стоят под знаком абсолютной величины. Для решения таких уравнений обычно необходимо избавиться от знака абсолютной величины. Наиболее часто применяется для решения таких уравнений так называемый *метод интервалов*.

Суть этого метода состоит в следующем. На числовой оси отмечаются все те значения x, при которых хотя бы одна из функций, стоящих под знаком абсолютной величины, обращается в нуль.

Таким образом, вся числовая ось разбивается на некоторое число промежутков, затем на каждом таком промежутке уравнение заменяется на другое уравнение, не содержащее знаков абсолютной величины и равносильное исходному уравнению на этом промежутке. На каждом таком промежутке ищутся корни соответствующего уравнения, отбираются из них те, которые попадают в данный промежуток. Они и будут корнями исходного уравнения на рассматриваемом промежутке. Наконец, для того чтобы выписать все корни исходного уравнения, собирают вместе (объединяют) все его корни, найденные на всех промежутках.

Продемонстрируем этот способ на нескольких примерах.

1. Решить уравнение

$$|x-3|=x^2-x-7. \tag{18}$$

В этом уравнении лишь одна функция $y=x-3$ находится под знаком абсолютной величины. Эта функция обращается в нуль в точке 3, поэтому вся числовая ось разбивается на два промежутка: $(-\infty;\ 3)$ и $[3;\ +\infty)$. Рассмотрим решение уравнения (18) на каждом из промежутков.

Пусть $x\in(-\infty;\ 3)$, в этом случае для любого числового значения x функция $y=(x-3)$ принимает отрицательные значения, поэтому по определению абсолютной величины $|x-3|=-(x-3)$ и уравнение (18) примет вид $-(x-3)=x^2-x-7$ или $x^2-10=0$. Решая это квадратное уравнение, находим его корни $x_1=-\sqrt{10}$ и $x_2=\sqrt{10}$. Из этих корней в рассматриваемый промежуток входит лишь один корень $x_1=-\sqrt{10}$, который и будет корнем уравнения (18) на рассматриваемом промежутке.

Пусть $x\in[3;\ +\infty)$, в этом случае для любого числового значения x функция $y=x-3$ принимает неотрицательные значения, поэтому по определению абсолютной величины $|x-3|=x-3$, и уравнение (18) примет вид $x-3=x^2-x-7$ или $x^2-2x-4=0$. Решая это квадратное уравнение, находим его корни $x_3=1-\sqrt{5}$ и $x_4=\sqrt{5}+1$. Из этих корней в рассматриваемый промежуток входит лишь один корень $x_4=1+\sqrt{5}$, который и будет корнем уравнения (18) на рассматриваемом промежутке. Наконец, множество всех корней уравнения (18) будет объединением решений на рассмотренных промежутках, т. е. уравнение (18) имеет два корня: $x_1 = -\sqrt{10}$ и $x_2 = 1 + \sqrt{5}$.

2. Решить уравнение

$$|x^2 - 1| - |x| + |2x + 3| - 4x + 6 = 0. \qquad (19)$$

Числовые значения, обращающие в нуль функции, стоящие под знаком абсолютной величины, следующие: $-3/2$, -1, 0, 1. Значит, вся числовая ось разбивается на пять промежутков: $(-\infty; -3/2)$, $[-3/2; -1]$, $(-1; 0)$, $[0; 1)$, $[1; +\infty)$. Рассмотрим решение уравнения (19) на каждом из промежутков.

Пусть $x \in (-\infty; -3/2)$, тогда уравнение (19) принимает вид $(x^2 - 1) - (-x) + (-2x - 3) - 4x + 6 = 0$ или $x^2 - 5x + 2 = 0$. Решая это уравнение, находим его корни $x_1 = (5 - \sqrt{17})/2$ и $x_2 = (5 + \sqrt{17})/2$. Ни один из этих корней не входит в рассматриваемый промежуток, поэтому уравнение (19) не имеет корней на этом промежутке.

Пусть $x \in [-3/2; -1]$, тогда уравнение (19) принимает вид $(x^2 - 1) - (-x) + (2x + 3) - 4x + 6 = 0$ или $x^2 - 3x + 8 = 0$. Это квадратное уравнение не имеет действительных корней, поэтому уравнение (19) не имеет корней на этом промежутке.

Пусть $x \in (-1; 0)$, тогда уравнение (19) принимает вид $-(x^2 - 1) - (-x) + (2x + 3) - 4x + 6 = 0$ или $x^2 + x - 10 = 0$. Решая это уравнение, находим его корни $x_3 = (-1 - \sqrt{41})/2$ и $x_4 = (-1 + \sqrt{41})/2$. Ни один из этих корней не входит в рассматриваемый промежуток, поэтому уравнение (18) не имеет корней на этом промежутке.

Пусть $x \in [0; 1)$, тогда уравнение (19) примет вид $-(x^2 - 1) - (-x) + (2x + 3) - 4x + 6 = 0$ или $x^2 + 3x - 10 = 0$. Корни последнего уравнения $x_5 = -5$, $x_6 = 2$. Ни один из этих корней не входит в рассматриваемый промежуток, поэтому уравнение (19) не имеет корней на этом промежутке.

Пусть $x \in [1; +\infty)$, тогда уравнение (19) принимает вид $(x^2 - 1) - (x) + (2x + 3) - 4x + 6 = 0$ или $x^2 - 3x + 8 = 0$. Это квадратное уравнение не имеет действительных корней, поэтому уравнение (19) не имеет корней на этом промежутке.

Итак, уравнение (19) не имеет корней на всей действительной оси.

§ 4
НЕРАВНОСИЛЬНЫЕ ПРЕОБРАЗОВАНИЯ УРАВНЕНИЙ

В предыдущем параграфе рассматривались лишь равносильные преобразования уравнений. Однако не любое уравнение удается решить при помощи лишь равносильных преобразований, гораздо чаще при решении уравнений приходится применять неравносильные преобразования. При этом надо помнить, что из-за неравносильности преобразований, вообще говоря, можно потерять некоторые корни исходного уравнения или приобрести так называемые «посторонние» корни (всякий корень последующего уравнения, не являющийся корнем исходного уравнения, будем называть посторонним корнем). Поэтому надо знать, при каких преобразованиях это может произойти.

Прежде всего отметим, что это может произойти при взятии от обеих частей уравнения некоторых функций (возведение в степень, извлечение корня, логарифмирование, потенцирование и т. д.), а также при применении некоторых формул.

Что касается взятия функций от обеих частей уравнения, то неравносильность в основном возникает за счет изменения ОДЗ уравнения. Естественно, неравносильность может возникнуть и по другим причинам, например, за счет перехода к уравнению, которое, на самом деле, есть совокупность двух уравнений: исходного и «постороннего».

Преобразования же, связанные с применением формул, приводят к неравносильности исходного и последующих уравнений потому, что для большинства формул левая и правая их части имеют смысл при разных значениях входящих в них букв. Таковы, например, формулы $(\sqrt{x})^2 = x$; $a^{\log_a x} = x$; $\log_a xy = \log_a x + \log_a y$; $\log_a x^n = n \log_a x$; $\operatorname{ctg} x = 1/\operatorname{tg} x$, $\sin 2x = 2 \operatorname{tg} x/(1+\operatorname{tg}^2 x)$ и т. д. Поэтому замена одной части формулы на другую приводит к расширению или сужению ОДЗ уравнения. Ясно, что из-за расширения ОДЗ исходного уравнения возможно приобретение посторонних корней, а из-за сужения ОДЗ — потеря его корней.

При решении уравнений нельзя применять преобразования, сужающие ОДЗ. Ибо, применяя такие преобразования, можно потерять корни исходного уравнения, т. е. найти не все его корни, а значит согласно определению исходное уравнение не будет решено. Что же касается преобразований, расширяющих ОДЗ, то их можно употреблять, но тогда обязательно надо проверить, нет ли среди найденных корней последующего уравнения таких, которые являются посторонними для исходного уравнения. После проверки посторонние корни исходного уравнения, естественно, надо отбросить.

Заметим еще, что при применении неравносильных преобразований можно еще пользоваться равносильностью уравнений на множестве. Тогда, если каждый раз эта равносильность на всей ОДЗ исходного уравнения или части ОДЗ оговаривалась, то найденные корни будут корнями и исходного уравнения и поэтому проверка не нужна.

Ниже приводятся примеры неравносильных преобразований, приводящие как к потере корней исходного уравнения, так и к приобретению посторонних корней.

Преобразования, связанные с логарифмическими формулами. Пусть число a таково, что $a>0$ и $a\neq 1$. Рассмотрим следующие логарифмические формулы:

$$f(x)=a^{\log_a f(x)}, \tag{1}$$

$$\log_a [f(x)]^2 = 2\log_a f(x), \tag{2}$$

$$\log_a [f(x)\, g(x)] = \log_a f(x) + \log_a \varphi(x). \tag{3}$$

Функции, записанные в левых частях этих формул, имеют области существования более широкие, чем области существования функций, стоящих в правых частях.

Действительно, пусть M — область существования функции $y=f(x)$, N — область существования функции $y=g(x)$ и $L = M\cap N$ (напомним, что знаком $\cap$ обозначается пересечение множеств), тогда левая часть формулы (1) имеет смысл для любого $x\in M$, а правая — лишь для тех $x\in M$, для каждого из которых функция $y=f(x)$ положительна. Левая часть формулы (2) имеет смысл для любого $x\in M$, кроме тех, для каждого из которых $f(x)=0$, а правая — лишь для тех $x\in M$, для каждого из которых функция $y=f(x)$ — положительна. Левая часть формулы (3) имеет смысл для тех x, для каждого из которых функции $y=f(x)$ и $y=g(x)$ либо одновременно положительны, либо одновременно отрицательны, а правая — лишь для тех x, для каждого из которых функции $y=f(x)$ и $y=g(x)$ одновременно положительны.

Если при решении уравнения применить к одной или нескольким функциям уравнения любую из рассматриваемых формул так, что левая часть этой формулы будет заменена правой частью, то возможна потеря корней исходного уравнения, поэтому такие преобразования **недопустимы**.

Если к одной или нескольким функциям уравнения применить любую из рассматриваемых формул так, что правая часть формулы будет заменена левой, то, вообще говоря, не всякий корень последующего уравнения будет являться корнем исходного уравнения и поэтому, если такие преобразования применялись, то в конце решения обязательно необходима **проверка**, т. е. необходимо каждый из найденных корней последующего уравнения подставить в исходное уравнение и убедиться, какие из них обращают исходное уравнение в верное числовое равенство. Те из них, при каждом из которых исходное уравнение превращается в неверное числовое равенство, нужно отбросить.

Ниже приводятся примеры, показывающие, что применение этих формул приводит как к потере корней исходного уравнения, так и к приобретению посторонних корней.

1. Решить уравнение $3^{\log_3(x^2-4x+3)}=x-3$. Применяя формулу (1), получим уравнение $x^2-4x+3=x-3$, корни которого $x_1=2$ и $x_2=3$. Поскольку преобразование было таким,

при котором могли появиться посторонние корни, то следует сделать проверку. Проверка показывает, что оба эти корня являются посторонними, так как каждый из них не удовлетворяет исходному уравнению.

2. Решить уравнение $\log_2(x+2)^2=6$. Если применить формулу (2), то получим уравнение $\log_2(x+2)=3$, откуда $x+2=8$, или $x=6$. Поскольку при таком преобразовании могли потеряться корни, то нельзя считать, что исходное уравнение решено. И действительно, при этом преобразовании потерян корень исходного уравнения $x=-10$.

Эти примеры показывают, что при решении уравнений применять логарифмические формулы надо очень внимательно, помня о том, что их применение может привести и к потере и к приобретению посторонних корней. Этого не может случиться, если применять каждую из формул (1)—(3) на том множестве M_1 из ОДЗ решаемого уравнения, на котором имеет смысл правая часть соответствующей формулы. Тогда такое преобразование приведет к уравнению, равносильному исходному на множестве M_1. При этом надо помнить, что таким образом уравнение решено не на всей ОДЗ, а только на множестве M_1. Поэтому надо найти еще его решения на той части ОДЗ, которая останется после выбрасывания множества M_1. Уравнения, для решения которых применимы рассматриваемые формулы, часто решаются так.

1. Отыскивают ОДЗ уравнения.

2. Разбивают ОДЗ на 2 части M_1 и M_2 (M_1 — та часть ОДЗ, где одновременно имеют смысл обе части применяемой формулы, M_2 — та часть ОДЗ, которая остается после выделения множества M_1).

3. Ищут решения на M_1, учитывая, что на M_1 применение формулы дает уравнение, равносильное исходному.

4. Затем ищут свой способ решения исходного уравнения на множестве M_2.

5. Объединяют корни, найденные на M_1 и M_2.

Например, ОДЗ уравнения $3^{\log_3(x^2-4x+3)}=x-3$ есть любое x, для которого $x^2-4x+3>0$, т. е. все $x>3$, а также все $x<1$. Поскольку для любого x из ОДЗ функция $y=x^2-4x+3$ положительна, то, применяя формулу (1), получим, что исходное уравнение равносильно на ОДЗ уравнению $x^2-4x+3=x-3$ или $x^2-5x+6=0$, которое имеет корни $x_1=2$, $x_2=3$. Так как ни один из этих корней не входит в ОДЗ исходного уравнения, а последнее и исходное уравнения равносильны на ОДЗ исходного, то отсюда вытекает, что исходное уравнение не имеет корней.

ОДЗ уравнения $\log_2(x+2)^2=6$ есть любое x, кроме $x=-2$. Разобьем ОДЗ на две части: $M_1=\{x>-2\}$ и $M_2=\{x<-2\}$ и решим уравнение отдельно на каждой из этих частей.

Пусть $x \in M_1$, тогда функция $y = x + 2$ положительна на этом множестве. Поэтому на множестве M_1 исходное уравнение равносильно уравнению $\log_2(x+3) = 3$, откуда $x + 2 = 8$, или $x = 6$. Так как $x = 6$ входит в рассматриваемое множество $M_1 = \{x > -2\}$, значит, это и есть корень исходного уравнения на множестве $M_1 = \{x > -2\}$.

Пусть $x \in M_2$. Тогда функция $y = -x - 2$ положительна на этом множестве и, кроме того, $(x+2)^2 = (-x-2)^2$. Поэтому на множестве $M_2 = \{x < -2\}$ исходное уравнение равносильно уравнению $\log_2(-x-2) = 3$, откуда $-x - 2 = 8$, или $x = -10$. Так как $x = -10$ входит в рассматриваемое множество $M_2 = \{x < -2\}$, то это и есть корень исходного уравнения на множестве M_2. Объединяя корни, найденные в рассмотренных случаях, получаем ответ: исходное уравнение имеет два корня: $x_1 = 6$ и $x_2 = -10$.

Преобразования, связанные с тригонометрическими формулами. Рассмотрим следующие тригонометрические формулы:

$$\operatorname{ctg} x = \frac{1}{\operatorname{tg} x}, \tag{4}$$

$$\sin 2x = \frac{2 \operatorname{tg} x}{1 + \operatorname{tg}^2 x}, \tag{5}$$

$$\cos 2x = \frac{1 - \operatorname{tg}^2 x}{1 + \operatorname{tg}^2 x}, \tag{6}$$

$$\operatorname{tg}(\alpha - \beta) = \frac{\operatorname{tg}\alpha - \operatorname{tg}\beta}{1 + \operatorname{tg}\alpha \operatorname{tg}\beta}. \tag{7}$$

Левые части этих формул имеют смысл в более широких областях, чем правые части. Действительно, левая часть формулы (4) имеет смысл для любого x, кроме $x = \pi k$, а правая — для любого x, кроме $x = \pi k$ и $x = {}^{\pi}/_2 + \pi p$, где k и p — любые целые числа. Левые части формул (5) и (6) имеют смысл для любого x, а правые — для любого x, кроме $x = {}^{\pi}/_2 + \pi m$, где m — любое целое число. Левая часть формулы (7) имеет смысл для любого α и любого β, кроме тех, для которых $\alpha - \beta = {}^{\pi}/_2 + \pi n$, а правая — для любого α и любого β, кроме тех, для которых $\alpha = {}^{\pi}/_2 + \pi l$, $\beta = {}^{\pi}/_2 + \pi q$, $\alpha - \beta = {}^{\pi}/_2 + \pi n$, где n, l, q — любые целые числа.

Если при решении уравнения применить к одной или нескольким функциям уравнения любую из рассматриваемых формул так, что левая часть этой формулы будет заменена правой частью, то возможны потери корней исходного уравнения, поэтому такие преобразования **недопустимы**. Если к одной или нескольким функциям уравнения применить любую из рассматриваемых формул так, что правая часть этой формулы будет заменена левой, то можно **приобрести** посторонние корни, а поэтому в конце решения необходима **проверка**. Приведем примеры, показывающие, что применение этих формул при решении уравнения приводит как к потере корней исходного уравнения, так и к приобретению посторонних корней.

1. Решить уравнение $\sin 2x - 3\cos 2x = 3$.

Применяя формулы (5) и (6), получим

$$\frac{2\,\mathrm{tg}\,x}{1+\mathrm{tg}^2 x} - \frac{3\,(1-\mathrm{tg}^2 x)}{1+\mathrm{tg}^2 x} = 3.$$

Перепишем это уравнение в виде

$$\frac{2\,(\mathrm{tg}\,x - 3)}{1+\mathrm{tg}^2 x} = 0,$$

откуда очевидно, что решением этого уравнения будет любое $x = \mathrm{arctg}\,3 + \pi t$, где $t = 0, \pm 1, \pm 2, \pm \ldots$. Однако нельзя утверждать, что найдены все корни исходного уравнения, ибо преобразования были такими, что корни могли быть потеряны. Действительно, числа $x + {}^{\pi}/_{2} + \pi l$ при любом $l = 0, \pm 1, \pm 2, \pm \ldots$ также являются корнями исходного уравнения. Эти корни были утеряны из-за применения формул (5) и (6). Значит, применяя формулы (5) и (6), исходное уравнение решить нельзя.

2. Решить уравнение $\frac{\mathrm{tg}\,x - 1}{\mathrm{tg}\,x + 1} = 1$. Поскольку $1 = \mathrm{tg}\,{}^{\pi}/_{4}$, то уравнение можно переписать так:

$$\frac{\mathrm{tg}\,x - \mathrm{tg}\,{}^{\pi}/_{4}}{1 + \mathrm{tg}\,x\,\mathrm{tg}\,{}^{\pi}/_{4}} = 1.$$

Применяя теперь формулу (7), получаем уравнение $\mathrm{tg}\,(x - {}^{\pi}/_{4}) = 1$. Решая это простейшее уравнение, находим его корни $x = {}^{\pi}/_{2} + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$. Поскольку преобразование было таким, при котором могли появиться посторонние корни, то необходима проверка. Проверка показывает, что ни один из этих корней не будет корнем данного уравнения. Посторонние корни возникли из-за применения формулы (7).

Эти примеры показывают, что при решении уравнений применять тригонометрические формулы надо внимательно, помня о том, что их применение может привести и к потере и к приобретению корней. Этого не может случиться, если применять каждую из формул (4)—(7) на том множестве M_1 из ОДЗ, на котором имеет смысл правая часть соответствующей формулы. Тогда такое преобразование приведет к уравнению, равносильному исходному на множестве M_1.

Поэтому уравнения, при решении которых применимы эти тригонометрические формулы, часто решаются по той же схеме, которая приведена при рассмотрении логарифмических формул.

1. Отыскивают ОДЗ уравнения.

2. Разбивают ОДЗ на 2 части M_1 и M_2 (M_1 — та часть ОДЗ, где одновременно имеют смысл обе части применяемой

формулы, M_2 — та часть ОДЗ, которая остается после выделения множества M_1).

3. Ищут решение на M_1, учитывая, что на M_1 применение формулы дает уравнение, равносильное исходному.

4. Ищут свой способ решения исходного уравнения на множестве M_2.

5. Объединяют корни, найденные на M_1 и M_2. Этим способом приведенные выше примеры могут быть решены так.

ОДЗ уравнения $\sin 2x - 3\cos 2x = 3$ есть любое действительное x. Разобьем множество всех действительных чисел на две части: любое x, кроме $x = {}^{\pi}/_2 + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$, и $x = {}^{\pi}/_2 + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$. Затем решим исходное уравнение на каждой из этих частей.

Пусть $x \neq {}^{\pi}/_2 + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$, тогда для любого такого x исходное уравнение равносильно уравнению

$$\frac{2\,\mathrm{tg}\,x}{1+\mathrm{tg}^2 x} - \frac{3\,(1-\mathrm{tg}^2 x)}{1+\mathrm{tg}^2 x} = 3.$$

Решение этого уравнения есть $x = \mathrm{arctg}\,3 + \pi t$, где $t = 0, \pm 1, \pm 2, \pm \ldots$. Все эти значения x входят в рассматриваемое множество, значит, все они и есть корни исходного уравнения на этом множестве.

Пусть теперь $x = {}^{\pi}/_2 + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$. Проверкой убеждаемся, что все эти числа являются корнями исходного уравнения. Объединяя корни, найденные в рассмотренных случаях, получаем ответ: исходное уравнение имеет две серии решений: $x = \mathrm{arctg}\,3 + \pi t$, где $t = 0, \pm 1, \pm 2, \pm \ldots$; $x = {}^{\pi}/_2 + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$.

ОДЗ уравнения $\frac{(\mathrm{tg}\,x - 1)}{\mathrm{tg}\,x + 1} = 1$ есть любое x, кроме $x = \pi/2 + \pi l$, где $l = 0, \pm 1, \pm 2, t \ldots$, и $x = -\pi/4 + \pi q$, где $q = 0, \pm 1, \pm 2, \pm \ldots$. На ОДЗ исходное уравнение равносильно уравнению $\mathrm{tg}\,(x - \pi/4) = 1$, которое имеет корни $x = \pi/2 + \pi m$, где $m = 0, \pm 1, \pm 2, \pm \ldots$. Так как ни один из этих корней не входит в ОДЗ исходного уравнения, то отсюда вытекает, что исходное уравнение не имеет корней.

Преобразования, связанные с возведением в натуральную степень. Как следует из утверждения 5 § 1, при возведении в натуральную степень обеих частей уравнения **нельзя потерять** корни, а можно лишь **приобрести** лишние. Поэтому, если применять это преобразование, то в конце необходима проверка найденных корней.

Приведем пример, показывающий, что применение этого преобразования приводит к приобретению посторонних корней.

1. Решить уравнение $2\cos x = 3\sin x - 2$. Возведя в квадрат это уравнение и применяя основное тригонометрическое тождество, получим $4\,(1 - \sin^2 x) = 9\sin^2 x - 12\sin x + 4$. Перепишем это уравнение в виде $13\sin^2 x - 12\sin x = 0$, откуда

$\sin x = 0$ или $\sin x = 12/13$, т. е. $x = \pi n$, где $n = 0, \pm 1, \pm 2, \pm \ldots$, или $x = (-1)^n \operatorname{arctg} {}^{12}/_{13} + \pi k$, где $k = 0, \pm 1, \pm 2, \pm \ldots$. Поскольку возводя в квадрат можно приобрести посторонние корни, то необходима проверка. Проверка показывает, что из первой серии корнями исходного уравнения будут $x = (2q+1)\pi$, где $q = 0, \pm 1, \pm 2, \pm \ldots$, а из второй серии корнями исходного уравнения будут $x = \arcsin 12/13 + 2\pi p$, где $p = 0, \pm 1, \pm 2, \pm \ldots$. Заметим еще, что очень часто действие возведения в натуральную степень сопровождается следующим преобразованием:

$$\left[\sqrt[n]{f(x)}\right]^n = f(x).$$

Ясно, что область существования функции, стоящей в правой части, вообще говоря, шире, чем область существования функции, стоящей в левой части. Поэтому, если при решении уравнения заменить, например, $\left[\sqrt{f(x)}\right]^2$ на $f(x)$, то можно приобрести посторонние корни, и поэтому в конце решения необходима проверка.

2. Решить уравнение $\sqrt{x^2+x-5} = \sqrt{x-1}$. Возводя это уравнение в квадрат и заменяя $\left(\sqrt{x^2+x-5}\right)^2$ на x^2+x-5, а $\left(\sqrt{x-1}\right)^2$ на $(x-1)$, получаем уравнение $x^2+x-5 = x-1$ или $x^2-4=0$. Корни последнего уравнения $x_1 = +2$ и $x_2 = -2$. Поскольку обе части исходного уравнения возводились в квадрат и проводилась замена $\left(\sqrt{f(x)}\right)^2$ на $f(x)$, то могли появиться посторонние корни, а поэтому необходима проверка. Проверка показывает, что исходное уравнение имеет только один корень $x_1 = +2$.

Заметим, что обычно замена $\left[\sqrt{f(x)}\right]^2$ на $f(x)$ при возведении уравнения в квадрат не оговаривается.

Вместо этого способа решения можно предложить и другой, основанный на применении утверждения 7 § 1.

1. Отыскивают ОДЗ уравнения.

2. Разбивают ОДЗ на две части M_1 и M_2, причем так, что на M_1 неотрицательны обе части уравнения, на M_2 правая и левая части уравнения имеют разные знаки.

Тогда очевидно, что на M_2 нет решений уравнения, а все они содержатся на M_1.

3. Решают уравнение на множестве M_1. Все найденные на M_1 корни и будут решениями исходного уравнения (здесь учитывается, что возведение в степень обеих частей исходного уравнения на множестве M_1 приводит к равносильному на M_1 уравнению).

Решим таким способом уравнение $\sqrt{2x+29} = 3-x$. ОДЗ этого уравнения $x \geqslant -29/2$. Разобьем ОДЗ на 2 части: $-29/2 \leqslant x \leqslant 3$ и $x > 3$. На множестве $x > 3$ нет решений исходного уравнения, ибо для любого x из этого множества слева стоит положительное число, а справа — отрицательное. Остается решить это уравнение на множестве $-29/2 \leqslant x \leqslant 3$. На этом множестве обе части уравнения неотрицательны, поэтому после возведения в квадрат получим уравнение $2x + 29 = (3-x)^2$, равносильное исходному на множестве $-29/2 \leqslant x \leqslant 3$. Последнее уравнение имеет корни $x = -2$ и $x = 10$. Из них в множество $-29/2 \leqslant x \leqslant 3$ входит лишь корень $x = -2$. Итак, исходное уравнение имеет только один корень $x = -2$.

Преобразования, связанные с освобождением от знаменателя. Пусть задано уравнение $f(x)/g(x) = \varphi(x)$. Замена этого уравнения на уравнение $f(x) = g(x)\varphi(x)$ называется освобождением уравнения от знаменателя. Ясно, что ОДЗ второго из этих уравнений шире, чем первого. Это расширение ОДЗ произошло за счет добавления в ОДЗ первого уравнения тех значений x, при которых функция $g(x)$ обращается в нуль. При применении этого преобразования возможно появление посторонних корней, а потому в конце необходима проверка.

Вот один из примеров применения этого преобразования. Решить уравнение $\frac{2(x-7)}{x^2-6x-7} = 1$. Освобождаясь от знаменателя, приходим к уравнению $2x - 14 = x^2 - 6x - 7$, или $x^2 - 8x + 7 = 0$, которое имеет корни $x_1 = 1$ и $x_2 = 7$. Поскольку при освобождении от знаменателя можно приобрести посторонние корни, то необходима проверка. Проверка показывает, что исходное уравнение имеет только один корень $x = 1$. Заметим, что согласно утверждению 9 § 1 уравнение $f(x)/g(x) = \varphi(x)$ равносильно на своей ОДЗ уравнению $f(x) = g(x)\varphi(x)$. Поэтому подобные уравнения чаще решаются так.

1. Отыскивают ОДЗ уравнения $f(x)/g(x) = \varphi(x)$.
2. Решают на ОДЗ этого уравнения равносильное ему уравнение $f(x) = g(x)\varphi(x)$, корни которого, попавшие в ОДЗ исходного уравнения, и будут корнями исходного уравнения.

Решим этим способом уравнение $\frac{2(x-7)}{x^2-6x-7} = 1$.

ОДЗ этого уравнения любое x, кроме $x = -1$ и $x = 7$. На ОДЗ исходное уравнение равносильно уравнению $x^2 - 8x + 7 = 0$, которое имеет корни $x_1 = 1$ и $x_2 = 7$. Из этих корней в ОДЗ исходного уравнения попадает лишь корень $x = 1$. Итак, исходное уравнение имеет только один корень $x = 1$.

Решим этим способом уравнение $\frac{\sin x^4 - 1}{\cos^4 x} = 1$.

ОДЗ этого уравнения любое x, кроме $x = \pi/2 + \pi m$, где $m = 0, \pm 1, \pm 2, \pm \ldots$. На ОДЗ исходное уравнение равносильно уравнению $\sin^4 x - 1 = \cos^4 x$, которое можно переписать в виде $\cos^4 x - \sin^4 x = -1$. Поскольку $\cos^4 x - \sin^4 x = (\cos^2 x + \sin^2 x)(\cos^2 - \sin^2 x) = \cos 2x$ при любом x, то остается на ОДЗ исходного уравнения решить уравнение $\cos 2x = -1$, которое имеет корни $x = \pi/2 + \pi l$, где $l = 0, \pm 1, \pm 2, \pm \ldots$.

Ни один из этих корней не входит в ОДЗ исходного уравнения. Значит, исходное уравнение не имеет корней.

Преобразования, связанные с сокращением уравнения на общий множитель. Пусть дано уравнение $\varphi(x) f(x) = \varphi(x) g(x)$. Часто это уравнение заменяют уравнением $f(x) = g(x)$, т. е. сокращают исходное уравнение на общий множитель $\varphi(x)$ и при этом обычно теряют корни исходного уравнения. Это грубая ошибка. Подобные уравнения надо решать только следующим образом.

1. Найти ОДЗ уравнения $\varphi(x) f(x) = \varphi(x) g(x)$.
2. Переписать уравнение в виде $\varphi(x)[f(x) - g(x)] = 0$.
3. Перейти от этого уравнения к равносильной ему на ОДЗ исходного уравнения совокупности уравнений

$$\varphi(x) = 0, \quad f(x) - g(x) = 0.$$

4. Решить эту совокупность уравнений на ОДЗ исходного уравнения.

Решим этим способом уравнение $\sqrt{x-2}\,(x^2+3) = 4x\sqrt{x-2}$.

ОДЗ уравнения — все действительные $x \geqslant 2$. Перепишем уравнение в виде $\sqrt{x-2}\,(x^2 - 4x + 3) = 0$ и перейдем к совокупности уравнений $\sqrt{x-2} = 0$, $x^2 - 4x + 3 = 0$, помня о том, что она равносильна исходному уравнению на множестве $x \geqslant 2$. Решим эту совокупность на множестве $x \geqslant 2$. Корень первого уравнения $x_0 = 2$ входит в ОДЗ исходного уравнения и, значит, является его корнем. Второе уравнение имеет корни $x_1 = 1$ и $x_2 = 3$. Из этих корней только корень $x_2 = 3$ входит в ОДЗ исходного уравнения, значит, только он является его корнем.

Итак, исходное уравнение имеет два корня: $x_0 = 2$ и $x_2 = 3$.

Преобразования, связанные с потенцированием уравнения. Замена уравнения $\log_a f(x) = \log_a g(x)$ $(a > 0,\ a \neq 1)$ на уравнение $f(x) = g(x)$ называется *потенцированием уравнения*. Как следует из утверждения 6 § 1, при потенцировании уравнения потерять корни нельзя, а можно лишь приобрести посторонние. Поэтому если при решении уравнения пришлось его потенцировать, то в конце решения необходима проверка.

Решим таким способом уравнение $\log_2(x^2 - 4) = \log_2(4x - 7)$. Потенцируя, приходим к уравнению $x^2 - 4 = 4x - 7$, которое имеет корни $x_1 = 3$, $x_2 = 1$. Поскольку при решении исходного уравнения применялось потенцирование, то могли появиться посторонние корни, поэтому необходима проверка. Проверка показывает, что исходное уравнение имеет только один корень $x = 3$.

Заметим, что подобные уравнения можно решать и несколько иначе. Для этого надо воспользоваться утверждением 8 § 1. В этом случае уравнения решают так.

Отыскивают множество M_1, на котором обе функции $y=f(x)$ и $y=g(x)$ одновременно положительны. На этом множестве уравнения $\log_a f(x)=\log_a g(x)$ и $f(x)=g(x)$ равносильны. Поэтому, решая уравнение $f(x)=g(x)$ на этом множестве, тем самым решают и исходное уравнение.

Решим этим способом предыдущий пример. Функции $y=x^2-4$ и $y=4x-7$ одновременно положительны на множестве $x>2$. На этом множестве уравнение $\log_2(x^2-4)=\log_2(4x-7)$ равносильно уравнению $x^2-4=4x-7$. Последнее уравнение имеет два корня: $x_1=3$ и $x_2=1$. Из них в множество $x>2$ входит лишь один корень $x=3$. Значит, он и есть решение исходного уравнения.

Преобразования, связанные с логарифмированием уравнения. Замена уравнения $f(x)=g(x)$ на уравнение $\log_a f(x)=\log_a g(x)$ $(a>0,\ a\neq 1)$ называется *логарифмированием уравнения*. Как следует из утверждения 6 § 1, при логарифмировании уравнения возможна потеря корней. Поэтому формальное применение этого преобразования **запрещается**.

Логарифмировать уравнение $f(x)=g(x)$ можно только на том множестве M_1 из ОДЗ, на котором обе функции $y=f(x)$ и $y=g(x)$ одновременно положительны. Тогда, как следует из утверждения 8 § 1, на множестве M_1 уравнения $f(x)=g(x)$ и $\log_a f(x)=\log_a g(x)$ $(a>0,\ a\neq 1)$ будут равносильны. На оставшейся части ОДЗ исходное уравнение следует решать своим способом.

Пример. Решить уравнение $3^{x^2-x}=2^{1-(\sqrt{x})^2}$.

ОДЗ этого уравнения все $x\geqslant 0$. Обе части этого уравнения положительны для всех $x\geqslant 0$, поэтому для любого $x\geqslant 0$ исходное уравнение равносильно уравнению $x^2-x=(1-x)\log_3 2$. Решая последнее уравнение, находим его корни $x_1=1$ и $x_2=-\log_3 2$. Условию $x\geqslant 0$ из этих двух корней удовлетворяет только $x_1=1$, он и будет корнем исходного уравнения.

Подведем итог рассуждениям этого параграфа.

Преобразования, при применении которых возможно появление посторонних корней, следующие:

возведение уравнения в натуральную степень;

освобождение уравнения от знаменателя;

потенцирование уравнения;

применение логарифмических, тригонометрических и других формул, расширяющих ОДЗ уравнения.

Если при решении уравнения применялись эти преобразования и при этом оговаривалось, что при их применении нельзя потерять корни, то в конце решения необходима проверка, ибо посторонние корни все же могли появиться, а их надо отбросить.

Преобразования, при применении которых возможна потеря корней уравнения, следующие:

извлечение корня из обеих частей уравнения;

сокращение обеих частей уравнения на множитель, содержащий x;

логарифмирование уравнения;

применение логарифмических, тригонометрических и других формул, сужающих ОДЗ исходного уравнения.

Применять эти преобразования при решении уравнений, вообще говоря, нельзя, ибо при формальном их применении возможна потеря корней.

Все отмеченные выше преобразования можно применять при решении уравнения, если все время следить за равносильностью исходного и последующих уравнений на соответствующих множествах ОДЗ исходного уравнения.

Конечно, последний способ решения всегда приведет к цели, но часто его применение осложняется трудоемкими вычислениями той области (множества), на которой преобразования равносильны. Поэтому лучше пользоваться преобразованиями, при применении которых нельзя потерять корни, а в конце решения сделать проверку найденных корней.

Отметим, что в этом параграфе рассмотрены не все возможные неравносильные преобразования, а лишь наиболее часто употребляемые. Конечно, можно привести примеры и других неравносильных преобразований, как, например, уничтожение подобных членов, переход к новому основанию логарифмов, содержащему неизвестную величину и другие, и для каждого из них указать причины потери или появление лишних корней. Не будем останавливаться на этом подробно, а сформулируем лишь общее правило.

При решении уравнений надо пользоваться одним из двух следующих способов.

1. Замена данного уравнения на уравнение, равносильное ему на некотором множестве; при этом явно указывается и само множество и то, что уравнения равносильны именно на этом множестве (в этом случае надо рассмотреть все множества, на которые разбивается ОДЗ).

2. Замена данного уравнения на уравнение, являющееся его **следствием**; при этом указывается, почему новое уравнение есть следствие предыдущего, и в конце решения обязательно делается проверка.

§ 5
ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И УТВЕРЖДЕНИЯ РАВНОСИЛЬНОСТИ НЕРАВЕНСТВ

Областью допустимых значений (ОДЗ) неравенства $f(x) > g(x)$ называется общая часть (пересечение) областей существования функций $y=f(x)$ и $y=g(x)$, т. е. множество всех числовых значений неизвестного x, при каждом из которых имеют смысл (определены) левая и правая части неравенства.

Число α из ОДЗ неравенства называется *решением неравенства* $f(x)>g(x)$, если при подстановке его вместо неизвестного x неравенство превращается в верное числовое неравенство $f(\alpha)>g(\alpha)$.

Решить неравенство — значит найти **все** его решения, т. е. найти **все** действительные числа, принадлежащие ОДЗ этого неравенства, каждое из которых обращает данное неравенство в верное числовое неравенство, или доказать, что таких чисел нет.

Пусть даны два неравенства: $f(x)>g(x)$ и $p(x)>\varphi(x)$. Если любое решение первого неравенства является решением второго неравенства и, наоборот, любое решение второго неравенства является решением первого неравенства, то такие два неравенства называются *равносильными*. Замена одного неравенства другим неравенством, ему равносильным, называется *равносильным переходом* от одного неравенства к другому.

Пусть даны два неравенства: $f(x)>g(x)$ и $p(x)>\varphi(x)$ и пусть дано некоторое множество M значений неизвестного x. Если любое решение первого неравенства, принадлежащее множеству M, является решением второго неравенства и, наоборот, любое решение второго неравенства, принадлежащее множеству M, является решением первого неравенства, то такие два неравенства называются *равносильными на множестве* M. Замена одного неравенства другим неравенством, равносильным ему на множестве M, называется *равносильным переходом на множестве* M от одного неравенства к другому.

Отметим, что аналогично формулируются задачи решения неравенств $f(x)<g(x)$, $f(x)\geqslant g(x)$ и $f(x)\leqslant g(x)$ и основные определения для них.

Приведем несколько примеров, иллюстрирующих введенные понятия.

Пусть надо решить неравенство $\sqrt{x+2}+\sqrt{x-5}>\sqrt{3-x}$. ОДЗ этого неравенства — пересечение областей существования функций $y=\sqrt{x+2}$, $y=\sqrt{x-5}$, $y=\sqrt{3-x}$, т. е. пересечение множеств $[-2;\ +\infty)$, $[5;\ +\infty)$, $(-\infty;\ 3]$. Это пересечение пусто. Следовательно, неравенство решено, так как нет ни одного значения неизвестного, при котором все функции, входящие в данное неравенство, имели бы смысл. Таким образом, данное неравенство решений не имеет.

Пусть надо решить неравенство $\sqrt{\sin x}<-1$. Так как при любом числовом значении неизвестного, взятого из ОДЗ неравенства, оно становится неверным числовым неравенством, то, следовательно, данное неравенство решений не имеет.

Неравенства $x+5>0$ и $(x^4+1)(x+5)>0$ равносильны на множестве всех действительных чисел; неравенства $\sqrt{x}>1$ и $x^2>1$ не являются равносильными на множестве всех действительных чисел, но равносильны, например, на множестве положительных чисел.

Приведем некоторые утверждения равносильности неравенств.

1. Неравенства $f(x)>g(x)$ и $f(x)-g(x)>0$ равносильны.

2. Неравенства $f(x)>g(x)$ и $f(x)+\alpha>g(x)+\alpha$ равносильны при любом действительном α.

3*а*. Неравенства $f(x)>g(x)$ и $\alpha f(x)>\alpha g(x)$ равносильны для любого положительного числа α.

3*б*. Неравенства $f(x)>g(x)$ и $\alpha f(x)<\alpha g(x)$ равносильны для любого отрицательного числа α.

4*а*. Неравенства $a^{f(x)}>a^{g(x)}$ и $f(x)>g(x)$ равносильны для любого фиксированного числа a из промежутка $(1; +\infty)$.

4*б*. Неравенства $a^{f(x)}>a^{g(x)}$ и $f(x)<g(x)$ равносильны для любого фиксированного числа a из промежутка $(0; 1)$. Справедливость этих утверждений доказывается аналогично, поэтому приведем доказательство лишь утверждения 3а.

Пусть число x_1 есть некоторое решение неравенства $f(x)>g(x)$, т. е. пусть существуют числа $f(x_1)$ и $g(x_1)$, для которых справедливо числовое неравенство $f(x_1)>g(x_1)$. Умножив это числовое неравенство на положительное число α, получим, что справедливо числовое неравенство $\alpha f(x_1)>\alpha g(x_1)$, а это означает, что число x_1 есть решение неравенства $\alpha f(x)>\alpha g(x)$. Такое рассуждение можно провести для любого решения неравенства $f(x)>g(x)$. Значит, любое решение неравенства $f(x)>g(x)$ является решением неравенства $\alpha f(x)>\alpha g(x)$.

Покажем теперь обратное. Пусть число x_2 есть некоторое решение неравенства $\alpha f(x)>\alpha g(x)$, т. е. пусть существуют числа $f(x_2)$ и $g(x_2)$, для которых справедливо числовое неравенство $\alpha f(x_2)>\alpha g(x_2)$. Из справедливости этого неравенства на основании свойств числовых неравенств вытекает справедливость числового неравенства $f(x_2)>g(x_2)$, а это означает, что число x_2 есть решение неравенства $f(x)>g(x)$. Такое рассуждение можно провести для любого решения неравенства $\alpha f(x)>\alpha g(x)$. Значит, любое решение неравенства $\alpha f(x)>\alpha g(x)$ является решением неравенства $f(x)>g(x)$. Утверждение 3*а* доказано.

Приведем теперь несколько утверждений равносильности неравенств на множествах.

5. Пусть n — натуральное число и пусть на некотором множестве M одновременно обе функции $y=f(x)$ и $y=g(x)$ неотрицательны, тогда на этом множестве равносильны неравенства

$$f(x)>g(x) \text{ и } [f(x)]^n>[g(x)]^n.$$

6*а*. Пусть a — фиксированное число из промежутка $(1; \infty)$ и пусть на некотором множестве M одновременно обе функции $y=f(x)$ и $y=g(x)$ положительны, тогда на этом множестве равносильны неравенства $f(x)>g(x)$ и $\log_a f(x)>\log_a g(x)$.

6б. Пусть a — фиксированное число из промежутка $(0; 1)$ и пусть на некотором множестве M одновременно обе функции $y=f(x)$ и $y=g(x)$ положительны, тогда на этом множестве равносильны неравенства $f(x)>g(x)$ и $\log_a f(x)<\log_a f(x)$.

7. Пусть на некотором множестве M функция $y=\varphi(x)$ положительна, тогда на этом множестве равносильны неравенства

$$f(x)>g(x) \text{ и } f(x)\varphi(x)>g(x)\varphi(x).$$

Справедливость утверждений 5—7 доказывается аналогично, поэтому приведем здесь доказательство лишь утверждения 5.

Если $n=1$, то утверждение 5 очевидно. Поэтому будем дальше считать, что $n\geqslant 2$. Пусть число x_1 принадлежит множеству M и является некоторым решением неравенства $f(x)>g(x)$, т. е. пусть существуют неотрицательные числа $f(x_1)$ и $g(x_1)$, для которых справедливо числовое неравенство $f(x_1)>g(x_1)$. Из справедливости этого неравенства следует, в частности, что число $f(x_1)$ положительно. Но тогда для любого натурального числа k число $[f(x_1)]^k$ положительно, а число $[g(x_1)]^k$ — неотрицательно. Значит, сумма

$$[f(x_1)]^{n-1}+[f(x_1)]^{n-2}g(x_1)+\ldots+f(x_1)[g(x_1)]^{n-2}+[g(x_1)]^{n-1}$$

положительна, так как первое ее слагаемое — положительно, а остальные — неотрицательны. Из числового неравенства $f(x_1)>g(x_1)$ вытекает еще, что число $f(x_1)-g(x_1)$ положительно. Так как произведение положительных чисел положительно, то положительно и число

$$[f(x_1)-g(x_1)]\{[f(x_1)]^{n-1}+[f(x_1)]^{n-2}g(x_1)+\ldots+[g(x_1)]^{n-1}\}.$$

Применяя теперь формулу сокращенного умножения, приходим к справедливости числового неравенства $[f(x_1)]^n-[g(x_1)]^n>0$, откуда вытекает справедливость числового неравенства $[f(x_1)]^n>[g(x_1)]^n$.

Итак, показано, что для любого числа x_1 из множества M из справедливости числового неравенства $f(x_1)>g(x_1)$ следует справедливость числового неравенства $[f(x_1)]^n>[g(x_1)]^n$. Значит, любое решение неравенства $f(x)>g(x)$, принадлежащее множеству M, является решением неравенства $[f(x)]^n>[g(x)]^n$.

Покажем теперь обратное. Пусть теперь число x_2 принадлежит множеству M и есть некоторое решение неравенства $[f(x)]^n>[g(x)]^n$, т. е. пусть существуют неотрицательные числа $f(x_2)$ и $g(x_2)$, для которых справедливо числовое неравенство $[f(x_2)]^n>[g(x_2)]^n$. Покажем, что число $f(x_2)$ положительно. Предположим, что число $f(x_2)$ равно нулю, тогда из неравенства $[f(x_2)]^n>[g(x_2)]^n$ вытекает, что число $[g(x_2)]^n$ отрицательно. Но так как число $g(x_2)$ неотрицательно, то неотрицательно, число $[g(x_2)]^n$. Полученное противоречие означает, что число $f(x_2)$ положительно. Но тогда положительна сумма

$$[f(x_2)]^{n-1}+[f(x_2)]^{n-2}g(x_2)+\ldots+[g(x_2)]^{n-1}.$$

Кроме того, из справедливости неравенства $[f(x_2)]^n>[g(x_2)]^n$ следует, что число $[f(x_2)]^n-[g(x_2)]^n$ положительно. Рассмотрим теперь числовое равенство

$$[f(x_2)]^n-[g(x_2)]^n=[f(x_2)-g(x_2)]\times$$
$$\times\{[f(x_2)]^{n-1}g(x_2)+[f(x_2)]^{n-2}g(x^2)+\ldots+[g(x_2)]^{n-1}\}.$$

В этом равенстве слева стоит положительное число, а справа — произведение двух чисел, одно из которых положительно, значит, и второе число положительно, т. е. справедливо числовое неравенство $f(x_2)-g(x_2)>0$. Из справедливости этого числового неравенства вытекает справедливость неравенства $f(x_2)>g(x_2)$. Итак, показано, что для любого числа x_2 из множества M из справедливости числового неравенства $[f(x_2)]^n>[g(x_2)]^n$ следует справедливость числового неравенства $f(x_2)>g(x_2)$. Значит, любое решение неравенства $[f(x)]^n>[g(x)]^n$, принадлежащее множеству M, является решением неравенства $f(x)>g(x)$, чем и завершается доказательство утверждения 5.

Говорят, что дана *система m неравенств*

$$\begin{cases} f_1(x)>g_1(x), \\ f_2(x)>g_2(x), \\ \cdots\cdots\cdots \\ f_m(x)>g_m(x), \end{cases} \tag{1}$$

если требуется найти **все** такие числа α, каждое из которых является решением каждого неравенства системы (1) (неравенства системы обычно записывают в столбик и объединяют фигурной скобкой).

Таким образом, решить систему неравенств (1) — значит сначала решить каждое неравенство $f_i(x)>g_i(x)$, $i=1, 2, \ldots, m$, т. е. найти множества $N_1, N_2, \ldots, N_m$ всех решений каждого из этих неравенств, а затем взять пересечение всех этих множеств $N=N_1\cap N_2\cap\ldots\cap N_m$. Любое число из множества N и будет решением системы неравенств (1).

Говорят, что дана *совокупность k систем неравенств*

$$\begin{cases} f_{11}(x)>g_{11}(x), \\ f_{21}(x)>g_{21}(x), \\ \cdots\cdots\cdots \\ f_{m1}(x)>g_{m1}(x), \end{cases} \quad \begin{cases} f_{12}(x)>g_{12}(x), \\ f_{22}(x)>g_{22}(x), \\ \cdots\cdots\cdots \\ f_{m2}(x)>g_{m2}(x), \end{cases} \quad \ldots \quad \begin{cases} f_{1k}(x)>g_{1k}(x), \\ f_{2k}(x)>g_{2k}(x), \\ \cdots\cdots\cdots \\ f_{mk}(x)>g_{mk}(x), \end{cases} \tag{2}$$

если требуется найти **все** числовые значения неизвестного, каждое из которых является решением хотя бы одной из систем совокупности (2). Каждое такое числовое значение называется *решением совокупности систем* (2). Заметим, что если каждая из k систем совокупности (2) имеет только одно неравенство, то говорят, что дана *совокупность k неравенств*; если $k=1$, то совокупность (2) является системой неравенств.

Говорят, что *неравенство*

$$f(x)>g(x) \tag{3}$$

равносильно совокупности систем неравенств (2), если любое решение неравенства (3) является решением совокупности (2) и, наоборот, любое решение совокупности (2) является решением неравенства (3). В случае, если в совокупности систем (2) $m_1=m_2=\ldots=m_k=1$, то говорят, что *неравенство* (3) *равносильно совокупности неравенств* (2). В случае, если в совокупности систем (2) $k=1$, то говорят, что *неравенство* (3) *равносильно системе неравенств* (2). В случае, если в совокупности систем (2) окажется систем бесконечно много, то говорят, что дана *бесконечная совокупность систем неравенств*.

Наконец, приведем понятие *смешанной совокупности* — совокупности уравнений и неравенств. Говорят, что дана смешанная совокупность

$$f_1(x)=g_1(x),\ f_2(x)=g_2(x),\ \ldots,\ f_k(x)=g_k(x),\ \ f_{k+1}(x)>g_{k+1}(x),$$
$$f_{k+2}(x)>g_{k+2}(x),\ \ldots,\ f_{k+m}(x)>g_{k+m}(x), \tag{4}$$

если требуется найти **все** числовые значения неизвестного x, каждое из которых является решением хотя бы одного из k уравнений или хотя бы одного из m неравенств из совокупности (4). Следовательно, *решить смешанную совокупность* (4) — значит решить все уравнения $f_i(x)=g_i(x)$, $i=1,\ 2,\ \ldots,\ k$, и тем самым найти множества $N_1,\ N_2,\ \ldots,\ N_k$ всех решений этих уравнений, затем решить все неравенства $f_{k+j}(x)>g_{k+j}(x)$, $j=1,\ 2,\ \ldots,\ m$, т. е. найти множества $N_{k+1},\ N_{k+2},\ \ldots,\ N_{k+m}$ всех решений этих неравенств и, наконец, взять объединение этих множеств: $N=N_1\cup N_2\cup\ldots\cup N_k\cup N_{k+1}\cup N_{k+2}\cup\ldots\cup N_{k+m}$. Каждое число α из множества N будет называться *решением смешанной совокупности* (4).

Будем говорить, что *неравенство*

$$f(x)\geqslant g(x) \tag{5}$$

равносильно смешанной совокупности (4), если любое решение неравенства (5) является решением смешанной совокупности (4) и, наоборот, любое решение смешанной совокупности (4) является решением неравенства (5). Например, нестрогое неравенство $f(x)\geqslant g(x)$ равносильно смешанной совокупности $f(x)>g(x)$, $f(x)=g(x)$. В связи с этим обычно рассматриваются лишь строгие неравенства, ибо решение нестрогого неравенства есть объединение решений соответствующего уравнения и строгого неравенства.

§ 6
ПРОСТЕЙШИЕ НЕРАВЕНСТВА

Простейшими неравенствами называются неравенства вида

$$f(x) > b \quad (f(x) < b), \tag{1}$$

где $y=f(x)$ — основная элементарная функция, b — некоторое фиксированное действительное число. Как уже отмечалось выше, нестрогое неравенство

$$f(x) \geqslant b \quad (f(x) \leqslant b)$$

равносильно совокупности строгого неравенства и уравнения

$$f(x) > 0, \ f(x) = b \quad (f(x) < 0, \ f(x) = b).$$

Поэтому будем рассматривать лишь неравенства вида (1).

Прежде всего следует сказать, что при решении неравенства нельзя формально записать решение соответствующего уравнения и затем заменить знак равенства на знак неравенства. Как будет показано ниже для решения простейших неравенств, надо хорошо знать свойства функций, входящих в состав неравенства и уметь пользоваться ими.

Большую помощь при решении простейших неравенств оказывают графики функций, входящих в состав неравенств. Обычно они хорошо иллюстрируют сам процесс отыскания решения данного неравенства. При этом пользуются следующим очевидным утверждением: если надо решить неравенство $f(x) > b$ или $f(x) < b$, то строят на одном чертеже графики функций $y=f(x)$ и $y=b$. Тогда решением неравенства $f(x) > b$ будут те значения x, для каждого из которых график функции $y = f(x)$ лежит **выше** графика функции $y=b$ (рис. 113), а решением неравенства $f(x) < b$ будут те значения x, для каждого из которых график функции $y=f(x)$ лежит **ниже** графика функции $y=b$ (рис. 114).

Перечислим простейшие неравенства и рассмотрим их решение.

Алгебраические неравенства. Пусть n — данное натуральное число. Рассмотрим решение неравенств

$$x^n > b, \tag{2}$$

$$x^n < b. \tag{3}$$

Как известно (см. § 3 гл. VI), областью существования функции $y=x^n$ является вся числовая ось, области монотонного изменения этой функции различны для четного и нечетного показателя n. Поэтому рассмотрим два случая.

Пусть n — нечетное натуральное число, т. е. $n=2m-1$, $m=1, 2, \ldots$. Построим (рис. 115) графики функций $y=x^{2m-1}$ и $y=b$. Вследствие того что функция $y=x^{2m-1}$ является возрастающей функцией на всей числовой прямой, каждое численное значение она принимает лишь один раз. Поэтому графики

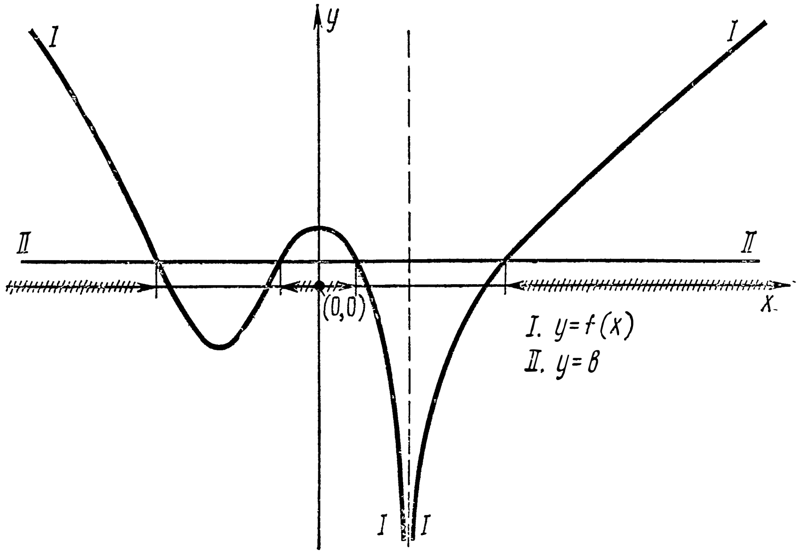


Рис. 113

функций $y=x^{2m-1}$ и $y=b$ пересекаются лишь в одной точке x_0, причем $x_0=\sqrt[2m-1]{b}$, если $b\geqslant 0$, и $x_0=-\sqrt[2m-1]{|b|}$, если $b<0$. Кроме того, вследствие возрастания функции $y=x^{2m-1}$

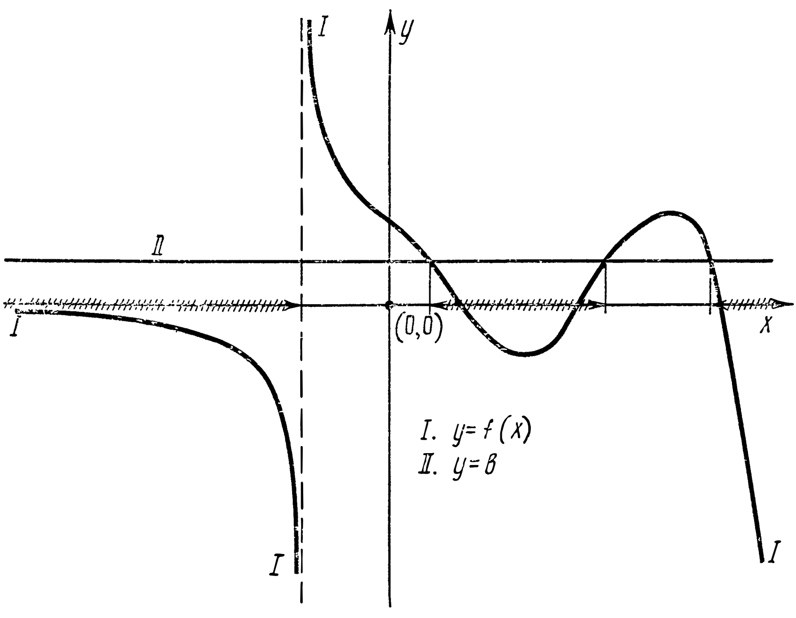


Рис. 114

на всей числовой прямой, для любого x, расположенного правее x_0, эта функция принимает значения бо́льшие, чем число b, а для любого x, расположенного левее x_0, эта функция принимает значения меньшие, чем число b. Поэтому правее точки x_0 график функции $y=x^{2m-1}$ лежит выше графика функции $y=b$, а левее — ниже.

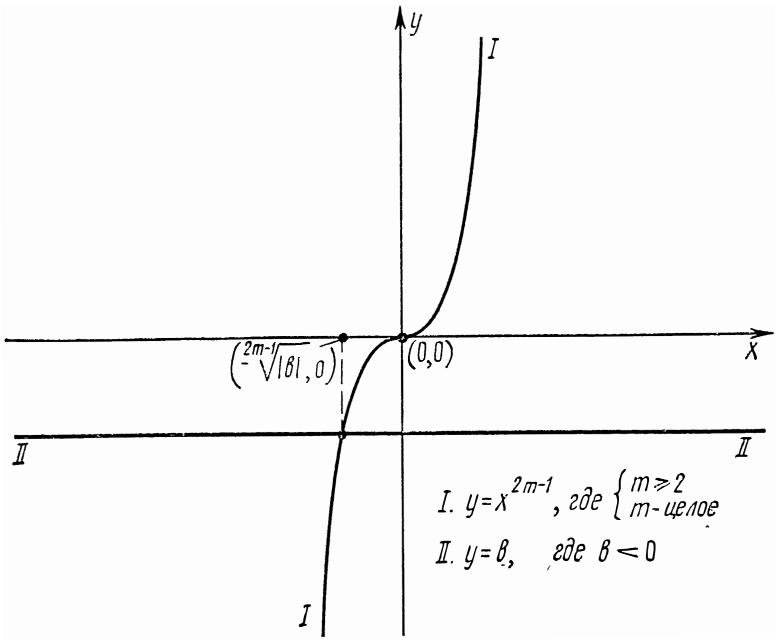


Рис. 115

Значит, решением неравенства (2) будут все $x \in (x_0;\ +\infty)$, а решением неравенства (3) будут все $x \in (-\infty;\ x_0)$.

Пусть n — четное натуральное число, т. е. $n=2m$. Функция $y=x^{2m}$ (рис. 116) имеет областью существования всю числовую ось, и на всей этой области она является неотрицательной. Поэтому, если число b отрицательно, то неравенство (2) справедливо при любом значении x, т. е. все действительные числа есть решения неравенства (2), другими словами, решением неравенства (2) в этом случае будут все $x \in (-\infty;\ +\infty)$. Если же $b=0$, то при $x=0$ функция $y=x^{2m}$ принимает значение — нуль, а для всех остальных x эта функция положительна, поэтому решением неравенства (2) в этом случае будут все действительные x, кроме $x=0$; другими словами, решением неравенства (2) в этом случае будут все x из объединения двух лучей: $(-\infty;\ 0)$ и $(0;\ +\infty)$. Наконец, рассмотрим случай $b>0$. Сначала рассмотрим решение неравенства (2) в области неотрицательных чисел. Для $x \geqslant 0$ функция $y = x^{2m}$ возрастает и в точке $x_0 = \sqrt[2m]{b}$ принимает значение b. Вследствие возрастания для всех $x > x_0$ эта функция будет принимать значения бо́льшие, чем b, а для всех x из промежутка $0 \leqslant x < x_0$ — меньшие, чем b.

Следовательно, решением неравенства (2) в области неотрицательных чисел будут те и только те x, которые больше,

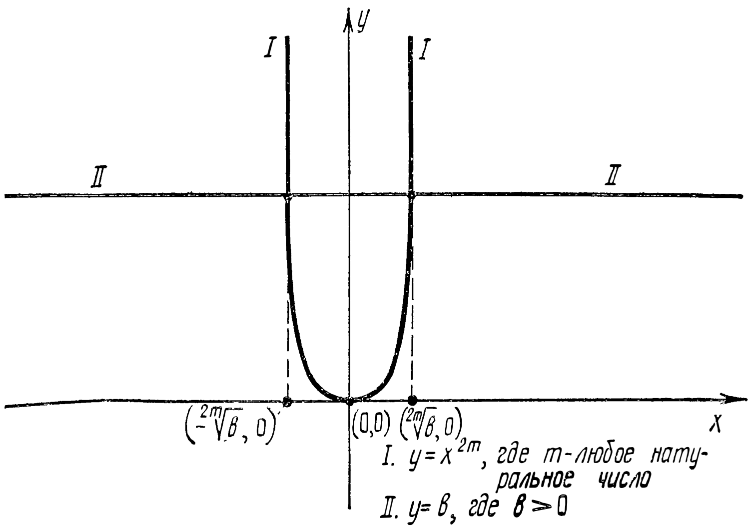


Рис. 116

чем x_0, т. е. все $x \in (x_0;\ +\infty)$. Так как функция $y = x^{2m}$ является четной функцией, то в области отрицательных чисел решением неравенства (2) будут все $x \in (-\infty;\ -x_0)$. Итак, решением неравенства (2) в случае $b > 0$ будут все x из объединения двух лучей: $(-\infty;\ -x_0) \cup (x_0;\ +\infty)$. Вышеизложенное решение неравенства (2) можно свести в следующую таблицу:

	$b>0$	$b=0$	$b<0$
$x^{2m-1} > b$	$x \in \left(\sqrt[2m-1]{b}; +\infty\right)$	$x \in (0;\ +\infty)$	$x \in \left(-\sqrt[2m-1]{\lvert b \rvert};\ +\infty\right)$
$x^{2m} > b$	$x \in \left(-\infty; -\sqrt[2m]{b}\right) \cup \left(\sqrt[2m]{b}; +\infty\right)$	$x \in (-\infty; 0) \cup (0; +\infty)$	$x \in (-\infty; +\infty)$

Аналогично рассматривается неравенство (3). Решение неравенства (3) также можно записать в виде таблицы.

	$b>0$	$b=0$	$b<0$
$x^{2m-1}<b$	$x\in\left(-\infty;\ \sqrt[2m-1]{b}\right)$	$x\in(-\infty,\ 0)$	$x\in\left(-\infty;\ -\sqrt[2m-1]{\lvert b\rvert}\right)$
$x^{2m}<b$	$x\in\left(-\sqrt[2m]{b};\ \sqrt[2m]{b}\right)$	нет решений	нет решений

Заметим, что запоминать эти таблицы, как и все следующие, не надо. Они приводятся здесь только как итог предыдущих рассуждений. При решении конкретного неравенства надо не вспоминать соответствующую таблицу, а провести соответствующее рассуждение.

Дробные неравенства. Пусть n — данное натуральное число. Рассмотрим решение неравенств

$$x^{-n}>b, \tag{4}$$

$$x^{-n}<b. \tag{5}$$

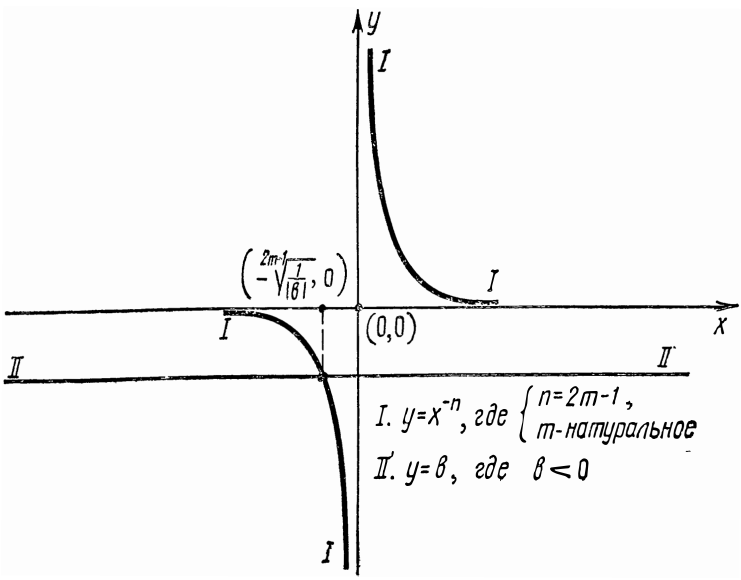


Рис. 117

ОДЗ неравенств (4) и (5) — вся числовая прямая, за исключением одной точки — нуля. Рассмотрим два случая. Пусть n — нечетное натуральное число, т. е. $n=2m-1$, $m=1, 2, \ldots$.

Если $x > 0$, то $x^{-(2m-1)} > 0$, а если $x < 0$, то $x^{-(2m-1)} < 0$ (рис. 117). Поэтому, если $b = 0$, то неравенство (4) имеет решение $x \in (0; +\infty)$, а неравенство (5) имеет решение $x \in \in (-\infty; 0)$. Если $b > 0$, то рассмотрим сначала решение неравенств на положительной полуоси. Так как для положительных x функция $y = x^{-(2m-1)}$ убывает и в точке $x_0 = \sqrt[2m-1]{1/b}$ принимает значение, равное b, то решением неравенства (4) для этих x будет интервал $(0; x_0)$, а решением неравенства (5) будет луч $(x_0; +\infty)$. Так как для отрицательных x функция $y = x^{-(2m-1)}$ принимает отрицательные значения (график функции $y = x^{-(2m-1)}$ лежит ниже оси OX и, следовательно, ниже прямой $y = b$), то очевидно, что на всей числовой прямой неравенство (4) имеет решение $x \in (0; x_0)$, а неравенство (5) имеет решение $x \in (-\infty; 0) \cup (x_0; +\infty)$.

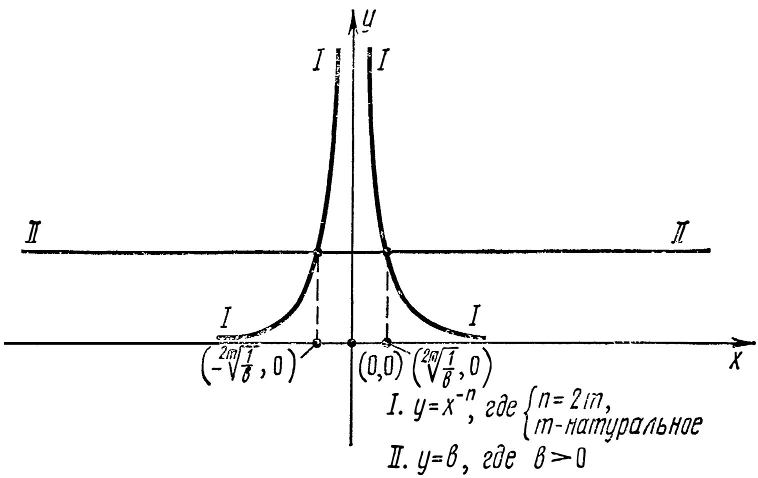


Рис. 118

Аналогично ищется решение при $b < 0$, а именно: неравенство (4) имеет решение $x \in (-\infty; x_1) \cup (0; +\infty)$, а неравенство (5) имеет решение $x \in (x_1; 0)$, где $x_1 = -\sqrt[2m-1]{1/|b|}$ (см. рис. 117). Пусть n — четное натуральное число, т. е. $n = 2m$. Тогда $x^{-2m} > 0$ для всех x из ОДЗ и, следовательно, при $b \leqslant 0$ решением неравенства (4) является вся область допустимых значений, а неравенство (5) не имеет решений. Если $b > 0$, то рассмотрим сначала решение неравенств на положительной полуоси. Так как функция $y = x^{-2m}$ убывает при $x > 0$, и в точке $x_0 = \sqrt[2m]{1/b}$ принимает значение b, то решением неравенства (4) для этих x является интервал $(0; x_0)$, а решением неравенства (5) — луч $(x_0; +\infty)$. На всей области допустимых значений в силу четности функции $y = x^{-2m}$ неравенство (4) имеет решение $x \in (-x_0;\ 0) \cup (0;\ x_0)$, неравенство (5) имеет решение $x \in (-\infty;\ -x_0) \cup (x_0;\ +\infty)$ (рис. 118).

Подведем итоги и запишем все решения дробного неравенства в виде таблицы.

	$b<0$	$b=0$	$b>0$
$x^{-(2m-1)}>b$	$x \in \left(-\infty;\ -\sqrt[2m-1]{1/\lvert b\rvert}\right) \cup$ $\cup (0;\ +\infty)$	$x \in (0;\ +\infty)$	$x \in \left(0;\ \sqrt[2m-1]{1/b}\right)$
$x^{-(2m-1)}<b$	$x \in \left(-\sqrt[2m-1]{1/\lvert b\rvert};\ 0\right)$	$x \in (-\infty;\ 0)$	$x \in (-\infty;\ 0) \cup$ $\cup \left(\sqrt[2m-1]{1/b};\ +\infty\right)$
$x^{-2m}>b$	$x \in (-\infty;\ +\infty)$	$x \in (-\infty;\ +\infty)$	$x \in \left(-\sqrt[2m]{1/b};\ 0\right) \cup$ $\cup \left(0;\ \sqrt[2m]{1/b}\right)$
$x^{-2m}<b$	нет решений	нет решений	$x \in \left(-\infty;\ -\sqrt[2m]{1/b}\right) \cup$ $\cup \left(\sqrt[2m]{1/b};\ +\infty\right)$

Степенные неравенства. Пусть α — данное нецелое действительное число. Рассмотрим решение неравенств

$$x^\alpha > b, \tag{6}$$
$$x^\alpha < b. \tag{7}$$

Рассмотрим два случая.

Пусть $\alpha > 0$. В этом случае ОДЗ неравенств (6) и (7) является множество всех неотрицательных чисел, т. е. $x \in [0;\ +\infty)$, и функция $y = x^\alpha$ неотрицательна на этом множестве и возрастает. Следовательно, при $b < 0$ неравенство (6) справедливо на всей области допустимых значений, а неравенство (7), наоборот, не имеет решений. При $b = 0$ неравенство (7) также не имеет решений, а неравенство (6) имеет решением луч $(0;\ +\infty)$, т. е. только $x = 0$ не является решением, а все остальные числа из ОДЗ удовлетворяют неравенству (6).

Если $b > 0$, то в точке $x_0 = b^{1/\alpha}$ функция $y = x^\alpha$ принимает значение, равное b (рис. 119). Поэтому во всех точках, расположенных правее точки x_0, значения функции $y = x^\alpha$ будут больше b, т. е. решением неравенства (6) является луч $(x_0;\ +\infty)$. Решением же неравенства (7) будет промежуток $[0;\ x_0)$, так как значение функции $y = x^\alpha$ в любой точке этого множества меньше b, а на отрицательной полуоси эта функция не определена.

Пусть $\alpha < 0$. ОДЗ неравенств (6) и (7) в этом случае является множество всех положительных чисел, т. е. $x \in$

$\in (0; +\infty)$ и функция $y = x^{\alpha}$ положительна на этом множестве и убывает. Поэтому при $b \leqslant 0$ неравенство (6) имеет решением всю ОДЗ, а неравенство (7) не имеет решения. При $b > 0$,

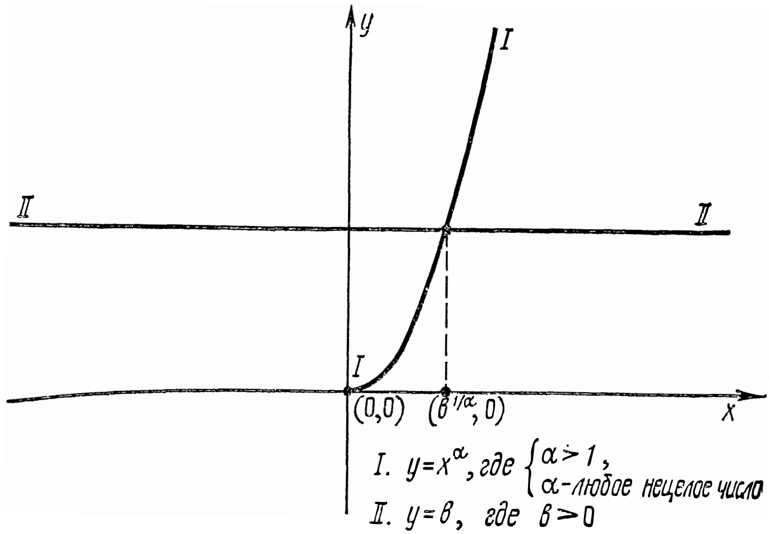


Рис. 119

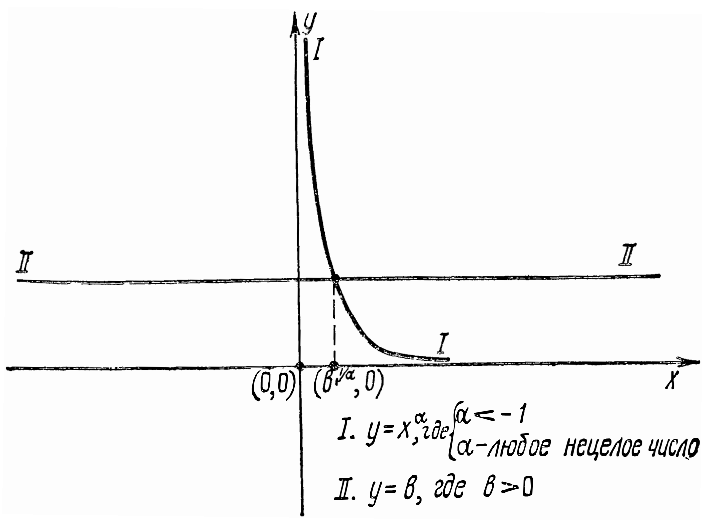


Рис. 120

рассуждая аналогично, получим решение неравенства (6): множество всех $x \in (0; b^{1/\alpha})$, а решение неравенства (7): множество всех $x \in (b^{1/\alpha}, +\infty)$ (рис. 120).

Вышеизложенные решения можно записать с помощью следующей таблицы:

	$b<0$	$b=0$	$b>0$
$x^\alpha > b$ $(\alpha>0)$	$x\in[0;\ +\infty)$	$x\in(0;\ +\infty)$	$x\in(b^{1/\alpha};\ +\infty)$
$x^\alpha < b$ $(\alpha>0)$	нет решения	нет решения	$x\in[0;\ b^{1/\alpha})$
$x^\alpha > b$ $(\alpha<0)$	$x\in(0;\ +\infty)$	$x\in(0;\ +\infty)$	$x\in(0;\ b^{1/\alpha})$
$x^\alpha < b$ $(\alpha<0)$	нет решения	нет решения	$x\in(b^{1/\alpha};\ +\infty)$

Показательные неравенства. Пусть a — фиксированное число, такое, что $a>0$ и $a\neq 1$. Рассмотрим неравенства

$$a^x > b, \tag{8}$$

$$a^x < b. \tag{9}$$

ОДЗ этих неравенств совпадает со всей числовой прямой, функция $y=a^x$ положительна и строго монотонна. Следовательно, при $b\leqslant 0$ неравенство (8) выполняется при любом x из ОДЗ, а неравенство (9) не имеет решений. При $b>0$ приходится рассмотреть два случая: $a>1$ и $0<a<1$.

Пусть $a>1$, тогда на всей числовой прямой функция $y=a^x$ является возрастающей (рис. 121). Значение, равное b, она принимает в единственной точке $x_0=\log_a b$, и потому решением неравенства (8) являются все $x\in(x_0;\ +\infty)$, а решением неравенства (9) — все $x\in(-\infty;\ x_0)$.

Пусть $0<a<1$, тогда на всей числовой прямой функция $y=a^x$ является убывающей (рис. 122), и потому решением неравенства (8) являются все $x\in(-\infty;\ x_0)$, а решением неравенства (9) — все $x\in(x_0;\ +\infty)$, где $x_0=\log_a b$.

Решения неравенств (8) и (9) можно представить в виде следующей таблицы:

	$b\leqslant 0$	$b>0$
$a^x>b$ $(a>1)$	$x\in(-\infty;\ +\infty)$	$x\in(\log_a b;\ +\infty)$
$a^x>b$ $(0<a<1)$	$x\in(-\infty;\ +\infty)$	$x\in(-\infty;\ \log_a b)$
$a^x<b$ $(a>1)$	нет решений	$x\in(-\infty;\ \log_a b)$
$a^x<b$ $(0<a<1)$	нет решений	$x\in(\log_a b;\ +\infty)$

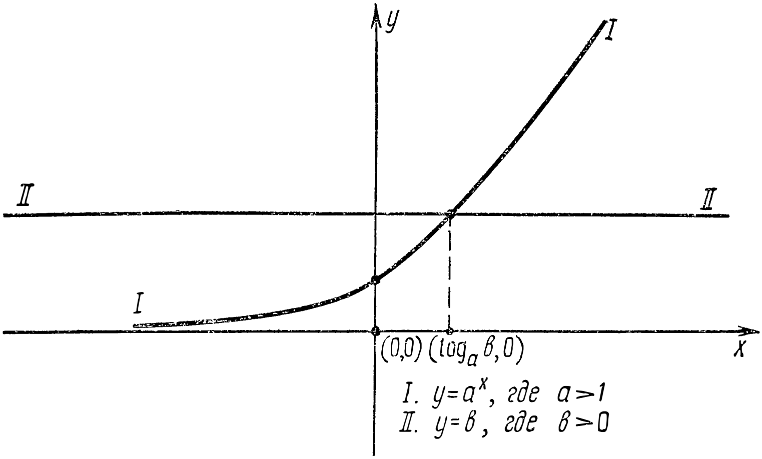


Рис. 121

Логарифмические неравенства. Пусть a — фиксированное число, такое, что $a > 0$ и $a \neq 1$. Рассмотрим неравенства

$$\log_a x > b, \qquad (10)$$

$$\log_a x < b. \qquad (11)$$

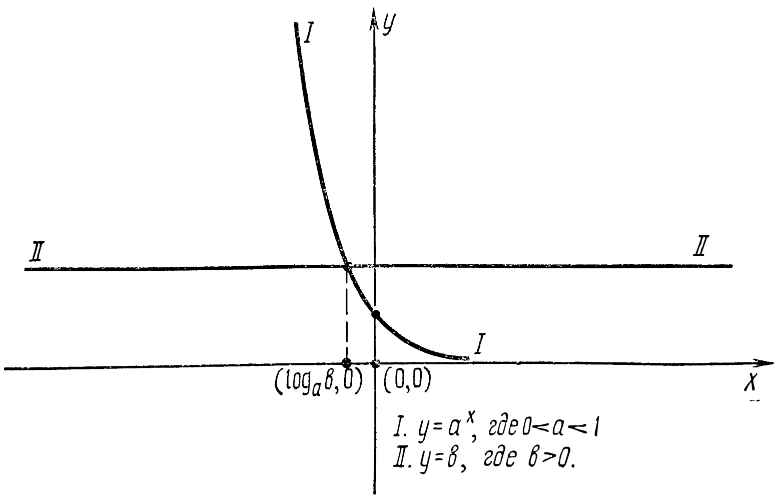


Рис. 122

ОДЗ этих неравенств является положительная полуось. Поскольку свойства логарифмической функции различны при основаниях, меньших и больших единицы, то рассмотрим случаи $a > 1$ и $0 < a < 1$.

Пусть $a > 1$. Тогда функция $y = \log_a x$ возрастает на всей области допустимых значений $(0;\ +\infty)$ (рис. 123) и принимает в точке $x_0 = a^b$ значение, равное b. Поэтому в каждой точке,

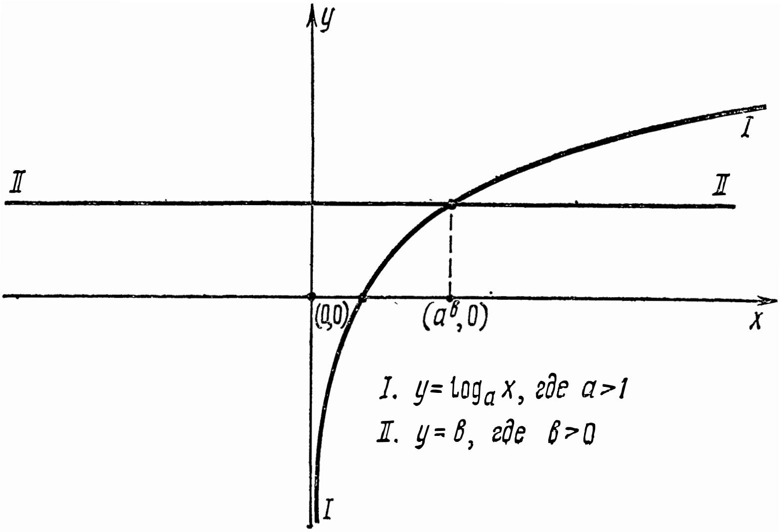


Рис. 123

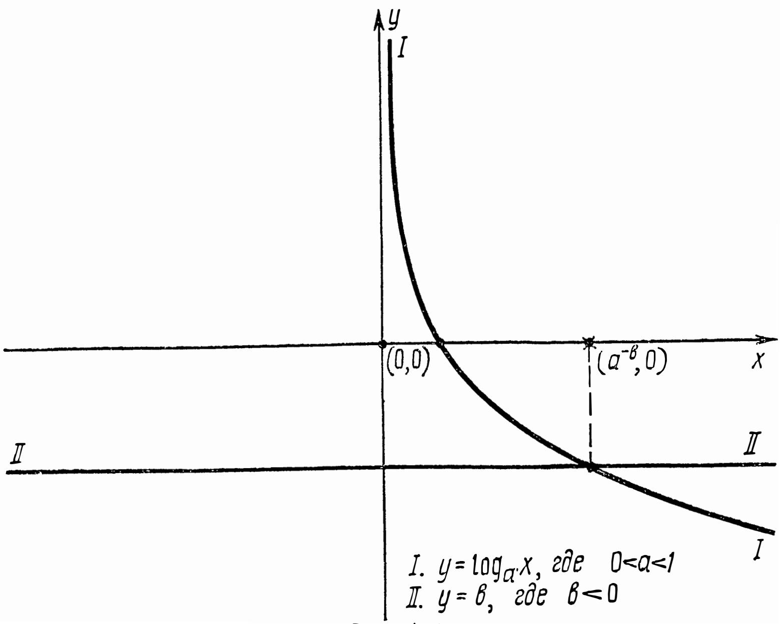


Рис. 124

лежащей правее точки x_0, значения функции $y = \log_a x$ будут больше, чем b, а в каждой точке ОДЗ, лежащей левее точки x_0, значения этой функции будут меньше, чем b. А это означает, что неравенство (10) выполняется для любого $x \in (x_0; +\infty)$, а неравенство (11) выполняется для любого $x \in (0; x_0)$.

Пусть $0 < a < 1$. Тогда функция $y = \log_a x$ убывает на всей области допустимых значений $(0; +\infty)$ (рис. 124). Рассуждая, как в случае $a > 1$, находим решения неравенств (10) и (11). Решение неравенства (10): $0 < x < a^b$, решение неравенства (11): $x > a^b$. При любом действительном b решения неравенств (10) и (11) можно записать в виде следующей таблицы:

	$a>1$	$0<a<1$
$\log_a x > b$	$x \in (a^b; +\infty)$	$x \in (0; a^b)$
$\log_a x < b$	$x \in (0; a^b)$	$x \in (a^b; +\infty)$

Тригонометрические неравенства. Прежде чем рассматривать простейшие тригонометрические неравенства, следует сделать несколько предварительных замечаний. Так как тригонометрические функции являются периодическими функциями, то решение неравенства следует находить на интервале длиной в главный период, а затем выписывать решения на всей области допустимых значений, используя периодичность. Интервал длиной в главный период может быть взят любым, но обычно выбирают так, чтобы он удовлетворял двум условиям: он должен содержать интервал, на котором для данной функции определена обратная тригонометрическая функция (см. § 4 гл. VI), и форма записи решения должна быть по возможности простой.

В случае тригонометрических функций $y = \operatorname{tg} x$ и $y = \operatorname{ctg} x$ вопрос о выборе интервала длиной в период решается просто: берется интервал, где определена соответствующая обратная функция. В случае тригонометрических функций $y = \cos x$ и $y = \sin x$ этот выбор лучше проводить в каждом конкретном случае по-своему.

Рассмотрим тригонометрическое неравенство (рис. 125)

$$\cos x > b. \qquad (12)$$

Если $b \geqslant 1$, то неравенство (12) не имеет решений. Если $b < -1$, то решением является вся числовая прямая. Рассмотрим решение при $-1 \leqslant b < 1$. В качестве отрезка длиной в период здесь можно взять либо отрезок $(-\pi; \pi]$, либо отрезок $[0; 2\pi)$. Эти отрезки длиной в главный период функции $y = \cos x$ полностью содержат отрезок $[0; \pi]$, на котором определена обратная функция. Из рисунка видно, что в случае неравенства (12) лучше взять отрезок $(-\pi; \pi]$, чем отрезок

[0; 2π), так как в первом случае решение можно представить в виде одного промежутка, а во втором случае решение является объединением двух промежутков. Итак, найдем решение неравенства (12) на множестве $(-\pi; \pi]$. Сначала решим неравенство (12) на отрезке $[0; \pi]$. Так как на отрезке $[0; \pi]$

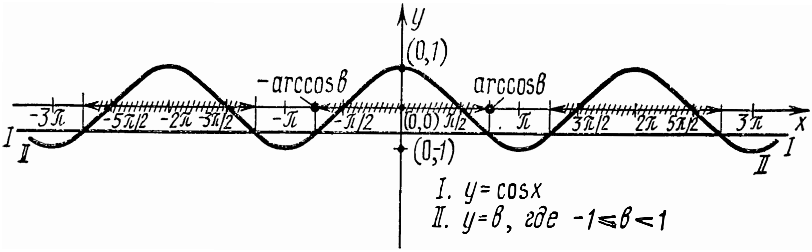


Рис. 125

функция $y = \cos x$ убывает и в точке $x_0 = \arccos b$ значение функции равно b, то решением неравенства (12) на отрезке $[0; \pi]$ будет промежуток $[0; \arccos b)$. Если теперь рассмотрим отрезок $(-\pi; \pi]$ длиной в период, то в силу четности функции $y = \cos x$ решением неравенства (12) на отрезке $(-\pi; \pi]$ будет интервал $(-\arccos b; \arccos b)$. Используя периодичность функции $y = \cos x$, получим решение неравенства на всей числовой прямой: это промежутки, отстоящие от исходного промежутка $(-\arccos b; \arccos b)$ на величину, кратную главному периоду 2π и имеющие такую же длину. Кратко все эти интервалы обычно записывают так:

$$-\arccos b + 2\pi k < x < \arccos b + 2\pi k, \text{ где } k = 0, \pm 1, \pm 2, \pm \ldots.$$

Отметим, что если $b = -1$, то решением уравнения (12) является множество всех действительных x, кроме $x = \pi + 2\pi k$, $k = 0, \pm 1, \pm 2, \pm \ldots$.

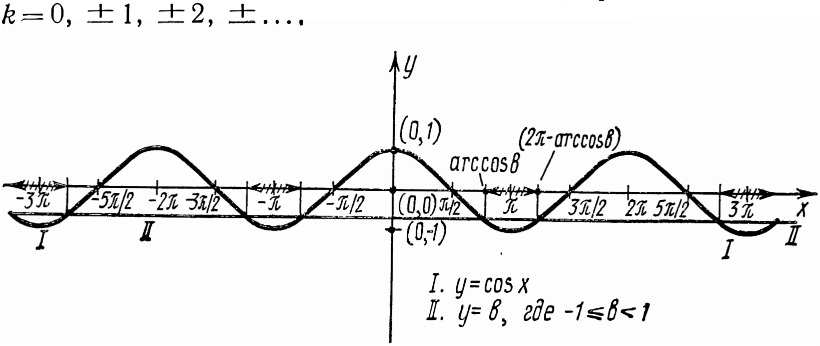


Рис. 126

Рассмотрим теперь неравенство (рис. 126)

$$\cos x < b. \tag{13}$$

При $b \leqslant -1$ неравенство (13) не имеет решения, при $b > 1$ решением неравенства (13) является вся числовая прямая. Для того чтобы решить неравенство (13) при $b \in (-1;\ 1]$ выберем отрезок длиной в период. Из рисунка видно, что в качестве такого отрезка лучше взять отрезок $[0;\ 2\pi)$, так как решение на нем можно записать в виде одного промежутка. Действительно, в силу строгой монотонности на отрезке $[0;\ \pi]$ и симметрии функции $y = \cos x$ относительно вертикальной прямой, проходящей через точку $(\pi;\ 0)$, решением на отрезке $[0;\ 2\pi)$ будет промежуток $(\arccos b;\ (2\pi - \arccos b)$.

Рассуждая аналогично предыдущему, получим, что на отрезке $(-\pi;\ \pi]$ решением неравенства (13) будут два промежутка $(\arccos b;\ \pi]$ и $(-\pi;\ -\arccos)$. Поэтому в первом случае решение на всей числовой прямой имеет вид

$$\arccos b + 2\pi k < x < (2\pi - \arccos b) + 2\pi k,$$
$$\text{где } k = 0,\ \pm 1,\ \pm 2,\ \pm \ldots .$$

Во втором случае решение на всей числовой прямой имеет более сложный вид $\arccos b + 2\pi m < x \leqslant \pi + 2\pi m$, где $m = 0,\ \pm 1,\ \pm 2,\ \pm \ldots$; $-\pi + 2\pi p \leqslant x < -\arccos b + 2\pi p$, где $p = 0,\ \pm 1,\ \pm 2,\ \pm \ldots$. Запишем полученные решения в виде таблицы.

	$b < -1$	$b = -1$	$-1 < b < 1$	$b = 1$	$b > 1$
$\cos x > b$	любое x	любое x, кроме $x = \pi + 2\pi n$, где $n = 0,\ \pm 1, \pm 2, \pm \ldots$	$-\arccos b + 2\pi k < x < \arccos b + 2\pi k$, где $k = 0,\ \pm 1,\ \pm 2,\ \pm \ldots$	нет решений	нет решений
$\cos x < b$	нет решений	нет решений	$-\arccos b + 2\pi m < x < (2\pi - \arccos b) + 2\pi m$, где $m = 0,\ \pm 1,\ \pm 2,\ \pm \ldots$	любое x, кроме $x = 2\pi t$, где $t = 0,\ \pm 1,\ \pm 2,\ \pm \ldots$	любое x

Рассмотрим тригонометрические неравенства (рис. 127, 128)

$$\sin x < b, \tag{14}$$

$$\sin x > b. \tag{15}$$

Проведя аналогичные рассуждения, можно получить следующую таблицу решений:

	$b<-1$	$b=-1$	$-1<b<1$	$b=1$	$b>1$
$\sin x<b$	нет решений	нет решений	$(-\pi-\arcsin b)+$ $+2\pi m<x<$ $<\arcsin b+2\pi m$, где $m=0, \pm 1,$ $\pm 2, \pm \ldots$	все x, кроме $x=\pi/2+2\pi t$, где $t=0,$ $\pm 1, \pm 2, \pm \ldots$	все x
$\sin x>b$	все x	все x, кроме $x=-\pi/2+$ $+2\pi k$, $k=0,$ $\pm 1, \mp 2, \pm \ldots$	$\arcsin b+2\pi n<x<$ $<(\pi-\arcsin b)+2\pi n$, где $n=0, \pm 1,$ $\pm 2, \pm \ldots$	нет решений	нет решений

Рассмотрим тригонометрические неравенства

$$\operatorname{tg} x<b, \tag{16}$$

$$\operatorname{tg} x>b. \tag{17}$$

Область допустимых значений этих неравенств — вся числовая прямая, за исключением точек $x=\pi/2+\pi k$, где $k=0, \pm 1,$

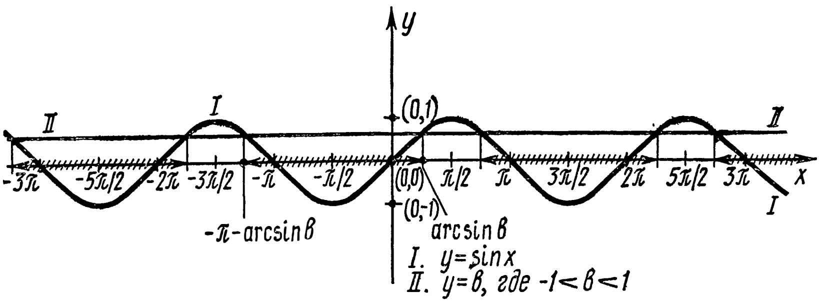


Рис. 127

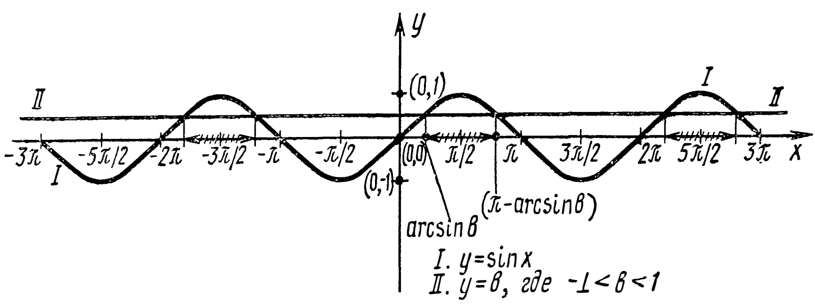


Рис. 128

$\pm 2, \pm \ldots$. Воспользовавшись периодичностью функции $y=$ $=\operatorname{tg} x$, рассмотрим решение неравенств (16), (17) на интервале

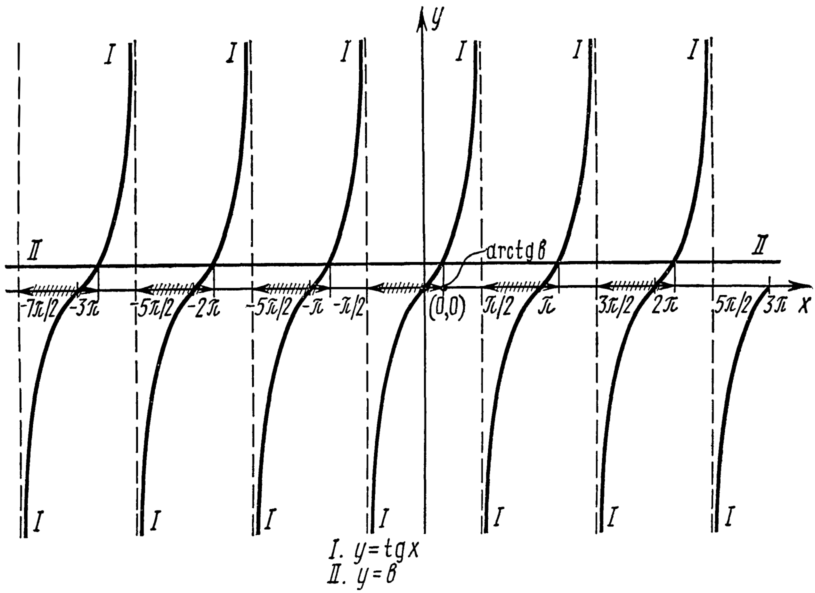


Рис. 129

$(-\pi/2;\ \pi/2)$ длиной в период (рис. 129). На этом интервале функция $y = \operatorname{tg} x$ возрастает и в точке $x_0 = \operatorname{arctg} b$ принимает значение, равное b. Следовательно, на интервале $(-{}^{\pi}/_{2};\ {}^{\pi}/_{2})$

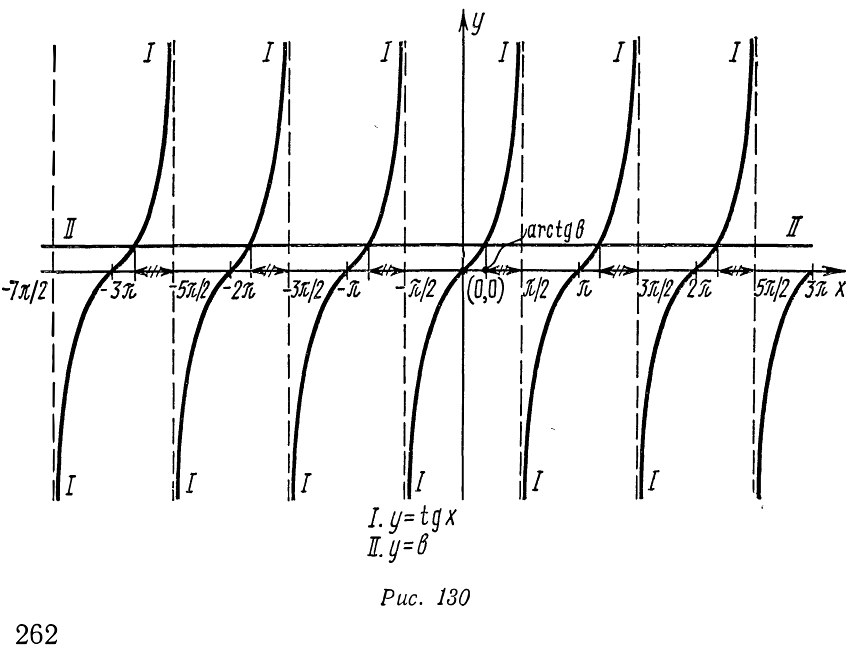


Рис. 130

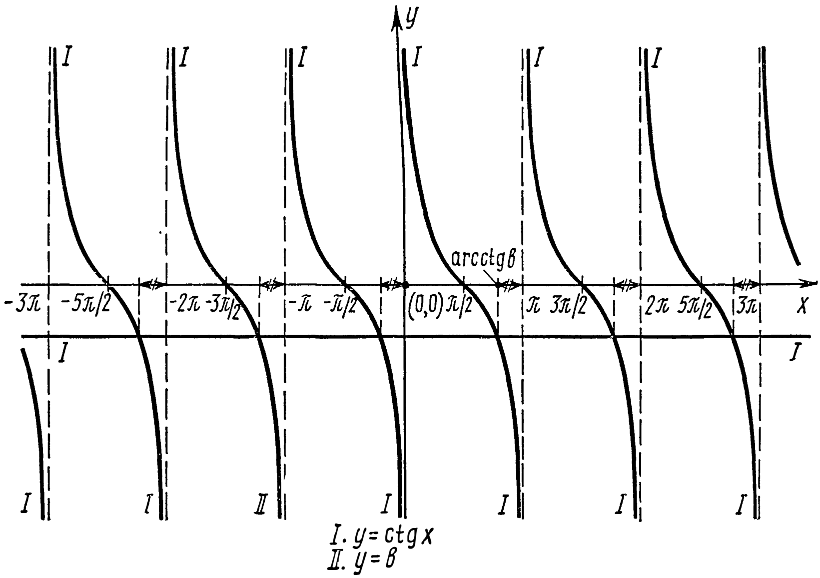


Рис. 131

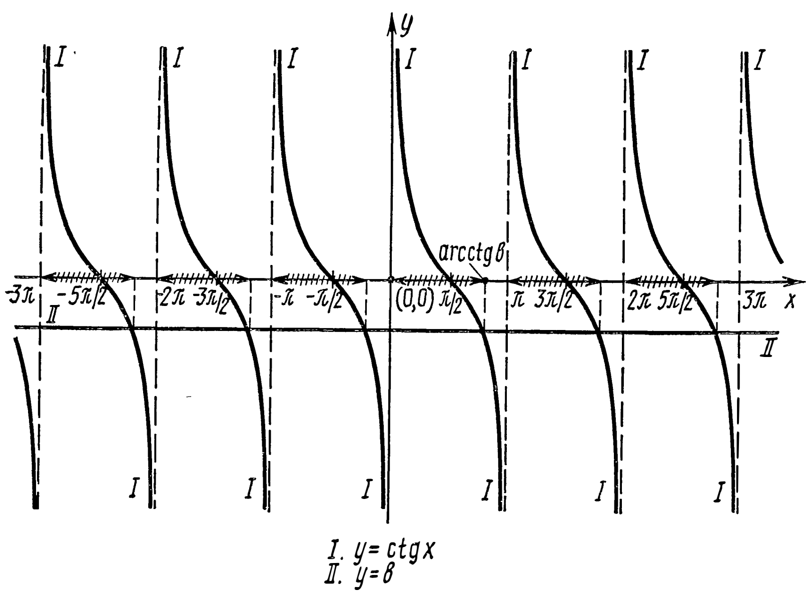


Рис. 132

неравенство (16) имеет решением любое $x \in (-\pi/2;\ x_0)$ а неравенство (17) любое $x \in (x_0,\ \pi/2)$. На всей же области допустимых значений решение неравенства (16) имеет вид (см. рис. 129) $-\pi/2+\pi m<x<x_0+\pi m$, где $m=0,\ \pm 1,\ \pm 2,\ \pm \ldots$, неравенства (17) (рис. 130) $x_0+\pi k<x<\pi/2+\pi k$, где $k=0,\ \pm 1,\ \pm 2,\ \pm \ldots$, а $x_0=\operatorname{arctg} b$.

Наконец, рассмотрим тригонометрические неравенства (рис. 131, 132)

$$\operatorname{ctg} x<b, \tag{18}$$

$$\operatorname{ctg} x>b. \tag{19}$$

Рассуждая аналогично предыдущему, получим решение неравенства (18) $x_0+\pi k<x<\pi+\pi k$ и неравенства (19) $\pi k<x<$ $<x_0+\pi k$, где $k=0,\ \pm 1,\ \pm 2,\ \ldots$, а $x_0=\operatorname{arctg} b$.

Следует еще раз отметить, что, несмотря на то что для решения простейших неравенств использовались решения простейших уравнений, основную роль все же сыграло знание свойств функций и умение ими пользоваться.

§ 7
РЕШЕНИЕ НЕРАВЕНСТВ

В § 3, 4 этой главы рассматривались соответственно равносильные и неравносильные преобразования уравнений. Причем в случае неравносильных преобразований отмечалось, что возможны два способа решения уравнений. Первый способ — это совершать равносильные переходы на множестве, принадлежащем области допустимых значений, но не обязательно совпадающем с ней. Второй способ — это переходы к уравнению, являющемуся следствием предыдущего, с обязательной проверкой полученных решений.

При решении неравенств второй способ на самом деле невозможен, так как множество действительных чисел, удовлетворяющих неравенству чаще всего бесконечно и в связи с этим проверка решений практически неосуществима. Поэтому при решении неравенств надо совершать только равносильные переходы. Кроме того, необходимо каждый раз отмечать, на каком множестве совершается равносильный переход.

Поэтому, в частности, формулы, отмеченные в § 4, можно применять при решении неравенств лишь на том множестве изменения неизвестной, на котором одновременно имеют смысл левая и правая части применяемой формулы.

Исходя из этого, неравенства, при решении которых хотят воспользоваться той или иной формулой, обычно решаются по такой схеме.

1. Отыскивают ОДЗ неравенства.

2. Разбивают ОДЗ на 2 части: M_1 и M_2. M_1 — та часть ОДЗ, где одновременно имеют смысл левая и правая части применяемой формулы; M_2 — та часть ОДЗ, которая остается после выделения множества M_1.

3. Решают неравенство на M_1, учитывая, что на нем применение формулы дает неравенство, равносильное на M_1 исходному.

4. Ищут свой способ решения неравенства на M_2.

5. Объединяют решения, найденные на M_1 и M_2.

Пример. Решить неравенство

$$\log_2 (x+2)^2 + \log_2 (x+2)^4 < 3. \qquad (1)$$

ОДЗ этого неравенства — все действительные числа x, кроме $x=-2$. Разобьем ОДЗ на 2 части: $M_1=(-2;\ +\infty)$ и $M_2=(-\infty;\ -2)$. Для любого x из M_1 справедливы формулы

$$\log_2 (x+2)^2 = 2\log_2 (x+2), \quad \log_2 (x+2)^4 = 4\log_2 (x+2). \quad (2)$$

Значит, на M_1 неравенство (1) равносильно неравенству

$$\log_2 (x+2) < 1/2.$$

Решая это простейшее неравенство, получаем, что любое x, удовлетворяющее условию $0 < x+2 < \sqrt{2}$, т. е. любое $x \in (-2;\ -2+\sqrt{2})$ является его решением. Поскольку любое такое x входит в M_1, то все x из промежутка $(-2;\ -2+\sqrt{2})$ являются решениями исходного неравенства на M_1.

Для x из M_2 равенства (2) не выполняются. Так как справедливы следующие равенства $(x+2)^2=(-x-2)^2$ и $(x+2)^4=(-2-x)^4$, то для любого x из множества M_2 справедливы равенства

$$\log_2 (x+2)^2 = 2\log_2 (-x-2),$$
$$\log_2 (x+2)^4 = 4\log_2 (-x-2). \qquad (3)$$

Значит, на M_2 неравенство (1) равносильно неравенству

$$\log_2 (-x-2) < 1/2.$$

Любое x, удовлетворяющее условию $0 < -x-2 < \sqrt{2}$, т. е. любое $x \in (-2-\sqrt{2};\ -2)$ является решением этого простейшего неравенства. Так как любое такое x входит в M_2, значит, все x из промежутка $(-2-\sqrt{2};\ -2)$ являются решениями исходного неравенства на M_2.

Объединяя решения, найденные на множествах M_1 и M_2, получаем, что множеством всех решений исходного неравенства является множество $(-2-\sqrt{2};\ -2) \cup (-2;\ -2+\sqrt{2})$.

Если для решения неравенства применяются преобразование возведение в степень, или потенцирование, или логарифмирование, или какое-либо другое, то необходимо будет действовать по следующей схеме:

1. Отыскать ОДЗ неравенства.

2. Разбить ОДЗ на две части: M_1 и M_2; M_1 — та часть ОДЗ, где применимо соответствующее утверждение (см. § 5) о равносильном преобразовании неравенства на множестве; M_2 — та часть ОДЗ, которая остается после выделения множества M_1.

3. Решить неравенство на M_1, учитывая, что на нем применимо соответствующее утверждение о равносильном преобразовании неравенства на множестве.

4. Найти свой способ решения неравенства на M_2.

5. Объединить решения, найденные на M_1 и на M_2.

Покажем теперь на конкретных примерах, как решать неравенства по этой схеме.

Решить неравенство

$$\sqrt{x+2} > x. \tag{4}$$

ОДЗ этого неравенства — все действительные $x \geqslant -2$. Разобьем ОДЗ на две части: $M_1 = [0; +\infty)$ и $M_2 = [-2; 0)$. Для любого $x \in M_1$ обе части неравенства (4) неотрицательны, значит, по утверждению 5 § 5 на этом множестве неравенство (4) равносильно неравенству

$$(\sqrt{x+2})^2 > x^2. \tag{5}$$

Для любого x из ОДЗ, а значит, и для любого x из его части M_1 $(\sqrt{x+2})^2 = x+2$ и поэтому неравенство (5) равносильно неравенству $x+2 > x^2$. Применяя утверждение 1 § 5, получаем, что на множестве M_1 исходное неравенство равносильно неравенству $x^2 - x - 2 < 0$. Решая это квадратное неравенство (см. § 2 гл. III), находим его решения $-1 < x < 2$. Из этих решений решениями исходного неравенства будут лишь те x, которые принадлежат M_1, т. е. $x \in [0, 2)$.

Рассмотрим решение неравенства (4) на множестве M_2. Для любого $x \in M_2$ левая часть неравенства (4) неотрицательна, а правая — отрицательна, значит, любое $x \in M_2$ является решением неравенства (4).

Объединяя решения, найденные на M_1 и M_2, получаем, что множеством всех решений исходного неравенства является промежуток $[-2; 2)$.

Решить неравенство

$$x+1 > \sqrt{x+3}. \tag{6}$$

ОДЗ этого неравенства — все действительные $x \geqslant -3$. Разобьем ОДЗ на две части: $M_1 = [-1, +\infty)$ и $M_2 = [-3, -1)$. Для любого $x \in M_1$ обе части неравенства (6) неотрицательны, поэтому по утверждению 5 § 5 на этом множестве неравенство (6) равносильно неравенству $(x+1)^2 > x+3$ (так как $(\sqrt{x+3})^2 = x+3$ на всей ОДЗ). Воспользовавшись утверждением 1 § 5, получаем, что на множестве M_1 исходное неравенство равносильно неравенству $x^2+x-2>0$, которое имеет решением два промежутка: $(-\infty;\ -2)$ и $(1;\ +\infty)$. Но из этих решений множеству M_1 принадлежат лишь $x\in(1;\ +\infty)$, которые и будут решениями исходного неравенства на множестве M_1.

Рассмотрим решение неравенства (6) на множестве M_2. Для любого $x\in M_2$ левая часть неравенства (6) отрицательна, а правая — неотрицательна. Значит ни для одного $x\in M_2$ неравенство (6) не выполняется, т. е. на множестве M_2 исходное неравенство решений не имеет. Итак, решением исходного неравенства является промежуток $(1;\ +\infty)$.

Аналогично решаются и более общие неравенства вида

$$\sqrt[2m]{f(x)}>\varphi(x), \tag{7}$$

$$\sqrt[2m]{f(x)}<\varphi(x). \tag{8}$$

А именно сначала ищется ОДЗ неравенства, затем ОДЗ делится на две части: M_1 — ту часть ОДЗ, где функция $\varphi(x)$ неотрицательна, и M_2 — ту часть ОДЗ, где функция $\varphi(x)$ отрицательна.

Тогда для неравенства (7) уже известна часть решений — множество M_2, а остальные решения ищутся как решения неравенства $f(x)>[\varphi(x)]^{2m}$, принадлежащие множеству M_1. Что же касается неравенства (8), то все его решения есть решения неравенства $f(x)<[\varphi(x)]^{2m}$, принадлежащие множеству M_1.

Решить неравенство

$$\frac{x^2-2x-1}{x+1}<x. \tag{9}$$

ОДЗ этого неравенства — все действительные x, кроме $x=-1$. Разобьем ОДЗ на две части: $M_1=(-1;\ +\infty)$ и $M_2=(-\infty;\ -1)$. Для любого $x\in M_1$ выражение $(x+1)$ положительно, а потому на множестве M_1 неравенство (9) по утверждению 7 § 5 равносильно неравенству

$$x^2-2x-1<x(x+1). \tag{10}$$

Применяя утверждение 1 § 5, получаем, что неравенство (10) равносильно неравенству $3x+1>0$, откуда $x>-1/3$. Поскольку все эти x входят в множество M_1, то все они являются решениями неравенства (9) на множестве M_1.

Для любого $x\in M_2$ выражение $(x+1)$ отрицательно, а потому на множестве M_2 неравенство (9) по утверждению 7 § 5 равносильно неравенству

$$x^2-2x-1>x(x+1). \tag{11}$$

Применяя утверждение 1 § 5, получаем, что неравенство (11) равносильно неравенству $3x+1<0$, откуда $x<-1/3$. Из найденных решений в множество M_2 входят $x<-1$, т. е. все

$x \in M_2$. Следовательно, решением исходного неравенства будут все $x \in M_2$.

Объединяя решения, найденные на M_1 и M_2, получаем, что исходное неравенство (9) справедливо для всех $x \in \in (-\infty;\ -1) \cup (-1/3;\ +\infty)$.

Аналогично решаются неравенства вида

$$\frac{f(x)}{g(x)} > \varphi(x). \tag{12}$$

А именно сначала ищется ОДЗ неравенства, затем ОДЗ делится на две части: M_1 — ту часть ОДЗ, где функция $y=g(x)$ положительна, и M_2 — ту часть ОДЗ, где функция $y=g(x)$ отрицательна.

Тогда на множестве M_1 неравенство (12) равносильно неравенству $f(x) > \varphi(x)g(x)$, а на множестве M_2 неравенство (12) равносильно неравенству $f(x) < \varphi(x)g(x)$. Далее решают эти неравенства на соответствующих множествах и объединяют получившиеся решения. Отметим, что если функции $y=f(x)$, $y=\varphi(x)$ и $y=g(x)$ будут многочленами, то решение неравенства (12) можно провести иначе, а именно неравенство (12) надо сначала переписать в равносильном виде

$$\frac{f(x)-\varphi(x)g(x)}{\varphi(x)} > 0. \tag{13}$$

Затем воспользоваться утверждением 7 § 5 и записать неравенство

$$\varphi(x)[f(x)-\varphi(x)g(x)] > 0, \tag{14}$$

равносильное неравенству (13) на его ОДЗ. Наконец, неравенство (14) решить методом интервалов (см. § 2 гл. III). Его решения и будут решениями неравенства (12).

Решить неравенство

$$\frac{x^2-2x-1}{x+1} < x. \tag{9}$$

Перенося x в левую часть неравенства, перепишем это неравенство в равносильном виде $\frac{-3x-1}{x+1} < 0$. Пользуясь утверждением 7 § 5 получаем неравенство $-(x+1)\times(3x+1) < 0$, равносильное неравенству (9) для всех x, кроме $x=-1$. Решая это неравенство методом интервалов, получаем ответ: неравенство (9) справедливо для всех x из промежутков $(-\infty;\ -1)$ и $(-1/3;\ +\infty)$.

Решить неравенство

$$\log_{x^2}(2+x) < 1. \tag{15}$$

ОДЗ этого неравенства определяется условиями $\begin{cases} 2+x>0 \\ x^2>0,\ x^2 \neq 1, \end{cases}$ т. е. ОДЗ этого неравенства состоит из четырех интервалов: $(-2;\ -1)$, $(-1;\ 0)$, $(0;\ 1)$, $(1;\ +\infty)$.

Разобьем ОДЗ на две части: $M_1=(-2;\ -1)\cup(1;\ +\infty)$ и $M_2=(-1;\ 0)\cup(0;\ 1)$. Для любого $x\in M_1$ получаем, что $x^2>1$, а потому на множестве M_1 неравенство (15) по утверждению 6а § 5 равносильно неравенству

$$2+x<x^2. \qquad (16)$$

Решая квадратное неравенство (16), находим, что оно справедливо для всех x из промежутков $(-\infty;\ -1)$ и $(2;\ +\infty)$. Из этих x в множество M_1 входят $x\in(-2;\ -1)\cup(2;\ +\infty)$. Они и дают решение неравенства (15) на множестве M_1.

Для любого $x\in M_2$ получаем, что $0<x^2<1$, а потому на множестве M_2 неравенство (15) по утверждению 6*б* § 5 равносильно неравенству

$$2+x>x^2, \qquad (17)$$

которое имеет решение $x\in(-1;\ 2)$. Из этих x в множество M_2 входят $x\in(-1;\ 0)\cup(0;\ 1)$, они и дают решение неравенства (15) на множестве M_2.

Объединяя решения, найденные на множестве M_1 и M_2, получаем, что исходное неравенство имеет решением все

$$x\in(-2;\ -1)\cup(-1;\ 0)\cup(0;\ 1)\cup(2;\ +\infty).$$

Аналогично решаются неравенства вида

$$\log_{\varphi(x)} g(x)<b, \qquad (18)$$

где b — данное число. А именно сначала ищется ОДЗ неравенства. Затем ОДЗ делится на две части: M_1 — ту часть ОДЗ, где $\varphi(x)>1$, и M_2 — ту часть ОДЗ, где $0<\varphi(x)<1$. Тогда неравенство (18) на множестве M_1 равносильно неравенству

$$g(x)<[\varphi(x)]^b, \qquad (19)$$

а на множестве M_2 неравенство (18) равносильно неравенству

$$g(x)>[\varphi(x)]^b. \qquad (20)$$

Далее остается решить неравенства (19) и (20) на соответствующих множествах и объединить получившиеся решения.

Решить неравенство

$$(x^2+x+1)^x<1. \qquad (21)$$

Для всех действительных x квадратный трехчлен x^2+x+1 положителен, поэтому ОДЗ неравенства (21) есть все действительные x. Логарифмируя неравенство (21) по любому основанию, например по основанию 10, получим на основании утверждения 6*а* § 5, что неравенство (21) равносильно неравенству $\lg(x^2+x+1)^x<\lg 1$. Используя свойства логарифмов, перепишем это неравенство в виде $x\lg(x^2+x+1)<0$. Это неравенство равносильно совокупности двух систем неравенств:

$$\begin{cases} x > 0, \\ \lg(x^2+x+1) < 0, \end{cases} \quad \begin{cases} x < 0, \\ \lg(x^2+x+1) > 0, \end{cases}$$

которая, в свою очередь, равносильна следующей совокупности двух систем неравенств:

$$\begin{cases} x > 0, \\ x^2+x+1 < 1, \end{cases} \quad \begin{cases} x < 0, \\ x^2+x+1 > 1. \end{cases}$$

Решая сначала первую систему, приходим к выводу, что она не имеет решений, так как ее решение есть пересечение двух множеств: $(0; +\infty)$ — решения первого неравенства и $(-1; 0)$ — решения второго неравенства, а это пересечение пусто.

Вторая система имеет решение $x \in (-\infty; -1)$. Следовательно, исходное неравенство имеет решение $x \in (-\infty; -1)$.

Аналогично решаются неравенства вида

$$[f(x)]^{\varphi(x)} < b, \tag{22}$$

где b — данное положительное число. А именно сначала ищут ОДЗ неравенства, затем выбирают любое число $a > 1$. Тогда на ОДЗ неравенство (22) равносильно неравенству

$$\varphi(x) \log_a f(x) < \log_a b.$$

Далее остается решить последнее неравенство на ОДЗ исходного неравенства (22).

Решить неравенство $4\cos^2 2x + 8\cos 2x - 5 > 0$. (23)
Для решения этого неравенства рассмотрим сначала квадратное неравенство $4v^2 + 8v - 5 < 0$. Оно имеет решение: все $v \in (-5/2; 1/2)$. Следовательно, неравенство (23) равносильно системе неравенств

$$\begin{cases} \cos 2x > -5/2, \\ \cos 2x < 1/2. \end{cases}$$

Первое неравенство в этой системе выполняется при любом действительном значении x. Решение второго неравенства $x \in (\pi k + \pi/6; \pi/3 + \pi k)$, где $k = 0, \pm 1, \pm 2, \pm \ldots$. Следовательно, неравенство (23) имеет решение $\pi/6 + \pi k < x < \pi/3 + \pi k$, где $k = 0, \pm 1, \pm 2, \pm \ldots$.

Аналогично решается неравенство вида

$$P[g(x)] > 0, \tag{24}$$

где $P(g)$ — квадратный трехчлен: $P(g) = ag^2 + bg + c$, а $g(x)$ — основная элементарная функция. Отметим, что в зависимости от дискриминанта и коэффициента a этого квадратного трехчлена возможны следующие четыре случая:

а) неравенство $ag^2 + bg + c > 0$ не имеет решений (если $a < 0$, $D = b^2 - 4ac \leqslant 0$). Тогда и неравенство (24) не имеет решений;

б) решением неравенства $ag^2 + bg + c > 0$ является интервал

(g_1, g_2) (если $a<0$, $D=b^2-4ac>0$). Тогда неравенство (24) равносильно системе неравенств

$$\begin{cases} g(x)<g_2, \\ g(x)>g_1; \end{cases}$$

в) решением неравенства $ag^2+bg+c>0$ является объединение лучей $(-\infty, g_1)$ и $(g_2, +\infty)$ (если $a>0$, $D=b^2-4ac\geqslant 0$). Тогда неравенство (24) равносильно совокупности неравенств

$$g(x)<g_1, \ g(x)>g_2;$$

г) решением неравенства $ag^2+bg+c>0$ является любое g, (если $a>0$, $D=b^2-4ac<0$). Тогда неравенство (24) имеет решением всю ОДЗ уравнения, т. е. все те x, при каждом из которых функция $g(x)$ существует.

Равносильный переход к системе или совокупности неравенств возможен и в более общем случае, чем рассмотренное неравенство (24). Например, пусть требуется решить неравенство

$$\operatorname{tg}[1/(1+x^2)]>1. \qquad (25)$$

Здесь имеется суперпозиция двух функций $y(g)=\operatorname{tg} g$ и $g(x)=$ $=1/(1+x^2)$. Решив простейшее неравенство $\operatorname{tg} g>1$, получим, что его решением являются все $g\in(\pi k+\pi/4, \ \pi/2+\pi k)$, где $k=0, \pm 1, \pm 2, \pm\ldots$, или, подставляя вместо g его значения, получим равносильную бесконечную совокупность систем неравенств

$$\begin{cases} 1/(1+x^2)<\pi/2+\pi k, \\ 1/(1+x^2)>\pi/4+\pi k, \end{cases} \qquad (26)$$

где $k=0, \pm 1, \pm 2, \pm\ldots$. Рассмотрим все системы в этой бесконечной совокупности. При $k=0$ имеем систему

$$\begin{cases} \dfrac{1}{1+x^2}<\pi/2, \\ \dfrac{1}{1+x^2}>\pi/4. \end{cases} \qquad (27)$$

Первое неравенство в системе (27) выполняется при любом действительном значении x, так как $\frac{1}{1+x^2}\leqslant 1<\pi/2$. Второе неравенство в системе (27) имеет решение все $x\in(-\sqrt{4/\pi-1}; \sqrt{4/\pi-1})$. Таким образом, все эти x и дают решения системы (27).

Рассмотрим теперь системы (26) при положительных значениях k. Ни одна из этих систем не имеет решения, так как $\pi/4+\pi k>1$ при любом натуральном k, а $\frac{1}{1+x^2}\leqslant 1$ и поэтому второе неравенство в системе (26) не имеет решений. При отрицательных k также ни одна из систем не имеет решения, так как $\frac{1}{1+x^2}>0$ при любом x, а $\pi/2+\pi k<0$ при любом отрицательном k.

Окончательно получаем, что неравенство (25) имеет решением всех x из промежутка $(-\sqrt{4/\pi-1};\ \sqrt{4/\pi-1})$.

В заключение рассмотрим решение неравенств, в которых некоторые функции, участвующие в неравенстве, стоят под знаком абсолютной величины. Для решения таких неравенств обычно употребляется метод интервалов. Описание этого метода для неравенств практически полностью повторяет описание этого метода для уравнений (см. § 3) с заменой слова «уравнение» на «неравенство». Продемонстрируем применение метода интервалов на примере решения неравенства

$$x^2+|x+1|-3>0. \tag{28}$$

В этом неравенстве лишь одна функция $y=x+1$ находится под знаком абсолютной величины. Эта функция обращается в нуль в точке $x=-1$, поэтому вся числовая ось разбивается на два промежутка $(-\infty,\ -1)$ и $[-1,\ +\infty)$. Рассмотрим решение неравенства (28) на каждом из промежутков. Пусть $x\in(-\infty,\ -1)$, тогда неравенство (28) равносильно на этом множестве неравенству $x^2-(x+1)-3>0$ или $x^2-x-4>0$. Решая это квадратное неравенство, получаем, что оно справедливо для всех $x\in[-\infty,\ (1-\sqrt{17})/2]\cup[(1+\sqrt{17})/2,\ +\infty]$. Из этих решений в промежуток $(-\infty,\ -1)$ входят лишь все $x\in[-\infty,\ (1-\sqrt{17})/2]$, которые являются решениями неравенства (28) на рассматриваемом промежутке.

Пусть $x\in[-1,\ +\infty)$. Тогда неравенство (28) равносильно на этом множестве неравенству $x^2+(x+1)-3>0$, или $x^2+x-2>0$. Решением этого квадратного неравенства является множество $(-\infty;\ -2)\cup(1;\ +\infty)$. Из этого множества в рассматриваемый промежуток входят лишь $x\in(1;\ +\infty)$, которые являются решениями неравенства (28) на рассматриваемом промежутке. Объединяя решения на двух рассматриваемых промежутках, находим, что решением неравенства (28) является множество всех $x\in[-\infty;\ (1-\sqrt{17})/2]\cup(1;\ +\infty)$.

Глава VIII

ПРЕДЕЛ ПОСЛЕДОВАТЕЛЬНОСТИ И ПРЕДЕЛ ФУНКЦИИ

§ 1

ПРЕДЕЛ ЧИСЛОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

Если каждому натуральному числу n поставлено в соответствие число x_n, то говорят, что дана *числовая последовательность* $x_1, x_2, \ldots, x_n, \ldots$, или, короче, последовательность $\{x_n\}$. Чтобы задать числовую последовательность, надо задать закон (правило), по которому каждому натуральному числу ставится в соответствие некоторое число, т. е. каждая числовая последовательность может рассматриваться как функция, область определения которой — множество всех натуральных чисел.

Если функция $y=f(x)$ такова, что множество всех натуральных чисел содержится в области ее существования, тогда с помощью функции $y=f(x)$ с областью определения — множеством всех натуральных чисел — можно задать числовую последовательность $f(1), f(2), \ldots, f(n), \ldots$, или $\{f(n)\}$. Например, множество всех натуральных чисел содержится в области существования каждой из следующих функций:

1. $y=x$;
2. $y=2^{1-x}$;
3. $y=-2-3(x-1)$;
4. $y=\operatorname{tg}(-\pi/4+\pi x/2)$;
5. $y=1^x$;
6. $y=(x+1)/x$;
7. $y=\{1+\sin(\pi/2+\pi x)\}/x$;
8. $y=1/x$;
9. $y=2^{x-1}$;
10. $y=x^2$.

На области определения — множестве всех натуральных чисел — эти функции задают соответственно следующие числовые последовательности:

1. $1, 2, 3, \ldots, n, \ldots$;
2. $1, 1/2, 1/4, \ldots, (1/2)^{n-1} \ldots$;
3. $-2, -5, -8, \ldots, -2-3(n-1), \ldots$;
4. $1, -1, +1, \ldots, (-1)^{n-1}, \ldots$;
5. $1, 1, 1, \ldots, (1)^n, \ldots$;
6. $2, 3/2, 4/3, \ldots, (n+1)/n, \ldots$;
7. $0, 1, 0, 1/2, \ldots, [1+(-1)^n]/n, \ldots$;
8. $1, 1/2, 1/3, \ldots, 1/n, \ldots$;
9. $1, 2, 4, \ldots, 2^{n-1}, \ldots$;
10. $1, 4, 9, \ldots, n^2, \ldots$.

Наиболее удобно задавать числовую последовательность a_1, a_2, a_3, ..., a_n, ... при помощи формулы для общего члена. Запишем, например, формулы общего члена в примерах 1 — 10:

1. $a_n = n$;
2. $a_n = 1/2^{n-1}$;
3. $a_n = -2 - 3(n-1)$;
4. $a_n = (-1)^{n-1}$;
5. $a_n = 1^n$;
6. $a_n = (n+1)/n$;
7. $a_n = [1+(-1)^n]/n$;
8. $a_n = 1/n$;
9. $a_n = 2^{n-1}$;
10. $a_n = n^2$.

Нередко последовательность задают рекуррентным соотношением, т. е. формулой, выражающей a_n через члены последовательности с меньшим номером, например:

11. Последовательность чисел Фибоначчи: 1, 1, 2, 3, 5, 8, ..., a_n, ... задается формулой $a_n = a_{n-1} + a_{n-2}$ для $n > 2$ и $a_1 = a_2 = 1$;

12. Арифметическая прогрессия с первым членом a_1 и разностью d определяется рекуррентным соотношением $a_n = a_{n-1} + d$ $(n \geqslant 2)$;

13. Геометрическая прогрессия с первым членом u_1 и знаменателем q задается рекуррентным соотношением $u_n = u_{n-1}q$ $(n \geqslant 2)$.

Последовательности могут задаваться и другими способами.

14. Последовательность десятичных приближений числа π с недостатком 3; 3,1; 3,14; 3,141; ... .

Поскольку всякая числовая последовательность может рассматриваться как функция натурального аргумента, то на числовые последовательности переносятся свойства монотонности и ограниченности функций.

Числовая последовательность $\{a_n\}$ называется *возрастающей*, если для любых натуральных чисел n_1 и n_2 из условия, что $n_1 < n_2$, следует, что $a_{n_1} < a_{n_2}$. Возрастающими являются, например, последовательности 1, 9, 10, 14.

Числовая последовательность $\{a_n\}$ называется *неубывающей*, если для любых натуральных чисел n_1 и n_2 из условия, что $n_1 < n_2$ следует, что $a_{n_1} \leqslant a_{n_2}$. Неубывающими являются, например, последовательности 5 и 11.

Числовая последовательность $\{a_n\}$ называется *убывающей*, если для любых натуральных чисел n_1 и n_2 из условия, что $n_1 < n_2$ следует, что $a_{n_1} > a_{n_2}$. Убывающими последовательностями являются, например, последовательности 2, 3, 6, 8.

Числовая последовательность $\{a_n\}$ называется *невозрастающей*, если для любых натуральных чисел n_1 и n_2 из условия, что $n_1 < n_2$ следует, что $a_{n_1} \geqslant a_{n_2}$. Например, последовательность $\{a_n\}$, где $a_n = 1/[\sqrt{n}]$ ($[a]$ — целая часть числа a, т. е. наибольшее целое число, не превосходящее a), является невозрастающей, так как $a_1 = 1$, $a_2 = 1$, $a_3 = 1$, $a_4 = 1/2$ и т. д.

Числовая последовательность называется *монотонной*, если она или убывает, или возрастает, или не убывает, или не возрастает. Монотонными являются все вышеприведенные последовательности, кроме 4, 7. Что касается геометрической прогрессии, то до задания чисел u_1 и q о ней нельзя сказать, монотонна она или нет.

Последовательность $\{a_n\}$ называется *ограниченной сверху*, если существует число B, такое, что для любого натурального числа n справедливо неравенство $a_n \leqslant B$. Ограниченными сверху являются, например, последовательности 2, 3, 4, 5, 6, 7, 8, 14 и последовательность $\{a_n\}$, где $a_n = \frac{1}{[\sqrt{n}]}$.

Последовательность $\{a_n\}$ называется *ограниченной снизу*, если существует число A, такое, что для любого натурального числа n справедливо неравенство $a_n \geqslant A$. Ограниченными снизу являются, например, последовательности 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 14 и последовательность $\{a_n\}$, где $a_n = 1/[\sqrt{n}]$.

Последовательность $\{a_n\}$ называется *ограниченной*, если она ограничена и снизу и сверху. Ограниченными являются, например, последовательности 2, 4, 5, 6, 7, 8.

Число a называется *пределом числовой последовательности* $\{a_n\}$, если для любого положительного числа ε найдется номер N, такой, что для любого $n > N$ справедливо неравенство $|a_n - a| < \varepsilon$. Тот факт, что число a является пределом последовательности $\{a_n\}$ записывают так: $\lim\limits_{n \to +\infty} a_n = a$, или $a_n \to a$.

Примеры. 1. Доказать, что $\lim\limits_{n \to +\infty} a_n = 0$, где $a_n = q^n$ и $|q| < 1$, $\neq 0$.

Возьмем произвольное число $\varepsilon > 0$.

а) Если $\varepsilon < 1$, то положим $N = [\log_{|q|} \varepsilon] + 1$. Ясно, что $N > \log_{|q|} \varepsilon$ и для любого $n > N$: $|q^n - 0| = |q|^n < |q|^N < |q|^{\log_{|q|} \varepsilon} = \varepsilon$.

б) Если $\varepsilon \geqslant 1$, то положим $N = 1$. Легко видеть, что для любого $n > 1$: $|q^n - 0| = |q|^n < 1 \leqslant \varepsilon$.

Итак, для любого $\varepsilon > 0$ найдется номер N, такой, что для любого $n > N$ справедливо неравенство $|q^n - 0| < \varepsilon$, т. е. $\lim\limits_{n \to +\infty} a_n = 0$.

2. Доказать, что $\lim\limits_{n \to +\infty} a_n = 1$, где $a_n = (n+1)/n$.

Возьмем произвольное число $\varepsilon > 0$. Положим $N = [1/\varepsilon] + 1$, тогда $N > 1/\varepsilon$ и для любого $n > N$: $\left|\frac{n+1}{n} - 1\right| = \frac{1}{n} < \frac{1}{N} < \frac{1}{1/\varepsilon} = \varepsilon$, что и требовалось доказать.

3. Доказать, что $\lim\limits_{n \to +\infty} a_n = 0$, где $a_n = 1/n$.

Возьмем произвольное число $\varepsilon > 0$. Положим $N = [1/\varepsilon] + 1$. Тогда для любого $n > N$ $\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \frac{1}{N} < \frac{1}{1/\varepsilon} = \varepsilon$, что и требовалось доказать.

Будем говорить, что последовательность $\{a_n\}$ имеет *пределом* $+\infty$ и писать $\lim\limits_{n\to+\infty} a_n = +\infty$, или $a_n \to +\infty$, если для любого сколь угодно большого положительного числа A найдется номер N, такой, что для любого $n > N$ справедливо неравенство $a_n > A$.

Примеры. 1. Доказать, что $\lim\limits_{n\to+\infty} a_n = +\infty$, где $a_n = n$.

Для любого положительного числа A положим $N = [A+1]+1$. Тогда для любого $n > N$: $a_n = n > N = [A+1]+1 \geqslant A+1 > A$, т. е. $\lim\limits_{n\to+\infty} a_n = +\infty$.

2. Доказать, что $\lim\limits_{n\to+\infty} a_n = +\infty$, где $a_n = n^2$.

Для любого положительного числа A положим $N = [\sqrt{A+1}]+1$. Тогда для любого $n > N$ $a_n = n^2 > N^2 = ([\sqrt{A+1}]+1)^2 \geqslant (\sqrt{A+1})^2 = A+1 > A$, т. е. $\lim\limits_{n\to+\infty} a_n = +\infty$.

3. Доказать, что $\lim\limits_{n\to+\infty} a_n = +\infty$, где $a_n = 2^{n-1}$.

Для любого положительного числа A положим $N = [\log_2(A+1)]+2$. Тогда для любого $n > N$ $a_n = 2^{n-1} > 2^{N-1} = 2^{[\log_2(A+1)]+1} \geqslant 2^{\log_2(A+1)} = A+1 > A$, т. е. $\lim\limits_{n\to+\infty} a_n = +\infty$.

Будем говорить, что последовательность $\{a_n\}$ имеет своим *пределом* $-\infty$ и писать $\lim\limits_{n\to+\infty} a_n = -\infty$ или $a_n \to -\infty$, если для любого отрицательного числа B, такого, что $|B|$ — сколь угодно большое число, найдется номер N, такой, что для любого $n > N$ справедливо неравенство $a_n < B$.

Примеры. 1. Доказать, что $\lim\limits_{n\to+\infty} a_n = -\infty$, где $a_n = -2-(n-1)3$.

Для любого отрицательного числа B положим $N = \left[\frac{|B-1|}{3}\right]$. Тогда для любого $n > N$: $a_n = -2-(n-1)3 < -2-(N-1)3 = -2-3\left(\left[\frac{|B-1|}{3}\right]-1\right) < 2-3\left(\frac{|B-1|}{3}-1\right) = -2-3\left(\frac{1-B}{3}-1\right) = B$, что и требовалось доказать.

2. Доказать, что $\lim\limits_{n\to+\infty} a_n = -\infty$, где $a_n = -2^n$.

Для любого отрицательного числа B положим $N = [\log_2|B|+1]$. Тогда для любого $n > N$: $a_n = -2^n < -2^N \leqslant -2^{\log_2|B|} = -|B| = B$, что и требовалось доказать.

3. Доказать, что $\lim\limits_{n\to+\infty} a_n = -\infty$, где $a_n = \log_{1/2} n$.

Для любого отрицательного числа B положим $N = [2^{|B|}+1]$. Тогда для любого $n > N$: $a_n = \log_{1/2} n < \log_{1/2} N = \log_{1/2} 2^{|B|} = -|B| = B$, что и требовалось доказать.

Отметим, что *существуют последовательности, не имеющие предела*. Такой последовательностью является, например, последовательность $\{a_n\}$, где $a_n = (-1)^n$.

§ 2
ТЕОРЕМЫ О ПРЕДЕЛАХ ЧИСЛОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

В этом параграфе будем рассматривать только те последовательности, которые имеют конечный предел.

Теорема 1. Если последовательность $\{a_n\}$ имеет предел A и число $p<A$, то найдется номер N, такой, что для любого $n>N$ справедливо неравенство $p<na_n$.

Доказательство. Поскольку A есть предел последовательности $\{a_n\}$, то для любого $\varepsilon>0$ найдется номер N, такой, что для любого $n>N$ справедливо неравенство $|a_n-A|<\varepsilon$. Перепишем это неравенство в форме

$$A-\varepsilon<a_n<A+\varepsilon. \tag{1}$$

Положим $\varepsilon=A-p>0$. Для этого $\varepsilon>0$ найдется номер N, такой, что для любого $n>N$ будет справедливо неравенство (1). Подставляя в левую часть неравенства (1) $\varepsilon=A-p$, получаем, что $p<a_n$, что и требовалось доказать.

Теорема 2. Если последовательность $\{a_n\}$ имеет предел A и число $q>A$, то найдется номер N, такой, что для любого $n>N$ справедливо неравенство $q>a_n$.

Доказательство теоремы 2 аналогично доказательству теоремы 1.

Замечание 1. Если $\lim\limits_{n\to+\infty} a_n=A$ и $A>0$, то найдется номер N, такой, что $a_n>0$ для любого $n>N$.

2. Если $\lim\limits_{n\to+\infty} a_n=A$ и $A<0$, то найдется номер N, такой, что $a_n<0$ для любого $n>N$.

Теорема 3. Если последовательность $\{a_n\}$ имеет предел A, то найдется положительное число M, такое, что $|a_n|\leqslant M$, т. е. последовательность, имеющая предел, ограничена.

Доказательство. Выберем число $M_1>0$ так, чтобы оно было больше $|A|$, т. е. чтобы было справедливо неравенство $-M_1<A<M_1$. Обозначив $p=-M_1$, $q=M_1$, получим, что $A>p$ и $A<q$. По теореме 1 существует номер N_1, такой, что для любого $n>N_1$ справедливо неравенство $p<a_n$. По теореме 2 существует номер N_2, такой, что для любого $n>N_2$ справедливо неравенство $q>a_n$. Выберем номер $N=\max\{N_1, N_2\}$. Тогда для любого $n>N$ одновременно справедливо двойное неравенство $p<a_n<q$ или $-M_1<a_n<M_1$, т. е. $|a_n|<M_1$. Неравенство $|a_n|<M_1$ выполняется для любого $n>N$, т. е. для $n=N+1$, $n=N+2$, $n=N+3$, $n=N+4$ и так далее. Значит, неравенство $|a_n|<M_1$ может не выполняться лишь для первых N членов последовательности.

Выберем среди чисел $|a_1|$, $|a_2|$, $|a_3|$, $\ldots$, $|a_{N-1}|$, $|a_N|$, M_1 наибольшее число и обозначим его через M. Тогда ясно, что для любого n будет справедливо неравенство $|a_n| \leqslant M$, что и требовалось доказать.

Теорема 4. Если последовательность $\{a_n\}$ имеет предел, то этот предел единственный.

Доказательство. Предположим противное, т. е. пусть одновременно $a_n \to A$ и $a_n \to B$ и пусть $A < B$. Возьмем любое число C между A и B, т. е. $A < C < B$. Поскольку $a_n \to A$ и $A < C$, то по теореме 2 найдется такой номер N_1, что для любого $n > N_1$ будет выполняться неравенство $a_n < C$. Поскольку $a_n \to B$ и $C < B$, то по теореме 1 найдется такой номер N_2, что для любого $n > N_2$ будет выполняться неравенство $a_n > C$. Выберем номер $N = \max\{N_1, N_2\}$. Тогда член последовательности a_N одновременно удовлетворяет двум неравенствам: $a_N > C$ и $a_N < C$, что невозможно. Следовательно, предположение неверно, а верно утверждение теоремы 4.

Теорема 5. Если $\lim\limits_{n \to +\infty} a_n = A$, то $\lim\limits_{n \to +\infty} |a_n| = |A|$.

Доказательство. Пусть $A = 0$. Тогда условие $\lim\limits_{n \to +\infty} a_n = 0$ означает, что для любого $\varepsilon > 0$ найдется номер N, такой, что для любого $n > N$ справедливо неравенство $|a_n - 0| < \varepsilon$. Поскольку это неравенство можно переписать так: $||a_n| - 0| < \varepsilon$, то получаем, что $\lim\limits_{n \to +\infty} |a_n| = 0$.

Пусть $A \neq 0$. Тогда условие $\lim\limits_{n \to +\infty} a_n = A$ означает, что для любого $\varepsilon > 0$ найдется номер N_1, такой, что для любого $n > N_1$ справедливо неравенство $|a_n - A| < \varepsilon$.

Если $A > 0$, то по замечанию 1 к теореме 2 найдется номер N_2, такой, что $a_n > 0$. Взяв номер $N = \max\{N_1, N_2\}$, получим, что для любого $n > N$ справедливо неравенство $||a_n| - |A|| < \varepsilon$, что и означает, что $\lim |a_n| = |A|$. Если же $A < 0$, то по замечанию 2 к теореме 2 найдется номер N_3, такой, что $a_n < 0$. Взяв номер $N = \max\{N_1, N_3\}$ и учитывая, что $a_n = -|a_n|$, $A = -|A|$, получим, что для любого $n > N$ справедливо неравенство $||a_n| - |-A|| < \varepsilon$, которое можно записать $|-|a_n| - (-|A|)| < \varepsilon$, т. е. $\lim\limits_{n \to +\infty} |a_n| = |A|$.

Арифметические операции над последовательностями $\{a_n\}$ и $\{b_n\}$ и *сравнение последовательностей* $\{a_n\}$ и $\{b_n\}$ производятся так же, как и над функциями, т. е. при одинаковых значениях аргумента, другими словами, почленно.

Теорема 6. Если последовательности $\{a_n\}$ и $\{b_n\}$ таковы, что $a_n = b_n$ для любого n и $a_n \to A$, а $b_n \to B$, то $A = B$, т. е. если $a_n = b_n$ для любого n, то $\lim\limits_{n \to +\infty} a_n = \lim\limits_{n \to +\infty} b_n$.

Доказательство. Условие $a_n = b_n$ для любого n означает, что на самом деле у нас одна последовательность, а по теореме 4 она не может иметь двух пределов, значит, $A = B$ и теорема 6 доказана.

Теорема 7. Если последовательности $\{a_n\}$ и $\{b_n\}$ таковы, что $a_n \geqslant b_n$ для любого n и $a_n \to A$, а $b_n \to B$, то $A \geqslant B$, т. е. если $a_n \geqslant b_n$ для любого n, то $\lim\limits_{n \to +\infty} a_n \geqslant \lim\limits_{n \to +\infty} b_n$.

Доказательство. Предположим противное. Пусть $A < B$. Выберем число C, такое, что $A < C < B$. Так как $a_n \to A$ и $A < C$, то по теореме 2 найдется номер N_1, такой, что для любого $n > N_1$ справедливо неравенство $a_n < C$. Так как $b_n \to B$ и $B > C$, то по теореме 1 найдется номер N_2, такой, что для любого $n > N_2$ справедливо неравенство $b_n > C$. Выберем номер $N = \max\{N_1, N_2\}$. Тогда для любого $n > N$ будут одновременно выполняться неравенства $a_n < C$ и $C < b_n$. Из справедливости этих неравенств следует, что $a_n < b_n$, что противоречит условию теоремы. Значит, наше предположение неверное и теорема 7 справедлива.

Замечание. Теорему 7 нельзя усилить, т. е. из строгого неравенства $a_n > b_n$ для членов последовательностей не обязательно вытекает строгое неравенство для пределов, например, если $a_n = {}^1/_{2^n}$, $b_n = -{}^1/_{3^n}$, то $a_n > b_n$ для любого n, но $\lim\limits_{n \to +\infty} a_n = \lim\limits_{n \to +\infty} b_n = 0$.

Теорема 8. Если последовательности $\{a_n\}$, $\{b_n\}$ и $\{c_n\}$ таковы, что $a_n \leqslant b_n \leqslant c_n$ для любого n и $a_n \to a$, и $c_n \to a$, то $b_n \to a$.

Доказательство. Возьмем произвольное $\varepsilon > 0$. Поскольку $a_n \to a$, то найдется номер N_1, такой, что для любого $n > N_1$ справедливо неравенство $a - \varepsilon < a_n < a + \varepsilon$. Поскольку $c_n \to a$, то найдется номер N_2, такой, что для любого $n > N_2$ справедливо неравенство $a - \varepsilon < c_n < a + \varepsilon$. Выберем номер $N = \max\{N_1, N_2\}$. Тогда для любого $n > N$ будут одновременно выполняться неравенства $a - \varepsilon < a_n < a + \varepsilon$, $a - \varepsilon < c_n < a + \varepsilon$. Добавим к этим неравенствам неравенство $a_n \leqslant b_n \leqslant c_n$, верное для любого n. По свойству транзитности неравенств для любого $n > N$ будет справедливо неравенство $a - \varepsilon < b_n < a + \varepsilon$, которое можно записать $|b_n - a| < \varepsilon$. Итак, для произвольно выбранного $\varepsilon > 0$ найдем номер N, такой, что для любого $n > N$ справедливо неравенство $|b_n - a| < \varepsilon$, т. е. $\lim\limits_{n \to +\infty} b_n = a$. Теорема доказана.

Замечание. Если $a \leqslant b_n \leqslant c_n$ для любого n и $c_n \to a$, то $b_n \to a$.

Теорема 9. Если последовательности $\{a_n\}$ и $\{b_n\}$ таковы, что $\lim\limits_{n \to +\infty} a_n = A$ и $\lim\limits_{n \to +\infty} b_n = B$, то последовательность $\{a_n + b_n\}$ такова, что $\lim\limits_{n \to +\infty} (a_n + b_n) = A + B$, т. е. если $a_n \to A$, $b_n \to B$, то $a_n + b_n \to A + B$.

Доказательство. Поскольку $\lim\limits_{n \to +\infty} a_n = A$ и $\lim\limits_{n \to +\infty} b_n = B$, то для любого $\varepsilon_1 > 0$ найдется номер N_1, такой, что для любого $n > N_1$ справедливо неравенство $A - \varepsilon_1 < a_n < A + \varepsilon_1$ и найдется номер N_2, такой, что для любого $n > N_2$ справедливо неравенство $B-\varepsilon_1 < b_n < B+\varepsilon_1$. Выберем номер $N=\max\{N_1, N_2\}$. Тогда для любого $\varepsilon_1 > 0$ найдется номер N, такой, что для любого $n > N$ одновременно справедливы неравенства $A-\varepsilon_1 < a_n < A+\varepsilon_1$ и $B-\varepsilon_1 < b_n < B+\varepsilon_1$. Тогда на основании свойств неравенств справедливо неравенство $A+B-2\varepsilon_1 < a_n+b_n < A+B+2\varepsilon_1$. Возьмем произвольное $\varepsilon > 0$ и обозначим $\varepsilon_1=\varepsilon/2$. Тогда, как следует из предыдущего, найдется номер N, такой, что для любого $n > N$ будет справедливо неравенство $(A+B)-\varepsilon < a_n+b_n < (A+B)+\varepsilon$, т. е. по определению предела $\lim\limits_{n\to+\infty}(a_n+b_n)=A+B$. Теорема доказана.

Теорема 10. Если $a_n \to A$, $b_n \to B$, то $a_n-b_n \to A-B$.

Доказательство этой теоремы аналогично доказательству теоремы 9 и поэтому опускается.

Теорема 11. Если $a_n \to A$, $b_n \to B$, то $a_nb_n \to AB$.

Доказательство. Поскольку $\lim\limits_{n\to+\infty} a_n=A$, то для любого наперед заданного числа $\varepsilon_1 > 0$ найдется номер N_1, такой, что для любого $n > N_1$ будет справедливо неравенство

$$|a_n-A| < \varepsilon_1. \tag{2}$$

Поскольку $\lim\limits_{n\to+\infty} b_n=B$, то для любого наперед заданного числа $\varepsilon_2 > 0$ найдется номер N_2, такой, что для любого $n > N_2$ будет справедливо неравенство

$$|b_n-B| < \varepsilon_2. \tag{3}$$

Рассмотрим неравенство

$$|a_nb_n-AB|=|a_nb_n-a_nB+a_nB-AB| \leqslant \leqslant |a_n|\,|b_n-B|+|B|\,|a_n-A|. \tag{4}$$

Оно справедливо для любого n. Поскольку по теореме 3 существует положительное число M, такое, что $|a_n| \leqslant M$ для любого n, то, обозначая $d=|B|+1 > 0$, из неравенства (4) получаем неравенство

$$|a_nb_n-AB| \leqslant M|b_n-B|+d|a_n-A|, \tag{5}$$

справедливое для любого n. Теперь берем произвольное $\varepsilon > 0$ и обозначим $\varepsilon_1=\varepsilon/2d$ и $\varepsilon_2=\varepsilon/2M$. Тогда для выбранного ε_1 существует номер N_1, такой, что для любого $n > N_1$ справедливо неравенство (2), а для выбранного ε_2 существует номер N_2 такой, что для любого $n > N_2$ справедливо неравенство (3). Выберем номер $N=\max\{N_1, N_2\}$. Тогда для любого $n > N$ одновременно справедливы неравенства (2), (3) и (5). Подставляя в (5) оценки (2) и (3), получаем, что для любого $n > N$ справедливо неравенство

$$|a_nb_n-AB| < \varepsilon. \tag{6}$$

Итак, для произвольного $\varepsilon > 0$ найден номер N, такой, что для любого $n > N$ справедливо неравенство (6) т. е.

по определению предела $\lim\limits_{n\to+\infty} a_nb_n=AB=\lim\limits_{n\to+\infty} a_n \lim\limits_{n\to+\infty} b_n$. Теорема доказана.

Теорема 12. Если последовательность $\{a_n\}$ имеет предел $A\neq 0$, то найдется номер N, такой, что для любого $n>N$ справедливо неравенство $|a_n|>|A|/2$.

Доказательство. Возьмем $\varepsilon=|A|/2$. Так как последовательность $\{a_n\}$ имеет предел, то по определению предела для этого $\varepsilon>0$ найдется номер N, такой, что для любого $n>N$ справедливо неравенство $|a_n-A|<|A|/2$.

Поскольку $|a_n-A|\geqslant|A|-|a_n|$, то получим, что для любого $n>N$ справедливо неравенство $|A|-|a_n|<|A|/2$, откуда $|A|/2<|a_n|$. Теорема доказана.

Теорема 13. Если $a_n\to A$, $b_n\to B$, где $B\neq 0$ для любого n, то $a_n/b_n\to A/B$.

Доказательство. Поскольку $\lim\limits_{n\to+\infty} a_n=A$, то для любого $\varepsilon_1>0$ найдется номер N_1, такой, что для любого $n>N_1$ справедливо неравенство

$$|a_n-A|<\varepsilon_1. \tag{7}$$

Поскольку $\lim\limits_{n\to+\infty} b_n=B$, то для любого $\varepsilon_2>0$ найдется номер N_2, такой, что для любого $n>N_2$ справедливо неравенство

$$|b_n-B|<\varepsilon_2. \tag{8}$$

По теореме 2 существует положительное число M, такое, что для любого n

$$|a_n|\leqslant M. \tag{9}$$

По теореме 12 существует номер N_3, такой, что для любого $n>N_3$ справедливо неравенство

$$|b_n|>|B|/2. \tag{10}$$

Ясно, что для любого n справедливо неравенство

$$\left|\frac{a_n}{b_n}-\frac{A}{B}\right|=\frac{|a_nB-b_nA|}{|b_nB|}=\frac{|a_nB-a_nb_n+a_nb_n-Ab_n|}{|a_n||B|}\leqslant$$
$$\leqslant\frac{|a_n||b_n-B|}{|b_n||B|}+\frac{|a_n-A|}{|B|}. \tag{11}$$

Возьмем теперь произвольное $\varepsilon>0$ и обозначим $\varepsilon_1=\varepsilon|B|/2$ и $\varepsilon_2=\varepsilon|B|^2/4M$. Тогда для выбранного ε_1 существует номер N_1, такой, что для любого $n>N_1$ справедливо неравенство (7), а для выбранного ε_2 существует номер N_2, такой, что для любого $n>N_2$ справедливо неравенство (8). Выберем номер $N=\max\{N_1, N_2, N_3\}$. Тогда для любого $n>N$ одновременно справедливы неравенства (7), (8), (9), (10) и (11). Подставляя оценки (7), (8), (9) и (10) в правую часть неравенства (11), получаем, что для любого $n > N$ справедливо неравенство

$$\left|\frac{a_n}{b_n} - \frac{A}{B}\right| < \varepsilon. \tag{12}$$

Итак, для произвольного $\varepsilon > 0$ найден номер N, такой, что для любого $n > N$ справедливо неравенство (12), а это и означает справедливость теоремы 13.

Теорема 14. Если $a_n \to A$, $b > 0$, $b \neq 1$, то $b^{a_n} \to b^A$.

Доказательство. Поскольку $a_n \to A$, то для любого $\varepsilon_1 > 0$ найдется номер N, такой, что для любого $n > N$ справедливо неравенство

$$A - \varepsilon_1 < a_n < A + \varepsilon_1. \tag{13}$$

Пусть $b > 1$. Поскольку функция $y = b^x$ возрастающая, то из неравенства (13) вытекает, что $b^{A-\varepsilon_1} < b^{a_n} < b^{A+\varepsilon_1}$, т. е. для любого $\varepsilon_1 > 0$ найдется номер N, такой, что для любого $n > N$ справедливо неравенство $b^{A-\varepsilon_1} < b^{a_n} < {}^{A+\varepsilon_1}$. Возьмем произвольное положительное число ε и обозначим $\varepsilon_1 = \log_b(1 + \varepsilon/b^A)$. Тогда для выбранного ε_1 существует номер N, такой, что для любого $n > N$ справедливо неравенство $b^{A-\varepsilon_1} < b^{a_n} < b^{A+\varepsilon_1}$. Ясно, что $b^{A+\varepsilon_1} = b^A(1 + \varepsilon/b^A) = b^A + \varepsilon$, $b^{A-\varepsilon_1} = b^A(1+\varepsilon/b^A)^{-1} = b^{2A}/(b^A + \varepsilon) > (b^{2A} - \varepsilon^2)/(b^A + \varepsilon) = b^A - \varepsilon$. Итак, мы показали, что для любого положительного ε найдется номер N, такой, что для любого $n > N$ справедливо неравенство $b^A - \varepsilon < b^{a_n} < b^A + \varepsilon$, т. е. $\lim\limits_{n \to +\infty} b^{a_n} = b^A$.

В случае $0 < b < 1$ доказательство теоремы проводится аналогично.

Теорема 15. Возрастающая и ограниченная сверху последовательность $\{a_n\}$ имеет предел.

Теорема 16. Убывающая и ограниченная снизу последовательность $\{a_n\}$ имеет предел.

Доказательство этих теорем опустим.

§ 3
ПРИМЕНЕНИЕ ТЕОРЕМ О ПРЕДЕЛАХ ЧИСЛОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

1. Найти предел последовательности $\{a_n\}$, если $a_n = \dfrac{(n+1)(n+2)}{(n+3)(n+4)}$.

Перепишем формулу для общего члена в виде $a_n = \dfrac{(1+{}^1/_n)(1+{}^2/_n)}{(1+{}^3/_n)(1+{}^4/_n)}$. Применяя последовательно теоремы о пределе частного, произведения и суммы, получаем

$$\lim_{n\to+\infty} a_n=\lim_{n\to+\infty}\frac{(1+{}^1/_n)(1+{}^2/_n)}{(1+{}^3/_n)(1+{}^4/_n)}=\frac{\lim\limits_{n\to+\infty}(1+{}^1/_n)(1+{}^2/_n)}{\lim\limits_{n\to+\infty}(1+{}^3/_n)(1+{}^4/_n)}=$$

$$=\frac{\lim\limits_{n\to+\infty}(1+{}^1/_n)\lim\limits_{n\to+\infty}(1+{}^2/_n)}{\lim\limits_{n\to+\infty}(1+{}^3/_n)\lim\limits_{n\to+\infty}(1+{}^4/_n)}=$$

$$=\frac{\left[\lim\limits_{n\to+\infty}1+\lim\limits_{n\to+\infty}({}^1/_n)\right]\left[\lim\limits_{n\to+\infty}1+\lim\limits_{n\to+\infty}({}^2/_n)\right]}{\left[\lim\limits_{n\to+\infty}1+\lim\limits_{n\to+\infty}({}^3/_n)\right]\left[\lim\limits_{n\to+\infty}1+\lim\limits_{n\to+\infty}({}^4/_n)\right]}=$$

$$=\frac{(1+0)(1+0)}{(1+0)(1+0)}=1.$$

2. Найти предел последовательности $\{a_n\}$, если $a_n=$ $=\frac{a_1n^2+b_1n+c_1}{a_2n^2+b_2n+c_2}$ $(a_2\neq 0$ и $a_2n^2+b^2n+c_2\neq 0$ для любого $n)$.

Разделив числитель и знаменатель в выражении для a_n на n^2 и применив те же теоремы, что и в п. 1, получим

$$\lim_{n\to+\infty} a_n=\lim_{n\to+\infty}\frac{a_1+b_1\frac{1}{n}+c_1\frac{1}{n^2}}{a_2+b_2\frac{1}{n}+c_2\frac{1}{n^2}}=\frac{\lim\limits_{n\to+\infty}\left(a_1+b_1\frac{1}{n}+c_1\frac{1}{n^2}\right)}{\lim\limits_{n\to+\infty}\left(a_2+b_2\frac{1}{n}+c_2\frac{1}{n^2}\right)}=$$

$$=\frac{\lim\limits_{n\to+\infty}a_1+\lim\limits_{n\to+\infty}\frac{b_1}{n}+\lim\limits_{n\to+\infty}\frac{c_1}{n^2}}{\lim\limits_{n\to+\infty}a_2+\lim\limits_{n\to+\infty}\frac{b_2}{n}+\lim\limits_{n\to+\infty}\frac{c_2}{n^2}}=\frac{a_1+0+0}{a_2+0+0}=\frac{a_1}{a_2}.$$

3. Выяснить, есть ли предел у последовательности $\{a_n\}$, если $a_n=(1+{}^1/_n)^n$.

При изучении бинома Ньютона была показана справедливость неравенств $\left(1+\frac{1}{n}\right)^n<3$ и $\left(1+\frac{1}{n}\right)^n<\left(1+\frac{1}{n+1}\right)^{n+1}$, т. е. последовательность $\{a_n\}$ такова, что $a_n<3$ и $a_n<a_{n+1}$ для любого n. Условие $a_n<3$ означает, что последовательность $\{a_n\}$ ограничена сверху; из условия $a_n<a_{n+1}$ легко получить, что $a_{n_1}<a_{n_2}$ для любых $n_1<n_2$, т. е. последовательность $\{a_n\}$ возрастающая. По теореме 15 возрастающая и ограниченная сверху последовательность имеет предел. Этот предел принято обозначать буквой e. Итак, $e=\lim\limits_{n\to+\infty}(1+{}^1/_n)^n$. Определение числа e позволяет вычислить его приближенно, а в курсе математического анализа показывается, что это число иррациональное.

4. Выяснить, есть ли предел у последовательности $\{a_n\}$, которая задана рекуррентным соотношением $a_1=\sqrt{2}$, $a_n=\sqrt{2+a_{n-1}}$ для любого $n\geqslant 2$.

Прежде всего выясним, ограничена ли эта последовательность. Докажем, что $a_n < 2$. Доказательство проведем методом математической индукции.

а) Для $n=1$ неравенство $a_1 < 2$ справедливо, поскольку $a_1 = \sqrt{2}$.

б) Предположим, что неравенство справедливо для $n=k$, т. е. $a_k < 2$.

в) Покажем, что из этого вытекает справедливость неравенства для $n=k+1$. Действительно, $a_{k+1} = \sqrt{2+a_k} < \sqrt{2+2} = 2$. Итак, неравенство $a_n < 2$ справедливо для любого n. Покажем, что последовательность $\{a_n\}$ возрастающая. Для этого достаточно показать, что $a_n < a_{n+1}$, т. е. доказать неравенство $a_n < \sqrt{2+a_n}$.

Это неравенство равносильно неравенству $a_n^2 < 2 + a_n$, которое можно переписать так: $(a_n+1)(a_n-2) < 0$. Поскольку $0 < a_n < 2$, то последнее неравенство очевидно, а значит справедливо и равносильное ему неравенство $a_n < a_{n+1}$. Итак, последовательность $\{a_n\}$ возрастает и ограничена сверху, значит, она имеет предел, который обозначим через c. Покажем, что $c=2$. Действительно, воспользовавшись рекуррентным соотношением $a_n = \sqrt{2+a_{n-1}}$, имеем $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} \sqrt{2+a_{n-1}}$.

Применяя теорему: если последовательность $\{b_n\}$ такова, что $b_n \geqslant 0$ и она имеет предел, то $\lim\limits_{n\to+\infty} \sqrt{b_n} = \sqrt{\lim\limits_{n\to+\infty} b_n}$, получаем: $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} \sqrt{2+a_{n-1}} = \sqrt{\lim\limits_{n\to+\infty} 2 + \lim\limits_{n\to+\infty} a_{n-1}}$. Поскольку $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} a_{n-1}$, то получаем, что c надо искать из условия $c = \sqrt{2+c}$. Возводя это равенство в квадрат, получаем, что $c=2$.

5. Пусть $a > 1$ и α — положительное иррациональное число. Вспомним определение a^α и уточним его. Поскольку любое иррациональное число есть бесконечная десятичная дробь, то, обрывая эту дробь на каком-то шаге, получаем приближенное значение этого числа с недостатком, а прибавляя к последней цифре единицу, получаем приближенное значение этого числа с избытком. Значит, для приближения числа α, равного $p, q_1q_2q_3\ldots q_n\ldots$ получаем две последовательности:

$$p,\ p+\frac{q_1}{10},\ p+\frac{q_1}{10}+\frac{q_2}{100},\ p+\frac{q_1}{10}+\frac{q_2}{100}+\frac{q_3}{1000},\ \ldots,$$

$$p+1,\ p+\frac{q_1+1}{10},\ p+\frac{q_1}{10}+\frac{q_2+1}{100},\ p+\frac{q_1}{10}+\frac{q_2}{100}+\frac{q_3+1}{1000},\ \ldots.$$

Обозначим общий член первой последовательности b_n, а общий член второй — c_n. Тогда очевидно, что последовательность $\{b_n\}$ возрастающая и ограниченная сверху (хотя бы числом $p+1$), а последовательность c_n убывающая и ограниченная снизу (хотя бы числом p). Рассмотрим последовательность $\{a^{b_n}\}$. Так как $a > 1$, то из условия $b_n < b_{n+1}$ вытекает, что $a^{b_n} < a^{b_{n+1}}$. Поскольку последовательность $\{b_n\}$ возрастающая, то последнее неравенство означает, что последовательность $\{a^{b_n}\}$ возрастающая. Кроме того, $a^{b_n} < a^{p+1}$, т. е. последовательность $\{a^{b_n}\}$ ограничена сверху. По теореме 15 последовательность $\{a^{b_n}\}$ имеет предел, который обозначим A. Аналогично показывается, что последовательность $\{a^{c_n}\}$ имеет предел, который обозначим B. Докажем, что $A = B$.

Рассмотрим последовательность $\{a^{c_n} - a^{b_n}\}$ и покажем, что $\lim\limits_{n \to +\infty} (a^{c_n} - a^{b_n}) = 0$. Действительно, применяя теоремы о пределах, имеем $\lim\limits_{n \to +\infty} (a^{c_n} - a^{b_n}) = \lim\limits_{n \to +\infty} [a^{b_n}(a^{c_n - b_n} - 1)] = $
$= \lim\limits_{n \to +\infty} a^{b_n} \lim (a^{c_n - b_n} - 1) = \lim a^{b_n} (a^{\lim\limits_{n \to +\infty} (c_n - b_n)} - 1) =$
$= A(a^0 - 1) = 0$, так как $\lim (c_n - b_n) = \lim \frac{1}{10^{n-1}} = 0$. Из условия $\lim\limits_{n \to +\infty} (a^{c_n} - a^{b_n}) = 0$ получаем, что $\lim\limits_{n \to +\infty} a^{c_n} - \lim\limits_{n \to +\infty} a^{b_n} = 0$, т. е. $A = B$.

Итак, обе последовательности имеют один предел, который и называется числом a^α.

В случае $a > 1$ и $\alpha < 0$ или $0 < a < 1$ и α — любое иррациональное число рассуждения аналогичны.

6. Пусть задана последовательность $\{a_n\}$. Тогда можно построить другую последовательность $\{S_n\}$ по правилу

$$S_1 = a_1,\ S_2 = a_1 + a_2, \ldots\ S_n = a_1 + a_2 + \ldots + a_n, \ldots.$$

Формально можно написать бесконечную сумму $a_1 + a_2 + a_3 + \ldots + a_{n-1} + a_n + \ldots$, которую называют *рядом*. Последовательность $\{S_n\}$ называется *последовательностью частичных сумм* этого ряда. В некоторых случаях последовательность $\{S_n\}$ может иметь конечный предел S. В таких случаях говорят, что ряд сходится, а число S называется суммой ряда. Приведем несколько примеров.

6.1. Пусть задана геометрическая прогрессия $\{a_n\}$ с первым членом a_1 и со знаменателем q, таким, что $|q| < 1$. Рассмотрим ряд $a_1 + a_1 q + a_1 q^2 + \ldots + a_1 q^{n-1} + \ldots$ и составим последовательность $\{S_n\}$ частичных сумм этого ряда. Формула для вычисления S_n для любого n следующая:

$$S_n = a_1 \frac{1 - q^n}{1 - q}.$$

Применяя теоремы о пределах и учитывая, что $|q|<1$, получаем, что $\lim\limits_{n\to+\infty} S_n = \lim\limits_{n\to+\infty}\left[\frac{a_1}{1-q}+\frac{a_1}{1-q}q^n\right]=\frac{a_1}{1-q}+$ $+\frac{a_1}{1-q}\lim q^n$. Поскольку $\lim q^n=0$ для $|q|<1$, то получаем, что $\lim\limits_{n\to+\infty} S_n=a/(1-q)$, т. е. по определению суммой написанного ряда будет число $a_1/(1-q)$, т. е. сумма геометрической прогрессии с $|q|<1$ равна первому члену, поделенному на разность единицы и знаменателя прогрессии.

6.2. Пользуясь определением суммы ряда, можно дать другое определение иррационального числа α.

Пусть положительное иррациональное число α равно $\overline{p,\ q_1q_2q_3\ldots q_n\ldots}$. Тогда последовательность p, $p+\frac{q_1}{10}$, $p+\frac{q_1}{10}+$ $+\frac{q_2}{100}$, ... — последовательность десятичных приближений числа α с недостатком, последовательность $p+1$, $p+\frac{q_1+1}{10}$, $p+\frac{q_1}{10}+$ $+\frac{q_2+1}{100}$, ... — последовательность десятичных приближений числа α с избытком. Эти последовательности имеют один предел, который называется числом α. Поэтому можно определить α как сумму ряда $\alpha=p+q_1/10+q_2/100+q_3/1000+\ldots+$ $+q_n/10^n+\ldots$. Заметим, что если α есть бесконечная периодическая дробь

$$\overline{\alpha=p_0,\ p_1p_2\ldots p_k\,(q_1q_2\ldots q_m)},$$

то α тоже можно определить как сумму ряда

$$\alpha=p_0+{}^{p_1}/_{10}+{}^{p_2}/_{10^2}+\ldots+{}^{p_k}/_{10^k}+{}^{q_1}/_{10^{k+1}}+\ldots+{}^{q_m}/_{10^{k+m}}+$$
$$+{}^{q_1}/_{10^{k+m+1}}+\ldots+q_m/10^{k+2m}+\ldots$$

6.3. Докажем правило перевода периодической десятичной дроби в обыкновенную. Пусть дана положительная периодическая десятичная дробь, целая часть которой, для простоты, равна нулю. Тогда эта десятичная дробь равна обыкновенной, у которой: числитель есть число, равное разности чисел, составленных цифрами, стоящими до второго периода, и цифрами, стоящими до первого периода; знаменатель есть число, в изображении которого цифра 9 повторяется столько раз, сколько цифр в периоде, а затем после девяток нуль повторяется столько раз, сколько цифр от запятой до периода.

Пример. $0{,}3\,(14)=\frac{314-3}{990}=\frac{311}{990}$; $0{,}127\,(31)=\frac{12\,731-127}{99\,000}=$ $=\frac{12\,604}{99\,000}$.

Доказательство. Пусть дробь α имеет вид

$$\alpha=\overline{0,\ p_1p_2\ldots p_k\,(q_1q_2\ldots q_m)}.$$

Запишем эту дробь в виде суммы ряда $\alpha = p_1/10 + p_2/10^2 + \ldots + p_k/10^k + q_1/10^{k+1} + q_2/10^{k+2} + \ldots + q_m/10^{k+m} + q_1/10^{k+m+1} + \ldots + q_m/10^{k+2m} + q_1/10^{k+2m+1} + \ldots + q_m/10^{k+3m} + \ldots$. Сделаем очевидные преобразования этого ряда

$$\alpha = \frac{1}{10^k}(10^{k-1}p_1 + 10^{k-2}p_2 + \ldots + p_k) +$$
$$\frac{1}{10^{k+m}}(10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m) +$$
$$+ \frac{1}{10^{k+2m}}(10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m) +$$
$$+ \frac{1}{10^{k+3m}}(10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m) + \ldots.$$

В этом ряде члены, начиная со второго, образуют геометрическую прогрессию со знаменателем $1/10^m$. Применяя формулу для суммы геометрической прогрессии с $|q| < 1$, получаем

$$\alpha = \frac{1}{10^k}(10^{k-1}p_1 + 10^{k-2}p_2 + \ldots + p_k) + \frac{\frac{1}{10^{k+m}}(10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m)}{1 - \frac{1}{10^m}} =$$

$$= \frac{1}{10^k}(10^{k-1}p_1 + 10^{k-2}p_2 + \ldots + p_k) + \frac{10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m}{10^k(10^m - 1)} =$$

$$= \frac{(10^m - 1)(10^{k-1}p_1 + 10^{k-2}p_2 + \ldots + p_k) + 10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m}{10^k(10^m - 1)} =$$

$$= \frac{(10^{m+k-1}p_1 + 10^{m+k-2}p_2 + \ldots + 10^m p_k + 10^{m-1}q_1 + 10^{m-2}q_2 + \ldots + q_m) - (10^{k-1}p_1 + 10^{k-2}p_2 + \ldots + p_k)}{10^k(10^m - 1)} =$$

$$= \frac{\overline{p_1p_2 \ldots p_k q_1 q_2 \ldots q_m} - \overline{p_1 p_2 \ldots p_k}}{\underbrace{99\ldots9}_{m \text{ раз}}\underbrace{00\ldots0}_{k \text{ раз}}},$$

и правило перевода доказано.

§ 4
ПРЕДЕЛ ФУНКЦИИ

Пусть задана функция $y = f(x)$. Точка a (a — конечное число) называется *точкой сгущения* области определения функции $y = f(x)$, если в любом сколь угодно малом промежутке оси OX, содержащем точку a, есть хотя бы одна точка области определения функции, отличная от точки a. Заметим, что точка a может и не принадлежать области определения функции.

Примеры. 1. Для функции $y = 2^x$ любая точка оси OX есть точка сгущения этой функции и все точки сгущения принадлежат области определения функции.

2. Для функции $y=1/x$ любая точка оси OX есть точка сгущения этой функции. Точка $x=0$ также есть точка сгущения, но она не принадлежит области определения функции.

3. Пусть задана функция $y=\sqrt{\log \sin x}$; область определения этой функции есть $x=\pi/2+2\pi k$, где $k=0, \pm 1, \pm 2, \pm 3, \ldots$. У этой функции нет точек сгущения.

Часто определение точки сгущения дается в несколько других терминах. Для любого $\delta>0$ промежуток $a-\delta<x<a+\delta$ оси OX называется δ (*дельта*)-*окрестностью точки* a.

Точка a (a — конечное число) называется *точкой сгущения* области определения функции $y=f(x)$, если в каждой δ-окрестности точки a содержится хотя бы одно, отличное от a, значение x из области определения функции.

Если точка a есть точка сгущения области определения функции $y=f(x)$, то найдется притом бесконечно много последовательностей $\{x_n\}$, $x_n \neq a$ из области определения функции $y=f(x)$, имеющих своим пределом точку a.

Примеры. 1. Пусть задана функция $y=2^x$ и точка $a=1$ — точка сгущения области определения этой функции. Последовательности точек x_n, например, такие: $x_n=1+1/n$, $x_n=1+1/2n$, $x_n=1+1/n^2$, $x_n=1-1/n$, $x_n=1-2/n$, $x_n=1-1/(n^2+n)$, $x_n=1+(-1)^n/n$ и т. д. — последовательности точек из области определения функции, причем в каждом случае $\lim\limits_{n\to+\infty} x_n=1$.

2. Пусть задана функция $y=1/x$ и точка $a=0$ — точка сгущения области определения этой функции. Последовательности точек x_n, например, такие: $x_n=1/n$, $x_n=(-1)^n/n$, $x_n=1/2^n$, $x_n=1/(4^n+n)$, $x_n=-1/n$, $x_n=2/n$, $x_n=1/10^n$ и т. д. — последовательности точек из области определения функции, причем $\lim\limits_{n\to+\infty} x_n=0$.

Покажем, что для любой функции $y=f(x)$ и любой точки сгущения a области определения этой функции можно построить хотя бы одну последовательность точек $\{x_n\}$ из области определения, такую, что $\lim\limits_{n\to+\infty} x_n=a$.

Выберем положительное число δ_1 и возьмем окрестность точки a, соответствующую числу δ_1, и в ней возьмем любую точку $x_1 \neq a$ из области определения этой функции. Возьмем теперь положительное число δ_2, такое, что $\delta_2<\delta_1$ и $\delta_2<|a-x_1|$. В окрестности точки a, соответствующей числу δ_2, возьмем любую точку $x_2 \neq a$ из области определения этой функции. Затем возьмем положительное число δ_3, такое, что $\delta_3<\delta_2$ и $\delta_3<|a-x_2|$. В окрестности точки a, соответствующей числу δ_3, возьмем любую точку $x_3 \neq a$ из области определения этой функции и т. д. В результате получим последовательность точек:

$$x_1,\ x_2,\ x_3,\ \ldots x_n,\ \ldots .$$

Последовательность положительных чисел $\{\delta_n\}$ берется такой, что кроме указанных условий она удовлетворяет условию $\lim\limits_{n\to+\infty}\delta_n=0$. Покажем, что $\lim\limits_{n\to+\infty}x_n=a$. Действительно, для любого $\varepsilon>0$ можно найти номер N, такой, что $\delta_N<\varepsilon$. Тогда для любого $n>N$ справедливо неравенство $|x_n-a|<\delta_n$, поскольку $\delta_n<\delta_N<\varepsilon$, тем более справедливо неравенство $|x_n-a|<\varepsilon$, т. е. $\lim x_n=a$.

Из построения видно, что для любой функции $y=f(x)$ и любой точки сгущения a области определения этой функции можно построить бесконечно много последовательностей точек из области определения, таких, что $\lim\limits_{n\to+\infty}x_n=a$.

Заметим, что каждой последовательности $x_1, x_2, x_3, \ldots, x_n, \ldots$ соответствует последовательность $f(x_1), f(x_2), f(x_3), \ldots, f(x_n), \ldots$

Пусть точка a — точка сгущения области определения функции $y=f(x)$. Число A называется *пределом* этой *функции при стремлении x к a*, если для любой последовательности точек из области определения функции $\{x_n\}$, $x_n\neq a$, имеющей пределом число a, последовательность $\{f(x_n)\}$ имеет пределом число A. При этом пишут $\lim\limits_{x\to a}f(x)=A$.

Пределом функции может являться как конечное число, так и $+\infty$ или $-\infty$.

Примеры. 1. Доказать, что $\lim\limits_{x\to1}2^x=2$. Для этого покажем, что $\lim\limits_{n\to+\infty}2^{x_n}=2$ для любой последовательности точек из области определения функции $\{x_n\}$, $x_n\neq1$, такой, что $\lim\limits_{n\to+\infty}x_n=$ $=1$, т. е. что для любой такой последовательности и любого $\varepsilon>0$ найдется номер N, такой, что для любого $n>N$ справедливы неравенства $2-\varepsilon<2^{x_n}<2+\varepsilon$. Заметим, что условие $\lim\limits_{n\to+\infty}x_n=1$ означает, что для любого числа $\varepsilon_1>0$ найдется номер N_1, такой, что для любого $n>N_1$ справедливы неравенства $1-\varepsilon_1<x_n<1+\varepsilon_1$.

Пусть теперь дано положительное число ε. Выберем число $\varepsilon_1=\log_2(1+\varepsilon/2)$. Ясно, что $\varepsilon_1>0$. Возьмем теперь любую последовательность точек из области определения функции $\{x_n\}$, $x_n\neq$ $\neq1$, такую, что $\lim\limits_{n\to+\infty}x_n=1$. Тогда для этой последовательности $\{x_n\}$ существует номер N_1, такой, что для любого $n>N_1$ справедливы неравенства $1-\varepsilon_1<x_n<1+\varepsilon_1$, а следовательно, и неравенства

$$2^{x_n}<2^{1+\varepsilon_1}=2\left[2^{\log_2(1+\varepsilon/2)}\right]=2(1+\varepsilon/2)=2+\varepsilon;$$
$$2^{x_n}>2^{1-\varepsilon_1}=2\left[2^{\log_2(1+\varepsilon/2)}\right]^{-1}=\frac{2}{1+\varepsilon/2}=\frac{4}{2+\varepsilon}>\frac{4-\varepsilon^2}{2+\varepsilon}=2-\varepsilon.$$

Итак, для любой последовательности точек из области определения функции $\{x_n\}$, $x_n\neq1$, такой, что $\lim x_n=1$, и для любого $\varepsilon>0$ найден номер $N=N_1$, такой, что для любого

$n>N$ справедливо неравенство $|2^{x_n}-2|<\varepsilon$, т. е. показано, что $\lim\limits_{n\to+\infty} 2^{x_n}=2$ для любой последовательности точек из области определения функции $\{x_n\}$, $x_n\neq 1$, такой, что $\lim\limits_{n\to+\infty} x_n=1$. Следовательно, $\lim\limits_{x\to 1} 2^x=2$.

2. Доказать, что $\lim\limits_{x\to 0} 1/|x|=+\infty$.

Пусть $\{x_n\}$, $x_n\neq 0$ — любая последовательность точек из области определения функции, такая, что $\lim x_n=0$. Это означает, что для любого числа $\varepsilon_1>0$ найдется номер N, такой, что для любого $n>N$ справедливы неравенства $0<|x_n|<\varepsilon_1$. Возьмем произвольное положительное число B и обозначим $\varepsilon_1=$ $=1/B$. Тогда для любой последовательности $\{x_n\}$, $x_n\neq 0$, такой, что $\lim\limits_{n\to+\infty} x_n=0$, найдется номер N, такой, что для любого $n>N$ справедливы неравенства $0<|x_n|<1/B$, т. е. справедливо неравенство $\frac{1}{|x_n|}>B$.

Итак, для любого $B>0$ нашелся номер N, такой, что для любого $n>N$ справедливо неравенство $\frac{1}{|x_n|}>B$, т. е. $\lim 1/|x_n|=+\infty$ для любой последовательности точек из области определения функции $\{x_n\}$, $x_n\neq 0$, такой, что $\lim\limits_{n\to+\infty} x_n=0$. Следовательно, $\lim\limits_{x\to 0} 1/|x|=+\infty$.

Будем говорить, что $+\infty$ *является точкой сгущения* для области определения функции $y=f(x)$, если для любого большого положительного числа B найдется хотя бы одно значение $x>B$, принадлежащее области определения функции.

Будем говорить, что $-\infty$ *является точкой сгущения* для области определения функции $y=f(x)$, если для любого отрицательного числа B, такого, что $|B|$ — сколь угодно большое число, найдется хотя бы одно значение $x<B$, принадлежащее области определения функции.

В случаях, когда $+\infty$ или $-\infty$ являются точками сгущения для области определения функции $y=f(x)$, также легко показать, что можно построить последовательность точек $\{x_n\}$ из области определения $y=f(x)$, которая будет иметь своим пределом $+\infty$ или $-\infty$ соответственно и определение предела функции $y=f(x)$ можно распространить на случаи таких точек сгущения.

Пример. Для функции $y=1/x+1$ $+\infty$ и $-\infty$ являются точками сгущения и $\lim\limits_{x\to+\infty}(1/x+1)=\lim\limits_{x\to-\infty}(1/x+1)=1$.

Пользоваться приведенным определением предела функции для его отыскания трудно. Поэтому чаще пользуются другим эквивалентным определением на так называемом языке «ε, δ».

Пусть точка a есть точка сгущения области определения функции $y=f(x)$ и a — конечное число. Число A называется *пределом функции* $y=f(x)$ *при стремлении* x *к* a, если для любого $\varepsilon>0$ найдется $\delta>0$, такое, что для любого x из области определения функции, и такого, что $0<|x-a|<\delta$, выполняется неравенство $|f(x)-A|<\varepsilon$.

При этом пишут $\lim\limits_{x\to a} f(x)=A$.

Пусть $+\infty$ есть точка сгущения области определения функции $y=f(x)$. Число A называется *пределом функции* $y=f(x)$ *при стремлении* x *к* $+\infty$, если для любого $\varepsilon>0$ найдется $M>0$, такое, что для любого x из области определения функции $y=f(x)$, и такого, что $M<x$, выполняется неравенство $|f(x)-A|<\varepsilon$. При этом пишут $\lim\limits_{x\to+\infty} f(x)=A$.

Пусть $-\infty$ есть точка сгущения области определения функции $y=f(x)$. Число A называется *пределом функции* $y=f(x)$ *при стремлении* x *к* $-\infty$, если для любого $\varepsilon>0$ найдется $M<0$, такое, что для любого x из области определения функции $y=f(x)$, и такого, что $x<M$, выполняется неравенство $|f(x)-A|<\varepsilon$. При этом пишут $\lim\limits_{x\to-\infty} f(x)=A$.

Аналогично можно дать определения $\lim\limits_{x\to a} f(x)=-\infty$, $\lim\limits_{x\to a} f(x)=+\infty$, $\lim\limits_{x\to-\infty} f(x)=-\infty$ и т. д. Для примера приведем одно из определений.

Пусть a есть точка сгущения области определения функции $y=f(x)$ и a — конечное число. Будем говорить, что *функция* $y=f(x)$ *имеет пределом* $+\infty$ *при стремлении* x *к* a, если для любого числа $B>0$ найдется число $\delta>0$, такое, что для любого x из области определения функции, и такого, что $0<|x-a|<\delta$, будет выполняться неравенство $f(x)>B$. При этом пишут $\lim\limits_{x\to a} f(x)=+\infty$.

Найдем пределы некоторых функций, используя определение предела на языке «ε, δ».

Пример. 1. Дана функция $y=2^x$ и точка $a=1$ — точка сгущения области определения этой функции. Доказать, что $\lim\limits_{x\to 1} 2^x=2$.

Возьмем произвольное положительное число ε и выберем число $\delta>0$, например $\delta=\log_2(1+\varepsilon/2)$. Ясно, что $\delta>0$. Возьмем теперь любое x, такое, что $0<|x-1|<\delta$, т. е. любое $x\neq 1$ из промежутка $1-\delta<x<1+\delta$.

Ясно, что $2^x<2^{1+\delta}=2\cdot 2^{\log_2(1+\varepsilon/2)}=2+\varepsilon$, $2^x>2^{1-\delta}=$

$$=2\left[2^{\log_2(1+\varepsilon/2)}\right]^{-1}=\frac{2}{1+\varepsilon/2}=\frac{4}{2+\varepsilon}>\frac{4-\varepsilon^2}{2+\varepsilon}=2-\varepsilon.$$

Эти неравенства означают, что $|2^x-2|<\varepsilon$. Итак, для любого $\varepsilon>0$ нашлось $\delta>0$, такое, что $|2^x-2|<\varepsilon$ для любого x из области определения функции, и такого, что $0<|x-1|<\delta$. По определению это и означает, что $\lim\limits_{x\to 1} 2^x=2$.

2. Дана функция $y={}^1/_{|x|}$ и точка $a=0$ — точка сгущения области определения этой функции. Доказать, что $\lim\limits_{x\to 0}({}^1/_{|x|})=$ $=+\infty$.

Возьмем произвольное число $B>0$ и выберем число $\delta>0$, например $\delta>{}^1/_B$. Ясно, что как только $0<|x-0|=|x|<\delta$, то $|{}^1/_x|>{}^1/_\delta=B$. Итак, для любого $B>0$ нашлось число $\delta>0$, такое, что ${}^1/_{|x|}>B$ для любого x из области определения функции и такого, что $0<|x-0|<\delta$. По определению это и означает, что $\lim\limits_{x\to 0} {}^1/_{|x|}=+\infty$.

Выше даны два разных определения предела функции: одно на языке последовательностей, другое — на языке «ε, δ». Эти определения равносильны, т. е. если функция имеет предел в смысле определения на языке последовательностей, то она имеет тот же предел в смысле определения на языке «ε, δ» и наоборот. Покажем равносильность двух определений предела функции в случае, когда $x=a$ — конечная точка сгущения и когда функция имеет конечный предел A.

1. Пусть A — предел функции $y=f(x)$ при x, стремящемся к точке a в смысле определения на языке «ε, δ».

Возьмем произвольное $\varepsilon>0$. Тогда найдется число $\delta>0$, такое, что для любого x, принадлежащего области определения функции и удовлетворяющего неравенствам $0<|x-a|<\delta$, выполняется неравенство $|f(x)-A|<\varepsilon$.

Рассмотрим любую последовательность $\{x_n\}$, сходящуюся к a, и такую, что $x_n\neq a$ и x_n принадлежит области определения функции $y=f(x)$ при любом n. Тогда для указанного числа $\delta>0$, зависящего от ε, существует число N, такое, что для любого $n>N$ выполняются неравенства $|x_n-a|<\delta$. Следовательно, при любом $n>N$ будет выполняться неравенство $|f(x_n)-A|<\varepsilon$. Но так как $\varepsilon>0$ выбиралось произвольно, то это и означает, что $f(x_n)\to A$, т. е. для любой последовательности $\{x_n\}$ (из области определения функции $y=f(x)$), сходящейся к a, последовательность $\{f(x_n)\}$ сходится к A, это и есть определение предела функции на языке последовательностей.

2. Пусть $\lim\limits_{x\to a} f(x)=A$ в смысле определения на языке последовательностей. Покажем, что $\lim\limits_{x\to a} f(x)=A$ в смысле определения на языке «ε, δ».

Доказательство проведем от противного. Пусть $f(x_n)\to A$ для любой последовательности $\{x_n\}$, $x_n\neq a$, точек из области определения функции, такой, что $x_n\to a$. Предположим, что для $f(x)$ число A не является пределом при $x\to a$ в смысле определения на языке «ε, δ», т. е. существует такое положительное число ε, что для любого положительного δ найдется хотя бы одно число $\bar{x}$, такое, что $0<|\bar{x}-a|<\delta$, но $|f(\bar{x})-A|\geqslant\varepsilon$.

Возьмем теперь последовательность $\{\delta_n\}$, такую, что $\delta_n \to 0$. По предположению для любого δ_n существует число $\bar{x}_n$, такое, что $0 < |\bar{x}_n - a| < \delta_n$, но $|f(\bar{x}_n) - A| \geqslant \varepsilon$. Рассмотрим последовательность $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n, \ldots$. Ясно, что $\bar{x}_n \to a$, но $f(\bar{x}_n)$ не стремится к A, так как $|f(\bar{x}_n) - A| \geqslant \varepsilon$ для всех n.

Следовательно, нашлась последовательность $\{\bar{x}_n\}$, $\bar{x}_n \neq a$ из области определения функции $y = f(x)$, такая, что $\lim\limits_{n \to +\infty} \bar{x}_n = a$, но последовательность $\{f(\bar{x}_n)\}$ не стремится к A. Это противоречит условию, что для любой последовательности точек из области определения функции $\{x_n\}$, $x_n \neq a$, такой, что $\lim\limits_{n \to +\infty} \bar{x}_n = a$ обязательно $\lim\limits_{n \to +\infty} f(\bar{x}_n) = A$, т. е. наше предположение неверно и равносильность двух определений предела функции доказана (в случае, когда точка сгущения a и $\lim\limits_{x \to a} f(x)$ — конечные числа). В остальных случаях равносильность определений доказывается аналогично.

Для пределов функций справедливы свойства, аналогичные свойствам пределов для последовательностей. Заметим, что их практически не надо доказывать, ибо из определения предела функций при помощи последовательностей вытекает, что все теоремы, доказанные для последовательностей, справедливы и для функций. Сформулируем некоторые из них.

Теорема 1. Пусть точка a есть точка сгущения общей части областей определения функций $y = f(x)$ и $y = g(x)$ и пусть существуют оба конечных предела $\lim\limits_{x \to a} f(x) = A$, $\lim\limits_{x \to a} g(x) = B$. Тогда функции $f(x) + g(x)$, $f(x) - g(x)$, $f(x)g(x)$, $f(x)/g(x)$ также имеют конечные пределы $A + B$, $A - B$, AB, A/B (в случае частного: $B \neq 0$).

Теорема 2. Пусть точка a (a — конечное число) есть точка сгущения области определения функции $y = f(x)$ и если $\lim\limits_{x \to a} f(x) = A$, где $A > 0$, то существует δ — окрестность точки a, такая, что для любого $x \neq a$ из области определения и принадлежащего этой окрестности, функция $y = f(x)$ положительна.

Аналогичное свойство справедливо, если $A < 0$.

Теорема 3. Если точка a (a — конечное число) есть точка сгущения области определения функции $y = f(x)$ и если $\lim\limits_{x \to a} f(x) = A$, где A — конечное число, то существует такая δ — окрестность точки a, что для любого $x \neq a$ из области определения и принадлежащего этой окрестности, функция $y = f(x)$ будет ограниченной, т. е. найдется такое число $M > 0$ и такое $\delta > 0$, что для любого x из области определения функции, такого, что $0 < |x - a| < \delta$ будет справедливо неравенство $|f(x)| \leqslant M$.

Теорема 4. Пусть точка a есть точка сгущения общей части областей определения функций $y = f(x)$, $y = p(x)$ и

$y=q(x)$ и для любого $x \neq a$, принадлежащего некоторой δ — окрестности точки a и общей части областей определения этих функций, справедливы неравенства $f(x) \leqslant p(x) \leqslant q(x)$. Если $\lim\limits_{x \to a} f(x) = \lim\limits_{x \to a} q(x) = A$, тогда и $\lim\limits_{x \to a} p(x) = A$.

Пример. Используя теорему 4, докажем, что $\lim\limits_{x \to 0} ({}^{\sin x}/_{x}) = 1$.

Функция $y = {}^{\sin x}/_{x}$ четная, поэтому рассмотрим ее на интервале $(0, {}^{\pi}/_{2})$. Докажем, что на этом интервале имеет место двойное неравенство

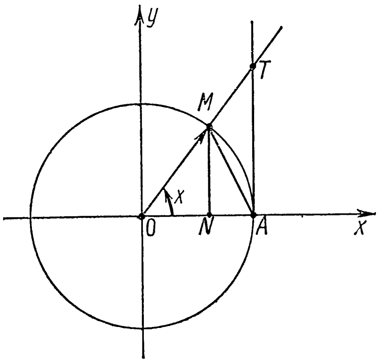


Рис. 133

$$\cos x < {}^{\sin x}/_{x} < 1. \qquad (1)$$

Возьмем дугу $\overset{\frown}{AM}$ единичной окружности, соответствующую углу, радианная мера которого равна x (рис. 133). Тогда $OA = 1$, $MN = \sin x$, $ON = \cos x$. Из подобия треугольников OAT и ONM находим $AT = \frac{AT}{OA} = \frac{MN}{ON} = \frac{\sin x}{\cos x} = \operatorname{tg} x$. Так как площадь треугольника ONM меньше площади сектора OAM, а площадь этого сектора меньше площади треугольника OAT, то имеем двойное неравенство $\frac{1}{2} AO \cdot MN < \frac{1}{2} OA \cdot \overset{\frown}{AM} < \frac{1}{2} OA \cdot AT$ или $\sin x < x < \operatorname{tg} x$. Разделим это двойное неравенство на $\sin x > 0$. Поскольку $\sin x > 0$, то знаки неравенства не изменятся $1 < \frac{x}{\sin x} < \frac{1}{\cos x}$.

Так как рассматривается интервал $0 < x < {}^{\pi}/_{2}$ и на нем $\cos x > 0$, то имеем равносильные неравенства

$$1 < \frac{x}{\sin x} \Leftrightarrow \frac{\sin x}{x} < 1; \ \frac{x}{\sin x} < \frac{1}{\cos x} \Leftrightarrow \cos x < \frac{\sin x}{x},$$

что и доказывает исходное неравенство (1).

В силу четности функций $\cos x$ и ${}^{\sin x}/_{x}$ двойное неравенство (1) будет иметь место и на интервале $(-{}^{\pi}/_{2}; 0)$. Действительно, если $x \in (-{}^{\pi}/_{2}; 0)$, то $0 < (-x) < {}^{\pi}/_{2}$ и неравенство (1) имеет вид

$$\cos(-x) < \frac{\sin(-x)}{(-x)} < 1.$$

Но $\cos(-x) = \cos x$, $\frac{\sin(-x)}{(-x)} = \frac{-\sin x}{-x} = \frac{\sin x}{x}$. Следовательно, неравенство (1) справедливо и при $x \in (-{}^{\pi}/_{2}, 0)$. Итак, для любого $x \in (-{}^{\pi}/_{2}, 0) \cup (0, {}^{\pi}/_{2})$ имеем $\cos x < \frac{\sin x}{x} < 1$. Для того чтобы теперь воспользоваться теоремой 4, необходимо доказать, что $\lim\limits_{x\to 0}\cos x=1$. Воспользуемся определением предела функции на языке «ε, δ», т. е. докажем, что для любого $\varepsilon>0$ найдется такое $\delta>0$, что как только $|x|<\delta$, то $|1-\cos x|\leqslant\varepsilon$. В качестве δ возьмем величину $\sqrt{2\varepsilon}$, тогда $|1-\cos x|=1-\cos x=2\sin^2 {}^{x}/_{2}=2(\sin {}^{|x|}/_{2})^2<2({}^{|x|}/_{2})^2={}^{x^2}/_{2}<$ $<{}^{2\varepsilon}/_{2}=\varepsilon$, что и требовалось доказать.

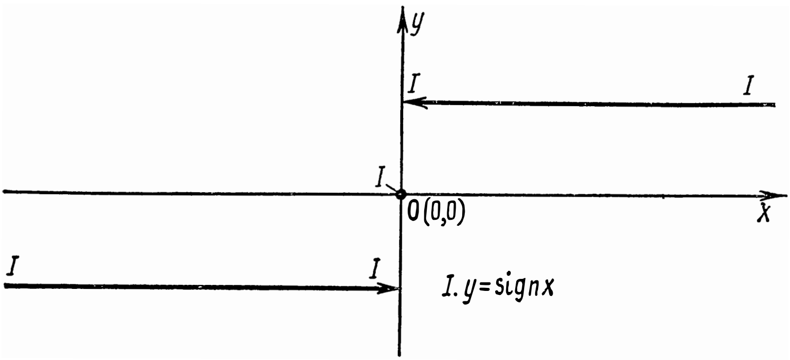


Рис. 134

Итак, функция ${}^{(\sin x)}/_{x}$ при $x\in(-{}^{\pi}/_{2},\ 0)\cup(0,\ {}^{\pi}/_{2})$ заключена между 1 и $\cos x$ и поскольку $\lim\limits_{x\to 0}\cos x=1$, то по теореме 4 получим $\lim\limits_{x\to 0}\dfrac{\sin x}{x}=1$.

В определении предела функции $y=f(x)$ в точке a x (из области определения функции) может приближаться к a по любому закону.

Иногда приходится рассматривать предел функции $y=f(x)$, когда x (из области определения функции), стремясь к a, остается все время справа от точки a. Тогда говорят, что функция $y=f(x)$ имеет *предел справа*, и записывают $\lim\limits_{\substack{x\to a+0\\ x\neq a}} f(x)=A$, где A — либо конечное число, либо $+\infty$ или $-\infty$. Аналогично определяется *предел слева* $\lim\limits_{\substack{x\to a-0\\ x\neq a}} f(x)=A$.

Примеры 1. $\lim\limits_{x\to 0+0}\operatorname{sign} x=1$, $\lim\limits_{x\to 0-0}\operatorname{sign} x=-1$ (рис. 134).

2. $\lim\limits_{x\to 2+0} {}^{1}/_{(x-2)}=+\infty$, $\lim\limits_{x\to 2-0} {}^{1}/_{(x-2)}=-\infty$ (рис. 135).

Функция $y=x^2$ имеет в любой точке a предел и справа и слева, при этом $\lim\limits_{x\to a+0} x^2=\lim\limits_{x\to a-0} x^2=\lim\limits_{x\to a} x^2=a^2$.

Пусть дана функция $y=f(x)$ и ее график. Если существуют точка $M[x_0,\ f(x_0)]$ и прямая, такая, что хотя при одном из условий: $x\to a$, $x\to a+0$, $x\to a-0$, $x\to+\infty$, $x\to-\infty$ начиная с точки M график и прямая не имеют общих точек

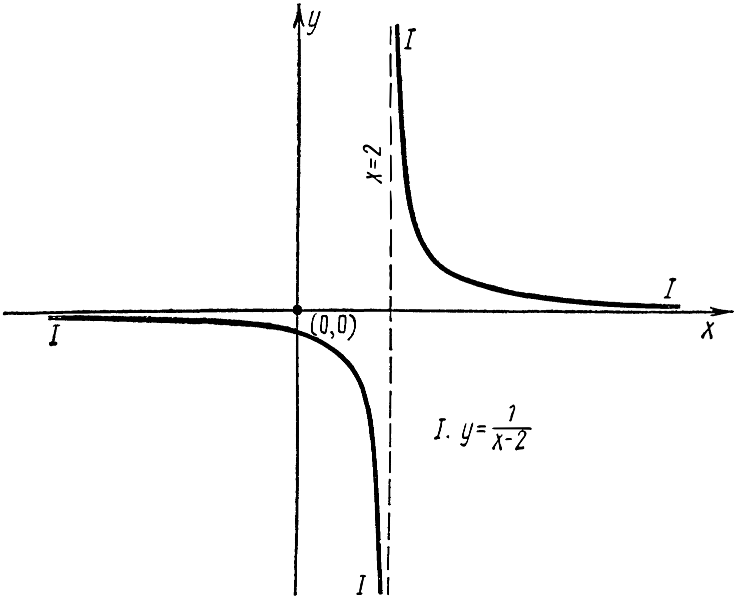


Рис. 135

и расстояние между точками графика и прямой неограниченно уменьшается, то эта прямая называется *асимптотой* функции $y=f(x)$.

Если выполняется хотя бы одно из условий $\lim\limits_{x\to a} f(x)=+\infty$, $\lim\limits_{x\to a} f(x)=-\infty$, $\lim\limits_{x\to a+0} f(x)=+\infty$, $\lim\limits_{x\to a+0} f(x)=-\infty$, $\lim\limits_{x\to a-0} f(x)=$

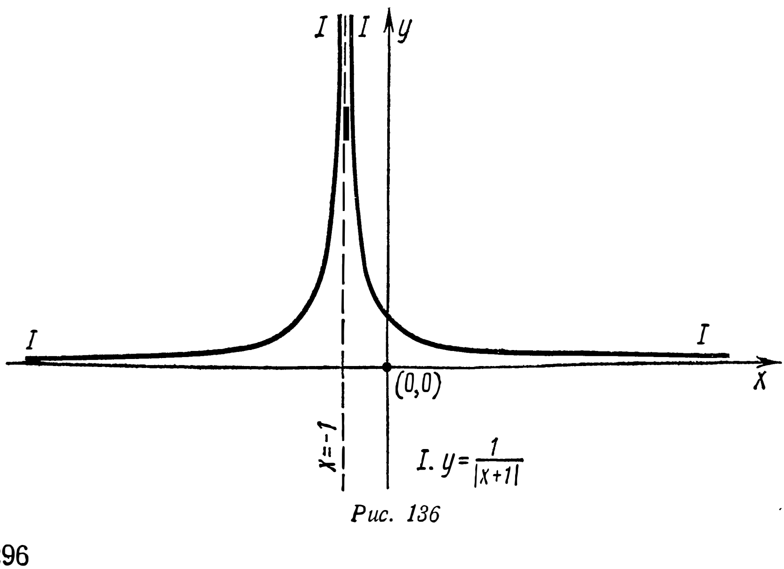


Рис. 136

$= +\infty$, $\lim\limits_{x \to a - 0} f(x) = -\infty$, то асимптота $x = a$ называется *вертикальной*.

Примеры. 1. Функция $y = 1/|x+1|$ имеет вертикальную асимптоту $x = -1$, так как $\lim\limits_{x \to -1} [1/|x+1|] = +\infty$ (рис. 136).

2. Функция $y = \log_2(x+3)$ имеет вертикальную асимптоту $x = -3$, так как $\lim\limits_{x \to -3+0} \log_2(x+3) = -\infty$ (рис. 137).

3. Функция $y = 1/(x-2)$ имеет вертикальную асимптоту $x = 2$, так как $\lim\limits_{x \to 2+0} [1/(x-2)] = +\infty$ и $\lim\limits_{x \to 2-0} [1/(x-2)] = -\infty$ (см. рис. 135).

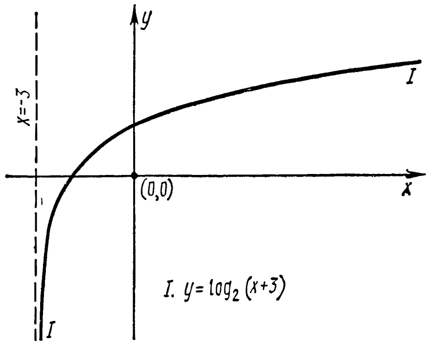


Рис. 137

Если выполняется хотя бы одно из условий $\lim\limits_{x \to +\infty} f(x) = b$ или $\lim\limits_{x \to -\infty} f(x) = b$, то асимптота $y = b$ называется *горизонтальной*.

Примеры. 1. Функция $y = 2 + 3^x$ имеет горизонтальную асимптоту $y = 2$, так как $\lim\limits_{x \to -\infty} (2 + 3^x) = 2$ (рис. 138).

2. Функция $y = 1/x$ имеет горизонтальную асимптоту $y = 0$, так как $\lim\limits_{x \to +\infty} [1/x] = 0$ и $\lim\limits_{x \to -\infty} [1/x] = 0$.

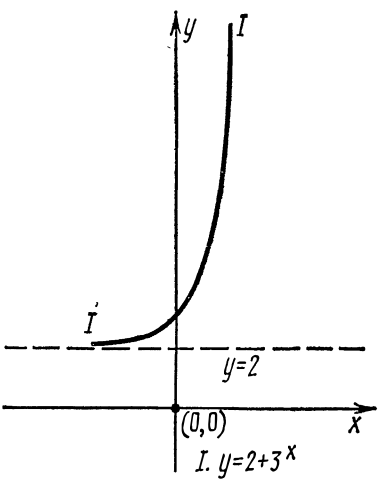


Рис. 138

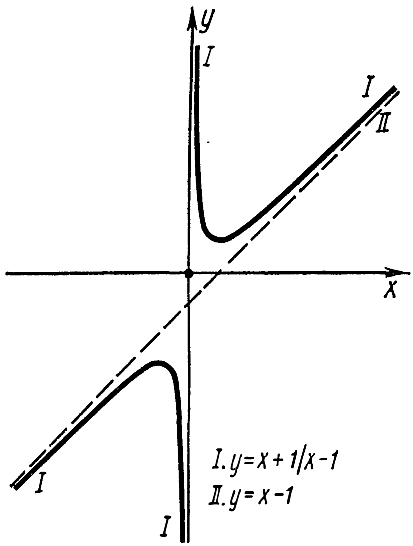


Рис. 139

Если выполняется хотя бы одно из условий $\lim\limits_{x\to+\infty}[f(x)-(kx+b)]=0$ или $\lim\limits_{x\to-\infty}[f(x)-(kx+b)]=0$, то асимптота $y=kx+b$ называется *наклонной*.

Пример. Функция $y=x+\frac{1}{x}-1$ имеет наклонную асимптоту $y=x-1$ (рис. 139).

§ 5
НЕПРЕРЫВНОСТЬ ФУНКЦИИ

Функция $y=f(x)$ непрерывна в точке x_0, если:

1) x_0 — точка сгущения области определения функции $y=f(x)$;

2) x_0 — принадлежит области определения функции $y=f(x)$;

3) существует $\lim\limits_{x\to x_0} f(x)$;

4) $\lim\limits_{x\to x_0} f(x)=f(x_0)$ [т. е. $\lim\limits_{x\to x_0} f(x)=y_0$, где $y_0=f(x_0)$].

Примеры. 1. Функция $y=2^x$ непрерывна, например, в точке $x_0=3$. Действительно: $x_0=3$ — точка сгущения области определения функции $y=2^x$; $x_0=3$ принадлежит области определения функции $y=2^x$, значит, $y_0=2^3=8$; существует $\lim\limits_{x\to 3} 2^x=8$; наконец, $\lim\limits_{x\to 3} 2^x=8=y_0$.

2. Функция $y=\log_3 x$ непрерывна, например, в точке $x_0=1$. Действительно: $x_0=1$ — точка сгущения области определения функции $y=\log_3 x$; $x_0=1$ принадлежит области определения функции $y=\log_3 x$, значит $y_0=\log_3 1=0$; существует $\lim\limits_{x\to 1}\log_3 x=0$; наконец, $\lim\limits_{x\to 1}\log_3 x=0=y_0$.

3. Функция $y=\frac{1}{x-2}$ в точке $x_0=2$ не является непрерывной, так как нарушено, например, условие 2, т. е. точка $x_0=2$ не входит в области определения функции $y=\frac{1}{x-2}$.

4. Функция $y=\operatorname{sing} x$ в точке $x_0=0$ не является непрерывной, так как нарушено, например, условие 3, т. е. не существует $\lim\limits_{x\to 0}\operatorname{sing} x$.

5. Легко проверить, что каждая основная элементарная функция является непрерывной в любой точке ее области существования.

Учитывая вышеприведенные эквивалентные определения предела функции в точке, приведем соответственно эквивалентные определения непрерывности функции в точке.

Функция $y=f(x)$ называется *непрерывной в точке* x_0, если для любого числа $\varepsilon>0$ найдется такое число $\delta>0$, что для любого x из области определения, удовлетворяющего неравенству $|x-x_0|<\delta$, выполняется неравенство $|f(x)-f(x_0)|<\varepsilon$.

Функция $y=f(x)$ называется *непрерывной в точке* x_0, если для любой последовательности $\{x_n\}$ значений аргумента из области определения, сходящейся в точке x_0, соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к значению $f(x_0)$.

Используя второе определение непрерывности функции в точке, покажем, например, что функция $y=2^x$ является непрерывной, например, в точке $x=2$. То есть покажем, что для любого числа $\varepsilon>0$ найдется такое число $\delta>0$, что для любого x, удовлетворяющего неравенству $|x-2|<\delta$ будет выполняться неравенство $|2^x-2^2|<\varepsilon$. Действительно, для любого числа $\varepsilon>0$ возьмем, например, число $\delta=\log_2(1+\varepsilon/4)$. Тогда для любого числа x, удовлетворяющего условию $|x-2|<\delta$, т. е. для любого x из промежутка $2-\delta<x<2+\delta$ будут справедливы следующие неравенства:

$$2^x<2^{2+\delta}<4\cdot 2^{\log_2(1+\varepsilon/4)}<4+\varepsilon,$$

$$2^x>2^{2-\delta}>4\left[2^{\log_2(1+\varepsilon/4)}\right]^{-1}=\frac{4}{1+\varepsilon/4}=\frac{16}{4+\varepsilon}>\frac{16-\varepsilon^2}{4+\varepsilon}=4-\varepsilon.$$

Эти неравенства означают, что $|2^x-2^2|<\varepsilon$ для любого x, такого, что $|x-2|<\delta$, т. е. согласно второму определению непрерывности функции в точке, функция $y=2^x$ непрерывна в точке $x=2$.

Используя третье определение непрерывности функции в точке, покажем, например, что функция $y=[x]$ не является непрерывной в точке $x=2$. Т. е. укажем хотя бы одну последовательность значений аргумента $\{x_n\}$, такую, что $\lim\limits_{n\to\infty} x_n=2$, но соответствующая последовательность значений функции $\{[x_n]\}$ такова, что $\lim\limits_{n\to+\infty}[x_n]\neq[2]$. Такова, например, последовательность значений аргумента с общим членом $x_n=2-1/n$. Действительно, $\lim\limits_{n\to+\infty}(2-1/n)=2$, но $\lim\limits_{n\to+\infty}[2-1/n]=1\neq 2$ и, следовательно $\lim\limits_{n\to+\infty}[2-1/n]\neq[2]$, что требовалось доказать.

Теорема 1. Пусть функции $y=f(x)$ и $y=g(x)$ непрерывны в точке x_0. Тогда в той же точке будут непрерывными и функции $f(x)+g(x)$, $f(x)-g(x)$, $f(x)g(x)$, $\frac{f(x)}{g(x)}$ (последняя при условии, что $g(x_0)\neq 0$).

Эта теорема является следствием теорем о пределе функций в точке. Например, функция $y=\sin x+x^2$ является непрерывной, например, в точке $x_0=1$. Действительно, точка $x_0=1$ принадлежит области существования и функции $y=\sin x$ и функции $y=x^2$. Так как любая основная элементарная функция непрерывна в каждой точке ее области существования, то и функция $y=\sin x$ и функция $y=x^2$ непрерывны в точке $x_0=1$. Тогда согласно теореме 1 и функция $y=\sin x+x^2$ непрерывна в точке $x_0=1$.

Наряду с уже рассмотренной непрерывностью функции в точке часто рассматривается так называемая односторонняя непрерывность функции в точке. Приведем соответствующие определения.

Функция $y=f(x)$ *непрерывна справа в точке* x_0, если:

1) x_0 — точка сгущения области определения функции $y=f(x)$;

2) x_0 входит в область определения функции $y=f(x)$;

3) существует $\lim\limits_{x \to x_0+0} f(x)$, обозначаемый $f(x_0+0)$;

4) $f(x_0+0)=f(x_0)$.

Аналогично определяется *непрерывность функции слева в точке* x_0. Ясно, что если функция $y=f(x)$ непрерывна в точке x_0, то она в этой точке одновременно непрерывна и справа и слева и $f(x_0+0)=f(x_0)=f(x_0-0)$.

Примеры. 1. Функция $y=[x]$ в точке $x_0=1$ непрерывна справа, так как $\lim\limits_{x \to 1+0} [x]=1=[1]$.

2. Функция $y=\text{sing}\, x$ в точке $x_0=0$ не является непрерывной ни справа ни слева. Действительно, хотя функция $y=\text{sing}\, x$ в точке $x_0=0$ имеет и предел справа ($\lim\limits_{x \to 1+0} \text{sing}\, x=1$) и предел слева ($\lim\limits_{x \to 1-0} \text{sing}\, x=-1$), но ни один из этих пределов не совпадает со значением функции $y=\text{sing}\, x$ в точке $x_0=0$ $(\text{sign}\, 0=0)$.

3. Функция $y=2^x$ в любой точке непрерывна и справа и слева.

Функция $y=f(x)$ имеет в точке x_0 разрыв, если не выполняется хотя бы одно из условий 1—4 определения непрерывности функции в точке x_0.

Пример. Функция $y=\text{sign}\, x$ имеет в точке $x_0=0$ разрыв. Действительно, равенство $\lim \text{sign}\, x=\text{sign}\, 0$ не выполняется, так как не существует предела функции $y=\text{sign}\, x$ в точке $x_0=0$ (хотя и существуют односторонние пределы, но они не равны друг другу и ни один из них не равен функции в точке $x_0=0$).

Пусть функция $y=f(x)$ имеет разрыв в точке $x_0=1$. По определению точка x_0 называется *точкой разрыва первого рода*, если в этой точке функция $y=f(x)$ имеет конечный предел как справа, так и слева. Во всех остальных случаях разрыва точка x_0 называется *точкой разрыва второго рода*.

Примеры. 1. Для функции $y=[x]$ точка $x_0=1$ является точкой разрыва первого рода. Действительно, функция $y=[x]$ имеет в точке $x_0=1$ разрыв, так как функция $y=[x]$ в точке $x_0=1$ не имеет предела, но имеет конечный предел и справа $(\lim\limits_{x \to 1+0} [x]=1)$ и слева $\lim\limits_{x \to 1-0} [x]=0$.

2. Для функции $y=\frac{1}{x}$ точка $x_0=0$ является точкой разрыва второго рода. Действительно, функция $y=\frac{1}{x}$ в точке

$x_0=0$ имеет разрыв, так как точка $x_0=0$ не принадлежит области существования функции. Кроме того, функция $y=\frac{1}{x}$ в точке $x_0=0$ не имеет конечного предела справа

$$\lim_{x\to 0+0}\left(\frac{1}{x}\right)=+\infty.$$

3. Для функции $y=\sin\left(\frac{1}{x}\right)$ точка $x_0=0$ является точкой разрыва второго рода. Действительно, функция $y=\sin\left(\frac{1}{x}\right)$ в точке $x_0=0$ имеет разрыв, так как точка $x_0=0$ не принадлежит области существования функции. Кроме того, функция $y=\sin\left(\frac{1}{x}\right)$ в точке $x_0=0$ не имеет предела справа, так как можно указать две последовательности значений аргумента справа $\{x_n\}$ и $\{x'_n\}$, такие, что $\lim\limits_{n\to+\infty} x_n=0$ и $\lim\limits_{n\to+\infty} x'_n=0$, но соответствующие последовательности значений функции $\left\{\sin\left(\frac{1}{x_n}\right)\right\}$ и $\left\{\sin\left(\frac{1}{x'_n}\right)\right\}$ таковы, что $\lim\limits_{n\to+\infty}\sin\left(\frac{1}{x_n}\right)\neq\lim\limits_{n\to+\infty}\sin\left(\frac{1}{x'_n}\right)$. Такими, например, будут последовательность с общим членом $x_n=\frac{1}{\pi/2+2\pi n}$ и последовательность с общим членом $x'_n=\frac{1}{\pi n}$.

Функция $y=f(x)$ называется *непрерывной на некотором промежутке* (a, b), если она непрерывна в каждой точке этого промежутка. Другими словами, функция $y=f(x)$ *непрерывна на промежутке* (a, b), если:

1) весь промежуток входит в область определения функции $y=f(x)$;

2) любая точка x_0 этого промежутка есть точка сгущения области определения функции $y=f(x)$;

3) в любой точке x_0 этого промежутка существует $\lim\limits_{x\to x_0} f(x)$;

4) в любой точке x_0 этого промежутка справедливо равенство $\lim\limits_{x\to x_0} f(x)=f(x_0)$ (т. е. $\lim\limits_{x\to x_0} f(x)=y_0$, где $y_0=f(x_0)$).

Примеры. 1. Функция $y=\cos x$ непрерывна на промежутке $(-\infty, +\infty)$.

2. Функция $y=\log_3 x$ непрерывна на промежутке $(0, +\infty)$.

3. Функция $y={}^1/_{(x-1)}$ непрерывна на промежутке $(1, +\infty)$, кроме того, она непрерывна на промежутке $(-\infty, 1)$, но она не является непрерывной ни на промежутке $[1, +\infty)$, ни на промежутке $(-\infty, 1]$, ни на промежутке $(-\infty, +\infty)$.

§ 6
ПРОИЗВОДНАЯ ФУНКЦИИ

Рассмотрим функцию $y=f(x)$. Пусть x — некоторая точка сгущения функции, принадлежащая ее области определения, $f(x)$ — значение функции в этой точке. От значения x переходим к другому значению аргумента $x_1 \neq x$, которое также принадлежит области определения. Разность $x - x_1$ (обозначим ее через Δx) называется *приращением аргумента в точке* x. Значение функции, соответствующее значению аргумента $x_1 = x + \Delta x$, обозначим $f(x+\Delta x)$.

Разность $f(x+\Delta x) - f(x)$ называется *приращением функции в точке* x, соответствующим приращению аргумента Δx в этой точке, и обозначается Δy или $\Delta f(x)$: $\Delta y = \Delta f(x) = f(x+\Delta x) - f(x)$. В зависимости от вида функции ее приращение Δy может быть равно нулю, положительным или отрицательным. Приращение аргумента Δx также может быть положительным или отрицательным, но $\Delta x \neq 0$.

Примеры. 1. Найти приращения функции $y=|x|$ в точках $x=0$, $x=1$ и $x=-1$, считая, что $|\Delta x| < 1$.

а) Если $x=0$, то

$$\Delta y = |0+\Delta x| - |0| = |\Delta x| = \begin{cases} \Delta x, \text{ если } \Delta x > 0, \\ -\Delta x, \text{ если } \Delta x < 0. \end{cases}$$

б) Если $x=1$, то

$\Delta y = |1+\Delta x| - |1| = 1+\Delta x - 1 = \Delta x$, если $|\Delta x| < 1$.

в) Если $x=-1$, то

$\Delta y = |-1+\Delta x| - |-1| = 1 - \Delta x - 1 = -\Delta x$, если $|\Delta x| < 1$.

2. Найти приращения функции $y=\sin x$ в точках $x=0$, $x=\pi/3$, $x=-\pi/3$ и в произвольной точке x.

а) Если $x=0$, то $\Delta y = \sin(0+\Delta x) - \sin 0 = \sin \Delta x$.

б) Если $x=\pi/3$, то

$\Delta y = \sin(\pi/3 + \Delta x) - \sin(\pi/3) = 2\sin(\Delta x/2)\cos(\pi/3 + \Delta x/2)$.

в) Если $x=-\pi/3$, то

$\Delta y = \sin(-\pi/3 + \Delta x) - \sin(-\pi/3) = 2\sin(\Delta x/2)\cos(-\pi/3 + \Delta x/2)$.

г) Если x — любая точка, то

$\Delta y = \sin(x+\Delta x) - \sin x = 2\sin(\Delta x/2)\cos(x + \Delta x/2)$.

3. Найти приращения функции $y=\sqrt{x}$ в точках $x=0$, $x=1$ и любой точке $x \geq 0$.

а) Если $x=0$, то приращение $\Delta x > 0$ (в силу области определения функции $y=\sqrt{x}$), поэтому $\Delta y = \sqrt{0+\Delta x} - \sqrt{0} = \sqrt{\Delta x}$.

б) Если $x=1$, то $\Delta y = \sqrt{1+\Delta x} - \sqrt{1} = \Delta x / \sqrt{1+\Delta x} + \sqrt{1}$, где $|\Delta x| < 1$.

в) Если x — любая точка $(x > 0)$, то $\Delta y = \sqrt{x+\Delta x} - \sqrt{x} = \Delta x / (\sqrt{x+\Delta x} + \sqrt{x})$, где $|\Delta x| < x$.

Пусть функция $y=f(x)$ определена в некоторой окрестности точки x. Дадим аргументу x приращение Δx (при этом предполагается, что точка $x+\Delta x$ принадлежит области определения функции). Тогда функция получит приращение $\Delta y = f(x+\Delta x) - f(x)$.

Производной функции $y=f(x)$ в точке x называется предел отношения приращения функции к приращению аргумента в этой точке при стремлении приращения аргумента к нулю, если этот предел существует и конечен.

Производная функции $y=f(x)$ в точке x обозначается y' (читается „игрек штрих"), или $\frac{dy}{dx}$ (читается „де игрек по де икс"), или $\frac{df(x)}{dx}$ (читается „де эф по де икс"). Нахождение производной от функции называется дифференцированием. Итак, по определению $y'=\lim\limits_{\Delta x\to 0}(\Delta y/\Delta x)$, т. е. $f'(x)=\lim\limits_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$.

Из определения следует, что производная функции $y=f(x)$ в точке x есть число, зависящее от рассматриваемого значения x, но не зависящее от Δx. Кроме того, производная функции $y=f(x)$ может существовать не во всех точках области определения этой функции. Рассмотрим множество M всех точек области определения функции $y=f(x)$, в которых эта функция имеет производную. Вычисляя производную функции $y=f(x)$ в любой точке $x_0\in M$, получаем, что каждому числу $x_0\in M$ ставится в соответствие одно число $f'(x_0)$. Другими словами, этим соответствием задана функция, обозначаемая y' или $f'(x)$ с областью определения — множеством M.

Найдем производные некоторых элементарных функций.

1. Пусть на некотором интервале (a, b) функция $y=f(x)$ имеет постоянное значение c, т. е. $y=c$ на (a, b). Тогда $y'=(c)'=0$ для любого x из (a, b).

Действительно, для любого x из (a, b): $\Delta y=c-c=0$, следовательно, $\Delta y/\Delta x=0$, откуда $y'=\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=0$.

2. Пусть $y=x$, тогда $y'=1$ для любого x. Действительно, для любого x: $\Delta y=(x+\Delta x)-x=\Delta x$,

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

3. Пусть $y=x^n$ (n — натуральное число), тогда $y'=nx^{n-1}$ для любого x. Действительно, для любого x по формуле бинома Ньютона (см. гл. II)

$$\Delta y=(x+\Delta x)^n-x^n=nx^{n-1}\Delta x+\frac{n(n-1)}{1\cdot 2}x^{n-2}(\Delta x)^2+\cdots+(\Delta x)^{n-1},$$

откуда

$$\frac{\Delta y}{\Delta x}=nx^{n-1}+\Delta x\left[\frac{n(n-1)}{1\cdot 2}x^{n-2}+\Delta x\frac{n(n-1)\,n-2}{1\cdot 2\cdot 3}x^{n-3}+\cdots+(\Delta x)^{n-2}\right]$$

и, следовательно, $y'=\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=nx^{n-1}$, т. е. $(x^n)'=nx^{n-1}$.

4. Пусть $y=\sin x$, тогда $y'=\cos x$ для любого x. Действительно, для любого x: $\Delta y=\sin(x+\Delta x)-\sin x=2\sin\left(\frac{\Delta x}{2}\right)\times$

$\cos\times\left(x+\frac{\Delta x}{2}\right)$ и, следовательно, $y'=\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0}\frac{\sin(\Delta x/2)}{(\Delta x/2)}\times$ $\times\lim\limits_{\Delta x\to 0}\cos(x+\Delta x/2)$.

Ранее было доказано, что $\lim\limits_{\Delta x\to 0}\frac{\sin(\Delta x/2)}{\Delta x/2}=1$. Докажем, что $\lim\limits_{\Delta x\to 0}\cos(x+\Delta x/2)=\cos x$, т. е. докажем, что для любого $\varepsilon>0$ найдется такое δ, что $|\cos(x+\Delta x/2)-\cos x|<\varepsilon$, как только $|\Delta x|<\delta$. Для доказательства возьмем $\delta=2\varepsilon$ и получим следующую цепочку неравенств:

$$|\cos(x+\Delta x/2)-\cos x|=|2\sin(x+\Delta x/4)\sin(\Delta x/4)|\leqslant 2\frac{\Delta x}{4}<\varepsilon,$$

следовательно, $(\sin x)'=\cos x$.

5. Пусть $y=\cos x$, тогда $y'=-\sin x$ для любого x. Действительно, для любого x

$$\Delta y=\cos(x+\Delta x)-\cos x=-2\sin(\Delta x/2)\sin(x+\Delta x/2).$$

Тогда

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=-\lim_{\Delta x\to 0}\frac{\sin(\Delta x/2)}{\Delta x/2}\cdot\lim_{\Delta x\to 0}\sin(x+\Delta x/2),$$

и поскольку $\lim\limits_{\Delta x\to 0}\frac{\sin(\Delta x/2)}{\Delta x/2}=1$ и $\lim\limits_{\Delta x\to 0}\sin(x+\Delta x/2)=\sin x$ (доказательство такое же, как и в предыдущем примере), то $y'=$ $=(\cos x)'=-\sin x$.

Если нужно найти производную функции $y=f(x)$ в некоторой фиксированной точке $x=x_0$, то обычно сначала находят производную этой функции в любой точке x, где она существует, т. е. находят $y'=f'(x)$, а затем вместо произвольного x подставляют $x=x_0$. Производную в фиксированной точке x_0 обозначают $f'(x_0)$, $y'(x_0)$, $y'|_{x=x_0}$ или $\frac{df(x_0)}{dx}$.

Пример. Найти производную функции $y=\sin x$ в точках $x_1=\pi/4$, $x_2=\pi$, $x_3=-\pi/2$.

Так как $y'=\cos x$, то $y'|_{x=\pi/4}=\cos(\pi/4)=\sqrt{2}/2$ т. е. $y'(\pi/4)=$ $=\sqrt{2}/2$; $y'|_{x=\pi}=\cos\pi=-1$, т. е. $y'(\pi)=-1$; $y'|_{x=-\pi/2}=$ $=\cos(-\pi/2)=0$, т. е. $y'(-\pi/2)=0$.

Решение примеров на нахождение производных значительно упрощается, если использовать правила дифференцирования, которые вытекают из теорем для пределов. Рассмотрим некоторые из них.

В дальнейшем предполагается, что аргумент x изменяется в общей части областей определения функций, участвующих в следующих утверждениях.

1. Если в точке x существует конечная производная функции $u=b(x)$ и функции $v=\varphi(x)$, то в точке x производная суммы этих функций существует и равна сумме производных этих функций, т. е. $(u+v)'=u'+v'$.

Действительно, дадим x приращение Δx. Тогда функции u и v получат приращения, соответственно равные Δu и Δv, и функция $y=(u+v)$ получит приращение $\Delta y=[(u+\Delta u)+(v+\Delta v)]-(u+v)=\Delta u+\Delta v$. Поскольку $\frac{\Delta y}{\Delta x}=\frac{\Delta u}{\Delta x}+\frac{\Delta v}{\Delta x}$, то $y'=\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0}\left(\frac{\Delta u}{\Delta x}+\frac{\Delta v}{\Delta x}\right)=u'+v'$, так как предел каждого слагаемого существует и конечен.

Пример. $(x^2+\sin x)'=(x^2)'+(\sin x)'=2x+\cos x$.

2. Если в точке x существует конечная производная функции $u=f(x)$ и функции $v=\varphi(x)$, то в точке x производная разности этих функций существует и равна разности производных этих функций, т. е. $(u-v)'=u'-v'$.

Доказательство утверждения 2 аналогично доказательству предыдущего утверждения.

Пример. $(x^4-\sin x)'=(x^4)'-(\sin x)'=4x^3-\cos x$. Используя формулы нахождения производной в точке для суммы и разности двух слагаемых, легко получить формулу для нахождения производной в точке для алгебраической суммы любого числа слагаемых. Напишем, например, эти формулы для алгебраической суммы пяти слагаемых: $(u+v-p+g-t)'=u'+v'-p'+g'-t'$.

Пример. $(x+x^2-x^7+\sin x-\cos x)'=1+2x-7x^6+\cos x+\sin x$.

3. Если в точке x существует конечная производная функции $u=f(x)$ и функции $v=\varphi(x)$, то в этой точке существует производная функции $y=uv$, при этом $y'=(uv)'=u'v+v'u$.

Действительно, дадим x приращение Δx. Тогда функции u, v и $y=uv$ получат приращения, при этом $\Delta y=(u+\Delta u)(v+\Delta v)-uv=v\,\Delta u+u\,\Delta v+\Delta u\,\Delta v$ и

$$\frac{\Delta y}{\Delta x}=\frac{\Delta u}{\Delta x}v+\frac{\Delta v}{\Delta x}u+\frac{\Delta u}{\Delta x}\Delta v.$$

Применяя теоремы о пределе суммы и произведения, имеем

$$y'=\left(\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}\right)v+\left(\lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x}\right)u+\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}\lim_{\Delta x\to 0}\Delta v.$$

Учитывая, что $\lim\limits_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=u'$, $\lim\limits_{\Delta x\to 0}\frac{\Delta v}{\Delta x}=v'$, $\lim\limits_{\Delta x\to 0}\Delta v=0$, получим $y'=(uv)'=u'v+v'u$, что и требовалось доказать.

В частности, если $v=c$ (c — константа), то $(cu)'=c'u+cu'=cu'$, т. е. постоянный множитель можно выписать за знак производной.

Пример. $y=5x^3+7x^2-4$.

$$y'=(5x^3)'+(7x^2)'-(4)'=5(x^3)'+7(x^2)'-0=$$
$$=15x^2+14x.$$

Используя формулу нахождения производной в точке x для двух сомножителей, легко получить формулу для случая произведения нескольких сомножителей.

4. Если в точке x существуют конечные производные функций $u=f(x)$ и $v=\varphi(x)$ и функция $v=\varphi(x)$ отлична от нуля

в этой точке, то в этой точке существует и производная функции $y=u/v$, при этом $y'=(u/v)'=(u'v-v'u)/v^2$.

Действительно, дадим x приращение Δx. Тогда функции u, v и $y=u/v$ получат приращения, при этом

$$\Delta y=\frac{u+\Delta u}{v+\Delta v}-\frac{u}{v}=\frac{v\,\Delta u-u\,\Delta v}{v\,(v+\Delta v)}$$

и

$$y'=\lim_{\Delta x\to 0}(\Delta y/\Delta x)=\lim_{\Delta x\to 0}\frac{v\dfrac{\Delta u}{\Delta x}-u\dfrac{\Delta v}{\Delta x}}{v\,(v+\Delta v)}.$$

Применяя теоремы о пределах частного и произведения и учитывая непрерывность функции v в точке x, получим, что $y'=\dfrac{vu'-uv'}{v^2}$.

Пример. $(\operatorname{tg}x)'=\left(\dfrac{\sin x}{\cos x}\right)'=\dfrac{(\sin x)'\cos x-(\cos x)'\sin x}{\cos^2 x}=\dfrac{\cos^2 x+\sin^2 x}{\cos^2 x}=\dfrac{1}{\cos^2 x}$; $(\operatorname{ctg}x)'=\left(\dfrac{\cos x}{\sin x}\right)'=\dfrac{(\cos x)'\sin x-(\sin x)'\cos x}{\sin^2 x}=\dfrac{-\cos^2 x-\sin^2 x}{\sin^2 x}=-\dfrac{1}{\sin^2 x}$.

Используя формулы производных основных элементарных функций и правила дифференцирования, можно найти производную любой функции, которая получается суперпозицией основных элементарных функций.

Приведем таблицу производных элементарных функций

$y=f(x)$	$y'=f'(x)$	$y=f(x)$	$y'=f'(x)$
$y=c$ ($c=\text{const}$)	$y'=0$	$y=\ln x$	$y'=\dfrac{1}{x}$
$y=x$	$y'=1$	$y=\sin x$	$y'=\cos x$
$y=x^2$	$y'=2x$	$y=\cos x$	$y'=-\sin x$
$y=x^n$ (n — натуральное число)	$y'=nx^{n-1}$	$y=\operatorname{tg}x$	$y'=\dfrac{1}{\cos^2 x}$
$y=x^{-n}$ (n — натуральное число)	$y'=-nx^{-n-1}$	$y=\operatorname{ctg}x$	$y'=-\dfrac{1}{\sin^2 x}$
$y=x^{\alpha}$ ($\alpha>0$)	$y'=\alpha x^{\alpha-1}$	$y=\arcsin x$	$y'=\dfrac{1}{\sqrt{1-x^2}}$
$y=x^{-\alpha}$ ($\alpha>0$)	$y'=-\alpha x^{-\alpha-1}$	$y=\arccos x$	$y'=\dfrac{-1}{\sqrt{1-x^2}}$
$y=a^x$ ($a>0$, $a\neq 1$)	$y'=a^x\ln a$	$y=\operatorname{arctg}x$	$y'=\dfrac{1}{1+x^2}$
$y=e^x$	$y'=e^x$	$y=\operatorname{arcctg}x$	$y'=\dfrac{-1}{1+x^2}$
$y=\log_a x$ ($a>0$, $a\neq 1$)	$y'=\dfrac{1}{x\ln a}$		

Глава IX

СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

В § 3 гл. III уже рассматривались системы алгебраических уравнений с несколькими переменными. В этой главе рассматривается общая теория решения систем линейных уравнений с несколькими переменными.

Отметим, что алгебраическое уравнение $P(x, y, \ldots, t)=0$ (см. гл. II) называется *линейным*, если $P(x, y, \ldots, t)$ есть многочлен первой степени, целый относительно переменных $x, y, \ldots, t$.

§ 1
МАТРИЦЫ

Матрицей размеров $m\times n$ называется совокупность mn чисел, расположенных в виде прямоугольной таблицы из m строк и n столбцов.

Например, прямоугольная таблица чисел $\begin{pmatrix} 1 & 3 & \sqrt{2} \\ -5 & 7 & 1 \end{pmatrix}$ есть матрица размеров 2×3; прямоугольная таблица чисел $\begin{pmatrix} 0 & 4 \\ -2 & 3/4 \\ 0 & 7 \end{pmatrix}$ есть матрица размеров 3×2.

Матрица размеров $1\times n$, т. е, состоящая из одной строки, называется *матрицей-строкой*. Например, матрица $(1\ 3\ 4\ 5)$ есть матрица-строка.

Матрица размеров $m\times 1$, т. е. матрица, состоящая из одного столбца, называется *матрицей-столбцом*. Например, матрица $\begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$ есть матрица-столбец.

Матрица размеров $n\times n$ называется *квадратной матрицей*, а число n называется ее порядком. Например, матрица $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 7 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ есть квадратная матрица третьего порядка.

Квадратная матрица порядка 1 есть просто число. Числа, из которых состоит матрица, называются *элементами матрицы*.

Для записи матрицы в общем виде элементы матрицы обозначаются буквами с двумя индексами, например a_{ij}; при этом первый индекс указывает номер строки, а второй индекс — номер столбца, в котором содержится этот элемент.

Например, матрица размеров 3×4 записывается в общем виде так:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}. \quad (1)$$

Часто матрицу, элементами которой являются, например, числа b_{ij}, обозначают одной заглавной буквой B и для записи этого факта употребляют знак равенства. Например, матрицу с элементами a_{ij} принято обозначать буквой A и записывать этот факт так:

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

Пусть дана матрица C размером $m\times n$:

$$C=\begin{pmatrix} c_{11} & c_{12} & c_{13}\ldots c_{1n} \\ c_{21} & c_{22} & c_{23}\ldots c_{2n} \\ c_{31} & c_{32} & c_{33}\ldots c_{3n} \\ \cdot & \cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ c_{m1} & c_{m2} & c_{m3}\ldots c_{mn} \end{pmatrix}.$$

Индексы i и j называются *допустимыми* для матрицы размеров $m\times n$, если i — индекс строки есть любое из чисел $1, 2, \ldots, m$, j — индекс столбца есть любое из чисел $1, 2, \ldots n$.

Часто у матрицы-столбца и матрицы-строки их элементы обозначаются одним индексом. Например,

$$X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},\; F=(f_1,\; f_2,\; f_3,\; f_4).$$

Матрицы A и B, каждая из которых размеров $m\times n$, называются *равными*, если равны их соответствующие элементы, т. е. если $a_{ij}=b_{ij}$ для любых допустимых индексов i и j. В таких случаях пишут $A=B$. Например, $\begin{pmatrix} 3 & 4 & \sqrt{2} \\ 0 & 1 & 2 \end{pmatrix}=\begin{pmatrix} \sqrt{9} & 4 & {}^2/_{\sqrt{2}} \\ 0 & 1 & {}^{10}/_5 \end{pmatrix}$.

Для любых матриц знаки сравнения $>$, $\geqslant$, $<$ или $\leqslant$ лишены смысла, а для матриц разных размеров, кроме того, лишен смысла знак равенства.

Суммой матриц A и B, каждая из которых размеров $m\times n$, называется матрица C тех же размеров $m\times n$, элементы которой равны суммам соответствующих элементов данных матриц, т. е. $C=A+B$, если $c_{ij}=a_{ij}+b_{ij}$ для любых допустимых индексов i и j. Например, $\begin{pmatrix} 2 & 3 \\ -1 & 1/_2 \\ 0 & 4 \end{pmatrix}+\begin{pmatrix} 1 & -2 \\ 0 & 3/_2 \\ 1 & -1 \end{pmatrix}=\begin{pmatrix} 3 & 1 \\ -1 & 2 \\ 1 & 3 \end{pmatrix}$.

Для матриц A и B с разными размерами операция сложения лишена смысла.

Легко видеть, что для любых матриц A, B, C, каждая из которых размеров $m\times n$, справедливы следующие утверждения:

а) сложение матриц коммутативно: $A+B=B+A$;

б) сложение матриц ассоциативно: $A+(B+C)=(A+B)+C$.

Из этих утверждений следует, что в любой сумме конечного числа матриц, каждая из которых размеров $m\times n$, слагаемые можно писать в любом порядке и скобки, указывающие порядок сложения, можно расставлять произвольно. Например, $[A+(C+B)]+D=[(A+C)+B]+D=A+B+C+D$. Матрица размеров $m\times n$, у которой каждый элемент равен нулю:

$$0=\begin{pmatrix} 0 & 0 & 0 \ldots 0 \\ 0 & 0 & 0 \ldots 0 \\ 0 & 0 & 0 \ldots 0, \\ \cdot & \cdot & \cdot \;\cdot\; \cdot \\ 0 & 0 & 0 \ldots 0 \end{pmatrix}$$

называется *нулевой матрицей* размеров $m\times n$ или *нуль-матрицей размеров* $m\times n$. Эта матрица обладает тем свойством, что для каждой матрицы A размеров $m\times n$ справедливы равенства $A+0=0+A=A$.

Например, матрица $0=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ есть нуль-матрица размеров 2×3, матрица $0=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ есть квадратная нуль-матрица второго порядка. Среди всех матриц, каждая из которых размеров $m\times n$, существует единственная нуль-матрица. Нуль-матрицы разных размеров принято обозначать одним и тем же символом 0, что не приводит к недоразумениям, ибо из контекста ясно, какие размеры имеет нуль-матрица в рассматриваемом случае.

Пусть дана матрица A размеров $m\times n$. Матрица $(-A)$, тех же размеров, каждый элемент которой есть элемент матрицы

A, взятый с противоположным знаком, обладает тем свойством, что $A+(-A)=(-A)+A=0$. Матрица $(-A)$ называется *противоположной матрицей к матрице* A. Например, матрица $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ противоположная к матрице $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$; матрица $\begin{pmatrix} -5 & -6 & 1 \\ 0 & 1 & -2 \end{pmatrix}$ противоположная к матрице $\begin{pmatrix} 5 & 6 & -1 \\ 0 & -1 & 2 \end{pmatrix}$.

Можно показать, что для любой матрицы A существует единственная противоположная матрица, т. е. для матрицы A только матрица $(-A)$ такова, что $A+(-A)=(-A)+A=0$.

Сложение матриц обладает обратной операцией — вычитанием. Это значит, что для любых двух матриц A и B, каждая из которых размеров $m\times n$, существует единственная матрица C тех же размеров и такая, что $A+C=B$. Матрица C называется *разностью* матриц A и B и обозначается $A-B$. Например,

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

Произведением матрицы A размеров $m\times n$ *на некоторое число* α называется матрица C тех же размеров $m\times n$, элементы которой получаются из соответствующих элементов матрицы A умножением на это число α, т. е. $C=\alpha A$, если $c_{ij}=\alpha a_{ij}$ для любых допустимых индексов i и j. Например, $2\begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 2 & 4 \end{pmatrix}$; $-3\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$. Из определения умножения матрицы на число следует, что для любых матриц A и B, каждая из которых размеров $m\times n$, и любых чисел α и β справедливы равенства:

а) $(\alpha+\beta)A=\alpha A+\beta A$;

б) $(\alpha\beta)A=\alpha(\beta A)$;

в) $\alpha(A+B)=\alpha A+\alpha B$.

Заметим, что противоположная матрица $(-A)$ к матрице A равна $(-1)A$, т. е. $(-A)=(-1)A$.

Пусть дана матрица-строка F размеров $1\times r$, т. е. $F=(f_1, f_2, \dots f_r)$, и матрица-столбец X размеров $r\times 1$, т. е. $X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix}$. *Произведением матрицы-строки* F *размеров* $1\times r$ *на матрицу-столбец* X *размеров* $r\times 1$ называется матрица, обозначаемая FX, размеров 1×1, т. е. число, равное сумме произведений соответствующих элементов этих матриц, т. е. по определению $FX=f_1x_1+f_2x_2+f_3x_3+\dots+f_{r-1}x_{r-1}+f_rx_r$.

Пример. $(1\ \ 2)\begin{pmatrix}-4\\ 2\end{pmatrix}=1(-4)+2\cdot 2=0;$

$$(^1/_5\ ^1/_5\ 0\ 1\ 3)\begin{pmatrix}2\\ 3\\ -7\\ 1\\ -1\end{pmatrix}=\frac{1}{5}\cdot 2+\frac{1}{5}\cdot 3+0(-7)+1\cdot 1+3(-1)=-1.$$

С помощью произведения матрицы-строки размеров $1\times r$ и матрицы-столбца размеров $r\times 1$ определяется произведение матрицы A размеров $m\times r$ и матрицы B размеров $r\times n$, т. е. таких матриц, что первый сомножитель, матрица A, имеет столько столбцов, сколько второй сомножитель, матрица B, имеет строк.

По определению *произведением матрицы A размеров $m\times r$ на матрицу B размеров $r\times n$* называется матрица C размеров $m\times n$, обозначаемая AB, у которой элемент c_{ij} равен произведению i-й строки первого сомножителя, матрицы A, на j-й столбец второго сомножителя, матрицы B, т. е. $C=AB$, если $c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{ir\ 1}b_{r-1j}+a_{ir}b_{rj}$ для любых допустимых индексов i и j матрицы C.

Пример. $(3\ \ 2\ \ -1)\begin{pmatrix}0&1\\ 1&0\\ 1&1\end{pmatrix}=(3\cdot 0+2\cdot 1-1\cdot 1\ \ 3\cdot 1+2\cdot 0-1\cdot 1)=$
$=(1\ \ 2).$

$$\begin{pmatrix}0&1\\ 1&0\\ 1&1\end{pmatrix}\begin{pmatrix}3\\ -3\end{pmatrix}=\begin{pmatrix}0\cdot 3+1\cdot(-3)\\ 1\cdot 3+0\cdot(-3)\\ 1\cdot 3+1\cdot(-3)\end{pmatrix}=\begin{pmatrix}-3\\ 3\\ 0\end{pmatrix}.$$

$$\begin{pmatrix}3&1&4\\ 2&0&5\end{pmatrix}\begin{pmatrix}1&3&0&0\\ 1&2&0&0\\ 0&0&1&1\end{pmatrix}=$$
$$=\begin{pmatrix}3\cdot 1+1\cdot 1+4\cdot 0 & 3\cdot 3+1\cdot 1+4\cdot 0 & 3\cdot 0+1\cdot 0+4\cdot 1 & 3\cdot 0+1\cdot 0+4\cdot 1\\ 2\cdot 1+0\cdot 1+5\cdot 0 & 2\cdot 3+0\cdot 1+5\cdot 0 & 2\cdot 0+0\cdot 0+5\cdot 1 & 2\cdot 0+0\cdot 0+5\cdot 1\end{pmatrix}=$$
$$=\begin{pmatrix}4&10&4&4\\ 2&6&5&5\end{pmatrix}.$$

Итак, операция умножения двух прямоугольных матриц выполнима только в том случае, когда число столбцов в первом сомножителе равно числу строк во втором множителе, в остальных случаях произведение матриц не определяется. В частности, если матрицы A и B — квадратные матрицы одного и того же порядка, то умножение матриц всегда выполнимо.

Однако заметим, что даже в этом частном случае не всегда $AB\neq BA$. Например, пусть $A=\begin{pmatrix}1&0\\ 1&0\end{pmatrix}$, $B=\begin{pmatrix}0&1\\ 0&1\end{pmatrix}$, тогда $AB=$

$$=\begin{pmatrix}1&0\\1&0\end{pmatrix}\begin{pmatrix}0&1\\0&1\end{pmatrix}=\begin{pmatrix}1\cdot 0+0\cdot 0 & 1\cdot 1+0\cdot 1\\1\cdot 0+0\cdot 0 & 1\cdot 1+0\cdot 1\end{pmatrix}=\begin{pmatrix}0&1\\0&1\end{pmatrix},\ BA=\begin{pmatrix}0&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\1&0\end{pmatrix}=$$

$$=\begin{pmatrix}0\cdot 1+1\cdot 1 & 0\cdot 0+1\cdot 0\\0\cdot 1+1\cdot 1 & 0\cdot 0+1\cdot 0\end{pmatrix}=\begin{pmatrix}1&0\\1&0\end{pmatrix},\ \text{т. е. }\ AB\neq BA.$$

Итак, умножение матриц некоммутативно, т. е., в общем случае $AB \neq BA$. Поэтому при перемножении матриц во избежание ошибок следует строго придерживаться заданного порядка следования множителей.

Если $AB = BA$, то матрицы A и B называются *перестановочными,* или *коммутирующими*, между собой.

Например, матрицы $A=\begin{pmatrix}1&2\\-2&0\end{pmatrix}$ и $B=\begin{pmatrix}-3&2\\-2&-4\end{pmatrix}$ перестановочны между собой, так как $AB=\begin{pmatrix}-7&-6\\6&-4\end{pmatrix}$, $BA=\begin{pmatrix}-7&-6\\6&-4\end{pmatrix}$, т. е. $AB = BA$.

Операция умножения матриц является *ассоциативной,* а также *дистрибутивной относительно сложения*, т. е. для любых прямоугольных матриц A, B, C, для которых имеют смысл соответствующие произведения, справедливы равенства:

а) $(AB)\,C = A\,(BC)$;

б) $(A+B)\,C = AC+BC,\ C\,(A+B) = CA+CB.$

Операция умножения матриц соответствующим образом распространяется на случай нескольких сомножителей. В силу определения произведения матриц умножать матрицу A на себя можно только в том случае, если это квадратная матрица. Пусть дана квадратная матрица A порядка n и натуральное число $p > 1$.

Матрица $\underbrace{A\ A\ldots A}_{p\text{ раз}}$ называется p-й степенью матрицы A и обозначает A^p, т. е. для $p > 1$ по определению $A^p = \underbrace{AA\ldots A}_{p\text{ раз}}.$

Кроме того, по определению $A^1 = A$.

Пример. Найти куб матрицы $A=\begin{pmatrix}1&0\\-1&1\end{pmatrix}$.

$$A^2 = AA = \begin{pmatrix}1&0\\-1&1\end{pmatrix}\begin{pmatrix}1&0\\-1&1\end{pmatrix}=\begin{pmatrix}1&0\\-2&1\end{pmatrix};$$

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

Из ассоциативного свойства умножения матриц следует, что $A^pA^q = A^{p+q}$, где p и q — произвольные натуральные числа.

Отметим, что для любой матрицы A размеров $m\times r$, произведение матрицы A на нуль-матрицу 0 размеров $r\times n$ есть нуль-матрица размеров $m\times n$, и произведение нуль-матрицы 0 размеров $l \times p$ на матрицу A размеров $p \times k$ есть нуль-матрица 0 размеров $l \times k$. Например, $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$; $\begin{pmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, т. е. в тех случаях, когда произведение матриц имеет смысл, для любой матрицы A справедливы равенства $A0 = 0$ и $0A = 0$. В частности, для любой квадратной матрицы A порядка n и квадратной нуль-матрицы 0 порядка n: $A0 = 0A = 0$. Известно, что произведение двух чисел равно нулю тогда и только тогда, когда хотя бы один из сомножителей равен нулю. В отличие от чисел произведение двух матриц может быть равно нулю даже тогда, когда каждый из сомножителей не является нуль-матрицей, т. е. существуют матрицы $A \neq 0$ и $B \neq 0$, такие, что $AB = 0$.

Например, матрицы $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ и $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ таковы, что произведение AB есть нуль-матрица. Действительно $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 0 \cdot 1 & 1 \cdot 0 + 0 \cdot 1 \\ 0 \cdot 0 + 0 \cdot 1 & 0 \cdot 0 + 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$ Заметим, что деление матриц не рассматривается, т. е. символ ${}^{B}/_{A}$ для матриц не существует.

Квадратная матрица порядка n

$$E = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

у которой элементы $a_{ii} = 1$, для всех $i = 1, 2, \dots, n$, а остальные элементы равны нулю, обладает следующим свойством $EA = A$ для каждой матрицы A размеров $n \times m$ и $AE = A$ для каждой матрицы A размеров $m \times n$. В частности $AE = EA = A$ для каждой квадратной матрицы A порядка n, $FE = F$ для каждой матрицы-строки F размеров $1 \times n$, $EX = X$ для каждой матрицы-столбца X размеров $n \times 1$. Матрица E в произведении матриц играет ту же роль, что число 1 в произведении чисел, поэтому матрицу E принято называть *единичной матрицей* порядка n.

Транспонированием матрицы A называется перемена ролями строк и столбцов с сохранением их номеров. Таким образом, строки данной матрицы A будут в той же последовательности столбцами транспонированной матрицы, обозначаемой A^T, а столбцы матрицы A будут в той же последовательности строками транспонированной матрицы A^T.

Пример. Если $A=\begin{pmatrix} 1 & 4 & 3 \\ -1 & 5 & 6 \end{pmatrix}$, то $A^T=\begin{pmatrix} 1 & -1 \\ 4 & 5 \\ 3 & 6 \end{pmatrix}$. Ясно, что при транспонировании матрица A размеров $m\times n$ переходит в матрицу A^T размеров $n\times m$. Если A — квадратная матрица порядка n, то и транспонированная матрица A^T является квадратной матрицей порядка n. Заметим, что единичная матрица не меняется при транспонировании. Легко проверяется, что транспонирование матриц обладает следующими свойствами:

а) $(A^T)^T=A$;

б) $(A+B)^T=A^T+B^T$, $(A-B)^T=A^T-B^T$;

в) $(\lambda A)^T=\lambda A^T$ (λ — число);

г) $(AB)^T=B^TA^T$.

§ 2
ОПРЕДЕЛИТЕЛИ

Пусть дана квадратная матрица второго порядка

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{1}$$

Определителем второго порядка, соответствующим матрице (1), называют число $a_{11}a_{22}-a_{12}a_{21}$. Обозначается этот определитель символом

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \tag{2}$$

или одной буквой Δ, или символом $|A|$. Таким образом, по определению

$$|A|=\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}. \tag{3}$$

Например, $\begin{vmatrix} 2 & 4 \\ -1 & -5 \end{vmatrix}=2(-5)-(-1)\cdot 4=-6.$

Элементы матрицы (1) принято называть *элементами определителя* (2), строки и столбцы матрицы (1) — *строками и столбцами определителя* (2).

Рассмотрим простейшие свойства определителей второго порядка.

1. Определитель квадратной матрицы (1) равен определителю транспонированной матрицы.

Действительно, пусть

$$\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta^T=\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}.$$

Так как $\Delta=a_{11}a_{22}-a_{21}a_{12}$, а $\Delta^T=a_{11}a_{22}-a_{12}a_{21}$, то $\Delta=\Delta^T$.

Замена определителя матрицы (1) на определитель транспонированной матрицы называется *транспонированием определителя* (2). Поэтому свойство 1 можно сформулировать так: определитель (2) не меняется при его транспонировании.

Замечание. Так как по свойству 1 строки определителя можно, не меняя его значения, заменить столбцами, то, если доказана справедливость некоторого утверждения для столбцов определителя, то тем самым будет доказана справедливость этого утверждения и для строк. Исходя из этого рассматриваемые ниже свойства определителей будут устанавливаться только для столбцов определителя.

2. При перестановке столбцов (строк) определитель меняет только знак.

Действительно, пусть $\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $\Delta_1=\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix}$.

Тогда $\Delta=a_{11}a_{22}-a_{21}a_{12}$, $\Delta_1=a_{12}a_{21}-a_{22}a_{11}=-(a_{11}a_{22}-a_{21}a_{12})$, т. е. $\Delta_1=-\Delta$.

3. Если все элементы хотя бы одного столбца (строки) определителя равны нулю, то определитель равен нулю.

Действительно, если $\Delta=\begin{vmatrix} 0 & a_{12} \\ 0 & a_{22} \end{vmatrix}$, то $\Delta=0\cdot a_{22}-0\cdot a_{12}=0$.

Если $\Delta_1=\begin{vmatrix} a_{11} & 0 \\ a_{21} & 0 \end{vmatrix}$, то по свойству 2 $\Delta_1=-\Delta$ и поэтому $\Delta_1=0$.

4. Определитель, имеющий два одинаковых столбца (строки), равен нулю.

Действительно, если $\Delta=\begin{vmatrix} a_{11} & a_{11} \\ a_{21} & a_{21} \end{vmatrix}$, то $\Delta=a_{11}a_{21}-a_{21}a_{11}=0$.

Если $\Delta_1=\begin{vmatrix} a_{12} & a_{12} \\ a_{22} & a_{22} \end{vmatrix}$, то $\Delta_1=a_{12}a_{22}-a_{22}a_{12}=0$.

5. Если все элементы некоторого столбца (строки) определителя умножить на одно и то же число α, то определитель умножится на это число.

Действительно, пусть например, $\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $\Delta_1=\begin{vmatrix} a_{11} & \alpha a_{12} \\ a_{21} & \alpha a_{22} \end{vmatrix}$.

Тогда $\Delta_1=a_{11}(\alpha a_{22})-a_{21}(\alpha a_{12})=\alpha(a_{11}a_{22}-a_{21}a_{12})=\alpha\Delta$.

Свойство 5 иногда формулируется следующим образом: если все элементы какого-либо столбца (строки) определителя содержат общий множитель α, то его можно вынести за знак определителя. Так например, по доказанному $\begin{vmatrix} a_{11} & \alpha a_{21} \\ a_{21} & \alpha a_{22} \end{vmatrix}=\alpha\begin{vmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{vmatrix}$.

Свойство 5 выражает правило умножения определителя на некоторое число.

6. Определитель, у которого элементы двух столбцов (строк) соответственно пропорциональны, равен нулю.

Действительно, пусть $\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ и пусть, например, $a_{11}=\beta a_{12}$, $a_{21}=\beta a_{22}$. Тогда на основании свойств 5 и 4 определителя, имеем

$$\Delta=\begin{vmatrix} \beta a_{12} & a_{12} \\ \beta a_{22} & a_{22} \end{vmatrix}=\beta\begin{vmatrix} a_{12} & a_{12} \\ a_{22} & a_{22} \end{vmatrix}=0.$$

7. Если каждый элемент какого-либо столбца (строки) определителя есть сумма двух слагаемых, то определитель равен сумме двух определителей, у одного из которых элементами соответствующего столбца (строки) являются первые слагаемые, у другого — вторые, а остальные элементы этих двух определителей те же, что и у данного.

Действительно, пусть например, $\Delta=\begin{vmatrix} a_{11}+b_{11} & a_{12} \\ a_{21}+b_{21} & a_{22} \end{vmatrix}$, $\Delta_1=\begin{vmatrix} a_{11} & a_{22} \\ a_{21} & a_{22} \end{vmatrix}$, $\Delta_2=\begin{vmatrix} b_{11} & a_{12} \\ b_{21} & a_{22} \end{vmatrix}$. Тогда $\Delta=(a_{11}+b_{11})a_{22}-(a_{21}+b_{21})a_{12}=(a_{11}a_{22}-a_{21}a_{12})+(b_{11}a_{22}-b_{21}a_{12})=\Delta_1+\Delta_2$. Свойство 7 выражает правило сложения определителей.

8. Определитель не изменится, если к элементам какого-либо столбца (строки) прибавить соответствующие элементы другого столбца (строки), умноженные на одно и то же число.

Действительно, пусть, например,

$$\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\quad \Delta_1=\begin{vmatrix} a_{11}+\alpha a_{12} & a_{12} \\ a_{21}+\alpha a_{22} & a_{22} \end{vmatrix}.$$

Покажем, что $\Delta=\Delta_1$. На основании свойств 7 и 6 определителей имеем $\Delta_1=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}+\begin{vmatrix} \alpha a_{12} & a_{12} \\ \alpha a_{22} & a_{22} \end{vmatrix}=\Delta+0=\Delta$.

Пусть дана матрица третьего порядка

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{4}$$

Определителем третьего порядка, соответствующим матрице (4), называется число

$$a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}+a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}. \tag{5}$$

Обозначается этот определитель символом

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \qquad (6)$$

или одной буквой Δ, или символом $|A|$. Таким образом, по определению

$$|A| = \Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}. \qquad (7)$$

Например, $\begin{vmatrix} 2 & 4 & -1 \\ 0 & 1 & 0 \\ 2 & -3 & 1 \end{vmatrix} = 2 \begin{vmatrix} 1 & 0 \\ -3 & 1 \end{vmatrix} - 0 \begin{vmatrix} 4 & -1 \\ -3 & 1 \end{vmatrix} + 2 \begin{vmatrix} 4 & -1 \\ 1 & 0 \end{vmatrix} =$

$= 2(1+0) - 0(4-3) + 2(0+1) = 4$. Элементы матрицы (4) называются *элементами определителя* (6), строки и столбцы матрицы (4) — *строками и столбцами определителя* (6).

Для изучения свойств определителя третьего порядка введем некоторые новые понятия.

Минором любого элемента определителя третьего порядка (6) называется определитель второго порядка, соответствующий матрице, полученной из данной матрицы (4) вычеркиванием столбца и строки, в которых содержится взятый элемент. Минор элемента a_{ij} принято обозначать M_{ij}. Например, минор элемента a_{11} обозначают M_{11}, элемента a_{31} — M_{31} и т. д. Таким образом, по определению

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

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

$M_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$. *Алгебраическим дополнением* любого элемента a_{ij} определителя третьего порядка (4) называют минор этого элемента M_{ij}, умноженный на $(-1)^{i+j}$. Алгебраическое дополнение элемента a_{ij} обозначают заглавной буквой того же наименования и с теми же двумя индексами, что и у данного элемента.

Например, алгебраическое дополнение элемента a_{21} обозначают A_{21}, элемента a_{23} — A_{23} и т. д. Таким образом, по определению

$$A_{11}=-(1)^{1+1}M_{11}=(-1)^{1+1}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix},$$

$$A_{21}=(-1)^{2+1}M_{21}=(-1)^{2+1}\begin{vmatrix}a_{12} & a_{13}\\ a_{32} & a_{33}\end{vmatrix},$$

$$A_{31}=-(1)^{3+1}M_{31}=(-1)^{3+1}\begin{vmatrix}a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix},$$

$$A_{12}=(-1)^{1+2}M_{12}=(-1)^{1+2}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix},$$

$$A_{23}=(-1)^{2+3}M_{23}=(-1)^{2+3}\begin{vmatrix}a_{11} & a_{12}\\ a_{31} & a_{32}\end{vmatrix},$$

$$A_{33}=(-1)^{3+3}M_{33}=(-1)^{3+3}\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}.$$

Теорема 1. Определитель (4) равен сумме произведений элементов любого его столбца (строки) на соответствующие им алгебраические дополнения, т. е.

$$\begin{aligned}
&\Delta=a_{11}A_{11}+a_{21}A_{21}+a_{31}A_{31}, \quad \Delta=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13},\\
&\Delta=a_{21}A_{21}+a_{22}A_{22}+a_{32}A_{32}, \quad \Delta=a_{21}A_{21}+a_{22}A_{22}+a_{23}A_{23},\\
&\Delta=a_{31}A_{31}+a_{23}A_{23}+a_{33}A_{33}, \quad \Delta=a_{31}A_{31}+a_{32}A_{32}+a_{33}A_{33}.
\end{aligned} \tag{8}$$

Докажем, например, четвертое из равенств (8). Для этого преобразуем правую часть этого равенства

$$a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}=a_{11}(-1)^{1+1}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}+$$
$$+a_{12}(-1)^{1+2}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix}+a_{13}(-1)^{1+3}\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}=$$
$$=a_{11}(a_{22}a_{33}-a_{32}a_{23})-a_{12}(a_{21}a_{33}-a_{31}a_{23})+$$
$$+a_{13}(a_{21}a_{32}-a_{31}a_{22})=a_{11}a_{22}a_{33}-a_{11}a_{32}a_{23}-a_{12}a_{21}a_{33}+$$
$$+a_{12}a_{31}a_{23}+a_{13}a_{21}a_{32}-a_{13}a_{31}a_{22}=a_{11}(a_{22}a_{33}-a_{32}a_{23})-$$
$$-a_{21}(a_{12}a_{33}-a_{32}a_{13})+a_{31}(a_{12}a_{23}-a_{22}a_{13})=$$
$$=a_{11}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}-a_{21}\begin{vmatrix}a_{12} & a_{13}\\ a_{32} & a_{33}\end{vmatrix}+a_{31}\begin{vmatrix}a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}.$$

Сравнивая полученное выражение с выражением (5) определителя Δ, заключаем, что $a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}=\Delta$. Аналогично устанавливается справедливость остальных равенств (8). Запись определителя (6) по любой из указанных формул (8) называется разложением этого определителя по элементам соответствующего столбца или строки. Эти разложения удобно использовать для вычисления определителей третьего порядка. Например, разложив определитель

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

по элементам второй строки, получим

$$\Delta=1(-1)^{2+1}\begin{vmatrix}3 & 1\\ 1 & 2\end{vmatrix}+0(-1)^{2+2}\begin{vmatrix}2 & 1\\ 4 & 2\end{vmatrix}+(-2)(-1)^{2+3}\begin{vmatrix}2 & 3\\ 4 & 1\end{vmatrix}=$$
$$=-(6-1)+2(2-12)=-25.$$

Для определителей третьего порядка остаются справедливыми свойства 1—8, рассмотренные для определителей второго порядка. Доказательство этих свойств непосредственно следует из теоремы 1 о разложении определителя третьего порядка по элементам строки (столбца) и соответствующих свойств определителей второго порядка.

Так, свойство 1, заключающееся в том, что определитель не изменится, если его строки заменить столбцами, можно доказать следующим образом. Пусть $\Delta=\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}$,

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

Раскладывая определитель Δ^T по элементам первой строки, имеем

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

Учитывая, что определитель второго порядка от замены строк столбцами не меняется, получим

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

Сравнивая полученное выражение с выражением (5) определителя Δ, заключаем, что $\Delta^T=\Delta$. Аналогично устанавливаются справедливость свойств 2—8. Свойства 1—8 определителей в сочетании с теоремой 1 о разложении определителя по элементам строк и столбцов часто позволяют упростить вычисления определителя третьего порядка.

Пример. Вычислить определитель $\Delta = \begin{vmatrix} 5 & 20 & 15 \\ 2 & 4 & 8 \\ 1 & 4 & 7 \end{vmatrix}$. По свойству 4 определителей третьего порядка, вынося за знак определителя общий множитель 5 первой строки и общий множитель 4 второго столбца, имеем

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

Прибавим теперь к элементам второго столбца элементы первого столбца, умноженные на -1, а к элементам третьего столбца элементы первого столбца, умноженные на -3. По свойству 7 определитель не изменится и равен

$$\Delta = 20 \begin{vmatrix} 1 & 0 & 0 \\ 2 & -1 & 2 \\ 1 & 3 & 4 \end{vmatrix}.$$

Разложив полученный определитель по элементам первой строки, имеем

$$\Delta = 20 \cdot 1\,(-1)^{1+1} \begin{vmatrix} -1 & 2 \\ 3 & 4 \end{vmatrix} = 20\,(-4-6) = -200.$$

С помощью определителей третьего порядка можно квадратной матрице четвертого порядка поставить в соответствие число — определитель четвертого порядка. Затем с помощью определителей четвертого порядка ввести определители пятого порядка и т. д. Таким образом, для введения определителей n-го порядка надо иметь определитель $(n-1)$-го порядка. Такое введение определителей называется введением определителей по индукции. Приведем теперь введение определителей n-го порядка индукцией по n.

Пусть $n = 1$; определитель квадратной матрицы порядка 1 (т. е. числа a_{11}) считается само число a_{11}. Будем считать, что определители квадратных матриц порядка $(n-1)$ уже введены и доказана справедливость их основных свойств (свойств типа 1—8, доказанных выше для определителей порядка 2). Рассмотрим квадратную матрицу n-го порядка

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}. \qquad (9)$$

Если в этой матрице вычеркнуть i-ю строку и j-й столбец, а порядок элементов оставить прежним, то получится квадратная матрица порядка $(n-1)$

$$\begin{pmatrix} a_{11} & a_{12} & \dots a_{1j-1} & a_{1j+1} & \dots a_{1n} \\ a_{21} & a_{22} & \dots a_{2j-1} & a_{2j+1} & \dots a_{2n} \\ \cdot & \cdot & \cdots \cdot & \cdot & \cdots \cdot \\ a_{i-11} & a_{i-12} & \dots a_{i-1\,j-1} & a_{i-1\,j+1} & \dots a_{i-1\,n} \\ a_{i+11} & a_{i+12} & \dots a_{i+1\,j-1} & a_{i+1\,j+1} & \dots a_{i+1\,n} \\ \cdot & \cdot & \cdots \cdot & \cdot & \cdots \cdot \\ a_{n1} & a_{n2} & \dots a_{nj-1} & a_{nj+1} & \dots a_{nn} \end{pmatrix}.$$

Определитель этой матрицы обозначается M_{ij} и называется минором элемента a_{ij} матрицы (9).

Определителем матрицы (9) называется число

$$a_{11}M_{11}-a_{21}M_{21}+a_{31}M_{31}-a_{41}M_{41}+\ldots+(-1)^{n+1}a_{n1}M_{n1}. \qquad (10)$$

Определитель матрицы (9) обозначается символом

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}, \qquad (11)$$

или одной буквой Δ, или символом $|A|$. Таким образом, по определению

$$|A| = \Delta = \begin{vmatrix} a_{11}, & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} =$$

$$= a_{11}M_{11}-a_{21}M_{21}+\ldots(-1)^{n+1}a_{n1}M_{n1}. \qquad (12)$$

Элементы матрицы (9) называются *элементами определителя* (11), строки и столбцы матрицы (9) называются *строками и столбцами определителя* (11). Миноры элемента a_{ij} матрицы (9) называются *минорами элемента* a_{ij} *определителя* (11). Отметим, что определители второго порядка, введенные по формуле (3) и введенные по формуле (10), будут одинаковыми. Определитель, получаемый из определителя (11) заменой строк на столбцы, а столбцов на строки, называется *определителем, транспонированным* с определителем Δ, и обозначается Δ^T, а замена определителя Δ на определитель Δ^T называется *транспонированием определителя* Δ.

Алгебраическим дополнением любого элемента a_{ij} определителя n-порядка (11) называют минор M_{ij} этого элемента, умноженный на $(-1)^{i+j}$. Алгебраическое дополнение элемента a_{ij}

обозначают заглавной буквой того же наименования и с теми же двумя индексами, что и у данного элемента, т. е. по определению $A_{ij}=(-1)^{i+j}M_{ij}$.

Теорема 2. Определитель n-го порядка (11) равен сумме произведений элементов любого его столбца (строки) на соответствующее им алгебраическое дополнение, т. е.

$$\begin{aligned}\Delta &= a_{1j}A_{1j}+a_{2j}A_{2j}+a_{3j}A_{3j}+\ldots+a_{nj}A_{nj}\\ &\qquad\qquad \text{для любого } j=1,\ 2,\ \ldots,\ n,\\ \Delta &= a_{i1}A_{i1}+a_{i2}A_{i2}+a_{i3}A_{i3}+\ldots+a_{in}A_{in}\\ &\qquad\qquad \text{для любого } i=1,\ 2,\ \ldots,\ n.\end{aligned} \tag{13}$$

Доказательство теоремы 2 опускается. С помощью теоремы 2 о разложении определителя n-го порядка по элементам столбца (строки) и справедливости свойств 1—8 для определителей $(n-1)$ порядка доказываются свойства 1—8 определителей n-го порядка.

Не останавливаясь на доказательстве, перечислим эти свойства.

1. При транспонировании определителя его значение не изменяется, т. е. $\Delta=\Delta^T$.

2. От перестановки двух столбцов (двух строк) определитель меняет знак.

3. Если все элементы хотя бы одного столбца (строки) определителя равны нулю, то определитель равен нулю.

4. Определитель, имеющий два одинаковых столбца (строки), равен нулю.

5. Если все элементы какого-либо столбца (строки) определителя умножить на одно и то же число α, то определитель умножится на это число (т. е. общий множитель элементов какого-либо столбца (строки) можно вынести за знак определителя).

6. Определитель, у которого элементы двух столбцов (строк) соответственно пропорциональны, равен нулю.

7. Если каждый элемент какого-либо столбца (строки) определителя есть сумма двух слагаемых, то определитель равен сумме двух определителей, у одного из которых элементами соответствующего столбца (строки) являются первые слагаемые, у другого — вторые, а остальные элементы этих двух определителей те же, что и у данного.

8. Определитель не изменится, если к элементам какого-либо столбца (строки) прибавить соответствующие элементы другого столбца (строки), умноженные на одно и то же число.

Сформулируем и докажем еще одно свойство определителей.

9. Сумма элементов некоторого столбца (строки) определителя порядка n, $n\geqslant 2$, умноженных на алгебраические дополнения соответствующих элементов другого столбца (строки), равна

нулю, т. е. для любого $n \geq 2$ и любого $i \neq j$ $\sum_{k=1}^{n} a_{ki}A_{kj} = 0$ $\left(\text{или } \sum_{k=1}^{n} a_{ik}A_{jk} = 0\right)$.

Доказательство. Пусть для определенности $j > i$, тогда, разложив определитель по элементам j-го столбца, получим

$$\sum_{k=1}^{n} a_{ki}A_{kj} = \begin{vmatrix} a_{11}a_{12} & \dots & a_{1i} & \dots & a_{1j-1}a_{1i}a_{1j+1} & \dots & a_{1n} \\ a_{21}a_{22} & \dots & a_{2i} & \dots & a_{2j-1}a_{2i}a_{2j+1} & \dots & a_{2n} \\ a_{31}a_{32} & \dots & a_{3i} & \dots & a_{3j-1}a_{3i}a_{3j+1} & \dots & a_{3n} \\ \cdot\ \cdot & \cdots & \cdot & \cdots & \cdot \quad \cdot\ \cdot & \cdots & \cdot \\ a_{n1}a_{n2} & \dots & a_{ni} & \dots & a_{nj-1}a_{ni}a_{nj+1} & \dots & a_{nn} \end{vmatrix}. \qquad (14)$$

У определителя, стоящего в правой части равенства (14), элементы i-го и j-го столбцов соответственно равны. По свойству 4 определитель (14) равен нулю, следовательно, $\sum_{k=1}^{n} a_{ki}A_{kj} = 0$. Для $i > j$ доказательство аналогично. Свойство 9 доказано.

Пользуясь свойствами 1 — 8, вычислим определитель единичной матрицы. Единичную матрицу порядка k обозначим E_k. Пусть дана матрица E_n. Разлагая определитель этой матрицы по первому столбцу, получаем

$$|E_n| = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = 1 \cdot A_{11} + 0 \cdot A_{21} + \dots + 0 \cdot A_{n1} = |E_{n-1}|.$$

Повторяя это рассуждение для определителя $|E_{n-1}|$, затем для $|E_{n-2}|$ и т. д., приходим к выводу, что $|E_n| = 1$. Итак, определитель единичной матрицы любого порядка n равен 1.

§ 3
ОБРАТНАЯ МАТРИЦА. РАНГ МАТРИЦЫ

Если две квадратные матрицы A и B одного и того же порядка n таковы, что $AB = BA = E$ (где E — единичная матрица порядка n), то говорят, что матрицы A и B являются обратными друг другу и используют обозначение $B = A^{-1}$ и $A = B^{-1}$. Например, матрицы $A = \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix}$ и $B =$

$= \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix}$ являются обратными друг к другу, поскольку $AB = BA = E$, где $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, т. е. $A = B^{-1}$, $B = A^{-1}$. По определению *обратной матрицей* для данной квадратной матрицы A порядка n называется квадратная матрица A^{-1} порядка n, обладающая свойствами $AA^{-1} = A^{-1}A = E$, где E — единичная матрица порядка n.

Для каждой матрицы A порядка n, имеющей обратную матрицу A^{-1}, *обратная матрица A^{-1} единственна.*

Действительно, пусть некоторая матрица A порядка n имеет обратную матрицу A^{-1}. Пусть существует матрица C порядка n, такая, что $CA = AC = E$, тогда

$$C = EC = (A^{-1}A)C = A^{-1}(AC) = A^{-1}E = A^{-1}.$$

Легко видеть, что обратные матрицы обладают следующими свойствами: a) $(A^{-1})^{-1} = A$, б) $(AB)^{-1} = B^{-1}A^{-1}$, c) $(A^T)^{-1} = (A^{-1})^T$.

Выясним, какие матрицы имеют обратные матрицы и если матрица A имеет обратную A^{-1}, то как ее находить.

Пусть дана квадратная матрица A порядка n $(n \geqslant 2)$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

Составим квадратную матрицу $\tilde{A}$ порядка n следующим образом: на место каждого элемента матрицы A поставим алгебраическое дополнение этого элемента, т. е.

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \dots & A_{3n} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ A_{n1} & A_{n2} & A_{n3} & \dots & A_{nn} \end{pmatrix}.$$

Транспонируем матрицу $\tilde{A}$:

$$\tilde{A}^T = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \dots & A_{n1} \\ A_{12} & A_{22} & A_{32} & \dots & A_{n2} \\ A_{13} & A_{23} & A_{33} & \dots & A_{n3} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ A_{1n} & A_{2n} & A_{3n} & \dots & A_{nn} \end{pmatrix}.$$

Матрица $\tilde{A}^T$ называется *присоединенной матрицей*. Например, найдем присоединенную матрицу $\tilde{A}^T$, если $A=\begin{pmatrix}1&0&0\\6&2&0\\5&4&3\end{pmatrix}$.

Так как $A_{11}=(-1)^{1+1}\begin{vmatrix}2&0\\4&3\end{vmatrix}=6$, $A_{12}=(-1)^{1+2}\begin{vmatrix}6&0\\5&3\end{vmatrix}=-18$,

$$A_{13}=(-1)^{1+3}\begin{vmatrix}6&2\\5&4\end{vmatrix}=14,\ A_{21}=(-1)^{2+1}\begin{vmatrix}0&0\\4&3\end{vmatrix}=0,$$

$$A_{22}=(-1)^{2+2}\times\begin{vmatrix}1&0\\5&3\end{vmatrix}=3,\ A_{23}=(-1)^{2+3}\begin{vmatrix}1&0\\5&4\end{vmatrix}=-4,$$

$$A_{31}=(-1)^{3+1}\begin{vmatrix}0&0\\2&0\end{vmatrix}=0,\ A_{32}=(-1)^{3+2}\begin{vmatrix}1&0\\6&0\end{vmatrix}=0,\ A_{33}=(-1)^{3+3}\begin{vmatrix}1&0\\6&2\end{vmatrix}=2,$$

то $\quad \tilde{A}=\begin{pmatrix}6&-18&14\\0&3&-4\\0&0&2\end{pmatrix}$ и $\tilde{A}^T=\begin{pmatrix}6&0&0\\-18&3&0\\14&-4&2\end{pmatrix}$.

Квадратная матрица A называется *вырожденной (особенной)*, если ее определитель равен нулю ($|A|=0$), и *невырожденной*, если $|A|\neq 0$.

Теорема 3. Для любой невырожденной матрицы A ($|A|\neq 0$) существует обратная матрица A^{-1}.

Доказательство. Рассмотрим невырожденную матрицу первого порядка, т. е. число a_{11}, отличное от нуля, обратная матрица к этой матрице есть число $\frac{1}{a_{11}}$.

Пусть дана некоторая невырожденная матрица A второго порядка. Найдем матричное произведение $A\tilde{A}^T$ и $\tilde{A}^T A$, где $\tilde{A}^T$ — присоединенная матрица. Так как сумма произведений элементов некоторого столбца (строки) матрицы A на алгебраические дополнения этих элементов равна определителю $|A|$ (см. теорему 2), а сумма произведений элементов некоторого столбца (строки) матрицы A на алгебраические дополнения элементов другого столбца (строки) (по свойству 9 определителей) равна нулю, то

$$A\tilde{A}^T=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}\begin{pmatrix}A_{11}&A_{21}\\A_{12}&A_{22}\end{pmatrix}=\begin{pmatrix}a_{11}A_{11}+a_{12}A_{12}&a_{11}A_{21}+a_{12}A_{22}\\a_{21}A_{11}+a_{22}A_{12}&a_{21}A_{21}+a_{22}A_{22}\end{pmatrix}=$$
$$=\begin{pmatrix}|A|&0\\0&|A|\end{pmatrix},$$

$$\tilde{A}^T A=\begin{pmatrix}A_{11}&A_{21}\\A_{12}&A_{22}\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}=\begin{pmatrix}A_{11}a_{11}+A_{21}a_{21}&A_{11}a_{12}+A_{21}a_{22}\\A_{12}a_{11}+A_{22}a_{21}&A_{12}a_{12}+A_{22}a_{22}\end{pmatrix}=$$
$$=\begin{pmatrix}|A|&0\\0&|A|\end{pmatrix}.$$

Каждый элемент полученных матриц имеет общим множителем число $|A|$, следовательно,

$$A\tilde{A}^T = |A| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = |A| E; \quad \tilde{A}^T A = |A| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = |A| E.$$

Отсюда в силу невырожденности матрицы A ($|A| \neq 0$) имеем $A\left(\frac{1}{|A|}\tilde{A}^T\right) = E = \left(\frac{1}{|A|}\tilde{A}^T\right)A$. Следовательно, для невырожденной матрицы A второго порядка матрица, элементами которой являются соответствующие элементы присоединенной матрицы $\tilde{A}^T$, деленные на определитель $|A|$, является обратной матрицей A^{-1}:

$$A^{-1} = \frac{1}{|A|}\tilde{A}^T = \begin{pmatrix} \frac{A_{11}}{|A|} & \frac{A_{21}}{|A|} \\ \frac{A_{12}}{|A|} & \frac{A_{22}}{|A|} \end{pmatrix}.$$

Итак, для невырожденной матрицы A второго порядка не только доказано существование обратной матрицы A^{-1}, но и указан способ ее построения.

Аналогично доказывается, что для любой невырожденной матрицы A порядка n, где $n \geqslant 3$, матрица, элементами которой являются соответствующие элементы присоединенной матрицы $\tilde{A}^T$, деленные на определитель $|A|$, является обратной матрицей A^{-1}:

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{|A|} & \frac{A_{21}}{|A|} & \cdots & \frac{A_{n1}}{|A|} \\ \frac{A_{12}}{|A|} & \frac{A_{22}}{|A|} & \cdots & \frac{A_{n2}}{|A|} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{A_{1n}}{|A|} & \frac{A_{n2}}{|A|} & \cdots & \frac{A_{nn}}{|A|} \end{pmatrix}.$$

Теорема 3 доказана.

Пример. Матрица $A = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 2 & 0 \\ 5 & 4 & 3 \end{pmatrix}$ невырожденная, так как $|A| = 6 \neq 0$. Ее присоединенная матрица $\tilde{A}^T = \begin{pmatrix} 6 & 0 & 0 \\ -18 & 3 & 0 \\ 14 & -4 & 2 \end{pmatrix}$,

следовательно, $A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1/2 & 0 \\ 7/3 & -2/3 & 1/3 \end{pmatrix}$.

Теорема 4. Определитель произведения двух квадратных матриц A и B порядка n равен произведению их определителей, т. е. $|AB| = |A||B|$ (аналогичное свойство справедливо и для любого конечного числа сомножителей).

Доказательство теоремы 4 опускается.

Так как определитель единичной матрицы E порядка n равен 1, то определители двух взаимно-обратных матриц являются числами взаимно-обратными, т. е. $|A||A^{-1}|=1$. *Любая вырожденная матрица A порядка n не имеет обратной*, так как по теореме 4 произведение вырожденной матрицы A на любую матрицу порядка n будет снова вырожденной матрицей.

Обратные матрицы часто применяются для решения *матричных уравнений.*

Пусть даны матрицы A, B и C. Если поставлена задача найти матрицу X, такую, что справедливо матричное равенство $AX=B$, то говорят, что дано матричное уравнение

$$AX=B \tag{1}$$

с неизвестной матрицей X. Матрицу X, удовлетворяющую уравнению (1), называют решением уравнения (1).

Если поставлена задача найти матрицу Y, такую, что справедливо матричное равенство $YA=C$, то говорят, что дано матричное уравнение

$$YA=C \tag{2}$$

с известной матрицей Y. Матрицу Y, удовлетворяющую уравнению (2), называют решением уравнения (2).

Рассмотрим частный случай решения матричных уравнений (1) и (2). Пусть матрица A невырожденная, квадратная матрица порядка n; B — матрица размеров $n\times m$; C — матрица размеров $k\times n$, где m, k — любые натуральные числа. В этом случае уравнения (1) и (2) имеют решения.

Действительно, умножим обе части первого уравнения слева на матрицу A^{-1}, что можно сделать в силу выбранных размеров матрицы B. Проделав соответствующие алгебраические действия: $A^{-1}AX=A^{-1}B$, $EX=A^{-1}B$, $X=A^{-1}B$, получим, что существует матрица $X=A^{-1}B$ размеров $n\times m$, удовлетворяющая уравнению (1). Умножим обе части второго уравнения справа на матрицу A^{-1}, что можно сделать в силу выбранных размеров матрицы C, проделав соответствующие алгебраические действия $YAA^{-1}=CA^{-1}$, $YE=CA^{-1}$, $Y=CA^{-1}$, получим, что существует матрица $Y=CA^{-1}$, размеров $k\times n$, удовлетворяющая уравнению (2).

Отметим, что даже если $B=C$, но матрицы A^{-1} и B не перестановочны, то $X\neq Y$, т. е. для одних и тех же матриц A и $B=C$ уравнения (1) и (2) имеют, вообще говоря, разные решения.

Пример. Пусть $A=\begin{pmatrix}7 & 3\\ 2 & 1\end{pmatrix}$, $B=C=\begin{pmatrix}1 & 2\\ 0 & -1\end{pmatrix}$. Решим уравнения (1) и (2). Прежде всего найдем $A^{-1}=\begin{pmatrix}1 & -3\\ -2 & 7\end{pmatrix}$. Тогда $X=$

$$= A^{-1}B = \begin{pmatrix} 1 & -3 \\ -2 & 7 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ -2 & -11 \end{pmatrix}, \ Y = BA^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \times$$

$$\times \begin{pmatrix} 1 & -3 \\ -2 & 7 \end{pmatrix} = \begin{pmatrix} -3 & 11 \\ 2 & -7 \end{pmatrix}, \text{ т. е. } X \neq Y.$$

Ранее (см. § 2) было дано понятие минора элемента a_{ij} квадратной матрицы n-го порядка. В общем случае матрицы размеров $m \times n$ понятие минора элементов матрицы не вводится, а вводится понятие минора порядка k данной матрицы.

Пусть дана матрица A размеров $m \times n$ и пусть k — натуральное число, такое, что $1 \leqslant k \leqslant \min(m, n)$. Выберем произвольно k строк и k столбцов матрицы A. Из элементов, стоящих на пересечении этих строк и столбцов (не меняя их порядка), образуем квадратную матрицу M порядка k. Определитель матрицы M называется *минором порядка k матрицы A*.

Пример. Пусть дана матрица $A = \begin{pmatrix} 1 & 3 \\ 3 & 4 \\ -3 & 0 \end{pmatrix}$ размеров 3×2. Тогда миноры первого порядка матрицы есть элементы этой матрицы, т. е. числа 1, 3, 3, 4, —3, 0. Миноры второго порядка есть определители $\begin{vmatrix} 1 & 3 \\ 3 & 4 \end{vmatrix}$, $\begin{vmatrix} 1 & 3 \\ -3 & 0 \end{vmatrix}$, $\begin{vmatrix} 3 & 4 \\ -3 & 0 \end{vmatrix}$. Миноров более высоких порядков эта матрица не имеет.

Если дана квадратная матрица A порядка n, то ее минор порядка n есть определитель $|A|$, миноры порядка $(n-1)$ есть миноры соответствующих элементов этой матрицы, миноры порядка 1 есть элементы этой матрицы. Кроме того, эта матрица имеет миноры порядка k, где $1 < k < n-1$.

Рангом матрицы A размеров $m \times n$ называется наивысший порядок отличных от нуля миноров матрицы A. Если все элементы матрицы A есть нули, то говорят, что ранг матрицы равен нулю. Ранг невырожденной квадратной матрицы A порядка n равен n. Например, ранг матрицы $A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 2 & 4 & 0 & 8 \end{pmatrix}$ равен единице, так как все миноры второго порядка этой матрицы равны нулю в силу пропорциональности строк матрицы, но есть миноры первого порядка — элементы матрицы A — отличные от нуля.

Пусть дана матрица A размеров $m \times n$ и пусть дано натуральное число $l < \min(m, n)$. *Если все миноры порядка l матрицы A равны нулю, то и все миноры этой матрицы порядка $l+1$ также равны нулю.*

Действительно, возьмем любой минор этой матрицы порядка $l+1$ и разложим его по элементам некоторой строки. Каждое слагаемое разложения содержит сомножителем некоторый определитель порядка l — минор порядка l соответствующего элемента выбранной строки определителя порядка $l+1$. Поскольку каждый минор порядка l равен нулю, то и сумма разложения тоже равна нулю, поэтому выбранный минор порядка $l+1$ равен нулю. Следовательно, если все миноры порядка l равны нулю, то и все миноры порядка, большего, чем l, также равны нулю.

Таким образом, для определения ранга матрицы размеров $m\times n$ надо сначала найти хотя бы один отличный от нуля минор первого порядка, т. е. надо выяснить, есть ли среди элементов матрицы хотя бы один, отличный от нуля. Если такого элемента нет, то ранг матрицы равен нулю. Если у матрицы есть элемент, отличный от нуля, то надо найти хотя бы один отличный от нуля минор второго порядка. Если такого минора нет, то ранг матрицы равен 1. Если есть отличный от нуля минор второго порядка, то надо искать отличный от нуля минор третьего порядка и т. д. до тех пор, пока найдется отличный от нуля минор r-го порядка, но либо все миноры порядка $r+1$ равны нулю, либо $r=\min(m, n)$. Тогда ранг матрицы равен r.

Пример. Найти ранг матрицы $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 10 & 13 \end{pmatrix}$. Минор первого порядка (число 1), лежащий на пересечении первого столбца и первой строки, отличен от нуля. Минор второго порядка, лежащий на пересечении первых двух столбцов и первых двух строк $\begin{vmatrix} 1 & 2 \\ 3 & 5 \end{vmatrix}=-1$ также отличен от нуля. Все примеры третьего порядка этой матрицы равны нулю. Значит, ранг матрицы равен двум.

§ 4
СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Рассмотрим систему m линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m, \end{cases} \qquad (1)$$

где $x_1, x_2 \ldots, x_n$ — неизвестные, a_{ij} (где $i=1, 2, \ldots, m$, $j=1, 2, \ldots, n$) — коэффициенты, $b_1, b_2, \ldots, b_m$ —свободные члены системы уравнений (1). Напомним, что числовой набор $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, соответствующий набору неизвестных $(x_1, x_2, \ldots, x_n)$, называется решением системы (1), если при подстановке в каждое уравнение системы (1) соответствующих чисел вместо неизвестных получается m верных числовых равенств. Решить систему (1) — это значит найти все ее решения или доказать, что их нет.

Система линейных уравнений называется *совместной*, если она имеет хотя бы одно решение, и *несовместной (противоречивой)*, если она не имеет решений. Совместная система называется *определенной*, если она имеет единственное решение, и *неопределенной*, если она имеет более одного решения. Легко видеть, что неопределенная система линейных уравнений всегда имеет бесконечное множество решений.

Пример. Линейное уравнение с одним неизвестным $a_{11}x = b_1$ $(a_{11} \neq 0)$ имеет единственное решение $x = b_1/a_{11}$.

Линейное уравнение с двумя неизвестными $a_{11}x_1 + a_{12}x_2 = b_1$ $(a_{11} \neq 0,\ a_{12} \neq 0)$ имеет бесконечно много решений. Действительно, так как $a_{12} \neq 0$, то уравнение $a_{11}x_1 + a_{12}x_2 = b_1$ равносильно уравнению $x_2 = (b_1 - a_{11}x_1)/a_{12}$, которое имеет решением набор $\left(\alpha,\ -\frac{a_{11}}{a_{12}}\alpha + \frac{b_1}{a_{12}}\right)$ при любом действительном α.

Система двух линейных уравнений с одним неизвестным $\begin{cases} a_{11}x = b_1 & (a_{11} \neq 0), \\ a_{21}x = b_2 & (a_{21} \neq 0), \end{cases}$ такая, что $b_1/a_{11} \neq b_2/a_{21}$, несовместна. Действительно, пусть число α — решение уравнения $a_{11}x = b_1$, т. е. пусть справедливо числовое равенство $a_{11}\alpha = b_1$. Тогда $\alpha = b_1/a_{11}$. Если бы $\alpha = b_1/a_{11}$ было решением уравнения $a_{21}x = b_2$, то было бы справедливо числовое равенство $a_{21}b_1/a_{11} = b_2$ или, в силу того что $a_{21} \neq 0$, числовое равенство $b_1/a_{11} = b_2/a_{21}$, что противоречит условию. Итак, решение α уравнения $a_{11}x = b_1$ не является решением уравнения $a_{21}x = b_2$, а это значит, что система двух уравнений с одним неизвестным, такая, что $b_1/a_{11} \neq b_2/a_{21}$, несовместна.

Основной матрицей системы (1) называется матрица A размеров $m \times n$, элементами которой являются коэффициенты при неизвестных системы уравнений (1), т. е.

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

Матрицей неизвестных системы (1) называется матрица-столбец X, элементами которой являются неизвестные системы, т. е.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Матрицей свободных членов системы (1) называется матрица-столбец B, элементами которой являются свободные члены системы, т. е. $B=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.

Матричное уравнение

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

или

$$AX=B \tag{2}$$

называется *матричным уравнением* системы (1), если матрицы A, X и B — соответствующие матрицы системы (1).

Поскольку любой числовой набор $(\alpha_1, \alpha_2, \dots, \alpha_n)$, являющийся решением системы (1), записанный в виде матрицы-столбца $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$, является решением матричного уравнения (2), а любая числовая матрица-столбец $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$, являющаяся решением матричного уравнения (2) системы (1), записанная в виде числового набора $(\beta_1, \beta_2, \dots, \beta_n)$, является решением системы (1), то говорят, что *система линейных уравнений* (1) *равносильна матричному уравнению* (2) *системы* (1).

Рассмотрим частный случай системы (1). Пусть дана система n уравнений с n неизвестными

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3+\dots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+a_{23}x_3+\dots+a_{2n}x_n=b_2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \dots \quad \cdot \\ a_{n1}x_1+a_{n2}x_2+a_{n3}x_3+\dots+a_{nn}x_n=b_n \end{cases} \tag{3}$$

и соответствующее матричное уравнение этой системы

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

или
$$AX=B. \tag{4}$$

Определитель матрицы A системы (3) будем обозначать через Δ, т. е. $\Delta=|A|$. Предположим, что основная матрица A (являющаяся квадратной матрицей порядка n) системы (3) невырожденная, т. е. пусть $\Delta=|A|\neq 0$. Тогда, как показано в § 3, существует единственная обратная матрица A^{-1}. Умножив левую и правую части матричного уравнения (4) системы (3) слева на обратную матрицу A^{-1}, получим, что $X=A^{-1}B$, причем матрица X — матрица размеров $n\times 1$, т. е. матрица-столбец. Итак, матричное уравнение (4) системы (3) в силу единственности обратной матрицы A^{-1} имеет единственное решение — матрицу-столбец $A^{-1}B$.

Так как матричное уравнение (4) равносильно системе (3), то если определитель основной матрицы A линейной системы (3) n уравнений с n неизвестными не равен нулю (т. е. $|A|\neq 0$), то система (3) совместна и определена, т. е. имеет единственное решение. Найти это единственное решение системы (3) можно по *правилу Крамера*, формулировка которого будет приведена ниже.

Если в основной матрице A системы (3) n линейных уравнений с n неизвестными вместо j-го столбца коэффициентов системы (3) поставить столбец свободных членов этой системы, то полученную матрицу называют *дополнительной матрицей системы* (3) и обозначают ее $\overset{j}{A}$, т. е.

$$\overset{j}{A}=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j-1} & b_1 & a_{1j+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j-1} & b_2 & a_{2j+1} & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nj-1} & b_n & a_{nj+1} & \dots & a_{nn} \end{pmatrix}. \tag{5}$$

Определитель матрицы $\overset{j}{A}$ обозначают обычно Δ_j, т. е. $\Delta_j=|\overset{j}{A}|$.

Перейдем теперь к решению системы уравнений (3).

Пусть $n=1$, т. е. система (3) состоит из одного уравнения с одним неизвестным: $a_{11}x=b_1$, и пусть $a_{11}\neq 0$. Ясно, что это уравнение имеет единственное решение — число b_1/a_{11}.

Пусть теперь $n=2$. Рассмотрим систему двух уравнений с двумя неизвестными:

$$\begin{cases} a_{11}x_1+a_{12}x_2=b_1, \\ a_{21}x_1+a_{22}x_2=b_2. \end{cases} \tag{6}$$

Основной матрицей A, дополнительными матрицами $\overset{1}{A}$ и $\overset{2}{A}$ этой системы являются соответственно матрицы

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \overset{1}{A}=\begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}, \quad \overset{2}{A}=\begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}.$$

Пусть определитель основной матрицы A не равен нулю

($\Delta=|A|\neq 0$). Тогда существует единственная обратная матрица $A^{-1}=\frac{1}{|A|}\begin{pmatrix}A_{11} & A_{21}\\ A_{12} & A_{22}\end{pmatrix}$ и справедлива цепочка матричных равенств

$$\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\frac{1}{|A|}\begin{pmatrix}A_{11} & A_{21}\\ A_{12} & A_{22}\end{pmatrix}\begin{pmatrix}b_1\\ b_2\end{pmatrix}=\frac{1}{|A|}\begin{pmatrix}A_{11}b_1+A_{21}b_2\\ A_{12}b_1+A_{22}b_2\end{pmatrix}=\begin{pmatrix}\dfrac{A_{11}b_1+A_{21}b_2}{|A|}\\ \dfrac{A_{12}b_1+A_{22}b_2}{|A|}\end{pmatrix},$$

откуда

$$x_1=\frac{A_{11}b_1+A_{21}b_2}{|A|}=\frac{b_1(-1)^{1+1}M_{11}+b_2(-1)^{2+1}M_{21}}{|A|}=\frac{\begin{vmatrix}b_1 & a_{12}\\ b_2 & a_{22}\end{vmatrix}}{|A|}=\frac{|\overset{1}{A}|}{|A|}=\frac{\Delta_1}{\Delta},$$

$$x_2=\frac{A_{12}b_1+A_{22}b_2}{|A|}=\frac{b_1(-1)^{1+2}M_{12}+b_2(-1)^{2+2}M_{22}}{|A|}=\frac{\begin{vmatrix}a_{11} & b_1\\ a_{21} & b_2\end{vmatrix}}{|A|}=\frac{|\overset{2}{A}|}{|A|}=\frac{\Delta_2}{\Delta},$$

т. е. компоненты решения $(x_1,\ x_2)$ системы (6) находятся по формулам

$$x_1=\frac{\Delta_1}{\Delta},\qquad x_2=\frac{\Delta_2}{\Delta},\tag{7}$$

где Δ, Δ_1, Δ_2 — определители матриц A, $\overset{1}{A}$, $\overset{2}{A}$ системы (6).

Рассмотрим систему трех линейных уравнений с тремя неизвестными:

$$\begin{cases}a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1,\\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2,\\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3.\end{cases}\tag{8}$$

Основной матрицей A и дополнительными матрицами $\overset{1}{A}$, $\overset{2}{A}$, $\overset{3}{A}$ системы (8) являются соответственно матрицы

$$A=\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix},\quad \overset{1}{A}=\begin{pmatrix}b_1 & a_{12} & a_{13}\\ b_2 & a_{22} & a_{23}\\ b_3 & a_{32} & a_{33}\end{pmatrix},$$

$$\overset{2}{A}=\begin{pmatrix}a_{11} & b_1 & a_{13}\\ a_{21} & b_2 & a_{23}\\ a_{31} & b_3 & a_{33}\end{pmatrix},\quad \overset{3}{A}=\begin{pmatrix}a_{11} & a_{12} & b_1\\ a_{21} & a_{22} & b_2\\ a_{31} & a_{32} & b_3\end{pmatrix}.$$

Пусть определитель основной матрицы A не равен нулю ($\Delta=|A|\neq 0$). Тогда существует единственная обратная матрица $A^{-1}=\frac{1}{|A|}\begin{pmatrix}A_{11} & A_{21} & A_{31}\\ A_{12} & A_{22} & A_{32}\\ A_{13} & A_{23} & A_{33}\end{pmatrix}$ и справедлива цепочка матричных равенств

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} A_{11}b_1+A_{21}b_2+A_{31}b_3 \\ A_{12}b_1+A_{22}b_2+A_{32}b_3 \\ A_{13}b_1+A_{23}b_2+A_{33}b_3 \end{pmatrix} =$$

$$= \begin{pmatrix} \dfrac{A_{11}b_1+A_{21}b_2+A_{31}b_3}{|A|} \\ \dfrac{A_{12}b_1+A_{22}b_2+A_{32}b_3}{|A|} \\ \dfrac{A_{13}b_1+A_{23}b_2+A_{33}b_3}{|A|} \end{pmatrix},$$

откуда

$$x_1 = \frac{A_{11}b_1+A_{21}b_2+A_{31}b_3}{|A|} = \frac{b_1(-1)^{1+1}M_{11}+b_2(-1)^{2+1}M_{21}+b_3(-1)^{3+1}M_{31}}{|A|} =$$

$$= \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{|A|} = \frac{|\overset{1}{A}|}{|A|} = \frac{\Delta_1}{\Delta},$$

$$x_2 = \frac{A_{12}b_1+A_{22}b_2+A_{32}b_3}{|A|} = \frac{b_1(-1)^{1+2}M_{12}+b_2(-1)^{2+2}M_{22}+b_3(-1)^{3+2}M_{32}}{|A|} =$$

$$= \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{|A|} = \frac{|\overset{2}{A}|}{|A|} = \frac{\Delta_2}{\Delta},$$

$$x_3 = \frac{A_{13}b_1+A_{23}b_2+A_{33}b_3}{|A|} = \frac{b_1(-1)^{1+3}M_{13}+b_2(-1)^{2+3}M_{23}+b_3(-1)^{3+3}M_{33}}{|A|} =$$

$$= \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{|A|} = \frac{|\overset{3}{A}|}{|A|} = \frac{\Delta_3}{\Delta},$$

т. е. компоненты решения (x_1, x_2, x_3) системы (8) находятся по формулам

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad x_3 = \frac{\Delta_3}{\Delta}, \tag{9}$$

где Δ, Δ_1, Δ_2, Δ_3 — определители матриц A, $\overset{1}{A}$, $\overset{2}{A}$, $\overset{3}{A}$ системы (8).

Докажем, что если определитель основной матрицы A системы (3) не равен нулю, то j-я компонента x_j ($j = 1, 2, 3, \ldots, n$) единственного решения $(x_1, x_2, x_3, \ldots, x_n)$ системы (4) находится по формулам $x_j = \frac{|\overset{j}{A}|}{|A|} = \frac{\Delta_j}{\Delta}$, где $|\overset{j}{A}| = \Delta_j$ — определитель дополнительной матрицы (5) системы (3), а $\Delta = |A|$ — определитель основной матрицы A системы (3).

Рассмотрим j-ю компоненту матрицы-столбца $A^{-1}B$, являющегося единственным решением матричного уравнения (4) системы (3). Имеем

$$x_j = \frac{A_{1j}b_1 + A_{2j}b_2 + \ldots + A_{nj}b_n}{|A|} =$$

$$= \frac{1}{|A|}\left(b_1(-1)^{1+j}M_{1j} + b_2(-1)^{2+j}M_{2j} + \ldots + b_n(-1)^{n+j}M_{nj}\right) =$$

$$= \frac{1}{|A|}\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1j-1} & b_1 & a_{1j+1} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2j-1} & b_2 & a_{2j+1} & \ldots & a_{2n} \\ a_{31} & a_{32} & \ldots & a_{3j-1} & b_3 & a_{3j+1} & \ldots & a_{3n} \\ \cdot & \cdot & \ldots & \cdot & \cdot & \cdot & \ldots & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nj-1} & b_n & a_{nj+1} & \ldots & a_{nn} \end{vmatrix} = \frac{|\overset{j}{A}|}{|A|} = \frac{\Delta_j}{\Delta},$$

что и требовалось доказать.

Правило, по которому находится решение системы (4), называется правилом Крамера. Сформулируем его.

Правило Крамера. Если система (3) n уравнений с n неизвестными такова, что определитель ее основной матрицы не равен нулю ($\Delta \neq 0$), то система имеет единственное решение $(x_1, x_2, \ldots, x_n)$, каждая компонента которого находится по формулам Крамера, т. е. $x_j = \frac{\Delta_j}{\Delta}$ $(j = 1, 2, \ldots, n)$, где Δ_j — определитель дополнительной матрицы $\overset{j}{A}$, полученной из основной матрицы A системы (3) заменой j-го столбца на столбец свободных членов системы.

Пример. Решить систему уравнений $\begin{cases} 7x_1 + 3x_2 + 2x_3 = 1, \\ 3x_1 + x_2 + 2x_3 = 2, \\ 10x_1 + 12x_2 + 8x_3 = 4. \end{cases}$

Определитель основной матрицы $\Delta = \begin{vmatrix} 7 & 3 & 2 \\ 3 & 1 & 2 \\ 10 & 12 & 8 \end{vmatrix} = -72 \neq 0$. Следовательно, система имеет единственное решение. Найдем его по правилу Крамера:

$$\Delta_1 = \begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 2 \\ 4 & 12 & 8 \end{vmatrix} = 0, \quad \Delta_2 = \begin{vmatrix} 7 & 1 & 2 \\ 3 & 2 & 2 \\ 10 & 4 & 8 \end{vmatrix} = 36, \quad \Delta_3 = \begin{vmatrix} 7 & 3 & 1 \\ 3 & 1 & 2 \\ 10 & 12 & 4 \end{vmatrix} = -90.$$

Следовательно, $x_1 = \Delta_1/\Delta = 0$, $x_2 = \Delta_2/\Delta = -1/2$, $x_3 = \frac{\Delta_3}{\Delta} = 5/4$, т. е. система имеет единственное решение $(0, -{}^1/_2, {}^5/_4)$.

В случае, когда определитель $|A|$ основной матрицы A системы (3) равен нулю, правило Крамера неприменимо.

Перейдем к изучению систем m линейных уравнений с n неизвестными.

Расширенной матрицей системы уравнений (1) называется матрица, получаемая из основной матрицы A системы (1) добавлением столбца свободных членов, располагая его последним столбцом и отделяя его вертикальной чертой, т. е. матрица

$$B=\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & b_n \end{array}\right).$$

Вопрос о совместности системы (1) решает *критерий Кронекера — Капелли*: система (1) m линейных уравнений с n неизвестными совместна тогда и только тогда, когда ранг расширенной матрицы B равен рангу основной матрицы A системы (1).

Доказательство этой теоремы опускается.

Пусть ранг основной матрицы системы (1) равен r, причем $1 \leqslant r \leqslant \min(m, n)$. В этом случае любой отличный от нуля минор порядка r основной матрицы системы (1) называют *главным минором системы*.

Решение системы линейных уравнений состоит в следующем. Вычисляем ранг основной матрицы A системы (1) и расширенной матрицы B. Если ранг основной матрицы A системы (1) не равен рангу расширенной матрицы B, то согласно критерию Кронекера — Капелли система несовместна, т. е. система (1) не имеет ни одного решения. На этом решение системы (1) заканчивается. Если ранги основной и расширенной матрицы равны между собой и равны r, т. е. система (1) совместна, то берут любой отличный от нуля минор основной матрицы порядка r и рассматривают r уравнений, коэффициенты которых входят в этот главный минор, а остальные уравнения системы отбрасывают. Неизвестные, коэффициенты которых входят в этот главный минор, объявляют главными, а остальные неизвестные — свободными. Новую систему переписывают так, что в левых частях всех уравнений остаются только члены, содержащие r главных неизвестных, а все остальные члены уравнений, содержащие $n-r$ неизвестных, переносятся в правые части уравнений. Затем по правилу Крамера находят главные неизвестные. Легко видеть, что при этом главные неизвестные выражаются через свободные неизвестные, каждое из которых может принимать любое числовое значение. Полученные выражения называются общим решением. Придавая всем свободным неизвестным некоторые числовые значения, из общего решения находят соответствующие числовые значения главных неизвестных и тем самым находят решение исходной системы уравнений (1), которое называется частным решением при данных числовых значениях свободных неизвестных. Указанным путем можно получить любое решение системы (1), т. е. решить ее.

Глава X

КОМПЛЕКСНЫЕ ЧИСЛА

При рассмотрении действительных чисел отмечалось, что в множестве действительных чисел нельзя, например, найти число, квадрат которого равен -1. Чтобы подобные задачи были разрешимы, понятие числа расширяется с помощью введения комплексных чисел.

§ 1
ПОНЯТИЕ КОМПЛЕКСНОГО ЧИСЛА

Пусть даны выражения вида $a+bi$, где a и b — действительные числа, i — некоторый символ, смысл и связь которого с действительным числом b, а также смысл знака «$+$», соединяющего a и bi, будут выяснены позже. Введем определения равенства, суммы и произведения таких выражений.

Выражения $a+bi$ и $c+di$ считаются *равными* в том и только том случае, если одновременно $a=c$ и $b=d$. Равенство выражений $a+bi$ и $c+di$ принято записывать в виде $a+bi=c+di$.

Из определения равенства вытекает, что два выражения $a+bi$ и $c+di$ различны, т. е. $a+bi\neq c+di$, если выполнено хотя бы одно из неравенств $a\neq c$ или $b\neq d$.

Замечание. Из всех знаков неравенств $\neq$, $<$, $\geqslant$, $\leqslant$, $>$ для выражений вида $a+bi$ употребляется только знак $\neq$.

Суммой выражений $a+bi$ и $c+di$ называется выражение $(a+c)+(b+d)\,i$. Сумму выражений $a+bi$ и $c+di$ принято обозначать через $(a+bi)+(c+di)$, т. е. по определению $(a+bi)+(c+di)=(a+c)+(b+d)\,i$.

Произведением выражений $a+bi$ и $c+di$ называется выражение $(ac-bd)+(ad+bc)\,i$. Произведение выражений $a+bi$ и $c+di$ принято обозначать через $(a+bi)(c+di)$, т. е. по определению $(a+bi)(c+di)=(ac-bd)+(ad+bc)\,i$.

Выражения $a+bi$, различающиеся по определению равенства, складывающиеся по определению суммы и перемножающиеся по определению произведения, называются *комплексными числами*. Множество всех этих выражений называется *множеством комплексных чисел*.

Часто комплексное число $a+bi$ обозначают одной буквой, например буквой $z: z=a+bi$. Пусть даны комплексные числа $z_1=a_1+b_1i$, $z_2=a_2+b_2i$, $z_3=a_3+b_3i$, $\ldots$, $z_n=a_n+b_ni$. Для того чтобы найти сумму комплексных чисел $z_1, z_2, z_3, \ldots, z_n$, надо найти сумму первых чисел, затем к полученной сумме прибавить третье число, к полученной сумме прибавить четвертое и т. д., пока не переберем все слагаемые. Аналогично определяется и произведение нескольких комплексных чисел.

Если комплексное число z взято сомножителем n раз ($n \geqslant 2$), то произведение $\underbrace{zz \ldots z}_{n \text{ раз}}$ называют n-й степенью числа z и обозначают z^n, т. е. по определению $\underbrace{zz \ldots z}_{n \text{ раз}} = z^n$. Кроме того, по определению $z^1 = z$.

Приведем *основные законы* сложения и умножения комплексных чисел:

а) $z_1 + z_2 = z_2 + z_1$ (коммутативность сложения);

б) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ (ассоциативность сложения);

в) $z_1 z_2 = z_2 z_1$ (коммутативность умножения);

г) $(z_1 z_2) z_3 = z_1 (z_2 z_3)$ (ассоциативность умножения);

д) $(z_1 + z_2) z_3 = z_1 z_3 + z_2 z_3$ (дистрибутивность умножения относительно сложения).

Справедливость этих законов следует из определения суммы и произведения комплексных чисел и из справедливости аналогичных законов для сложения и умножения действительных чисел (проверка их справедливости опускается).

Для действий сложения и умножения комплексных чисел вводятся действия, обратные им.

Разностью комплексных чисел z_1 и z_2 называется такое комплексное число z_3, которое в сумме с z_2 дает z_1.

Покажем, что для любых комплексных чисел $z_1 = a_1 + b_1 i$ и $z_2 = a_2 + b_2 i$ их разность $z_3 = z_1 - z_2$ существует, единственна и вычисляется по правилу $z_3 = (a_1 - a_2) + (b_1 - b_2) i$, т. е. существует единственное комплексное число $z_3 = x + yi$, которое в сумме с z_2 дает z_1. По определению суммы комплексных чисел $z_2 + z_3 = (a_2 + x) + (b_2 + y) i$. По определению равенства комплексных чисел числа z_1 и $z_2 + z_3$ равны тогда и только тогда, когда одновременно справедливы равенства

$$\begin{cases} a_2 + x = a_1, \\ b_2 + y = b_1. \end{cases}$$

Из этих равенств числа x и y всегда определяются и притом единственным образом: $x = a_1 - a_2$, $y = b_1 - b_2$, т. е. существует и притом единственное комплексное число $z_3 = (a_1 - a_2) + (b_1 - b_2) i$, которое и является разностью z_1 и z_2, т. е. $z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2) i$.

Частным от деления комплексного числа z_1 на комплексное число $z_2 \neq 0 + 0i$ называется такое комплексное число z_3, которое при умножении на z_2 дает z_1.

Можно показать, что для любых комплексных чисел $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$ ($z_2 \neq 0 + 0i$) их частное $z_3 = z_1 / z_2$ существует, единственно и вычисляется по правилу $z_3=\frac{a_1a_2+b_1b_2}{a_2^2+b_2^2}+$ $+\frac{b_1a_2-a_1b_2}{a_2^2+b_2^2}i$. Доказательство этого факта опускается.

Рассмотрим комплексные числа вида $a+0i$. Каждое такое число принято рассматривать как действительное число a, т. е. каждое число вида $a+0i$ отождествляется с действительным числом a. Обычно числа a и $a+0i$ не различаются и даже принято писать равенство $a+0i=a$. В частности, числа 0 и $0+0i$ не различают, число $0+0i$ также называют нулем и пишут $0+0i=0$.

Рассмотрим теперь комплексные числа вида $0+bi$. Принято такие числа обозначать просто bi и писать равенство $0+bi=bi$. В частности, комплексное число $0+1i$ принято обозначать просто i и писать равенство $0+1i=i$. Комплексное число i называется *мнимой единицей*. Покажем, что мнимая единица обладает свойством $i^2=-1$. Действительно, на основании принятых соглашений справедливы следующие равенства: $i^2=$ $=(0+1i)^2=(0+1i)(0+1i)=(0\cdot 0-1\cdot 1)+(0\cdot 1+1\cdot 0)\,i=$ $=-1+0i=-1$. Комплексные числа вида $0+bi$ называют *чисто мнимыми* числами.

Согласно принятым соглашениям, любое чисто мнимое число bi представляет собой произведение двух комплексных чисел: числа b и мнимой единицы i. Действительно,

$$(b+0i)(0+1i)=(b\cdot 0-0\cdot 1)+(b\cdot 1+0\cdot 0)\,i=0+bi=bi.$$

Любое комплексное число $a+bi$ представляет собой сумму двух комплексных чисел: числа a и чисто мнимого числа bi. Действительно, $(a+0i)+(0+bi)=(a+0)+(0+b)\,i=a+bi$.

Из определения действий сложения, вычитания и умножения комплексных чисел вытекает, что эти действия можно проводить по правилам действий над многочленами (см. гл. II), заменяя i^2 на -1 и объединяя затем члены, содержащие и не содержащие i.

Пример. $(2+3i)-4+(7-13i)+4i=(2-4+7)+$ $+(3-13+4)\,i=5-6i$; $(4+5i)+(x+yi)-(a^2-bi)=(4+x-a^2)+$ $+(5+y+b)\,i$; $(a+bi)+(x+yi)=ax+ayi+bxi+byi^2=(ax-by)+$ $+(ay+bx)\,i$; $(2+4i)(7-i)=14-2i+28i-4i^2=18+26i$.

Пусть z_1 и z_2 — любые комплексные числа, n — любое натуральное число. Пользуясь правилами действий над комплексными числами, легко доказать справедливость формул сокращенного умножения:

$$(z_1+z_2)^2=z_1^2+2z_1z_2+z_2^2;$$

$$(z_1+z_2)^n=z_1^n+C_n^1z_1^{n-1}z_2+C_n^2z_1^{n-2}z_2^2+\ \dots\ +C_n^{n-1}z_1z_2^{n-1}+z_2^n;$$

$$z_1^2-z_2^2=(z_1+z_2)(z_1-z_2);$$

$$z_1^n-z_2^n=(z_1-z_2)(z_1^{n-1}+z_1^{n-2}z_2+z_1^{n-3}z_2^2+\ \dots\ +z_1z_2^{n-2}+z_2^{n-1}).$$

Комплексное число $\bar{z}=a-bi$ называют *числом, сопряженным комплексному числу* $z=a+bi$.

По формулам сокращенного умножения $z\bar{z}=(a+bi)\times\times a-bi)=a^2-(bi)^2=a^2+b^2$. Это свойство сопряженных комплексных чисел позволяет деление комплексного числа z_1 на комплексное число $z_2\neq 0$ свести к умножению комплексных чисел z_1 и $\bar{z}_2$. Действительно, пусть $z_1=a_1+b_1i$, $z_2=a_2+b_2i$ и $z_2\neq 0$, тогда

$$\frac{z_1}{z_2}=\frac{z_1\bar{z}_2}{z_2\bar{z}_2}=\frac{(a_1+b_1i)(a_2-b_2i)}{a_2^2+b_2^2}=\frac{a_1a_2-a_1b_2i+b_1a_2i-b_1b_2i^2}{a_2^2+b_2^2}=$$
$$=\frac{a_1a_2+b_1b_2}{a_2^2+в_2^2}+\frac{b_1a_2-a_1b_2}{a_2^2+b_2^2}i.$$

Пример. $\frac{2+i}{4-i}=\frac{(2+i)(4+i)}{(4+i)(4-i)}=\frac{8+2i+4i+i^2}{16+1}=\frac{7}{17}+\frac{6}{17}i.$

По определению, если $z\neq 0$, то $z^0=1$. Если $z\neq 0$ и n — натуральное число, то $z^{-n}=1/z^n$.

Пример. $(2+i)^0=1$; $i^{-1}=\frac{1}{i}=\frac{-i}{i(-i)}=\frac{-i}{-i^2}=-i$; $(2+i)^{-2}=$
$=\frac{1}{(2+i)^2}=\frac{1}{4+4i+i^2}=\frac{1}{3+4i}=\frac{1\cdot(3-4i)}{(3+4i)(3-4i)}=\frac{3-4i}{9+16}=\frac{3}{25}-\frac{4}{25}i.$

§ 2
ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ КОМПЛЕКСНЫХ ЧИСЕЛ

Пусть на плоскости дана прямоугольная система координат. Каждому комплексному числу $z=a+bi$ поставим в соответствие точку на плоскости с координатами a и b, т. е. точку $M(a, b)$ (рис. 140). Легко видеть, что тем самым между множеством комплексных чисел и множеством точек на плоскости установлено такое соответствие, что каждому числу соответствует только одна точка и разным числам соответствуют разные точки. При этом на плоскости нет ни одной точки, которая бы не соответствовала какому-нибудь комплексному числу. Значит, между множеством комплексных чисел и множеством точек на плоскости установлено взаимно-однозначное соответствие и поэтому можно считать, что комплексное число $z=a+bi$ есть точка на плоскости с координатами (a, b).

Модулем комплексного числа $z=a+bi$ называется действительное число $r=|z|=\sqrt{a^2+b^2}$.

Поскольку комплексному числу $z=a+bi$ можно поставить в соответствие точку M на плоскости с координатами a и b, то модулю комплексного числа можно придать следующий геометрический смысл: $|z|$ — расстояние соответствующей точки $M(a, b)$ до начала координат. Рассмотрим следующие задачи.

1. Найти множество точек плоскости, соответствующих комплексным числам z, таким, что $|z|=1$.

Согласно геометрическому смыслу модуля комплексного числа это будут точки, находящиеся от начала координат на расстоянии единицы, т. е. точки, которые лежат на окружности радиуса, равного единице, и с центром в начале координат (рис. 141).

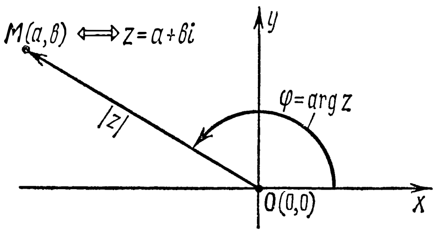


Рис. 140

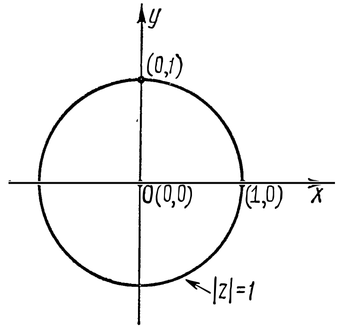


Рис. 141

2. Найти множество точек плоскости, соответствующих комплексным числам z, таким, что $2 \leqslant |z| < 3$.

Согласно геометрическому смыслу модуля комплексного числа это будут точки, расположенные внутри кольца, образованного окружностями с радиусами 2 и 3 и с центром в начале координат, включая окружность радиуса 2 (рис. 142).

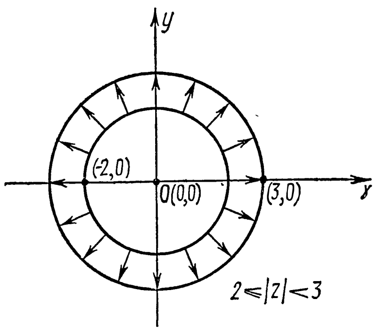


Рис. 142

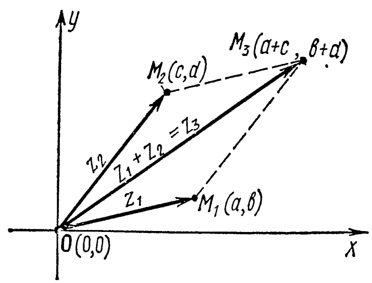


Рис. 143

Комплексное число $z = a + bi$ можно рассматривать как вектор z, начало которого находится в начале координат, а конец — в точке $M(a, b)$, изображающей это число (см. рис. 140). Далее, говоря о векторах, изображающих комплексные числа, будем подразумевать, что начало этих векторов находится в начале координат.

Пусть даны числа $z_1 = a + bi$, $z_2 = c + di$ и $z_3 = (a + c) + (b + d)i$. Рассмотрим вектор z_1, конец которого находится в точке $M_1(a, b)$, вектор z_2, конец которого—точка $M_2(c, d)$, и вектор z_3, конец которого — точка $M(a+c,\ b+d)$. Легко видеть, что вектор z_3 является диагональю параллелограмма $OM_1M_3M_2$ (рис. 143). Поскольку число z_3 есть сумма чисел z_1 и z_2, то отсюда вытекает, что сложение двух комплексных чисел можно геометрически интерпретировать как сложение по правилу параллелограмма векторов z_1 и z_2, начала которых находятся в начале координат, а концы — в точках $M_1(a,\ b)$ и $M_2(c,\ d)$, изображающих эти числа.

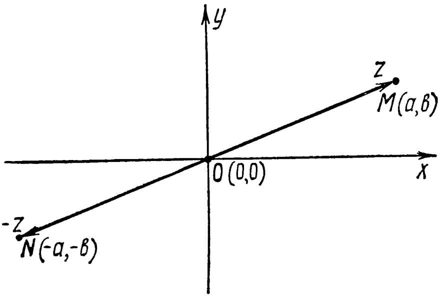


Рис. 144

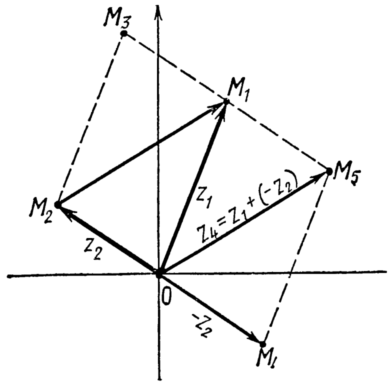


Рис. 145

Векторы, изображающие комплексные числа $z=a+bi$ и $(-z)=-a-bi$, расположены симметрично относительно начала координат, поскольку концы этих векторов, точки $M\ (a,\ b)$ и $N\,(-a,\ -b)$, симметричны относительно начала координат (рис. 144).

Пусть даны числа $z_1=a+bi$ и $z_2=c+di$. Рассмотрим вектор z_1, конец которого находится в точке $M_1(a,\ b)$, и вектор z_2, конец которого находится в точке $M_2(c,\ d)$ (рис. 145). На этих векторах построим параллелограмм $OM_1M_3M_2$. Рассмотрим вектор $(-z_2)$, конец которого находится в точке $M_4(-c,\ -d)$. Складывая векторы z_1 и $(-z_2)$ по правилу параллелограмма, получаем их сумму — вектор z_4, конец которого находится в точке $M_5(a-c,\ b-d)$. Очевидно, что длина вектора z_4 равна длине отрезка M_1M_2. Поскольку длина вектора z_4 равна $|z_4|=|z_1-z_2|$, то отсюда вытекает, что длина диагонали M_1M_2 равна $|z_1-z_2|$ и модуль разности двух комплексных чисел z_1 и z_2 представляет собой расстояние между точками M_1 и M_2, изображающими эти числа.

Такое геометрическое толкование суммы и модуля разности двух комплексных чисел часто применяется при решении задач.

Примеры. 1. Найти множество точек плоскости, соответствующих комплексным числам z, таким, что $|z-i|=|z+2|$.

Расстояния от искомых точек до точек, соответствующих комплексным числам i и (-2), равны. Значит, искомое множество состоит из точек прямой, перпендикулярной к отрезку, соединяющему точки $(0, 1)$ и $(-2,0)$, и проходящей через середину этого отрезка (рис. 146).

2. Найти точки z, удовлетворяющие условию $|z-1| = |z-2| = |z-i|$.

Точки, соответствующие числам 1, 2 и i, образуют треугольник. Условию задачи удовлетворяет единственная точка — центр описанной окружности этого треугольника. Так как она равноудалена от точек (1,0) и (2,0), то ее координата $x=3/2$, а так как она равноудалена от точек (1,0) и (0,1), то $y=x=3/2$, т. е. условию задачи удовлетворяет единственное число $z={}^3/_2+{}^3/_2 i$ (рис. 147).

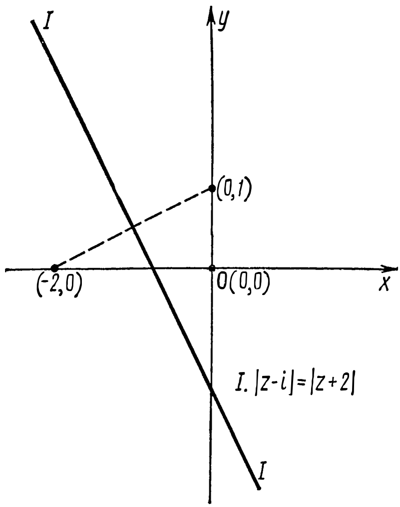


Рис. 146

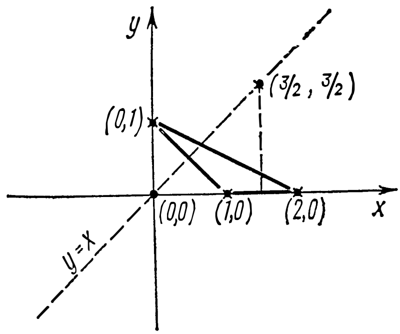


Рис. 147

Аргументом комплексного числа $z=a+bi$, отличного от нуля, называется любое из чисел φ, являющихся решением системы уравнений

$$\begin{cases} \cos\varphi = \dfrac{a}{\sqrt{a^2+b^2}}, \\ \sin\varphi = \dfrac{b}{\sqrt{a^2+b^2}}. \end{cases}$$

Для числа $z=0$ аргумент не определяется. Данная система имеет бесконечное множество решений, причем если φ_0 есть одно из ее решений, то все остальные решения получаются из него по формуле $\varphi=\varphi_0+2\pi k$, где k — любое целое число. Таким образом, любое комплексное число $z\neq 0$ имеет бесконечно много аргументов, отличающихся друг от друга на число, кратное 2π.

Главным аргументом комплексного числа $z=a+bi\neq0$ принято называть его аргумент из интервала $[0;\ 2\pi)$ и обозначать этот аргумент $\arg z$. Аргумент комплексного числа z имеет следующий геометрический смысл. Если комплексное число $z=a+bi\neq0$ рассматривать как вектор z, конец которого — точка $M(a,\ b)$, то величина угла φ, на который нужно повернуть против часовой стрелки положительную полуось OX до первого совмещения с вектором z, является главным аргументом комплексного числа z (см. рис. 140). Величина любого угла, отличающегося от главного аргумента $\arg z$ на целое число полных углов, является также аргументом рассматриваемого числа z.

Пример. Найти на плоскости точки z, удовлетворяющие условию $\pi/4\leqslant\arg z<5\pi/3$.

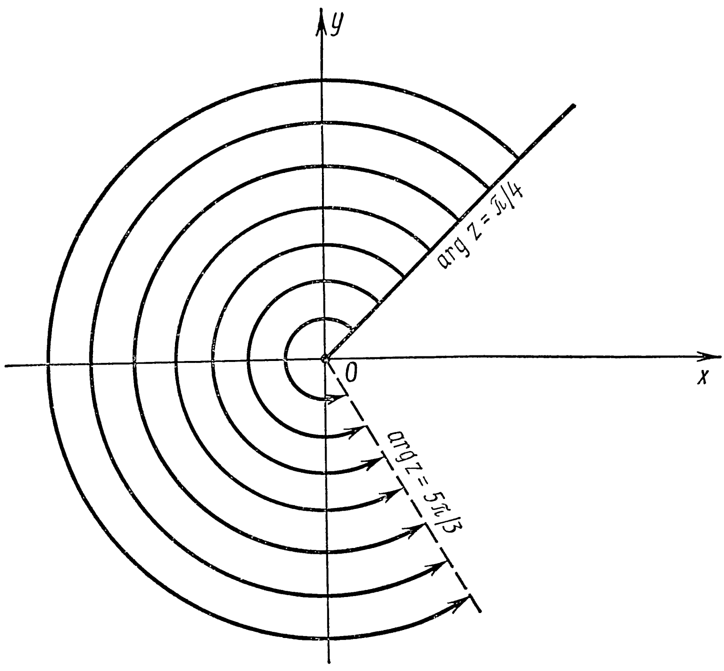


Рис. 148

Все точки, расположенные на каком-либо луче, выходящем из начала координат, имеют один и тот же главный аргумент, поэтому условию задачи удовлетворяют все точки той части плоскости, которая расположена между лучами, выходящими из начала координат под углами $\pi/4$ и $5\pi/3$, а также точки, лежащие на первом луче (рис. 148).

§ 3
ТРИГОНОМЕТРИЧЕСКАЯ ФОРМА КОМПЛЕКСНЫХ ЧИСЕЛ

Пусть $z=a+bi$ — некоторое комплексное число, отличное от нуля. Обозначим через r его модуль, а через φ — один из его аргументов. Тогда число z можно записать в виде

$$z=r(\cos\varphi+i\sin\varphi). \qquad (1)$$

Правая часть равенства (1) называется *тригонометрической формой комплексного числа* z. Тригонометрическая форма комплексного числа, отличного от нуля, определена однозначно: это запись комплексного числа z в виде (1), где r — положительное число, равное модулю числа z, косинус и синус берутся от одного и того же угла φ, равного аргументу числа z, при этом между косинусом и синусом стоит знак плюс. Ясно, что следующие комплексные числа записаны не в тригонометрической форме: $z_1=\cos{}^{\varphi}/_2+i\sin(\pi-{}^{\varphi}/_2)$; $z_2=-2(\cos{}^{\pi}/_3+i\sin{}^{\pi}/_3)$; $z_3=\sin{}^{\pi}/_6+i\cos{}^{\pi}/_6$; $z_4=\cos{}^{\pi}/_4-i\sin{}^{\pi}/_4$. Тригонометрическая форма этих комплексных чисел следующая: $z_1=\cos{}^{\varphi}/_2+i\sin{}^{\varphi}/_2$; $z_2=2(\cos{}^{4\pi}/_3+i\sin{}^{4\pi}/_3)$; $z_3=\cos{}^{\pi}/_3+i\sin{}^{\pi}/_3$; $z_4=\cos{}^{7\pi}/_4+$ $+i\sin{}^{7\pi}/_4$.

Действия умножения, деления, возведения в целую степень над любыми комплексными числами удобнее производить, если они записаны в тригонометрической форме.

Теорема 1. Модуль произведения двух комплексных чисел равен произведению их модулей, а аргумент равен сумме аргументов.

Доказательство. Пусть даны комплексные числа $z_1=$ $=r_1(\cos\varphi_1+i\sin\varphi_1)$ и $z_2=r_2(\cos\varphi_2+i\sin\varphi_2)$. Рассмотрим число $z_3=z_1z_2$. Применяя правила действия над комплексными числами и формулы для косинуса и синуса суммы двух углов, имеем $z_3=z_1z_2=[r_1(\cos\varphi_1+i\sin\varphi_1)][r_2(\cos\varphi_2+i\sin\varphi_2)]=$ $=r_1r_2[(\cos\varphi_1+i\sin\varphi_1)(\cos\varphi_2+i\sin\varphi_2)]=r_1r_2[(\cos\varphi_1\cos\varphi_2-$ $-\sin\varphi_1\sin\varphi_2)+(\sin\varphi_1\cos\varphi_2+\cos\varphi_1\sin\varphi_2)i]=$ $=r_1r_2[\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)]$.
Итак, $z_3=r_1r_2[\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)]$, т. е. число z_3 записано в тригонометрической форме и его модуль равен r_1r_2, а аргумент равен $(\varphi_1+\varphi_2)$, что и доказывает теорему.

Теорема 2. Модуль частного двух комплексных чисел равен частному их модулей, а аргумент равен разности аргументов.

Доказательство этой теоремы проводится легко, если воспользоваться тем, что деление — операция, обратная умножению.

Теорема 3 (формула Муавра). Пусть z — любое, отличное от нуля, комплексное число, n — любое целое число, тогда

$$z^n=[r(\cos\varphi+i\sin\varphi)]^n=r^n(\cos n\varphi+i\sin n\varphi). \qquad (2)$$

1. Докажем эту формулу для любого натурального n методом математической индукции.

Для $n=1$ формула верна. Предположим, что формула (2) верна для $n=k$, т. е. справедливо равенство

$$z^k=[r(\cos\varphi+i\sin\varphi)]^k=r^k(\cos k\varphi+i\sin k\varphi). \qquad (3)$$

Докажем, что из справедливости равенства (3) вытекает, что формула (2) справедлива и для $n=k+1$. Применяя формулу (3), правила действий над комплексными числами, формулы для синуса и косинуса суммы двух углов, имеем

$$z^{k+1}=z^kz=[r^k(\cos k\varphi+i\sin k\varphi)][r(\cos\varphi+i\sin\varphi)]=$$
$$=r^{k+1}[(\cos k\varphi\cos\varphi-\sin k\varphi\sin\varphi)+i(\sin k\varphi\cos\varphi+$$

$+\cos k\varphi\sin\varphi)]=r^{k+1}[\cos(k+1)\varphi+i\sin(k+1)\varphi]$, т. е. формула (2) доказана для $n=k+1$. Следовательно, по методу математической индукции формула (2) справедлива для любого натурального n.

2. Если $n=0$ и $z\neq0$, то по определению $z^0=1$, поэтому $z^0=1(\cos 0\varphi+i\sin 0\varphi)$, т. е. формула (2) верна для $n=0$.

3. Пусть $n=-1$. Применяя определение степени с целым отрицательным показателем и теорему 2, получим, что

$$z^{-1}={}^1/_z={}^{(\cos 0+i\sin 0)}/_{[r(\cos\varphi+i\sin\varphi)]}=\frac{1}{r}[\cos(-\varphi)+i\sin(-\varphi)],$$

т. е. для $n=-1$ формула (2) верна.

4. Пусть n — любое целое отрицательное число, тогда $n=-m$, где $m=|n|$ — натуральное число. Применяя сначала определение степени с целым отрицательным показателем, затем справедливость формулы (2) сначала для $n=-1$, затем для любого натурального m, имеем

$$z^n=z^{-m}=({}^1/_z)^m=\{({}^1/_r)[\cos(-\varphi)+i\sin(-\varphi)]\}^m=$$
$$=({}^1/_r)^m[\cos(-m\varphi)+i\sin(-m\varphi)]=r^n(\cos n\varphi+i\sin n\varphi).$$

Итак, формула (2) верна для любых целых n.

Приведем пример на применение формулы Муавра. Вычислим $\sin 3x$ и $\cos 3x$ через $\sin x$ и $\cos x$. По формуле Муавра $(\cos x+i\sin x)^3=\cos 3x+i\sin 3x$. В то же время по формуле бинома Ньютона $(\cos x+i\sin x)^3=\cos^3 x+i\,3\cos^2 x\sin x+$ $+3\cos x(i\sin x)^2+(i\sin x)^3=(\cos^3 x-3\cos x\sin^2 x)+$ $+i(3\cos^2 x\sin x-\sin^3 x)$. Итак, по правилу равенства комплексных чисел имеем

$$\sin 3x=3\cos^2 x\sin x-\sin^3 x,$$
$$\cos 3x=\cos^3 x-3\cos x\sin^2 x.$$

Замечание. Эти же формулы были получены в гл. V другим способом. Используя формулы Муавра и бинома Ньютона, можно вычислить $\cos nx$ и $\sin nx$ для любого натурального числа n. Тригонометрическая форма комплексных чисел позволяет доказать свойства модулей комплексных чисел:

а) $|z_1 z_2| = |z_1||z_2|$;

б) $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$, если $z_2 \neq 0$;

в) $|z_1 + z_2| \leqslant |z_1| + |z_2|$;

г) $|z_1 - z_2| \leqslant |z_1| + |z_2|$;

д) $||z_1| - |z_2|| \leqslant |z_1 + z_2|$;

е) $||z_1| - |z_2|| \leqslant |z_1 - z_2|$.

Докажем эти свойства. Пусть $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_2)$, $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$. По теореме 1 $z_1 z_2 = r_1 r_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)]$. По определению модуля комплексного числа $|z_1 z_2| = \sqrt{(r_1 r_2)^2[\cos^2(\varphi_1 + \varphi_2) + \sin^2(\varphi_1 + \varphi_2)]} = r_1 r_2 = |z_1|\,z_2|$. Свойство а) доказано.

Свойство б) доказывается аналогично.

Докажем свойство в). Поскольку $(z_1 + z_2) = (r_1\cos\varphi_1 + r_2\cos\varphi_2) + i(r_1\sin\varphi_1 + r_2\sin\varphi_2)$, то

$$z_1 + z_2| = \sqrt{r_1^2\cos^2\varphi_1 + 2r_1r_2\cos\varphi_1\cos\varphi_2 + r_2^2\cos^2\varphi_2 + r_1^2\sin^2\varphi_1 + \sqrt{2r_1r_2\sin\varphi_1\sin\varphi_2 + r_2^2\sin^2\varphi_2}} = \sqrt{r_1^2 + r_2^2 + 2r_1r_2\cos(\varphi_1 - \varphi_2)}.$$

Поскольку $\cos(\varphi_1 - \varphi_2) \leqslant 1$, то $|z_1 + z_2| \leqslant \sqrt{r_1^2 + r_2^2 + 2r_1r_2} = r_1 + r_2 = |z_1| + |z_2|$. Свойство в) доказано.

Свойство г) доказывается аналогично.

Свойство д) вытекает из свойства г). Действительно,

$|z_1| = |(z_1 + z_2) - z_2| \leqslant |z_1 + z_2| + |z_2|$, откуда

$$|z_1| - |z_2| \leqslant |z_1 + z_2|. \tag{4}$$

Аналогично $|z_2| = |(z_2 + z_1) - z_1| \leqslant |z_2 + z_1| + |z_1|$, откуда

$$|z_2| - |z_1| \leqslant |z_1 + z_2|. \tag{5}$$

Из справедливости неравенств (4) и (5) и вытекает справедливость свойства д).

Свойство е) доказывается аналогично.

§ 4
СВОЙСТВА КОРНЕЙ ИЗ КОМПЛЕКСНЫХ ЧИСЕЛ

В главе IV решалась задача: для любого данного числа a и любого данного натурального числа n найти число b, такое, чтобы $b^n = a$. Такие числа b принято называть корнями степени n из числа a. Там было показано, что если число a — действительное положительное число, а n — четное натуральное число, то существуют два действительных числа b_1 и b_2, таких, что $b_1^n = a$ и $b_2^n = a$; если a — действительное число, а n — нечетное натуральное число, то существует одно действительное число b, такое, что $b^n = a$.

Однако найти число, четная степень которого была бы равна отрицательному числу, в области действительных чисел уже нельзя. В области комплексных чисел это сделать можно. Справедливо более общее утверждение: в множестве комплексных чисел можно найти корень любой натуральной степени из любого комплексного числа. Это утверждение является следствием следующей теоремы.

Теорема 1. Пусть z — комплексное число, $z \neq 0$, n — натуральное число. Существует n различных комплексных чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$, таких, что $\alpha_i^n = z$, где $i = 1, 2, \ldots, n$.

Эти числа называются *корнями степени n из комплексного числа z*. Для обозначения этих корней нет специальных символов, подобных символу, которым обозначается арифметический корень ($\sqrt[n]{a} = b \Leftrightarrow a = b^n$, $a > 0$, $b > 0$, n — натуральное число).

Доказательство. Для $n = 1$ теорема очевидна. Пусть $n \geqslant 2$ и пусть $z = r(\cos\varphi + i\sin\varphi)$, $z \neq 0$. Будем искать комплексное число $\alpha = \rho(\cos\psi + i\sin\psi)$, такое, что $\alpha^n = z$. Покажем, что такое число α существует и что таких чисел α, различных между собой, ровно n штук. По формуле Муавра $z = \alpha^n = \rho^n(\cos n\psi + i\sin n\psi)$. По определению модуля комплексного числа $|z| = \rho^n$, т. е. $r = \rho^n$, откуда $\rho = \sqrt[n]{r}$ (по определению модуля комплексного числа $z \neq 0$ числа ρ и r положительны, поэтому символ арифметического корня здесь употреблен правильно). Применяя определение равенства двух комплексных чисел, получаем, что одновременно справедливы равенства $\begin{cases} \cos n\psi = \cos\varphi, \\ \sin n\psi = \sin\varphi. \end{cases}$

Эти равенства одновременно выполнены тогда и только тогда, когда $n\psi = \varphi + 2\pi k$, где k — любое целое число, т. е. для $\psi = (\varphi + 2\pi k)/n$, где $k = 0, \pm 1, \pm 2\ldots$. Значит, числа α, такие, что каждое из них удовлетворяет равенству $\alpha^n = r(\cos\varphi + i\sin\varphi)$, существуют и могут быть записаны в виде

$$\alpha = \sqrt[n]{r}\left\{\cos\left[\frac{\varphi + 2\pi k}{n}\right] + i\sin\left[\frac{\varphi + 2\pi k}{n}\right]\right\}, \tag{6}$$

где $k = 0, \pm 1, \pm 2, \ldots$.

Итак, корней n-й степени из комплексного числа, отличного от нуля, бесконечно много. Но среди этих корней различных будет только n штук. Действительно, обозначая через α_p корень, вычисляемый по формуле (6) для $k=p$, получаем:

$$\alpha_0=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi}{n}\right]+i\sin\left[\frac{\varphi}{n}\right]\right\},$$

$$\alpha_1=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi+2\pi}{n}\right]+i\sin\left[\frac{\varphi+2\pi}{n}\right]\right\},$$

$$\alpha_2=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi+2\pi\cdot 2}{n}\right]+i\sin\left[\frac{\varphi+2\pi\cdot 2}{n}\right]\right\},$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\alpha_{n-1}=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi+2\pi(n-1)}{n}\right]+i\sin\left[\frac{\varphi+2\pi(n-1)}{n}\right]\right\},$$

$$\alpha_n=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi+2\pi n}{n}\right]+i\sin\left[\frac{\varphi+2\pi n}{n}\right]\right\}=\alpha_0,$$

$$\alpha_{n+1}=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi+2\pi(n+1)}{n}\right]+i\sin\left[\frac{\varphi+2\pi(n+1)}{n}\right]\right\}=\alpha_1,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\alpha_{2n}=\alpha_0,$$

$$\alpha_{-1}=\alpha_{n-1},$$

$$\alpha_{-2}=\alpha_{n-2},$$

$$\dots\dots$$

$$\alpha_{-n}=\alpha_0,$$

$$\dots\dots$$

Отсюда легко видеть, что для любого целого p справедливы равенства $\alpha_0=\alpha_{pn}$, $\alpha_1=\alpha_{pn+1}$, $\alpha_2=\alpha_{pn+2}$, $\dots$, $\alpha_{n-1}=\alpha_{pn+n-1}$.

Итак, различных корней n штук: α_0, α_1, α_2, $\dots$, α_{n-1}. Они могут быть вычислены по формуле

$$\alpha_k=\sqrt[n]{r}\left\{\cos\left[\frac{\varphi+2\pi k}{n}\right]+i\sin\left[\frac{\varphi+2\pi k}{n}\right]\right\}, \tag{7}$$

где $k=0,\ 1,\ 2,\ \dots,\ (n-1)$. Заметим, что из формулы (7) можно получить все n различных корней, если вместо k подставлять в нее любые n целых чисел, идущих подряд.

По определению, если $z=0$, то считается, что существует n чисел $\alpha_0=\alpha_1=\alpha_2=\dots=\alpha_{n-1}=0$, которые называются корнями n-й степени из нуля.

Пример. Найти корни третьей степени из числа $z=\sqrt{3}+i$.

Поскольку $r=2$, а $\varphi={}^{\pi}/_{6}$, то $\alpha_k=\sqrt[3]{2}\,\{\cos[{}^{(\pi/6+2\pi k)}/_{3}]+i\sin[{}^{(\pi/6+2\pi k)}/_{3}]\}$, где $k=1,\ 2,\ 3$, т. е. $\alpha_1=\sqrt[3]{2}\,[\cos({}^{13\pi}/_{18})+i\sin({}^{13\pi}/_{18})]$, $\alpha_2=\sqrt[3]{2}\,[\cos({}^{25\pi}/_{18})+i\sin({}^{25\pi}/_{18})]$, $\alpha_3=\sqrt[3]{2}\times[\cos({}^{\pi}/_{18})+i\sin({}^{\pi}/_{18})]$.

Рассмотрим частный случай: отыскания корня n-й степени из единицы, т. е. найдем числа α_k, такие, что $\alpha_k^n=1$. Поскольку $r=1$ и $\varphi=0$, то

$$\alpha_k=\cos({}^{2\pi k}/_{n})+i\sin({}^{2\pi k}/_{n}), \tag{8}$$

где $k=0,\ 1,\ 2,\ \dots,\ n-1$. Если $n=2m$ (четное число), то среди этих корней существует два действительных корня $\alpha_0=1$ и

$\alpha_m=-1$. Если $n=2m+1$ (нечетное число), то существует один действительный корень $\alpha_0=1$. Приведем еще и другие свойства корней n-й степени из единицы,

а) $|\alpha_k|=1$;

б) $\alpha_k\alpha_m=\alpha_{k+m}$;

в) $\alpha_k/\alpha_m=\alpha_{k-m}$;

г) $\alpha_k^m=\alpha_{km}$, где m — любое целое число (корень α_m находится по формуле (8), где вместо k следует взять m).

Докажем эти свойства. Свойство а) вытекает из определения модуля комплексного числа.

Для доказательства свойства б) воспользуемся теоремой о произведении комплексных чисел в тригонометрической форме:

$$\alpha_k\alpha_m=\cos(2\pi k/n+2\pi m/n)+i\sin(2\pi k/n+2\pi m/n)=\alpha_{k+m}.$$

Свойство в) доказывается аналогично.

Докажем свойство г). По формуле Муавра $\alpha_k^m=\cos[(2\pi k/n)m]+$ $+i\sin[(2\pi k/n)m]=\alpha_{km}$.

Рассмотрим теперь геометрическую интерпретацию корней n-й степени из единицы: $\alpha_0=1$,

$$\alpha_1=\cos(2\pi/n)+i\sin(2\pi/n),\ \alpha_2=\cos(4\pi/n)+i\sin(4\pi/n),$$
$$\alpha_3=\cos(6\pi/n)+i\sin(6\pi/n),\ \ldots,\ \alpha_{n-2}=$$
$$=\cos[2\pi(n-2)/n]+i\sin[2\pi(n-2)/n],$$
$$\alpha_{n-1}=\cos[2\pi(n-1)/n]+i\sin[2\pi(n-1)/n].$$

Очевидно, что точки $\alpha_0,\ \alpha_1,\ \ldots,\ \alpha_{n-1}$ будут вершинами правильного n-угольника, вписанного в единичную окружность, одна из вершин которого — точка $A_0(1,\ 0)$.

Пример. 1. Пусть $n=3$, тогда $\alpha_0=1$, $\alpha_1=\cos(2\pi/3)+$ $+i\sin(2\pi/3)$, $\alpha_2=\cos(4\pi/3)+i\sin(4\pi/3)$. Точки $A_0(1,\ 0)$, $A_1\left(-\frac{1}{2},\ \sqrt{3}/2\right)$, $A_2\left(-\frac{1}{2},\ -\sqrt{3}/2\right)$ являются вершинами правильного треугольника $A_0A_1A_2$, вписанного в единичную окружность (рис. 149).

2. Пусть $n=4$, тогда $\alpha_0=1$, $\alpha_1=i$, $\alpha_2=-1$, $\alpha_3=-i$. Точки $A_0(1,\ 0)$, $A_1(0,\ 1)$, $A_2=(-1,\ 0)$ и $A_3(0,\ -1)$ являются вершинами квадрата $A_0A_1A_2A_3$, вписанного в единичную окружность (рис. 150).

Приведем формулу для корня n-й степени из числа -1. Поскольку $r=1$ и $\varphi=\pi$, то $\alpha_k=\cos[(\pi+2\pi k)/n]+i\sin[(\pi+2\pi k)/n]$, где $k=0,\ 1,\ 2,\ \ldots,\ n-1$. Отсюда видно, что если $n=2m$ (четное число), то среди α_k нет ни одного действительного числа. Если $n=2m+1$ (нечетное число), то существует одно действительное число $\alpha_m=1$.

Вообще для любого положительного числа a и любого четного натурального числа n существует только два действительных числа b_1 и b_2, таких, что $b_1^n=b_2^n=a$. Действительно, поскольку $a=|a|\cdot 1$, то все корни n-й степени из этого числа вычисляются по формуле $b_k=\sqrt[n]{|a|}\,[\cos(2\pi k/n)+i\sin(2\pi k/n)]$.

Если $n=2m$, то среди этих чисел будут только два действительных числа $b_0=\sqrt[n]{|a|}$ и $b_m=-\sqrt[n]{|a|}$, что и утверждалось выше. Аналогично можно показать справедливость следующих утверждений.

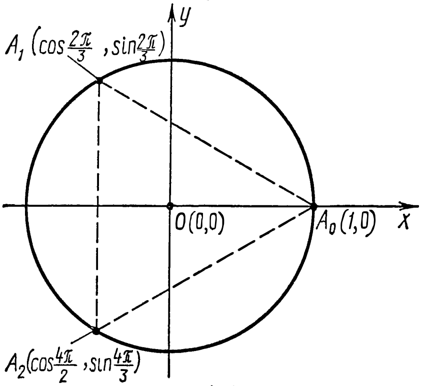


Рис. 149

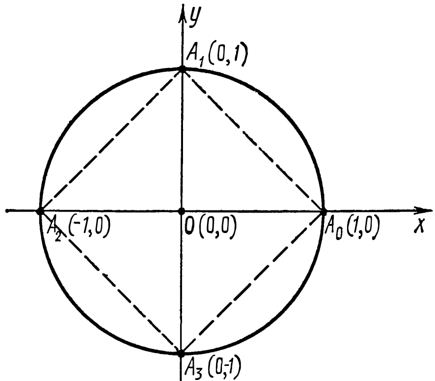


Рис. 150

Для любого положительного числа a и любого нечетного натурального числа n существует только одно действительное число $b=-\sqrt[n]{|a|}$, такое, что $b^n=a$.

Для любого отрицательного числа a и любого нечетного натурального числа n существует только одно действительное число b, такое, что $b^n=a$.

Для любого отрицательного числа a и любого четного натурального числа n не существует ни одного действительного числа b, такого, что $b^n=a$.

§ 5
СОПРЯЖЕННЫЕ КОМПЛЕКСНЫЕ ЧИСЛА

В § 1 приводилось определение числа, сопряженного к данному комплексному числу. Напомним его: комплексное число $\bar{z}=a-bi$ называется числом, *сопряженным* к комплексному числу $z=a+bi$. Легко видеть, что число, сопряженное к числу $\bar{z}$ есть число z, т. е. $\bar{\bar{z}}=z$ и поэтому числа $\bar{z}$ и z называются взаимно сопряженными. В геометрической интерпретации комплексных чисел взаимно сопряженные комплексные числа представляют собой точки, симметричные относительно действительной оси OX (рис. 151). Отметим ряд свойств взаимно сопряженных комплексных чисел.

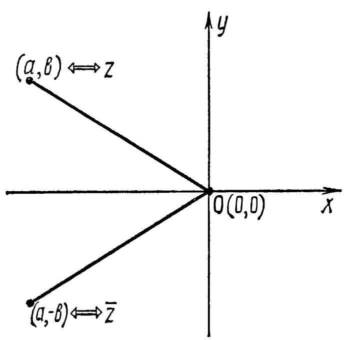


Рис. 151

а) $|\bar{z}|=|z|$.

Действительно, $|z|=\sqrt{a^2+b^2}$, $|\bar{z}|=\sqrt{a^2+(-b)^2}=\sqrt{a^2+b^2}$, т. е. $|\bar{z}|=|z|$.

б) Если $z=a$, где a — действительное число, то $\bar{z}=z$, $\arg\bar{z}=\arg z=0$.

в) Если $\bar{z}\neq z$, то $\arg\bar{z}=2\pi-\arg z$.

Действительно, если φ есть решение системы уравнений

$$\begin{cases}\cos\varphi=\dfrac{a}{\sqrt{a^2+b^2}},\\ \sin\varphi=\dfrac{b}{\sqrt{a^2+b^2}},\end{cases}$$

то число $\varphi_1=2\pi-\varphi$ есть решение системы уравнений

$$\begin{cases}\cos\varphi={}^{a}/_{\sqrt{a^2+b^2}},\\ \sin\varphi={}^{(-b)}/_{\sqrt{a^2+b^2}}.\end{cases}$$

г) $|z|=\sqrt{z\bar{z}}$.

Действительно, так как $|z|=\sqrt{a^2+b^2}$ и так как $z\bar{z}=$ $=(a+bi)(a-bi)=a^2+b^2$, то $|z|=\sqrt{z\bar{z}}$.

д) Сумма и произведение взаимно сопряженных комплексных чисел есть действительное число.

Действительно, если $z=a+bi$, то $z+\bar{z}=2a$ (действительное число) и $z\bar{z}=a^2+b^2$ (действительное число).

е) Число, сопряженное к сумме двух любых комплексных чисел, равно сумме чисел, сопряженных к слагаемым, т. е. $\overline{(z_1+z_2)}=\bar{z}_1+\bar{z}_2$.

Действительно, если $z_1=a+bi$, $z_2=c+di$, то $\bar{z}_1+\bar{z}_2=$ $=(a-bi)+(c-di)=(a+c)-(b+d)i=\overline{[(a+c)+(b+d)i]}=$ $=\overline{z_1+z_2}$.

ж) Число, сопряженное к произведению двух комплексных чисел, равно произведению чисел, сопряженных к сомножителям, т. е. $\overline{(z_1z_2)}=\bar{z}_1\bar{z}_2$.

Действительно, если $z_1=a+bi$, $z_2=c+di$, то $z_1z_2=(ac-bd)+$ $+i(ad+bc)$ и $\overline{(z_1z_2)}=(ac-bd)-i(ad+bc)$, а $\bar{z}_1\bar{z}_2=(a-bi)\times$ $\times(c-di)=(ac-bd)+i(-ad-bc)=(ac-bd)-i(ad+bc)$, т. е. $\overline{(z_1z_2)}=\bar{z}_1\bar{z}_2$.

з) Число, сопряженное к разности двух комплексных чисел, равно разности сопряженных чисел, т. е. $\overline{(z_1-z_2)}=\bar{z}_1-\bar{z}_2$.

Свойство з) доказывается так же, как и свойство е).

и) Если $z_2\neq 0$, то $\overline{(z_1/z_2)}=\bar{z}_1/\bar{z}_2$.

Свойство и) доказывается так же, как и свойство ж).

к) Число, сопряженное n-й степени комплексного числа z, равно n-й степени числа, сопряженного к числу z, т. е. $\overline{(z^n)}=$ $=(\bar{z})^n$, где n — натуральное число.

Свойство к) можно доказать методом математической индукции.

Глава XI
МНОГОЧЛЕНЫ, КОЛЬЦА, ПОЛЯ, ГРУППЫ

Понятия кольца, поля и группы широко используются в математике. Поэтому в § 1, 3 и 4 этой главы приводятся определения этих понятий и примеры множеств, являющихся либо кольцом, либо полем, либо группой.

В школе в основном изучаются многочлены с действительными коэффициентами. Однако свойства многочленов зависят от поля, над которым дан многочлен. Поэтому в § 2 этой главы изучаются свойства многочленов с коэффициентами из данного числового поля.

§ 1
ЧИСЛОВЫЕ ПОЛЯ И КОЛЬЦА

В этом параграфе рассматриваются *множества чисел*, причем под множеством чисел понимаются либо все комплексные числа, либо какая-нибудь их часть.

Ниже под *операцией* понимается одна из четырех арифметических операций: сложение, умножение, вычитание и деление.

Некоторое множество называется *замкнутым* относительно данной операции, если применение этой операции к любой паре чисел из этого множества дает число из этого же множества.

Примеры. 1. Множество натуральных чисел замкнуто относительно операции сложения, так как сумма любых двух натуральных чисел есть натуральное число.

2. Множество натуральных чисел не замкнуто относительно операции вычитания, так как разность двух натуральных чисел не всегда есть натуральное число, например $2-3=-1$ (-1 не натуральное число).

3. Множество, состоящее из чисел 0, 1 и -1, замкнуто относительно операции умножения.

4. Множество целых чисел замкнуто относительно операции вычитания, но не замкнуто относительно операции деления.

Множество чисел, замкнутое относительно операций сложения, умножения, вычитания и деления (кроме деления на нуль), называется *числовым полем*.

Примеры. 1. Множество всех комплексных чисел образует поле, называемое полем комплексных чисел.

2. Множество всех действительных чисел образует поле, называемое полем действительных чисел.

3. Множество всех рациональных чисел образует поле, называемое полем рациональных чисел.

4. Множество целых чисел не образует числового поля, так как оно не замкнуто относительно операции деления.

5. Множество чисел вида $p+q\sqrt{2}$, где p, q — любые рациональные числа, образует числовое поле.

Доказательство. а) Поскольку $(p_1+q_1\sqrt{2})+(p_2+q_2\sqrt{2})=(p_1+p_2)+(q_1+q_2)\sqrt{2}$, а (p_1+p_2) и (q_1+q_2) — рациональные числа, то замкнутость этого множества относительно операции сложения доказана. Аналогично доказывается замкнутость относительно вычитания.

б) Поскольку $(p_1+q_1\sqrt{2})(p_2+q_2\sqrt{2})=(p_1p_2+2q_1q_2)+(p_1q_2+p_2q_1)\sqrt{2}$, а $(p_1p_2+2q_1q_2)$ и $(p_1q_2+p_2q_1)$ — рациональные числа, то доказана замкнутость этого множества относительно операции умножения.

в) Поскольку $\frac{(p_1+q_1\sqrt{2})}{p_2+q_2\sqrt{2}}=\frac{(p_1+q_1\sqrt{2})(p_2-q_2\sqrt{2})}{(p_2+q_2\sqrt{2})(p_2-q_2\sqrt{2})}=\frac{p_1p_2-2q_1q_2}{p_2^2-2q_2^2}+\frac{q_1p_2-p_1q_2}{p_2^2-2q_2^2}\sqrt{2}$ и $p_2^2-2q_2^2\neq 0$, то замкнутость относительно операции деления очевидна.

Итак, множество чисел вида $p+q\sqrt{2}$, где p, q — любые рациональные числа, является полем.

6. Множество чисел вида $p+q\sqrt{2}$, где p и q — любые натуральные числа, не является полем, так как множество натуральных чисел не замкнуто относительно операции вычитания.

Множество чисел, замкнутое относительно операций сложения, вычитания и умножения, называется *числовым кольцом*.

Примеры. 1. Любое числовое поле есть числовое кольцо.

2. Множество целых чисел является числовым кольцом, так как это множество замкнуто относительно операций сложения, вычитания и умножения.

3. Множество четных чисел является числовым кольцом, так как при сложении, вычитании и умножении четных чисел вновь получается четное число.

4. Множество нечетных чисел не будет кольцом, так как уже при сложении нечетных чисел получается четное число, т. е. такое числовое множество не является замкнутым относительно операции сложения.

Рассмотрим *кольцо целых чисел.* Не выходя за пределы этого множества, не всегда удается разделить нацело целое число на натуральное число. Однако всегда возможно деление с остатком.

Разделить целое число n на натуральное число m с остатком — это значит найти два целых числа q и r, таких, что справедливо равенство $n=mq+r$, причем число r должно удовлетворять условию $0 \leqslant r < m$. Если $r=0$, то говорят, что целое число n делится нацело на натуральное число m.

Теорема 1. Пусть n — любое целое число, m — натуральное число. Тогда в кольце целых чисел существует единственная пара чисел q и r, удовлетворяющая условиям $n=mq+r$ и $0 \leqslant r < m$ (число q называется частным, а число r — остатком).

Доказательство. Рассмотрим сначала случай, когда число n есть натуральное число.

Пусть $n < m$, тогда пара чисел $q=0$, $r=n$ удовлетворяет условиям теоремы.

Пусть $n=m$, тогда пара чисел $q=1$, $r=0$ удовлетворяет условиям теоремы.

Пусть $n > m$, тогда пара чисел $q=1$ и $r_1=n-m$ удовлетворяет условиям $n=m \cdot 1+r_1$ и $r_1 > 0$. Если $r_1 < m$, то пара чисел $q=1$ и $r=r_1$ удовлетворяет условиям теоремы. Если $r_1=m$, то равенство $n=m \cdot 1+r_1$ можно записать в виде $n=$ $=m \cdot 2$ и тогда пара чисел $q=2$ и $r=0$ удовлетворяет условиям теоремы. Если $r_1 > m$, то число $r_2=r_1-m$ таково, что $r_1=m+r_2$ и $0 < r_2 < r_1$, и поэтому справедливо равенство $n=$ $=2m+r_2$. Если $r_2 < m$, то пара чисел $q=2$ и $r=r_2$ удовлетворяет условиям теоремы. Если $r_2=m$, то равенство $n=2m+r_2$ можно записать в виде $n=3m$, и тогда пара чисел $q=3$ и $r=0$ удовлетворяет условиям теоремы. Если $r_2 > m$, то повторяем тот же процесс до тех пор, пока на каком-то k-м шагу окажется, что $n=km+r_k$, при этом $0 < r_k \leqslant m$. Существование такого k-го шага вытекает из следующего свойства натуральных чисел: для любых двух натуральных чисел n и m, таких, что $n > m$, всегда найдется натуральное число p, такое, что $n < mp$. Если $0 < r_k < m$, тогда пара чисел $q=k$ и $r=r_k$ удовлетворяет условиям теоремы. Если $r_k=m$, то равенство $n=$ $=mk+r_k$ запишется в виде $n=m(k+1)$ и тогда пара чисел $q=(k+1)$ и $r=0$ удовлетворяет условиям теоремы.

Итак, если n — натуральное число, то доказано существование пары целых чисел q и r, удовлетворяющей условиям теоремы.

Пусть $n=0$, тогда пара чисел $q=0$ и $r=0$ удовлетворяет условиям теоремы.

Пусть $n < 0$. Тогда $|n|$ — натуральное число и по уже доказанному получаем, что существует пара целых чисел q_1 и r_1, такая, что $|n|=mq_1+r_1$ и $0 \leqslant r_1 < m$. В случае, если $r_1=0$, то из равенства $|n|=mq_1+r_1$ вытекает равенство $n=mq_0$, где $q_0=-q_1$, т. е. пара чисел $q=-q_1$ и $r=0$ удовлетворяет условиям теоремы. Если $0 < r_1 < m$, то из равенства $|n|=mq_1+r_1$ вытекает равенство $n=(-q_1-1)m+(m-r_1)$ и тогда пара чисел $q=-q_1-1$ и $r=m-r_1$ удовлетворяет условиям теоремы.

Итак, для любого целого числа n и любого натурального числа m существует пара целых чисел q и r, таких, что $n=$ $=mq+r$ и $0 \leqslant r < m$.

Покажем единственность такой пары целых чисел. Предположим, что есть две пары чисел q, r и q_1, r_1, удовлетворяющие условиям теоремы, т. е. такие, что

$$n = mq + r \text{ и } 0 \leqslant r < m, \quad n = mq_1 + r_1 \text{ и } 0 \leqslant r_1 < m.$$

Отсюда вытекает равенство $mq + r = mq_1 + r_1$ или равенство $m(q - q_1) = r_1 - r$.

Предположим для определенности, что $r_1 > r$, тогда $m > r_1 - r > 0$. Но тогда разность $q - q_1 > 0$, и поэтому получим, что в равенстве $m(q - q_1) = r_1 - r$ в правой части стоит натуральное число, меньшее, чем m, а в левой части — большее, чем m, и, следовательно, равенство $m(q - q_1) = r_1 - r$ является неверным. Аналогично, рассматривая случай $r_1 < r$, приходим к противоречию. Следовательно, $r = r_1$, тогда из равенства $m(q - q_1) = r_1 - r$ следует равенство $q = q_1$, т. е. если пара чисел q и r, удовлетворяющая условиям теоремы, существует, то она единственна.

Итак, теорема доказана.

Приведем некоторые следствия теоремы.

а) Любое четное число n может быть записано в виде $n = 2q$, где q — некоторое целое число.

б) Любое нечетное число n может быть записано в виде $n = 2q + 1$, где q — некоторое целое число.

в) Любое число n, делящееся нацело на три, может быть записано в виде $n = 3q$, где q — некоторое целое число.

г) Любое целое число n, не делящееся нацело на три, может быть записано в одном из следующих видов: $n = 3q + 1$ или $n = 3q + 2$, где q — некоторое целое число.

д) Любое целое число n, делящееся нацело на некоторое натуральное число k, может быть записано в виде $n = kq$, где q — некоторое целое число.

е) Любое целое число n, не делящееся нацело на некоторое натуральное число k, может быть записано в виде $n = kq + r$, где r — одно из чисел $1, 2, \ldots, k - 1$, а q — некоторое целое число.

В зависимости от делимости целых чисел на данное натуральное число k множество целых чисел можно разбить на k классов. Например, множество всех целых чисел разбивается на два класса: четные числа и нечетные числа. Множество всех целых чисел можно разбить на три класса:

а) числа, кратные числу три, т. е. числа вида $n = 3q$, где q — любое целое число;

б) числа, имеющие при делении на три остаток единицу, т. е. числа вида $n = 3q + 1$, где q — любое целое число;

в) числа, имеющие при делении на три остаток два, т. е. числа вида $n = 3q + 2$, где q — любое целое число.

Из приведенных примеров ясно, как разбить множество целых чисел на 4 класса, 5 классов и т. д.

Приведем примеры, показывающие, как разбиение целых чисел на классы помогает решать ряд задач.

1. Доказать, что ни для какого целого числа m число m^2+1 не делится нацело на три.

Доказательство. Разобьем множество всех целых чисел на три класса: а) $n=3q$; б) $n=3q+1$; в) $n=3q+2$, где q — любое целое число.

Пусть m — любое число из класса а). Тогда $m^2+1=9q^2+1= =3(3q^2)+1$, откуда видно, что при любом целом q число $3(3q^2)+1$ при делении на 3 дает остаток единица. Значит, при любом m из класса а) число m^2+1 не делится нацело на три.

Пусть m — любое число из класса б). Тогда $m^2+1=(3q+ +1)^2+1=3(3q^2+2q)+2$, откуда видно, что при любом целом q число $3(3q^2+2q)+2$ при делении на три дает остаток два. Значит, при любом m из класса б) число m^2+1 не делится нацело на три.

Аналогично показывается, что при любом m из класса в) число m^2+1 не делится нацело на три.

2. Доказать, что среди любых k последовательных целых чисел есть число, делящееся нацело на k.

Доказательство. Все целые числа можно разбить на следующие k классов: $n=kq$, $n=kq+1$, $n=kq+2$, ..., $n= =kq+(k-2)$, $n=kq+(k-1)$, где q — любое целое число.

Пусть даны k последовательных целых чисел, начинающихся с некоторого целого числа m, т. е. m, $(m+1)$, $(m+2)$, ..., $[m+(k-2)]$, $[m+(k-1)]$, и пусть число m содержится в классе $n=kq+i$ для некоторого i $(i=0, 1, 2, \ldots, k-2, k-1)$, т. е. пусть $m=kq_1+i$, где q_1 — некоторое целое число. Среди k последовательных чисел есть число $[m+(k-i)]-(kq_1+ +i+k-i)=k(q_1+1)$, которое делится нацело на k.

3. Доказать, что при любом целом числе m число m^3-m делится нацело на три.

Доказательство. Представим число m^3-m в виде $(m+1)m(m-1)$ и рассмотрим три последовательных целых числа: $(m-1)$, m и $(m+1)$. Среди них есть число, делящееся на три, поэтому и число m^3-m делится на три.

§ 2
МНОГОЧЛЕНЫ

Многочленом степени n (n — данное неотрицательное целое число) *от переменного x над числовым полем K* называется выражение вида

$$a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n,$$

где коэффициенты a_0, a_1, a_2, ..., a_{n-1}, a_n — данные числа из поля K, причем $a_0\neq 0$.

Из этого определения, в частности, вытекает, что многочлены нулевой степени — это отличные от нуля числа из поля K. Число нуль также считается многочленом, причем это единственный многочлен, степень которого не определена. Для сокращенной записи многочленов обычно употребляют следующие символы: $P(x)$, $Q(x)$, $T(x)$, $R(x)$, $p(x)$, $q(x)$, $r(x)$ и другие; при этом если хотят подчеркнуть, что многочлен $P(x)$ степени n, то пишут $P_n(x)$.

В этом параграфе многочлены рассматриваются в основном над полем комплексных чисел, поэтому дальше под многочленом понимается *многочлен над полем комплексных чисел*. В тех случаях, когда многочлены рассматриваются над другими числовыми полями, об этом будет сказано особо.

Два многочлена $P(x)$ *и* $Q(x)$ считаются *тождественно равными* (иногда говорят кратко — равными) тогда и только тогда, когда равны их степени и равны коэффициенты при x в одинаковых степенях. Для записи тождественного равенства многочленов употребляется знак равенства, если многочлены $P(x)$ и $Q(x)$ тождественно равны, то пишут $P(x)=Q(x)$. Значит, если

$$P_n(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n,$$
$$Q_m(x)=b_0x^m+b_1x^{m-1}+b_2x^{m-2}+\ldots+a_{m-1}x+a_m,$$

то $P_n(x)=Q_m(x)$ тогда и только тогда, когда $n=m$ и $a_{n-i}=$ $=b_{m-i}$ для всех $i=0,\ 1,\ 2,\ \ldots,\ n-1,\ n$.

Суммой двух многочленов

$$P_n(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a^n,$$
$$Q_m(x)=b_0x^m+b_2x^{m-1}+b_2x^{m-1}+\ldots+a_{m-1}x+a_m,$$

где $n\geqslant m$, называется многочлен $T(x)=c_0x^n+c_1x^{n-1}+c_2^{n-2}+\ldots$ $\ldots+c_{n-1}x+c_n$, такой, что $c_{n-i}=a_{n-i}+b_{m-i}$ для каждого $i=0,$ $1,\ 2,\ \ldots,\ n$, причем $b_{m-i}=0$ для каждого $i=m+1,\ m+2,\ \ldots$ $\ldots,\ n$, т. е. суммой многочленов $P_n(x)$ и $Q_m(x)$ называется многочлен $T(x)$, коэффициенты которого при каждой степени переменного x равны сумме коэффициентов при той же степени x в многочленах $P_n(x)$ и $Q_m(x)$, причем если $n>m$, то коэффициенты $b_{m+1},\ b_{m+2},\ \ldots,\ b_n$ следует считать равными нулю. Для нахождения суммы многочленов $P_n(x)$ и $Q_m(x)$ нужно записать подряд все члены этих двух многочленов и затем сделать приведение подобных членов.

Произведением двух многочленов

$$P_n(x)=a_0x^n+a_1x^{n-1}+a_2a^{n-2}+\ldots+a_{n-1}x+a_n,$$
$$Q_m(x)=b_0x^m+b_1x^{m-1}+b_2x^{m-2}+\ldots+b_{m-1}x+b_m$$

называется многочлен

$$R(x)=d_0x^{m+n}+d_1x^{n+m-1}+d_2x^{n+m-2}+\ldots+d_{n+m-1}x+d_{n+m},$$

такой, что $d_i=\sum\limits_{p+q=i}a_{n-p}b_{m-q}$ для каждого $i=0,\ 1,\ 2,\ \ldots$

$\ldots, n+m$, т. е. произведением многочленов $P_n(x)$ и $Q_m(x)$ называется многочлен $R(x)$, коэффициент которого d_i есть результат сложения всех произведений, в каждом из которых перемножаются такие коэффициенты a_k и b_e многочленов $P_n(x)$ и $Q_m(x)$, сумма индексов $k+e$ которых равна $n+m-i$. Для нахождения произведения многочленов $P_n(x)$ и $Q_m(x)$ нужно каждый член многочлена $P_n(x)$ умножить на каждый член многочлена $Q_m(x)$, сложить полученные произведения и привести подобные члены.

Нетрудно проверить, что справедливы следующие основные законы сложения и умножения многочленов:

1) $P(x)+Q(x)=Q(x)+P(x)$ (коммутативность сложения);

2) $[P(x)+Q(x)]+T(x)=P(x)+[Q(x)+T(x)]$ (ассоциативность сложения);

3) $P(x)Q(x)=Q(x)P(x)$ (коммутативность умножения);

4) $[P(x)Q(x)]T(x)=P(x)[Q(x)T(x)]$ (ассоциативность умножения);

5) $[P(x)+Q(x)]T(x)=P(x)T(x)+Q(x)T(x)$ (дистрибутивность сложения относительно умножения).

Вычесть из многочлена $P(x)$ *многочлен* $T(x)$ — это значит найти такой многочлен $Q(x)$, что $P(x)=T(x)+Q(x)$.

Нетрудно проверить, что для любых двух многочленов $P(x)$ и $T(x)$ такой многочлен $Q(x)$ существует и единствен, он называется *разностью многочленов* $P(x)$ и $T(x)$ и обозначается $Q(x)=P(x)-T(x)$.

Если $T(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n$,

$$T^*(x)=-T(x)=(-a_0)x^n+(-a_1)x^{n-1}+(-a_2)x^{n-2}+\ldots$$
$$\ldots+(-a_{n-1})x+(-a_n),$$

то $Q(x)=P(x)+T^*(x)$. Значит, в множестве многочленов всегда выполнена операция вычитания, обратная к операции сложения.

Разделить нацело многочлен $P(x)$ *на многочлен* $T(x)$, отличный от нуля, — это значит найти многочлен $Q(x)$, такой, что $P(x)=T(x)Q(x)$.

Если такой многочлен $Q(x)$ существует, то говорят, что многочлен $T(x)$ является делителем многочлена $P(x)$, а многочлен $Q(x)$ называется частным от деления многочлена $P(x)$ на многочлен $T(x)$. Не всегда многочлен $P(x)$ можно разделить нацело на многочлен $T(x)$. Например, многочлен x^2+1 не делится нацело на многочлен $x+1$. Значит, в множестве многочленов не всегда выполнена операция деления нацело, обратная к операции умножения. Зато, как будет показано ниже, в множестве многочленов всегда выполнена операция деления с остатком.

Разделить с остатком многочлен $P(x)$ на многочлен $T(x)$, отличный от нуля, — это значит найти два многочлена $q(x)$ и $r(x)$, такие, что

$$P(x) = T(x)\,q(x) + r(x), \tag{1}$$

причем либо степень многочлена $r(x)$ строго меньше степени многочлена $T(x)$, либо $r(x)$ есть нуль.

В случае, если выполнено равенство (1), говорят, что многочлен $P(x)$ делится на многочлен $T(x)$ с остатком $r(x)$ и частным $q(x)$; если $r(x)=0$, то говорят, что многочлен $P(x)$ делится на многочлен $T(x)$ с остатком нуль или многочлен $P(x)$ делится нацело на многочлен $T(x)$.

Пример. Пусть $P(x)=x^7-x^6+2x^5+x^2$, $T(x)=x^2-x+2$. Тогда $P(x)=T(x)(x^5+1)+x-2$, т. е. многочлен $P(x)$ делится на многочлен $T(x)$ с остатком $r(x)=x-2$ и частным x^5+1. Отметим, что из равенства $P(x)=T(x)\,x^5+x^2$ не вытекает, что многочлен $P(x)$ делится на многочлен $T(x)$ с остатком x^2, ибо нарушено условие: степень остатка $r(x)$ должна быть строго меньше степени многочлена $T(x)$.

Теорема 2. Для любых двух многочленов $P(x)$ и $T(x)$, где $T(x)\neq 0$, существует пара многочленов $q(x)$ и $r(x)$, таких, что $P(x)=T(x)\,q(x)+r(x)$, причем либо степень многочлена $r(x)$ строго меньше степени многочлена $T(x)$, либо $r(x)$ есть нуль.

Доказательство. Пусть $P(x)=0$, а $T(x)$ — любой отличный от нуля многочлен, тогда многочлены $q(x)=0$ и $r(x)=0$ удовлетворяют условиям теоремы.

Пусть $P(x)\neq 0$, а многочлен $T(x)$ имеет степень большую, чем степень многочлена $P(x)$, тогда многочлены $q(x)=0$ и $r(x)=P(x)$ удовлетворяют условиям теоремы.

Наконец, пусть $P(x)\neq 0$, а многочлен $T(x)$ имеет степень, меньшую или равную степени многочлена $P(x)$. Если $T(x)=c$, где c — константа, отличная от нуля, то многочлены $q(x)=(1/c)\,P(x)$ и $r(x)=0$ удовлетворяют условиям теоремы.

Остается рассмотреть случай, когда многочлен $P(x)$ имеет степень n, причем $n\geqslant 1$, а многочлен $T(x)$ имеет степень m, причем $0<m\leqslant n$. Пусть $P(x)=P_n(x)$, $T(x)=T_m(x)$, где $0<m\leqslant n$, $n\geqslant 1$, т. е.

$$P_n(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n,$$
$$T_{mm}(x) = b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m,$$

где $a_0\neq 0$, $b_0\neq 0$. Построим последовательность многочленов $Q_{n_k}(x)$ следующим образом. Положим $Q_{n_1}(x)=P_n(x)-\frac{a_0}{b_0}x^{n-m}T_m(x)$. Тогда многочлен $Q_{n_1}(x)$ можно записать в виде

$$Q_{n_1}(x) = a_0^{(1)}x^{n_1} + a_1^{(1)}x^{n_1-1} + \ldots + a_{n_1-1}^{(1)}x + a_{n_1}^{(1)},$$

причем $a_0^{(1)}\neq 0$ и степень многочлена $Q_{n_1}(x)$ меньше, чем n, т. е. $n_1<n$; если окажется, что $n_1<m$, то многочлены $q(x)=\frac{a_0}{b_0}x^{n-m}$ и $r(x)=Q_{n_1}(x)$ удовлетворяют условиям теоремы; если

же $n_1 \geqslant m$, то делаем следующий шаг: положим $Q_{n_2}(x) = Q_{n_1}(x) - \frac{a_0^{(1)}}{b_0} x^{n_1 - m} T_m(x)$. Ясно, что $n_2 < n_1 < n$ и $Q_{n_2}(x)$ можно записать в виде

$$Q_{n_2}(x) = a_0^{(2)} x^{n_2} + a_1^{(2)} x^{n_2 - 1} + \ldots + a_{n_2 - 1}^{(2)} x + a_{n_2}^{(2)},$$

причем $a_0^{(2)} \neq 0$. Если окажется, что $n_2 < m$, то многочлены $q(x) = \frac{a_0}{b_0} x^{n-m} + \frac{a_0^{(1)}}{b_0} x^{n_1 - m}$ и $r(x) = Q_{n_2}(x)$ удовлетворяют условиям теоремы; если же $n_2 \geqslant m$, то делаем следующий шаг и продолжаем этот процесс.

Поскольку на каждом шагу степень уменьшается: $n > n_1 > n_2 > \ldots$, то на некотором k-м шагу натуральное число n_k станет меньше натурального числа m и процесс закончится. В результате получим, что $P_n(x) = \left(\frac{a_0}{b_0} x^{n-m} + \frac{a_0^{(1)}}{b_0} x^{n_1 - m} + \ldots + \frac{a_0^{(k-1)}}{b_0} x^{n_{k-1} - m}\right) T_m(x) + Q_{n_k}(x)$. Тогда многочлены $r(x) = Q_{n_k}(x)$ и $q(x) = \left(\frac{a_0}{b_0} x^{n-m} + \frac{a_0^{(1)}}{b_0} x^{n_1 - m} + \ldots + \frac{a_0^{(k-1)}}{b_0} x^{n_{k-1} - m}\right)$ удовлетворяют условиям теоремы.

Итак, утверждение теоремы о существовании многочленов $q(x)$, $r(x)$ доказано.

Теорема 3. Пара многочленов $q(x)$, $r(x)$, удовлетворяющая условиям теоремы 2, единственна.

Доказательство. Предположим противное, т. е. предположим, что существуют две пары многочленов $q(x)$, $r(x)$ и $q_1(x)$, $r_1(x)$, таких, что $P(x) = T(x) q(x) + r(x)$ и $P(x) = T(x) q_1(x) + r_1(x)$. Пользуясь определением равенства многочленов, имеем $T(x)[q(x) - q_1(x)] = r_1(x) - r(x)$. Возможны два случая: либо $r_1(x) = r(x)$, тогда так как $T(x) \neq 0$, то $q(x) - q_1(x) = 0$, и единственность доказана; либо $r_1(x) - r(x) \neq 0$, тогда так как степень разности двух многочленов не больше степени любого из многочленов разности, то степень $r_1(x) - r(x)$ меньше степени многочлена $T(x)$, но степень многочлена $T(x)[q(x) - q_1(x)]$ либо больше, либо равна степени многочлена $T(x)$. Полученное противоречие означает, что $r_1(x) - r(x) = 0$, а в этом случае единственность уже доказана. Итак, теорема доказана.

Объединяя теоремы 2 и 3, получаем справедливость следующей теоремы.

Теорема 4. Для любых двух многочленов $P(x)$ и $T(x)$, где $T(x) \neq 0$ существует и притом единственная пара многочленов $q(x)$ и $r(x)$, таких, что $P(x) = T(x) q(x) + r(x)$, причем либо степень многочлена $r(x)$ строго меньше степени многочлена $T(x)$, либо $r(x)$ есть нуль.

Для определения коэффициентов многочленов $q(x)$ и $r(x)$ существует несколько способов. Наиболее распространенным среди них является *метод неопределенных коэффициентов*. Рассмотрим его. Напишем многочлены

$$q(x) = c_0 x^{n-m} + c_1^{n-m-1} + \ldots + c_{n-m} \qquad (c_0 \neq 0),$$

$$r(x) = d_0 x^{m-1} + d_1 x^{m-2} + \ldots + d_{m-1},$$

где коэффициенты c_i и d_j пока неопределенны (отметим, что их всего $n+1$), и потребуем, чтобы было справедливо равенство

$$P_n(x) = T_m(x)\, q(x) + r(x).$$

Раскрывая скобки и приравнивая коэффициенты при одинаковых степенях x в левой и правой частях, получим систему $n+1$ линейных уравнений с $n+1$ неизвестными, откуда и определим коэффициенты c_0, c_1, c_2, $\ldots$, c_{n-m}, d_0, d_1, d_2, $\ldots$, d_{m-1}, и тем самым найдем многочлены $q(x)$ и $r(x)$.

Рассмотрим пример. Пусть $P(x) = 2x^5 - 5x^2 + 2$, $T(x) = 2x^3 - 3x$. Полагая $q(x) = c_0 x^2 + c_1 x + c_2$ $(c_0 \neq 0)$, $r(x) = d_0 x + d_1$, напишем равенство $2x^5 - 5x^3 + 2 = (2x^3 - 3x)(c_0 x^2 + c_1 x + c_2) + (d_0 x + d_1)$, которое можно переписать в виде $2x^5 - 5x^2 + 2 = 2c_0 x^5 + 2c_1 x^4 + (2c_2 - 3c_0) x^3 - 3c_1 x^2 + (d_0 - 3c_2) x + d_1$, откуда получим систему уравнений

$$\begin{cases} 2c_0 = 2, \\ 2c_1 = 0, \\ 2c_2 - 3c_0 = -5, \\ -3c_1 = 0, \\ d_0 - 3c_1 = 0, \\ d_1 = 2. \end{cases}$$

Решая ее, получаем, что $q(x) = x^2 - 1$, $r(x) = -3x + 2$.

Рассмотрим деление многочлена на двучлен $(x - \alpha)$.

Пусть дан многочлен $P_n(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n$, где $a_0 \neq 0$, $n \geqslant 1$, и двучлен $(x - \alpha)$. По теореме 1 существует многочлен $q(x)$ и число r (т. е. многочлен степени нуль), такие, что $P_n(x) = (x - \alpha)\, q(x) + r$. Степень многочлена $q(x)$ равна $(n-1)$. Поэтому $q(x) = b_0 x^{n-1} + b_1 x^{n-2} + \ldots + b_{n-2} x + b_{n-1}$, где $b_0 \neq 0$. Найдем числа b_0, b_1, b_2, $\ldots$, b_{n-1} и r методом неопределенных коэффициентов. Подставим $q(x)$ в равенство $P(x) = (x - \alpha)\, q(x) + r$, получим, что

$$a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n = b_0 x^n + (b_1 - \alpha b_0) x^{n-1} + (b_2 - \alpha b_1) x^{n-2} + \ldots + (b_{n-1} - \alpha b_{n-2}) x + (r - \alpha b_{n-1}).$$

По правилу равенства многочленов, отсюда получаем, что

$$\begin{cases} a_0 = b_0, \\ a_1 = b_1 - \alpha b_0, \\ a_2 = b_2 - \alpha b_1, \\ \cdots\cdots \\ a_{n-1} = b_{n-1} - \alpha b_{n-2}, \\ a_n = r - \alpha b_{n-1}, \end{cases}$$

откуда

$$\begin{cases} b_0 = a_0, \\ b_1 = a_1 + \alpha b_0, \\ b_2 = a_2 + \alpha b_1, \\ \cdots\cdots\cdots \\ b_{n-1} = a_{n-1} + \alpha b_{n-2}, \\ r = a_n + \alpha b_{n-1}. \end{cases} \tag{2}$$

Итак, коэффициенты частного — многочлена $q(x)$, и остаток r выражаются через коэффициенты многочлена $P(x)$ и число α при помощи действий сложения и умножения по формулам (2), откуда следует:

а) если $a_0, a_1, a_2, \ldots, a_n$ и α — вещественные числа, то $b_0, b_1, b_2, \ldots, b_{n-1}$ и r — также вещественные числа;

б) если $a_0, a_1, a_2, \ldots, a_n$ и α — рациональные числа, то $b_0, b_1, b_2, \ldots, b_{n-1}$ и r — также рациональные числа;

в) если $a_0, a_1, a_2, \ldots, a_n$ и α — целые числа, то $b_0, b_1, b_2, \ldots, b_{n-1}$ и r — также целые числа.

Из формул (2) вытекает следующее правило для вычисления коэффициентов $b_0, b_1, b_2, \ldots, b_{n-1}$ и остатка r.

Выписать подряд, начиная с a_0, в строку все коэффициенты многочлена $P_n(x)$, во второй строке под a_0 написать коэффициент b_0, равный коэффициенту a_0, слева от b_0 во второй строке написать α; затем, умножая α на b_0 и прибавляя произведение к a_1, получим коэффициент b_1, который и напишем под a_1 во второй строке, затем, умножая α на b_1 и прибавляя произведение к a_2, получим коэффициент b_2, который напишем под a_2 во второй строке; продолжая этот процесс, получим, наконец, коэффициент b_{n-1}. Умножая теперь α на b_{n-1} и прибавляя произведение к a_n, получим остаток r, который и поместим под a_n во второй строке.

Это правило записывается в виде следующей таблицы, которая называется *схемой Горнера*:

$$\begin{array}{ccccccccc} a_0 & & a_1 & & a_2 & & a_3 & \ldots & a_{n-1} & & a_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ a_0=b_0 & - & \alpha b_0+a_1=b_1 & - & \alpha b_1+a_2=b_2 & - & \alpha b_2+a_3=b_3 & \ldots & \alpha b_{n-2}+a_{n-1}=b_{n-1} & - & \alpha b_{n-1}+a_n=r \end{array}$$

Схема Горнера позволяет легко разделить многочлен $P(x)$ на двучлен $x-\alpha$, т. е. найти коэффициенты частного $q(x)$ и остаток r.

Пример. Применяя схему Горнера, найдем частное $q(x)$ и остаток r при делении многочлена $P(x)=2x^5-x^4-3x^3-x-3$ на многочлен $T(x)=x-3$.

Схема Горнера имеет вид

$$\begin{array}{ccccccc}
 & 2 & -1 & -3 & 0 & 1 & -3 \\
3 & 2= & -1+3\cdot 2= & -3+3\cdot 5= & 0+3\cdot 12= & 1+3\cdot 36= & -3+3\cdot 109= \\
 & =2 & =5 & -12 & =36 & =109 & =324
\end{array}$$

Таким образом, $2x^5-x^4-3x^3-x-3=(x-3)(2x^4+5x^3+12x^2+$ $+36x+109)+324$.

Теорема 5 (*теорема Безу*). Остаток от деления многочлена $P(x)$ на двучлен $(x-\alpha)$ равен значению многочлена $P(x)$ при $x=\alpha$, т. е. $r=P(\alpha)$.

Доказательство. Подставив в равенство $P(x)=(x-\alpha)\times$ $\times q(x)+r$ вместо x значение α, получим $P(\alpha)=(\alpha-\alpha)q(\alpha)+r$, откуда и вытекает, что $r=P(\alpha)$.

Теорема 6. Многочлен $P(x)$ делится нацело на двучлен $(x-\alpha)$ тогда и только тогда, когда значение многочлена при $x=\alpha$ равно нулю, т. е. $P(\alpha)=0$.

Доказательство. Необходимость. Пусть многочлен $P(x)$ делится нацело на двучлен $(x-\alpha)$. Это значит, что остаток r равен нулю. С другой стороны, по теореме Безу остаток $r=P(\alpha)$. Следовательно, $P(\alpha)=0$. Достаточность. Пусть $P(\alpha)=0$. С другой стороны, по теореме Безу $r=P(\alpha)$. Значит, $r=0$, т. е. $P(x)$ делится нацело на $(x-\alpha)$.

Приведем несколько следствий из этой теоремы.

1. Многочлен $P_n(x)=x^n-\alpha^n$ делится нацело на двучлен $(x-\alpha)$ при любом натуральном n. Действительно, $P_n(\alpha)=$ $=\alpha^n-\alpha^n=0$.

2. Многочлен $P_n(\alpha)=x^n-\alpha^n$ делится нацело на двучлен $(x+\alpha)$ при любом четном $n=2m$. Действительно, $P_{2m}(-\alpha)=$ $=(-\alpha)^{2m}-\alpha^{2m}=0$.

3. Многочлен $P_n(x)=x^n+\alpha^n$ делится нацело на двучлен $(x+\alpha)$ при любом нечетном $n=2m+1$. Действительно, $P_{2m+1}(-\alpha)=(-\alpha)^{2m+1}+\alpha^{2m+1}=0$. Теперь перейдем к изучению корней многочленов.

Число α называется *корнем многочлена* $P(x)$, если $P(\alpha)=0$. Сформулируем теорему 6, используя определение корня многочлена.

Теорема 7. Число α является корнем многочлена $P(x)$ тогда и только тогда, когда многочлен $P(x)$ делится нацело на двучлен $(x-\alpha)$.

Возникает вопрос: всякий ли многочлен имеет корень? Ответ на этот вопрос дает основная теорема алгебры.

Основная теорема алгебры. Любой многочлен степени n ($n \geqslant 1$) над полем комплексных чисел имеет хотя бы один корень.

Эта теорема принимается здесь без доказательства. Докажем в качестве следствия этой теоремы такую теорему.

Теорема 8. Любой многочлен $P_n(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n$ ($a_0 \neq 0$) степени n ($n \geqslant 1$) над полем комплексных чисел разлагается в произведение n линейных множителей, т. е.

$$P_n(x) = a_0 (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \ldots (x - \alpha_n). \qquad (3)$$

Доказательство. Пусть P — многочлен степени n. Тогда по основной теореме алгебры у него есть корень. Обозначим его α_1. Тогда по теореме 7 $P_n(x) = (x - \alpha_1) q_1(x)$, где $q_1(x)$ — многочлен степени $(n-1)$. Для многочлена $q_1(x)$ опять применима основная теорема алгебры, поэтому $q_1(x)$ имеет корень α_2. Тогда по теореме 7 $q_1(x) = (x - \alpha_2) q_2(x)$, где $q_2(x)$ есть многочлен степени $(n-2)$. Продолжая этот процесс, получим, что $q_{n-1}(x) = (x - \alpha_{n-1}) q_n(x)$, где $q_n(x)$ есть многочлен первой степени, т. е. $q_n(x) = b_0 (x - \alpha_n)$ и $b_0 \neq 0$. Из этих рассуждений вытекает, что

$$P_n(x) = b_0 (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \ldots (x - \alpha_n). \qquad (4)$$

Раскрыв скобки в произведении, стоящем в правой части равенства (4), получим, что у многочлена $P_n(x)$ коэффициент при x^n будет b_0, а на самом деле он a_0, поэтому $b_0 = a_0$.

Итак, показано, что $P_n(x) = a_0 (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \ldots$ $\ldots (x - \alpha_n)$. Теорема доказана.

Заметим, что в равенстве (3) некоторые из чисел $\alpha_1, \alpha_2, \ldots \alpha_n$ могут быть равными. Объединяя вместе одинаковые линейные множители, равенство (3) можно переписать в виде

$$P_n(x) = a_0 (x - \alpha_1)^{k_1} (x - \alpha_2)^{k_2} (x - \alpha_3)^{k_2} \ldots (x - \alpha_m)^{k_m}, \qquad (5)$$

где $k_1 + k_2 + k_3 + \ldots + k_m = n$, при этом предполагается, что среди чисел $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_m$ уже нет равных.

Если многочлен $P(x)$ делится нацело на $(x - \alpha)^k$, где k — некоторое фиксированное число, но не делится нацело на $(x - \alpha)^{k+1}$, то число α называется *корнем кратности k многочлена $P_n(x)$*. Корни кратности единица называются *простыми корнями многочлена*. Можно показать, что в равенстве (5) число α_i является корнем кратности k_i многочлена $P_n(x)$. Из теоремы 8 вытекают следующие теоремы.

Теорема 9. Любой многочлен степени n над полем комплексных чисел имеет n корней, если каждый из корней считать столько раз, какова его кратность.

Теорема 10 (*формула Виета*). Если $P_n(x) = x^n + a_1 x^{n-1} + \ldots$

$\ldots + a_n$, т. е. если старший коэффициент многочлена равен единице ($a_0 = 1$), а числа $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ — его корни, то справедливы равенства

$$a_1 = -(\alpha_1 + \alpha_2 + \alpha_3 + \ldots + \alpha_n),$$
$$a_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \ldots \alpha_1\alpha_n + \alpha_2\alpha_3 + \ldots \alpha_2\alpha_n + \ldots + \alpha_{n-1}\alpha_n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{n-1} = (-1)^{n-1}(\alpha_1\alpha_2\ldots\alpha_{n-1} + \alpha_1\alpha_2\ldots\alpha_{n-2}\alpha_n + \ldots + \alpha_2\alpha_3\ldots\alpha_n),$$
$$a_n = (-1)^n(\alpha_1\alpha_2\alpha_3\ldots\alpha_{n-1}\alpha_n).$$

Доказательство. По теореме 8 $P_n(x) = (x - \alpha_1)(x - \alpha_2) \times \times (x - \alpha_3)\ldots(x - \alpha_n)$. Раскрывая скобки и пользуясь правилом равенства многочленов, получим справедливость этих равенств, которые называются формулами Виета.

Теорема 11. Если многочлен

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n \qquad (a_0 \neq 0)$$

с действительными коэффициентами $a_0, a_1, \ldots, a_n$ имеет комплексный корень $a + bi$, то он обязательно имеет и корень $a - bi$, т. е. число, сопряженное к корню многочлена с действительными коэффициентами, является также корнем этого многочлена.

Доказательство. Пусть число $a + bi$ — корень многочлена $P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$ $(a_0 \neq 0)$. Тогда $P_n(a + bi) = 0$, т. е. справедливо равенство $a_0(a + bi)^n + + a_1(a + bi)^{n-1} + a_2(a + bi)^{n-2} + \ldots + a_{n-1}(a + bi) + a_n = 0$. Применяя формулу бинома Ньютона, левую часть этого равенства можно записать $A + Bi$, откуда по правилу равенства нулю комплексного числа получаем, что $A = 0$ и $B = 0$.

Теперь рассмотрим $P_n(a - bi) = a_0(a - bi)^n + a_1(a - bi)^{n-1} + + a_2(a - bi)^{n-2} + \ldots + a_{n-1}(a - bi) + a_n$.

В гл. X доказано, что произведение, степень и алгебраическая сумма чисел, сопряженных с данными комплексными числами, есть комплексное число, сопряженное соответственно к произведению, степени и алгебраической сумме данных чисел, поэтому $P_n(a - bi) = A - Bi$, но так как $A = B = 0$, то $P_n(a - bi) = 0$, а это и значит, что число $(a - bi)$ является корнем многочлена $P_n(x)$.

Теорема доказана.

Теорема 12. Если многочлен с действительными коэффициентами имеет комплексный корень $a + bi$, то он делится нацело на квадратный трехчлен $x^2 + px + q$, где $p = -2a$, $q = a^2 + b^2$.

Доказательство. Если число $x_1 = a + bi$ — корень многочлена $P(x)$, то согласно теореме 11 число $x_2 = a - bi$ — также корень этого многочлена. Тогда $P(x)$ делится нацело и на двучлен $(x - x_1)$ и на двучлен $(x - x_2)$, значит, он делится на их произведение

$$(x-x_1)(x-x_2)=(x-a-bi)(x-a+bi)=$$
$$=x^2-2ax+a^2+b^2=x^2+px+q,$$

что и требовалось доказать.

Теорема 13. Всякий многочлен с действительными коэффициентами разлагается на произведение либо двучленов $(x-\alpha_k)$, либо трехчленов $x^2+p_mx+q_m$, либо двучленов $(x-\alpha_k)$ и трехчленов $x^2+p_mx+q_m$, где α_k, p_m, q_m — действительные числа, а трехчлены $x^2+p_mx+q_m$ не имеют действительных корней. Теорема 13 есть следствие теорем 8 и 12.

Пример. Известно, что многочлен $P(x)=x^4-2x^3+4x-4$ имеет корень $x_1=1+i$. Разложить его на произведение двучленов и трехчленов с действительными коэффициентами.

Поскольку многочлен $P(x)$ имеет корень $x_1=1+i$, то по теореме 12 он делится на трехчлен x^2-2x+2. Разделив многочлен $P(x)$ на этот трехчлен, получим $x^4-2x^3+4x-4=$ $=(x^2-2x+2)(x^2-2)$. Разлагая теперь многочлен x^2-2 на линейные множители, получаем ответ: $P(x)=(x^2-2x+2)\times$ $\times(x-\sqrt{2})(x+\sqrt{2})$.

Пусть дан многочлен

$$P_n(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n; \quad (a_0\neq 0). \tag{6}$$

Выше было показано:

1. Многочлен (6) над полем комплексных чисел может быть записан в виде

$$P_n(x)=a_0(x-\alpha_1)(x-\alpha_2)(x-\alpha_3)\ldots(x-\alpha_n), \tag{7}$$

где числа a_0, α_1, α_2, α_3, $\ldots$, α_n из этого же поля.

2. Многочлен (6) над полем действительных чисел может быть записан в виде

$$P_n(x)=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_k)(x^2+p_1x+q_1)\ldots(x^2+p_mx+q_m), \tag{8}$$

где $k+2m=n$, числа a_0, α_1, α_2, $\ldots$, α_k, p_1, q_1, $\ldots$, p_m, q_m — действительные числа. При этом в разложении (8) квадратные трехчлены нельзя представить как произведение двух множителей $(x-\alpha_i)(x-\alpha_j)$ с действительными α_i и α_j. Другими словами, эти квадратные трехчлены неразложимы в поле действительных чисел на множители степени 1.

Многочлен $P_n(x)$ $(n\geqslant 1)$ над полем K называется *приводимым в поле* K, если он может быть разложен в произведение двух многочленов $Q_m(x)$ и $T_k(x)$ с коэффициентами из поля K, степени которых m и k таковы, что $0<m<n$, $0<k<n$ и $m+k=n$. Многочлен $P_n(x)$ $(n\geqslant 1)$ над полем K называется *неприводимым в поле* K, если его нельзя разложить в произведение двух многочленов $Q_m(x)$ и $T_k(x)$ с коэффициентами из

поля K, степени каждого из которых меньше, чем n, но больше нуля.

Из этого определения следует, что любой многочлен первой степени над любым полем K неприводим в этом поле.

Из вышеизложенного вытекает, что в поле комплексных чисел неприводимыми многочленами являются только многочлены первой степени; в поле действительных чисел неприводимыми многочленами, кроме того, могут быть и многочлены второй степени.

Значит, теоремы 8 и 13 можно сформулировать так: любой многочлен над полем комплексных чисел разлагается в произведение неприводимых в этом поле многочленов первой степени; любой многочлен над полем действительных чисел разлагается в произведение неприводимых в этом поле многочленов либо первой, либо второй, либо и первой и второй степеней.

Заметим, что в поле рациональных чисел неприводимыми могут быть многочлены любой степени. Например, многочлены x^8+2 и $x^{10^{10}}+2$ неприводимы в поле рациональных чисел.

Перейдем теперь к *вычислению корней многочленов*. Начнем с многочлена первой степени $P_1(x)=a_0x+a_1$ $(a_0\neq 0)$. В любом числовом поле этот многочлен имеет один корень $x_1=-a_1/a_0$.

Рассмотрим многочлен второй степени $P_2(x)=a_0x^2+a_1x+a_2$ $(a_0\neq 0)$. Как следует из теоремы 9, многочлен $P_2(x)$ над полем комплексных чисел имеет два корня. Для нахождения этих корней преобразуем многочлен $P_2(x)$, выделив полный квадрат:

$$P_2(x)=a_0\left[x^2+2x\frac{a_1}{2a_0}+\left(\frac{a_1}{2a_0}\right)^2-\left(\frac{a_1}{2a_0}\right)^2+\frac{a_2}{a_0}\right]=$$
$$=a_0\left[\left(x+\frac{a_1}{2a_0}\right)^2-\frac{D}{4a_0^2}\right],$$

где $D=a_1^2-4a_0a_2$. Найдем (см. § 5 гл. X) комплексное число β, такое, что $\beta^2=D/4a_0^2$. Тогда $P_2(x)=a_0\left(x+\frac{a_1}{2a_0}-\beta\right)\left(x+\frac{a_1}{2a_0}+\beta\right)$, откуда очевидно, что многочлен $P_2(x)$ имеет два корня: $x_1=-\frac{a_1}{2a_0}+\beta$, $x_2=-\frac{a_1}{2a_0}-\beta$.

Заметим, что если $D\neq 0$, то существует два комплексных числа β_1 и β_2, таких, что $\beta_1^2=\beta_2^2=D/(4a_0^2)$. Так как $\beta_1=-\beta_2$ (см. § 5 гл. X), то для нахождения корней многочлена $P_2(x)$ все равно, какой из них взять за β. Если $D=0$, то $x_1=x_2=a_1/(2a_0)$ и многочлен $P_2(x)$ имеет один корень кратности два.

В случае, если коэффициенты a_0, a_1, a_2 — действительные числа, то D также действительное число и поэтому:

а) если $D>0$, то многочлен $P_2(x)$ имеет два действительных неравных корня: $x_1=(-a_1+\sqrt{D})/(2a_0)$, $x_2=(-a_1-\sqrt{D})/(2a_0)$;

б) если $D=0$, то многочлен $P_2(x)$ имеет один действительный корень кратности два: $x_1=x_2=-a_1/(2a_0)$;

в) если $D<0$, то многочлен $P_2(x)$ имеет два комплексных корня: $x_1=-a_1/(2a_0)+[\sqrt{\bar{D}}/(2a_0)]i$ и $x_2=-a_1/(2a_0)-[\sqrt{\bar{D}}/(2a_0)]i$.

Для любого многочлена третьей или четвертой степени над полем комплексных чисел существуют способы нахождения его корней, причем эти корни можно выразить через коэффициенты многочлена при помощи четырех арифметических действий и действия нахождение корня. Эти способы здесь не приводятся, так как они весьма трудоемки.

Что касается многочленов степени пять и выше, то известно, что для любого $n\geqslant 5$ существует многочлен степени n, корни которого невозможно выразить через его коэффициенты при помощи четырех арифметических действий и действия нахождения корня.

Поэтому ниже приводятся методы нахождения корней многочленов степени n, где $n\geqslant 3$, только для некоторых частных случаев.

Теорема 14. Если все коэффициенты многочлена степени n, где $n\geqslant 1$ — целые числа и корень многочлена α — также целое число, то число α — делитель свободного члена многочлена.

Доказательство. Пусть дан многочлен

$$P_n(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n \qquad (a_0\neq 0)$$

степени n, где $n\geqslant 1$, и пусть α — корень этого многочлена. Разделим с остатком многочлен $P_n(x)$ на двучлен $(x-\alpha)$, тогда частное есть многочлен $q(x)=b_0x^{n-1}+b_1x^{n-2}+\ldots+b_{n-1}$, а остаток — число r. Как показано выше, если все коэффициенты многочлена $P_n(x)$ и α — целые числа, то числа b_0, b_1, ..., b_{n-1} и r — также целые числа. По схеме Горнера $r=a_n+\alpha b_{n-1}$, а по теореме 7 если α — корень многочлена, то $r=0$. Поэтому имеем равенство $a_n+\alpha b_{n-1}=0$, откуда $a_n=\alpha(-b_{n-1})$. Так как a_n, α, $-b_{n-1}$ — целые числа, то отсюда и вытекает, что α — делитель числа a_n, и теорема доказана.

Следствие. Целыми корнями многочлена с целыми коэффициентами могут быть лишь делители свободного члена многочлена.

Это следствие позволяет находить все целые корни многочлена с целыми коэффициентами, применяя схему Горнера.

Пример. Выяснить, имеет ли целые корни многочлен

$$P_4(x)=x^4+2x^3-2x^2-6x+5. \tag{9}$$

Делители свободного члена: 1, —1, 5, —5. Найдем значения многочлена в этих точках:

$$P_4(1)=1+2-2-6+5=0,$$
$$P_4(-1)=1-2-2+6+5=8\neq 0,$$

$$P_4(5) = 625 + 250 - 50 - 30 + 5 = 700 \neq 0,$$
$$P_4(-5) = 625 - 250 - 50 + 30 + 5 = 290 \neq 0.$$

Итак, многочлен (9) имеет целый корень $x_1 = 1$, а числа 5, -5 и -1 не являются его корнями. Применив схему Горнера, разложим многочлен (9) на множители. Схема Горнера имеет вид

$$\begin{array}{cccccc}
 & 1 & 2 & -2 & -6 & 5 \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
1 & 1 = & 2 + 1 \cdot 1 = & -2 + 1 \cdot 3 = & -6 + 1 \cdot 1 = & 5 + 1 \cdot (-5) = \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & = 1 & = 3 & = 1 & = -5 & = 0
\end{array}$$

Следовательно, $x^4 + 2x^3 - 2x^2 - 6x + 5 = (x - 1)(x^3 + 3x^2 + x - 5)$.

Теперь будем искать корни многочлена $P_3(x) = x^3 + 3x^2 + x - 5$. Делители его свободного члена: $1, -1, +5, -5$. Нет необходимости искать значение многочлена $P_3(x)$ в точках -1, $+5$, -5, так как эти числа заведомо не являются корнями многочлена $P_4(x)$, а значит, и многочлена $P_3(x)$ в силу того, что многочлен $P_4(x)$ в них не обращается в нуль. Поэтому проверим только число 1:

$$P_3(1) = 1 + 3 + 1 - 5 = 0.$$

Применив опять схему Горнера:

$$\begin{array}{ccccc}
 & 1 & 3 & 1 & -5 \\
 & \downarrow & \downarrow & \downarrow & \downarrow \\
1 & 1 = & 3 + 1 \cdot 1 = & 1 + 1 \cdot 4 = & -5 + 1 \cdot 5 = \\
 & \downarrow & \downarrow & \downarrow & \downarrow \\
 & = 1 & = 4 & = 5 & = 0
\end{array}$$

получим $P_3(x) = (x - 1)(x^2 + 4x + 5)$, а потому $P_4(x) = (x - 1)^2 (x^2 + 4x + 5)$.

Так как квадратный трехчлен $x^2 + 4x + 5$ целых корней не имеет, то, следовательно, многочлен $P_4(x)$ имеет два целых корня $x_1 = 1$, $x_2 = 1$, или, точнее, один корень $x_1 = 1$ кратности два.

Замечание. Если найден один корень $x_1 = \alpha$ многочлена $P(x)$, то этот многочлен можно записать $P(x) = (x - \alpha) q(x)$, где коэффициенты многочлена $q(x)$ легко вычисляются по схеме Горнера. Чтобы найти другие корни многочлена $P(x)$, следует найти корни многочлена $q(x)$. Важно отметить, что многочлен $q(x)$ может иметь корнем то же число α, который находится также по схеме Горнера.

Если не искать корни многочлена $q(x)$, а отыскивать корни многочлена $P(x)$, то корень, который уже найден, во второй раз этим же способом не будет обнаружен. Поэтому после нахождения одного корня, надо искать корни частного, т. е. корни многочлена $q(x)$.

Пример. Найти все корни многочлена $P(x) = x^4 - x^3 - 4x^2 - 5x - 3$. Свободный член многочлена $P(x)$ имеет делители

$+1$, -1, -3, $+3$. Найдем значение многочлена в этих точках:

$$P(1)=1-1-4-5-3=-12\neq 0,$$
$$P(-1)=1+1-4+5-3=0,$$
$$P(-3)=81+27-36+15-3=84\neq 0,$$
$$P(3)=81-27-36-15-3=0.$$

Таким образом, данный многочлен имеет два целых корня $x_1=-1$ и $x_2=3$. Применим схему Горнера для разложения многочлена

$$\begin{array}{c|c|c|c|c|c}
 & 1 & -1 & -4 & -5 & -3 \\
1 & 1= & -1+(-1)\,1= & -4+(-1)\,(-2)= & -5+(-1)\,(-2)= & -3+(-1)\,(-3)= \\
 & =1 & =-2 & =-2 & =-3 & =0
\end{array}$$

Следовательно, $P(x)=(x+1)(x^3-2x^2-2x-3)$.

Применим второй раз схему Горнера для разложения полученного многочлена x^3-2x^2-2x-3, зная, что $x_2=3$ является его корнем:

$$\begin{array}{c|c|c|c|c}
 & 1 & -2 & -2 & -3 \\
3 & 1= & -2+3\cdot 1= & -2+3\cdot 1= & -3+3\cdot 1= \\
 & =1 & =1 & =1 & =0
\end{array}$$

Получим, что $P(x)=(x+1)(x-3)(x^2+x+1)$.

Многочлен x^2+x+1 имеет комплексные корни

$$x_3=(-1+\sqrt{3}\,i)/2 \quad \text{и} \quad x_4=(-1-\sqrt{3}\,i)/2.$$

Таким образом, многочлен $P(x)=x^4-x^3-4x^2-5x+3$ имеет четыре корня: $x_1=-1$, $x_2=3$, $x_3=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$, $x_4=-\frac{1}{2}-\frac{\sqrt{3}}{2}i$.

Теорема 15. Если многочлен

$P_n(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n$

с целыми коэффициентами и со старшим коэффициентом, равным единице, имеет рациональный корень, то этот корень— целое число.

Доказательство. Доказательство этой теоремы проведем методом от противного. Предположим, что многочлен $P_n(x)$ имеет корень $\alpha=p/q$, где p и q взаимно простые целые числа. Так как число p/q корень многочлена $P_n(x)$, то справедливо равенство

$$\frac{p^n}{q^n}+a_1\frac{p^{n-1}}{q^{n-1}}+a_2\frac{p^{n-2}}{q_{n-2}}+\ldots+a_{n-1}\frac{p}{q}+a_n=0,$$

которое можно записать в равносильной форме

$$\frac{p^n}{q^n} = -\left(a_1 \frac{p^{n-1}}{q^{n-1}} + a_2 \frac{p^{n-2}}{q^{n-2}} + \ldots + a_{n-1} \frac{p}{q} + a_n\right).$$

Умножая это равенство на q^{n-1}, получим равносильное равенство $p^n/q = -a_1p^{n-1} - a_2p^{n-2}q - \ldots - a_{n-1}pq^{n-2} - a_nq^{n-1}$. Так как числа p и q просты, то число p^n/q — дробь, а справа в последнем равенстве стоит целое число, чего быть не может. Значит, предположение неверно, а верна теорема.

Следствие. Если у многочлена все коэффициенты целые числа, а старший коэффициент равен единице, то все рациональные корни этого многочлена — целые числа.

Рассмотрим многочлен $P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$ с целыми коэффициентами и многочлен $Q_n(x) = a_0^{n-1}P_n(x) = (a_0x)^n + a_1(a_0x)^{n-1} + \ldots + a_na_0^{n-1}$. Ясно, что многочлены $P_n(x)$ и $Q_n(x)$ имеют одинаковые корни. Обозначим $y = a_0x$, тогда $Q_n(x) = T_n(y) = y^n + a_1y^{n-1} + a_2a_0y^{n-2} + a_3a_0^2y^{n-3} + \ldots + a_na_0^{n-1}$. Многочлен $T_n(y)$ имеет по теореме 15 только целые корни, которые можно найти. Пусть это будут числа y_k, тогда числа $x_k = y_k/a_0$ и только они будут рациональными корнями многочлена $P_n(x)$. Итак, у *любого многочлена с целыми коэффициентами можно найти все его рациональные корни.*

Если коэффициенты многочлена рациональные числа, то после приведения их к общему знаменателю можно искать лишь корни числителя, который есть многочлен с целыми коэффициентами. Значит, *в поле рациональных чисел можно найти все рациональные корни многочлена, заданного над полем рациональных чисел.*

Пример. Найти корни многочлена $P_3(x) = x^3 + \frac{1}{2}x^2 + \frac{1}{2}x - \frac{1}{2}$.

Рассмотрим многочлен $Q_3(x) = 8P_3(x) = (2x)^3 + (2x)^2 + 2(2x) - 4$ или $T_3(t) = t^3 + t^2 + 2t - 4$, где $t = 2x$. Делители свободного члена многочлена $T_3(t)$: $+1, -1, +2, -2, +4, -4$. Найдем значения многочлена $T_3(t)$ в этих точках:

$$T_3(1) = 1 + 1 + 2 - 4 = 0,$$
$$T_3(-1) = -1 + 1 - 2 - 4 = -6 \neq 0,$$
$$T_3(2) = 8 + 4 + 4 - 4 = 12 \neq 0,$$
$$T_3(-2) = -8 + 4 - 4 - 4 = -12 \neq 0,$$
$$T_3(4) = 64 + 16 + 8 - 4 = 84 \neq 0,$$
$$T_3(-4) = -64 + 16 - 8 - 4 = -60 \neq 0.$$

Применим схему Горнера:

	1	1	2	-4
1	$1=$	$1+1\cdot 1=$	$2+1\cdot 2=$	$-4+1\cdot 4=$
	$=1$	$=2$	$=4$	$=0$

Итак, $T_3(t)=(t-1)(t^2+2t+4)$.

Квадратный трехчлен t^2+2t+4 имеет комплексные корни $t_2=-1+\sqrt{3}\,i$, $t_3=-1-\sqrt{3}\,i$. Следовательно, исходный многочлен $P_3(x)=x^3+\frac{1}{2}x^2+\frac{1}{2}x-\frac{1}{4}$ имеет один рациональный корень $x_1=\frac{1}{2}$ и два комплексных корня $x_2=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$, $x_3=\frac{-1}{2}-\frac{\sqrt{3}}{2}i$.

§ 3
КОЛЬЦА И ПОЛЯ

Пусть дано некоторое непустое множество элементов. Будем говорить, что в этом множестве определена *алгебраическая операция,* если указан закон, по которому любой паре элементов a и b из этого множества однозначным образом ставится в соответствие некоторый элемент c, также принадлежащий этому множеству.

Если эта операция называется *сложением*, то элемент c будет называться *суммой* элементов a и b и обозначаться $c=a+b$.

Если эта операция называется *умножением*, то элемент c будет называться *произведением* элементов a и b и обозначаться $c=ab$.

Непустое множество элементов называется *кольцом*, если в нем определены две операции — сложение и умножение, обладающие свойствами:

1. Сложение коммутативно: $a+b=b+a$;
2. Сложение ассоциативно: $a+(b+c)=(a+b)+c$;
3. Сложение и умножение связаны левым и правым законами дистрибутивности: $c(a+b)=ca+cb$, $(a+b)c=ac+bc$;
4. Существует такой элемент этого множества, что прибавление его к любому элементу этого множества не меняет последнего. Этот элемент называется нулем кольца и обозначается символом 0; другими словами, существует элемент 0 такой, что $a+0=0+a=a$;
5. Для любого элемента a этого множества существует так называемый противоположный элемент, принадлежащий этому же множеству, и такой, что сумма a и этого элемента равна

нулю кольца; обозначая этот элемент через $(-a)$, запишем это свойство в виде $a+(-a)=0$.

Заметим, что из приведенного определения вытекают следующие свойства кольца.

1. Любое кольцо имеет единственный нуль.

2. В любом кольце для каждого элемента a существует единственный противоположный элемент.

3. Для любого элемента a кольца справедливы равенства $0\cdot a=a\cdot 0=0$.

Если в данном кольце операция умножения будет обладать свойством коммутативности, т. е. если для любых двух элементов кольца a и b справедливо равенство $ab=ba$, то *кольцо* называется *коммутативным*.

Если в данном кольце операция умножения будет обладать свойством ассоциативности, т. е. если для любых трех элементов кольца a, b и c справедливо равенство $(ab)c=a(bc)$, то *кольцо* называется *ассоциативным*.

Числовые кольца, рассмотренные в § 1, представляют собой простейшие примеры коммутативных и ассоциативных колец.

Приведем другие примеры колец.

1. Множество многочленов, целых относительно одной буквы x, с коэффициентами из данного числового поля, образует коммутативное и ассоциативное кольцо, которое называется кольцом многочленов над данным полем.

Действительно, как указано в § 2, сумма и произведение многочленов есть многочлен, т. е. в множестве многочленов определены две операции: сложение и умножение. Причем в § 2 указано, что обе эти операции коммутативны, ассоциативны и связаны законом дистрибутивности. Роль нуля в этом кольце играет многочлен, все коэффициенты которого равны нулю, т. е. число нуль. Противоположный элемент для каждого многочлена есть многочлен, который получается умножением этого многочлена на число -1.

2. Множество всех функций, непрерывных на данном отрезке $[a; b]$, также образует коммутативное и ассоциативное кольцо.

Действительно, в гл. VIII дано определение непрерывной на данном отрезке функции и указано, что сумма и произведение двух непрерывных функций есть непрерывная функция. Легко проверить, что операции сложения и умножения непрерывных функций коммутативны, ассоциативны и связаны законом дистрибутивности. Роль нуля играет функция, тождественно равная нулю, т. е. число нуль; роль противоположного элемента для функции $f(x)$ играет функция $[-f(x)]$.

3. Множество всех квадратных матриц данного порядка n образует кольцо. Действительно, в гл. IX показано, что сумма и произведение квадратных матриц данного порядка есть квадратная матрица того же порядка. При этом роль нуля играет квадратная матрица порядка n, все элементы которой есть

нули. Роль противоположного элемента для данной матрицы играет матрица, все элементы которой получены из элементов данной матрицы умножением на число -1.

В гл. IX показано, что операция умножения матриц является ассоциативной, но не является коммутативной. Поэтому кольцо квадратных матриц данного порядка n будет ассоциативным, но не коммутативным кольцом.

Множество элементов называется *полем*, если это множество состоит не менее чем из двух элементов, и является коммутативным и ассоциативным кольцом и если в нем существует элемент (называемый единицей поля), такой, что произведение любого элемента a поля на эту единицу равно этому элементу a, и если в нем для любого элемента a, отличного от нуля, существует элемент (называемый обратным элементом), такой, что произведение любого элемента a на его обратный элемент равно единице поля.

Примерами полей могут служить все числовые поля, рассмотренные в § 1.

Приведем другие примеры полей и колец.

1. Множество матриц вида

$$\begin{pmatrix} a & b \\ 2b & a \end{pmatrix}, \tag{1}$$

где числа a и b — любые числа из данного числового кольца, образует коммутативное и ассоциативное кольцо относительно обычных операций сложения и умножения матриц. Действительно, рассмотрим сумму и произведение матриц вида (1). Ясно, что

$$\begin{pmatrix} a & b \\ 2b & a \end{pmatrix} + \begin{pmatrix} c & d \\ 2d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ 2(b+d) & a+c \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ 2b & a \end{pmatrix} \begin{pmatrix} c & d \\ 2d & c \end{pmatrix} = \begin{pmatrix} ac+2bd & ad+bc \\ 2(ad+bc) & ac+2bd \end{pmatrix}.$$

Поскольку числа, являющиеся элементами матриц, стоящих в правых частях этих равенств, есть числа из данного кольца, то это означает, что сумма и произведение матриц вида (1) есть матрицы вида (1). Другими словами, показано, что на множестве матриц вида (1) определены операции сложения и умножения.

Столь же легко показывается, что справедливы свойства коммутативности, ассоциативности и дистрибутивности сложения и умножения матриц вида (1).

Матрица $\begin{pmatrix} 0 & 0 \\ 2\cdot 0 & 0 \end{pmatrix}$ является нулем этого кольца, а матрица $\begin{pmatrix} -a & -b \\ -2b & -a \end{pmatrix}$ является противоположным элементом для матрицы $\begin{pmatrix} a & b \\ 2b & a \end{pmatrix}$.

Таким образом, показано, что множество матриц вида (1), где числа a и b — любые числа из данного числового кольца, образует коммутативное и ассоциативное кольцо.

2. Множество матриц вида

$$\begin{pmatrix} a & b \\ 2b & a \end{pmatrix}, \tag{2}$$

где a и b — рациональные числа, образует поле.

Так как множество рациональных чисел есть числовое кольцо, то, как показано выше, множество матриц вида (2) с рациональными a и b образует коммутативное и ассоциативное кольцо. Роль единицы будет играть матрица $\begin{pmatrix} 1 & 0 \\ 2\cdot 0 & 1 \end{pmatrix}$.

Пусть теперь матрица $\begin{pmatrix} a & b \\ 2b & a \end{pmatrix}$ не является нулем кольца, т. е. пусть либо $a \neq 0$, либо $b \neq 0$, либо одновременно $a \neq 0$, $b \neq 0$. Покажем, что для любой такой матрицы найдется матрица вида (2), т. е. матрица $\begin{pmatrix} x & y \\ 2y & x \end{pmatrix}$, такая, что

$$\begin{pmatrix} a & b \\ 2b & a \end{pmatrix} \times \begin{pmatrix} x & y \\ 2y & x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

т. е. покажем, что у любого элемента этого кольца, отличного от нуля, есть обратный элемент. Применяя правило перемножения матриц, получаем, что справедливость равенства (3) равносильна справедливости равенства

$$\begin{pmatrix} ax+2by & ay+bx \\ 2(bx+ay) & (ax+2by) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4}$$

Равенство (4) справедливо тогда и только тогда, когда справедлива система уравнений

$$\begin{cases} ax+2by=1, \\ bx+ay=0. \end{cases} \tag{5}$$

Вычислим определитель системы (5)

$$\begin{vmatrix} a & 2b \\ b & a \end{vmatrix} = a^2-2b^2. \tag{6}$$

Поскольку a и b — рациональные числа, то выражение $a^2 - 2b^2 \neq 0$. Значит, для данных рациональных чисел a и b существует единственная пара чисел x и y, удовлетворяющая системе (5).

Таким образом, показано, что существует и притом только одна матрица вида (2), удовлетворяющая условию (3). Другими словами, показано, что любой отличный от нуля элемент имеет единственный обратный элемент.

Значит, действительно, множество матриц вида (2) с рациональными a и b образует поле.

Замечание. Множество матриц вида (2) с действительными a и b поля не образует, так как для действительных чисел определитель (6) может обращаться в нуль, например, если $a = b\sqrt{2}$, и тогда у элемента, отличного от нуля, нет обратного элемента.

3. Рассмотрим множество элементов, где каждый элемент — упорядоченная пара целых чисел (n, m). Введем следующие определения.

Два элемента (n, m) и (p, q) *равны* тогда и только тогда, когда $n = p$ и $m = q$.

Суммой двух элементов (n, m) и (k, l) называется элемент $(n+k, m+l)$, т. е. $(n, m)+(k, l)=(n+k, m+l)$.

Произведением двух элементов (n, m) и (k, l) называется элемент (nk, ml), т. е. $(n, m)(k, l)=(nk, ml)$. Покажем, что так введенное множество есть коммутативное и ассоциативное кольцо. Будем обозначать элементы этого множества одной буквой A, т. е. $A=(n, m)$. Проверим, что операции сложения и умножения обладают свойствами:

1) $A+B=B+A$;
2) $(A+B)+C=A+(B+C)$;
3) $AB=BA$;
4) $(AB)C=A(BC)$;
5) $(A+B)C=AC+BC$.

Действительно, пусть $A=(n, m)$, $B=(k, l,)$ $C=(p, q)$. Тогда, используя свойства коммутативности, ассоциативности и дистрибутивности сложения и умножения целых чисел, получаем, что

$$A+B=(n, m)+(k, l)=(n+k, m+l),$$
$$B+A=(k, l)+(n, m)=(k+n, l+m)=(n+k, m+l),$$

откуда и вытекает, что $A+B=B+A$, т. е. свойство 1 доказано. Докажем свойство 5. По определению

$$(A+B)C=(n+k, m+l)(p, q)=(np+kp, mq+lq),$$
$$AC=(n, m)(p, q)=(np, mq),$$
$$BC=(k, l)(p, q)=(kp, lq),$$
$$AC+BC=(np, mq)+(kp, lq)=(np+kp, mq+lq),$$

т. е. $(A+B)C=AC+BC$. Остальные свойства доказываются аналогично.

Роль нуля в этом кольце играет элемент (0, 0). Для элемента (n, m) роль противоположного элемента играет элемент

$(-n, -m)$. Покажем, что это кольцо не является полем. Легко видеть, что роль единицы играет элемент $(1, 1)$. Рассмотрим, например, элемент $(3, 2)$ и покажем, что для него нет обратного элемента.

Предположим противное: что такой элемент есть. Пусть это будет элемент (x, y). Тогда должно быть справедливо равенство $(3, 2)(x, y)=(1, 1)$, откуда следует, что должны быть справедливы равенства $3x=1$ и $2y=1$. Но в множестве целых чисел эти равенства не выполняются. Значит, наше предположение было неверным, а это и означает, что рассматриваемое кольцо не является полем.

4. Рассмотрим множество, состоящее из трех элементов: A_0, A_1, A_2. Под элементом A_0 понимается класс всех целых чисел, делящихся нацело на число 3; под элементом A_1 понимается класс всех целых чисел, которые при делении на 3 дают остаток, равный 1: под элементом A_2 понимается класс всех целых чисел, которые при делении на 3 дают остаток, равный 2.

Определим в множестве (A_0, A_1, A_2) операции сложения и умножения следующими правилами:

$$A_k+A_l=\begin{cases} A_{k+l}, & \text{если } k+l<3, \\ A_{k+l-3}, & \text{если } k+l\geqslant 3, \end{cases}$$

$$A_kA_l=\begin{cases} A_{kl}, & \text{если } kl<3, \\ A_{kl-3}, & \text{если } kl\geqslant 3. \end{cases}$$

Легко видеть, что операция сложения элементов A_k и A_l означает сложение любых чисел из классов A_k и A_l и отнесение их суммы в соответствующий класс $A_m (k, l, m=0, 1, 2)$; операция умножения элементов A_k и A_l означает умножение любых чисел из классов A_k и A_l и отнесение их произведения в соответствующий класс $A_m (k, l, m=0, 1, 2)$.

Легко также проверить, что операции сложения и умножения элементов A_k и A_l обладают свойствами коммутативности, ассоциативности и дистрибутивности.

Роль нуля в этом множестве играет класс A_0. Противоположным элементом для элемента A_k является элемент A_{3-k}.

Роль единицы в этом множестве играет класс A_1. Для элемента A_1 обратным элементом будет он сам, т. е. A_1, для элемента A_2 обратным элементом будет он сам, т. е. A_2. Значит, у любого элемента, отличного от нуля, этого множества есть обратный элемент.

Все эти свойства легко проверяются на основании правил сложения и умножения этих классов. Значит, это множество есть поле.

Замечание. Если рассмотреть множество, состоящее из k элементов $(k\geqslant 2)$: $A_0, A_1, A_2, \ldots, A_{k-1}$, где под элементом A_l понимается класс всех целых чисел, которые при делении на k дают остаток, равный l (где $l=0, 1, 2, \ldots, k-1$), то в каждом таком множестве можно определить операции сложения и умножения аналогично рассмотренному примеру. Для любого натурального k $(k \geqslant 2)$ множество элементов $A_0, A_1, A_2, \ldots, A_{k-1}$ с операциями сложения и умножения элементов, определенными указанным образом, будет кольцом, а если число k будет простым числом, то такое множество будет полем.

5. Рассмотрим множество упорядоченных пар натуральных чисел (n, m). Разобьем это множество на классы, относя в один класс A те и только те пары (n, m) и (k, l), которые обладают свойством $n+l=m+k$. По определению пары, принадлежащие одному и тому же классу, называются равными и считается, что каждая такая пара определяет класс A. В множестве, элементами которого будут классы равных пар, определим операции сложения и умножения по следующим правилам.

Сложить два элемента A и B — значит сложить любую пару (n, m) из класса A и любую пару (p, q) из класса B по правилу $(n, m)+(p, q)=(n+p, m+q)$. Пара $(n+p, m+q)$ попадет в некоторый класс C, который и называется *суммой* элементов A и B и обозначается $A+B$, т. е. $C=A+B$.

Умножить два элемента A и B — это значит умножить любую пару (n, m) из класса A и любую пару (p, q) из класса B по правилу $(n, m)(p, q)=(np+mq, nq+mp)$. Пара $(np+mq, nq+mp)$ попадет в некоторый класс D, который и называется *произведением* элементов A и B и обозначается AB, т. е. $D=AB$.

Введенное множество классов пар есть коммутативное и ассоциативное кольцо. Действительно, легко проверить, что операции сложения и умножения обладают свойствами коммутативности, ассоциативности и дистрибутивности.

Докажем, например, свойство дистрибутивности: $(A+B)M = AM+BM$. Пусть пара $(n, m) \in A$, пара $(k, l) \in B$, пара $(p, q) \in M$. Если $A+B=C$, то класс C определяется парой $(n+k, m+l)$. Если $CM=D$, то класс D определяется парой $(np+kp+mq+lq, nq+kq+mp+lp)$. Если $AM=N_1$, то класс N_1 определяется парой $(np+mq, nq+mp)$, если $BM=N_2$, то класс N_2 определяется парой $(kp+lq, kq+lp)$. Если $N_1+N_2=N$, то класс N определяется парой
$(np+mq+kp+lq, nq+mp+kq+lp)$.

Теперь очевидно, что классы D и N определяются одной парой, т. е. это один класс, таким образом показано, что $(A+B)M=AM+BM$. Аналогично доказываются и другие свойства сложения и умножения.

Роль нуля в этом множестве играет класс пар вида (k, k).

Покажем, что у каждого класса A, определяемого парой (n, m), имеется противоположный класс. Выберем натуральное число p, такое, что $p>n$ и $p>m$. Тогда числа $x=p-n$ и $y=p-m$ есть натуральные числа. Покажем, что класс B, определяемый парой (x, y), и есть класс, противоположный классу A. Действительно, $A+B=C$, где класс C определяется парой (p, p). Но пара (p, p) как раз и определяет класс, являющийся нулем этого множества. Итак, у каждого элемента этого множества есть противоположный элемент. Значит, рассматриваемое множество есть коммутативное и ассоциативное кольцо.

Замечание. При построении этого кольца были использованы только свойства натуральных чисел. Это кольцо можно использовать для конструктивного построения кольца целых чисел на базе натуральных чисел. Делается это так: класс A, определяемый парой (n, m), отождествляется:

а) с натуральным числом k, если $n>m$ и $n-m=k$;

б) с числом нуль, если $n=m$;

в) с отрицательным числом $-k$, если $n<m$ и $m-n=k$.

Полученное множество и будет кольцом целым чисел.

6. Рассмотрим множество упорядоченных пар целых чисел $[a, b]$, таких, что $b \neq 0$. Разобьем это множество на классы, относя в один класс α те и только те пары $[a, b]$ и $[c, d]$, которые обладают свойством $ad=bc$. По определению пары, принадлежащие одному и тому же классу, называются равными и считается, что каждая такая пара определяет класс α.

В множестве, элементами которого будут классы равных пар, определим операции сложения и умножения по следующим правилам.

Сложить два элемента α и β — значит сложить любую пару $[a, b]$ из класса α и любую пару $[c, d]$ из класса β по правилу $[a, b]+[c, d]=[ad+bc, bd]$. Пара $[ad+bc, bd]$ попадет в некоторый класс γ, который и называется *суммой* элементов α и β и обозначается $\alpha+\beta$, т. е. $\gamma=\alpha+\beta$.

Умножить два элемента α и β — значит умножить любую пару $[a, b]$ из класса α и любую пару $[c, d]$ из класса β по правилу $[a, b][c, d]=[ac, bd]$. Пара $[ac, bd]$ попадет в некоторый класс δ, который и называется *произведением* элементов α и β и обозначается $\alpha\beta$, т. е. $\delta=\alpha\beta$.

Введенное множество классов пар есть поле. Действительно, легко проверить, что операции сложения и умножения обладают свойствами коммутативности, ассоциативности и дистрибутивности.

Докажем, например, свойство дистрибутивности: $(\alpha+\beta)\mu = \alpha\mu+\beta\mu$. Пусть пара $[a, b] \in \alpha$, пара $[c, d] \in \beta$, пара $[l, f] \in \mu$. Если $\alpha+\beta=\gamma$, то класс γ определяется парой $[ad+bc, bd]$, если $\gamma\mu=\nu$, то класс ν определяется парой $[adl+bcl, bdf]$. Если $\alpha\mu=\delta_1$, то класс δ_1 определяется парой $[al, bf]$, если $\beta\mu=\delta_2$, то класс δ_2 определяется парой $[cl, df]$. Если $\delta_1+\delta_2=\delta$, то класс δ определяется парой $[aldf + bfcl, bfdf]$. Так как пары $[aldf+bfcl, bfdf]$ и $[adl+bcl, bdf]$ удовлетворяют условию равенства пар, то они относятся к одному и тому же классу, т. е. класс ν и класс δ есть один и тот же класс. Таким образом показано, что $(\alpha+\beta)\mu=\alpha\mu+\beta\mu$. Аналогично доказываются и другие свойства сложения и умножения.

Роль нуля в этом множестве играет класс γ_0, определяемый парой $[0, 1]$.

Покажем, что для любого класса α имеется противоположный класс β, т. е. такой, что $\alpha+\beta=\gamma_0$. Пусть пара $[a, b]\in\alpha$, тогда пара $[-a, b]$ определяет класс β, который и является противоположным классу α. Действительно, если $\alpha+\beta=\sigma$, то класс σ определяется парой $[0, b^2]$, где $b^2\neq 0$. Так как пары $[0, b^2]$ и $[0, 1]$ удовлетворяют условию равенства пар, то они относятся к одному и тому же классу, т. е. класс σ и класс γ_0 есть один и тот же класс, что и означает, что $\alpha+\beta=\gamma_0$.

Итак, у каждого элемента рассматриваемого множества есть противоположный элемент. Значит, рассматриваемое множество есть коммутативное и ассоциативное кольцо. Роль единицы в этом кольце играет класс E, определяемый парой $[1, 1]$.

Покажем, что у каждого класса α, отличного от нуля, есть обратный класс β, т. е. такой, что $\alpha\beta=E$. Пусть пара $[a, b]\in\alpha$ и α не нуль кольца, т. е. пусть $a\neq 0$ и $b\neq 0$. Рассмотрим класс β, определяемый парой $[b, a]$. Если $\alpha\beta=\mu$, то класс μ определяется парой $[ab, ab]$. Так как пары $[ab, ab]$ и $[1, 1]$ удовлетворяют условию равенства пар, то они относятся к одному и тому же классу. Значит, класс μ совпадает с классом E, т. е. $\alpha\beta=E$.

Итак, у любого отличного от нуля элемента рассматриваемого множества есть обратный элемент. Значит, это множество есть поле.

Замечание. При построении этого поля были использованы только свойства целых чисел. Это поле можно использовать для конструктивного построения поля рациональных чисел на базе целых чисел. Делается это так: класс τ, определяемый парой $[a, b]$, отождествляется с рациональным числом a/b; при этом если $a=bn$, где n — целое число, то класс, определяемый парой $[bn, b]$, отождествляется с целым числом n. Полученное множество и будет полем рациональных чисел.

7. Рассмотрим множество упорядоченных пар действительных чисел $\{a, b\}$. Будем считать, что два элемента этого множества $\{a, b\}$ и $\{c, d\}$ равны тогда и только тогда, когда $a=c$ и $b=d$. Введем в этом множестве пар действительных чисел операции сложения и умножения:

$$\{a, b\}+\{c, d\}=\{a+c, b+d\},$$
$$\{a, b\}\{c, d\}=\{ac-bd, ad+bc\}.$$

Можно показать, что это множество является полем. Если отождествить пару $\{a, b\}$ с комплексным числом $a+bi$, то получится поле комплексных чисел, рассмотренное в главе X.

§ 4
ГРУППЫ

В § 3 рассмотрены множества, в которых определены две операции над элементами этих множеств: сложение и умножение элементов.

В этом параграфе рассматриваются множества, в которых определена лишь одна операция. Если эта операция обладает некоторыми свойствами, то это множество принято называть группой. А именно, имеет место следующее определение.

Непустое множество элементов называется группой, если в нем определена операция, обозначаемая $\ast$ и со свойствами:

а) ассоциативность: $(a \ast b) \ast c = a \ast (b \ast c)$;

б) в этом множестве существует так называемый нейтральный элемент, обозначаемый символом e, такой, что $a \ast e = e \ast a = a$ для любого элемента a этого множества;

в) для любого элемента a из этого множества существует элемент b из этого же множества, такой, что $a \ast b = b \ast a = e$.

Поскольку наиболее распространены две операции — сложение и умножение, то наряду с этим общим определением приведем еще два частных случая — определение группы по сложению и группы по умножению.

Множество элементов называется *группой по сложению*, если в нем определена операция, называемая сложением и обладающая свойствами:

а) ассоциативность сложения $(a+b)+c=a+(b+c)$;

б) в этом множестве существует элемент, называемый нулем и обозначаемый символом 0, такой, что для любого элемента a этого множества $0+a=a+0=a$;

в) для любого элемента a из этого множества существует элемент b из этого же множества, такой, что $b+a=a+b=0$.

Элемент b называется противоположным для элемента a и обозначается через $(-a)$.

Множество элементов называется *группой по умножению*, если в нем определена операция, называемая умножением и обладающая свойствами:

а) ассоциативность умножения: $(ab)c = a(bc)$;

б) в этом множестве существует элемент, называемый единицей группы и обозначаемый символом e, такой, что для любого элемента a этого множества $ae = ea = a$;

в) для любого элемента a из этого множества существует элемент c из этого же множества, такой, что $ac = ca = e$.

Элемент c называется обратным для элемента a и обозначается через a^{-1}.

Если операция, определенная в группе, будет коммутативной, то группа называется *коммутативной группой*.

Приведем примеры групп.

1. Отметим, что каждое кольцо является коммутативной

группой по сложению, каждое поле является коммутативной группой по сложению.

Любое кольцо или любое поле не являются группой по умножению, однако если рассмотреть любое поле без нулевого элемента, то это новое множество уже является коммутативной группой по умножению.

2. Рассмотрим множество, состоящее из числа нуль и всех многочленов, целых относительно одной буквы x степени не выше, чем n, с коэффициентами из данного числового поля. На этом множестве определена лишь одна операция — сложение многочленов (операция умножения многочленов выводит произведение за пределы этого множества). Легко видеть, что в этом множестве операция сложения многочленов является коммутативной и ассоциативной, число нуль — нулевой элемент множества и у любого многочлена есть противоположный элемент.

Значит, рассматриваемое множество образует коммутативную группу по сложению.

3. Рассмотрим множество матриц, имеющих n строк и m столбцов, причем $n \neq m$. Легко видеть, что в этом множестве определена лишь одна операция — сложение матриц (операция умножения для таких матриц не определена) (см. гл. IX). Легко видеть, что в этом множестве: а) операция сложения является коммутативной и ассоциативной; б) существует нулевой элемент — матрица, у которой все элементы равны нулю; в) у любой такой матрицы есть противоположный элемент — матрица, все элементы которой получены из элементов данной матрицы умножением на число -1.

Значит, множество таких матриц образует коммутативную группу по сложению.

4. Рассмотрим множество, состоящее из всех чисел вида 2^k, где k — любое целое число. Очевидно, что операция умножения этих чисел не выводит за пределы этого множества, другими словами, на этом множестве определена операция умножения элементов (обычная операция умножения чисел). Ясно также, что эта операция является коммутативной и ассоциативной. Роль единичного элемента играет число 1, принадлежащее этому множеству, так как $1=2^0$; у каждого элемента 2^k есть обратный элемент 2^{-k}. Значит, это множество образует коммутативную группу по умножению.

5. Рассмотрим множество всех положительных рациональных чисел. На этом множестве определена операция умножения. Причем эта операция является коммутативной и ассоциативной. Число единица играет роль нейтрального элемента этого множества, и у каждого элемента p/q этого множества есть обратный элемент q/p.

Значит, множество всех положительных рациональных чисел образует коммутативную группу по умножению.

6. Рассмотрим множество всех положительных рациональ-

ных чисел и выясним, образует ли это множество группу относительно операции $*$, где $*$ — деление рациональных чисел. Хотя операция $*$ не выводит за пределы этого множества и в этом множестве есть нейтральный элемент — число 1, ибо у каждого элемента p/q есть элемент p/q, такой, что $\frac{p}{q} * \frac{p}{q} = 1$, это множество относительно введенной операции $*$ не является группой, так как легко проверить, что эта операция $*$ не будет ассоциативной.

7. Рассмотрим множество всех решений уравнения $x^n = 1$, или, другими словами, рассмотрим множество всех комплексных корней n-й степени из числа единица. Как показано в гл. X, произведение любых двух таких корней есть также корень степени n из числа единица, причем эта операция является коммутативной и ассоциативной. Роль единичного элемента играет число 1, роль обратного элемента для элемента α_k — элемент α_{-k}.

Значит, это множество образует коммутативную группу относительно операции умножения.

8. Рассмотрим множество вращений правильного треугольника вокруг его центра, совмещающих треугольник с самим собой. Если за умножение двух вращений принять их последовательное выполнение, то это множество образует коммутативную группу относительно этой операции. Можно проверить, что выполнены все свойства, определяющие коммутативную группу.

9. Рассмотрим множество вращений шара вокруг его центра. В качестве произведения вращений естественно взять последовательное выполнение двух вращений. Можно показать, что это множество относительно введенной операции образует некоммутативную группу.

М. К. ПОТАПОВ, В. В. АЛЕКСАНДРОВ, П. И. ПАСИЧЕНКО

ЛЕКЦИИ ПО АЛГЕБРЕ И ЭЛЕМЕНТАРНЫМ ФУНКЦИЯМ

Редактор *О В. Семененко* Художник *В А Щорц*
Художественный редактор *Н Ф Зыков*
Технический редактор *З С Кондрашова*
Корректоры *С. С Мазурская, Л. А Айдарбекова, И. С. Хлыстова*

БЗ № 29—16—77. ИБ № 343

Сдано в набор 25 07.77. Подп. в печать 10 05 78. Л-78316. Формат 60×90 1/16. Бум. тип. № 3. Гарнитура литературная. Печать высокая. Усл. печ. л. 24,0. Уч.-изд. л. 23,7. Тираж 55 000 экз. (1-й завод 1—30 000 экз.) Зак. № 1432. Цена 1 р. 30 к. Изд. № 3135

Издательство Московского университета. Москва, К 9, ул. Герцена, 5/7.

Ордена Октябрьской Революции, ордена Трудового Красного Знамени Ленинградское производственно-техническое объединение «Печатный Двор» имени А. М. Горького Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств, полиграфии и книжной торговли. 197136, Ленинград, П-136, Гатчинская ул., 26.