manual de **matemáticas** para la enseñanza media

A.G. Tsipkin

ИЗДАТЕЛЬСТВО • МИР •

А. Г. Цыпкин

Справочник по математике для средней школы

ИЗДАТЕЛЬСТВО «НАУКА»

A. G. Tsipkin

manual de matemáticas para la enseñanza media

Editorial Mir Moscú

Traducido del ruso por T. I. Shapovalova

A nuestros lectores:

Mir edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas extranjeros. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia-ficción. Dirijan sus opiniones a la Editorial Mir, 1 Rizhski per., 2, 129820, Moscú, I-110, GSP, URSS.

Impreso en la URSS
На испанском языке

Índice

Capítulo 5. **Método de coordenadas**

Capítulo 6. **Geometría**

Prólogo

El manual está destinado para las escuelas de enseñanza media y los centros docentes medios especializados y contiene conceptos, definiciones, fórmulas, teoremas y métodos de resolución de los problemas que se incluyen en los cursos de matemáticas para la enseñanza media. Algunos de los apartados que integran este libro no forman parte actualmente del programa escolar, pero son imprescindibles para una mejor comprensión de los fundamentos de las matemáticas. Entre ellos cabe señalar: la divisibilidad de los números enteros y polinomios, los algoritmos de Euclides, los números complejos y el teorema fundamental del álgebra, las curvas de segundo orden, etc.

En una serie de párrafos la exposición del material se distingue de la exposición que se usa actualmente en los manuales de escuela. Por ejemplo, el material de geometría se expone en base a la axiomática de Hilbert, con el empleo de la terminología tradicional. Además, se definen de otra manera el vector, la sucesión numérica y la integral definida.

En el manual se da una exposición sistemática de la teoría de los números reales, que finaliza con un capítulo sobre los números complejos.

Se ha ampliado el círculo de cuestiones referidas al cálculo aproximado. Junto con los conocimientos elementales que tienen carácter general, en el manual se examinan las propiedades aproximadas de las fracciones continuas y se da uno de los más simples y conocidos métodos del cálculo aproximado, el método de tangentes de Newton. Las designaciones en el manual, a rara excepción, coinciden con las designaciones que suelen utilizarse en los manuales de escuela.

El manual tiene fundamentalmente un carácter teórico y puede servir no sólo como guía, sino que también como un libro de repaso y para sistematizar los conocimientos, siendo muy útil para la preparación de los exámenes de ingreso en los centros de enseñanza superior.

CAPÍTULO 1

Elementos de la teoría de conjuntos

La teoría de conjuntos es la parte de las matemáticas que estudia las propiedades generales de los conjuntos (principalmente infinitos). La separación de la teoría de conjuntos en una parte autónoma de las matemáticas se produjo hace comparativamente poco tiempo, en el tope de los siglos XIX y XX. La teoría de conjuntos ejerció una gran influencia sobre el desarrollo de las matemáticas contemporáneas, sirvió de base para la aparición de una serie de partes nuevas de esta ciencia, permitió enfocar de una manera nueva las matemáticas clásicas y entender más profundamente esta asignatura.

§ 1. Conjuntos y operaciones con ellos

1. 1. Conjuntos y subconjuntos. Ciertos conceptos de la matemáticas son primarios, indefinibles. Estos conceptos son: el número natural, el punto, la recta, etc.

Uno de estos conceptos indefinibles es el concepto «conjunto». A este concepto no puede dársele una definición formal que no se convierta en un simple cambio de la palabra «conjunto» por sus sinónimos: «población», «lista de elementos», etc. Los *conjuntos* se pueden componer sobre la base de distintos criterios de los más diferentes objetos (a los que en adelante vamos a denominar como *elementos* de un conjunto). Se pueden examinar no sólo los conjuntos, elementos de los cuales son los objetos materiales, sino también los conjuntos, elementos de los cuales son ciertos conceptos abstractos (números, figuras geométricas, símbolos, etc.).

Acordemos de antemano que el concepto «conjunto» no puede comprenderse literalmente e interpretarse como un conjunto de «muchos» elementos. Por conjunto se entiende también un grupo de objetos que puede estar compuesto, por ejemplo, de uno, dos y muchos más elementos. Además, resulta muy cómodo por conjunto entender también un conjunto *vacío*, es decir, un conjunto que no contiene ni un sólo elemento.

Los conjuntos suelen designarse con las letras mayúsculas del alfabeto latino A, B, . . ., X, y sus elementos, con las minúsculas: a, b, . . ., x. El conjunto vacío se designa con el símbolo $\varnothing$.

Si el conjunto A se compone de n elementos a_1, a_2, ..., a_n, entonces se escribe

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

Se dice que «el elemento a pertenece al conjunto A» y se escribe: $a \in A$ ó $A \ni a$ (A contiene a); la inscripción $a \notin A$ ó $A \not\ni a$ significa que el elemento a no pertenece al conjunto A (A no contiene a).

El conjunto B se llama *subconjunto* del conjunto A, si todos los elementos del conjunto B pertenecen al conjunto A, y se escribe

$$B \subset A.$$

Por ejemplo, sea A un conjunto de números racionales; B, un conjunto de números naturales. En este caso $B \subset A$.

Cualquier conjunto A tiene como subconjuntos un conjunto vacío y el propio conjunto A.

Si para dos conjuntos A y B son válidas al mismo tiempo las afirmaciones

$$A \subset B \text{ y } B \subset A,$$

entonces los conjuntos A y B se componen de los mismos elementos. Los conjuntos A y B se llaman *iguales* (o *coincidentes*) y se escriben

$$A = B.$$

El subconjunto no vacío B del conjunto A se llama *propio*, si B no coincide con A.

1.2. Operaciones con los conjuntos. Con los conjuntos se pueden efectuar distintas operaciones, de las cuales las más simples son: la unión, la intersección y el complemento. Para mayor claridad representemos los conjuntos en forma de figuras geométricas; por ejemplo, los elementos del conjunto, representado en la fig. 1.1, son los puntos de la parte rayada del plano.

Se llama *unión* de dos conjuntos, al conjunto compuesto de todos los elementos, que pertenecen, por lo menos, a uno

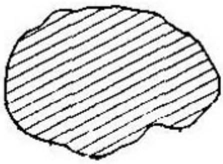


Fig. 1.1.

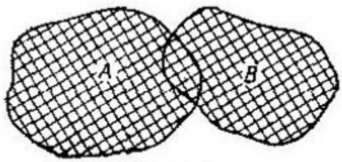


Fig. 1.2.

de estos conjuntos. La unión de los conjuntos A y B se designa $A \cup B$. En la fig. 1.2 el conjunto $A \cup B$ está indicado con un rayado doble.

Un conjunto compuesto de todos los elementos que pertenecen al mismo tiempo a dos conjuntos, se llama *intersec-*

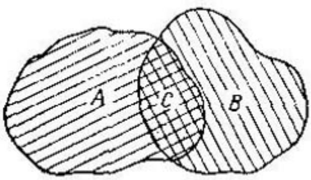


Fig. 1.3.

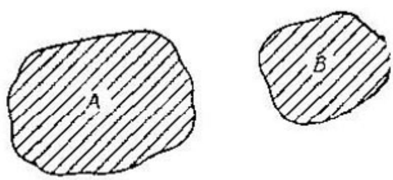


Fig. 1.4.

ción de estos conjuntos. La intersección de los conjuntos A y B se anota $A \cap B$. En la fig. 1.3 el conjunto $C = A \cap B$ está representado con un rayado doble.

Los conjuntos A y B pueden ser tales que su intersección será un conjunto vacío: $A \cap B = \varnothing$ (fig. 1.4).

Las propiedades de las operaciones de unión e intersección son:

1) *conmutatividad*:

$$A \cup B = B \cup A,$$
$$A \cap B = B \cap A;$$

2) *asociatividad*;

$$(A \cup B) \cup C = A \cup (B \cup C),$$
$$(A \cap B) \cap C = A \cap (B \cap C);$$

3) *distributividad*.

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

Se llama *complemento* del conjunto B hasta el conjunto A, que contiene B, al conjunto de todos los elementos de A, que no pertenecen a B. El complemento del conjunto B hasta el conjunto A se designa $A \setminus B$. En la fig. 1.5 el complemento $A \setminus B$ está representado con un rayado doble.

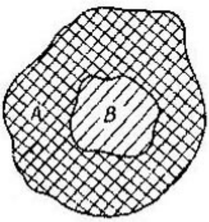


Fig. 1.5.

Aclaremos el sentido de estas definiciones con el siguiente ejemplo. Sea A el conjunto de todos los números naturales, múltiplos de dos, y B, el conjunto de todos los números naturales múltiplos de tres. La unión de estos conjuntos es el conjunto de todos los números naturales múltiplos de 2 ó de 3. Su intersección es el conjunto de todos los números naturales múltiplos tanto de 2 como de 3, es decir, múltiplos de 6.

Sea ahora A el conjunto de todos los números naturales múltiplos de 2 y B el conjunto de todos los números naturales múltiplos de 6 (es decir, múltiplos tanto de 2 como de 3). El conjunto B es un subconjunto del conjunto A. El conjunto de todos los números naturales múltiplos de 2, pero que no son múltiplos de 3, es el complemento del conjunto B hasta el conjunto A.

Observemos que las nociones de unidad e intersección dadas para el caso de dos conjuntos pueden ser válidas también para el caso de cualquier número finito de conjuntos.

Pues bien, el conjunto compuesto de elementos, cada uno de los cuales pertenece, por lo menos, a uno de los conjuntos A_i, se llama unión del número finito de conjuntos A_i y se escribe

$$\bigcup_{i=1}^{n} A_i.$$

§ 2. Correspondencia entre los conjuntos y aplicación de los mismos

2.1. Correspondencia y aplicación. A veces se puede demostrar la *correspondencia* entre dos conjuntos, es decir, se puede introducir una regla, según la cual a cada elemento de un conjunto se asigna un elemento definido o un conjunto de elementos de otro conjunto. Al mismo tiempo se admite que a ciertos elementos del primer conjunto les puede corresponder un conjunto vacío.

Sobre la base del concepto de correspondencia entre conjuntos se introduce el concepto de aplicación de los conjuntos. Al mismo tiempo se distinguen la aplicación «en» el conjunto y la aplicación del conjunto «sobre» el conjunto.

La correspondencia, para la cual a cada elemento del conjunto A le corresponde el único elemento del conjunto B, se llama *aplicación del conjunto A en el conjunto B.*

La correspondencia, para la cual a cada elemento del conjunto A le corresponde el único elemento del conjunto B y, además, a cada elemento del conjunto B le corresponde, por lo menos, un elemento del conjunto A se llama *aplicación del conjunto A «sobre» el conjunto B.*

A las aplicaciones de conjuntos se les asignan las letras f, g, h, . . . y se escribe

$$A \xrightarrow{f} B, \quad A \xrightarrow{g} B, \quad A \xrightarrow{h} B$$

o

$$f\colon A \to B, \quad g\colon A \to B, \quad h\colon A \to B.$$

Si en la aplicación f al elemento $a \in A$ le corresponde el elemento $b \in B$, entonces al elemento b se le llama *imagen* del elemento a y al elemento a, *preimagen* del elemento b y se escribe

$$b = f(a).$$

El conjunto de todas las preimágenes del elemento b se llama *preimagen completa* del mismo. En el caso de la aplicación del conjunto A «sobre» el conjunto B se escribe también

$$B = f(A).$$

2.2. Aplicación biunívoca. La aplicación del conjunto A sobre el conjunto B, para la cual a cada elemento del conjunto B le corresponde el único elemento del conjunto A, se llama *aplicación biunívoca* del conjunto A sobre el conjunto B. A la aplicación biunívoca entre los dos conjuntos la llaman también *biyección*.

Sea que el conjunto A se aplica biunívocamente sobre el conjunto B:

$$A \xrightarrow{f} B,$$

es decir, a cada elemento $a \in A$ le corresponde el único elemento b del conjunto B y a cada elemento $b \in B$ le corresponde el único elemento $a \in A$. La aplicación f^{-1} que pone en la correspondencia a cada elemento $b \in B$ su preimagen $a \in A$, la llaman *aplicación inversa* a la aplicación f y se escribe

$$B \xrightarrow{f^{-1}} A \text{ o } f^{-1}: B \to A.$$

Si f es una aplicación biunívoca, entonces la aplicación inversa f^{-1} será una aplicación también biunívoca; la aplicación inicial f será inversa a la aplicación f^{-1}.

2.3. Equivalencia de conjuntos. Si el conjunto A se aplica biunívocamente sobre el conjunto B, entonces los conjuntos A y B se llaman *equivalentes*. La equivalencia de los conjuntos se escribe mediante el signo:

$$A \sim B$$

(se lee: A es equivalente a B).

De los conjuntos equivalentes A y B se dice también que entre ellos está establecida una *relación de equivalencia*.

La relación de equivalencia de los conjuntos tiene las propiedades siguientes:

1) *reflexividad*: $A \sim A$ (cualquier conjunto es equivalente a sí mismo);

2) *simetría*: $A \sim B \Leftrightarrow B \sim A$ (si el conjunto A es equivalente al conjunto B, entonces el conjunto B es equivalente al conjunto A, y a la inversa);

3) *transitividad*: $A \sim B$ y $B \sim C \Rightarrow A \sim C$ (si el conjunto A es equivalente al conjunto B, y el conjunto B es equivalente al conjunto C, entonces y el conjunto A es equivalente al conjunto C).

2.4. Clasificación de conjuntos. Los conjuntos se dividen en conjuntos finitos e infinitos.

Un conjunto equivalente a un segmento de una serie natural*) se llama *conjunto finito* (el conjunto vacío también se considera como conjunto finito). En otras palabras, el conjunto finito (si no es vacío) es un conjunto, cuyos elementos se pueden «calcular» por un número finito de pasos. Si para eso damos un número natural cualquiera desde 1 hasta n, de tal manera que todos los números desde la unidad hasta n sean utilizados y los distintos elementos del conjunto reciban estos números, entonces el número n indica la cantidad de elementos en el conjunto A dado.

La cantidad de elementos en el conjunto finito A se llama *potencia* de éste.

Para los conjuntos finitos es válido el teorema llamado teorema principal de los conjuntos finitos:

Cualquier conjunto finito no es equivalente a ninguno de sus propios subconjuntos.

Del teorema citado más arriba, en particular, se deduce también que cada conjunto finito, no vacío, es equivalente a uno y sólo a uno de los segmentos de una serie natural.

Un conjunto no finito se llama *infinito*. Como «patrón» para la comparación de los conjuntos infinitos se puede escoger el conjunto infinito más simple, el conjunto de todos los números naturales **N**. Con todo eso, para la comparación de conjuntos nuevamente utilizaremos el concepto de la equivalencia de conjuntos.

Si realizamos la comparación de cualesquiera conjuntos infinitos con el conjunto de todos los números naturales **N**, resulta que todos los conjuntos infinitos se dividen en dos clases: la clase de conjuntos equivalentes al conjunto de todos los números naturales (tales conjuntos se llaman *numerables*) y la clase de conjuntos no equivalentes al conjunto de los números naturales (tales conjuntos se llaman *innumerables*).

De esta manera, el conjunto numerable es el conjunto infinito, cuyos elementos pueden «enumerarse» con ayuda del conjunto de números naturales, es decir, se puede indicar tal vía de numeración de elementos del conjunto, para la cual cada elemento del conjunto reciba su único número.

Los conjuntos numerables son, por ejemplo, el conjunto de todos los números naturales pares, el conjunto de todos los números enteros, el conjunto de todos los números racionales, etc.

Los ejemplos citados más arriba de los distintos conjuntos numerables parecen ser, a primera vista, un poco extraños: es evidente que el conjunto de todos los números naturales pares es el subconjunto del conjunto de todos los números naturales y, al mismo tiempo, el conjunto de los números pares es equivalente al conjunto de los números naturales. Esta paradoja es imaginaria y surge porque estamos acostumbrados a considerar los conjuntos finitos, es decir, los conjuntos

---

*) El conjunto de todos los números naturales iguales o menos que cierto número natural n se llama *segmento de una serie natural* y se designa $|1, n|$ ó $\overline{1, n}$

que se componen de un número finito de elementos, para los cuales la sustracción, por lo menos de un sólo elemento, produce la disminución de su número. De resultas de tal sustracción obtenemos un conjunto finito nuevo que ya no es equivalente al inicial, porque los conjuntos finitos son equivalentes sólo en el caso de igualdad del número de sus elementos. Todo lo contrario sucede con los conjuntos numerables. La sustracción de un conjunto numerable de un número finito (a veces incluso infinito) de elementos sigue dejando al conjunto par.

Mostremos que el conjunto de todos los números naturales y el conjunto de todos los números naturales pares son equivalentes entre sí. Tracemos dos ejes numéricos paralelos Ox y $O'x'$ (fig. 1.6), escogiendo los segmentos de unidad y de escala de tal manera que el segmento de unidad del eje Ox sea dos veces más grande que el segmento de unidad del eje $O'x'$. El conjunto de números naturales vamos a representarlo mediante los puntos en los ejes numéricos Ox y $O'x'$. Pongamos en correspondencia al número 2 del eje numérico $O'x'$, el número 1 del eje numérico Ox; al número 4 del eje numérico $O'x'$, el número 2 del eje numérico Ox; al número 6, el número 3, etc. Con todo esto resulta que a cualquier número natural par le corresponde un único número natural y, a la inversa, cualquier número natural corresponde a un único número natural par, es decir, entre el conjunto de todos los números naturales pares y el conjunto de todos los números naturales se establece una correspondencia biunívoca, y, por consiguiente, estos conjuntos son equivalentes.

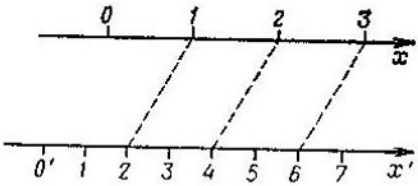


Fig. 1.6.

El ejemplo, recién citado, muestra que el conjunto de números naturales contiene un subconjunto propio equivalente a él (el conjunto de números pares). Tal propiedad de contener un subconjunto propio equivalente al mismo conjunto, la tienen todos los conjuntos infinitos. La propiedad indicada puede tomarse como definición del conjunto infinito.

Propiedades de los conjuntos infinitos:

1) Cualquier subconjunto del conjunto numerable es finito o numerable.
2) La suma de cualquier conjunto, finito o numerable, de los conjuntos numerables es un conjunto numerable.
3) Cualquier conjunto infinito contiene un subconjunto numerable.

La última propiedad dice que entre los conjuntos infinitos los conjuntos numerables son «los más pequeños».

Existen conjuntos infinitos que no son numerables; estos *conjuntos* son *innumerables*. Observemos que esta «definición» del conjunto innumerable no puede considerarse suficientemente satisfactoria por la causa siguiente: introduciendo el concepto del conjunto infinito, estábamos seguros de que por lo menos existe tal conjunto. Este conjunto es el conjunto de números naturales. Para «la definición» dada del conjunto innumerable ya tenemos que demostrar, en calidad de teorema, que existen tales conjuntos. Este *teorema* se enuncia del modo siguiente:

El conjunto de los números reales comprendidos entre 0 y 1 es innumerable.

No vamos a demostrar este teorema fundamental de la teoría de conjuntos, pero observemos, que la demostración se basa en la representación de los números reales en forma de fracciones decimales infinitas. La esencia de la demostración consiste en que, de cualquier manera que numeremos los números situados entre el cero y la unidad, utilizando todos los números naturales, siempre se halla tal número real, mayor que cero y menor que la unidad, que puede ser omitido.

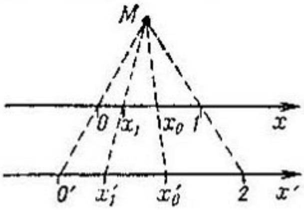


Fig. 1.7.

Si un conjunto es equivalente al conjunto de todos los números reales, mayores que cero y menores que la unidad, se dice que el primero *tiene potencia de continuo.*

El conjunto de todos los puntos de cualquier segmento de una recta, el conjunto de todos los puntos de la recta, el conjunto de todas las rectas en el plano, etc., sirven de ejemplo a los conjuntos equivalentes al de los números reales comprendidos entre cero y la unidad, es decir, los que tienen potencia de continuo.

Es muy fácil observar que los ejemplos citados crean la misma situación «paradoja» que surge también en el caso de los conjuntos numerables, es decir, el conjunto de números que se encuentran entre 0 y 1 resulta equivalente al conjunto de todos los números mayores que cero y menores, por ejemplo, que dos. En el caso de los conjuntos innumerables «la paradoja» se explica de la misma manera que en el caso de los conjuntos numerables, por eso no vamos a examinarlo detalladamente. Aquí mostraremos solamente la manera en que se puede demostrar la correspondencia biunívoca entre un conjunto de números mayores que cero y menores que la unidad, y entre el conjunto de números mayores que cero y menores que dos, y al mismo tiempo demostrar su equivalencia. Tracemos dos ejes paralelos numéricos Ox y $O'x'$, escogiendo en ellos segmentos de escala iguales (fig. 1.7). Ahora describamos el procedimiento que permite demostrar la correspondencia biunívoca entre los puntos pertenecientes al segmento $[0; 1]$ del eje Ox y los puntos que pertenecen al segmento $[0; 2]$ del eje $O'x'$. Tracemos a través de los puntos O y O' y de los puntos 1 y 2, unas rectas que se intersecan en cierto punto M. Tomemos un punto cualquiera $x_0 \in [0; 1]$ en el eje Ox y tracemos una recta que pase por los puntos M y x_0. Esta recta se cruza con el eje $O'x'$ en cierto punto x_0', el cual representa la imagen del punto x_0 para la aplicación del intervalo $[0; 1]$ del eje Ox sobre el intervalo $[0; 2]$ del eje $O'x'$. Es evidente que tal punto x_0' es el único punto de intersección, debido a lo cual las rectas Mx_0 y $O'x'$ no son paralelas. Y a la inversa, tomando un punto cualquiera $x_1' \in [0; 2]$ en el eje $O'x'$ y trazando una recta que pase por el punto M y el punto x_1' que pertenece al eje $O'x'$, obtendremos el único punto de intersección de la recta dada con el eje Ox, el cual designamos x_1. El punto x_1 es la imagen del punto x_1' para la aplicación del conjunto $[0; 2]$ del eje $O'x'$ sobre el conjunto $[0; 1]$ del eje Ox. De esta manera hemos demostrado la correspondencia biunívoca entre nuestros conjuntos y al mismo tiempo hemos demostrado su equivalencia.

§ 3. Conjuntos ordenados

3.1. Concepto del conjunto ordenado. Introduciendo las operaciones con los conjuntos y clasificando los distintos conjuntos, no teníamos en cuenta que los conjuntos mismos pueden tener una estructura interior, propia a ellos, es decir, suponíamos que todos los elementos del conjunto son «iguales». Pero en matemática estos conjuntos «puros» resultan poco interesantes y más a menudo se estudian los conjuntos, entre cuyos elementos existen unas u otras relaciones. Una de las relaciones más importantes entre los elementos del conjunto es la relación de orden. Supongamos que A es cierto conjunto. El conjunto A se llama *conjunto ordenado*, si para cualesquiera de sus dos elementos a y b existe una de las siguientes *relaciones de orden*:

bien $a \leqslant b$ (a no supera a b),

o bien $b \leqslant a$ (b no supera a a),

que tienen las propiedades siguientes:

1) *reflexividad*: cualquier elemento no se supera a sí mismo;

2) *antisimetría*: si a no supera a b y b no supera a a, entonces los elementos a y b coinciden;

3) *transitividad*: si a no supera a b y b no supera a c, entonces a no supera a c.

Hemos considerado al conjunto vacío como ordenado.

Dos conjuntos compuestos de unos mismos elementos, pero con distintas relaciones de orden, se consideran conjuntos ordenados diferentes; por eso, para la representación del conjunto ordenado a través de sus elementos es necesario también representar su orden, es decir, indicar la regla, conforme a la cual respecto a cualquier par de elementos a, b se puede decir, cual de los elementos del par dado no supera a tal elemento, con el cumplimiento posterior de las propiedades 1)—3).

Observemos, que el mismo conjunto se puede ordenar por los distintos procedimientos, obteniendo al mismo tiempo distintos conjuntos ordenados. Por ejemplo, examinemos un conjunto, los elementos del cual son distintos polígonos convexos: un triángulo, cuadrilátero, pentágono, hexágono, etc. Un procedimiento de la formación de un conjunto ordenado a partir del conjunto desordenado dado consiste en que como primer elemento del conjunto ordenado tomamos el triángulo; como segundo, el cuadrilátero; como tercero, el pentágono, etc., es decir, disponemos el conjunto en orden de crecimiento del número de ángulos interiores de los polígonos. El otro procedimiento de ordenación del conjunto de polígonos puede ser su enumeración en el orden de crecimiento de sus áreas, es decir, como primero tomamos el polígono con un área más pequeña, como segundo, un polígono con un area que no supera la de los demás, menos el elegido, etc.

Los conjuntos ordenados (finitos o numerables) se escriben muy a menudo situando sus elementos según el orden dado entre paréntesis. Por ejemplo, las inscripciones

$$(1;\ 2;\ 3) \text{ y } (2;\ 1;\ 3) \tag{1}$$

representan distintos conjuntos ordenados, los cuales se puede obtener del mismo conjunto $\{1;\ 2;\ 3\}$, ordenándolo mediante los dos distintos procedimientos. Para escribir un conjunto numerable ordenado en forma análoga a la (1), es necesario indicar el primer elemento del conjunto ordenado y señalar el orden (la regla) de la posición de los elementos sucesivos.

3.2. Reordenaciones. Sea dado cierto conjunto finito A compuesto de n elementos distintos:

$$A = \{a_1;\ a_2;\ a_3;\ \ldots\ ;\ a_n\}.$$

Formemos de los elementos del conjunto A conjuntos ordenados. Tomemos como primer conjunto ordenado uno, cuya totalidad de elementos está dispuesta en orden de crecimiento de sus números:

$$(a_1;\ a_2;\ a_3;\ \ldots;\ a_n).$$

El segundo conjunto ordenado se puede formar cambiando de sitios los elementos a_1 y a_2 y dejando todos los demás elementos del primer conjunto en sus sitios:

$$(a_2;\ a_1;\ a_3;\ \ldots\ a_n).$$

Cambiando de sitio los elementos a_1 y a_3 y dejando en sus sitios todos los demás elementos, en el primer conjunto ordenado, obtenemos un conjunto ordenado que se distingue tanto del primer, como también del segundo conjunto ordenado. Análogamente a los elementos del conjunto A se pueden construir también otros conjuntos ordenados.

Los conjuntos finitos ordenados de toda clase que contienen n elementos distintos, los cuales se puede obtener de cierto conjunto desordenado, compuesto de n elementos distintos, se llaman *reordenaciones* de n elementos.

De tal manera, la reordenación no es otra cosa que el procedimiento de ordenación de elementos de cierto conjunto finito. Al mismo tiempo, dos reordenaciones distintas cualesquiera representan por sí mismas dos procedimientos distintos de formación de un conjunto ordenado (del conjunto desordenado dado).

Surge el problema de cuántos conjuntos ordenados distintos se pueden formar del conjunto finito A dado que contiene n elementos distintos, es decir, a qué es igual el número de reordenaciones de los elementos del conjunto

$$A = \{a_1;\ a_2;\ a_3;\ \ldots;\ a_n\}.$$

Cualquier conjunto ordenado, obtenido del conjunto A, se puede representar condicionalmente en la forma de:

$$(i_1;\ i_2;\ \ldots;\ i_k;\ \ldots;\ i_n),$$

donde cada uno de los elementos i_k es uno de los elementos $a_1, a_2, a_3, \ldots, a_n$; al mismo tiempo ninguno de los elementos $a_1, a_2, \ldots, a_n$ se encuentra más de una sola vez. Como i_1 se puede tomar cualquiera de los elementos $a_1, a_2, a_3, \ldots, a_n$ lo que proporciona posibilidades distintas. Ahora, cuando ya hemos elegido i_1, como i_2 se puede tomar cualquiera de los $n - 1$ de elementos restantes del conjunto A, es decir, el número de procedimientos distintos de elección de los elementos i_1, i_2 será igual a $(n - 1)\cdot n$. Continuando el proceso indicado, de resultas obtendremos que el último elemento restante i_n puede ser eligido sólo mediante un procedimiento.

El número P_n de reordenaciones de n elementos es igual al producto de los n números naturales sucesivos, comenzando por la unidad. Este producto tiene una denominación especial n (se lee: *n factorial*). De tal manera,

$$P_n = 1 \cdot 2 \cdot 3 \ldots (n-1) \cdot n = n!$$

Para el conjunto vacío se acuerda lo siguiente: el conjunto vacío se puede ordenar sólo mediante un procedimiento; por eso se considera que

$$0! = 1.$$

Si en la reordenación dada cambiamos de sitios dos elementos cualesquiera, el resto de los elementos los dejaremos en sus sitios, entonces se obtiene una reordenación nueva.

Tal transformación de reordenación se denomina *transposición*.

Todas $n!$ reordenaciones de n elementos se pueden ordenar en tal orden que cada reordenación siguiente se obtenga de la precedente mediante una sola transposición; al mismo tiempo, en calidad de reordenación inicial, se puede eligir cualquiera de las $n!$ reordenaciones. En particular, de cualquier reordenación de n elementos se puede pasar a cualquiera otra reordenación de estos mismos elementos mediante algunas transposiciones.

La reordenación $(i_1;\ i_2;\ \ldots;\ i_n)$ se llama *par*, si se obtiene de la reordenación $(a_1;\ a_2;\ \ldots;\ a_n)$ mediante un número par de transposiciones e *impar* en el caso contrario.

Ejemplo. Sea $A = \{1;\ 2;\ 3;\ 4\}$. Según la definición, la reordenación $(1;\ 2;\ 3;\ 4)$ es par. La reordenación $(4;\ 2;\ 1;\ 3)$ es par también, por lo que ella puede ser obtenida de la reordenación $(1;\ 2;\ 3;\ 4)$ mediante un número par de transposiciones: cambiamos de sitios en la reordenación $(4;\ 2;\ 1;\ 3)$ el primer y el cuarto elementos:

$$(\underrightarrow{\overleftarrow{4;\ 2;\ 1;\ 3}}) \to (3;\ 2;\ 1;\ 4);$$

en la reordenación obtenida cambiamos de sitios el primer y tercer elementos:

$$(\underrightarrow{\overleftarrow{3;\ 2;\ 1}};\ 4) \to (1;\ 2;\ 3;\ 4).$$

De tal manera, la reordenación dada se transforma en la inicial como resultado de dos transposiciones y, por consiguiente, es par. Por analogía podemos convencernos de que la reordenación $(4;\ 1;\ 2;\ 3)$ es impar, élla se transforma en la reordenación $(1;\ 2;\ 3;\ 4)$ después de tres transposiciones:

$$(\underrightarrow{\overleftarrow{4;\ 1}};\ 2;\ 3) \to (1;\ \underrightarrow{\overleftarrow{4;\ 2}};\ 3) \to (1;\ 2;\ \underrightarrow{\overleftarrow{4;\ 3}}) \to (1;\ 2;\ 3;\ 4),$$

Cualquier transposición cambia la paridad de la reordenación: el número de reordenaciones pares de n elementos ($n \geqslant 2$) es igual a un número impar, es decir, es igual a $n!/2$.

3.3. Sustituciones. Sea que A es un conjunto finito, compuesto de n elementos distintos. Para mayor simplicidad consideremos que éstos son los primeros n números naturales, es decir,

$$A = \{1;\ 2;\ 3;\ \ldots;\ n\}.$$

Cualquier aplicación biunívoca F del conjunto A sobre sí se denomina *sustitución* de potencia n.

Para denominar la sustitución F se usa la inscripción

$$F = \begin{pmatrix} i_1 & i_2 & i_3 & \ldots & i_n \\ k_1 & k_2 & k_3 & \ldots & k_n \end{pmatrix}, \tag{2}$$

donde $(i_1;\ i_2;\ i_3;\ \ldots;\ i_n)$, $(k_1;\ \ldots;\ k_n)$ son las dos reordenaciones de n elementos del conjunto A ($k_1, k_2, k_3, \ldots, k_n$ son las imágenes de los elementos $i_1, i_2, i_3, \ldots, i_n$ para la aplicación F).

La misma sustitución puede ser escrita de distintas maneras. Pues bien, la sustitución (2) también puede ser escrita en forma de

$$\begin{pmatrix} i_2 & i_1 & i_3 & \ldots & i_n \\ k_2 & k_1 & k_3 & \ldots & k_n \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} i_n & i_{n-1} & \ldots & i_3 & i_2 & i_1 \\ k_n & k_{n-1} & \ldots & k_3 & k_2 & k_1 \end{pmatrix}.$$

En particular, cualquier sustitución de potencia n puede ser escrita en la forma siguiente:

$$F = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ l_1 & l_2 & l_3 & \ldots & l_n \end{pmatrix},$$

es decir, en el orden de crecimiento de los números naturales en la línea superior. Para tal inscripción las distintas sustituciones se distinguen la una de la otra por las reordenaciones que se encuentran en la línea inferior y por eso el número de sustituciones de la potencia n es igual al de reordenaciones de n elementos, es decir, es igual a $n!$.

La sustitución

$$E = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 1 & 2 & 3 & \ldots & n \end{pmatrix}$$

se llama *sustitución idéntica* de la potencia n.

Examinemos ahora una sustitución arbitraria de la potencia n:

$$F = \begin{pmatrix} i_1 & i_2 & i_3 & \ldots & i_n \\ k_1 & k_2 & k_3 & \ldots & k_n \end{pmatrix}.$$

Las reordenaciones de n elementos que forman las líneas superior e inferior de la sustitución F, o tienen paridades iguales (es decir, ambas son pares, o impares) o tienen paridades contrarias (es decir, una es par, y la otra, impar), al mismo tiempo, la coincidencia o no coincidencia de paridades de las reordenaciones no depende de la manera de escribir la sustitución dada.

Si las paridades de las líneas superior e inferior de la sustitución F coinciden, entonces la sustitución se llama *par*; si son opuestas, la sustitución se llama *impar*.

El número de sustituciones pares de la potencia n es igual al de sustituciones impares, es decir, es igual a $n!/2$.

Multiplicación de sustituciones. Según la definición, la sustitución de la potencia n es la aplicación biunívoca F del conjunto finito de n elementos distintos sobre sí. La realización sucesiva de las dos sustituciones de la potencia n nuevamente es la aplicación biunívoca del conjunto A sobre sí y provoca una tercera sustitución, en sumo grado definida, de la potencia n que se llama producto de la primera de las sustituciones dadas sobre la segunda.

Se sabe también que el producto de dos sustituciones dadas F y G se obtiene como resultado de la multiplicación de la sustitución F por la sustitución G.

Sean dadas dos sustituciones de cuarta potencia

$$F=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \quad G=\begin{pmatrix} 2 & 3 & 1 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}.$$

El producto de las sustituciones F y G es la sustitución

$$FG=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

Efectivamente, en la sustitución F el símbolo 1 pasa al 4, y en la sustitución G el símbolo 4 pasa al 2, por eso para la aplicación FG el símbolo 1 pasa al 2, etc.

La multiplicación de las sustituciones está definida sólo para las sustituciones de igual potencia y tiene las *propiedades* siguientes:

1) La multiplicación de las sustituciones de la potencia n para $n \geqslant 3$ no es conmutativa, es decir, la sustitución FG no siempre es igual a la sustitución GF.

2) La multiplicación de las sustituciones es asociativa, es decir, $F(GH) = (FG)H$; por eso podemos hablar del producto de cualquier número finito de sustituciones de la potencia n.

3) El producto de cualquier sustitución F por una sustitución idéntica E, así como el producto de E por F es igual a F:

$$EF = FE = F.$$

Se denomina *sustitución inversa* a la sustitución de la n-ésima potencia de F tal sustitución de la n-ésima potencia de G para la cual

$$FG = GF = E.$$

La sustitución, inversa a la sustitución F, se escribe por lo común F^{-1}; así pues

$$FF^{-1} = F^{-1}F = E.$$

La sustitución obtenida de F mediante el cambio de sitios de líneas

$$F^{-1}=\begin{pmatrix} k_1 & k_2 & k_3 & \dots & k_n \\ i_1 & i_2 & i_3 & \dots & i_n \end{pmatrix}$$

es una sustitución inversa a la sustitución

$$F=\begin{pmatrix} i_1 & i_2 & i_3 & \dots & i_n \\ k_1 & k_2 & k_3 & \dots & k_n \end{pmatrix}$$

3.4. Variaciones. Sea que cierto conjunto finito A, está compuesto de n elementos distintos. Elijamos de cierta manera de entre estos n elementos, m elementos distintos y compongamos de ellos distintos conjuntos ordenados.

Las *variaciones* de n elementos a m elementos en cada una se llaman subconjuntos ordenados finitos que contienen m elementos cada uno, eligidos de n elementos del conjunto principal. El número de variaciones posibles de n elementos a razón de m elementos en cada una se denomina A_n^m.

Es muy fácil convencerse de que

$$A_n^1 = n,$$

es decir, un elemento de n se puede elegir mediante n procedimientos, y de un elemento se puede formar sólo un conjunto ordenado.

Hallemos a que es igual el número de variaciones de n elementos a razón de m elementos en cada una. Para distribuir $m+1$ elementos, tomados de n elementos por $m+1$ sitios, se pueden eligir primero m elementos y distribuirlos por las m posiciones primeras. Esto se puede efectuar mediante A_n^m procedimientos. Para cada elección de m elementos de n quedan «libres» $n-m$ elementos, cualquiera de los cuales se puede poner en la posición $(m+1)$ libre. De esta manera, de cada A_n^m ocupaciones de las m posiciones primeras obtendremos $n-m$ ocupaciones posibles de la $(m+1)$-ésima posición. Por consiguiente, el número de variaciones tomadas por $m+1$ elementos, eligidos de entre los n elementos dados se relaciona con el número de variaciones de n elementos por m mediante la igualdad

$$A_n^{m+1}=(n-m)\,A_n^m.$$

Teniendo en cuenta que $A_n^1 = n$, escribimos sucesivamente

$$A_n^1 = n,$$
$$A_n^2 = n(n-1),$$
$$A_n^3 = n(n-1)(n-2),$$
$$\dots\dots\dots\dots\dots\dots$$
$$A_n^m = n(n-1)(n-2)\dots(n-m+1).$$

La última igualdad se puede escribir en una forma más cómoda:

$$A_n^m = \frac{n!}{(n-m)!}.$$

Para $m = 0$ obtenemos $A_n^0 = 1$, es decir, de cualquier conjunto se puede extraer un conjunto vacío mediante un único procedimiento, el cual, según la definición, se puede ordenar mediante un único procedimiento.

El número de variaciones y el número de reordenaciones se relacionan por medio de la fórmula

$$A_n^n = P_n = n!$$

3.5. Combinaciones. Los conjuntos *desordenados* finitos que contienen m elementos distintos, eligidos de entre n elementos del conjunto dado se llaman combinaciones de n elementos por m. El número de combinaciones de n elementos por m elementos se designa C_n^m o $\left(\frac{n}{m}\right)$.

Hallemos a que es igual el número de combinaciones de n elementos por m. Debido a la definición de variaciones del conjunto dado que contiene n elementos se pueden formar A_n^m diferentes conjuntos ordenados que contienen m elementos distintos. Del conjunto que contiene m elementos diferentes se pueden formar P_m distintos conjuntos ordenados, por eso el número C_n^m de distintos conjuntos desordenados, que contienen m elementos distintos cada uno, eligidos de entre n elementos, será calculado según la fórmula

$$C_n^m = \frac{A_n^m}{P_m}.$$

Utilizando la fórmula para el cálculo del número de reordenaciones P_m y del número de variaciones A_n^m, obtendremos

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

Para el número de combinaciones son válidas las igualdades

$$\begin{gathered} C_n^m = C_n^{n-m}, \\ C_{n+1}^{m+1} = C_n^m + C_n^{m+1}, \end{gathered} \tag{3}$$

y también

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^{n-1} + C_n^n = 2^n.$$

La última propiedad se formula a veces en forma del *teorema* siguiente *sobre los conjuntos finitos*:

El número total de subconjuntos de un conjunto, compuesto de n elementos, es igual a 2^n.

Citemos aquí una tabla de valores C_n^m para $n = 10$ y $m \leqslant n$. Esta tabla se llama *triángulo de Pascal*. Anote-

Tabla 1

n \ m	0	1	2	3	4	5	6	7	8	9	10
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	
10	1	10	45	120	210	252	210	120	45	10	1

mos que la segunda de las fórmulas (3) permite seguir completando sucesivamente las líneas del triángulo de Pascal, utilizando el hecho de que al principio y al final de cada línea se encuentran unidades. Calculemos, por ejemplo,

C_{11}^4 no citado en la tabla:

$$C_{11}^4 = C_{10}^3 + C_{10}^4 = 120 + 210 = 330.$$

3.6. Binomio de Newton. La potencia natural de la suma de dos variables se calcula según la fórmula

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

que se llama *binomio de Newton*. Los coeficientes C_n^m que son iguales al número de combinaciones de n elementos por m se llaman *coeficientes binomiales*.

Propiedades de los coeficientes binomiales:

1) El número de los coeficientes binomiales (y, por consiguiente, el número de sumandos en el desarrollo de la potencia binomial) es igual a $n+1$.

2) Los coeficientes de los términos equidistantes del principio y del final del binomio son iguales entre sí (puesto que $C_n^m = C_n^{n-m}$).

3) La suma de los coeficientes binomiales es igual a 2^n.

4) La suma de los coeficientes binomiales que se encuentran en los lugares impares es igual a la suma de los coeficientes que se encuentran en los lugares pares.

La potencia natural de diferencia de dos variables se calcula por una fórmula, análoga al binomio de Newton:

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

§ 4. Método de inducción matemática

Muchas afirmaciones matemáticas se demuestran por el método de inducción matemática, el cual se basa en el principio de la *inducción matemática*:

Supongamos que $A(n)$ es cierta afirmación que tiene sentido para los números naturales n. Y sea que esta afirmación es auténtica para $n = 1$. Entonces, si de la autenticidad de esta afirmación para $n = k$ (k es cualquier número natural mayor que uno) se desprende la autenticidad de la afirmación para $n = k+1$, la afirmación $A(n)$ es auténtica para cualquier número natural n.

La demostración mediante el *método de la inducción matemática* consiste en lo siguiente:

1) se demuestra que la afirmación $A(1)$ es auténtica;

2) se presupone que la afirmación $A(n)$ es auténtica para $n = k$ y se demuestra su equidad para $n = k + 1$.

Después, sobre la base del principio de la inducción matemática, se deduce que la afirmación $A(n)$ es auténtica para cualquier n natural.

Como ilustración de la aplicación del método de la inducción matemática demostremos que para cualquier número natural n y cualquier número real $a > -1$ tiene lugar una desigualdad que se llama *desigualdad de Bernoulli*:

$$(1 + a)^n \geqslant 1 + an. \tag{1}$$

Si $n = 1$, entonces es evidente que la desigualdad (1) es cierta

$$(1 + a)^1 \geqslant 1 + a.$$

Supongamos que la desigualdad (1) es válida para $n = k$:

$$(1 + a)^k \geqslant 1 + ak.$$

Multipliquemos ambos miembros de la última desigualdad por el número positivo $1 + a$; de resultas obtenemos

$$(1 + a)^{k+1} \geqslant 1 + ak + a + a^2k. \tag{2}$$

Omitiendo el último sumando en el miembro segundo de la desigualdad (2), disminuimos el miembro segundo de esta desigualdad, y por eso

$$(1 + a)^{k+1} \geqslant 1 + a(k + 1).$$

El resultado obtenido muestra que la desigualdad (1) es válida también para $n = k + 1$.

Hemos efectuado ambas partes de la demostración por el método de la inducción matemática y, por consiguiente, la desigualdad (1) es válida para cualquier n natural.

Ilustremos la aplicación del método de inducción matemática, examinando una propiedad de los números que forman la sucesión de Fibonacci. La sucesión de los números $F_0, F_1, F_2, \ldots, F_n, \ldots$, donde $F_0 = 0$, $F_1 = 1$ y cada término siguiente es la suma de los dos antecedentes se llama la *sucesión de Fibonacci*, es decir, la sucesión indicada tiene la forma de 0, 1, 1, 2, 3, 5, 8, ...

La sucesión citada presenta la solución del problema que primeramente propuso en el año 1202 Leonardo Fibonacci y se formulaba de la manera siguiente: «¿Cuántos pares de conejos da una pareja durante un año? Al mismo tiempo se presupone que cada par da un par de conejos cada mes, cada nuevo par puede dar conejitos a la edad de un mes y, además, estos conejos no se mueren nunca».

Demostremos que si designamos el número $(1+\sqrt{5})/2$ por medio de Φ*), entonces

$$F_n \leqslant \Phi^{n-1} \tag{3}$$

para todos n naturales.

Si $n=1$, entonces $F_1 = 1 = \Phi^0$, es decir, se cumple una desigualdad no rigurosa (3).

Supongamos que para cualquier $k>1$ natural se cumplen las desigualdades

$$F_2 \leqslant \Phi^1,\ F_3 \leqslant \Phi^2,\ \ldots,\ F_k \leqslant \Phi^{k-1}.$$

Demostremos que entonces la desigualdad (3) es válida también para $n=k+1$.

Realmente, $F_{k+1} = F_{k-1} + \Phi_k \leqslant \Phi^{k-2} + \Phi^{k-1} =$ $= \Phi^{k-2}(1+\Phi)$.

Observando que $1+\Phi = \Phi^2$, definitivamente obtenemos

$$F_{k+1} \leqslant \Phi^{k-2}\Phi^2 = \Phi^k.$$

De esta manera, de la presuposición de que la desigualdad (3) es válida para $n=k$ se desprende que la misma es válida también para $n=k+1$.

Están efectuadas ambas partes de la demostración por el método de la inducción matemática y, por consiguiente, dicha afirmación es válida para cualquier n natural.

Mediante el método de la inducción matemática se puede demostrar también que son válidos todos n naturales de las relaciones siguientes:

1) $F_{n+1} \geqslant \Phi^{n-1}$;

2) $F_{n+1}\cdot F_{n-1} - F_n^2 = (-1)^n$;

*) Al número Φ lo llaman *relación de un extremal y un medio*, puesto que se define por medio de la condición: si $\Phi = A:B$, entonces $A:B = (A+B):A$. Son conocidas también otras denominaciones de este número: *proporción divina*, *proporción justa*. La misma designación Φ procede de la primera letra del nombre del escultor griego Fidias, el cual utilizaba frecuentemente la proporción justa en las proporciones de sus esculturas.

3) $F_{n+m} = F_m \cdot F_{n+1} + F_{m-1} \cdot F_n$ para cualquier m natural.

En una serie de casos puede suceder que la afirmación $A(n)$ para ciertos valores naturales $n < m$ (m es un número natural fijo) sea falsa o no tenga ningún sentido. Por otra parte, en una serie de casos resulta posible demostrar que cierta afirmación $A(n)$ es cierta no sólo para los valores naturales n, sino también para ciertos valores negativos n. En estos casos se usa la *generalización siguiente del principio de la inducción matemática*:

Si la afirmación $A(n)$, en la cual n es un número entero, es cierta para $n = m$ y si de la veracidad de esta afirmación para el número $n = k$, donde k es cualquier número entero mayor o igual a m, se desprende que ella es cierta para el número siguiente $n = k + 1$, entonces la afirmación $A(n)$ será cierta también para cualquier valor entero $n \geqslant m$.

§ 5. Conjuntos con operaciones binarias

5.1. Operaciones binarias en los conjuntos. Para los números enteros están definidas las operaciones de adición y multiplicación. Operaciones análogas se pueden introducir también en conjuntos con elementos de naturaleza arbitraria.

Sea A cierto conjunto finito o infinito con los elementos

$$a, b, c, d, \ldots$$

La aplicación del conjunto A en sí, la cual a cada par de elementos a, b del conjunto A asigna respectivamente cierto tercer elemento (la imagen de los elementos a y b) del mismo conjunto A, se llama *operación binaria* (o simplemente *operación*).

En algunos casos a la operación binaria la llaman *multiplicación*, y a la imagen de los elementos a y b la llaman *producto* de ésta y la designan ab o $a \cdot b$. El producto puede depender del orden de la sucesión de factores, es decir, no es obligatorio que $a \cdot b = b \cdot a$.

En otros casos a la operación binaria la llaman *adición*; a la imagen de elementos a y b la llaman suma y la designan $a + b$. La suma de elementos a y b puede depender también del orden de la sucesión de sumandos, es decir, no es obligatorio que $a + b = b + a$.

El producto de tres elementos a, b, $c \in A$ se puede hacer mediante dos procedimientos distintos:

$$(ab)\,c \quad \text{y} \quad a\,(bc).$$

Si al mismo tiempo $(ab)\,c = a\,(bc)$, se dice que la operación binaria es *asociativa*.

El elemento $e \in A$ es tal que

$$ex = xe = x$$

para cualquier $x \in A$ se llama *elemento unidad* del conjunto A respecto a la operación binaria elegida. (Cuando la operación binaria es una adición, entonces el elemento unidad se designa con el símbolo 0 y se llama *elemento nulo*.)

5.2. Isomorfismo de conjuntos. Sean dados dos conjuntos A y A' y cada uno de ellos está definido por una operación binaria. Los conjuntos A y A' se llaman *isomorfos*, si entre los elementos de estos conjuntos se puede demostrar la aplicación biunívoca f que conserva la operación binaria, es decir:

Si los elementos a' y b' del conjunto A' son las imágenes de los elementos a y b del conjunto A para la aplicación f, entonces $a'b'$ es la imagen del elemento ab.

Por ejemplo, examinemos dos conjuntos de números naturales: el conjunto de todos los números pares y el conjunto de todos los números múltiples al número 5. La suma de dos números pares cualesquiera es un número par, y la suma de dos números cualesquiera múltiplos de 5 es un número múltiplo de cinco. Asignemos al número par $2n$ el número $5n$ ($n \in \mathbf{N}$). Esta correspondencia entre el conjunto de todos los números pares y el conjunto de todos los números múltiplos de 5 es la aplicación biunívoca que posee todas las propiedades enumeradas más arriba y, por consiguiente, los conjuntos dados son isomorfos respecto a la operación de adición.

Los conjuntos isomorfos con las operaciones binarias pueden diferenciarse uno del otro tanto por la naturaleza de sus elementos, como también por el nombre de la operación binaria.

Por ejemplo, examinemos dos conjuntos de números: el conjunto de todos los números reales $\mathbf{R}$ y el conjunto de todos los números reales positivos $\mathbf{R}_+$. Asignemos al número positivo $a \in \mathbf{R}_+$ su logaritmo natural $\ln a$. El conjunto de los logaritmos naturales de todos los números reales positivos forma el conjunto de todos los números reales $\mathbf{R}$. De acuerdo con las propiedades de la función logarítmica, la correspondencia demostrada es la aplicación biunívoca del conjunto de los números reales positivos $\mathbf{R}_+$ sobre el conjunto de todos los números reales $\mathbf{R}$.

La igualdad

$$\ln (a \cdot b) = \ln a + \ln b,$$

donde $a, b \in \mathbf{R}_+$, muestra que el conjunto de números reales positivos con la operación de multiplicación es isomorfo al conjunto de todos los números reales con la operación de adición.

Anotemos que los conjuntos isomorfos son indistinguibles desde el punto de vista de las propiedades de las operaciones: todo lo que puede ser demostrado para un conjunto con cierta operación binaria, sobre la base de las propiedades de esta operación, pero sin la utilización de la naturaleza de los elementos del conjunto, se traslada automáticamente a todos los conjuntos isomorfos. De tal manera el concepto de isomorfismo permite abstraerse de la naturaleza de los elementos de los conjuntos, prestando mayor atención al estudio de las propias operaciones binarias.

5.3. Grupos. Un conjunto no vacío $G = \{a;\ b;\ c;\ \ldots\}$ se llama *grupo*, si en él está definida la operación binaria (la cual suele llamarse *multiplicación*), pues bien

1) la operación binaria en el conjunto G es asociativa, es decir.

$$a(bc) = (ab)c;$$

2) el conjunto G contiene tal elemento e que para cada elemento $a \in G$

$$ea = a;$$

el elemento e se llama *unidad izquierda*;

3) para cada elemento $a \in G$ en el conjunto G existe tal elemento b que

$$ba = e;$$

el elemento b se llama *elemento inverso izquierdo* y se designa a^{-1}:

$$a^{-1}a = e.$$

La operación en el grupo G no debe ser obligatoriamente conmutativa. Si es conmutativa, entonces el grupo G se llama *conmutativo* (o *abeliano*).

De la definición del grupo se deducen sus *propiedades* más simples:

1) Cada grupo G tiene una sola unidad izquierda y una sola unidad derecha y estas unidades son iguales:

$$ea = ae = a$$

2) Cada elemento del grupo G tiene un solo elemento inverso izquierdo y un solo elemento inverso derecho y estos elementos son iguales:

$$a^{-1}a = aa^{-1} = e.$$

De la última propiedad se deducen las reglas llamadas *leyes de abreviación*: para tres elementos cualesquiera a, b, c del grupo G de $ca = cb$, así como también de $ac = bc$, se desprende que $a = b$.

Si el grupo G se compone de un número finito de elementos, entonces este grupo se llama grupo *finito* y la cantidad de elementos en este grupo se llama *orden* del grupo; en caso contrario el grupo se llama *infinito*. El grupo infinito puede ser numerable e innumerable.

Si a la operación binaria de grupo la llaman multiplicación (lo que aceptamos más arriba), entonces el grupo se llama también *multiplicativo*. Si a la operación de grupo la llaman adición, entonces el grupo se llama *aditivo*. En este caso, en vez de la unidad del grupo se habla del *cero* del grupo y lo designan con el símbolo 0; al elemento, inverso al elemento a, lo llaman *opuesto* y lo designan $-a$.

Ejemplos. 1. El conjunto de todos los números enteros forma un grupo de operación de adición, es decir, un grupo aditivo de números enteros. Este grupo es abeliano debido a la conmutatividad de la operación de adición de los números enteros. El número cero desempeña aquí un papel de unidad.

Los conjuntos de números racionales, reales y complejos de la operación de adición también forman grupos aditivos abelianos.

2. El conjunto de todos los números reales positivos de la multiplicación forma un grupo abeliano.

3. En el p. 3.3 ya hemos introducido el concepto de sustituciones de n-ésima potencia, es decir, las aplicaciones del conjunto de n elementos diferentes sobre sí. Todas las sustituciones de n-ésima potencia con la multiplicación definida de sustituciones en calidad de la operación de grupo forman un grupo no conmutativo. Esto es muy fácil de demostrar, verificando el cumplimiento de las propiedades 1)—3) que definen el grupo: el producto de las sustituciones será de nuevo una sustitución: el papel de unidad lo desempeña una sustitución idéntica; para cualquier sustitución de n-ésima potencia existe una sustitución inversa; la multiplicación de las sustituciones es asociativa. En el ejemplo de sustituciones de cuarta potencia podemos convencernos de que la multiplicación de sustituciones no es conmutativa. Sean a y b dos sustituciones de cuarta potencia:

$$a=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix},\quad b=\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix}.$$

Entonces

$$ab=\begin{pmatrix}1&2&3&4\\1&3&4&2\end{pmatrix},\quad ba=\begin{pmatrix}1&2&3&4\\4&2&1&3\end{pmatrix}.$$

Comparando las sustituciones ab y ba, es muy fácil ver que $ab \neq ba$, es decir, la multiplicación de sustituciones no es conmutativa.

El grupo de sustituciones de n-ésima potencia se llama *grupo simétrico* de la potencia n. Es el grupo finito del orden $n!$ y se designa S_n.

5.4. Anillos. El conjunto R se llama *anillo*, si están definidas en él la *adición* y *multiplicación* que satisfacen las condiciones:

1) el conjunto R es un grupo *conmutativo* por adición, es decir,

$$a+b=b+a,\quad a+(b+c)=(a+b)+c,$$
$$a+0=a,\quad a+(-a)=0;$$

2) para cualesquiera $a, b, c \in R$

$$a(bc)=(ab)c;$$

3) la adición y multiplicación se relacionan por medio de las *leyes de distributividad*:

$$a(b+c)=ab+ac,\quad (b+c)a=ba+ca,$$

El conjunto R se llama anillo *conmutativo*, si a estas condiciones le añadimos la de conmutatividad de multiplicación.

El anillo que contiene un elemento e (unidad multiplicativa) en el cual

$$ea=a$$

para todas $a \in R$ se llama *anillo con unidad*. Si R es un anillo con unidad puede suceder que el elemento dado a del anillo R puede tener o no el elemento inverso multiplicativo a^{-1}.

Ejemplos. 1. El conjunto de todos los números enteros, conjunto de todos los números reales, conjunto de todos los números complejos son anillos.

2. El conjunto de números enteros pares es un anillo sin unidad, pero el conjunto de números impares no es un anillo, puesto que la suma de dos números impares es un número par (no cerrado respecto a la operación de adición).

5.5. Campos. Un anillo con unidad, compuesto no sólo de un cero, y el cual para cada elemento a suyo, diferente de cero, contiene también su elemento inverso multiplicativo a^{-1}, se llama *campo*.

Ahora, sabiendo las definiciones de los conceptos de anillo y de campo, aclaremos como desde el punto de vista de estos conceptos se clasifican los conjuntos numéricos más frecuentes.

El conjunto de todos los números naturales con operaciones binarias de adición y de multiplicación definidas en él no es ni un anillo, ni un campo.

El conjunto de todos los números enteros es un anillo conmutativo con unidad.

El conjunto de todos los números racionales, obtenido mediante la extensión del conjunto de números enteros y asociación al conjunto de los cocientes de la división de dos números enteros cualesquiera entre sí (a excepción de la división entre cero), es un campo.

El conjunto de todos los números reales, obtenido mediante la extensión del conjunto de números racionales y asociación al conjunto de elementos nuevos, números irracionales, también es un campo.

El conjunto de números complejos que forman un campo representa de sí una extensión del campo de números reales, obtenida mediante la asociación al campo de los números reales de la raíz de la ecuación

$$x^2 + 1 = 0.$$

Observemos que de todos los campos numéricos, el campo de números racionales es «el más pequeño», puesto que no existen campos numéricos que se distingan del de números racionales y que se encuentren enteramente en este campo, y, además, el campo de los números racionales se halla en cualquier campo númerico.

§ 6. Matrices. Determinantes

La tabla rectangular

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & a_{s3} & \dots & a_{sn} \end{Vmatrix}, \tag{1}$$

compuesta por $n \cdot s$ números, se llama *matriz* de s filas y de n columnas o *matriz de dimensión* $s \times n$, y también $s \times n$ — *matriz*.

Los números a_{ij} ($i = 1, 2, \dots, s$; $j = 1, 2, \dots, n$) se llaman *elementos de la matriz*; el primer índice i del elemento de la matriz indica el número de fila, en la cual se encuentra el elemento, y el segundo índice j indica el número correspondiente de la columna.

La matriz (1) puede designarse también

$$\| a_{ij} \| \quad (i=1, 2, \dots, s; \quad j=1, 2, \dots, n).$$

Además, para las mátrices se utilizan las designaciones

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & a_{s3} & \dots & a_{sn} \end{pmatrix} \quad \text{o} \quad (a_{ij});$$

$$\left|\begin{matrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & a_{s3} & \dots & a_{sn} \end{matrix}\right| \quad \text{o} \quad [a_{ij}].$$

Si una matriz tiene n filas y n columnas, entonces ella se llama *matriz cuadrada de orden* n.

Formemos los productos de todo género por n elementos de esta matriz que se encuentran en distintas filas y en distintas columnas, es decir, los productos del tipo

$$a_{1i_1} a_{2i_2} \dots a_{ni_n}, \tag{2}$$

donde los índices $i_1, i_2, \dots, i_n$ forman cierta permutación de los números $1, 2, 3, \dots, n$. El número de tales permutaciones (y, por consiguiente, de los productos de todo género del tipo (2)) es igual a $n!$

La suma algebraica de los distintos productos de todo género

$$\sum (-1)^p a_{1i_1} a_{2i_2} \dots a_{ni_n},$$

donde $p = 1$, se llama *determinante* de la matriz cuadrada de n-ésimo orden,

$$\left\|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix}\right\|$$

si la sustitución

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ i_1 & i_2 & i_3 & \dots & i_n \end{pmatrix}$$

es impar, y $p = 2$, si la sustitución indicada es par.

El determinante de la matriz cuadrada de n-ésimo orden se denomina

$$\left|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix}\right| \quad \text{o det} \ \| a_{ij} \| \quad (i, j = 1, 2, \dots, n).$$

En particular, el determinante de la matriz cuadrada de segundo orden

$$\left\|\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}\right\|, \quad \text{o} \ \| a_{ij} \| \quad (i, j = 1, 2),$$

es igual a $D = a_{11}a_{22} - a_{12}a_{21}$; el determinante de la matriz cuadrada de tercer orden

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}, \quad \text{o} \ \| a_{ij} \| \quad (i, j = 1, 2, 3),$$

es igual a

$$D = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{32}a_{23}a_{11}.$$

La regla de cálculo del determinante de la matriz de tercer orden se puede representar esquemáticamente de la manera siguiente: en la fig. 1.8 los elementos de una matriz de tercer orden se conectan con

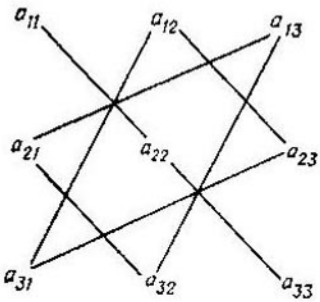


Fig. 1.8.

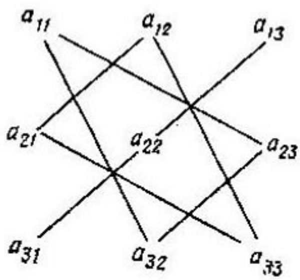


Fig. 1.9.

ayuda de unas rectas, y su producto forma parte del determinante con el signo positivo, y en la fig. 1.9 con ayuda de las rectas se enlazan los elementos de esta matriz, el producto de los cuales forma parte del determinante con el signo negativo.

En conclusión, basándose en los ejemplos sobre los determinantes de tercer orden, formulamos una serie de propiedades muy simples de los determinantes.

1) Un determinante no cambia, si sustituimos sus filas por las columnas con el mismo número y, a la inversa, si cambiamos sus columnas por las filas con el mismo número, por ejemplo;

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

2) Un determinante es igual a cero, si todos los elementos de una de sus filas (o columna) son iguales a cero.

3) Un determinante es igual a cero, si los elementos de dos de sus filas (o columnas) son iguales o proporcionales, por ejemplo:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ ka_{11} & ka_{12} & ka_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0.$$

4) El factor común a todos los elementos de una fila (o columna) se puede sacar del signo del determinante, por ejemplo:

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

5) Si dos filas (o dos columnas) de un determinante son intercambiados de lugar, el determinante cambiará el signo por el opuesto.

6) El determinante no cambia, si añadimos a los elementos de una de sus filas (una de las columnas) los elementos respectivos de otra fila (otra columna) multiplicados por un mismo número, por ejemplo:

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

CAPÍTULO 2

Números reales

El número es el concepto matemático más importante. El surgimiento del concepto de número natural se refiere a la sociedad primitiva y fue acondicionado a la necesidad de calcular la actividad práctica del hombre. Primeramente el concepto de número abstracto no existía, el número fue «atado» a aquellos sujetos que calculaban, y en la lengua de los pueblos primitivos existían unas expresiones verbales para designar los números de los distintos objetos. El concepto abstracto del número natural (es decir, del número no asociado al cálculo de objetos concretos) surgió junto a la aparición de la lengua escrita y a la introducción, para la designación de números, de determinados símbolos.

La aparición de los números fraccionarios (positivos racionales) fue acondicionada a la necesidad de efectuar mediciones, es decir, procesos, en los cuales un valor se compara con otro del mismo género, eligido como patrón de referencia (de unidad de medición). Pero puesto que la unidad de medición no siempre se contenía un número entero de veces en la magnitud que medían, y en muchos casos resultaba imposible despreciar esta circunstancia, entonces surgió la necesidad práctica de introducir números «más pequeños» que los naturales. Esto sirvió de base para el surgimiento de las fracciones más «simples», tales como la mitad, el tercio, el cuarto, etc. El desarrollo ulterior del concepto de número fue acondicionado ya no sólo por la actividad práctica directa del hombre, sino también como consecuencia del desarrollo de las matemáticas.

La introducción de los números negativos fue provocada por el desarrollo del álgebra como ciencia que da los procedimientos generales de solución de los problemas aritméticos, independientemente de su contenido concreto y de sus datos numéricos iniciales. Los números negativos eran utilizados sistemáticamente por los matemáticos de la India ya en los siglos VI—XI. En la ciencia europea los números negativos definitivamente empezaron a utilizarse sólo después de los trabajos de R. Descartes en el siglo XVII el cual dio su representación geométrica.

El conjunto de números racionales es suficiente para satisfacer la mayoría de las necesidades prácticas; mediante los números racionales se pueden realizar mediciones con cualquier grado de precisión.

El desarrollo ulterior del concepto de número se produjo en el siglo XVII en el período del nacimiento de las matemáticas modernas cuando surgió la necesidad de introducir una definición clara del concepto de número. Tal definición fue dada por uno de los fundadores del análisis matemático, I. Newton, en «La aritmética universal»: «Por número entendemos no tanto un conjunto de unidades, como la relación abstracta entre una magnitud cualquiera y otra del mismo género tomada por nosotros como unidad». Esta formulación da una definición única del número real, tanto racional, como irracional. (Los sabios de la Grecia Antigua ya sabían de la existencia de los segmentos inconmensurables, la relación de los cuales es un número irracional.) Más tarde en los años 70 del siglo XIX, fue desarrollada una teoría rigurosa del número real en los trabajos de R. Dedekind, G. Cantor y K. Weierstrass.

§ 1. Números naturales

1.1. Conjunto de números naturales. Los *números naturales* son los que se utilizan para el cálculo:

$$1, 2, 3, 4, \ldots, n, \ldots \qquad (1)$$

De dos números vecinos cualesquiera en la inscripción (1) el número que se encuentra a la derecha se llama *sucesivo* respecto al que se encuentra a la izquierda.

Los números naturales (1) forman un conjunto que se llama *conjunto de números naturales.* El conjunto de todos los números naturales se designa con el símbolo **N**:

$$\mathbf{N} = \{1;\ 2;\ 3;\ \ldots;\ n;\ \ldots\}.$$

El conjunto de números naturales es un conjunto ordenado, es decir, para dos números naturales cualesquiera m y n tiene lugar una de las relaciones siguientes:

ó $m = n$ (m es igual a n),

ó $m < n$ (m es menor que n),

ó $n < m$ (n es menor que m).

El número natural más pequeño es 1 (unidad).

En el conjunto de números naturales se introducen dos operaciones aritméticas principales: la adición y la multiplicación. Para la designación de estas operaciones se utilizan respectivamente los símbolos $+$ y $\cdot$ ó $\times$.

Adición de números naturales. A cada par de números naturales $(n;\ p)$ se le asigna el número natural s que se llama *suma* de éstos. La suma s se compone de tantas unidades cuantas contienen los números n y p. Del número s se dice que se obtuvo como resultado de la adición de los números n y p, y se escribe

$$s = n + p. \tag{2}$$

Los números n y p en la expresión (2) se llaman *sumandos*. La operación de adición de los números naturales se somete a las reglas de

1) *conmutatividad*: $n + p = p + n$;

2) *asociatividad*: $(n + p) + k = n + (p + k)$.

Multiplicación de números naturales. A cada par ordenado de números naturales $(n;\ p)$ se le asigna el número natural m que se llama *producto* de éstos. El producto m se compone de tantas unidades, cuantas contiene el número n, tomadas tantas veces, cuantas unidades contiene el número p. Del número m se dice que es el resultado de la multiplicación de los números n y p, y se escribe

$$m = n\cdot p \text{ ó } m = n \times p. \tag{3}$$

Los números n y p en la expresión (3) se llaman *factores*

La multiplicación de los números naturales se somete a las reglas siguientes:

1) *conmutatividad*: $n\cdot p = p\cdot n$;

2) *asociatividad*: $(n\cdot p)\cdot k = n\cdot(p\cdot k)$.

Las operaciones de adición y multiplicación de los números naturales se relacionan con la *ley de la distributividad de multiplicación respecto a la adición*:

$$(n+p)\cdot k = n\cdot k + p\cdot k.$$

De esta manera, la suma y el producto de dos números naturales cualesquiera son dos números naturales. Por eso se dice que el conjunto de todos los números naturales está cerrado respecto a las operaciones de adición y multiplicación.

Sustracción de números naturales. La sustracción de números naturales es la operación inversa a la adición, es decir, la correspondencia que a un par de números naturales $(n;\ p)$ le asigna tal número natural r que

$$n = p + r.$$

Del número r se dice que se ha obtenido como resultado de la sustracción de p del número n, y se escribe

$$r = n - p.$$

El número r se llama *diferencia* de los números n y p; el número n en la expresión se llama *minuendo*, y el número p se llama *substraendo*.

En el conjunto de números naturales la diferencia de dos números naturales $r = n - p$ existe cuando y sólo cuando $n > p$; por eso se dice que el conjunto de números naturales no está cerrado respecto a la sustracción.

Ejemplo. El número natural 5 es más que el número natural 3. Su diferencia existe y es igual al número natural 2:

$$5 - 3 = 2.$$

El número natural 6 es menos que el número natural 8. Su diferencia 6—8 ya no es el número natural.

División de números naturales. La división de números naturales es la operación inversa a la multiplicación, es decir, la correspondencia que a un par ordenado de números naturales $(n;\ p)$ le asigna tal número natural q que

$$n = p\cdot q.$$

Del número q se dice que se ha obtenido como resultado de la división del número n entre el número p y se escribe

$$q=\frac{n}{p}, \quad \text{o } q=n/p, \quad \text{o } q=n:p.$$

El número q se llama *cociente* de los números naturales n y p; el número n se llama *dividendo* y el número p se llama *divisor*.

En el conjunto de números naturales el cociente está definido no para cualquier par de números naturales $(n;\ p)$, es decir, el conjunto de números naturales no está cerrado respecto a la operación de división. Por ejemplo, supongamos que $n = 7$ y $p = 2$. Para este par de números naturales es imposible elegir tal número natural q, que se cumpla la igualdad

$$7 = 2\cdot q.$$

Potencia natural de un número. La propiedad de asociatividad de la multiplicación de los números naturales permite introducir el concepto de potencia natural de un número natural: la n-ésima potencia del número natural m es el número natural k obtenido como resultado de la multiplicación del número m por sí n veces:

$$\underbrace{k=m\cdot m\cdot m\cdot\ \ldots\ \cdot m}_{n\ \text{factores}}.$$

En general se utiliza la expresión más corta:

$$k = m^n.$$

El número m se llama *base de la potencia* y el número n se llama *exponente de la potencia*.

1.2. Construcción axiomática de un conjunto de números naturales. Ya hemos citado más arriba ciertas propiedades de los números naturales y hemos expuesto ya las operaciones con números naturales que se someten a las reglas de la conmutatividad, asociatividad y distributividad.

Es muy fácil observar que, expresando todos estos conocimientos sobre los números naturales y las operaciones con ellos, nos dirigíamos hacia la comprensión intuitiva de muchos de los conceptos que utilizamos, porque utilizamos estos conceptos en la práctica diaria y obtenemos así resultados correctos. Tal, por ejemplo, parece ser muy natural que $3 + 2 = 2 + 3$ y no surge ninguna duda acerca de la procedencia de esta propiedad de la adición de números naturales. En las matemáticas, naturalmente, surge la pregunta acerca de cuántas y cuáles precisamente afirmaciones primarias (axiomas) sobre los números naturales es necesario usar para que de estos axiomas en forma de teoremas se puedan obtener todas las propiedades de los números naturales y operaciones con ellos que sabemos de nuestra experiencia práctica.

Resulta ser que todas las propiedades de los números naturales pueden ser introducidas como teoremas de cinco axiomas y fórmulas, los cuales definen las operaciones de adición y multiplicación de los números naturales.

Axiomas de los números naturales (*axiomas de Peano*):

I. 1 (unidad) es un número natural.

II. Para cada número natural n existe precisamente un número natural que se llama *sucesivo* de aquél y que se designa $S(n)$.

III. Siempre $S(n) \neq 1$.

IV. De la igualdad $S(n) = S(m)$ se deduce que $n = m$.

V. *Principio de inducción completa.* El conjunto de números naturales, que contiene 1 y para cada uno de n elementos, el elemento sucesivo $S(n)$, contiene todos los números naturales.

La adición y multiplicación de números naturales se define por las fórmulas

$$n + 1 = S(n),$$

$$m + S(n) = S(m + n),$$

$$n \cdot 1 = n,$$

$$n \cdot S(m) = n \cdot m + n.$$

De los axiomas de Peano y de la definición de las operaciones de adición y multiplicación de los números naturales como teorema se deducen las leyes de conmutatividad y asociatividad de la adición y multiplicación, la propiedad de distributividad de la multiplicación respecto a la adición.

De los axiomas de Peano y de la definición de la operación de adición también se deduce la *propiedad de ordenación del conjunto de los números naturales*: para dos números naturales cualesquiera m y n:

ó $m = n$ (m es igual a n);

ó $m < n$ (m es menor que n);

ó $n < m$ (n es menor que m).

La ordenación del conjunto de números naturales se demuestra de la manera siguiente. En primer lugar, como teorema se demuestra que para m y n naturales cualesquiera tiene lugar uno de los tres casos siguientes:

ó $m = n$;

o existe un número natural k único que satisface la condición

$$n = m + k;$$

o existe un número natural p único tal que

$$m = n + p.$$

En segundo lugar, se introduce la definición del signo $>$ (mayor) y del signo $<$ (menor).

Al número natural m lo consideran mayor que el número natural n (y se escribe: $m > n$), si existe tal número natural k para el cual

$$m = n + k.$$

Al número natural m lo consideran menor que el número natural n (y se escribe: $m < n$), si existe tal número natural p para el cual

$$m + p = n.$$

De las definiciones dadas y del teorema formulado más arriba se desprende la propiedad de ordenación del conjunto de los números naturales.

Puesto que el conjunto de números naturales es un conjunto ordenado, entonces para los números naturales resulta válida una serie de afirmaciones que se relacionan con los conceptos «mayor» y «menor». A tales afirmaciones pertenecen, por ejemplo, los teoremas siguientes:

1) De $m > n$, para cualquier k se deduce que $m + k > n + k$;
de $m = n$, para cualquier k se deduce que $m + k = n + k$;
de $m < n$, para cualquier k se deduce que $m + k < n + k$.

2) De $m > n$, para cualquier k se deduce que $m \cdot k > n \cdot k$;
de $m = n$, para cualquier k se deduce que $m \cdot k = n \cdot k$;
de $m < n$, para cualquier k se deduce que $m \cdot k < n \cdot k$.

3) En cada conjunto no vacío de números naturales existe un número más pequeño.

1.3. Números primos. Teorema principal de la aritmética. Si para unos números naturales dados n y p se halla un número natural q tal, que $n = p \cdot q$, entonces se dice que el número n se *divide enteramente* entre el número p. El número p se llama *divisor* del número n, y sobre el número natural n se dice que es *múltiplo* de p. Si el número natural p es el divisor del número natural n, entonces el número natural q también será divisor del número n.

El número natural, cuyos únicos divisores son él mismo y la unidad, se llama número *primo*. Todos los demás números naturales se llaman *compuestos*. El número natural 1 no se considera número primo.

La presentación del número natural n en forma del producto de dos números naturales $p \cdot q$ se llama *descomposición en factores*. Al mismo tiempo se considera que si el número n es primo, entonces tiene una descomposición en factores que consta sólo del número n. Por ejemplo, el número primo 37 tiene una descomposición en factores que consta sólo del factor 37 (y no la descomposición $1 \cdot 37$).

Sea que el número natural n es compuesto, es decir,

$$n = p \cdot q,$$

donde $p \neq 1$ y $q \neq 1$. Entonces son posibles los casos siguientes:

1) Si los números naturales p y q son primos, el número n se representa como el producto de dos números primos p y q.

2) Si, por lo menos, uno de los números naturales p y q es compuesto, entonces el número compuesto (o los dos números compuestos p y q) lo descomponen en un producto de números naturales aún menores, para los cuales son posibles las mismas variantes. Puesto que existe sólo un conjunto finito de números naturales que son menores que n, entonces el proceso indicado de descomposicion termina dentro de un número finito de pasos. De resultas obtenemos la descomposición del número n en factores, cada uno de los cuales es un número primo. La presentación del número n en forma del producto de los números primos se llama *descomposición en factores primos.*

Debido a la descomposición examinada del número natural en factores primos surge la pregunta si existe o no cualquier otra descomposición, distinta de la obtenida, del número natural en factores primos. El teorema llamado *teorema principal de la aritmética* nos da la respuesta:

Cada número natural, distinto de 1, puede ser descompuesto en factores primos, y al mismo tiempo, mediante un único procedimiento (si vamos a identificar las descomposiciones $p \cdot q$ y $q \cdot p$, donde p y q son los números primos). Uniendo en la descomposición del número n los factores primos idénticos $p_1, p_2, \ldots, p_s$, obtenemos la así llamada *descomposición canónica* del número n:

$$n = p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s}$$

($k_1, k_2, \ldots, k_s$ — son números naturales).

Un número natural se llama *par*, si entre sus factores primos existe el número 2, y es *impar*, si entre sus factores primos el número 2 no existe.

1.4. Ciertos criterios de divisibilidad de los números naturales.

Criterio de divisibilidad entre 2. Un número se divide entre 2, si su última cifra es un número par o cero.

Criterio de divisibilidad entre 4.. Un número se divide entre 4, si sus dos últimas cifras son ceros o forman un número que se divide entre 4.

Criterio de divisibilidad entre 8. Un número se divide entre 8, si sus tres últimas cifras son ceros o forman un número que se divide entre 8.

Criterios de divisibilidad entre 3 y entre 9. Un número se divide entre 3, si la suma de cifras del número se divide entre 3. Un número se divide entre 9, si la suma de sus cifras se divide entre 9.

Criterio de divisibilidad entre 5. Un número se divide entre 5, si termina ó en cero ó en 5.

Criterios de divisibilidad entre 25. Un número se divide entre 25, si sus dos últimas cifras son ceros o forman un número que se divide entre 25.

Criterio de divisibilidad entre 11. Un número se divide entre 11, si la suma de sus cifras que ocupan los lugares pares es igual a la suma de las cifras que ocupan los lugares impares, o se distingue de la misma en un número que se divide en 11.

1.5. Mínimo común múltiplo. Máximo común divisor. Algoritmo de Euclides. Un número natural que es múltiplo de cada uno de una serie de números naturales se llama *común múltiplo* de los mismos. Si existen varios múltiplos, entonces el mínimo de ellos se llama *mínimo común múltiplo* (en forma abreviada: MCM)*)

Para hallar el mínimo común múltiplo de varios números es necesario:

1) copiar las descomposiciones canónicas de los números dados;

2) enumerar todos los factores primos que entran, por lo menos, en una de las descomposiciones canónicas de los números dados;

3) elevar cada uno de los factores primos enumerados a la máxima potencia, con la cual este factor primo entra en la descomposición canónica de los números dados.

---

*) El mínimo común múltiplo de dos números m y n se designa también mediante llaves: $\{m, n\}$.

El producto de las potencias obtenidas de los factores primos da el número que es el mínimo común múltiplo de los números dados.

Ejemplo 1. Hallar el MCM de los números 49 896 y 26 460.

1) Escribimos las descomposiciones canónicas de los números dados

$$49\,896 = 2^3\cdot 3^4\cdot 7\cdot 11, \quad 26\,460 = 2^2\cdot 3^3\cdot 5\cdot 7^2;$$

2) copiamos los factores primos que entran, por lo menos, en una descomposición canónica:

$$2;\ 3;\ 5;\ 7;\ 11;$$

3) la potencia máxima, con la cual el factor 2 entra en las descomposiciones canónicas de los números dados, es igual a 3; escribimos el número 2^3.

La potencia máxima, con la cual el factor 3 entra en las descomposiciones canónicas, es igual a 4; escribimos el número 3^4.

Análogamente, las potencias máximas, con las cuales los factores 5, 7 y 11 entran en las descomposiciones canónicas de los números dados, son respectivamente iguales a 1, 2 y 1; escribimos los números $5^1 = 5$, 7^2 y $11^1 = 11$.

El producto $2^3\cdot 3^4\cdot 5\cdot 7^2\cdot 11 = 1\,746\,360$ es el mínimo común múltiplo de los números dados.

Un número que es divisor de cada uno de los números dados se llama *divisor común* de varios números naturales. En caso de que existan varios divisores comunes, el máximo de ellos se llama *máximo común divisor* (en forma abreviada: MCD)*).

Si el máximo común divisor de varios números naturales es igual a la unidad, entonces estos números se llaman *recíprocamente simples*.

Para hallar el máximo común divisor de varios números es necesario:

1) copiar las descomposiciones canónicas de los números dados;

2) enumerar todos los factores comunes primos que entran en las descomposiciones canónicas de cada uno de los números dados;

---

*) El máximo común divisor de dos números m y n se designa también mediante paréntesis: (m, n).

3) elevar cada uno de los factores primos enumerados a la potencia mínima, con la cual este factor primo entra en las descomposiciones canónicas de los números dados.

El producto de las potencias obtenidas de los factores primos da un número que es el máximo común divisor de los números dados.

Ejemplo 2. Hallar el MCD de los números 49 896 y 26 460.

1) Escribimos las descomposiciones canónicas de los números dados:

$$49\,896 = 2^3 \cdot 3^4 \cdot 7 \cdot 11 \text{ y } 26\,460 = 2^2 \cdot 3^3 \cdot 5 \cdot 7^2$$

2) copiamos los factores primos que entran en ambas descomposiciones canónicas.

$$2;\ 3;\ 7;$$

3) la potencia más pequeña, con la cual el factor 2 entra en las descomposiciones canónicas de los números dados, es igual a 2; escribimos el número 2^2.

La potencia más pequeña, con la cual el factor 3 entra en las descomposiciones canónicas, es igual a 3; escribimos el número 3^3.

La potencia más pequeña, con la cual el número 7 entra en las descomposiciones canónicas, es igual a 1; escribimos el número $7^1 = 7$.

El producto $2^2 \cdot 3^3 \cdot 7 = 756$ es el máximo común divisor de los dos números naturales dados.

Propiedad del MCD. El máximo común divisor de dos números se divide entre cualesquiera de los comunes divisores de estos números.

El máximo común divisor (m, n) de los números m y n, y el mínimo común múltiplo $\{m, n\}$ se relacionan mediante la igualdad

$$m \cdot n = (m, n) \cdot \{m, n\}.$$

Algoritmo de Euclides para la obtención del MCD. Hallar el máximo común divisor de dos números naturales suficientemente grandes es bastante difícil. Sin embargo, existe un procedimiento para la obtención del MCD que no requiere el conocimiento de todos los factores primos de estos números. Este procedimiento se llama *algoritmo de Euclides.*

Antes de pasar a la demostración del algoritmo de Euclides, introduzcamos la operación de división con resto.

Dividir un número natural m entre un número natural n $(m \geqslant n)$ *con resto*, significa obtener tal número natural k y tal número entero no negativo $r < n$, que

$$m = n \cdot k + r.$$

Los números k y r se llaman respectivamente *cociente* y *resto* de la división del número m entre el número n.

En caso de que $r = 0$, también se dice que el resto de la división es igual a cero o que el número m se divide entre el número n.

Pasamos ahora a la demostración del algoritmo de Euclides. Sean m y n dos números naturales y sea $m > n$. Designemos mediante m_1 y r_1 respectivamente el cociente y el resto de la división de m entre n:

$$m = n \cdot m_1 + r_1, \quad 0 \leqslant r_1 < n. \tag{4}$$

Si resulta que $r_1 > 0$, entonces dividamos n entre r_1, designando mediante m_2 y r_2 respectivamente el cociente y el resto de la división:

$$n = r_1 \cdot m_2 + r_2, \quad 0 \leqslant r_2 < r_1. \tag{5}$$

Si r_2 aún no es cero, dividamos r_1 entre r_2 y obtenemos análogamente

$$r_1 = r_2 \cdot m_3 + r_3, \quad 0 \leqslant r_3 < r_2. \tag{6}$$

Puesto que $n, r_1, r_2, \ldots$ son números naturales y $n > r_1 > > r_2 > r_3 > \ldots$, entonces el proceso de división después de un número finito de pasos debe cesarse, es decir, debemos obtener un resto igual a cero. Supongamos que tal resto sea $r_{n+1} = 0$, puesto que

$$r_{n-1} = r_n \cdot m_{n+1}. \tag{*}$$

El número natural r_n es el máximo común divisor de los números m y n. Para convencerse de esto, mostremos que m y n se dividen entre r_n. En vigor de la igualdad (*) r_{n-1} se divide entre r_n. Entonces de la igualdad

$$r_{n-2} = r_{n-1} \cdot m_n + r_n,$$

que es anterior a la igualdad (*), se deduce que r_{n-2} se divide entre r_n.

Siguiendo la cadena de las igualdades construidas (desde abajo hacia arriba), podemos convencernos de que entre r_n se dividen también $r_{n-3}, r_{n-4}, \ldots, r_2; r_1$, y de resultas, en vigor de las igualdades (5) y (4), los números m y n. Queda convencernos sólo de que cualquier divisor común k de dos números m y n es también el divisor del número r_n (por lo que evidentemente será demostrado que k es el máximo común divisor de los números m y n). Para esto tendremos que pasar de nuevo toda la cadena de igualdades, pero esta vez desde arriba hacia abajo. Primero reescribiremos la cadena de las igualdades del algoritmo de Euclides en forma de

$$\begin{aligned}
r_1 &= m - nm_1, \\
r_2 &= n - r_1 m_2, \\
r_3 &= r_1 - r_2 m_3, \\
&\cdots\cdots\cdots\cdots \\
r_n &= r_{n-2} - r_{n-1} m_n.
\end{aligned}$$

De la primera igualdad se deduce que r_1 se divide entre el común divisor k de los números m y n. Pero de la segunda igualdad se desprende que puesto que r_1 se divide entre k y n se divide entre k, entonces r_2 también se divide entre k. Siguiendo adelante este proceso, obtendremos que también r_{n-2}, y r_{n-1} se dividen entre k. Pero entonces, en vigor de la última igualdad, r_n se divide también entre k.

El procedimiento de obtención del máximo común divisor de dos números mediante el algoritmo de Euclides en la mayoría de los casos resulta el más corto y por eso es el más ventajoso en la práctica.

Podemos observar que en el razonamiento citado más arriba se demuestra, al mismo tiempo, la propiedad del máximo común divisor.

Ejemplo 3. Mediante el algoritmo de Euclides hallar el máximo común divisor de los números 13 172 y 261.

Dividimos el número 13 172 entre 261:

$$\begin{array}{r|l} 13172 & 261 \\ -\underline{1305} & \overline{50} \\ 122 & \end{array}$$

Según el algoritmo de Euclides dividimos el número 261 entre el resto de la división (el número 122):

$$\begin{array}{r|l} 261 & 122 \\ -\underline{244} & \overline{2} \\ 17 & \end{array}$$

Dividimos el número 122 entre el resto (el número 17):

$$\begin{array}{r|l} 122 & 17 \\ -\underline{119} & \overline{7} \\ 3 & \end{array}$$

Continuando dividiendo sucesivamente cada resto anterior entre cada resto posterior, de resultas obtendremos un resto igual a cero:

$$\begin{array}{r|l} 17 & 3 \\ -\underline{15} & \overline{5} \\ 2 & \end{array} \qquad \begin{array}{r|l} 3 & 2 \\ -\underline{2} & \overline{1} \\ 1 & \end{array} \qquad \begin{array}{r|l} 1 & 1 \\ -\underline{1} & \overline{1} \\ 0 & \end{array}$$

El penúltimo resto es igual a la unidad. Por consiguiente, la unidad es el máximo común divisor de los dos números dados, es decir, los números 13 172 y 261 son recíprocamente primos.

§ 2. Números enteros

2.1. Conjunto de números enteros. El conjunto de números enteros es un conjunto obtenido como resultado de la adición al conjunto de todos los números naturales de los objetos nuevos (a los cuales en lo posterior les vamos a designar los números) del números cero y de los números negativos enteros.

El número *cero* que se designa por medio del símbolo 0 y los *números negativos enteros* se introducen del modo siguiente.

La suma de cualquier número natural n y el número 0 da el número n:

$$n + 0 = n.$$

A cualquier número natural n le corresponde el único número entero negativo n tal que la suma de los números n y $-n$ resulta igual a cero:

$$n + (-n) = 0.$$

El número $-n$ se llama *opuesto* al número n. El número opuesto al número $-n$ es el número n: $-(-n) = n$. Los números naturales en el conjunto de números enteros se llaman *números enteros positivos**). El conjunto de todos los números enteros frecuentemente se designa por **Z**.

El conjunto de los números naturales **N** está cerrado respecto a la adición y multiplicación, pero no está cerrado respecto a la sustracción: la suma y el producto de dos números naturales cualesquiera es un número natural, mientras que la diferencia $n - p$ de dos números naturales en el conjunto de números naturales se define sólo cuando $n > p$. El conjunto de números enteros se obtiene como una extensión del conjunto de números naturales hasta el conjunto **Z**, en el cual

1) el conjunto **N** es un subconjunto propio;

2) la adición y multiplicación de números naturales en **Z** coincide con las operaciones idénticas en **N**;

3) la sustracción en **Z** siempre es posible, es decir, la diferencia de dos elementos cualesquiera de números de **Z** es un elemento (número) de **Z**;

4) el conjunto extendido **Z** es mínimo en el sentido que no posee subconjunto propio que satisfaga las condiciones 1)—3).

El conjunto de números enteros es un conjunto ordenado, es decir, para dos números enteros cualesquiera m y n es válida una y sólo una relación de las siguientes:

ó $m = n$,

ó $m < n$,

ó $n < m$.

*) A continuación denominaremos los números enteros con letras latinas minúsculas m, n, k, ...

Para los números positivos n se escribe $n > 0$, para los números negativos se escribe $n < 0$. Si se quiere indicar que el número puede ser positivo o nulo, se escribe $n \geqslant 0$ y se dice que es no *negativo*; análogamente la inscripción $n \leqslant 0$ significa que n es negativo o es igual a cero.

2.2. Operaciones matemáticas con números enteros. El número $|n|$ que se calcula según la regla siguiente

$$|n| = \begin{cases} n, & \text{si } n > 0, \\ 0, & \text{si } n = 0, \\ -n, & \text{si } n < 0. \end{cases}$$

se llama *valor absoluto* (o *módulo*) del número n.

El valor absoluto del número n es positivo para n positivos y para n negativos y es igual a cero sólo para $n = 0$.

Ejemplo 1. Hallar los valores absolutos de números 4 y -3.

Puesto que el número 4 es natural, entonces $|4| = 4$. Puesto que -3 es un número negativo, entonces según la regla de cálculo del valor absoluto del número $|-3| = -(-3) = 3$.

Adición de números enteros. Un número entero s que se calcula según la regla:

si $n > 0$, y $p > 0$, entonces $s = n + p$;

si $n < 0$ y $p < 0$, entonces $s = -(|n| + |p|)$;

si $n > 0$, y $p < 0$ y $|n| > |p|$, entonces $s = |n| - |p|$;

si $n > 0$, y $p < 0$ y $|n| = |p|$, entonces $s = 0$;

si $n > 0$, y $p < 0$ y $|n| < |p|$, entonces $s = -(|p| - |n|)$;

si $n < 0$, $p > 0$ y $|n| > |p|$, entonces $s = -(|n| - |p|)$;

si $n < 0$, $p > 0$ y $|n| = |p|$, entonces $s = 0$;

si $n < 0$, $p > 0$ y $|n| < |p|$, entonces $s = |p| - |n|$;

si $n = 0$, entonces $s = p$;

si $p = 0$, entonces $s = n$, se llama *suma* de los números enteros n y p.

La suma s de los números enteros n y p se escribe mediante el símbolo $+$:

$$s = n + p.$$

La regla de cálculo de la suma de dos números enteros, citada más arriba, se basa en el cálculo de la suma o de diferencia de dos números naturales.

Ejemplo 2. Calcular la suma de los números enteros 6 y -8.

El par de números dado satisface la quinta condición de la regla de adición de dos números enteros ($6 > 0$, $-8 < 0$, $|6| < |-8|$). La suma de estos números es igual a

$$s = -(|-8| - |6|) = -(8-6) = -2.$$

La adición de números enteros, así como también la de números naturales, posee las propiedades de conmutatividad y de asociatividad.

Multiplicación de números enteros. Un número entero m *que se calcula según la regla:*

si $n > 0$ y $p > 0$, entonces $m = n \cdot p$;

si $n < 0$ y $p < 0$, entonces $m = |n| \cdot |p|$;

si $n < 0$, y $p > 0$ o si $n > 0$, y $p < 0$, entonces $m = -(|n| \cdot |p|)$;

si $n = 0$ ó $p = 0$, entonces $m = 0$,

se llama *producto* de los números enteros n y p.

El producto de los números enteros n y p se escribe mediante el símbolo $\cdot$ ó $\times$:

$$m = n \cdot p \quad (\text{ó } n \times p).$$

El cálculo del producto de dos números enteros se basa en el de dos números naturales.

Ejemplo 3. Calcular el producto de los números enteros -2 y -7.

El par dado satisface la segunda condición de la regla de multiplicación de dos números enteros. Los valores absolutos de estos números son iguales a 2 y 7, y su producto es igual a 14.

La multiplicación de números enteros, así como también la de números naturales, posee las propiedades de conmutatividad y de asociatividad.

Además, las operaciones de adición y multiplicación de números enteros, como también en el caso de números naturales, están relacionadas por la ley de distributividad de la multiplicación respecto a la adición, *Sustracción de números enteros.* Un número entero r que se calcula según la regla:

$$r = n + (-p), \tag{1}$$

se llama *diferencia* de los números enteros n y p, es decir, la diferencia de los números enteros n y p es la suma del número entero n y el número opuesto al número p. Por consiguiente, la diferencia se calcula según la regla de cálculo de la suma de dos números enteros.

Del número r se dice que ha sido obtenido como resultado de la sustracción del número p a partir del número n y se escribe

$$r = n - p. \tag{2}$$

En la expresión (2) el número n se llama *minuendo*, y el número p se llama *sustraendo*.

Ejemplo 4. Calcular la diferencia de los números enteros 2 y —7. El número opuesto al número —7 es igual a 7, y por eso según la regla (1) la diferencia de estos números es $r = 2 + 7 = 9$.

El conjunto de números enteros está cerrado respecto a las operaciones de adición, multiplicación y sustracción, es decir, para dos números enteros dados cualesquiera existe un solo número entero, el tercero que representa la suma de los dos números enteros dados; existe un solo número entero que es su diferencia y, por fin, un solo número entero que es su producto.

División de números enteros. Un número entero p que satisface la igualdad

$$m = n \cdot p, \tag{3}$$

se llama *cociente* de la división del número entero m entre el número entero n.

La división entre el número cero está prohibida. Del número p se dice que ha sido obtenido como resultado de la división del número m entre el número n y se escribe

$$p = m : n, \text{ ó } p = \frac{m}{n}, \text{ ó } p = m/n.$$

En el conjunto de números enteros, así como en el conjunto de los números naturales, la división no siempre es posible, porque no para cualquier par de números enteros m y n existe un cociente. Por eso se dice que el conjunto de números enteros no está cerrado respecto a la operación de división. Sin embargo, entre las operaciones de división en el conjunto de números naturales y en el conjunto de números enteros existe una diferencia esencial: si en un conjunto de números naturales existía el cociente de dos números naturales, entonces este cociente era el único. En el conjunto de números enteros sucede otro proceso: sea m un número entero arbitrario y $n = 0$; entonces la igualdad (3) obtiene la forma

$$m = 0 \cdot p. \qquad (4)$$

Tratemos de hallar tal número p que satisfaga la igualdad (4). Existen dos posibilidades:

si $m \neq 0$, entonces no existe tal número entero p, para el cual se cumple esta igualdad;

si $m = 0$, entonces p puede ser cualquier número entero.

De esta manera, el cociente de la división de un número entero entre cero o no existe, o se define por más de una forma. Para evitar tal indeterminación, es necesario prohibir la división de un número entero entre cero.

§ 3. Números racionales

3.1. Fracciones racionales. Un par ordenado de números enteros $(m;\ n)$, donde el número n se distingue de cero, se llama *fracción racional*. La fracción racional se designa por medio del símbolo $\frac{m}{n}$ ó m/n. El número m se llama *numerador* de la fracción, y el número n se llama *denominador* de la misma. Dos fracciones racionales $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$ son equivalentes, si $m_1 \cdot n_2 = n_1 \cdot m_2$, y se escribe $\frac{m_1}{n_1} \sim \frac{m_2}{n_2}$.

1) Cualquier fracción racional es equivalente a sí misma:

$$\frac{m_1}{n_1} \sim \frac{m_1}{n_1} \quad \text{(reflexividad).}$$

2) Si la fracción racional $\frac{m_1}{n_1}$ es equivalente a la fracción racional $\frac{m_2}{n_2}$, entonces la fracción racional $\frac{m_2}{n_2}$ es también equivalente a la fracción $\frac{m_1}{n_1}$ *):

$$\frac{m_1}{n_1} \sim \frac{m_2}{n_2} \Longleftrightarrow \frac{m_2}{n_2} \sim \frac{m_1}{n_1} \quad \text{(simetría).}$$

---

*) Los símbolos $\Rightarrow$ y $\Longleftrightarrow$ significan «sigue» y «sigue a ambas partes» respectivamente.

3) Si la fracción racional $\frac{m_1}{n_1}$ es equivalente a la fracción racional $\frac{m_2}{n_2}$, y la fracción racional $\frac{m_2}{n_2}$ es equivalente a la fracción $\frac{m_3}{n_3}$, entonces y la fracción $\frac{m_1}{n_1}$ es equivalente a la fracción $\frac{m_3}{n_3}$:

$$\frac{m_1}{n_1} \sim \frac{m_2}{n_2} \text{ y } \frac{m_2}{n_2} \sim \frac{m_3}{n_3} \Rightarrow \frac{m_1}{n_1} \sim \frac{m_3}{n_3} \quad \text{(transitividad).}$$

De la definición de la equivalencia de fracciones se deduce que la fracción racional $\frac{m_1}{n_1}$ es equivalente a la fracción $\frac{m_1 \cdot k}{n_1 \cdot k}$:

$$\frac{m_1}{n_1} \sim \frac{m_1 \cdot k}{n_1 \cdot k},$$

donde k es cualquier número entero, distinto de cero. El paso de la fracción $\frac{m_1 \cdot k}{n_1 \cdot k}$ a la fracción equivalente $\frac{m_1}{n_1}$ se llama *simplificación* de la fracción en el número k. Esta propiedad de la equivalencia de fracciones racionales permite dar otra definición de la fracción racional.

Un par de números $(m; n)$ que se designa con el símbolo $\frac{m}{n}$, donde m es un número entero, y n es un número natural se llama *fracción racional*. En este párrafo vamos a utilizar precisamente esta definición, es decir, considerar n como un número natural.

La fracción racional $\frac{m_1}{n_1}$ se considera *mayor* que la fracción racional $\frac{m_2}{n_2}$:

$$\frac{m_1}{n_1} > \frac{m_2}{n_2},$$

si $m_1 \cdot n_2 > n_1 \cdot m_2$.

La fracción racional $\frac{m_1}{n_1}$ se considera *menor* que la fracción racional $\frac{m_2}{n_2}$:

$$\frac{m_1}{n_1} < \frac{m_2}{n_2},$$

si $m_1 \cdot n_2 < n_1 \cdot m_2$.

Un conjunto de fracciones racionales es un conjunto ordenado para las dos fracciones cualesquiera $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$

$$\text{o} \quad \frac{m_1}{n_1} \sim \frac{m_2}{n_2}, \quad \text{o} \quad \frac{m_1}{n_1} > \frac{m_2}{n_2}, \quad \text{o} \quad \frac{m_1}{n_1} < \frac{m_2}{n_2}.$$

Al mismo tiempo se cumplen las condiciones siguientes:

1) Cualquier fracción racional es menor o equivalente a sí misma:

$$\frac{m_1}{n_1} \lesssim \frac{m_1}{n_1} \quad \text{(reflexividad).}$$

2) Si la fracción $\frac{m_1}{n_1}$ es menor o equivalente a la fracción $\frac{m_2}{n_2}$, y la fracción $\frac{m_2}{n_2}$ es menor o equivalente a la fracción $\frac{m_1}{n_1}$, entonces estas fracciones son equivalentes:

$$\frac{m_1}{n_1} \sim \frac{m_2}{n_2} \quad \text{(antisimétrica).}$$

3) Si la fracción $\frac{m_1}{n_1}$ es menor o equivalente a la fracción $\frac{m_2}{n_2}$, y la fracción $\frac{m_2}{n_2}$ es menor o equivalente a la fracción $\frac{m_3}{n_3}$, entonces la fracción $\frac{m_1}{n_1}$ es menor o equivalente a la fracción $\frac{m_3}{n_3}$:

$$\frac{m_1}{n_1} \lesssim \frac{m_2}{n_2} \quad \text{y} \quad \frac{m_2}{n_2} \lesssim \frac{m_3}{n_3} \Rightarrow \frac{m_1}{n_1} \lesssim \frac{m_3}{n_3} \quad \text{(transitividad).}$$

Adición de fracciones. La fracción $\frac{m_1 \cdot n_2 + n_1 \cdot m_2}{n_1 \cdot n_2}$ se llama *suma* de las fracciones racionales $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$. Se dice que esta fracción ha sido obtenida como resultado de la adición de fracciones $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$ y se escribe

$$\frac{m_1}{n_1}+\frac{m_2}{n_2}\sim\frac{m_1\cdot n_2+n_1\cdot m_2}{n_1\cdot n_2}. \tag{1}$$

En la expresión (1) las fracciones $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$ se llaman *sumandos.*

Multiplicación de fracciones. La fracción $\frac{m_1\cdot m_2}{n_1\cdot n_2}$ se llama *producto* de las fracciones racionales $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$. Se dice que esta fracción ha sido obtenida como resultado de la multiplicación de la fracción $\frac{m_1}{n_1}$ por la fracción $\frac{m_2}{n_2}$ y se escribe

$$\frac{m_1}{n_1}\cdot\frac{m_2}{n_2}\sim\frac{m_1\cdot m_2}{n_1\cdot n_2}. \tag{2}$$

En la expresión (2) las fracciones $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$ se llaman *factores.*

Las operaciones de adición y multiplicación de fracciones racionales poseen las propiedades de conmutatividad y asociatividad; las operaciones de adición y multiplicación se relacionan mediante la ley de distributividad de la multiplicación respecto a la adición.

Sustracción de fracciones. La adición de fracciones tiene una operación inversa que se llama *sustracción*. La fracción $\frac{m_1\cdot n_2-m_2\cdot n_1}{n_1\cdot n_2}$ se llama *diferencia* de dos fracciones racionales $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$. De la fracción racional $\frac{m_1\cdot n_2-m_2\cdot n_1}{n_1\cdot n_2}$ se dice que ha sido obtenida como resultado de la sustracción de la fracción $\frac{m_2}{n_2}$ a partir de la fracción $\frac{m_1}{n_1}$ y se escribe

$$\frac{m_1}{n_1}-\frac{m_2}{n_2}\sim\frac{m_1\cdot n_2-m_2\cdot n_1}{n_1\cdot n_2}.$$

División de fracciones. La multiplicación de fracciones tiene una operación inversa que se llama *división*. La fracción $\frac{m_1\cdot n_2}{n_1\cdot m_2}$ se llama *cociente* de dos fracciones racionales $\frac{m_1}{n_1}$ y $\frac{m_2}{n_2}$ $(m_2\neq 0)$. Se dice que la fracción racional $\frac{m_1\cdot n_2}{n_1\cdot m_2}$

es el resultado de la división de la fracción $\frac{m_1}{n_1}$ entre la fracción $\frac{m_2}{n_2}$ y se escribe

$$\frac{m_1}{n_1} : \frac{m_2}{n_2} \sim \frac{m_1 \cdot n_2}{n_1 \cdot m_2} \left(\text{o } \frac{m_1}{n_1} \Big/ \frac{m_2}{n_2} \sim \frac{m_1 \cdot n_2}{n_1 \cdot m_2} \right).$$

La fracción $\frac{m}{n}$, donde m es un número entero y n, un número natural, se llama *positiva*, si m es positivo, y *negativa*, si m es negativo.

La fracción racional negativa $\frac{m}{n}$ (m es un número entero negativo; n, un número natural) se suele escribir en la forma $-\frac{|m|}{n}$.

Por ejemplo, la fracción racional negativa $\frac{-5}{7}$ se escribe en la forma $-\frac{5}{7}$.

La fracción racional positiva $\frac{m}{n}$ se denomina *propia*, si su numerador es menor que su denominador ($m < n$), y es *impropia*, si su numerador es mayor o igual a su denominador ($m \geqslant n$).

Si la fracción racional positiva es impropia, su numerador se expresará en la forma $m = n \cdot k + r$, donde k es un número natural y r, un número entero no negativo que satisface la condición $r < n$. Entonces la fracción $\frac{m}{n}$ se escribe en la forma

$$\frac{m}{n} = k + \frac{r}{n}.$$

El número k se denomina la *parte entera de la fracción.*

Si resulta que $r \neq 0$, entonces la fracción racional impropia $\frac{m}{n}$ a veces se escribe en la siguiente forma:

$$k\frac{r}{n}. \qquad (3)$$

La expresión (3) se denomina expresión de una fracción impropia en forma de una *fracción mixta.*

Por ejemplo, la fracción racional impropia $\frac{10}{3}$ puede ser escrita en forma de la fracción mixta $3\frac{1}{3}$.

Una fracción se llama *irreductible*, si su numerador y denominador son números recíprocamente primos.

Cualquier fracción racional se puede escribir en forma de una fracción irreductible.

Ejemplo. Escribir la fracción racional $\frac{15}{75}$ en forma de fracción irreductible.

Desarrollemos los números 15 y 75 en el producto de los factores primos: $15 = 3 \cdot 5$, $75 = 3 \cdot 5 \cdot 5$.

La fracción $\frac{15}{75}$ puede ser escrita en forma de $\frac{3\cdot 5}{3\cdot 5\cdot 5}$. Simplificando los mismos factores en el numerador y denominador, obtenemos la inscripción de la fracción $\frac{15}{75}$ en la forma de una fracción irreductible $\frac{1}{5}$.

3.2. Números racionales. El conjunto de todas las fracciones racionales equivalentes entre sí se llama *número racional.* En vigor de la definición del número racional las fracciones distintas equivalentes entre sí representan solamente distintas expresiones de un mismo número racional. Así, por ejemplo, tres fracciones racionales equivalentes distantas.

$$\frac{2}{3}; \quad \frac{4}{6}: \quad \frac{-6}{-9}$$

tienen diferentes expresiones de un mismo número racional.

Se puede dar también otra definición del número racional, identificándolo no con el conjunto de todas las fracciones racionales equivalentes entre sí, sino con cierta fracción fija del mismo conjunto. Uno de los procedimientos posibles de separación de tal fracción fija consiste en lo siguiente.

Tomemos cualquier fracción $\frac{m}{n}$ del conjunto de fracciones, equivalentes entre sí (aquí m es un número entero, distinto de cero y n, un número natural). Si los números $|m|$ y n son recíprocamente primos, entonces consideremos que la fracción $\frac{m}{n}$ es precisamente la fracción fija que buscamos.

Si los números $|m|$ y n no son respectivamente números primos, entonces dividimos el numerador y denominador de la fracción entre el máximo común divisior de los números $|m|$ y n. Como resultado de la división obtenemos la fracción $\frac{m_1}{n_1}$, el numerador y denominador de la cual son recíprocamente números primos. La fracción obtenida $\frac{m_1}{n_1}$ y es la fracción fija que necesitamos*).

Ahora se puede dar la definición siguiente de un (distinto del cero) número:

Un *número racional* es tal número que puede ser representado en la forma $\frac{m}{n}$, donde $|m|$ y n son recíprocamente números primos (irreducibles) naturales.

Introduzcamos el concepto de igualdad de dos números racionales. Si usamos la definición de número racional como conjunto de fracciones racionales equivalentes, la definición de igualdad de dos números racionales tendrá la forma siguiente.

A dos números racionales α y β los llaman *iguales*:

$$\alpha = \beta,$$

si los dos conjuntos de fracciones racionales que proporcionan estos números coinciden.

La ordenación del conjunto de números racionales, del concepto de suma, del producto, de la diferencia y del cociente de dos números racionales se introducen de la misma manera que los respectivos conceptos para las fracciones racionales. La expresión matemática de la ordenación del conjunto de números racionales, de operaciones de adición, de multiplicación, de sustracción y de división de dos números racionales es la misma que la expresión de las operaciones respectivas con fracciones racionales, sólo que en vez del signo de equivalencia se debe escribir el signo de igualdad.

La propiedad de asociatividad de la multiplicación de números racionales permite introducir el concepto de potencia natural de un número racional:

---

*) El número racional $\frac{m_1}{n_1}$ puede ser considerado también como el resultado de la división del número m_1 entre el número n_1.

un número racional $\frac{p}{q}$ se llama *potencia k del número racional* $\frac{m}{n}$, obtenido como resultado de la multiplicación del número $\frac{m}{n}$ por sí mismo k veces:

$$\frac{p}{q}=\underbrace{\frac{m}{n}\cdot\frac{m}{n}\cdot\frac{m}{n}\ \ldots\ \frac{m}{n}}_{k \text{ factores}}.$$

De costumbre se utiliza una expresión más corta:

$$\frac{p}{q}=\left(\frac{m}{n}\right)^k. \tag{4}$$

En la expresión (4) el número $\frac{m}{n}$ se llama *base de una potencia* y el número k, *exponente de la potencia*.

Un número denominado $\left|\frac{m}{n}\right|$ y calculado según la regla:

$$\left|\frac{m}{n}\right|=\begin{cases}\frac{m}{n}, & \text{si } m>0;\\ 0, & \text{si } m=0; \ (n \text{ es natural}).\\ -\frac{m}{n}, & \text{si } m<0;\end{cases}$$

se llama *valor absoluto* (*módulo*) del número racional $\frac{m}{n}$.

3.3. Números enteros y racionales. Un número racional n se llama *entero*, si en el conjunto de todas las fracciones equivalentes que dan este número existe una fracción que tiene la forma de $\frac{n}{1}$.

Para calcular la suma, diferencia, producto y cociente de un número racional y entero n es suficiente escribir este número entero en forma de una fracción, cuyo denominador es igual a la unidad, es decir, en forma de $\frac{n}{1}$, y después utilizar las definiciones de la suma, diferencia, producto y cociente de dos números racionales.

Es muy fácil convencerse de que estas operaciones de adición, sustracción, multiplicación y división coinciden con las mismas operaciones para los números enteros escritos en forma de fracciones racionales con denominadores iguales a la unidad.

3.4. Conjunto de números racionales como extensión del conjunto de números enteros. El conjunto de todos los números enteros Z fue obtenido como una extensión del conjunto de números naturales mediante la adición al mismo de nuevos objetos numéricos: del número 0 y los números negativos enteros. El conjunto de números enteros obtenido está cerrado respecto a las operaciones de adición, multiplicación y sustracción, pero no está cerrado respecto a la división. El conjunto de todos los números racionales puede ser obtenido también como una extensión natural de un conjunto de números enteros mediante la adición de elementos nuevos, tales, que:

1) el conjunto extendido contenga como subconjunto propio el conjunto de todos los números enteros;

2) las operaciones aritméticas, definidas para los números enteros sean definidas también para los elementos de un conjunto extendido; al mismo tiempo el signo de estas operaciones para los números enteros, que examinamos como elementos de un conjunto extendido, debe coincidir con el signo de los mismos en el conjunto de números enteros antes de la extensión;

3) en el conjunto extendido sea posible realizar una operación de división (menos la división entre cero), la cual en un conjunto de números enteros no siempre es realizable, es decir, el cociente de dos elementos del conjunto extendido debe ser un elemento de este conjunto;

4) el conjunto extendido sea el menor en sentido de que él mismo no debe contener un subconjunto propio que satisfaga las condiciones 1) —3).

Existe un sólo conjunto que satisface las condiciones 1)—4): el conjunto ordenado de todos los números racionales con operaciones aritméticas introducidas en él. El conjunto de todos los números racionales frecuentemente se designa por Q.

Es muy fácil verificar que el conjunto de números racionales es un campo. Un campo de números racionales es de Arquimedes, es decir, para cualesquiera α y β, donde $\beta > 0$, existe tal número natural n, que $n \cdot \beta > \alpha$.

§ 4. Números reales

4.1. Conjunto de números reales como extensión de un conjunto de números racionales. En los §§ 2, 3 fue consecutivamente efectuada la extensión del conjunto de números naturales hasta el conjunto de números enteros y la extensión del conjunto de números enteros hasta el conjunto de números racionales. El conjunto de todos los números racionales representa por sí mismo un conjunto cerrado respecto a las operaciones de adición, multiplicación, sustracción y división (menos la división entre cero); la suma, producto, diferencia y cociente de dos números racionales igualmente es un número racional.

Sin embargo, existen problemas algebraicos y geométricos que no tienen solución dentro de conjunto de números racionales. Así, el problema que muy a menudo no tiene solución en el conjunto de números racionales es la extracción de una raíz de un número entero positivo. Por ejemplo, el número $\sqrt{2}$ (véase p. 4.6) no es un número racional, es decir, no se puede escribir en forma de $\frac{m}{n}$, donde m y n son números enteros y $n \neq 0$. Es muy fácil también citar aquí otros ejemplos de números que no pueden ser representados en forma de $\frac{m}{n}$, es decir, que no son números racionales.

Al número que no puede representarse en forma de $\frac{m}{n}$, donde m y n son números enteros y $n \neq 0$, lo llaman *número irracional*.

El conjunto de todos los números reales también puede ser obtenido como la extensión natural del conjunto de todos los números racionales. Sin embargo, a diferencia de los procedimientos relativamente simples de las extensiones del conjunto de números naturales hasta el conjunto de números enteros y el conjunto de números enteros hasta el conjunto de números racionales, el método de extensión (o de completación) del conjunto de todos los números racionales hasta el conjunto de números reales es más complicado. Una rigurosa teoría, desde el punto de vista matemático, de los números reales fue desarrollada sólo a mediados del siglo pasado en las obras de R. Dedekind y G. Cantor y para su desarrollo fue utilizada una serie de resultados bastante finos del análisis matemático.

Uno de los métodos de completación del conjunto de números racionales hasta un conjunto de números reales se basa en el concepto de la sucesión fundamental de los números racionales. Mostremos esquemáticamente en que consiste el contenido de este método.

La sucesión (x_n) de los números racionales x_n se llama *fundamental*, si para cualquier racional $\varepsilon > 0$ existe tal n_0 que

$$| x_p \ldots x_q | < \varepsilon$$

para todos p y q mayores que n_0. Cualquier sucesión convergente es fundamental. Por otra parte, cualquier sucesión fundamental de números racionales tiene un límite, pero puede resultar que este límite no sea un número racional. La sucesión de números racionales positivos (r_n), cuyos cuadrados pueden aproximarse tanto como se quiera a 2:

$$| r_1^2 - 2 | < \frac{1}{10}; \quad | r_2^2 - 2 | < \frac{1}{10^2}; \quad \ldots, \quad | r_n^2 - 2 | < \frac{1}{10^n}; \quad \ldots$$

puede servir de ejemplo de tal sucesión. De la representación de la sucesión (r_n) se deduce que su límite es un número no perteneciente al conjunto de números racionales; éste es el número irracional $\sqrt{2}$.

El conjunto de todos los números reales se obtiene mediante la completación del conjunto de números racionales por números irracionales que son los límites de cualesquiera sucesiones fundamentales de números racionales y que de costumbre se designa con la letra **R**.

4.2. Construcción axiomática de un conjunto de números reales. El conjunto de todos los números reales puede ser representado como un conjunto, cuyos elementos satisfacen las propiedades citadas más abajo I—VI. A los elementos de este conjunto, aunque ésto no esté bien adecuado, vamos a designarlos como *números*.

I. *Propiedad de ordenación.*

Para dos números cualesquiera a y b está definida la *relación del orden*, es decir, para dos números reales cualesquiera a y b

ó $a = b$ (a es igual a b),

ó $a < b$ (a es menor que b),

ó $b < a$ (b es menor que a);

al mismo tiempo, si $a < b$ y $b < c$, entonces $a < c$.

II. *Propiedades de la operación de adición.*

En un conjunto de número reales está definida la operación binaria de *adición*, es decir, a cualquier par ordenado de números $(a; b)$ se le asigna un único número que se llama *suma* de los números a y b y que se designa $a + b$; al mismo tiempo

1) para cualquiera terna a, b, c

$(a + b) + c = a + (b + c)$ (ley asociativa de la adición);

2) para cualquier par de números a y b

$a + b = b + a$ (ley conmutativa de la adición);

3) existe un número que se designa con el símbolo 0 y se llama *cero*, tal que para cualquier número a

$$a + 0 = a;$$

4) para cualquier número a existe un número que se designa $-a$, tal que

$$a + (-a) = 0;$$

el número $-a$ se llama *opuesto* al número a;

5) si $a < b$, entonces para cualquier número c

$$a + c < b + c.$$

El número $a > 0$ se llama *positivo*, y el número $a < 0$, *negativo*.

Para cualquier par ordenado de números $(a; b)$ el número

$$a + (-b)$$

se llama *diferencia* de los números a y b, y se designa $a - b$:

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

III. *Propiedades de la operación de multiplicación.*

En un conjunto de números reales está definida una operación binaria llamada *multiplicación*, es decir, a cualquier par ordenado de números $(a; b)$ se le asigna un único número que se llama *producto* de éstos y que se designa $a \cdot b$, al mismo tiempo

1) para cualquier terna de números a, b, c

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \text{ (asociatividad);}$$

2) para cualquier par de números a, b

$$a \cdot b = b \cdot a \text{ (conmutatividad);}$$

3) existe un número que se designa con el símbolo 1 y que se llama *unidad*, tal que para cualquier número a

$$a \cdot 1 = a;$$

4) para cualquier número a, distinto del cero, existe un número que se designa $\frac{1}{a}$, tal que

$$a \cdot \frac{1}{a} = 1;$$

al número $\frac{1}{a}$ lo llaman *inverso* del número a;

5) si $a < b$ y $c > 0$, entonces $a \cdot c < b \cdot c$; si $a < b$ y $c < 0$, entonces $a \cdot c > b \cdot c$.

Para cualquier par ordenado de número a y b (b es distinto de cero) el número $a \cdot \frac{1}{b}$ se llama *cociente* de la división de a entre b y se designa $\frac{a}{b}$:

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

IV. *Conexión de las operaciones de adición y multiplicación.*

Para cualquier terna de números a, b y c

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

(distributividad de la multiplicación respecto a la adición).

V. *Propiedad de Arquímedes.*

Para cualquier número a existe tal número entero n que $n > a$.

De esta propiedad se desprende, en particular, que para cualesquiera números a y b si $a > 0$ existe un número natural n tal que

$$n \cdot a > b.$$

VI. *Propiedad de continuidad.*

Las propiedades enumeradas más arriba I—V son propias también para algunos otros conjuntos numéricos (por ejemplo, para el conjunto de todos los números racionales). Un conjunto de números reales, a diferencia de un conjunto de números racionales, tiene una propiedad más, esencialmente nueva, la continuidad.

Existen diferentes definiciones de la propiedad de continuidad del conjunto de números reales. Una de ellas será citada a continuación y se denomina *principio de segmentos encajados* o *axioma de continuidad* del conjunto de números reales (según Cantor).

Si se dan dos números a y b, $a \leqslant b$, entonces el conjunto de todos los números x para los cuales $a \leqslant x \leqslant b$, se llama *segmento numérico* y se designa $[a; b]$. El número $b - a$ se llama *longitud* del segmento numérico.

Un sistema de segmentos numéricos

$$[a_1; b_1], [a_2; b_2], \ldots, [a_n; b_n]$$

se llama *sistema de segmentos encajados*, si

$$a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant b_n \leqslant b_{n-1} \leqslant \ldots \leqslant b_1.$$

Respecto al sistema de segmentos encajados $[a_n; b_n]$, $n = 1, 2, 3, \ldots$ se dice que la longitud de los mismos tiende a cero con el crecimiento de n, si para cualquier número $\varepsilon > 0$ existe tal número n_0 que para todos los números $n \geqslant n_0$ se cumple la desigualdad

$$b_n - a_n < \varepsilon.$$

Principio de segmentos encajados. Para cualquier sistema de segmentos encajados existe por lo menos un número que pertenece a todos los segmentos del sistema dado.

Del principio de segmentos encajados se deduce, en particular, que para cualquier sistema de segmentos encajados que tiende a cero en la longitud tiene un único número que pertenece a todos los segmentos del sistema dado.

4.3. Representación de números reales por fracciones decimales. En vigor de la propiedad de Arquímedes para cualquier número no negativo a se halla un número entero no negativo N tal que

$$N \leqslant a < N + 1.$$

Partamos un segmento numérico $[N;\ N+1]$ en diez partes iguales y examinemos los segmentos

$$[N;\ N,1],\quad [N,1;\ N,2],\ \ldots,\ [N,9;\ N+1].$$

Existen dos casos: a no coincide con ninguno de los puntos de la división o a coincide con uno de los puntos de la división. En el primer caso a pertenece sólo a uno de los segmentos enumerados, que designaremos I_1:

$$I_1 = [N,n_1;\ N,n_1+1],$$

donde n_1 es una de las cifras 0, 1, 2, . . ., 9. Si a es un punto de división, entonces como segmento I_1 eligiremos aquél, para el cual a es el extremo izquierdo.

Partamos el segmento I_1 en diez partes iguales y escojamos de los diez segmentos obtenidos aquél que contiene el punto a y para el cual e no es el extremo derecho. Designemos el segmento elegido por I_2. Continuando este proceso, obtenemos un sistema de segmentos encajados

$$I_k = [N,n_1n_2\ldots n_k;\quad N,n_1n_2\ldots(n_k+1),$$

donde n_i $(i = 1, 2, 3, \ldots, k)$ es una de las cifras 0, 1, 2, . . ., 9. Cada uno de los segmentos I_k contiene el punto a y para ninguno de estos segmentos el punto a es el extremo derecho.

A las expresiones

$$N,n_1n_2\ldots n_k \text{ y } N,n_1n_2\ldots(n_k+1)$$

las llaman respectivamente *fracciones decimales convenientes inferior* y *superior* del orden k y las designan $\underline{a}_k$ y $\overline{a}_k$. Estas fracciones satisfacen las relaciones

$$\begin{aligned} &\underline{a}_k < \overline{a}_k, \\ &\underline{a}_k \leqslant \underline{a}_{k+1}, \\ &\overline{a}_k \geqslant \overline{a}_{k+1}, \\ &\overline{a}_k - \underline{a}_k = \frac{1}{10^k} \end{aligned} \qquad (1)$$

Las sucesiones $(\underline{a}_k)$ y $(\overline{a}_k)$, formadas respectivamente por las fracciones convenientes inferiores y superiores, tienen los límites $\underline{a}$ y $\overline{a}$

$$\lim_{k\to\infty} N,\ n_1 n_2 \ldots n_k = \underline{a}; \qquad \lim_{k\to\infty} N,\ n_1 n_2 \ldots (n_k + 1) = \overline{a}.$$

En vigor de la última de las relaciones (1) las fracciones decimales $\underline{a}$ y $\overline{a}$ son iguales, es decir, representan una misma fracción decimal.

De otra parte, a es el único punto que pertenece a todos los segmentos I_k con longitudes que tienden a cero.

De tal manera, al número a le está asignada la fracción decimal $N,\ n_1 n_2 \ldots n_k \ldots$, la cual se llama *inscripción decimal* (o *representación decimal*) del número a y se escribe

$$a = N,\ n_1 n_2 \ldots n_k \ldots$$

La correspondencia entre el conjunto de todos los números reales y el conjunto de todas las fracciones decimales no es una correspondencia biunívoca: a distintas fracciones decimales les puede corresponder un sólo número, es decir, a fracciones con la forma

$$N,\ n_1 n_2 \ldots n_k (9) \text{ y } N,\ n_1 n_2 \ldots (n_k + 1) \qquad (0)$$

les corresponde un mismo número racional. Por ejemplo, el número racional $\frac{1}{4}$ admite la inscripción en forma de dos fracciones decimales distintas: 0,25 y 0,24(9), de lo que es muy fácil convencerse, utilizando el algoritmo de la conversión de una fracción periódica infinita a una fracción racional (véase p. 5.5 del capítulo presente).

Si examinamos el conjunto de fracciones decimales cuyo período se compone no sólo de la cifra 9 (a tales fracciones les llaman *admisibles*), entonces entre el conjunto de todos los números reales y el conjunto de todas las fracciones decimales admisibles se puede establecer una correspondencia biunívoca, en vigor de la cual se puede identificar el mismo número y su representación decimal.

Partiendo de la definición de la suma, diferencia, producto y cociente de los números racionales, se puede introducir el concepto de suma, del producto, de la diferencia y del cociente de dos números reales cualesquiera (menos la división entre cero), y el concepto del valor absoluto del número real.

Sean a y b dos números reales:

$$a = N_0,\ n_1 n_2 \dots n_k \dots;\quad b = M_0,\ m_1 m_2 \dots m_k \dots$$

Los números

$$\lim_{k\to\infty} (\underline{a}_k + \underline{b}_k) = a + b,$$

$$\lim_{k\to\infty} (\underline{a}_k \underline{b}_k) = a \cdot b,$$

$$\lim_{k\to\infty} (\underline{a}_k - \underline{b}_k) = a - b.$$

se llaman respectivamente *suma*, *producto* y *diferencia* de los números reales a y b.

El número

$$\lim_{k\to\infty} \frac{\underline{a}_k}{\underline{b}_k} = \frac{a}{b}$$

se llama *cociente* de los números a y b si $b \neq 0$.

Para tal definición del cociente de dos números reales puede resultar que para ciertos $k\underline{b}_k = 0$. Pero en vigor de la condición $b \neq 0$ siempre se halla tal número k_0 que $\underline{b}_k \neq 0$ para $k \geqslant k_0$ y entonces la sucesión $\underline{a}_k/\underline{b}_k$ debe ser examinada no para todos $k \in \mathbf{N}$, sino para $k = k_0, k_0 + 1, k_0 + 2, \dots$

El número no negativo

$$|a| = \lim_{k\to\infty} |\underline{a}_k|.$$

se llama *valor absoluto* (*módulo*) del número real a.

4.4. Representación geométrica de un conjunto de números reales. Examinemos una línea recta (que vamos a llamar *eje*) con un punto arbitrario elegido en ella, que designaremos con la letra O (por ahora precisamente con la letra y no con el número cero). Elijamos cualquier otro punto E que se encuentre en la recta a la izquierda del punto O. El punto O parte la recta en dos rayos: un rayo, al cual pertenece el punto E, se llama *semieje positivo* y un segundo rayo que se llama *semieje negativo*. Examinemos ahora el segmento de la recta con los extremos O y E (fig. 2.1).

Tomemos como unidad (al segmento OE también lo llaman segmento de escala) la longitud del segmento OE que designaremos $|OE|$. Demostremos ahora la correspondencia entre los números 0 y 1 y los puntos O y E, suponiendo que el punto designado con la letra O corresponde al número cero y el punto designado con la letra E corresponde al número 1 (al punto O lo llaman también *origen de las coordenadas*). Ahora es muy fácil demostrar la correspondencia entre cualquier número natural fijo y cierto punto

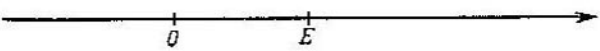


Fig. 2.1.

de la recta, bien determinado. Por ejemplo un punto que corresponde al número 5 se encuentra a la derecha del origen de las coordenadas a una distancia igual a cinco longitudes del segmento OE; al número negativo entero n le asignamos el punto que se encuentra a la izquierda del punto O a una distancia igual a n longitudes del segmento OE: $n \cdot |OE|$. Por ejemplo, el punto que corresponde al número 5 se encuentra a la izquierda del punto O a una distancia cinco veces mayor que la unidad de longitud.

Así sucesivamente, si tomamos el segmento unidad OE, lo partimos en n partes iguales y trazamos a la derecha del punto O la parte n del segmento OE, entonces el punto, cual es el extremo derecho del segmento trazado, se considera como el punto que representa el número $\frac{1}{n}$. Si trazamos a la derecha del punto O la suma m de segmentos de longitud $|OE|$ entonces, el extremo derecho de la suma será un punto que representa el número positivo racional $\frac{m}{n}$. Análogamente se pueden construir los puntos que corresponden a los números racionales negativos.

Queda por señalar el procedimiento de correspondencia de los números irracionales a los puntos de la recta. Supongamos que el número positivo a es irracional. Construyamos las sucesiones de fracciones decimales convenientes para el número irracional a:

$$\underline{a}_1;\ \underline{a}_2;\ \underline{a}_3;\ \ldots;\ \underline{a}_k;\ \ldots \text{ y } \overline{a}_1;\ \overline{a}_2;\ \overline{a}_3;\ \ldots,\ \overline{a}_k;\ \ldots$$

Los términos de estas sucesiones son los números racionales, a los cuales corresponden los puntos adecuados de la recta. Construyamos un sistéma de segmentos $[\underline{a}_1;\ \overline{a}_1]$; $[\underline{a}_2;\ \overline{a}_2]$; ... $[\underline{a}_k;\ \overline{a}_k]$; ..., encajados uno en otro. La longitud de estos segmentos tiende a cero con el crecimiento de k. Estos segmentos tienen un sólo punto común, precisamente el punto correspondiente al número irracional a.

De tal manera se puede aplicar el conjunto de todos los números reales **R** sobre el conjunto de los puntos de la recta. Es válido también lo inverso: a cada punto de la recta se puede asignar un único número real.

Entre el conjunto de números reales y el conjunto de puntos de la recta existe una correspondencia biunívoca y con frecuencia no distinguen estos dos conjuntos, diciendo simplemente «la recta numérica».

4.5. Representaciones decimales de los números racionales e irracionales. Un número racional se define como un número que puede ser escrito en la forma m/n, donde m y n son números enteros y $n > 0$.

Examinemos tres ejemplos de representación de los distintos números racionales en forma de fracciones decimales.

Ejemplo 1. Escribir el número racional $1/4$ en forma de fracción decimal:

$$\begin{array}{r|l} 1 & 4 \\ \hline 10 & 0{,}25 \\ -\ 8 & \\ \hline 20 & \\ -\ 20 & \\ \hline 0 & \end{array}$$

De los cálculos citados más arriba se ve que el proceso de división termina después de un número finito de pasos y el número $1/4$ puede ser escrito en forma de la fracción decimal $0{,}25$.

Ejemplo 2. Escribir el número racional 5/11 en forma de fracción decimal:

$$\begin{array}{rl}
{*}\ 5 & \big|\ 11 \\
50 & \big|\ 0,4545 \\
44 & \\
\hline
60 & \\
55 & \\
\hline
{*}\ 50 & \\
44 & \\
\hline
60 & \\
55 & \\
\hline
{*}\ 5 &
\end{array}$$

Al examinar los cálculos citados más arriba, es muy fácil observar que si continuamos el proceso de división, seguiremos obteniendo en el resto los números 6 y 5 sucesivamente y cada vez, obteniendo como resto el número 5, empezamos de nuevo el ciclo de división (en el ejemplo citado más arriba marcamos con estrellitas el comienzo de los tres primeros de estos ciclos). En este ejemplo el proceso de división no termina nunca, por consiguiente, el número racional 5/11 se escribe como fracción periódica decimal infinita 0,(45).

Ejemplo 3. Escribir el número racional 131/990 en forma de una fracción decimal:

$$\begin{array}{rl}
131 & \big|\ 990 \\
\hline
-\,1310 & \big|\ 0,132 \\
990 & \\
\hline
{*}\ 3200 & \\
-\,2970 & \\
\hline
2300 & \\
-\,1980 & \\
\hline
{*}\ 320 &
\end{array}$$

En este ejemplo, al igual que en el anterior el proceso de división no termina nunca. Pero aquí la reiteración del ciclo de división comienza con un número, distinto al número del comienzo de la división (aquí son los números 320), este número apareció en uno de los restos. Por consiguiente, el número racional 131/990 se escribe en forma de fracción decimal periódica mixta 0,1(32).

Los tres ejemplos recién citados describen todos los casos posibles con los que tropezamos al escribir un número racional en forma de fracción decimal. Examinemos un número racional positivo arbitrario m/n, donde m y n son números enteros positivos (el caso de una fracción racional negativa puede ser examinado análogamente). Durante la división del número natural m entre el número natural n en el resto pueden aparecer sólo los números siguientes:

$$0,\ 1,\ 2,\ \ldots,\ n-1.$$

Si durante el proceso de división en el resto aparece el número 0, el proceso de división se termina y, por consiguiente, el número racional m/n se representa como una fracción decimal finita. Si en el proceso de división el número cero en el resto no aparece, entonces por lo menos en el transcurso de $n-1$ paso surge inevitablemente la reiteración del resto; a partir de este momento comienza un ciclo nuevo. La fracción decimal periódica infinita (mejor decir, una fracción decimal periódica admisible) es el resultado de la división.

De tal manera, cualquier número racional m/n es representable tanto en forma de una fracción finita como infinita periódica; y a la inversa, cualquier fracción finita y también cualquier fracción decimal periódica infinita es la expresión de cierto número racional.

El número racional m/n (m, n son números enteros recíprocamente simples y $n > 1$) puede ser escrito en forma de fracción decimal finita cuando y sólo cuando el número n tiene como divisores simples suyos sólo los números 2 y 5. Además el número n no debe obligatoriamente tener entre sus divisores simples ambos números (2 y 5); este número puede dividirse sólo entre uno de ellos. Si $n = 1$, entonces la fracción $\frac{m}{n}$ también se escribe en forma de una fracción decimal finita. Por ejemplo, los números racionales 1/25, 1/16 y 7/1, donde n es igual respectivamente a 25, 16 y 1, se representan en forma de fracciones decimales finitas:

$$\frac{1}{25} = 0{,}04; \qquad \frac{1}{16} = 0{,}0625; \qquad \frac{7}{1} = 7.$$

Puesto que cualquier número real puede ser escrito en forma de fracción decimal y cualquier número racional, o en forma de una fracción decimal finita, o en forma de una fracción decimal periódica infinita, entonces cualquier número irracional puede ser escrito en forma de fracción decimal aperiódica infinita.

Esta propiedad de los números irracionales a veces es considerada como definición de los números irracionales:

Un número real, cuya inscripción decimal es una fracción decimal aperiódica infinita, se llama *número irracional.*

Propiedades de los números irracionales. A diferencia del conjunto de números racionales que era cerrado respecto a las operaciones de adición, sustracción, multiplicación y división (menos la división entre cero) el conjunto de números irracionales no tiene carácter cerrado para ninguna de las operaciones citadas. Para convencerse de que los conjuntos de números irracionales no tienen caracter cerrado, por ejemplo, respecto a la operación de adición, es suficiente indicar un sólo par de números irracionales, la suma de los cuales es racional. Como par de tales números pueden ser tomados, por ejemplo, los números 0,1010010001 ... y 0,0101101110..., el primero de los cuales se forma mediante la sucesión de unidades, separadas respectivamente por un cero, dos ceros, tres ceros, etc., el segundo, mediante la sucesión de ceros, entre los cuales se encuentran una unidad, dos unidades, tres unidades, etc. Cada uno de estos números es un número irracional, mientras que su suma es un número racional, cuya representación decimal tiene la forma

$$0,111\ldots1\ldots=0,(1)=\frac{1}{9}.$$

La suma, diferencia, producto y resto del número irracional a y del número racional r son números irracionales.

De esta propiedad de los números irracionales, en particular, se deduce que, teniendo sólo un número irracional, se puede construir mediante números racionales una cantidad infinita de números irracionales.

4.6. Algunos modos de demostración de la irracionalidad de los números.

Demostración de la irracionalidad de los números basada en la definición del número racional. La irracionalidad de ciertos números puede ser demostrada mediante el método de demostración por reducción «al absurdo».

Supongamos que, por ejemplo, necesitamos demostrar la irracionalidad del número $\sqrt{2}$. Supongamos que $\sqrt{2}$ es un número racional, es decir, es un número que puede ser representado en forma de

$$\sqrt{2}=\frac{m}{n}, \tag{2}$$

donde m y n — son números naturales reciprocamente primos. Para demostrar la imposibilidad de la representación del número $\sqrt{2}$ en forma (2) utilicemos la simplicidad recíproca de los números m y n. O, mejor dicho, utilicemos el que los números m y n no son pares, en caso contrario la fracción $\frac{m}{n}$ sería reducible. Elevando los dos miembros de la igualdad (2) al cuadrado, obtendremos

$$2=\frac{m^2}{n^2} \Leftrightarrow 2\cdot n^2=m^2.$$

El número $2\cdot n^2$ es par. Por eso m^2 y, por consiguiente, m también debe ser par. Suponiendo $m = 2\cdot k$, la igualdad $2\cdot n^2 = m^2$ se puede escribir en la siguiente forma

$$2\cdot n^2 = (2\cdot k)^2 \Leftrightarrow 2\cdot n^2 = 4k^2 \Leftrightarrow n^2 = 2\cdot k^2.$$

De la última igualdad vemos que el número n^2 es también par; por consiguiente, n es par también. Llegamos a la conclusión que tanto m como n son números pares, mientras que la fracción m/n según la suposición es irreducible. La contradicción obtenida demuestra que el número $\sqrt{2}$ no se puede representar en forma de m/n, y, por consiguiente, es irracional.

Análogamente puede ser demostrada la irracionalidad del número $\sqrt{3}$. Pero, a diferencia del caso anterior, el factor decisivo aquí es que los números m y n no se dividen al mismo tiempo entre 3.

Demostración de la irracionalidad de los números mediante el teorema principal de la aritmética. La irracionalidad de ciertos números se puede demostrar mediante el teorema principal de la aritmética (véase p. 1.3). Según el teorema principal de la aritmética cualquier número natural únicamente se descompone en el producto de factores primos.

Por ejemplo, demostremos que el número lg 2 es irracional. Supongamos que esto no es así, es decir, que existen tales m y n naturales para los cuales

$$\lg 2=\frac{m}{n}, \quad ó \quad 2=10^{\frac{m}{n}}.$$

Elevando ambos miembros de la última igualdad a la potencia n, obtenemos

$$2^n=10^m \Longleftrightarrow 2^n=2^m\cdot 5^m.$$

Según el teorema principal de aritmética la igualdad $2^n = 2^m\cdot 5^m$ es imposible, puesto que 2^n es un número natural que para ningún valor de n se divide entre 5, puesto que m es número natural. Por consiguiente, el número lg 2 es irracional.

Un método general de demostración de la irracionalidad de una serie de números (incluso la irracionalidad de ciertos valores de funciones trigonométricas). Un método bastante general de demostración de la irracionalidad de los números se basa en *el teorema* siguiente:

Si la ecuación algebraica

$$n_0x^k+n_1x^{k-1}+\ldots+n_{k-1}x+n_k=0$$

cuyos coeficientes son números enteros tiene una raíz racional $\frac{m}{n}$ (los números m y n son recíprocamente primos), entonces el número m es divisor del número n_k y el número n es divisor del número n_0.

Examinemos ahora la ecuación que tiene la forma

$$x^k+n_1x^{k-1}+\ldots+n_{k-1}x+n_k=0$$

cuyos coeficientes son números enteros y el mayor de los coeficientes es igual a 1. Si esta ecuación tiene una raíz racional, entonces según el teorema citado esta raíz es entera y es divisor del número n_k. El número $\sqrt[k]{a}$, donde k y a son números positivos enteros, o es irracional, o es entero. En el último caso el número a es la potencia k de un número entero.

La demostración de la irracionalidad de los números mediante el teorema formulado se realiza de la manera siguiente*):

1) Se escribe una ecuación algebraica de la potencia natural más pequeña con coeficientes expresados con números enteros, una de las

*) El método que exponemos a continuación puede utilizarse sólo para demostrar la irracionalidad de los números algebraicos (véase p. 4.7).

raíces de esta ecuación es conocida de antemano, el número a, cuya irracionalidad hay que demostrar.

2) Se hallan los divisores primos del primer coeficiente n_0 de la ecuación obtenida y del término libre n_k. De los números enteros obtenidos se componen todos los números racionales posibles cuyo numerador es el divisor del número n_k y el denominador es el divisor del número n_0. Entonces, según el teorema formulado más arriba, sólo estos números racionales pueden ser las raíces racionales de la ecuación dada.

3) Comparando estas raíces «potenciales» con el número dado a, muestran que ninguno de los números racionales construidos es igual al número a, lo que demuestra que el número a es irracional.

Ejemplos. 1. Demostremos que el número $\sqrt[3]{2} - \sqrt{3}$ es irracional.

Supongamos que $x = \sqrt[3]{2} - \sqrt{3}$. Entonces $x + \sqrt{3} = \sqrt[3]{2}$. Elevando ambos miembros de la igualdad al cubo, después de unas transformaciones no complicadas obtendremos la ecuación

$$x^3 + 9x - 2 = -3\sqrt{3}(x^2 + 1).$$

Elevando ambos miembros de la última ecuación al cuadrado, después de la reducción de términos semejantes obtendremos la ecuación

$$x^6 - 9x^4 - 4x^3 + 27x^2 - 36x - 23 = 0.$$

Del procedimiento de construcción de esta ecuación se desprende que el número $\sqrt[3]{2} - \sqrt{3}$ es su raíz. Por otra parte, las únicas raíces racionales posibles de esta ecuación son los números — dividores enteros del número -23, es decir, $+1$, -1, $+23$, -23. La sustitución directa de estos números en la ecuación muestra que éstos no son sus raíces. De esta manera, nuestra ecuación no tiene raíces racionales y el número $\sqrt[3]{2} - \sqrt{3}$ es de antemano irracional.

2. Demostremos que $\cos 20°$ es un número irracional.

Según la fórmula del argumento triple $\cos 60° = 1/2$, $\cos 20°$ se relaciona con la fórmula

$$\cos 60° = 4\cos^3 20° - 3\cos 20°.$$

Al denominar $\cos 20° = x$, escribiremos la última igualdad en la siguiente forma

$$8x^3 - 6x - 1 = 0. \qquad (3)$$

El número $x = \cos 20°$ es la raíz de esta ecuación. Aplicando a esta ecuación el teorema formulado más arriba, vemos que sus únicas raíces racionales posibles son los números ± 1, $\pm 1/2$, $\pm 1/4$, $\pm 1/8$. Sustituyendo estos números en la ecuación (3), podemos convencernos que ninguno de ellos es raíz de esta ecuación. Por consiguiente, nuestra ecuación no tiene raíces racionales y por eso el número $\cos 20°$ es irracional. Es necesario subrayar que el procedimiento de la utilización directa de la fórmula del argumento triple no sirve para la demostración, por ejemplo, de la irracionalidad del número $\cos 10°$, puesto que la ecuación algebraica de la tercera potencia, cuya raíz es $x = \cos 10°$, tiene los coeficientes irracionales. La irracionalidad de cos 10° se puede demostrar en dos etapas: primeramente demostrar que cos 20° es un número irracional, después, utilizando la fórmula

$$1 + \cos 20° = 2 \cos^2 10°,$$

mostrar, haciendo uso del método de la demostración por reducción al absurdo que cos 10° no puede ser un número racional.

Indiquemos además un *principio* simple que permite demostrar la irracionalidad de muchas funciones trigonométricas:

Si el ángulo θ es tal que el número cos 2θ es un valor irracional, entonces cos θ, sen θ, tg θ también son irracionales.

Para convencerse de que esta afirmación es válida es suficiente utilizar el método de la demostración por reducción al absurdo, utilizando previamente las fórmulas del argumento medio y la identidad trigonométrica conocida

$$1+\mathrm{tg}^2\,\theta=\frac{1}{\cos^2\theta}.$$

Así pues, supongamos que sen θ es un número racional; entonces 2 sen² θ es también un número racional. Puesto que sen² θ y cos 2θ se relacionan por la igualdad

$$1 - \cos 2\theta = 2\,\mathrm{sen}^2\,\theta,$$

de la racionalidad de 2 sen² θ se deduce la racionalidad de cos 2θ. La contradicción obtenida demuestra que tanto sen θ, como cos 2θ, son irracionales.

4.7. Números algebraicos y trascendentes. Un conjunto de números reales, además de estar dividido en un conjunto de números racionales y un conjunto de números irracionales, puede ser dividido en otros dos conjuntos: en un conjunto de números algebraicos y un conjunto de números trascendentes.

Un número real que es raíz de cierta ecuación algebraica con los coeficientes enteros $n_0, n_1, \ldots, n_{k-1}, n_k$ cuya forma es

$$n_0x + n_1x^{k-1} + \ldots + n_{k-1}x + n_k = 0$$

se llama *número algebraico.*

Un número real que no es raíz de ninguna ecuación algebraica con coeficientes enteros se llama *número trascendente.*

El conjunto de todos los números racionales es un subconjunto del conjunto de números algebraicos, puesto que cualquier número racional $\frac{m}{n}$ es la raíz de la ecuación de primer grado

$$nx - m = 0.$$

De aquí se deduce también que cualquier número trascendente es irracional.

La *partición* de un conjunto de números reales en distintos subconjuntos puede ser representada esquemáticamente como sigue:

$$\text{Números reales} — \begin{cases} \text{Racionales (todos ellos son algebraicos)} \\ \text{Irracionales} — \begin{cases} \text{Algebraicos} \\ \text{Trascendentes} \end{cases} \end{cases}$$

$$\text{Números reales} — \begin{cases} \text{Algebraicos} — \begin{cases} \text{Racionales} \\ \text{Irracionales} \end{cases} \\ \text{Trascendentes (todos ellos son irracionales)} \end{cases}$$

Ciertos números trascendentes. En el p. 1.6 fue demostrada la irracionalidad de ciertos números, en particular, la irracionalidad del lg 2. Al mismo tiempo el lg 2 es número trascendente. Por primera vez la suposición acerca de la trascendencia del número lg 2 fue mencionada por L. Euler en el siglo XVIII. La demostración de la trascendencia del número lg 2 es más complicada que la demostración de su irracionalidad y se basa en métodos mucho más genéricos y profundos que aquellos que se utilizan para la demostración de su irracionalidad. La trascendencia del número lg 2 se desprende del siguiente *teorema* general:

El número a^b es trascendente, si los números a y b son algebraicos (los casos $a = 0$, $a = 1$ y el caso de b racional están eliminados).

Los bien conocidos números e y π son también trascendentes, su trascendencia fue demostrada a fines del siglo diecinueve. Anotemos que el problema de algebraidad o de trascendencia del número π se relaciona estrechamente con la solución de un problema geométrico formulado algunos siglos antes de nuestra era. Este problema se llama *problema de la cuadratura del círculo* y se formula de la manera siguiente:

¿Es posible construir un cuadrado cuya área sea igual a la de un círculo con un radio igual a la unidad utilizando para ello sólo un compás y una regla?

Se sabe que mediante un compás y una regla se pueden construir sólo tales figuras geométricas, cuyas áreas se expresan mediante un número algebraico. Por eso la trascendencia del número π muestra la imposibilidad de solución del problema sobre la cuadratura del círculo.

Y a la inversa, los valores de las funciones trigonométricas para ciertos valores del argumento (por ejemplo, los valores de $\cos 20°$ y $\operatorname{sen} 10°$, la irracionalidad de los cuales fue demostrada en el p. 4.6) son números algebraicos, lo que se desprende del *teorema* siguiente:

Para cualquier número racional r los números $\operatorname{sen}(90° \cdot r)$, $\cos(90° \cdot r)$ y $\operatorname{tg}(90° \cdot r)$ son algebraicos. La única restricción que es necesario tener en cuenta consiste en lo siguiente: en el caso de $\operatorname{tg}(90° \cdot r)$ el número r debe ser tal que el número $\operatorname{tg}(90° \cdot r)$ exista.

Anotemos en conclusión que el conjunto de todos los números algebraicos forma un conjunto numerable y el conjunto de todos los números trascendentes forma un conjunto innumerable.

4.8. Potencias y raíces. La propiedad de asociatividad de la multiplicación de números reales permite introducir el concepto de potencia natural de cualquier número real: el número real b que se obtiene como resultado de la multiplicación del número a por sí mismo n veces:

$$b = \underbrace{a \cdot a \cdot a \cdot \ \ldots \ \cdot a}_{n \text{ factores}}$$

se llama *potencia* natural n.
La potencia n del número a se denota a^n y se escribe

$$b = a^n.$$

El número a se llama *base de una potencia* y el número n, *exponente de una potencia*.

Para $a \neq 0$, según la definición $a^0 = 1$; 0^0 no está definido.

Para $a \neq 0$, según la definición $a^{-n} = \frac{1}{a^n}$ (n es un número natural).

La única solución positiva de la ecuación $x^n = a$ se llama *raíz* de n-ésima potencia (o *raíz aritmética* de n-ésima potencia) del número real positivo a.

La raíz de n-ésima potencia del número a se designa con $\sqrt[n]{a}$ ó $a^{\frac{1}{n}}$ *). La raíz de segunda potencia (raíz cuadrada) suele designarse simplemente $\sqrt{a}$.

$$\text{Para } a = 0, \quad \sqrt{0} = 0 \text{ ó } 0^{\frac{1}{n}} = 0.$$

Si un número real a es negativo, la raíz de n-ésima potencia del número a se halla sólo para n impar como solución única real negativa de la ecuación $x^n = a$.

La raíz de n-ésima potencia del número real negativo a se designa con el mismo símbolo $\sqrt{a}$ (n es impar).

El número $(\sqrt{a})^m$ se llama *potencia racional* $\frac{m}{n}$ (m es entero, n es natural) del número real positivo a.

---

*) El símbolo $\sqrt[n]{\ }$ también se llama *radical* de n-ésima potencia.

La potencia racional del número a también se escribe de la siguiente forma:

$$(\sqrt[n]{a})^m = (a^{\frac{1}{n}})^m = (a^m)^{\frac{1}{n}} = \sqrt[n]{a^m} = a^{\frac{m}{n}}.$$

De la definición de la potencia racional de un número positivo se deducen las igualdades

$$\sqrt[m]{a}\cdot\sqrt[n]{a} = \sqrt[m\cdot n]{a^{m+n}},$$

$$\sqrt[m]{\sqrt[n]{a}} = \sqrt[m\cdot n]{a},$$

$$\sqrt[n]{a\cdot b} = \sqrt[n]{a}\cdot\sqrt[n]{b},$$

$$\sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}.$$

Una potencia con cualquier valor real del exponente puede ser introducida, por ejemplo, de la manera siguiente.

Supongamos que b es un número real escrito en forma de fracción decimal infinita:

$$b = N,\ n_1,\ n_2,\ \ldots,\ n_k\ldots,$$

y sea $(\underline{b}_k)$ una sucesión de fracciones decimales convenientes inferiores del orden k para el número b, es decir,

$$\underline{b}_1 = N,\ n_1;\ \underline{b}_2 = N,\ n_1,\ n_2;\ \ldots;\ \underline{b}_k = N,\ n_1,\ n_2\ldots\ldots n_k;\ \ldots$$

Para cualquier número positivo a se puede formar una sucesión infinita

$$a^{\underline{b}_1};\ a^{\underline{b}_2};\ \ldots;\ a^{\underline{b}_k};\ \ldots$$

El límite de esta sucesión se denota a^b y representa una potencia real del número a.

De la definición del número a^b se deducen las igualdades

$$a^{\alpha+\beta} = a^{\alpha} \cdot a^{\beta},$$

$$(a^{\alpha})^{\beta} = a^{\alpha \cdot \beta},$$

$$\frac{a^{\alpha}}{a^{\beta}} = a^{\alpha-\beta},$$

$$(a \cdot b)^{\alpha} = a^{\alpha} \cdot b^{\alpha},$$

$$\left(\frac{a}{b}\right)^{\alpha} = \frac{a^{\alpha}}{b^{\alpha}}.$$

4.9. Logaritmos. Sea a un número real positivo que se distingue de la unidad y M, cualquier número real positivo. El número designado $\log_a M$, tal que

$$a^{\log_a M} = M$$

se llama *logaritmo* del número M respecto a la base a.

El logaritmo $\log_a M$ del número positivo M para la base positiva a distinta de la unidad se puede definir también como la solución de la ecuación

$$a^x = M.$$

Las propiedades principales de los logaritmos:

$$\log_a a = 1 \qquad (a > 0,\ a \neq 1);$$
$$\log_a (a^k) = k \qquad (a > 0,\ a \neq 1);$$
$$\log_a (M_1 M_2) = \log_a M_1 + \log_a M_2 \ (M_1 > 0,\ M_2 > 0);$$
$$\log_a \left(\frac{M_1}{M_2}\right) = \log_a M_1 - \log_a M_2 \ (M_1 > 0,\ M_2 > 0);$$
$$\log_a (b^c) = c \log_a b;$$
$$\log_a c = \frac{\log_b c}{\log_b a}; \quad \log_a c = \frac{1}{\log_c a}.$$

Los *logaritmos decimales* (logaritmos respecto a la base 10) $\log_{10} a$ suelen denotarse $\lg a$.

Los *logaritmos naturales* (es decir, los logaritmos respecto a la base $e = 2{,}718281828\ \ldots$) $\log_e a$ suelen designarse $\ln a$.

Los logaritmos decimales y naturales se relacionan por las siguientes igualdades

$$\ln a = \frac{\ln a}{\lg e} = \ln 10 \cdot \lg a = (2{,}30259\ \ldots) \lg a,$$

$$\lg a = \frac{\ln a}{\ln 10} = \lg e \cdot \ln a = (0{,}43429\ \ldots) \ln a.$$

§ 5. Fracciones decimales

5.1. Sistema decimal posicional de numeración. El modo más difundido para escribir los números es el empleado en el sistema decimal posicional de numeración)*. La esencia de este modo de escribir los números consiste en lo siguiente:

1. Todos los números naturales desde la unidad hasta el nueve se designan con los siguientes símbolos individuales:

$$1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9.$$

A estos símbolos se une adicionalmente el décimo signo 0, que es el símbolo del cero. Los diez símbolos enumerados se llaman cifras del *sistema decimal de numeración.*

2. Una misma cifra tiene distinto sentido según la posición de esta cifra respecto a las demás en la escritura del número (en esto consiste el carácter posicional del sistema de numeración). Así, por ejemplo, en el sistema decimal posicional de numeración las combinaciones 4521 y 4125 de cuatro cifras distintas 1, 2, 4, 5 dan la expresión de dos números naturales distintos.

La cifra que se encuentra en el primer lugar a la derecha en la expresión del número natural indica la cantidad de unidades que contiene el número dado; la cifra que se encuentra en el segundo lugar indica la cantidad de decenas; en el tercer lugar, la cantidad de centenas; en el cuarto lugar, la cantidad de millares, etc.

Para escribir un número natural en el sistema decimal posicional de numeración se dice de la primera cifra situada a la derecha, que ella se encuentra en el orden de unidades, de la segunda se dice, que se encuentra en el orden de decenas; de la tercera, que se encuentra en el orden de centenas, etc.

*) En la práctica se suele omitir la palabra «posicional» y se dice sólo «sistema decimal de numeración».

En el sistema decimal de numeración la expresión

$$n_k n_{k-1} n_{k-2} \ldots n_1 n_0,$$

donde n_k, n_{k-1}, n_{k-2}, ..., n_1, n_0 son cifras, es la expresión condicional del número

$$n_k \cdot 10^k + n_{k-1} \cdot 10^{k-1} + n_{k-2} \cdot 10^{k-2} + \ldots \\ \ldots + n_1 \cdot 10 + n_0. \qquad (1)$$

Para escribir un número natural en el sistema decimal de numeración suele emplearse también la regla siguiente: cualesquiera de las diez cifras, menos el cero, puede ser la última cifra a la izquierda de la inscripción. Así, por ejemplo, se suele escribir el número doce en forma de 12 y no 012 ó 0012. Gracias a esta regla desaparece la posible multiformidad en la escritura de un mismo número.

El sistema decimal de numeración es uno de los muchos sistemas posicionales de numeración posibles. Así pues, todos los cálculos realizados por las computadoras electrónicas se efectúan mediante el así llamado sistema posicional binario de numeración. Los números en el sistema posicional binario de numeración se escriben mediante dos cifras: el cero (0) y la unidad (1).

En el sistema posicional binario de numeración la escritura

$$a_n a_{n-1} a_{n-2} \ldots a_1 a_0,$$

donde a_n, a_{n-1}, a_{n-2}, ..., a_1, a_0 son cifras del sistema binario de numeración que representan la escritura abreviada del número

$$a_n \cdot 2^n + a_{n-1} \cdot 2^{n-1} + a_{n-2} \cdot 2^{n-2} + \ldots \\ \ldots + a_1 \cdot 2^1 + a_0 \cdot 2^0. \qquad (2)$$

Ejemplo. En el sistema posicional binario de numeración el número 9 se escribe como 1001.

Es muy fácil convencerse en esto, si escribimos el número 1001 en forma de (2):

$$1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 9.$$

5.2. Concepto de fracción decimal. La *fracción decimal* es una forma de escribir el número real en el sistema decimal posicional de numeración. La fracción decimal

$$N,\ n_1,\ n_2,\ n_3,\ \ldots,\ n_k \ldots, \tag{3}$$

donde N es un número entero y $n_1, n_2, n_3, \ldots, n_k, \ldots$, cifras del sistema decimal de numeración, es la expresión condicional del número real

$$N + \frac{n_1}{10^1} + \frac{n_2}{10^2} + \frac{n_3}{10^3} + \ldots + \frac{n_k}{10^k} + \ldots$$

El número entero N se llama *parte entera* de la fracción decimal (3). De la cifra n_1, que se encuentra en la fracción decimal (3) en el primer lugar después del punto, se dice que se encuentra en el orden de décimas partes de la unidad; de la cifra n_2, que se encuentra en el segundo lugar, se dice que se encuentra en el orden de centésimas partes de la unidad; de la cifra n_3 que se encuentra en el orden de milésimas partes, etc.

Conforme a la definición de fracción decimal dada más arriba, en caso que N sea un número entero negativo, el signo menos no se escribe delante de toda la fracción, sino sobre el número $|N|$. Por ejemplo, una fracción decimal 2, 135... es la expresión condicional del número real

$$-2 + 0{,}135\ldots$$

Debemos subrayar que para N negativo tal expresión de la fracción decimal en una serie de casos (por ejemplo, para efectuar acciones aritméticas) es bastante incómoda. Por eso, muy a menudo, se usa otra forma de inscripción:

Sea a un número real positivo que tiene la representación decimal

$$a = 3{,}521.$$

El número $-a$ inverso del número a, conforme a la designación citada más arriba tendrá la representación decimal

$$-a = 4{,}479.$$

Para evitar la desconformidad en las designaciones de los números reales y sus representaciones decimales, también se escribe el número $-a$ en la forma siguiente:

$$-a = -3{,}521 = -3 - \frac{5}{10^1} - \frac{2}{10^2} - \frac{1}{10^3}.$$

La fracción decimal $N, n_1, n_2, n_3 \ldots n_k \ldots$ se llama *infinita*, si para cualquier k natural existe un número $l > k$ tal que $n_l \neq 0$.

La fracción decimal $N, n_1 n_2 n_3 \ldots n_k \ldots$ se llama *finita*, si existe tal número natural k que $n_k \neq 0$ y $n_l = 0$ para todos $l > k$. Las cifras $N, n_1, n_2, \ldots, n_k$ se llaman *cifras significativas*.

En la escritura de una fracción finita suelen omitir los ceros que se encuentran después de la última cifra significativa, es decir, la fracción

$$N, n_1 n_2 n_3 \ldots n_k 00 \ldots 0 \ldots,$$

donde $n_k \neq 0$, se escribe en la forma siguiente:

$$N, n_1 n_2 n_3 \ldots n_k.$$

Fracciones decimales periódicas. Una fracción decimal $N, n_1 n_2 n_3 \ldots n_k \ldots$ se llama *periódica*, si existen tales números naturales p, q que $n_{k+p} = n_k$ para todos $k > q$. Para la designación de una fracción decimal periódica infinita se usa la expresión $N, n_1 n_2 \ldots n_q (n_{q+1} n_{q+2} \ldots n_{q+p})$, donde un conjunto de cifras $e_{q+1} n_{q+2} \ldots n_{q+p}$ se llama *período de la fracción*.

A veces las fracciones periódicas suelen dividirse en fracciones periódicas puras, las cuales se escriben

$$N, (n_1 n_2 \ldots n_p),$$

y las fracciones *periódicas mixtas* que se escriben

$$N, n_1 n_2 \ldots n_k (n_{k+1} n_{k+2} \ldots n_{k+p}) \quad (k \geqslant 1).$$

Por ejemplo, $2{,}131313\ldots = 2{,}(13)$ es una fracción periódica pura y $2{,}41313\ldots = 2{,}4(13)$, una fracción periódica mixta.

La partición de un conjunto de fracciones decimales en subconjuntos de fracciones finitas e infinitas es bastante condicional, puesto que cualquier fracción decimal finita (menos el número cero) puede ser escrita como una fracción periódica infinita. Esto se puede hacer con un procedimiento evidente, por ejemplo, representando una fracción decimal finita $0{,}25$ como $0{,}25000\ldots 0 \ldots = 0{,}25(0)$, es decir añadiendo a la fracción finita una cantidad infinita de ceros.

Las fracciones decimales infinitas, cuyo período consta no sólo de una cifra 9, se llaman *fracciones decimales admisibles*.

En un conjunto de fracciones decimales admisibles puede ser introducida la relación del orden, es decir, el conjunto de todas las fracciones decimales admisibles es un conjunto ordenado.

Las dos fracciones decimales infinitas admisibles

$$N,\ n_1n_2n_3\ldots n_k\ldots \text{ y } M,\ m_1m_2m_3\ldots m_k\ldots$$

se consideran *iguales*, si la parte entera de la primera fracción es igual a la parte entera de la segunda fracción ($N=M$) y las cifras que se encuentran en las mismas posiciones de estas fracciones son iguales:

$$n_1=m_1,\quad n_2=m_2,\ \ldots,\quad n_k=m_k,\ \ldots$$

La fracción decimal infinita admisible $0,\ n_1n_2n_3\ldots$ $\ldots n_k\ldots$ se considera menor que la fracción decimal infinita admisible $0,\ m_1m_2m_3\ldots m_k\ldots$

$$0,\ n_1n_2n_3\ldots n_k\ldots < 0,\ m_1m_2m_3\ldots m_k\ldots$$

si se halla tal número k para el cual

$$n_1=m_1,\ n_2=m_2,\ \ldots,\ n_{k-1}=m_{k-1},\ \text{pero } n_k<m_k,$$

es decir, la comparación de las fracciones decimales se efectúa por el primer par de cifras desiguales. Por ejemplo, la fracción 0,72159. . . es menor que la fracción 0,72160. . ., puesto que las tres primeras cifras que se encuentran después del punto coinciden en ambas fracciones y la cuarta cifra de la primera fracción es menor que la cuarta cifra de la segunda fracción.

La fracción decimal admisible $N,\ n_1n_2n_3\ldots n_k\ldots$ se considera menor que la fracción decimal admisible $M,\ m_1m_2m_3\ldots m_k\ldots$, si la parte entera de la primera fracción es menor que la parte entera de la segunda fracción ($N<M$), o, si las partes enteras de estas dos fracciones son iguales ($N=M$), la fracción decimal $0,\ n_1n_2n_3\ldots n_k\ldots$ es menor que la fracción decimal $0,\ m_1m_2m_3\ldots m_k\ldots$

La comparación de dos fracciones finitas y también de una fracción finita y una fracción decimal infinita admisible se efectúa por las mismas reglas que la comparación de dos fracciones decimales infinitas admisibles.

5.3. Operaciones aritméticas con las fracciones decimales finitas. Antes de empezar la descripción de las reglas de las operaciones aritméticas con las fracciones decimales es necesario hacer una advertencia. Vamos a escribir las fracciones decimales negativas en la forma que hemos aceptado en el p. 4.3, es decir, vamos a considerar que si un número positivo a puede ser escrito en forma de una fracción decimal finita

$$N_0, n_1 n_2 \ldots n_k,$$

entonces el número negativo a se escribe en forma de fracción decimal negativa

$$-N_0, n_1 n_2 \ldots n_k.$$

Por ejemplo, un número negativo $-\frac{37}{25}$ se escribe $-1,48$ y no $\bar{2},52$.

La utilización de tal forma para escribir las fracciones decimales negativas permite realizar operaciones aritméticas con las fracciones negativas en muchos casos análogamente a las operaciones aritméticas con los números enteros negativos.

La adición y sustracción de fracciones decimales finitas se efectúa según las mismas reglas que la adición y sustracción de números enteros; sólo es necesario escribir cada orden de una fracción bajo el orden de la misma denominación de la segunda fracción y en lugar de los órdenes que faltan escribir ceros.

Ejemplo 1. Sumar las fracciones 2,14 y 0,151.

Escribamos la fracción 2,14 en la forma 2,140 y realicemos la adición de números enteros 2140 y 151:

$$2140 + 151 = 2291;$$

separando a la derecha del número entero 2291 el mismo número de signos que fue separado en las fracciones 2,140 y 0,151 (precisamente tres signos), de resultas obtendremos la fracción 2,291, la cual es la suma de las fracciones dadas.

La multiplicación de las fracciones decimales finitas se efectúa de la manera siguiente: sin fijar la atención en las comas, las fracciones finitas se multiplican como los números enteros; en el producto obtenido separan a partir de la derecha un número de signos igual a la suma del número de signos después de la coma para todos los factores.

Ejemplo 2. Hallar el producto de los fracciones decimales finitas 2,1 y 0,27.

Multiplicando los números enteros 21 y 27 (se puede despreciar el cero en el número 0,27), obtendremos el número 567. Separando de la derecha tres signos, obtendremos que el producto de las dos fracciones dadas es igual a 0,567.

La división de una fracción decimal finita entre un número entero se efectúa de la manera siguiente:

1) Si el dividendo es menor que el divisor, escribimos en la parte entera del cociente un cero y después la coma. A continuación, sin fijar la atención en la coma del dividendo añadimos a la parte entera del dividendo la primera cifra de su parte fraccionaria; si después de tal adición se obtiene un número que es menor que el divisor, esribimos en el cociente después de la coma un cero y añadimos al número anterior obtenido la cifra siguiente del dividendo; si después de esta acción obtenemos un número menor que el divisor, colocamos otro cero más, etc., hasta obtener un número mayor que el divisor. En adelante la división se efectúa de la misma manera que con los números enteros, al mismo tiempo, el dividendo se puede «extender» infinitamente hacia la derecha del punto, escribiendo ceros al final.

Ejemplo 3. Dividir la fracción decimal 0,525 entre 2.

Según la regla citada el cálculo del cociente va a tener el siguente aspecto:

$$\begin{array}{r|l}
0{,}525 & 2 \\ \cline{2-2}
-5 & 0{,}2625 \\
4 & \\ \cline{1-1}
-12 & \\
12 & \\ \cline{1-1}
-5 & \\
4 & \\ \cline{1-1}
-10 & \\
10 & \\ \cline{1-1}
0 &
\end{array}$$

Puede suceder que el proceso de división no termine nunca. En tal caso el cociente no se puede expresar mediante una fracción decimal finita.

2) Si el dividendo es mayor que el divisor, dividimos primeramente la parte entera, escribimos en el cociente el resultado de la división y ponemos el punto. Después continuamos la división como en el caso anterior.

La división de una fracción decimal finita entre una fracción decimal finita se efectúa según la regla siguiente:

Para dividir una fracción decimal (o un número entero) entre una fracción decimal, omitimos la coma en el divisor; pasamos del dividendo hacia la derecha en tantos signos, cuantos contiene la parte fraccional del divisor (en caso de necesidad añadimos al dividendo algunos ceros al final). Después realizamos la división de una fracción decimal entre un número entero de la manera descrita anteriormente.

Así, por ejemplo, la división de la fracción decimal 0,525 entre la fracción 0,2 se reduce a la división de la fracción decimal 5,25 entre 2.

Las operaciones aritméticas con las fracciones decimales periódicas infinitas son bastante complicadas y voluminosas, por eso es mucho más cómodo hacer lo siguiente: transformar las fracciones decimales periódicas infinitas, con las cuales tenemos que realizar operaciones aritméticas, en fracciones racionales; efectuar con éstas las operaciones necesarias; si es necesario, transformar en decimal la fracción obtenida como resultado de los cálculos.

5.4. Conversión de una fracción decimal finita en fracción racional. Sea N, $n_1\, n_2\, \ldots\, n_k$ una fracción decimal finita; N, la parte entera de esta fracción y $n_1, n_2, \ldots, n_k$, sus cifras significativas. Construyamos una fracción racional, cuyo numerador es el número $n_1 n_2 \ldots n_k$ *) y el denominador, el número 10^k, es decir, la fracción

$$\frac{n_1 n_2 \ldots n_k}{10^k}$$

Añadimos a ésta la parte entera de la fracción decimal dada. De resultas obtenemos la fracción racional buscada.

Ejemplo. La fracción decimal finita 2,135 se escribe en forma de la siguiente fracción racional:

$$2{,}135=2+0{,}135=2+\frac{135}{10^3}=\frac{2135}{1000}$$

*) Aquí tenemos una escritura, y no una multiplicación.

y la fracción decimal finita 1,091, en la forma de

$$1,091 = -1 + 0,091 = -1 + \frac{91}{10^3} = -\frac{909}{1000}.$$

5.5. Conversión de una fracción periódica infinita en fracción racional.

1) Representemos la fracción decimal periódica infinita $0,\ n_1n_2 \ldots n_k\,(n_{k+1} \ldots n_{k+p})$ como la suma de una fracción decimal finita y otra periódica infinita:

$$0,\ n_1n_2 \ldots n_k\,(n_{k+1} \ldots n_{k+p}) =$$
$$= 0,\ n_1n_2 \ldots n_k + 0,\ \underbrace{0 \ldots 0}_{k}\,(n_{k+1} \ldots n_{k+p}).$$

2) Representemos la fracción periódica infinita $0,\ \underbrace{0 \ldots 0}_{k}$ $(n_{k+1} \ldots n_{k+p})$ en forma del producto:

$$0,\ \underbrace{00 \ldots 0}_{k}\,(n_{k+1} \ldots n_{k+p}) = \frac{1}{10^k} \cdot 0,\ (n_{k+1}n_{k+2} \ldots n_{k+p}). \quad (4)$$

3) Escribamos la fracción periódica pura $0,\ (n_{k+1}n_{k+2} \ldots$ $\ldots n_{k+p})$:

$$0,\ (n_{k+1}n_{k+2} \ldots n_{k+p}) = \frac{n_{k+1}n_{k+2} \ldots n_{k+p}}{10^p} +$$
$$+ \frac{n_{k+1}n_{k+2} \ldots n_{k+p}}{10^{2p}} + \ldots = \frac{n_{k+1}n_{k+2} \ldots n_{k+p}}{10^p} \times$$
$$\times \left[1 + \frac{1}{10^p} + \frac{1}{10^{2p}} + \ldots\right].$$

4) Calculemos la suma de los términos de la progresión geométrica decreciente infinita que se encuentran entre corchetes. La suma es igual a

$$\frac{10^p}{10^p - 1}$$

y, por consiguiente, la fracción periódica pura $0,\ (n_{k+1}n_{k+2} \ldots$ $\ldots n_{k+p})$ se escribe en forma de la fracción racional siguiente:

$$0,\ (n_{k+1}n_{k+2}\ldots n_{k+p})=\frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p}\times$$

$$\times\frac{10^p}{10^p-1}=\frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p-1}.$$

Ahora, teniendo en cuenta la igualdad (4), es muy fácil escribir la fracción decimal $0,\ n_1n_2\ldots n_k\ (n_{k+1}n_{k+2}\ldots n_{k+p})$ en forma de fracción racional:

$$0,\ n_1n_2\ldots n_k\ (n_{k+1}n_{k+2}\ldots n_{k+p})=$$

$$=0,\ n_1n_2\ldots n_k+\frac{1}{10^k}\cdot\frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p-1}=$$

$$=\frac{n_1n_2\ldots n_kn_{k+1}n_{k+2}\ldots n_{k+p}-n_1n_2\ldots n_k}{10^k\,(10^p-1)}.$$

De tal manera, la regla de conversión de una fracción decimal periódica mixta en una fracción racional se puede formular así:

Para convertir una fracción decimal periódica mixta en otra racional, es necesario del número formado por las cifras que se encuentran detrás de la coma antes del comienzo del segundo período restar el número formado por las cifras que se encuentran detrás de la coma y delante del primer período; tomar la diferencia obtenida como nominador de la fracción y en el denominador escribir la cifra nueve tantas veces, cuantas cifras haya en el período y con tantos ceros, cuantas cifras haya entre el punto y el período.

Ejemplos. 1. Escribir la fracción periódica pura 2, (13) en forma de fracción racional.

Representemos esta fracción en la forma

$$2,\ (13)=2+0,\ (13)=2+\frac{13}{100}+\frac{13}{10\,000}+\ldots=$$

$$=2+\frac{13}{10^2}+\frac{13}{10^4}+\ldots=2+\frac{13}{10^2}\left(1+\frac{1}{10^2}+\frac{1}{10^4}+\ldots\right).$$

Entre paréntesis se encuentra la suma de los términos de una progresión geométrica decreciente infinita con el primer término igual a 1 y el denominador $q=\frac{1}{10^2}$.

Si utilizamos la fórmula para la suma de los términos de la progresión geométrica decreciente infinita, obtenemos

$$2,(13)=2+\frac{13}{10^2}\cdot\frac{1}{1-\frac{1}{10^2}}=2+\frac{13}{10^2}\cdot\frac{10^2}{10^2-1}=2+\frac{13}{99}=\frac{211}{99}.$$

De esta manera, la fracción decimal periódica 2, (13) puede ser escrita en forma de la fracción racional $\frac{211}{99}$.

2. Escribir la fracción periódica mixta 2,5 (13) en forma de una fracción racional.

Representemos la fracción 2,5 (13) en la forma siguiente:

$$2{,}5\,(13)=2+0{,}5+0{,}0\,(13)=2+0{,}5+\frac{1}{10}\cdot 0,\,(13).$$

Después es necesario escribir la fracción periódica pura 0, (13) en forma de fracción racional análogamente a como se hizo en el ejemplo 1 y efectuar la adición.

§ 6. Fracciones continuas

El algoritmo de Euclides (véase p. 1.5) por el que se halla el máximo común divisor de dos números naturales conduce a un procedimiento bastante interesante de representación de los números racionales.

Por ejemplo, la aplicación del algoritmo de Euclides a los números 840 y 611 da la siguiente serie de igualdades:

$$840=1\cdot 611+229,$$
$$611=2\cdot 229+153,$$
$$229=1\cdot 153+76,$$
$$153=2\cdot 76+1$$

ó

$$\frac{840}{611}=1+\frac{229}{611}=1+\frac{1}{\frac{611}{229}},\qquad \frac{229}{153}=1+\frac{76}{153}=1+\frac{1}{\frac{153}{76}},$$

$$\frac{611}{229}=2+\frac{153}{229}=2+\frac{1}{\frac{229}{153}},\qquad \frac{153}{76}=2+\frac{1}{76}.$$

Combinando las últimas igualdades, llegamos a la representación del número racional $\frac{840}{611}$ en la forma siguiente:

$$\frac{840}{611}=1+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{76}}}}$$

Examinemos ahora la cadena de igualdades del algoritmo de Euclides para la obtención del máximo común divisor de los números naturales m y n:

$$\begin{aligned} m &= n\cdot m_1 + r_1, \\ n &= r_1\cdot m_2 + r_2, \\ r_1 &= r_2\cdot m_3 + r_3, \\ &\dots\dots\dots\dots \\ r_{k-2} &= r_{k-1}\cdot m_k + r_k, \\ r_{k-1} &= r_k\cdot m_{k+1}, \end{aligned}$$

donde $n > r_1 > r_2 > r_3 > \ldots > r_k > 0$.

Reescribimos esta cadena de igualdades en la forma

$$\frac{m}{n}=m_1+\frac{r_1}{n},$$

$$\frac{n_1}{r_1}=m_2+\frac{r_2}{r_1},$$

$$\frac{r_1}{r_2}=m_3+\frac{r_3}{r_2},$$

$$\dots\dots\dots\dots$$

$$\frac{r_{k-2}}{r_{h-1}}=m_k+\frac{r_k}{r_{k-1}},$$

$$\frac{r_{k-1}}{r_k}=m_{k+1}.$$

Cada una de las igualdades citadas (a excepción de la última) describe la eliminación de una parte entera de una fracción impropia, es decir, representa una fracción impropia en forma de la suma de un número natural y de cierta fracción propia. Anotando que el miembro primero de cada igualdad es una magnitud inversa al segundo sumando del miembro segundo de la igualdad anterior, se puede escribir el número racional $\frac{m}{n}$, utilizando sólo los números m_1, m_2, m_3, ...

$\ldots, m_k, m_{k+1}$:

$$\frac{m}{n} = m_1 + \cfrac{1}{m_2 + \cfrac{1}{m_3 + \cfrac{1}{m_4 + \cdot \atop {\cdot \atop {\cdot \atop + \frac{1}{m_{k+1}}}}}}}, \tag{1}$$

donde $m_2, m_3, m_4, \ldots, m_{k+1}$ son números naturales y m_1, un número natural o cero.

La expresión que tiene la forma (1) se llama *fracción continua finita* (o *fracción de cadena*).

Si un número racional: $\frac{m}{n}$ es negativo $(m < 0, n > 0)$, entonces para escribir el número $\frac{m}{n}$ en forma de fracción continua se utiliza el método siguiente.

Representan el número racional negativo $\frac{m}{n}$ en la forma

$$\frac{m}{n} = M + \frac{k}{n},$$

donde M es un número negativo y k, un número natural. Se escribe la fracción racional positiva $\frac{k}{n}$ en forma de fracción continua:

$$\frac{k}{n} = \cfrac{1}{l_1 + \cfrac{1}{l_2 + \cfrac{1}{l_3 + \cdot \atop {\cdot \atop {\cdot \atop + \frac{1}{l_k}}}}}}.$$

Entonces la fracción racional negativa $\frac{m}{n}$ se escribe en forma de

$$\frac{m}{n} = M + \cfrac{1}{l_1 + \cfrac{1}{l_2 + \cfrac{1}{l_3 + \cdot \atop {\cdot \atop {\cdot \atop + \frac{1}{l_k}}}}}}$$

La fracción continua que se encuentra en el miembro segundo de la igualdad (1) se designa

$$[m_1;\ m_2,\ m_3,\ \ldots,\ m_{k+1}], \tag{2}$$

donde el punto y coma separa la parte entera del número racional de su parte fraccional.

La fracción continua

$$[m_1;\ m_2,\ m_3,\ \ldots,\ m_s] \quad (s \leqslant k+1)$$

se llama *fracción conveniente* de s-ésimo orden para la fracción continua (2) y se designa brevemente $\frac{p_s}{q_s}$.

Una fracción continua conveniente se llama fracción par, si s es par, e impar, si s es impar.

Todas las fracciones del orden par son menores que $\frac{m}{n}$ y su valor crece con el crecimiento del orden; todas las fracciones convenientes del orden impar son mayores que $\frac{m}{n}$ y su valor decrece con el crecimiento del orden:

$$\frac{p_s}{q_s} < \frac{m}{n} < \frac{p_{s+1}}{q_{s+1}} \quad (s \text{ es par}).$$

El módulo de diferencia entre las fracciones convenientes $\frac{p_s}{q_s}$ y $\frac{p_{s+1}}{q_{s+1}}$ es igual a $\frac{1}{p_s \cdot q_{s+1}}$ y por eso la diferencia entre la fracción conveniente de s-ésimo orden y el número $\frac{m}{n}$ se estima por la desigualdad

$$\left|\frac{m}{n} - \frac{p_s}{q_s}\right| < \frac{1}{q_s \cdot q_{s+1}}. \tag{3}$$

La teoría de las fracciones continuas históricamente surgió de la necesidad de representar aproximadamente una fracción racional, el numerador y el denominador de la cual son inmensamente grandes, mediante otra fracción racional considerablemente menor, así como de estimar el error de esta aproximación. Sean m y n unos números grandes iguales, por ejemplo, a 1355 y 946. Es necesario sustituir la fracción $\frac{1355}{946}$, de manera aproximada, por la fracción $\frac{p_k}{q_k}$, donde q_k debe ser menos de 100 y también estimar el error de esta aproximación. Representemos la fracción $\frac{1355}{946}$ en forma de una fracción continua y calculemos las fracciones convenientes. Al efectuar esto, vemos que

$$\frac{m}{n} = \frac{1355}{946} = [1;\ 2,\ 3,\ 5,\ 8,\ 3],$$

$$\frac{p_3}{q_3} = \frac{53}{37}, \qquad \frac{p_4}{q_4} = \frac{434}{303}.$$

A nuestra condición (el denominador <100) satisface la fracción $\frac{53}{37}$.

La desigualdad (3) permite simplemente estimar el error absoluto entre la fracción dada y la fra ción conveniente de tercer orden:

$$\left|\frac{1355}{946}-\frac{53}{37}\right|<\frac{1}{q_3\cdot q_4}=\frac{1}{11\,211}<0{,}0001.$$

En la teoría de las fracciones continuas se introduce también el concepto de fracción continua infinita y se demuestra que cualquier número real puede ser escrito únicamente en forma de fracción continua. Para esto los números racionales se escriben en forma de fracciones finitas y los números irracionales se escriben en forma de fracciones continuas infinitas.

Observemos que tanto en el caso de a racional, como tambien en el caso de a irracional, las fracciones continuas convenientes tienen tal propiedad que proporcionan una«mejor» aproximación de cualquier número real a un número racional, es decir:

Para cualquier fracción conveniente $\frac{p_k}{q_k}$ de k-ésimo orden es imposible hallar una fracción racional con un denominador menor que q_k, tal que ella dé una mayor aproximación del número real a que la fracción conveniente $\frac{p_k}{q_k}$.

§ 7. Procedimientos de cálculos

7.1. Valor aproximado de un número y errores. Para efectuar los cálculos es muy cómodo escribir los números reales en forma de fracciones decimales. Sin embargo, no cualquier número racional y tanto más un número real pueden ser escritos en forma de fracción finita, con la cual es muy cómodo efectuar los cálculos necesarios. Por eso muy a menudo nos vemos obligados a interrumpir la fracción decimal infinita, que representa la escritura del número real dado, en cierta cifra detrás de la coma y operar con la fracción decimal finita obtenida. El número representado por la fracción decimal finita que obtuvimos como resultado de la interrupción de una fracción decimal infinita se llama *valor aproximado* del número dado. Por ejemplo, diversos valores aproximadas del número $\pi = 3{,}14159\ldots$ son los números 3,14; 3,1415, etc.; se pueden tomar como valores aproximados del número $e = 2{,}71828\ldots$ los números 2,7; 2,71; 2,7182, etc. Los valores aproximados de los números pueden aparecer no sólo de resultas de la interrupción de la escritura del número como fracción decimal. Para la solución de cualquier problema práctico utilizamos los números obtenidos como resultado de ciertas mediciones. Sin embargo, durante cualesquiera mediciones surge un error debido a la imperfección de nuestros instrumentos de medición, forma incorrecta geométrica del objeto que medimos, etc.

Designemos el valor aproximado del número a a través de a^*.

El valor absoluto de la diferencia entre el número a y su valor aproximado a^*

$$| a - a^* |$$

se denomina *error absoluto* del valor aproximado a^*.

El error absoluto es una de las características del valor aproximado de un número. Muy a menudo este error representa sólo interés teórico, puesto que en la mayoría de los casos no conocemos el número a. En una serie de casos se pueden indicar los límites, dentro de los cuales varía el error absoluto. Estos límites suelen definirse mediante el procedimiento de obtención del valor aproximado del número. Por ejemplo, efectuando las mediciones de una longitud mediante una regla con una graduación de 1 mm, podemos garantizar que el error absoluto de las mediciones no va a superar 0,5 mm.

La posibilidad de estimar los límites de la variación del error absoluto de un valor aproximado de un número es la causa de la introducción de una estimación más comoda del valor aproximado de un número, del error límite absoluto.

El número A tal que

$$| a - a^* | < A$$

se llama *error límite absoluto* del valor aproximado de un número. Según su definición, el error absoluto puede ser polidígito. Así, en el caso examinado más arriba, en el ejemplo de la medición de longitudes con una regla, podemos tomar el error absoluto igual a 1 mm, 1,5 mm, etc.

Observemos que ni el error absoluto ni el error límite absoluto caracterizan por sí mismo la precisión del resultado, si no está indicado el resultado mismo. Por ejemplo, sea el error límite absoluto del resultado de cierta medición igual a 10 km. Si medimos la distancia entre dos ciudades, tal precisión no es satisfactoria. Si medimos la distancia entre los polos de la Tierra, entonces la precisión indicada es muy buena.

Los valores llamados error relativo y error límite relativo son unas características de la precisión del resultado de la medición, en las cuales junto con el error participa también el resultado de la medición.

La relación entre el error absoluto y la magnitud absoluta del valor aproximado del número:

$$\delta = \frac{|a - a^*|}{|a^*|}$$

se llama *error relativo.*

La relación entre el error límite absoluto y la magnitud absoluta del valor aproximado del número

$$\Delta = \frac{A}{|a^*|}$$

se llama *error límite relativo* del valor aproximado de un número.

Por ejemplo, el número 3,14 es un valor aproximado del número π. La magnitud absoluta de este valor aproximado es igual a 0,00159. . .; el error límite absoluto podemos estimarlo igual a 0,0016 y el error límite relativo

$$\frac{0{,}0016}{3{,}14} = 0{,}000509\ \ldots$$

podemos suponerlo igual a 0,00051.

7.2. Anotación decimal de los valores aproximados de un número. Cualquier número real positivo x puede ser escrito en forma de fracción decimal finita o infinita:

$$a = n_l 10^l + n_{l-1}\cdot 10^{l-1} + \ldots + n_{l-k+1}\cdot 10^{l-k+1} + \ldots,$$

donde n_l son las cifras del número a ($n_l = 0, 1, 2, \ldots, 9$), al mismo tiempo $n_l \neq 0$ y l es cierto número entero. Por ejemplo,

$$523{,}14\ldots = 5\cdot 10^2 + 2\cdot 10^1 + 3\cdot 10^0 + 1\cdot 10^{-1} + \\ + 4\cdot 10^{-2} + \ldots$$

Se puede obtener el valor aproximado a^* del número a, interrumpiendo la anotación decimal del número a:

$$a^* = n_l \cdot 10^l + n_{l-1} \cdot 10^{l-1} + \ldots + n_{l-k+1} \cdot 10^{l-k+1}.$$

Para escribir el número a^* en el sistema posicional decimal de numeración a veces tenemos que introducir ceros que sobran en el comienzo o en el final del número. Cualquier cifra de la anotación decimal del cero que se diferencia de éste y el cero si se encuentra entre las cifras significativas o si es un representante del orden decimal conservado, se llama *cifra significativa* del valor aproximado de un número. Por ejemplo, como valor aproximado del número

$$b = 2 \cdot 10^{-2} + 4 \cdot 10^{-3} + 0 \cdot 10^{-4} + 9 \cdot 10^{-5} + \ldots$$

se puede tomar el número

$$b^* = 0,\boxed{0}2409$$

y como valor aproximado del número

$$c = 1 \cdot 10^5 + 0 \cdot 10^4 + 7 \cdot 10^3 + 3 \cdot 10^2 + \ldots$$

se puede tomar

$$c^* = 1073\boxed{00}.$$

El cero que se encuentra delante de la coma en la escritura del número b^* y los ceros que se encuentran al final del número c^* no se consideran cifras significativas (en la escritura de los números b^* y c^* los ceros indicados se encuentran dentro de los cuadrados).

Sea, por ejemplo, el número 0,002080 un valor aproximado de cierto número. Los tres primeros ceros no son cifras significativas, puesto que sirven sólo para indicar el orden decimal de otras cifras. Los otros dos ceros son cifras significativas, puesto que el primero de ellos se encuentra entre las cifras significativas 2 y 8 y el segundo indica que el valor aproximado del número fue obtenido mediante la interrupción de la escritura decimal del número, conservando el sexto orden detrás de la coma. En el caso, si en el valor aproximado 0,002080 la última cifra no es significativa, este valor debe ser escrito en forma de 0,20208. De esta manera, los valores aproximados 0,00208 y 0,00208 no son equivalentes, puesto que el primero de ellos contiene cuatro cifras significativas y el segundo, tres.

Para la anotación de valores aproximados de números enteros, los ceros al final de la escritura decimal de un número pueden servir tanto para la designación de cifras significativas como también para la indicación del orden de las otras cifras. Por eso la inscripción común de los valores aproximados de números enteros puede entenderse polidígitamente. Por ejemplo, sea el número 134 000 un valor aproximado de cierto número. A primera vista no se puede definir cuantas cifras significativas contiene este número (aunque sí se puede decir que son no menos de tres). Para evitar tal indeterminación, el número entero se escribe mediante fracciones decimales y potencias del número 10. Por ejemplo, si se quiere indicar que el número 134000 tiene tres cifras significativas, entonces se escribe en forma de $1{,}34\cdot 10^5$ (ó $13{,}4\cdot 10^4$, ó $134\cdot 10^3$). Si se quiere indicar que el número 134000 tiene cinco cifras significativas, entonces se escribe en forma de $1{,}3400\cdot 10^5$ (ó $13{,}400\cdot 10^4$, $134{,}00\cdot 10^3$ respectivamente).

Es conocido que las primeras n cifras significativas del valor aproximado de un número son *justas*, si el error absoluto del valor aproximado del número no supera la mitad de la unidad del orden de la n-ésima cifra significativa, contando los órdenes de izquierda a derecha.

De esta manera, si para el valor aproximado a^* del número

$$a = n_l\cdot 10^l + n_{l-1}\cdot 10^{l-1} + \ldots + n_{l-k+1}\cdot 10^{l-k+1} + \ldots$$

se sabe que

$$|a - a^*| \leqslant 0{,}5\cdot 10^{l-k+1},$$

entonces las primeras k cifras n_l, n_{l-1}, . . ., n_{l-k+1} del valor aproximado a^* son justas.

Por ejemplo, para 121,57 el número 122,00 es un valor aproximado con tres cifras justas, puesto que

$$|121{,}57 - 122| = 0{,}43 < 0{,}5\cdot 10^0.$$

El resultado de las operaciones aritméticas con valores aproximados de números representa también un valor aproximado de cierto número. El error del resultado obtenido se estima mediante las reglas siguientes:

1) El error límite absoluto de una suma algebraica de los valores aproximados de números no es mayor que la suma de los errores límites absolutos de los sumandos.

2) El error relativo de una suma de valores aproximados de números se encuentra entre el error máximo y mínimo de los errores relativos de los sumandos.

Conforme a las reglas formuladas, con el fin de evitar cálculos excesivos, antes de calcular la suma de los valores aproximados de números se efectúa un redondeo previo de los sumandos. (véase p. 7.3). Para esto el sumando que tenga el número mínimo de signos exactos se deja sin redondear y en los demás casos dejan uno o dos signos más que en el número no redondeado.

7.3. Redondeo de números. Para redondear un número hasta la cantidad n de cifras significativas se omiten todas las cifras que se encuentran detrás del n-ésimo orden, o, si esto es necesario para conservar los órdenes, se sustituyen por ceros. Para esto:

1) Si a la última cifra que conservamos le sigue la cifra 0, 1, 2, 3 ó 4, entonces para el redondeo se conservan todas las cifras, inclusive la última que conservamos. Tal redondeo se llama *redondeo por defecto.*

2) Si a la última cifra que conservamos le sigue 9, 8, 7, 6 ó 5, entonces a la última cifra que conservamos se le añade una unidad. Tal redondeo se llama *redondeo por exceso.*

Ejemplos. 1. El redondeo del número 219,364 hasta tres cifras significativas da el número 219,4. La última cifra del número redondeado se aumenta en una unidad, puesto que la primera cifra que omitimos es mayor que cinco.

2. El redondeo del número 6,4502 hasta dos cifras significativas da el número 6,5. La última cifra del número redondeado fue aumentada en una unidad, puesto que la primera cifra que omitimos es igual a 5.

3. El redondeo del número 1836 hasta tres cifras significativas da el número $1{,}84 \cdot 10^3$.

Utilizando las reglas del redondeo vemos que un error absoluto de redondeo no supera la mitad de la unidad del orden de la última cifra significativa que dejamos, es decir, todas las cifras significativas de un valor aproximado del número, obtenido mediante las reglas del redondeo, son cifras exactas.

7.4. Método de las tangentes (método de Newton). El método *de las tangentes* es uno de los métodos numéricos más efectivos para la resolución de la ecuación.

Supongamos se sabe que cierta función $f(x)$ tiene su raíz en un intervalo $[a; b]$ (la función $f(x)$ es diferenciable en $[a; b]$ y su derivada no se anula en $[a; b]$). Se puede hallar esta raíz mediante el siguiente procedimiento. Tomemos un punto arbitrario $x_1 \in [a; b]$ y escribamos la ecuación de la tangente a la gráfica de la función $f(x)$ en el punto cuya asbscisa*) es x_1:

$$y = f(x_1) + f'(x_1)(x - x_1).$$

Resolviendo la ecuación

$$f(x_1) + f'(x_1)(x - x_1) = 0,$$

hallemos el punto de intersección de la tangente con el eje Ox:

$$x = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Designemos el punto hallado a través de x_2:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Escribamos la ecuación de la tangente a la gráfica de la función $f(x)$ en el punto cuya abscisa es x_2:

$$y = f(x_2) + f'(x_2)(x - x_2),$$

y hallemos el punto de intersección de esta tangente con el eje Ox:

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}.$$

Siguiendo este proceso, obtendremos la sucesión (x_n), que se representa mediante la fórmula recurrente

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \in \mathbf{N}. \tag{1}$$

Si la función $f(x)$ satisface algunas condiciones adicionales (por ejemplo, la segunda derivada de $f(x)$ es positiva (o negativa) en todos los puntos interiores $[a; b]$ y la raíz de la ecuación $f(x) = 0$ está situada dentro de $[a; b]$), entonces la sucesión (1) converge en el número coincidente con la raíz de la ecuación $f(x) = 0$.

El método de las tangentes puede ser utilizado para la resolución del problema de la extracción de la raíz de un número positivo

---

*) Es necesario que el punto x_0 sea ubicado lo suficientemente cerca de la raíz de la ecuación $f(x) = 0$, por eso para eligir el punto x_0 a veces es necesario efectuar una investigación previa de la función $f(x)$.

arbitrario a. Vamos a buscar el valor $\sqrt{a}$ como solución de la ecuación

$$f(x) = x^2 - a = 0.$$

La fórmula recurrente (1) en el caso dado adquiere la forma

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

Aquí como primer término de la sucesión (x_n) se puede tomar cualquier número $x_1 \in (0; +\infty)$.

Los errores absolutos δ_n de los valores aproximados x_n del número $\sqrt{a}$

$$\delta_n = x_n - \sqrt{a}$$

se relacionan mediante la desigualdad

$$\delta_{n+1} \leqslant \frac{\delta_n^2}{2\sqrt{a}}. \tag{2}$$

Esta estimación recurrente del error absoluto confirma la alta velocidad de convergencia de la sucesión de valores aproximados x_n.

Por ejemplo, calculemos mediante el método de las tangentes $\sqrt{33\,520}$. Por x_1 tomemos el número 200. Este es el valor de la raíz dada con exceso. El error absoluto δ_1 no es mayor que 20. Entonces según la fórmula (2)

$$\delta_2 < \frac{400}{2 \cdot 180} < 1{,}5; \qquad \delta_3 < \frac{2{,}25}{2 \cdot 180} < 0{,}003.$$

Aquí, para simplificar las estimaciones, $\sqrt{a}$ en la fórmula (2) está sustituido por un número menor: 180. De esta manera, para calcular el valor $\sqrt{33\,520}$ ya después del tercer paso se obtiene un valor aproximado de la raíz que se diferencia de su valor exacto menos que en 0,003.

7.5 Algoritmo de la extración de la raíz cuadrada de un número natural. Con mucha frecuencia para la resolución de las ecuaciones cuadradas es necesario extraer la raíz cuadrada de un número natural. Pongamos como ejemplo el algoritmo de la extracción de la raíz cuadrada para el caso, cuando el número que se encuentra bajo el signo de la raíz es el cuadrado de un número natural. Supongamos que es necesario calcular $\sqrt{33\,489}$. El número 33 489 lo partimos en grupos de cifras (de dos en dos), de derecha a izquierda:

3 34 89.

Buscamos el número más grande, cuyo cuadrado no supere el número 3 que es el que se encuentra en el primer grupo de cifras. Este número es el número 1. Lo escribimos en la respuesta. Elevamos la unidad al cuadrado y lo sustraemos del número 3. A la diferencia obtenida le añadimos el segundo grupo de cifras. Obtenemos el número 234.

Doblamos el número escrito en la respuesta (en nuestro caso el 1) y añadimos a la derecha del número obtenido la cifra máxima cuyo producto de la multiplicación del número de dos signos obtenido por esta cifra no supere 234. En nuestro caso esta cifra es la cifra 8:

$$28 \cdot 8 = 224 < 234.$$

Escribimos la cifra 8 detrás de la cifra 1 en la respuesta. Del número 234 sustraemos el número 224 y a la diferencia obtenida le añadimos el último grupo de cifras. Obtenemos el número 1089.

Doblamos el número escrito en la respuesta (en nuestro caso 18) y añadimos a la derecha del número obtenido (en nuestro caso el número 36) la cifra máxima cuyo producto de la multiplicación del número de tres cifras obtenido por la cifra añadida no supere el número 1089. En nuestro caso ésta es la cifra 3:

$$363 \cdot 3 = 1089.$$

Escribimos la cifra 3 en la respuesta. El proceso de extracción de la raíz cuadrada está terminado:

$$\sqrt{33\,489} = 183.$$

El proceso que describimos suele escribirse en la forma expuesta en el esquema siguiente:

$$\sqrt{-3\;34\;89} = 183$$

$$\begin{array}{r|r}
 & 1 \\ \hline
\times\; 28 & -\;234 \\
8 & 224 \\ \cline{1-1}
224 & \\ \hline
\times\; 363 & -\;1089 \\
3 & 1089 \\ \cline{1-1}
1089 & \\ \cline{2-2}
 & 0
\end{array}$$

CAPÍTULO 3

Números complejos

Históricamente la introducción de números complejos se relaciona con la obtención de la fórmula analítica para las raíces de la ecuación cúbica:

$$x^3 = px + q. \tag{1}$$

En el primer tercio del siglo XVI el matemático italiano N. Tartaglia mostró que la raíz de esta ecuación se representa siempre por la expresión

$$x = \sqrt[3]{u} + \sqrt[3]{v}, \tag{2}$$

donde u y v son las soluciones del sistema de ecuaciones

$$\begin{aligned} u+v &= q, \\ uv &= \left(\frac{p}{3}\right)^3. \end{aligned} \tag{3}$$

Así, por ejemplo, si es necesario hallar la raíz de la ecuación cúbica

$$x^3 = 9x + 28,$$

componiendo el sistema (3) para la ecuación dada y resolviéndola, obtenemos dos soluciones del sistema: $u = 27$, $v = 1$ y $u = 1$, $v = 27$.

Utilizando la relación (2), obtenemos $x = 4$, es decir, el número 4 es la raíz de la ecuación dada.

Sin embargo, resultó que existen ecuaciones cúbicas, para las cuales el sistema (3) no tiene soluciones en el conjunto de números reales, mientras que es conocido que la ecuación cúbica tiene una raíz real. Por ejemplo, la ecuación

$$x^3 = 15x + 4$$

tiene la raíz real $x = 4$, de lo que es muy fácil convencerse, sustituyendo en la ecuación dada x por el número 4. Si para la ecuación dada escribimos el sistema (3), entonces resulta que este sistema no tiene soluciones en el conjunto de números reales.

Este hecho incomprensible en aquellos tiempos lo explicó por primera vez el matemático italiano R. Bombelly en el año de 1572

8—01477

y su explicación, en lo esencial, fue basada en la introducción del concepto de número complejo y de las reglas de operaciones con los números complejos. Pero hasta el siglo XIX, a pesar de que los números complejos permitieron obtener muchos hechos importantes que se refieren también a los números reales, la misma existencia de los números complejos parecía bastante dudosa para muchos matemáticos. Sólo en el siglo XIX después de la aparición de las obras de K. Gauss, en las cuales se daba una representación geométrica muy clara de los números complejos (como puntos y vectores en el plano), la existencia de los números complejos llegó a ser un hecho universalmente reconocido.

Vamos a mencionar aquí un hecho más, el cual también provoca la idea sobre la necesidad de la extensión de un conjunto de números reales hasta un conjunto de números complejos.

Como es conocido, la potencia natural de cualquier número real de nuevo es un número real. Sin embargo, la operación de extracción de una raíz (operación inversa a la potenciación) no siempre es realizable en un conjunto de números reales; no existe un número real a, cuya potencia par sea un número negativo. Con otras palabras, en el conjunto de números reales no existe un número, el cual sea la raíz de la ecuación

$$x^n = b,$$

donde n es un número par y b, un número real negativo.

Siguiendo el plan general de la extensión de dominios numéricos, como ya se ha hecho muchas veces (por ejemplo, para la introducción de números negativos y números racionales), el conjunto de números reales se puede extender hasta un conjunto de números, el cual está cerrado respecto a la operación de extracción de la raíz. De antemano podemos decir que al mismo tiempo obtendremos un resultado esencialmente nuevo y para aquellos casos, cuando la operación de extracción de la raíz es realizable en el conjunto de números reales.

Uno de los procedimientos de construcción de un conjunto de números complejos consiste en que un conjunto de números reales se extiende mediante la adición al conjunto de números reales de un objeto numérico nuevo, la raíz de la ecuación

$$x^2 + 1 = 0.$$

El conjunto «extendido» obtenido se llama *conjunto de números complejos.*

§ 1. Conjunto de números complejos

1.1. Construcción axiomática de un conjunto de números complejos. Los números complejos no son números en el sentido elemental de esta palabra, que se usan para los cálculos y mediciones, sino que representan objetos matemáticos, los cuales se caracterizan por las propiedades citadas más abajo.

El *número complejo* se designa por el símbolo $a + bi$, donde a y b son números reales que se llaman respectivamente la *parte real* y la *parte imaginaria* del número complejo $a + bi$, y el símbolo i, que se define por la condición $i^2 = -1$, se llama *unidad imaginaria.*

El número complejo $a + bi$ suele designarse por una letra (más a menudo por z),

$$z = a + bi.$$

Las partes real e imaginaria del número complejo $z = a + bi$ se designan por $\operatorname{Re} z$ y $\operatorname{Im} z$ respectivamente*):

$$a = \operatorname{Re} z, \quad b = \operatorname{Im} z.$$

Dos números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$ se consideran iguales ($z_1 = z_2$), si sus partes reales e imaginarias ($a_1 = a_2$; $b_1 = b_2$) son iguales. El número complejo $z = a + bi$ se considera igual a cero ($z = 0$), si sus partes real e imaginaria son iguales a cero ($a = b = 0$). El número complejo $z = a + bi$ para $b = 0$ se considera coincidente con el número real a ($a + 0i = a$).

El número complejo $z = a + bi$ para $a = 0$ se llama *puro imaginario* y se designa bi ($0 + bi = bi$).

El número complejo z, cuya parte real es igual a la suma de las partes reales de los números z_1 y z_2, y la parte imaginaria es igual a la suma de las partes imaginarias de los números z_1 y z_2, es decir,

$$z = (a_1 + a_2) + (b_1 + b_2) i,$$

se llama *suma* de los números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$.

Sobre el número z se dice que está obtenido como resultado de la adición de los números complejos z_1 y z_2 y se escribe

$$z = z_1 + z_2.$$

Los números z_1 y z_2 se llaman *sumandos.*

Las propiedades de adición de los números complejos son:

1) la asociatividad: $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$;

2) la conmutatividad: $z_1 + z_2 = z_2 + z_1$.

El número complejo $-a - bi$ se llama *opuesto* al número complejo $a + bi$.

El número complejo que es opuesto al número complejo z se designa $-z$.

La suma de los números complejos z y $-z$ es igual a cero ($z + (-z) = 0$).

El número complejo z, el cual es la suma del número z y el número que es opuesto a z_2:

$$z = z_1 + (-z_2) = (a_1 - a_2) + (b_1 - b_2) i,$$

es decir, el número complejo, las partes real e imaginaria del cual son iguales respectivamente a la diferencia de las partes real e imaginaria del minuendo y sustraendo, es la *diferencia* de los números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$.

*) Re viene de la palabra réel (francés), real, Im, de la palabra imaginaire (francés), imaginario.

Del número z se dice que está obtenido como resultado de la sustracción del número complejo z_2 del número complejo z_1 y se escribe

$$z = z_1 - z_2.$$

El número complejo z_1 se llama *minuendo* y el número z_2 se llama *sustraendo.*

El número complejo

$$z = (a_1 \cdot a_2 - b_1 \cdot b_2) + (a_1 \cdot b_2 + a_2 \cdot b_1)\, i$$

se llama *producto* de los números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$.

Sobre el número z se dice que está obtenido como resultado de la multiplicación del número complejo z_1 por el número complejo z_2 y se escribe

$$z = z_1 \cdot z_2.$$

Los números z_1 y z_2 se llaman *factores.*

Las propiedades de multiplicación de los números complejos son:

1) la asociatividad: $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$;

2) la conmutatividad: $z_1 \cdot z_2 = z_2 \cdot z_1$.

Tal número complejo z que

$$z_1 = z \cdot z_2$$

se llama *cociente* de los números complejos z_1 y z_2 $(z_2 \neq 0)$.

El cociente de los números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$ se calcula según la fórmula

$$z = \frac{a_1 \cdot a_2 + b_1 \cdot b_2}{a_2^2 + b_2^2} + \frac{a_2 \cdot b_1 - a_1 \cdot b_2}{a_2^2 + b_2^2}\, i.$$

Sobre el número z se dice que está obtenido como resultado de la división del número complejo z_1 entre el número complejo z_2 y se escribe

$$z = \frac{z_1}{z_2} \quad \text{ó} \quad z = z_1 / z_2.$$

La adición y multiplicación de los números complejos se relaciona por la regla que se llama *ley de distributividad* de la multiplicación respecto a la adición:

$$(z_1 + z_2)\, z_3 = z_1 \cdot z_3 + z_2 \cdot z_3.$$

El número $\sqrt{a^2 + b^2}$ se llama *módulo* del número complejo $z = a + bi$. El módulo del número complejo se designa $| z |$.

Los módulos de dos números complejos cualesquiera z_1 y z_2 (en el caso del cociente se supone que $z_2 \neq 0$) satisfacen a las relaciones

$$|z_1+z_2| \leqslant |z_1|+|z_2|,$$

$$|z_1-z_2| \geqslant ||z_1|-|z_2||,$$

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|,$$

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|},$$

$$|z^n| = |z|^n = |\bar{z}|^n.$$

El número complejo $a - bi$ se llama *conjugado complejamente* con el número $a + bi$ y se designa $\bar{z}$.

Las propiedades de los números conjugados complejamente son:

1) El producto del número complejo $z = a + bi$ y el número conjugado complejamente con el $\bar{z} = a - bi$ es un número real:

$$z \cdot \bar{z} = a^2 + b^2.$$

2) Si $\bar{z}$ es un número conjugado complejamente con el número z, entonces el número conjugado complejamente con el número $\bar{z}$ es el número z:

$$\overline{(\bar{z})} = z.$$

3) Si z_1, $\bar{z}_1$ y z_2, $\bar{z}_2$ son dos pares de números conjugados complejamente, entonces la suma, el producto, la diferencia y el cociente de los números $\bar{z}_1$ y $\bar{z}_2$ son números conjugados complejamente a la suma, al producto, a la diferencia y al cociente de los números z_1 y z_2 respectivamente:

$$\bar{z}_1 + \bar{z}_2 = (\overline{z_1 + z_2}),$$

$$\bar{z}_1 \cdot \bar{z}_2 = (\overline{z_1 \cdot z_2}),$$

$$\bar{z}_1 - \bar{z}_2 = (\overline{z_1 - z_2}),$$

$$\frac{\bar{z}_1}{\bar{z}_2} = \overline{\left(\frac{z_1}{z_2}\right)} \qquad (z_2 \neq 0).$$

4) Si $z = a + bi$ y $\bar{z} = a - bi$ son un par de números conjugados complejamente, entonces

$$a = \operatorname{Re} z = \frac{z+\bar{z}}{2}, \qquad b = \operatorname{Im} z = \frac{z-\bar{z}}{2i}.$$

1.2. Conjunto de pares ordenados de los números reales y conjunto de números complejos. Para el procedimiento axiomático de introducción de números complejos como objetos matemáticos abstractos que satisfacen las propiedades enumeradas más arriba es necesario aclarar el sentido de los signos + y — en la expresión $a + bi$, los cuales antes se utilizaban para designar la suma y el producto de los números reales. Para aclarar esta cuestión realicemos otra construcción de un conjunto de números complejos, distinta de la construcción citada anteriormente.

Examinemos un conjunto de pares $(a;\ b)$, donde a, b son números reales. Dos pares $(a_1;\ b_1) = (a_2;\ b_2)$ los consideraremos iguales, si $a_1 = a_2$ y $b_1 = b_2$, y escribiremos $(a_1;\ b_1) = (a_2;\ b_2)$.

Al par $(a + c;\ b + d)$ lo vamos a llamar *suma* de los pares $(a;\ b)$ y $(c;\ d)$. Diremos que el par está obtenido como resultado de la adición de los pares $(a;\ b)$ y $(c;\ d)$ y escribiremos

$$(a + c;\ b + d) = (a;\ b) + (c;\ d).$$

Al par $(ac - bd;\ ad + bc)$ lo llamaremos *producto* de los pares $(a;\ b)$ y $(c;\ d)$. Diremos que este par está obtenido como resultado de la multiplicación del par $(a;\ b)$ por el par $(c;\ d)$ y escribiremos

$$(ac - bd;\ ad + bc) = (a;\ b)\cdot(c;\ d).$$

Las propiedades de la adición y multiplicación de pares son:

1) la asociatividad de adición:

$$[(a;\ b) + (c;\ d)] + (p;\ q) = (a;\ b) + [(c;\ d) + (p;\ q)];$$

2) la conmutatividad de adición:

$$(a;\ b) + (c;\ d) = (c;\ d) + (a;\ b);$$

3) la asociatividad de multiplicación:

$$[(a;\ b)\cdot(c;\ d)]\cdot(p;\ q) = (a;\ b)\cdot[(c;\ d)\cdot(p;\ q)];$$

4) la conmutatividad de multiplicación:

$$(a;\ b)\cdot(c;\ d) = (c;\ d)\cdot(a;\ b);$$

5) la distributividad de multiplicación respecto a la adición

$$[(a;\ b) + (c;\ d)]\cdot(p;\ q) = (a;\ b)\cdot(p;\ q) + (c;\ d)\cdot(p;\ q).$$

Al *par nulo* lo llamaremos el par $(0;\ 0)$ que satisface la condición

$$(a;\ b) + (0;\ 0) = (a;\ b).$$

Al *par de unidad* lo llamaremos el par $(1;\ 0)$ que satisface la condición

$$(a;\ b)\cdot(1;\ 0) = (a;\ b).$$

La *diferencia de los pares* $(a;\ b)$ y $(c;\ d)$ llamaremos al par

$$(a - c;\ b - d) = (a;\ b) - (c;\ d).$$

El *cociente de pares* $(a;\ b)$ y $(c;\ d)$ llamaremos al par

$$\left(\frac{a\cdot c + b\cdot d}{c^2 + d^2};\ \frac{b\cdot c - a\cdot d}{c^2 + d^2}\right) = \frac{(a;\ b)}{(c;\ d)}.$$

El cociente de dos pares está determinado para cualesquiera pares $(a;\ b)$ y $(c;\ d)$ menos el caso, cuando el par $(c;\ d)$ es nulo (es decir, el par $(0;\ 0)$).

Entre el conjunto de todos los pares que tienen la forma de $(a;\ 0)$ y el conjunto de todos los números reales existe una correspondencia biunívoca, si consideramos que al par $(a;\ 0)$ le corresponde el número real a y, a la inversa, a cualquier número real b le corresponde el par $(b;\ 0)$. Es muy fácil convencerse que para tal correspondencia a la suma de los pares $(a;\ 0)$ y $(b;\ 0)$ le corresponde la suma de los números reales a y b, y al producto de pares, el producto de los números reales a y b. Debido a tal correspondencia biunívoca entre el conjunto de todos los pares $(a;\ 0)$ y el conjunto de todos los números reales se puede no distinguir estos dos conjuntos y designar al par $(a;\ 0)$ por una sola letra a.

Tomemos ahora el par $(0;\ 1)$ y según la regla de multiplicación de pares lo multiplicamos por sí mismo:

$$(0;\ 1)\cdot(0;\ 1) = (-1;\ 0).$$

Designemos al par $(0;\ 1)$ por el símbolo i. Puesto que no distinguimos el par $(-1;\ 0)$ y el número -1, podemos escribir

$$i\cdot i = -1 \quad \text{ó} \quad i = \ldots 1.$$

La simple verificación muestra que cualquier par $(a;\ b)$ de números reales puede ser representado en la forma

$$(a;\ b) = (a;\ 0) + (b;\ 0)\cdot(0;\ 1)$$

o, en otras designaciones, en la forma

$$a + b\cdot i \quad \text{o} \quad a + bi.$$

De tal manera, la inscripción $a + bi$, la cual antes (véase p. 1.1) se consideraba como un símbolo, ahora adquiere un concreto sentido aritmético.

§ 2. Representación geométrica y forma trigonométrica de expresión de los números complejos

2.1. Representación geométrica del número complejo. Lo mismo que los números reales se pueden representar por puntos en la recta numérica, el número complejo puede representarse por puntos en el plano. La posibilidad de tal representación se basa en la identificación del conjunto de los números complejos $a + bi$ y el conjunto de pares de los números reales $(a;\ b)$, los cuales en el sistema rectangular de coordenadas Oxy pueden ser interpretados como coordenadas de los puntos del plano.

Con cada punto A del plano de coordenadas Oxy se puede relacionar el vector $\overrightarrow{OA}$ que sale del origen de coordenadas y termina en el punto A. Por eso los números complejos admiten también otra interpretación geométrica más: cada número complejo $a + bi$ se puede interpretar geométricamente como el vector $\overrightarrow{OA}$ con las coordenadas $(a;\ b)$ (fig. 3.1). Al mismo tiempo las coordenadas del vector $\overrightarrow{OA}$ son loss mismos que las coordenadas del punto A, es decir, $(a;\ b)$.

2.2. Representación geométrica de la suma y de la diferencia de los números complejos. La interpretación geométrica de los números complejos permite interpretar más claro la suma y la diferencia de dos números complejos.

Sean dados dos números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$. El número complejo $z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)\, i$ es su suma.

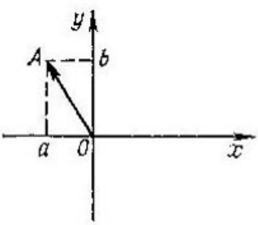


Fig. 3.1.

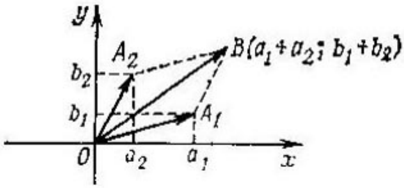


Fig. 3.2.

Por otra parte es sabido que para la adición de vectores sus respectivas coordenadas se suman. Por eso, si el vector $\overrightarrow{OA}_1$ tiene las coordenadas $(a_1;\ b_1)$ (fig. 3.2), y el vector $\overrightarrow{OA}_2$ tiene las coordenadas $(a_2;\ b_2)$, entonces su suma (el vector $\overrightarrow{OB}$) tiene las coordenadas $(a_1 + a_2;\ b_1 + b_2)$.

El vector $\overrightarrow{OB}$ es la representación geométrica de la suma de los números complejos z_1 y z_2.

Puesto que la diferencia de dos números complejos $z_1 = a_1 + b_1 i$ y $z_2 = a_2 + b_2 i$ es la suma del número complejo z_1 y el número opuesto al número complejo z_2, entonces, geométricamente, puede ser representada como la suma del vector $\overrightarrow{OA}_1$ con las coordenadas $(a_1;\ b_1)$ y el vector $\overrightarrow{OA}_2$ con las coordenadas $(-a_2;\ -b_2)$ (fig. 3.3), es decir, como el vector $\overrightarrow{OB}$ con las coordenadas $(a_1 - a_2;\ b_1 - b_2)$.

2,3. Forma trigonométrica de expresión del numero complejo. Sea que el número complejo $z = a + bi$ se representa por el vector $\overrightarrow{OA}$ con las coordenadas $(a;\ b)$ (fig. 3.4). Designemos la longitud del vector $\overrightarrow{OA}$ por la letra r:

$$r = |\overrightarrow{OA}|,$$

y el ángulo, que el vector forma con la dirección positiva del eje Ox, por φ (el ángulo φ se considera medido en radianes).

Utilizando las definiciones de las funciones sen φ y cos φ:

$$\operatorname{sen}\varphi=\frac{b}{r}, \qquad \cos\varphi=\frac{a}{r},$$

se puede escribir el número complejo $z = a + bi$ en la forma

$$z = r\cdot(\cos\varphi + i\operatorname{sen}\varphi), \tag{1}$$

donde $r = \sqrt{a^2 + b^2}$ y el ángulo φ se define de las condiciones

$$\operatorname{sen}\varphi=\frac{b}{\sqrt{a^2+b^2}}, \qquad \cos\varphi=\frac{a}{\sqrt{a^2+b^2}} \tag{2}$$

La expresión del número complejo en la forma (1) se llama *forma trigonométrica* de expresión del número complejo.

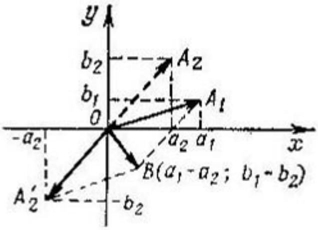


Fig. 3.3.

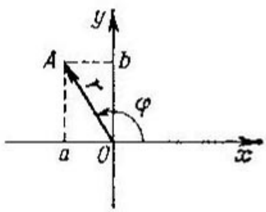


Fig. 3.4.

El número real r se llama *módulo* del número complejo y se designa $|z|$, y el ángulo φ, medido en radianes, *argumento* del número complejo z. El argumento φ del número complejo z se designa Arg z.

Si el número complejo no es igual a cero, entonces su módulo es positivo; si $z = 0$, es decir, $a = b = 0$, entonces también su módulo es igual a cero. El módulo de cualquier número complejo está definido univocamente.

Si el número complejo $z = a + bi$ no es igual a cero, entonces su argumento se define por las fórmulas (2) con una precisión de hasta el ángulo múltiple a 2π. Si $z = 0$, entonces $r = 0$ y de la fórmula (1) se ve que como argumento φ se puede eligir cualquier ángulo. Por eso el argumento del número complejo igual a cero no está definido.

Por lo general para eliminar la multiformidad que surge para calcular el argumento del número complejo suelen utilizar el concepto de *valor principal* del argumento del número complejo (la designación arg z), suponiendo que $\arg z \in (-\pi;\ \pi)$.

El argumento del número complejo z se relaciona con el valor principal del argumento por la correlación

$$\operatorname{Arg} z = \arg z + 2\pi k, \qquad k \in Z.$$

Sea que z_1 y z_2 son dos números complejos distintos de cero y escritos en la forma trigonométrica:

$$z_1 = r_1 \cdot (\cos \varphi_1 + i \operatorname{sen} \varphi_1),$$
$$z_2 = r_2 \cdot (\cos \varphi_2 + i \operatorname{sen} \varphi_2).$$

El *producto* de dos números complejos z_1 y z_2 es un número complejo, cuyo módulo es igual al producto de los módulos de los factores y el argumento, a la suma de los argumentos de los factores:

$$z_1 \cdot z_2 = r_1 \cdot r_2 \cdot [\cos (\varphi_1 + \varphi_2) + i \operatorname{sen} (\varphi_1 + \varphi_2)].$$

En la interpretación geométrica el vector que representa el producto de los números complejos z_1 y z_2 se obtiene mediante el giro del

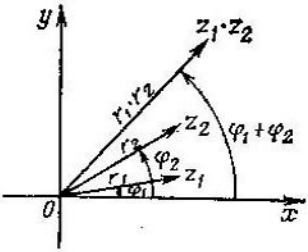


Fig. 3.5.

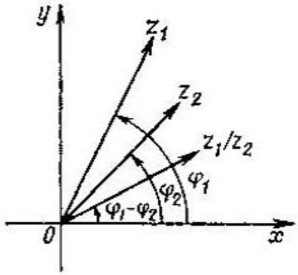


Fig. 3.6.

vector z_1 en sentido antihorario en un ángulo igual a φ_2, y mediante su estiramiento en $| z_2 |$ veces (para el caso $| z_2 | > 1$ véase la fig. 3.5).

El cociente de dos números complejos z_1 y z_2 representa un número complejo, cuyo módulo es igual al cociente de los módulos del dividendo y del divisor y el argumento del cociente de dos números complejos no iguales a cero es igual a la diferencia de los argumentos del dividendo y del divisor:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} \cdot [\cos (\varphi_1 - \varphi_2) + i \operatorname{sen} (\varphi_1 - \varphi_2)].$$

Geométricamente el vector que representa el cociente de dos números complejos z_1 y z_2, se obtiene girando el vector que representa el número complejo z_1 en sentido horario en un ángulo igual a φ_2 y mediante su contracción en $| z_2 |$ veces (para el caso $| z_2 | > 1$ véase la fig. 3.6).

§ 3. Potencia del número complejo

3.1. Potencia natural del número complejo. La ley de asociatividad de la multiplicación de los números complejos permite introducir el concepto de potencia natural del número complejo:

El número complejo w que se obtiene como resultado de la multiplicación del número z por sí mismo n veces:

$$w = \underbrace{z \cdot z \cdot \ldots \cdot z}_{n \text{ factores}} \qquad (1)$$

se llama n-ésima *potencia* del número complejo z.

Se suele utilizar la expresión más corta siguiente:

$$w = z^n. \qquad (2)$$

En la igualdad (2) el número z se llama *base de una potencia*, y el número natural n, *exponente de una potencia*.

La n-ésima potencia del número complejo z dado en la forma trigonométrica

$$z = r(\cos\varphi + i \operatorname{sen}\varphi)$$

se calcula mediante la *fórmula de Moivre*

$$z^n = r^n(\cos n\varphi + i \operatorname{sen} n\varphi).$$

3.2. Raíz de la n-ésima potencia de un número complejo. Tal número complejo w, cuya n-ésima potencia es igual a z, se llama *raíz de la n-ésima potencia del número complejo* z:

$$w^n = z.$$

La raíz de la n-ésima potencia del número complejo z se designa por el símbolo $\sqrt[n]{z}$. A diferencia de la raíz del número real, la raíz de la n-ésima potencia del número complejo se define multiformemente. Precisamente, en el conjunto de números complejos existen exactamente n raíces de la n-ésima potencia del número complejo dado (véase el teorema principal del álgebra).

Todas las raíces de la n-ésima potencia del número complejo z, dado en la forma trigonométrica

$$z = r(\cos\varphi + i \operatorname{sen}\varphi),$$

se calculan mediante la fórmula

$$\sqrt[n]{z} = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i \operatorname{sen}\frac{\varphi + 2\pi k}{n}\right),$$

donde $k = 0, 1, 2, \ldots, n-1$.

Geométricamente todas las raíces de la n-ésima potencia del número complejo $z = r(\cos\varphi + i \operatorname{sen}\varphi)$ se representan por los puntos que están situados en la circunferencia con el centro en el origen de coordenadas, cuyo radio es igual a $\sqrt{r}$, y los ángulos centrales, formados por radios trazados en los puntos cercanos, son iguales a φ/n.

Ejemplo. Calcular las raíces de la cuarta potencia del número -1. El número -1 en la forma trigonométrica puede ser escrito como sigue:

$$-1 = 1 \cdot (\cos\pi + i \operatorname{sen}\pi).$$

Las raíces de la cuarta potencia del número -1 son los números complejos

$$\sqrt[4]{-1}=\sqrt[4]{1}\left(\cos\frac{\pi+2\pi k}{4}+i\operatorname{sen}\frac{\pi+2\pi k}{4}\right),$$

donde $k=0, 1, 2, 3$, es decir, son los números complejos

$$\cos\frac{\pi}{4}+i\operatorname{sen}\frac{\pi}{4}=\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i,$$

$$\cos\frac{3\pi}{4}+i\operatorname{sen}\frac{3\pi}{4}=-\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i,$$

$$\cos\frac{5\pi}{4}+i\operatorname{sen}\frac{5\pi}{4}=-\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i,$$

$$\cos\frac{7\pi}{4}+i\operatorname{sen}\frac{7\pi}{4}=\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i$$

(véase la fig. 3.7).

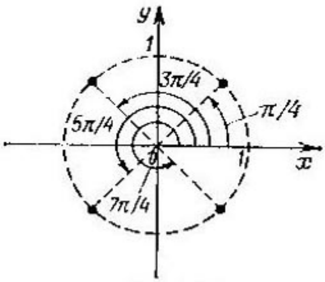


Fig. 3.7.

Análogamente en el conjunto de números complejos se puede calcular la raíz de la n-ésima potencia de cualquier número real. En este caso por lo menos una raíz del número real positivo será real.

CAPÍTULO 4

Álgebra

El término «álgebra» viene del título de la obra de Muhammad al-Jwarismi «Al-yabr wa-l-muqabala» (siglo IX) que contiene los métodos generales de solución de problemas, los cuales se reducen a las ecuaciones de primero y segundo grado. A mediados del siglo XVII fue establecido, en lo principal, el simbolismo algebraico contemporáneo. Hasta el siglo XVIII bajo el término de álgebra se entendía la ciencia sobre los cálculos con letras, idénticas a las transformaciones de las fórmulas con letras *), la solución de las ecuaciones de primero-cuarto grado, los logaritmos las progresiones, el análisis combinatorio, es decir, las mismas partes de las matemáticas, las cuales en el presente son la asignatura que se estudia en la escuela de enseñanza media. Actualmente a todas estas partes del álgebra se las llama álgebra elemental.

En los siglos XVIII—XIX el álgebra tenía por objeto, ante todo, el estudio de los polinomios, la teoría de las ecuaciones algebraicas con una incógnita, la teoría de los sistemas de ecuaciones lineales con varias incógnitas, así como la teoría de las matrices y determinantes.

La tercera etapa (moderna) del desarrollo del álgebra como ciencia de las operaciones algebraicas comenzó a mediados del siglo XIX y está relacionada con la aparición de los distintos ejemplos de operaciones algebraicas con objetos de naturaleza muy distinta a los números reales. Las multiplicaciones de sustituciones y las operaciones con los números complejos son los primeros de tales ejemplos.

*) En el álgebra, a *diferencia* de la *aritmética*, se examinan no los números, sino las expresiones con letras, denominándolas como expresiones algebraicas.

§ 1. Polinomios de una variable

1.1. Concepto de polinomio. Operaciones aritméticas con polinomios. La expresión de la forma

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n, \quad (1)$$

donde n es un número entero no negativo: $a_0, a_1, a_2, \ldots, a_n$ son los coeficientes del polinomio, además a_0, llamado coeficiente mayor, se considera no igual a cero, se denomina *polinomio* de n-ésimo grado respecto al valor variable x.

El polinomio de primer grado también se denomina polinomio *lineal*, el polinomio de segundo grado, polinomio *cuadrado* y el polinomio de tercer grado, polinomio *cúbico*.

La expresión del polinomio en la forma (1) se llama expresión de *grados decrecientes* x; precisamente esta expresión es la que se utilizará posteriormente, si no se advierte lo contrario.

Cada uno de los sumandos en la expresión (1) se llama *término* de un polinomio.

De la definición del polinomio de n-ésimo grado se deduce que existen polinomios de n-ésimo grado para cualquier n no negativo. En particular, cualquier número real distinto de cero es un polinomio de grado nulo. El número cero también se considera un polinomio; es el único polinomio, cuyo grado no está definido.

Dos polinomios $P(x)$ y $Q(x)$ se consideran iguales (o idénticamente iguales), si sus coeficientes son iguales para los mismos grados de la variable x. La igualdad de polinomios se escribe mediante el signo de igualdad:

$$P(x) = Q(x).$$

El polinomio $S(x)$, cuyos coeficientes para cada grado x son iguales a la suma de los coeficientes para este grado x de los polinomios $P(x)$ y $Q(x)$ se llama *suma* de los polinomios

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n,$$
$$Q(x) = b_0x^m + b_1x^{n-1} + b_2x^{n-2} + \ldots + b_{m-1}x + b_m.$$

Sobre el polinomio $S(x)$ se dice que está obtenido como resultado de la adición de los polinomios $P(x)$ y $Q(x)$ y se escribe

$$S(x) = P(x) + Q(x).$$

Ejemplo 1. El polinomio

$$S(x) = x^4 + x^3 - x + 5$$

es la suma de los polinomios $P(x) = x^4 - x^3 + 3$ y $Q(x) = 2x^3 - x + 2$.

El polinomio $M(x)$ de grado $n + m$, cuyos coeficientes $c_0, c_1, c_2, \ldots, c_{n+m-1}, c_{n+m}$ se calculan por las fórmulas

$$\begin{aligned}
c_0 &= a_0 b_0, \\
c_1 &= a_1 b_0 + a_0 b_1, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
c_k &= a_k b_0 + a_{k-1} b_1 + \ldots + \\
&+ a_1 b_{k-1} + a_0 b_k, \\
\\
c_{n+m} &= a_n b_m
\end{aligned}$$

se llama *producto* de los polinomios

$$P(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n,$$
$$D(x) = b_0 x^m + b_1 x^{m-1} + b_2 x^{m-2} + \ldots + b_{m-1} x + b_m.$$

Sobre el polinomio $M(x)$ se dice que está obtenido como resultado de la multiplicación del polinomio $P(x)$ por el polinomio $Q(x)$ y se escribe

$$M(x) = P(x) \cdot Q(x).$$

Ejemplo 2. El polinomio de sexto grado

$$M(x) = P(x) \cdot Q(x) = 2x^6 + 2x^4 - x^3 + 3x^2 - x + 3$$

es el producto del polinomio de segundo grado $P(x) = x^2 + 1$ por el polinomio de cuarto grado $Q(x) = 2x^4 - x + 3$.

Las operaciones de adición y multiplicación de polinomios son asociativas, conmutativas, y se conectan entre sí por la ley de distributividad.

El polinomio $-a_0 x^n - a_1 x^{n-1} - a_2 x^{n-2} - \ldots - a_{n-1} x - a_n$

se llama *opuesto* para el polinomio

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n.$$

El polinomio que es opuesto al polinomio $P(x)$ se designa con $P(x)$. La suma del polinomio $P(x)$ y el polinomio opuesto, $P(x)$, es igual a cero:

$$P(x) + (-P(x)) = 0.$$

El polinomio $L(x)$ que es la suma del polinomio $P(x)$ y el polinomio que es opuesto al polinomio $Q(x)$:

$$L(x) = P(x) + (-Q(x))$$

se llama diferencia de polinomios $P(x)$ y $Q(x)$. Sobre el polinomio $L(x)$ se dice que está obtenido como resultado de la sustracción del polinomio $Q(x)$ del polinomio $P(x)$ y se escribe

$$L(x) = P(x) - Q(x).$$

La suma, el producto y la diferencia de dos polinomios cualesquiera son también polinomios.

1.2. Divisores del polinomio. Si para los polinomios dados $P(x)$ y $Q(x)$ se halla un polinomio $G(x)$ tal que

$$P(x) = Q(x) \cdot G(x),$$

entonces se dice que el polinomio $P(x)$ se divide (o se divide exactamente) entre el polinomio $Q(x)$; al mismo tiempo, el polinomio $Q(x)$ se llama *divisor entero* (o simplemente *divisor*) del polinomio $P(x)$, y el polinomio $G(x)$, *cociente* de los polinomios $P(x)$ y $Q(x)$. Si el polinomio $Q(x)$ es el divisor del polinomio $P(x)$, entonces también el cociente $G(x)$ es el divisor del polinomio $P(x)$.

Principales propiedades de los divisores del polinomio:

1) Si el polinomio $P(x)$ se divide entre el polinomio $K(x)$, y $K(x)$ se divide entre el polinomio $Q(x)$, entonces también $P(x)$ se divide entre $Q(x)$.

2) Si los polinomios $P(x)$ y $K(x)$ se dividen entre el polinomio $Q(x)$, entonces su suma y diferencia también se dividen entre $Q(x)$.

3) Si el polinomio $P(x)$ se divide entre el polinomio $Q(x)$, entonces el producto $P(x)$ por cualquier polinomio $K(x)$ también se divide entre $Q(x)$.

4) Cualquier polinomio $P(x)$ se divide entre cualquier polinomio de grado nulo (es decir, entre un número, distinto de cero).

5) Sólo los polinomios $cP(x)$ son divisores del polinomio $P(x)$, los cuales tienen el mismo grado que $P(x)$.

6) Los polinomios $P(x)$ y $Q(x)$ se dividen al mismo tiempo uno entre otro cuando y sólo $P(x) = cQ(x)$ $(c \neq 0)$.

Máximo común divisor de dos polinomios. El polinomio $R(x)$ se llama *máximo común divisor* para el polinomio $P(x)$ y $Q(x)$, si sirve de divisor para cada uno de estos polinomios. La propiedad 4) (véase más arriba) muestra que todos los polinomios de grado cero (es decir, todos los números reales, a excepción del número cero) pertenecen al número de divisores comunes de los polinomios arbitrarios $P(x)$ y $Q(x)$. Si dos polinomios no tienen otros divisores comunes, entonces estos polinomios se llaman *recíprocamente simples.*

Se llama *máximo común divisor* de los polinomios $P(x)$ y $Q(x)$ tal polinomio $D(x)$, el cual es el divisor común de éstos y al mismo tiempo se divide entre cualquier divisor común de dichos polinomios.

Anotemos que si $D(x)$ es el máximo divisor común de los polinomios $P(x)$ y $Q(x)$, entonces el máximo común divisor de estos polinomios es también el polinomio $cD(x)$, donde c es un número arbitrario, diferente de cero. En otras palabras, hasta ahora el máximo común divisor de dos polinomios está determinado sólo con una precisión hasta el factor de la potencia cero. Para lograr la uniformidad completa en la definición del máximo común divisor de dos polinomios, se aprovecha la condición siguiente: de todos los polinomios $cD(x)$ el máximo común divisor es aquél, cuyo coeficiente mayor es igual a la unidad.

Ahora se puede decir que dos polinomios son recíprocamente simples, si su máximo común divisor es igual a la unidad.

1.3. División de polinomios. Esquema de Horner. Teorema de Bezout. Sea que $P(x)$ y $Q(x)$ son los polinomios dados y el grado del polinomio $P(x)$ es mayor o igual al grado del polinomio $Q(x)$. Si resulta que el polinomio $P(x)$ no se divide entre el polinomio $Q(x)$, es decir, si no existe el polinomio $G(x)$ tal que

$$P(x) = Q(x) \cdot G(x),$$

entonces se introduce la operación de la *división inexacta con resto.*

Dividir el polinomio $P(x)$ entre el polinomio $Q(x)$ con resto significa hallar dos polinomios $G(x)$ y $R(x)$ tales que

$$P(x) = Q(x) \cdot G(x) + R(x), \tag{2}$$

donde el grado del polinomio $R(x)$ debe ser menos que el grado del polinomio $Q(x)$.

El polinomio $P(x)$ se llama *dividendo,* $Q(x)$, *divisor,* $G(x)$, *cociente,* y $R(x)$ *resto.* Para dividir el polinomio $P(x)$ entre el polinomio $Q(x)$ los polinomios $G(x)$ y $R(x)$ se hallan uniformemente.

Para dividir un polinomio entre otro se suele aplicar la regla de la «división mediante un ángulo». Para esto se ordenan los polinomios $P(x)$ y $Q(x)$ según las potencias decrecientes x y se halla el término mayor del cociente $G(x)$ partiendo de la condición que para la multiplicación del mismo por el término mayor del divisor $Q(x)$ debe obtenerse el término mayor del dividendo $P(x)$. El término mayor del cociente hallado se multiplica luego por el divisor y el polinomio obtenido se sustrae del dividendo. De resultas de la sustracción se obtiene el polinomio $P_1(x)$. El término siguiente del cociente se define de la condición que este término, siendo multiplicado por el término mayor del divisor $Q(x)$, debe dar el término mayor del polinomio $P_1(x)$, etc. El proceso continúa hasta que el grado de la diferencia nueva (es decir, el grado de cierto polinomio $P_h(x)$) no sea menor que el grado del divisor. El polinomio $P_h(x)$ será el resto.

Como resultado de la división del polinomio de n-ésimo grado $P(x)$ entre el polinomio de k-ésimo grado $Q(x)$ $(k \leqslant n)$ en el cociente se obtiene el polinomio de grado $n - k$.

Ejemplo 3. Dividir el polinomio $P(x) = x^4 + x^3 + 3x^2 + 3x + 2$ entre el polinomio $Q(x) = x^3 + x + 1$.

El término mayor del polinomio $P(x)$ es el término x^4, y el término mayor del polinomio $Q(x)$ es el término x^2. Dividiendo x^4 entre x^2, obtenemos x^2. El monomio x^2 es el término mayor del cociente.

Multiplicamos el monomio x^2 por el polinomio $Q(x)$. De resultas de la multiplicación se obtiene el polinomio $x^4 + x^3 + x^2$. Sustraemos este polinomio del polinomio $P(x)$:

$$(x^4 + x^3 + 3x^2 + 3x + 2) - (x^4 + x^3 + x^2) = 2x^2 + 3x + 2 = P_1(x).$$

Observando que el grado del polinomio $P_1(x)$, obtenido como resultado de la sustracción, no es menor que el grado del polinomio $Q(x)$, continuamos el proceso de división; dividimos el término mayor $2x^2$ del polinomio $P_1(x)$ entre el término mayor del polinomio $Q(x)$:

$$\frac{2x^2}{x^2} = 2.$$

El polinomio de grado cero (el número 2) será el término siguiente del cociente.

Multiplicamos el polinomio $Q(x)$ por el número 2. De resultas de la multiplicación se obtiene el polinomio $2x^2 + 2x + 2$, el cual sustraemos del polinomio $P_1(x)$:

$$P_1(x) - (2x^2 + 2x + 2) = 2x^2 + 3x + 2 -$$

$$- (2x^2 + 2x + 2) = x = P_2(x).$$

Aquí el proceso de división se termina, puesto que el grado del polinomio $P_2(x)$ es menor que el grado del polinomio $Q(x)$. El polinomio $P_2(x)$ será el resto de la división del polinomio $P(x)$ entre el polinomio $Q(x)$. Todo el proceso descrito de la división del polinomio $P(x)$ entre el polinomio $Q(x)$ se refleja brevemente en la expresión siguiente:

$$\begin{array}{r|l}
-\,x^4 + x^3 + 3x^2 + 3x + 2 & x^2 + x + 1 \\
\underline{x^4 + x^3 + x^2 \qquad\qquad} & \overline{x^2 + 2} \\
-\,2x^2 + 3x + 2 & \\
\underline{2x^2 + 2x + 2} & \\
x &
\end{array}$$

De esta manera el cociente de la división del polinomio

$$P(x) = x^4 + x^3 + 3x^2 + 3x + 2$$

entre el polinomio $Q(x) = x^2 + x + 1$ es el polinomio $G(x) = x^2 + 2$, y el resto es el polinomio $R(x) = x$, es decir, el polinomio $P(x)$ puede ser escrito en la forma

$$x^4 + x^3 + 3x^2 + 3x + 2 = (x^2 + x + 1)(x^2 + 2) + x.$$

Esquema de Horner. El teorema de Bezout. Para dividir el polinomio de grado n $P(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ entre el polinomio de primer grado $x - c$ es cómodo emplear el método de división abreviada (denominado esquema de Horner). Este se obtiene como consecuencia de la definición de la operación de división de polinomios, de la cual se deduce que dividiendo el polinomio de n-ésimo grado entre el polinomio lineal $x - c$ se puede obtener en el resto bien el polinomio de grado cero (es decir, un número diferente cero) o bien el cero, y el grado del cociente el igual a $n - 1$.

Sea que el cociente de los polinomios $P(x)$ y $x - c$ tiene la forma

$$G(x) = \alpha_0x^{n-1} + \alpha_1x^{n-2} + \ldots + \alpha_{n-2}x + \alpha_{n-1},$$

y el resto $R(x)$ es igual al número β. En virtud de la fórmula (2) tenemos

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n =$$
$$= (x - c)(\alpha_0x^{n-1} + \alpha_1x^{n-2} + \ldots + \alpha_{n-2}x + \alpha_{n-1}) + \beta. \quad (3)$$

Suprimiendo los paréntesis y reduciendo los términos semejantes en el miembro segundo de la igualdad (3), de la condición de la igualdad de polinomios obtenemos un sistema de ecuaciones lineales para hallar los coeficientes $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{n-2}, \alpha_{n-1}, \beta$:

$$\begin{aligned} \alpha_0 &= a_0, \\ \alpha_1 &= a_1 + c\alpha_0, \\ \alpha_2 &= a_2 + c\alpha_1, \\ &\cdots\cdots \\ \alpha_{n-1} &= a_{n-1} + c\alpha_{n-2}, \\ \beta &= a_n + c\alpha_{n-1}. \end{aligned}$$

A esta tabla se denomina *esquema de Horner.*

La primera ecuación del sistema da el valor $\alpha_0 = a_0$. Sustituyendo el valor de α_0 en la segunda ecuación del sistema, obtenemos $\alpha_1 = a_1 + ca_0$. Sustituyendo este valor de α_1 en la tercera ecuación del sistema, obtenemos el valor α_2, etc. La última será la expresión para el resto β:

$$\beta = a_0c^n + a_1c^{n-1} + a_2c^{n-2} + \ldots + a_{n-1}c + a_n.$$

Esta igualdad es conocida bajo el nombre de *teorema de Bezout:*

Dividiendo el polinomio $a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ entre $x - c$ se obtiene un resto que es igual al valor de este polinomio para $x = c$.

1.4. Algoritmo de Euclides para hallar el máximo común divisor de dos polinomios. Sea que $P(x)$ y $Q(x)$ son dos polinomios arbitrarios, y el grado del polinomio $P(x)$ es más o igual al grado del polinomio $Q(x)$. Dividimos el polinomio $P(x)$ entre el polinomio $Q(x)$, es decir, hallamos el cociente $G(x)$ y el resto $R(x)$ de la división de estos polinomios. Luego dividimos $Q(x)$ entre $R(x)$ y obtenemos el resto $R_1(x)$; dividimos $R(x)$ entre $R_1(x)$ y obtenemos el resto $R_2(x)$; dividimos $R_1(x)$ entre $R_2(x)$, etc. Puesto que los grados del resto todo el tiempo decrecen, en esta cadena de divisiones sucesivas tenemos que llegar dentro de un número finito de pasos, hasta el momento, cuando la división se efectúa exactamente. El resto $R_n(x)$, entre el cual se divide exactamente el resto antecedente $R_{n-1}(x)$ será el máximo común divisor de polinomios $P(x)$ y $Q(x)$. Es necesario advertir que este máximo común divisor, en general, no tendrá un coeficiente mayor que sea igual a la unidad.

El máximo común divisor de dos polinomios con un coeficiente mayor igual a la unidad se obtiene, si dividimos todos los coeficientes del divisor común obtenido por el coeficiente para el término mayor. Para mayor claridad escribiremos el *algoritmo de Euclides* en la forma de una cadena de igualdades:

$$
\begin{aligned}
P(x) &= Q(x)\cdot G(x) + R(x),\\
Q(x) &= R(x)\cdot G_1(x) + R_1(x),\\
R(x) &= R_1(x)\cdot G_2(x) + R_2(x),\\
R_1(x) &= R_2(x)\cdot G_3(x) + R_3(x)\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
R_{n-2}(x) &= R_{n-1}(x)\cdot G_n(x) + R_n(x),\\
R_{n-1}(x) &= R_n(x)\cdot G_{n+1}(x).
\end{aligned}
$$

Anotemos otra vez que el máximo común divisor del polinomio $P(x)$ y $Q(x)$ es el polinomio $R_n(x)$ dividido por el coeficiente mayor. El algoritmo de Euclides para hallar el máximo común divisor de dos polinomios es completamente análogo al algoritmo de Euclides para hallar el máximo común divisor de dos números naturales.

Ejemplo. Hallar el máximo común divisor de los polinomios

$$P(x) = x^5 - x^4 - 2x^3 + 2x^2 + x - 1,$$
$$Q(x) = x^3 - 1.$$

Dividamos «mediante el ángulo» el polinomio $P(x)$ entre el polinomio $Q(x)$:

$$
\begin{array}{r|l}
-\;x^5 - x^4 - 2x^3 + 2x^2 + x - 1 & x^3 - 1 \\
\cline{2-2}
x^5 \qquad\qquad\quad - x^2 \qquad\quad & x^2 - x - 2 \\
\cline{1-1}
-\;- x^4 - 2x^3 + 3x^2 + x & \\
- x^4 \qquad\qquad\qquad + x & \\
\cline{1-1}
-\;- 2x^3 + 3x^2 - 1 & \\
- 2x^3 \qquad\quad + 2 & \\
\cline{1-1}
3x^2 - 3 &
\end{array}
$$

El cociente es $G(x) = x^2 - x - 2$, el resto, $R(x) = 3x^2 - 3$. Dividimos el polinomio $Q(x)$ entre el resto $R(x)$:

$$\begin{array}{r|l} -\,x^3-1 & 3x^2-3 \\ \underline{x^3-x} & \frac{1}{3}x \\ x-1 & \end{array}$$

Designamos el cociente con $G_1(x)$, y el resto, con $R_1(x)$. En virtud del algoritmo de Euclides dividimos el resto de la división anterior $R(x)$ entre el resto obtenido $R_1(x)$:

$$\begin{array}{r|l} -\,3x^2-3 & x-1 \\ \underline{3x^2-3x} & 3x+3 \\ 3x-3 & \end{array}$$

El polin mio $R(x)$ se dividió exactamente entre el polinomio $R_1(x) = x - 1$. Por consiguiente, el polinomio $x - 1$ es el máximo común divisor de dos polinomios dados $P(x)$ y $Q(x)$.

1.5. Raíces de un polinomio. Sea $P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$ cierto polinomio de grado n. Tomemos el número c y sustituyámoslo por la variable x en el polinomio $P(x)$. El número obtenido

$$a_0c^n + a_1c^{n-1} + a_2c^{n-2} + \ldots + a_{n-1}c + a_n$$

se llama *valor del polinomio* $P(x)$ para $x = c$ y se designa $P(c)$.

El número c se llama *raíz del polinomio* $P(x)$, si el número $P(c)$ es un cero.

El concepto de raíz de un polinomio se conecta estrechamente con la obtención del divisor lineal de un polinomio. Si el número c es la raíz del polinomio de grado n $P(x)$, entonces el polinomio $P(x)$ se divide entre el polinomio lineal $x - c$, es decir, el polinomio $P(x)$ puede representarse en la forma

$$P(x) = (x - c)\,Q(x),$$

donde $Q(x)$ es el polinomio de grado $n - 1$.

Puede resultar que el polinomio de n-ésimo grado $P(x)$ se divide no sólo entre el polinomio lineal $x - c$, sino también entre un grado más alto de éste, es decir, por el polinomio de la forma $(x - c)^k$, donde k es un número natural mayor que la unidad. El número c se llama *raíz de la multiplicidad* k del polinomio $P(x)$ (o *raíz múltiple* k), si el polinomio $P(x)$ se divide entre el polinomio $(x - c)^k$, pero no se divide por el polinomio $(x - c)^{k+1}$. Si el polinomio $P(x)$ se divide entre el polinomio lineal $x - c$, pero no se divide por $(x - c)^2$, entonces el número c se llama *raíz simple* del polinomio $P(x)$.

Para que el número c sea la raíz múltiple k del polinomio $P(x)$, es necesario y suficiente que el número c sea la raíz múltiple $(k - 1)$ de la variable del polinomio $P(x)$.

1.6. Fórmulas de la multiplicación abreviada. Es fácil convencerse que el número c es la raíz del polinomio $x^n - c^n$ (n es cualquiera) y el número c, la raíz de los polinomios $x^n - c^n$ (n es par) y $x^n + c^n$ (n es impar); por eso

1) el polinomio $x^n - c^n$ se divide por el binomio $x - c$ para cualquier n natural;

2) el polinomio $x^n - c^n$ se divide por el binomio $x + c$ para cualquier n par;

3) el polinomio $x^n + c^n$ se divide por el binomio $x + c$ para cualquier n impar.

Como resultado de la división se obtienen respectivamente los polinomios de grado $(n - 1)$:

$$\frac{x^n - c^n}{x - c} = x^{n-1} + x^{n-2}c + x^{n-3}c^2 + \ldots + xc^{n-2} + c^{n-1} =$$

$$= \sum_{k=0}^{n-1} x^{n-k-1}c^k \quad (n \text{ es cualquiera});$$

$$\frac{x^n - c^n}{x + c} = x^{n-1} - x^{n-2}c + x^{n-3}c^2 - \ldots + xc^{n-2} - c^{n-1} =$$

$$= \sum_{k=0}^{n-1} (-1)^k x^{n-k-1}c^k \quad (n \text{ es par});$$

$$\frac{x^n + c^n}{x + c} = x^{n-1} - x^{n-2}c + x^{n-3}c^2 - \ldots - xc^{n-2} + c^{n-1} =$$

$$= \sum_{k=0}^{n-1} (-1)^k x^{n-k-1}c^k \quad (n \text{ es impar}).$$

En particular, de las igualdades citadas más arriba se deducen las fórmulas más simples que se llaman *fórmulas de* *la multiplicación abreviada:*

$$(x+c)(x-c) = x^2 - c^2,$$
$$(x+c)(x^2 - xc + c^2) = x^3 + c^3,$$
$$(x-c)(x^2 + xc + c^2) = x^3 - c^3.$$

1.7. Fórmulas de Viète. Entre las raíces $c_1, c_2, \ldots c_n$ del polinomio de grado n

$$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1}x + a_n$$

con el coeficiente mayor que es igual a la unidad (cada raíz se toma tal número de veces, cual es su multiplicidad), existen las dependencias que se llaman *fórmulas de Viéte:*

$$c_1 + c_2 + c_3 + \ldots + c_{n-1} + c_n = -a_1,$$
$$c_1c_2 + c_1c_3 + \ldots + c_{n-1}c_n = a_2,$$
$$c_1c_2c_3 + c_1c_2c_4 + \ldots + c_{n-2}c_{n-1}c_n = -a_3,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$c_1c_2c_3 \ldots c_{n-1}c_n = (-1)^n a_n.$$

Si $P(x)$ es un trinomio de segundo grado, es decir,

$$P(x) = x^2 + px + q,$$

y los números c_1 y c_2 son sus raíces, entonces las fórmulas de Viéte tienen la forma

$$c_1 + c_2 = -p, \quad c_1c_2 = q.$$

Si $P(x)$ es un polinomio cúbico, es decir,

$$P(x) = x^3 + px^2 + qx + r,$$

y los números c_1, c_2, c_3 son sus raíces, entonces las fórmulas de Viéte tienen la forma

$$c_1 + c_2 + c_3 = -p, \quad c_1c_2 + c_2c_3 + c_1c_3 = q,$$
$$c_1c_2c_3 = -r.$$

1.8. Teorema fundamental del álgebra. Introduciendo el concepto de raíces del polinomio, no hemos planteado la cuestión sobre si cada polinomio con coeficientes reales tiene raíces. Es conocido que existen polinomios con coeficientes reales, que no tienen raíces reales ($x^2 + 1$ es uno de tales polinomios) y precisamente éste hecho es una de las causas de la introducción del concepto de número complejo. Se puede suponer que existen polinomios que no tienen raíces incluso en el conjunto de números complejos, sobre todo si examinamos los polinomios con cualesquiera coeficientes complejos (un caso particular de los cuales son los polinomios con coeficientes reales). Si esto fuera así, entonces el sistema de números complejos necesitaría de una extensión posterior. En realidad, sin embargo, es válido el siguiente *teorema fundamental del álgebra*:

Cualquier polinomio de grado n con coeficientes complejos en el conjunto de números complejos tiene exactamente n raíces, si cada raíz múltiple se cuenta tal número de veces, cual es su multiplicidad.

Todas las demostraciones de este teorema (las cuales son bastante numerosas) en mayor o menor grado utilizan las propiedades de los números reales y complejos relacionados con la continuidad, a consecuencia de lo cual todas estas demostraciones no se pueden considerar puramente algebraicas.

El teorema fundamental del álgebra es válido también para $n = 0$, puesto que el polinomio de grado cero no tiene raíces. El teorema fundamental del álgebra no es aplicable sólo al polinomio nulo, cuyo grado no está definido.

1.9. Descomposición del polinomio en factores. En virtud del teorema fundamental del álgebra el polinomio de n-ésimo grado

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$$

con cualesquiera coeficientes complejos (y, en particular, con reales) tiene n raíces (teniendo en cuenta sus multiplicidades) y, por consiguiente, se divide exactamente en polinomios lineales.

$$x - c_1, \quad x - c_2, \ldots, x - c_n,$$

donde $c_1, c_2, \ldots, c_n$ son las raíces del polinomio: es decir, el polinomio lo representamos en forma del producto de n factores lineales:

$$P(x) = a_0(x - c_1)(x - c_2)\ldots(x - c_n).$$

La descomposición del polinomio $P(x)$ en producto de polinomios lineales es únicamente con una precisión hasta el orden de sucesión de factores.

1.10. Ciertas consecuencias del teorema fundamental del álgebra para los polinomios con coeficientes reales. Si un número complejo (pero no real) $c = \alpha + \beta i$ es una raíz del polinomio de n-ésimo grado con coeficientes reales, entonces también el número $\bar{c} = \alpha - \beta i$ conjugado complejamente con el número $c = \alpha + \beta i$ será la raíz del polinomio dado.

Del carácter conjugado de las raíces complejas de un polinomio con coeficientes reales se deduce que cualquier polinomio de grado impar con coeficientes reales tiene por lo menos una raíz real.

Sea que $P(x)$ es un polinomio de grado n con coeficientes reales, y $c = \alpha + \beta i$, su raíz. Entonces el número $\bar{c} = \alpha - \beta i$ es también la raíz del polinomio $P(x)$, y, por consiguiente, el polinomio $P(x)$ puede ser representado en la forma

$$P(x) = (x - c)(x - \bar{c})\,Q(x),$$

donde $Q(x)$ es el polinomio de $(n-2)$-ésimo grado con coeficientes reales.

El producto

$$(x-c)(x-\bar{c}) = x^2 - (c+\bar{c})x + c\bar{c} = x^2 - 2\alpha x + \alpha^2 + \beta^2$$

representa un polinomio de segundo grado con coeficientes reales.

De esta manera, cualquier polinomio con coeficientes reales se representa y además mediante un único procedimiento (con una precisión hasta el orden de factores), en la forma del producto de su coeficiente mayor a_0, de ciertos polinomios con coeficientes reales del tipo $x-c$, que corresponden a sus raíces reales, y de los polinomios cuadrados de la forma $x^2-(c+\bar{c})x+c\bar{c}$, que corresponden a los pares de raíces complejas conjugadas.

Anotemos que entre los polinomios con coeficientes reales y con coeficiente mayor, igual a la unidad, sólo los polinomios lineales $x-c$ y los polinomios cuadrados que tienen la forma $x^2-(c+\bar{c})x+$ $+c\bar{c}$ son indescomponibles en factores de grado menor en el conjunto de números reales (o son *irreducibles*).

La propiedad demostrada de los polinomios con coeficientes reales es análoga a la propiedad de descomposición de cualquier número natural en factores primos. Para cualquier polinomio $P(x)$ de n-ésimo grado con coeficientes reales y el primer coeficiente, igual a la unidad, se puede escribir la descomposición, análoga al desarrollo canónico de los números naturales (véase p. 1.3 de capítulo 2):

$$P(x) = (x-c_1)^{k_1}(x-c_2)^{k_2}\dots(x-c_l)^{k_l}\times$$
$$\times[x^2-(c_{l+1}+\bar{c}_{l+1})x+c_{l+1}\bar{c}_{l+1}]^{k_{l+1}}\dots$$
$$\dots[x^2-(c_{l+m}+\bar{c}_{l+m})x+c_{l+m}\bar{c}_{l+m}]^{k_{l+m}}$$
$$k_1+k_2+\dots+k_l+2(k_{l+1}+\dots+k_{l+m})=m,$$

donde $c_1, c_2, \dots, c_l$ son las raíces reales del polinomio con las multiplicidades $k_1, k_2, \dots, k_l$, y $\{c_{l+1}; \bar{c}_{l+1}\}, \dots, \{c_{l+m}; \bar{c}_{l+m}\}$ son los pares de las raíces conjugadas complejamente con las multiplicidades $k_{l+1}, \dots, k_{l+m}$ respectivamente.

§ 2. Polinomios de varias variables

2.1. Monomios y polinomios de varias variables. La expresión de la forma

$$ax_1^{k_1}x_2^{k_2}\dots x_n^{k_n},$$

donde a es un coeficiente numérico distinto de cero, y $k_1, k_2, \dots, k_n$ son números enteros no negativos, se llama monomio de n variables

$x_1, x_2, \ldots, x_n$. Dos monomios de n variables

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \quad \text{y} \quad bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$$

se llaman *semejantes*, si $k_1 = l_1$, $k_2 = l_2$, $\ldots$, $k_n = l_n$.

La suma del número finito de un monomio que tiene la forma $ax_1^{k_1}x_2^{k_2}x_3^{k_3} \ldots x_n^{k_n}$ se llama *polinomio* $P(x_1, x_2, x_3, \ldots, x_n)$ de n variables $x_1, x_2, x_3, \ldots, x_n^{k_n}$. Los monomios que forman el polinomio se llaman términos del polinomio.

Al mismo tiempo se presupone que el polinomio $P(x_1, x_2, x_3, \ldots, x_n)$ no contiene términos semejantes y que se consideran sólo los términos con coeficientes distintos de cero. El exponente más alto con el cual la variable x_i entra en el polinomio se llama *grado del polinomio* $P(x_1, x_2, x_3, \ldots, x_n)$, respecto a cierta variable elegida x_i ($i = 1, 2, 3, \ldots, n$).

El número

$$k_1 + k_2 + k_3 + \ldots + k_n,$$

es decir, la suma de los exponentes de las potencias de todas las variables x_i que entran en el monomio dado se llama *grado del monomio*

$$x_1^{k_1}x_2^{k_2}x_3^{k_3} \ldots x_n^{k_n}$$

(o *grado de un conjunto, de variables*). El valor más grande de la suma $k_1 + k_2 + k_3 + \ldots + k_n$ que se encuentra en cualquiera de los términos del polinomio se llama *grado de un polinomio* del conjunto de todas las variables.

El polinomio $P(x_1, x_2, x_3, \ldots, x_n)$ se llama *homogéneo*, si los grados de todos los monomios que entran en el polinomio son iguales.

2.2. Ordenación lexicográfica de los términos de un polinomio. Mientras que para los polinomios de una variable existen dos modos naturales de ordenación de los términos de un polinomio, según las potencias decrecientes y crecientes de la variable, para los polinomios de varias variables tales modos naturales ya no son utilizables. Sin embargo, existe un modo que se utiliza frecuentemente de ordenación bastante determinada de los términos de un polinomio de varias variables, que se llama modo «lexicográfico». Este modo es análogo al modo de ordenación de palabras en el diccionario: considerando las letras ordenadas de tal manera como se hace en el alfabeto, se determina la posición mutua de dos palabras en el diccionario por sus primeras letras; si las primeras letras coinciden, entonces por las segundas, etc.

La esencia del modo «lexicográfico» de ordenación de los términos de un polinomio de varias variables consiste en lo siguiente.

Sea que $P(x_1, x_2, x_3, \ldots, x_n)$ es un polinomio de n variables $x_1, x_2, x_3, \ldots, x_n$, y supongamos que

$$x_1^{k_1}x_2^{k_2}x_3^{k_3} \ldots x_n^{k_n} \quad \text{y} \quad x_1^{l_1}x_2^{l_2}x_3^{l_3} \ldots x_n^{l_n}$$

son dos términos distintos del polinomio dado. Puesto que estos dos términos son distintos, por lo menos una de las diferencias de los exponentes de grados

$$k_1 - l_1,\ k_2 - l_2,\ \ldots,\ k_n - l_n \tag{1}$$

es distinta del cero. El término $x_1^{k_1}x_2^{k_2}x_3^{k_3}\ldots x_n^{k_n}$ se considera superior al término $x_1^{l_1}x_2^{l_2}x_3^{l_3}\ldots x_n^{l_n}$, si la primera diferencia distinta de cero en (1) es positiva. En caso contrario el término $x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ se considera inferior al término $x_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$. Dicho de otra manera, el término $x_1^{k_1}x_2^{k_2}x_3^{k_3}\ldots x_n^{k_n}$ es superior al término $x_1^{l_1}x_2^{l_2}x_3^{l_3}\ldots x_n^{l_n}$, si $k_1 > l_1$ o si $k_1 = l_1$, $k_2 > l_2$, etc. Es difícil advertir que de que el término $x_1^{k_1}x_2^{k_2}x_3^{k_3}\ldots x_n^{k_n}$ es superior al término $x_1^{l_1}x_2^{l_2}x_3^{l_3}\ldots x_n^{l_n}$ no se deduce que el grado del término $x_1^{k_1}x_2^{k_2}x_3^{k_3}\ldots x_n^{k_n}$ es mayor que el grado del término $x_1^{l_1}x_2^{l_2}x_3^{l_3}\ldots x_n^{l_n}$. Así, por ejemplo, el término $x_1^3x_2$ se considera superior al término $x_1^2x_2^6$, aunque su grado (por el conjunto de variables) es menor.

Comparando entre sí todos los términos del polinomio, destacamos de ellos aquel que es superior a todos los demás, y lo escribimos en el primer lugar. Comparamos los demás términos del polinomio y destacamos de ellos al superior, que lo escribimos en el segundo lugar, etc. De resultas obtenemos la expresión del polinomio de n variables, en la cual cada término, comenzando por el segundo, es inferior al anterior.

Esta forma de escritura se llama modo lexicográfico de expresión de un polinomio de varias variables. El término del polinomio que se encuentra en el primer lugar se considera término superior de un polinomio.

§ 3. Fracciones algebraicas racionales

3.1. Operaciones aritméticas con las fracciones algebraicas. La expresión que tiene la forma

$$\frac{P}{Q}, \tag{1}$$

donde P y Q son los polinomios de una o de varias variables*), se llama *fracción algebraica racional*. El polinomio P se llama *numerador* de una fracción, y el polinomio Q, su *denominador*.

Seguidamente, caso de que no haya advertencia especial, se estudiarán las fracciones, el numerador y denominador de las cuales son los polinomios de una variable.

La fracción algebraica racional (1) está definida para todos los valores de la variable x, a excepción de la lista finita de valores que anulan el polinomio $Q(x)$ (las raíces del polinomio $Q(x)$).

*) La fracción algebraica racional a veces también se denomina *expresión algebraica racional*.

El conjunto de valores de una variable x, para los cuales el denominador de la fracción se distingue de cero, se llama conjunto de *valores admisibles* de la variable x de la fracción algebraica dada. Dos fracciones racionales algebraicas $\frac{P(x)}{Q(x)}$ y $\frac{P_1(x)}{Q_1(x)}$ se llaman *iguales* (o *idénticamente iguales*), si $P(x)\cdot Q_1(x) = Q(x)\cdot P_1(x)$ para todos los valores x, para los cuales $Q(x) \neq 0$ y $Q_1(x) \neq 0$. De la definición de la igualdad de fracciones se deduce que la fracción algebraica no cambia, si el numerador y el denominador de la fracción se multiplica por un mismo polinomio $K(x)$:

$$\frac{P(x)}{Q(x)} = \frac{P(x)\cdot K(x)}{Q(x)\cdot K(x)}.$$

En esta propiedad se basa la regla de simplificación de las fracciones algebraicas, es decir, la división del numerador y del denominador por su común divisor y también la reducción de fracciones a un denominador común.

Como resultado de la simplificación de una fracción algebraica puede ser obtenida una fracción algebraica que tiene un conjunto de valores admisibles que no coincide con el conjunto de valores admisibles de la fracción original. Esta circunstancia debe tenerse en cuenta cuando se simplifican las fracciones algebraicas.

Por ejemplo, la simplificación de la fracción algebraica

$$\frac{x^2-1}{x(x-1)}$$

en un polinomio lineal $x-1$ reduce la fracción algebraica dada a la forma

$$\frac{x+1}{x}.$$

El número 1 entra en el conjunto de valores admisibles de la fracción obtenida, mientras que para la fracción original el número 1 no entraba en el conjunto de valores admisibles.

La fracción algebraica

$$\frac{P\cdot Q_1+P_1\cdot Q}{Q\cdot Q_1} \qquad (2)$$

se llama *suma* de dos fracciones algebraicas racionales

$$\frac{P}{Q} \text{ y } \frac{P_1}{Q_1}.$$

Se dice que la fracción algebraica (2) está obtenida de resultas de la adición de fracciones $\frac{P}{Q}$ y $\frac{P_1}{Q_1}$ y se escribe

$$\frac{P\cdot Q_1+P_1\cdot Q}{Q\cdot Q_1}=\frac{P}{Q}+\frac{P_1}{Q_1}.$$

La fracción

$$\frac{P\cdot P_1}{Q\cdot Q_1} \tag{3}$$

se llama *producto* de dos fracciones algebraicas racionales

$$\frac{P}{Q} \text{ y } \frac{P_1}{Q_1}.$$

Se dice que la fracción algebraica (3) está obtenida de resultas de la multiplicación de fracciones $\frac{P}{Q}$ y $\frac{P_1}{Q_1}$ y se escribe

$$\frac{P\cdot P_1}{Q\cdot Q_1}=\frac{P}{Q}\cdot\frac{P_1}{Q_1}.$$

La fracción

$$\frac{P\cdot Q_1-P_1\cdot Q}{Q\cdot Q_1} \tag{4}$$

se llama *diferencia* de dos fracciones algebraicas

$$\frac{P}{Q} \text{ y } \frac{P_1}{Q_1}.$$

Se dice que la fracción algebraica (4) está obtenida de resultas de la sustracción de la fracción $\frac{P_1}{Q_1}$ de la fracción $\frac{P}{Q}$ y se escribe

$$\frac{P\cdot Q_1-P_1\cdot Q}{Q\cdot Q_1}=\frac{P}{Q}-\frac{P_1}{Q_1}.$$

La fracción

$$\frac{P\cdot Q_1}{Q\cdot P_1} \tag{5}$$

se llama *cociente* de dos fracciones algebraicas racionales

$$\frac{P}{Q} \text{ y } \frac{P_1}{Q_1}.$$

Se dice que la fracción algebraica (5) está obtenida de resultas de la división de la fracción $\frac{P}{Q}$ entre la fracción $\frac{P_1}{Q_1}$ y se escribe

$$\frac{P\cdot Q_1}{Q\cdot P_1}=\frac{P}{Q}:\frac{P_1}{Q_1}.$$

El cociente de dos fracciones algebraicas racio ales está definido para todos aquellos valores admisibles de la variable x de las fracciones $\frac{P}{Q}$ y $\frac{P_1}{Q_1}$, para los cuales $P_1(x)\neq 0$.

Las operaciones de adición y multiplicación de las fracciones algebraicas racionales son asociativas y conmutativas y se relacionan por la ley de la distributividad de la multiplicación respecto a la adición.

3.2. Fracciones algebraicas propias. La fracción algebraica racional $\frac{P(x)}{Q(x)}$ se llama *propia*, si el grado de un polinomio que se encuentra en el numerador de la fracción es menor que el grado del polinomio que se encuentra en el denominador de la fracción, y se llama *impropia* en el caso contrario.

De esta manera, la fracción

$$\frac{x^2}{x^3+1}$$

es propia y las fracciones

$$\frac{x^2}{x^2-3} \quad \text{y} \quad \frac{x^3}{x+1}$$

son impropias.

Cualquier fracción algebraica racional impropia $\frac{P(x)}{Q(x)}$ puede ser representada en forma de suma de un polinomio y de una fracción propia. Para esto es necesario hallar el cociente $G(x)$ y el resto $R(x)$ de la división del polinomio $P(x)$ por el polinomio $Q(x)$. Entonces la fracción impropia $\frac{P(x}{Q(x}$ resulta ser representada en la forma

$$\frac{R(x)}{Q(x)} = G(x) + \frac{R(x)}{Q(x)}, \tag{6}$$

donde $\frac{K(x)}{Q(x)}$ es una fracción propia, y $G(x)$ es un polinomio llamado *parte entera* de la fracción algebraica racional.

La representación de la fracción impropia $\frac{P(x)}{Q(x)}$ en la forma (6) se llama *formación de la parte entera de la fracción algebraica impropia.*

Ejemplo. La fracción algebraica racional

$$\frac{x^3}{x+1}$$

puede ser escrita en forma de suma de un polinomio y de una fracción algebraica racional:

$$\frac{x^3}{x+1} = x^2 - x + 1 - \frac{1}{x+1},$$

donde el polinomio $x^2 - x + 1$ es la parte entera de la fracción algebraica impropia.

3.3. Fracciones simples. Una fracción algebraica racional propia $\frac{P(x)}{Q(x)}$ se llama *simple*, si su denominador $Q(x)$ es el grado natural de cierto polinomio irreducible $q(x)$:

$$Q(x) = q^k(x) \quad (k \geqslant 1),$$

y el grado del numerador $P(x)$ es menor que el grado del polinomio $q(x)$. Recordemos, que entre los polinomios con coeficientes reales y con un término mayor igual a la unidad, son irreducibles sólo los polinomios lineales $x - c$ y los polinomios cuadrados que tienen la forma $x^2 + px + q$, si los coeficientes del trinomio de segundo grado satisfacen la desigualdad $p^2 - 4q < 0$. A consecuencia de esto una fracción algebraica racional puede ser simple sólo en los casos, cuando su numerador $P(x)$ es o un polinomio de primer grado o un polinomio de grado cero (es decir, un número no igual a cero).

Ejemplos. 1. La fracción $\frac{x-1}{(x^2+1)^k}$ (k es natural) será una fracción algebraica racional simple, puesto que su denominador es el grado de un polinomio irreducible y el grado de un polinomio irreducible es mayor que el grado del numerador.

2. La fracción $\frac{x^3-3}{x^2+1}$ no es una fracción simple, puesto que el grado del polinomio que se encuentra en el numerador de la fracción

es mayor que el grado del polinomio que se encuentra en el denominador.

3. La fracción $\frac{x}{x^2-1}$ no es simple, puesto que su denominador se puede descomponerlo en factores.

En la teoría de las fracciones algebraicas racionales el lugar de mayor importancia le pertenece al *teorema* siguiente:

Cualquier fracción racional propia se descompone en una suma de fracciones simples y esta descomposición es única.

Ejemplo 4. Descompóngase en una suma de fracciones simples la fracción propia $\frac{P(x)}{Q(x)}$, donde

$$P(x) = 2x^4 - 10x^3 + 7x^2 + 4x + 3,$$

$$Q(x) = x^5 - 2x^3 + 2x^2 - 3x + 2.$$

El polinomio $Q(x)$ puede ser representado en la forma del producto de polinomios irreducibles:

$$x^5 - 2x^3 + 2x^2 - 3x + 2 = (x+2)(x-1)^2(x^2+1).$$

La descomposición buscada de la fracción $\frac{P(x)}{Q(x)}$ debe tener la forma de

$$\frac{P(x)}{Q(x)} = \frac{A}{x+2} + \frac{B}{(x-1)^2} + \frac{2}{x-1} + \frac{Dx+E}{x^2+1},$$

donde quedan por hallar los números A, B, C, D, E. Reduzcamos el miembro segundo de la fracción a un denominador común y, utilizando la definición de la igualdad de polinomios, obtendremos un sistema de cinco ecuaciones lineales con cinco incógnitas: A, B, C, D, E.

El sistema obtenido tendrá una única solución:

$A = 3$, $B = 1$, $C = -2$, $D = 1$, $E = -3$ y, por consiguiente, la descomposición buscada tiene la forma

$$\frac{P(x)}{Q(x)} = \frac{3}{x+2} + \frac{1}{(x-1)^2} - \frac{2}{x-1} + \frac{x-3}{x^2+1}.$$

3.4. Proporciones. La igualdad que tiene la forma

$$\frac{a}{b} = \frac{c}{d} \tag{7}$$

se llama *proporción* y a, b, c, d, *términos* de la proporción. En la proporción (7) a y d se llaman términos *extremos*, y b y c, términos *medios* de la proporción.

De la proporción (7) se deduce:

1) $a\cdot d = b\cdot c$;

2) $\frac{a \pm b}{b} = \frac{c \pm d}{d}$, $\frac{a-b}{a+b} = \frac{c-d}{c+d}$,

$$\frac{ma+nb}{pa+qb} = \frac{mc+nd}{pc+qd} \text{ (\textit{proporciones deducidas})},$$

donde m, n, p, q son números arbitrarios (p y q no son simultáneamente iguales a cero).

§ 4. Expresiones algebraicas irracionales

Una expresión algebraica se llama *irracional*, si con las variables que entran en la expresión algebraica se efectúa, junto con las operaciones de adición, sustracción, multiplicación y división, la operación de elevación a una potencia (no entera) racional.

Ejemplos. 1. La expresión algebraica

$$x^2 + a\sqrt{x} + b$$

es una expresión algebraica irracional, puesto que la variable x se encuenta bajo el signo de la raíz cuadrada.

2. La expresión algebraica

$$\sqrt{2}x^3 - \sqrt[3]{a}x - 1$$

respecto a la variable x no es una expresión algebraica irracional. Si consideramos a como variable en esta expresión algebraica, entonces la expresión es irracional. De esta manera, la respuesta a la cuestión sobre la irracionalidad de una expresión algebraica depende de qué magnitudes en la expresión algebraica dada se consideran variables y cuáles, coeficientes.

Las transformaciones de las expresiones algebraicas irracionales se efectúan de acuerdo con las leyes generales de las operaciones aritméticas y con las reglas de operaciones con radicales.

Eliminación de la irracionalidad del numerador (denominador) de una fracción algebraica irracional. La expresión algebraica que tiene la forma

$$\frac{f(x)}{g(x)} \qquad (1)$$

se llama expresión *irracional fraccionaria*, si por lo menos una de las expresiones algebraicas $f(x)$ o $g(x)$ es irracional respecto a la variable x.

Sea $S(x)$ cierta expresión algebraica irracional respecto a la variable x. La expresión algebraica $\overline{S}(x)$ que no es idénticamente igual a cero se llama *factor adicional* para la expresión algebraica $S(x)$, si $S(x)\,\overline{S}(x)$ es una expresión algebraica racional respecto a la variable x. El conocimiento de factores adicionales para las expresiones algebraicas que se encuentran en el numerador (o en el denominador) de la fracción (1), permite representar la fracción $\frac{f(x)}{g(x)}$ en forma de una fracción, el numerador (denominador) de la cual es la expresión algebraica racional:

$$\frac{f}{g}=\frac{f\cdot\overline{f}}{g\cdot\overline{f}}=\frac{f\cdot\overline{g}}{g\cdot\overline{g}},$$

donde $\overline{f}$ es el factor adicional del numerador y $\overline{g}$, el factor adicional del denominador. Tal transformación de la expresión irracional es la que se llama *eliminación de la irracionalidad* respectivamente del numerador y del denominador de la expresión fraccional irracional.

De esta manera, el proceso de eliminación de la irracionalidad se reduce a hallar un factor adicional para la expresión algebraica S que contiene radicales. Podemos decir, que es muy difícil indicar un método universal de obtención del factor adicional. A continuación citamos los factores adicionales para ciertas expresiones algebraicas de dos variables, x y y:

$S(x, y)$	$\overline{S}(x, y)$	$S(x, y)\,\overline{S}(x, y)$
$\sqrt[n]{x^k y^l}$	$\sqrt[n]{x^{n-k} y^{n-l}}$	xy
$\sqrt{x}\pm\sqrt{y}$	$\sqrt{x}\mp\sqrt{y}$	$x-y$
$\sqrt[3]{x}\pm\sqrt[3]{y}$	$\sqrt[3]{x^2}\mp\sqrt[3]{xy}+\sqrt[3]{y^2}$	$x\pm y$
$\sqrt[n]{x}-\sqrt[n]{y}$	$\sqrt[n]{x^{n-1}}+\sqrt[n]{x^{n-2}y}+\ldots+\sqrt[n]{y^{n-1}}$	$x-y$

(n es cualquier)

$$\sqrt[n]{x}+\sqrt[n]{y} \quad \sqrt[n]{x^{n-1}}-\sqrt[n]{x^{n-2}y}+\ldots+\sqrt[n]{y^{n-1}} \quad x+y$$

(n es impar).

Es muy fácil convencerse, que si la expresión algebraica $\overline{S}(x, y)$ es un factor adicional para la expresión algebraica $S(x, y)$, también la expresión algebraica $S(x, y)$ será un factor adicional para la expresión algebraica $\overline{S}(x, y)$.

Observemos que si el número de sumandos que contienen radicales es mayor que dos, entonces la eliminación de radicales del numerador (denominador) de la fracción es cómodo efectuarla en ciertos casos, utilizando sucesivamente las fórmulas de la multiplicación de irracionalidades.

Ejemplos. 1. Eliminar la irracionalidad del denominador de la fracción algebraica

$$\frac{x}{\sqrt[3]{x^2}+\sqrt[3]{x}+1}.$$

Observando que de factor adicional para el denominador sirve

$$\sqrt[3]{x}-1,$$

multiplicamos el numerador y el denominador de la fracción por la expresión $\sqrt[3]{x}-1$:

$$\frac{x}{\sqrt[3]{x^2}+\sqrt[3]{x}+1}=\frac{x(\sqrt[3]{x}-1)}{(\sqrt[3]{x}-1)(\sqrt[3]{x^2}+\sqrt[3]{x}+1)}=\frac{x(\sqrt[3]{x}-1)}{x-1}.$$

2. Elimínese la irracionalidad del denominador de la fracción

$$\frac{1}{\sqrt{x}+\sqrt{y}-\sqrt{z}}.$$

Designando $\sqrt{x}+\sqrt{y}=t$ y multiplicando el numerador y denominador de la fracción dada por la expresión $t+\sqrt{z}$, obtenemos

$$\frac{1}{t-\sqrt{z}}=\frac{t+\sqrt{z}}{t^2-z}=\frac{\sqrt{x}+\sqrt{y}+\sqrt{z}}{(\sqrt{x}+\sqrt{y})^2-z}=\frac{\sqrt{x}+\sqrt{y}+\sqrt{z}}{x+y-z+2\sqrt{xy}}.$$

Como resultado de las transformaciones de la fracción algebraica original efectuadas hemos obtenido una fracción algebraica que contiene en su denominador sólo un radical. Ahora, multiplicando el numerador y denominador de la fracción por la expresión

$$(x+y-z)-2\sqrt{xy},$$

obtenemos la fracción que ya no contiene la irracionalidad en el denominador:

$$\frac{(\sqrt{x}+\sqrt{y}+\sqrt{z})(x+y-z-2\sqrt{xy})}{(x+y-z)^2-4xy}.$$

§ 5. Ecuaciones. Ecuaciones algebraicas

5.1. Principales definiciones. En álgebra se estudian dos tipos de igualdades: identidades y ecuaciones.

La *identidad* es una igualdad que se cumple para todos los (admisibles) valores de las letras que entran en ella *). Para escribir la identidad junto con el signo $=$ se usa también el signo $\equiv$.

La *ecuación* es una igualdad que se cumple sólo para ciertos valores de letras que entran en ella. Las letras que entran en la ecuación, según la condición del problema, pueden ser no equivalentes: unas pueden adquirir todos sus valores admisibles (son los llamados *parámetros* o *coeficientes* de la ecuación y suelen designarse por las primeras letras del alfabeto latino: a, b, c, . . ., o por las mismas letras con los índices: a_1, a_2, . . ., o b_1, b_2, . . .); otras, cuyos valores es necesario hallar, son las llamadas *incógnitas* (suelen denominarse por las últimas letras del alfabeto latino: x, y, z, . . ., o por las mismas letras con los índices: x_1, x_2, . . . o y_1, y_2, . . .).

*) Bajo el término *admisibles* se entiende aquellos valores numéricos de las letras, para los cuales son realizables todas las operaciones que se efectúan con las letras que entran en la igualdad. Por ejemplo, los valores admisibles de las letras que entran en la igualdad

$$\frac{b+1}{a}=\sqrt{x-2}$$

son los siguientes: para a $(-\infty;\ 0)\cup(0;\ +\infty)$; para x $[2;\ +\infty)$, para b $(-\infty;\ +\infty)$.

En su forma general la ecuación puede ser escrita como sigue:

$$F(x_1, x_2, \ldots, x_n) = 0.$$

Según el número de incógnitas la ecuación se llama, ecuación con una, dos, etc. incógnitas.

El valor de las incógnitas que convierten la ecuación en identidad se llama *solución* de la ecuación.

Resolver una ecuación significa hallar el conjunto de sus soluciones o demostrar que las mismas no existen. Según el tipo de ecuación el conjunto de soluciones de una ecuación puede ser infinito, finito y vacío.

Si todas las soluciones de la ecuación $F = 0$ son soluciones de la ecuación $G = 0$, entonces se dice que la ecuación $G = 0$ es consecuencia de la ecuación $F = 0$ y se escribe

$$F = 0 \Rightarrow G = 0.$$

Dos ecuaciones

$$F = 0 \text{ y } G = 0$$

se llaman *equivalentes*, si cada una de ellas es consecuencia de la otra, y se escribe

$$F = 0 \Leftrightarrow G = 0.$$

De esta manera, dos ecuaciones se consideran equivalentes, si el conjunto de soluciones de estas ecuaciones coinciden.

La ecuación $F = 0$ se considera equivalente a las dos (o a varias) ecuaciones $F_1 = 0$, $F_2 = 0$, si el conjunto de soluciones de la ecuación $F = 0$ coincide con la unión de los conjuntos de soluciones de las ecuaciones $F_1 = 0$, $F_2 = 0$.

Ciertas ecuaciones equivalentes:

1) La ecuación $F + G = G$ es equivalente a la ecuación $F = 0$, que estudia sobre el conjunto de valores admisibles de la ecuación original.

2) La ecuación $\frac{F}{G} = 0$ es equivalente a la ecuación $F = 0$ que se estudia sobre el conjunto de valores admisibles de la ecuación original.

3) La ecuación $F \cdot G = 0$ es equivalente a las dos ecuaciones $F = 0$ y $G = 0$.

4) La ecuación $F^n = 0$ es equivalente a la ecuación $F = 0$.

5) La ecuación $F^n = G^n$ para n impar es equivalente a la ecuación $F = G$, y para n par es equivalente a las dos ecuaciones $F = G$ y $F = -G$.

La ecuación que tiene la forma

$$P_n = 0,$$

donde P_n es el polinomio de n-ésimo grado de una o varias variables, se llama *ecuación algebraica*.

Se llama *ecuación algebraica* con una incógnita la ecuación que se reduce a la forma

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0,$$

donde n es un número entero no negativo; los coeficientes del polinomio $a_0, a_1, a_2, \ldots, a_{n-1}, a_n$ se llaman *coeficientes* (o *parámetros*) de la ecuación y se consideran dados; x se llama *incógnita* y es buscada. El número n se llama *grado* de una ecuación.

Los valores de la incógnita x que convierten a la ecuación algebraica en identidad se llaman *raíces* (a veces *soluciones*) de una ecuación algebraica. Antes de exponer los conocimientos sobre los tipos particulares de la ecuación algebraica (lineal, cuadrática y otras ecuaciones) observemos, que los métodos de solución de las ecuaciones cuadráticas y lineales ya eran conocidos hace mucho tiempo, mientras que el descubrimiento de los métodos de solución de las ecuaciones de tercero y cuarto grado se refiere al siglo XVI. Después, casi tres siglos continuaron las tentativas infructuosas para hacer el siguiente paso, es decir, para hallar las fórmulas que expresen mediante los radicales las raíces de cualquiera ecuación de quinto grado a través de sus coeficientes. Estas tentativas se terminaron sólo cuando Abel en los años veinte del siglo pasado demostró que tales fórmulas, que dan la solución de las ecuaciones de n-ésimo grado para cualquier $n \geqslant 5$, no pueden ser halladas.

Este resultado de Abel no eliminó, sin embargo, las posibilidades de que las raíces de ciertos polinomios concretos con coeficientes numéricos pueden, no obstante, expresarse de una u otra manera a través de los coeficientes mediante cierta combinación de radicales. El problema sobre las condiciones, para las cuales la ecuación dada es resoluble en radicales, fue investigado por Galois en los años treinta del siglo pasado. En particular, él mostró que para cualquier $n \geqslant 5$ se pueden indicar las ecuaciones de n-ésimo grado que son irresolubles en radicales incluso con coeficientes numéricos enteros. Tal, por ejemplo, es la ecuación $x^5 = 4x - 2 = 0$.

5.2. Ecuación lineal. La ecuación de primer grado

$$ax + b = 0, \tag{1}$$

donde a y b son ciertos números reales, se llama *ecuación lineal*.

La ecuación lineal siempre tiene una sola raíz $x = -\frac{b}{a}$, la cual se halla de la manera siguiente.

Adicionando a ambos miembros de la ecuación (1) el número $-b$, obtenemos la ecuación

$$ax = -b, \tag{2}$$

que es equivalente a la ecuación (1). Dividiendo ambos miembros de la ecuación (2) por el valor $a \neq 0$, obtenemos la raíz de la ecuación (1):

$$x = -\frac{b}{a}.$$

5.3. Ecuación cuadrática. La ecuación algebraica de segundo grado

$$ax^2 + bx + c = 0, \tag{3}$$

donde a, b, c son ciertos números reales, se llama *ecuación cuadrática*. Si $a = 1$, entonces la ecuación cuadrática (3) se llama *reducida*.

Las raíces de la ecuación cuadrática se calculan por la fórmula

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

la cual puede ser obtenida como resultado de las transformaciones siguientes de la ecuación inicial *).

Escribimos el trinomio de segundo grado $ax^2 + bx + c$ que se encuentra en el primer miembro de la ecuación (3) en forma de cuadrado perfecto y efectuamos las transformacio-

*) La expresión $b^2 - 4ac$ se llama *discriminante* de la ecuación cuadrática.

nes, como resultado de las cuales cada vez se obtendrá una ecuación equivalente a la ecuación inicial:

$$a\left(x+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a}=0 \Leftrightarrow a\left(x+\frac{b}{2a}\right)^2=$$

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

$$\Leftrightarrow \sqrt{\left(x+\frac{b}{2a}\right)^2}=\sqrt{\frac{b^2-4ac}{4a^2}} \Leftrightarrow$$

$$\Leftrightarrow \left|x+\frac{b}{2a}\right|=\frac{\sqrt{b^2-4ac}}{2\,|a|}.$$

La última ecuación es equivalente a dos ecuaciones:

$$x+\frac{b}{2a}=\frac{\sqrt{b^2-4ac}}{2\,|a|}, \tag{4}$$

$$x+\frac{b}{2a}=-\frac{\sqrt{b^2-4ac}}{2\,|a|}. \tag{5}$$

Utilizando la definición del valor absoluto del número, es fácil convencerse que las ecuaciones (4) y (5) son equivalentes a las ecuaciones

$$x=-\frac{b}{2a}+\frac{\sqrt{b^2-4ac}}{2a},$$

$$x=-\frac{b}{2a}-\frac{\sqrt{b^2-4ac}}{2a}.$$

Uniendo estas dos fórmulas obtenemos la fórmula de las raíces:

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

En este caso:

si $b^2-4ac>0$, entonces la ecuación tiene dos distintas raíces reales;

si $b^2-4ac=0$, la ecuación tiene una raíz real de multiplicidad 2;

si $b^2-4ac<0$, la ecuación no tiene raíces reales, sino que tiene dos raíces conjugadas complejamente:

$$x_1=-\frac{b}{2a}+\frac{\sqrt{4ac-b^2}}{2a}\,i,\quad x_2=-\frac{b}{2a}-\frac{\sqrt{4ac-b^2}}{2a}\,i.$$

Las formas particulares de la ecuación cuadrática (3) son:

1) La ecuación cuadrática reducida (en el caso, si $a = 1$), la cual suele escribirse en la forma

$$x^2 + px + q = 0.$$

Las raíces de la ecuación cuadrática se calculan por la fórmula

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}. \tag{6}$$

2) La ecuación cuadrática con el segundo coeficiente par, la cual suele escribirse en la forma

$$ax^2 + 2kx + c = 0 \text{ (k es un número entero).}$$

Las raíces de esta ecuación cuadrática se calculan cómodamente por la fórmula

$$x_{1,2} = \frac{-k \pm \sqrt{k^2 - ac}}{a}. \tag{7}$$

Las fórmulas (6) y (7) son formas particulares de las fórmulas que se emplean para calcular las raíces de la ecuación cuadrática completa.

Las raíces de la ecuación cuadrática reducida

$$x^2 + px + q = 0$$

se relacionan con sus coeficientes por las fórmulas de Viete

$$x_1 + x_2 = -p,$$
$$x_1 x_2 = q.$$

En caso de que la ecuación cuadrática reducida tenga raíces reales, las fórmulas de Viète permiten juzgar tanto sobre los signos, como sobre el valor relativo de las raíces de la ecuación cuadrática, es decir,

si $q > 0$, $p > 0$, entonces ambas raíces son negativas;

si $q > 0$, $p < 0$, ambas raíces son positivas;

si $q < 0$, $p > 0$, la ecuación tiene raíces de distintos signos, además, la raíz negativa por su valor absoluto es mayor que la positiva;

si $q < 0$, $p < 0$, entonces la ecuación tiene raíces de distintos signos, además, la raíz negativa por su valor absoluto es menor que la raíz positiva

5.4. Ecuaciones binomias. La ecuación de n-ésimo grado

$$ax^n \pm b = 0 \tag{8}$$

se llama *ecuación binomia*. Para $a > 0$ y $b > 0$ por sustitución *) de

$$x = y\sqrt[n]{\frac{b}{a}},$$

donde $\sqrt[n]{\frac{b}{a}}$ es el valor aritmético de la raíz, la ecuación (8) se reduce a la ecuación

$$y^n \pm 1 = 0,$$

la cual examinaremos a continuación.

La ecuación binomia $y^n - 1 = 0$ para n impar tiene una raíz real $y = 1$. En el conjunto de números complejos esta ecuación tiene n raíces (de las cuales una es real y $n - 1$ son complejas):

$$y_k = \cos\frac{2\pi k}{n} + i \operatorname{sen} \frac{2\pi k}{n} \quad (k = 0, 1, 2, \ldots, n-1). \tag{9}$$

La ecuación binomia $y^n - 1 = 0$ para n par en el conjunto de números reales tiene dos raíces $\{y = 1;\ y = -1\}$, y en el conjunto de números complejos, n raíces que se calculan mediante la fórmula (9).

La ecuación binomia $y^n + 1 = 0$ para n impar tiene una raíz real $y = -1$, y en el conjunto de números complejos, n raíces que se calculan por la fórmula

$$y_k = \cos\frac{2\pi k + \pi}{n} + i \operatorname{sen} \frac{2\pi k + \pi}{n} \quad (k = 0, 1, 2, \ldots, n-1). \tag{10}$$

La ecuación binomia $y^n + 1 = 0$ para n par no tiene raíces reales. En el conjunto de números complejos la ecuación tiene n raíces que se calculan por la fórmula (10).

*) Si a y b tienen los signos distintos, entonces $x = y\sqrt[n]{\left|\frac{b}{a}\right|}$.

Demos una breve enumeración de los conjuntos de raíces de la ecuación binomia para ciertos valores concretos de n.

1) $y^2 - 1 = 0$ $(n = 2)$.

La ecuación tiene dos raíces reales $y_{1,2} = \pm 1$.

2) $y^3 - 1 = 0$ $(n = 3)$.

La ecuación tiene una raíz real $y_1 = 1$ y dos raíces complejas $y_{2,3} = \dfrac{-1 \pm i\sqrt{3}}{2}$.

3) $y^4 - 1 = 0$ $(n = 4)$.

La ecuación tiene dos raíces reales $y_{1,2} = \pm 1$ y dos raíces complejas $y_{3,4} = \pm i$.

4) $y^2 + 1 = 0$ $(n = 2)$.

La ecuación no tiene raíces reales. Las raíces complejas son: $y_{1/2} = \pm i$.

5) $y^3 + 1 = 0$ $(n = 3)$.

La ecuación tiene una raíz real $y_1 = -1$ y dos raíces complejas $y_{2,3} = \dfrac{1 \pm i\sqrt{3}}{2}$.

6) $y^4 + 1 = 0$ $(n = 4)$.

La ecuación no tiene raíces reales. Las raíces complejas son:

$$y_{1,2} = \frac{\sqrt{2}}{2}(1 \pm i), \quad y_{3,4} = -\frac{\sqrt{2}}{2}(1 \pm i).$$

5.5. Ecuación bicuadrada. La ecuación algebraica de cuarto grado

$$ax^4 + bx^2 + c = 0,$$

donde a, b, c son ciertos números reales, se llama *ecuación bicuadrada*. Por sustitución de $x^2 = y$ la ecuación se reduce a la ecuación cuadrática $ay^2 + by + c = 0$ con la solución posterior de dos ecuaciones binomias $x^2 = y_1$ y $x^2 = y_2$ (y_1 e y_2 son las raíces de la ecuación cuadrática respectiva).

Si $y_1 \geqslant 0$ e $y_2 \geqslant 0$, entonces la ecuación bicuadrada tiene cuatro raíces reales:

$$x_{1,2} = \pm\sqrt{y_1}, \quad x_{3,4} = \pm\sqrt{y_2}.$$

Si $y_1 > 0$, $y_2 < 0$ *), la ecuación bicuadrada tiene dos raíces reales $x_{1,2} = \pm\sqrt{y_1}$ y dos raíces conjugadas ima-

*) El caso $y_1 < 0$, $y_2 \geqslant 0$ es análogo al caso que hemos examinado.

ginarias puras:

$$x_{3,4} = \pm i\sqrt{-y_2},$$

Si $y_1 < 0$ $y_2 < 0$, la ecuación bicuadrada tiene cuatro raíces conjugadas por pares imaginarias puras.

$$x_{1,2} = \pm i\sqrt{-y_1}, \quad x_{3,4} = \pm i\sqrt{-y_2}.$$

Si y_1 y y_2 son raíces complejas de la ecuación cuadrada respectiva, entonces para obtener cuatro raíces de la ecuación bicuadrada es necesario extraer las raíces cuadradas de los números complejos y_1 o y_2. Para evitar esta operación que es bastante voluminosa, transformemos la ecuación bicuadrada dada de la manera siguiente. Escribamos la ecuación inicial $ax^4 + bx^2 + c = 0$ en la forma

$$x^4 + px^2 + q = 0, \tag{11}$$

donde $p = b/a$, $q = c/a$, además, los coeficientes p y q, en vigor de la suposición sobre el carácter complejo de las raíces de la ecuación cuadrática, satisfacen las condiciones $|p| < 2\sqrt{q}$ y $q > 0$.

Efectuemos las transformaciones siguientes del primer miembro de la ecuación (11):

$$x^4 + px^2 + q = (x^4 + q) + px^2 =$$
$$= (x^4 + 2\sqrt{q}\,x^2 + q) - (2\sqrt{q} - p)\,x^2 = (x^2 + \sqrt{q})^2 -$$
$$- (2\sqrt{q} - p)\,x^2 = (x^2 + x\sqrt{2\sqrt{q} - p} + \sqrt{q})(x^2 -$$
$$- x\sqrt{2\sqrt{q} - p} + \sqrt{q}).$$

De esta manera la solución de la ecuación bicuadrada $ax^4 + bx^2 + c = 0$ se reduce a la solución de dos ecuaciones cuadráticas con los coeficientes reales:

$$x^2 + x\sqrt{2\sqrt{q} - p} + \sqrt{q} = 0,$$
$$x^2 - x\sqrt{2\sqrt{q} - p} + \sqrt{q} = 0.$$

La solución de las ecuaciones del tipo $ax^{2k} + bx^k + c = 0$ ($a \neq 0$, k es un número natural) por medio de la sustitución de $x^k = y$ se reduce a la solución de la ecuación cuadrática $ay^2 + by + c = 0$ con la solución posterior de las ecuaciones binomias respectivas.

5.6. Algunas ecuaciones de cuarto grado reducibles a ecuaciones cuadráticas. La ecuación algebraica de cuarto grado del tipo

$$ax^4 + bx^3 + cx^2 + dx + e = 0 \tag{12}$$

se llama *recíproca* ($e \neq 0$), si hay tal número $\lambda \neq 0$, que entre los coeficientes de la ecuación a, b, d, e existen las relaciones $d = \lambda\, b$, $e = \lambda^2 a$ o, lo que es lo mismo, tiene lugar la igualdad

$$\left(\frac{d}{b}\right)^2 = \frac{e}{a}.$$

Utilizando esta relación entre los coeficientes, la ecuación (12) se puede escribir en la forma

$$ax^4 + bx^3 + cx^2 + \lambda bx + \lambda^2 a = 0. \tag{13}$$

Puesto que $x = 0$ no es la raíz de la ecuación (12), entonces, dividiendo término a término ambos miembros de la ecuación (13) por x^2 y realizando la agrupación respectiva de los términos en el primer miembro de la ecuación, obtenemos la ecuación que es equivalente a la ecuación (13):

$$a\left(x^2 + \frac{\lambda^2}{x^2}\right) + b\left(x + \frac{\lambda}{x}\right) + c = 0.$$

Ahora por sustitución de $x + \frac{\lambda}{x} = y$ (teniendo en cuenta que $x^2 + \frac{\lambda^2}{x^2} = y^2 - 2\lambda$) la última ecuación se reduce a la ecuación cuadrática respecto a y:

$$ay^2 + by + c - 2\lambda a = 0. \tag{14}$$

Resolviendo la última ecuación, obtenemos que las raíces de la ecuación recíproca (12) se hallan como raíces de dos ecuaciones cuadráticas:

$$x^2 - y_1 x + \lambda = 0,$$
$$x^2 - y_2 x + \lambda = 0,$$

donde y_1 e y_2 son raíces de la ecuación (14).

Las ecuaciones *simétricas* que corresponden a $\lambda = 1$)

$$ax^4 + bx^3 + cx^2 + bx + a = 0$$

y las ecuaciones *antisimétricas* (las que corresponden a $\lambda = -1$)

$$ax^4 + bx^3 + cx^2 - bx + a = 0,$$

son casos particulares de las ecuaciones recíprocas. Estas ecuaciones se reducen a las ecuaciones cuadráticas mediante la sustitución $x + \frac{1}{x} = y$ para la ecuación simétrica y la sustitución $x - \frac{1}{x} = y$ para la antisimétrica.

La ecuación de cuarto grado de la forma

$$(x^2 + bx + c)(x^2 + bx + d) = k, \tag{15}$$

donde b, c, d, k son ciertos números reales, también se reduce a la ecuación cuadrática

$$y^2 + (c + d)y - k = 0 \tag{16}$$

por sustitución de

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

Si la ecuación (16) tiene unas raíces reales y_1 e y_2, entonces las raíces de la ecuación (15) se hallan como las raíces de dos ecuaciones cuadráticas con coeficientes reales:

$$x^2 + bx - y_1 = 0,$$
$$x^2 + bx - y_2 = 0.$$

La solución de la ecuación que tiene la forma

$$x(x + a)(x + b)(x + a + b) = c \tag{17}$$

puede ser reducida a la solución de dos ecuaciones cuadráticas mediante el procedimiento siguiente.

Uniendo el primer factor que se encuentra en el primer miembro de la ecuación (17) con el cuarto, y el segundo con el tercero, obtenemos la ecuación

$$[x^2 + (a + b)x][x^2 + (a + b)x + ab] = c,$$

la cual mediante la sustitución

$$x^2 + (a + b)x = y$$

se reduce a la ecuación cuadrática respecto a la nueva incógnita y:

$$y^2 + aby - c = 0. \tag{18}$$

Si la ecuación (18) tiene las raíces y_1 e y_2, entonces el conjunto de raíces de la ecuación (17) se halla como el conjunto de raíces de las dos ecuaciones cuadráticas con coeficientes reales siguientes:

$$x^2 + (a + b)x - y_1 = 0,$$
$$x^2 + (a + b)x - y_2 = 0.$$

5.7. Solución de una ecuación algebraica con coeficientes enteros. Las raíces racionales de la ecuación algebraica de n-ésimo grado

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0, \tag{19}$$

donde $a_0, a_1, a_2, \ldots, a_{n-1}, a_n$ son números enteros, puede hallarse utilizando la regla siguiente:

Sólo los números $\frac{m}{p}$ (m es entero, p, natural), donde el número $|m|$ es el divisor del número $|a_n|$, y el número p, el divisor del número $|a_0|$, pueden ser raíces racionales de la ecuación (19).

Ejemplo. Hallar las raíces de la ecuación

$$4x^4 + 8x^3 - 3x^2 - 7x + 3 = 0. \tag{20}$$

Los divisores del número 3 serán los números 1, 3, y los divisores del número 4, los números 1, 2, 4. El conjunto de valores m será $\{1; -1; 3; -3\}$ y el conjunto de valores p, $\{1, 2, 4\}$. El conjunto de distintos números racionales cualesquiera será $\left\{\pm 1;\ \pm 2;\ \pm\frac{1}{2};\ \pm\frac{1}{4};\ \pm\frac{3}{2};\ \pm\frac{3}{4}\right\}$. Sustituyendo estos números en la ecuación, hallamos que los números $\frac{1}{2}$ y $-\frac{3}{2}$ son las raíces de la ecuación dada. Por consiguiente, según el teorema de Bezout el polinomio dado se divide por los polinomios lineales

$$\left(x-\frac{1}{2}\right) \text{ y } \left(x+\frac{3}{2}\right)$$

y, por consiguiente, se divide también por su producto

$$\left(x-\frac{1}{2}\right)\left(x+\frac{3}{2}\right) = x^2 + x - \frac{3}{4}.$$

Efectuando la división mediante el «ángulo», hallamos el polinomio del cociente:

$$4x^2 + 4x - 4.$$

Resolviendo la ecuación cuadrática

$$4x^2 + 4x - 4 = 0,$$

obtenemos dos raíces reales más de la ecuación (20):

$$x_3 = \frac{-1+\sqrt{5}}{2}, \qquad x_4 = \frac{-1-\sqrt{5}}{2}.$$

Así pues, el problema está completamente resuelto, es decir, hemos hallado todas las cuatro raíces de la ecuación inicial:

$$x_1 = \frac{1}{2}, \qquad x_2 = -\frac{3}{2}, \qquad x_3 = \frac{-1+\sqrt{5}}{2}, \qquad x_4 = \frac{-1-\sqrt{5}}{2}.$$

5.8. Ecuaciones algebraicas racionales. La ecuación que tiene la forma

$$\frac{P(x)}{Q(x)} = 0, \tag{21}$$

donde $P(x)$ y $Q(x)$ son polinomios, se llama ecuación algebraica *racional*. A continuación, para mayor precisión, vamos a suponer que $P(x)$ es un polinomio de m-ésimo grado, y $Q(x)$, un polinomio de n-ésimo grado.

Un conjunto de valores admisibles de la ecuación algebraica racional (21) está dado por la condición $Q(x) \neq 0$, es decir, $x \neq c_1$, $x \neq c_2$, . . ., $x \neq c_n$, donde $c_1, c_2, \ldots, c_n$ son las raíces del polinomio $Q(x)$.

El método de solución de la ecuación (21) consiste en lo siguiente. Resolvemos la ecuación

$$P(x) = 0,$$

cuyas raíces designamos a través de

$$x_1, x_2, x_3, \ldots, x_m.$$

Comparamos los conjuntos de raíces de los polinomios $P(x)$ y $Q(x)$. Si ninguna raíz del polinomio $P(x)$ es raíz del polinomio $Q(x)$, entonces todas las raíces del polinomio $P(x)$ son raíces de la ecuación (21). Si cualquier raíz del polinomio $P(x)$ es raíz del polinomio $Q(x)$, entonces es necesario comparar sus multiplicidades: si la multiplicidad de una raíz del polinomio $P(x)$ es mayor que la multiplicidad de una raíz del polinomio $Q(x)$, entonces esta raíz es la raíz de la ecuación (21) con la multiplicidad, igual a la diferencia de multiplicidades de las raíces del dividendo y del divisor;

en caso contrario la raíz del polinomio $P(x)$ no es raíz de la ecuación racional (21).

Ejemplo. Hallamos las raíces reales de la ecuación

$$\frac{P(x)}{Q(x)}=0,$$

donde $P(x)=x^4-1$, $Q(x)=x-1$.

El polinomio $P(x)$ tiene dos raíces reales (las dos son simples):

$$x_1=1, \quad x_2=-1.$$

El polinomio $Q(x)$ tiene una raíz simple, $c=1$. Por consiguiente, la ecuación tiene una raíz real $x=-1$.

Resolviendo la misma ecuación en el conjunto de números complejos, obtendremos, que la ecuación $\frac{P(x)}{Q(x)}=0$ tiene, además de la raíz real indicada, dos raíces conjugadas complejamente:

$$x_2=i, \; x_3=-2$$

5.9. Ecuaciones irracionales. La ecuación que contiene una incógnita (o bien una expresión racional algebraica de la incógnita) bajo el signo del radical se llama ecuación *irracional*. En las matemáticas elementales las soluciones de las ecuaciones irracionales se buscan en el conjunto de los números reales.

Cualquier ecuación irracional mediante las operaciones algebraicas (de multiplicación, división, elevación a una potencia entera de los dos miembros de la ecuación) puede ser reducida a una ecuación algebraica racional. Al mismo tiempo es necesario tener en cuenta que la ecuación algebraica racional obtenida puede resultar ser no equivalente a la ecuación irracional inicial, es decir, puede contener raíces «sobrantes», las cuales no son raíces de la ecuación irracional inicial. Por eso, hallando las raíces de la ecuación algebraica racional obtenida, es necesario verificar, si todas las raíces de la ecuación racional son raíces de la ecuación irracional.

En forma general es muy difícil indicar algún método universal de solución de cualquiera ecuación irracional, puesto que es deseable que como resultado de las transformaciones de la ecuación irracional inicial se obtenga no sólo cierta ecuación algebraica racional, entre las raíces de la cual se encuentren también las raíces de la ecuación irracional dada, sino una ecuación algebraica racional formada de polinomios del menor grado posible. El deseo de obtener una ecuación algebraica racional, formada de polinomios del menor grado posible es muy natural, puesto que la obtención de todas las raíces de la ecuación algebraica racional puede resultar ser por sí misma un problema bastante difícil, cuya resolución es posible sólo en un número limitado de casos. Citamos aquí algunos métodos estándar que se usan con mayor frecuencia para resolver las ecuaciones algebraicas irracionales.

1) Uno de los procedimientos más simples para resolver las ecuaciones irracionales es el método de liberación de radicales mediante la elevación sucesiva de ambos miembros de la ecuación a la potencia natural respectiva. En este caso es necesario tener en cuenta que cuando se elevan ambos miembros de la ecuación a una potencia impar, se obtiene una ecuación equivalente a la inicial, y cuando se elevan ambos miembros de la ecuación a una potencia par, se obtiene una ecuación que, generalmente, no es equivalente a la ecuación inicial. Es fácil convencerse de esto elevando ambos miembros de la ecuación

$$f(x) = g(x)$$

a cualquier potencia par. De resultas de esta operación se obtiene la ecuación

$$[f(x)]^{2n} = [g(x)]^{2n},$$

cuyo conjunto de soluciones representa la unión de los conjuntos de soluciones de dos ecuaciones:

$$f(x) = g(x) \text{ y } f(x) = -g(x).$$

Sin embargo, a pesar de este defecto, precisamente el proceso de elevación de ambos miembros de la ecuación a cierta potencia (muy a menudo, par) es el proceso más divulgado de reducción de una ecuación irracional a una ecuación racional.

Ejemplo 1. Resolver la ecuación

$$\sqrt{P(x)} + \sqrt{Q(x)} = R(x), \tag{22}$$

donde $P(x)$, $Q(x)$, $R(x)$ son ciertos polinomios.

En vigor de la definición de la operación de extracción de la raíz en el conjunto de números reales, los valores admisibles de la incógnita x se definen por las condiciones

$$P(x) \geqslant 0, \quad Q(x) \geqslant 0.$$

Elevando ambos miembros de la ecuación (22) al cuadrado, obtenemos la ecuación

$$2\sqrt{P(x)\,Q(x)} = R^2(x) - P(x) - Q(x).$$

Luego de una reiterada elevación al cuadrado la ecuación se convierte en la ecuación algebraica

$$4P(x)\,Q(x) = [R^2(x) - P(x) - Q(x)]^2. \qquad (23)$$

Puesto que ambos miembros de la ecuación (22) fueron elevados al cuadrado, puede resultar que no todas las raíces de la ecuación (23) sean resoluciones de la ecuación inicial, y que sea necesaria la verificación de las raíces.

2) Otro procedimiento de resolución de las ecuaciones irracionales es el método de introducción de nuevas incógnitas, respecto a las cuales se obtiene o una ecuación irracional más simple o una ecuación racional.

Ejemplo 2. Resolver la ecuación irracional

$$(3-x)\sqrt[3]{\frac{3-x}{x-1}} + (x-1)\sqrt[3]{\frac{x-1}{3-x}} = 2.$$

El conjunto de valores admisibles de esta ecuación es:

$$x \in (-\infty;\ 1) \cup (1;\ 3) \cup (3;\ +\infty).$$

Poniendo $\sqrt[3]{\frac{3-x}{x-1}} = y$, después de la sustitución obtenemos la ecuación

$$(3-x)\,y + \frac{x-1}{y} = 2$$

o la ecuación equivalente

$$(3-x)\,y^2 - 2y + x - 1 = 0,$$

la cual puede considerarse como ecuación cuadrática respecto a y. Resolviendo esta ecuación, obtenemos

$$y_1 = 1, \quad y_2 = \frac{x-1}{3-x}.$$

Por consiguiente, el conjunto de soluciones de una ecuación irracional inicial representa una unión de conjuntos de soluciones de las dos ecuaciones siguientes:

$$\sqrt[3]{\frac{3-x}{x-1}}=1, \quad \sqrt[3]{\frac{3-x}{x-1}}=\frac{x-1}{3-x}.$$

Elevando ambos miembros de cada una de estas ecuaciones al cubo, obtenemos dos ecuaciones algebraicas racionales:

$$\frac{3-x}{x-1}=1, \quad \frac{3-x}{x-1}=\left(\frac{x-1}{3-x}\right)^3.$$

Resolviendo estas ecuaciones, hallamos que la ecuación irracional dada tiene una sola raíz $x=2$.

Concluyendo podemos anotar que para resolver las ecuaciones irracionales no hace falta comenzar a resolver la ecuación con la elevación de ambos miembros de las ecuaciones a una potencia natural, tratando de reducir la solución de la ecuación irracional a la solución de la ecuación algebraica racional. Primeramente es necesario ver, si es posible efectuar alguna transformación idéntica de la ecuación, la cual puede sustancialmente simplificar su solución.

Ejemplo 3. Resolver la ecuación

$$a\sqrt[4]{1+x}+\frac{a}{x}\sqrt[4]{1+x}=\sqrt[4]{x}. \tag{24}$$

El conjunto de los valores admisibles de la ecuación dada es: $x\in(0;\ +\infty)$. Efectuemos las transformaciones siguientes de la ecuación dada:

$$a\sqrt[4]{1+x}+\frac{a}{x}\sqrt[4]{1+x}=\sqrt[4]{x} \Leftrightarrow a\sqrt[4]{1+x}\left(1+\frac{1}{x}\right)=$$
$$=\sqrt[4]{x} \Leftrightarrow a\sqrt[4]{1+\frac{1}{x}}\left(1+\frac{1}{x}\right)=1.$$

Luego, escribiendo la ecuación en la forma

$$a\left(1+\frac{1}{x}\right)^{5/4}=1,$$

obtenemos:

para $a=0$, la ecuación no tiene soluciones;

para $a \neq 0$, la ecuación puede ser escrita en la forma

$$\left(1+\frac{1}{x}\right)^{5/4}=\frac{1}{a}.$$

Para $a < 0$ la ecuación dada no tiene soluciones, puesto que para cualquier x que pertenece al conjunto de valores admisibles de la ecuación, la expresión que se encuentra en el primer miembro de la ecuación es positiva.

Para $a > 0$ la ecuación tiene la solución

$$x=\frac{1}{\sqrt[5]{1/a^4-1}}.$$

Teniendo en cuenta que el conjunto de soluciones admisibles de la ecuación se define por la condición $x > 0$, obtenemos definitivamente:

Para $0 < a < 1$ la solución de la ecuación irracional (24) será

$$x=\frac{1}{\sqrt[5]{1/a^4-1}}.$$

Para todos los demás valores a la ecuación no tiene soluciones, es decir, el conjunto de sus soluciones es un conjunto vacío.

5.10. Ecuaciones que contienen una incógnita bajo el signo del valor absoluto. Las ecuaciones que contienen una incógnita bajo el signo del valor absoluto pueden reducirse a las ecuaciones que no contienen el signo del valor absoluto, utilizando la definición del módulo. Así, por ejemplo, la solución de la ecuación

$$x^2-5\left|x-\frac{5}{2}\right|-\frac{13}{2}=0 \tag{25}$$

se reduce a la solución de dos ecuaciones con condiciones adicionales.

1) Si $x-\frac{5}{2} \geqslant 0$, entonces la ecuación (25) se reduce a la forma

$$x^2-5x+6=0. \tag{26}$$

Las soluciones de esta ecuación son: $x_1 = 2$, $x_2 = 3$. La condición $x-\frac{5}{2} \geqslant 0$ se satisface por la segunda raíz de la ecuación cuadrática (26), y el número 3 es la raíz de la ecuación (25).

2) Si $x - \frac{5}{2} < 0$, entonces la ecuación (25) se reduce a la forma

$$x^2 + 5x - 19 = 0.$$

Las raíces de esta ecuación son los números $x_1 = \frac{-5+\sqrt{101}}{2}$ y $x_2 = \frac{-5-\sqrt{101}}{2}$. La primera raíz $x_1 = \frac{-5+\sqrt{101}}{2}$ no satisface la condición $x - \frac{5}{2} < 0$ y por eso no es la solución de la ecuación dada (25).

De esta manera, las soluciones de la ecuación (25) serán los números 3 y $\frac{-5-\sqrt{101}}{2}$.

Observemos, que los coeficientes de la ecuación que contiene una incógnita bajo el signo del valor absoluto pueden

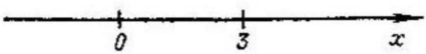


Fig. 4.1.

ser elegidos de tal manera, que las soluciones de la ecuación serán todos los valores de la incógnita que pertenecen a cierto intervalo del eje numérico. Por ejemplo, resolvamos la ecuación

$$|x| + |3 - x| = 3. \tag{27}$$

Examinemos el eje numérico Ox y marquemos sobre él los puntos 0 y 3 (ceros de las funciones situados bajo el signo del valor absoluto). Estos puntos partirán el eje numérico en tres intervalos (fig. 4.1.):

$$-\infty < x < 0, \quad 0 \leqslant x \leqslant 3, \quad 3 < x < +\infty.$$

1) Para $x \in (-\infty; 0)$ la ecuación (27) se reduce a la forma

$$3 - 2x = 3.$$

En el intervalo $(-\infty; 0)$ la última ecuación no tiene soluciones. Análogamente, para $x \in (3; +\infty)$ la ecuación (27) se reduce a la forma

$$2x - 3 = 3$$

y en el intervalo $(3; +\infty)$ no tiene soluciones.

2) Para $x \in [0; 3]$ la ecuación (27) se reduce a la forma

$$x + (3 - x) = 3,$$

es decir, se convierte en identidad. Por consiguiente, cualquier valor $x \in [0; 3]$ es la solución de la ecuación (27).

5.11. Solución de ecuaciones en el conjunto de números complejos. Examinemos dos ejemplos de solución de ecuaciones en el conjunto de números complejos. Las dos ecuaciones se resuelven mediante la utilización de las propiedades más simples de los números complejos.

Ejemplo 1. Hallar los números complejos z que satisfacen la ecuación*)

$$z|z| + az + i = 0, \tag{28}$$

donde a es cierto número real positivo.

Según la definición cualquier número complejo z se representa en la forma

$$z = x + iy, \tag{29}$$

donde x, y son números reales.

Según la definición del módulo de un número complejo

$$|z| = \sqrt{x^2 + y^2}. \tag{30}$$

Sustituyendo en la ecuación (28) z y $|z|$ por sus expresiones a través de x e y, representadas por las fórmulas (29) y (30), obtenemos la siguiente expresión de la ecuación (28):

$$x\sqrt{x^2+y^2} + ax + i\left[y\sqrt{x^2+y^2} + ay + 1\right] = 0 + 0i. \tag{31}$$

De la definición de la igualdad de dos números complejos se desprende que la ecuación (31) es equivalente al sistema de dos ecuaciones

$$\begin{aligned} x\sqrt{x^2+y^2} + ax &= 0, \\ y\sqrt{x^2+y^2} + ay + 1 &= 0, \end{aligned} \tag{32}$$

cuyas soluciones ya se hallan en el conjunto de números reales.

---

*) Designemos aquí el valor incógnito no por la letra x, sino por la letra z, subrayando de esta manera que las soluciones de las ecuaciones se hallan en el conjunto de los números complejos.

No es difícil ver que el conjunto de soluciones de la primera ecuación del sistema (32) puede ser hallado como la unión de los conjuntos de las soluciones de dos ecuaciones:

$$x=0 \quad \text{y} \quad \sqrt{x^2+y^2}+a=0.$$

La segunda de estas ecuaciones no tiene soluciones en vigor de la condición $a>0$; por eso ahora se puede deducir que las soluciones de la ecuación inicial (28) son sólo números imaginarios puros.

Sustituyendo $x=0$ en la segunda ecuación del sistema (32), obtenemos la ecuación para el número real y*):

$$y|y|+ay+1=0,$$

el conjunto de soluciones del cual se obtiene como la unión de los conjuntos de soluciones de dos sistemas:

$$\left.\begin{aligned} &y \geqslant 0, \\ &y^2+ay+1=0; \end{aligned}\right\} \quad (33) \qquad \left.\begin{aligned} &y \leqslant 0 \\ &-y^2+ay+1=0. \end{aligned}\right\} \quad (34)$$

Es fácil convencerse de que, teniendo en cuenta la condición $a>0$, el sistema (33) no tiene soluciones, y el sistema (34) tiene una sola solución

$$y=\frac{a-\sqrt{a^2+4}}{2}.$$

De este manera la solución de la ecuación (28) es un número imaginario puro

$$z=i\,\frac{a-\sqrt{a^2+4}}{2}.$$

Ejemplo 2. Hallar los números complejos z que satisfacen la ecuación

$$\bar{z}=z^{n-1}, \tag{35}$$

donde $\bar{z}$ es un número complejo conjugado z, y n, un número natural.

Para dos números cualesquiera $\bar{z}$ y z, conjugados complejamente, tenemos

$$|z|=|\bar{z}|.$$

Teniendo en cuenta esta igualdad obtenemos, que de la ecuaaión (35) se desprende la ecuación

$$|z|=|z|^{n-1}. \tag{36}$$

La ecuación (36) es una ecuación algebraica de $(n-1)$-ésimo grado respecto a la variable $|z|$. El conjunto de soluciones de esta ecuación satisface las condiciones:

1) para $n=1$, $|z|=1$;
2) para $n\neq 1$, $(n\in N)$ $|z|=1$ y $|z|=0$.

*) Aquí $|y|$ es la designación del valor absoluto del número real y.

Examinemos el caso, cuando $|z| = 1$. Multiplicando ambos miembros de la ecuación (35) por z, obtenemos la ecuación

$$z^n = 1, \tag{37}$$

la cual es equivalente a la ecuación (35), puesto que $|z| \neq 0$, por consiguiente, también $z \neq 0$.

El conjunto de soluciones de la ecuación (37) se describe por las fórmulas

$$z_k = \cos\frac{2\pi k}{n} + i \operatorname{sen}\frac{2\pi k}{n} \quad (k = 0, 1, 2, \ldots, n-1).$$

Así pues, definitivamente las soluciones de la ecuación (35) tienen la forma

1) $$z = \cos\frac{2\pi k}{n} + i \operatorname{sen}\frac{2\pi k}{n} \quad (k = 0, 1, 2, \ldots, n-1)$$

para cualquier n natural;

2) $z = 0$ para cualquier $n \neq 1$.

5.12. Ecuaciones diofánticas. La ecuación algebraica con una o varias incógnitas, cuyos coeficientes son todos números enteros y las soluciones se hallan en el conjunto de números enteros, se llama *ecuación diofántica.*

Tales ecuaciones o no tienen soluciones o tienen un número finito o infinito de soluciones.

La ecuación diofántica más simple es la ecuación lineal con dos incógnitas, x e y:

$$ax + by = c, \tag{38}$$

donde a, b, c son números enteros y es necesario hallar las soluciones enteras de la ecuación.

La solución completa de esta ecuación puede ser hallada mediante la siguiente importante consecuencia del algoritmo de Euclides (véase p. 1.5 del capítulo 2):

Si el número d es el máximo común divisor de los números enteros a y b, entonces existen tales números enteros k y l que $d = ka + lb$. El algoritmo de Euclides permite calcular los números enteros k y l.

La ecuación diofántica lineal (38) no tendrá soluciones, si los números c y d son recíprocamente primos. Si el número c es múltiplo del número d (es decir, es representable en la forma $c = pd$, donde p es un número entero, distinto de cero), entonces una de las soluciones de la ecuación (38) tendrá la forma

$$x^* = pk, \qquad y^* = pl.$$

El conjunto de todas las soluciones de la ecuación diofántica (38) está dado por las fórmulas

$$x = x^* + \frac{nb}{d}, \qquad y = y^* - \frac{na}{d},$$

donde $n \in \mathbf{Z}$.

§ 6. Ecuaciones trascendentes

La ecuación que no se reduce a la ecuación algebraica mediante las transformaciones algebraicas, se llama *ecuación trascendente* *).

Las ecuaciones trascendentes más simples son las trigonométrcas, logarítmicas y exponenciales.

6.1. Ecuaciones exponenciales. Se llama ecuación *exponencial* aquella en la cual la incógnita forma parte sólo de los exponentes de potencias para ciertas bases constantes.

Una de las ecuaciones exponenciales más simples, cuya solución se reduce a la de una ecuación algebraica, es la ecuación del tipo

$$a^{f(x)} = b, \tag{1}$$

donde a y b son ciertos números positivos ($a \neq 1$). La ecuación exponencial (1) es equivalente a la ecuación algebraica

$$f(x) = \log_a b.$$

En el caso más simple, cuando $f(x) = x$, la ecuación exponencial (1) tiene la solución

$$x = \log_a b.$$

El conjunto de soluciones de la ecuación exponencial que tiene la forma

$$R(a^x) = 0, \tag{2}$$

donde R es cierto polinomio, se halla de la manera siguiente.

Se introduce una variable nueva $y = a^x$ y la ecuación (2) se resuelve como una ecuación algebraica respecto a la incógnita y. Después la solución de la ecuación inicial (2) se reduce a la solución de las ecuaciones exponenciales más simples que tienen la forma (1).

---

*) Por *transformaciones algebraicas* de la ecuación

$$F = 0$$

se entiende las transformaciones siguientes:

1) la adición a ambos miembros de la ecuación una misma expresión algebraica;

2) la multiplicación de ambos miembros de la ecuación por una misma expresión algebraica;

3) la elevación de ambos miembros de la ecuación a una potencia racional.

Ejemplo 1. Resolver la ecuación

$$9\cdot 8^x - 18\cdot 4^x - 2\cdot 2^x + 4 = 0.$$

Escribiendo la ecuación en la forma

$$9\,(2^x)^3 - 18\,(2^x)^2 - 2\cdot 2^x + 4 = 0$$

e introduciendo una variable nueva $y = 2^x$, obtenemos la ecuación cúbica respecto a la variable y:

$$9y^3 - 18y^2 - 2y + 4 = 0.$$

Según los resultados del § 4, capítulo 2, es muy fácil convencerse que la ecuación cúbica dada tiene una sola raíz racional $y_1 = 2$ y dos raíces irracionales $y_2 = \frac{\sqrt{2}}{3}$ e $y_3 = -\frac{\sqrt{2}}{3}$.

De esta manera, la solución de la ecuación inicial se reduce a la solución de las ecuaciones exponenciales más simples:

$$2^x = 2, \quad 2^x = \frac{\sqrt{2}}{3}, \quad 2^x = -\frac{\sqrt{2}}{3}.$$

La última de las ecuaciones enumeradas no tiene soluciones. El conjunto de soluciones de las ecuaciones primera y segunda es:

$$x = 1 \quad \text{y} \quad x = \log_2 \frac{\sqrt{2}}{3}.$$

Algunas ecuaciones exponenciales más simples:

1) La ecuación que tiene la forma

$$\alpha a^{2x} + \beta a^x + \gamma = 0$$

por sustitución $a^x = y$ se reduce a la ecuación cuadrática

$$\alpha y^2 + \beta y + \gamma = 0.$$

2) La ecuación que tiene la forma

$$\alpha a^x + \beta a^{-x} + \gamma = 0$$

por sustitución $a^x = y$ se reduce a la ecuación cuadrática

$$\alpha y^2 + \gamma y + \beta = 0.$$

3) La ecuación del tipo

$$\alpha a^{2x} + \beta (ab)^x + \gamma b^{2x} = 0$$

por sustitución $\left(\frac{a}{b}\right)^x = y$ se reduce a la ecuación cuadrática

$$\alpha y^2 + \beta y + \gamma = 0.$$

6.2. Ecuaciones logarítmicas. La ecuación, en la cual la incógnita forma parte como argumento de la función logarítmica, se llama ecuación *logarítmica*.

La ecuación logarítmica más simple es la que tiene la forma

$$\log_a f(x) = b, \tag{3}$$

donde a es cierto número positivo distinto de la unidad, b es cualquier número real. La ecuación logarítmica (3) es equivalente a la ecuación algebraica

$$f(x) = a^b.$$

El conjunto de soluciones de la ecuación logarítmica que tiene la forma $R(\log_a x) = 0$, donde R es cierto polinomio de la incógnita indicada, se halla de la manera siguiente.

Se introduce una variable nueva $\log_a x = y$ y la ecuación (3) se resuelve como una ecuación algebraica respecto a y. Después se resuelven las ecuaciones logarítmicas más simples que tienen la forma (3).

Ejemplo 1. Resolver la ecuación

$$\log_2^2 x + \log_2 x - 2 = 0 \tag{4}$$

Respecto a la incógnita $\log_2 x = y$ la ecuación dada es cuadrática:

$$y^2 + y - 2 = 0.$$

Las raíces de esta ecuación son: $y_1 = 1$, $y_2 = -2$.

Resolviendo las ecuaciones logarítmicas

$$\log_2 x = 1, \quad \log_2 x = -2,$$

obtenemos las soluciones de la ecuación (4): $x_1 = 2$, $x_2 = 1/4$.

En ciertos casos, para reducir la solución de una ecuación logarítmica a la solución sucesiva de una ecuación algebraica y de las ecuaciones logarítmicas más simples, es necesario efectuar previamente las transformaciones de logaritmos convenientes que entran en la ecuación. Tales transformaciones pueden ser: la transformación de la suma de logaritmos de dos valores en el logaritmo del producto de estos valores, el paso de un logaritmo con una base a un logaritmo con otra base, etc.

Ejemplo 2. Resolver la ecuación

$$2\sqrt[3]{2\log_{16}^2 x}-\sqrt[3]{\log_2 x}-6=0. \tag{5}$$

Para reducir la solución de la ecuación dada a la solución sucesiva de las ecuaciones algebraica y logarítmicas más simples, es necesario ante todo reducir todos los logaritmos a una base (aquí, por ejemplo, a la base 2). Para esto utilicemos la fórmula

$$\log_a N=\frac{\log_b N}{\log_b a},$$

en vigor de la cual $\log_{16} x=\frac{\log_2 x}{\log_2 16}=\frac{\log_2 x}{4}$. Al sustituir en la ecuación (5) el $\log_{16} x$ por la magnitud igual a $\frac{\log_2 x}{4}$, obtenemos la ecuación

$$\sqrt[3]{\log_2^2 x}-\sqrt[3]{\log_2 x}-6=0.$$

Mediante la sustitución $\sqrt[3]{\log_2 x}=y$ esta ecuación se reduce a la ecuación cuadrática respecto a la incógnita y:

$$y^2-y-6=0.$$

Las raíces de esta ecuación son: $y_1=3$, $y_2=-2$. Resolvemos las ecuaciones $\sqrt[3]{\log_2 x}=3$ y $\sqrt[3]{\log_2 x}=-2$:

$$\sqrt[3]{\log_2 x}=3 \Longleftrightarrow \log_2 x=27 \Longleftrightarrow x=2^{27},$$

$$\sqrt[3]{\log_2 x}=-2 \Longleftrightarrow \log_2 x=-8 \Longleftrightarrow x=2^{-8}.$$

Ejemplo 3. Resolver la ecuación

$$\log_3(2x-1)-\log_3(x-1)=1.$$

Transformando la diferencia de los logaritmos de dos magnitudes en el logaritmo del cociente de estas magnitudes:

$$\log_3(2x-1)-\log_3(x-1)=\log_3\frac{2x-1}{x-1},$$

reducimos la ecuación dada a la ecuación logarítmica más simple

$$\log_3\frac{2x-1}{x-1}=1 \Leftrightarrow \frac{2x-1}{x-1}=3 \Leftrightarrow x=2.$$

§ 7. Sistemas de ecuaciones. Sistemas de ecuaciones lineales

7.1. Definiciones fundamentales. El conjunto de valores de las incógnitas que convierten al mismo tiempo todas las ecuaciones del sistema en identidades se llama *solución* de cierto conjunto (sistema) de ecuaciones

$$F_1(x_1, x_2, \ldots, x_n)=0, \ \ldots, \ F_m(x_1, x_2, \ldots, x_n)=0$$

con las incógnitas $x_1, x_2, \ldots, x_n$. El sistema de ecuaciones se considera resuelto, si se hallan todas tales soluciones.

Si el sistema no tiene soluciones, entonces se dice que es *incompatible*.

A veces el sistema de ecuaciones se escribe uniéndolas entre llaves.

Dos sistemas de ecuaciones se consideran equivalentes, si tienen el mismo conjunto de soluciones. (Según la definición dos sistemas incompatibles se consideran equivalentes).

Se dice, que el sistema de ecuaciones (S) es equivalente a dos sistemas de ecuaciones (S_1) y (S_2), si el conjunto de soluciones del sistema (S) coincide con la unión de los conjuntos de soluciones de los sistemas (S_1) y (S_2).

Las propiedades de los sistemas de ecuaciones son:

1) Si cambiamos cualquier ecuación del sistema por una ecuación equivalente, obtenemos un sistema equivalente.

2) Si una de las ecuaciones del sistema (S) es equivalente a ciertas dos ecuaciones, entonces el sistema inicial (S) es equivalente a dos sistemas (S_1) y (S_2), en cada uno de los cuales esta ecuación está sustituida por una de las ecuaciones del conjunto equivalente, y las demás quedan sin cambiar.

Por ejemplo, el sistema

$$x^3 + y^3 = 3(x + y),$$
$$x^2 + y^2 = 7,$$

en el cual la primera ecuación es equivalente a dos ecuaciones: $x + y = 0$ y $x^2 - xy + y^2 = 3$, es equivalente a dos sistemas:

$$x + y = 0, \qquad x^2 - xy + y^2 = 3,$$
$$x^2 + y^2 = 7; \qquad x^2 + y^2 = 7.$$

7.2. Sistemas de ecuaciones lineales. El sistema que tiene la forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + x_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n &= b_s \end{aligned} \tag{1}$$

se llama *sistema a de ecuaciones lineales* con n incógnitas $x_1, x_2, \ldots, x_n$. Los valores $a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots$ $\ldots, a_{2n}, \ldots, a_{s1}, a_{s2}, \ldots, a_{sn}$ se llaman *coeficientes* del sistema lineal de ecuaciones dado. Los índices que tienen los coeficientes del sistema lineal significan lo siguiente: el primer índice indica el número de la ecuación del sistema en la expresión (1), el segundo índice indica el número de la incógnita, con el cual se encuentra el coeficiente dado. Así, por ejemplo, a_{25} es un coeficiente que se encuentra en la segunda ecuación del sistema (1) para la incógnita x_5.

Los valores $b_1, b_2, \ldots, b_s$ se llaman *términos independientes* de la primera, segunda, ..., s-ésima ecuación del sistema (1).

El sistema de ecuaciones (1) se llama *homogéneo*, si todos los números b_i son iguales a cero ($i = 1, 2, 3, \ldots, s$), y *no homogéneo*, si por lo menos un b_i es distinto de cero.

Un conjunto ordenado n de los números $k_1, k_2, k_3, \ldots$ $\ldots, k_n$ se llama *solución del sistema* (1), si al sustituirlo en el sistema en lugar de las incógnitas $x_1, x_2, \ldots, x_n$ todas las ecuaciones del sistema se convierten en identidades. El sistema de ecuaciones se llama *compatible*, si tiene por lo menos una solución, y se llama *incompatible*, si no tiene ninguna solución.

Un sistema compatible de ecuaciones lineales se llama *determinado*, si tiene una sola solución, es decir, si existe sólo una lista n de números $k_1, k_2, \ldots, k_n$, la cual convierte todas las ecuaciones del sistema en identidades.

Un sistema compatible de ecuaciones lineales se llama *indeterminado*, si existen más soluciones que una.

El sistema de ecuaciones homogéneas siempre tiene una solución nula

$$x_1 = x_2 = x_3 = \ldots = x_n = 0.$$

Si el sistema de ecuaciones homogéneas tiene una solución distinta de cero $k_1, k_2, \ldots, k_n$ (es decir, por lo menos uno de los números k_i $(i = 1, 2, \ldots, n)$ es distinto de cero), entonces tal sistema tiene conjunto innumerable de soluciones del tipo $lk_1, lk_2, \ldots, lk_n$, donde l es un número cualquiera.

7.3. Método de eliminación sucesiva de incógnitas (método de Gauss). Uno de los métodos más divulgados de solución de sistemas de ecuaciones algebraicas lineales es el método de eliminación sucesiva de incógnitas, método de Gauss. Este método se basa en ciertas transformaciones del sistema de ecuaciones lineales, de resultas de las cuales se obtiene el sistema, equivalente al sistema inicial. Efectuemos estas transformaciones con un ejemplo del sistema de tres ecuaciones lineales con tres incógnitas

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1, \\ a_{21}x + a_{22}y + a_{23}z &= b_2, \\ a_{31}x + a_{32}y + a_{33}z &= b_3. \end{aligned} \tag{2}$$

1) Si ambos miembros de cualquiera ecuación del sistema los multiplicamos por un mismo número (que no es igual a cero), entonces el sistema obtenido es equivalente al inicial (es decir, ambos son incompatibles, o ambos son compatibles y los conjuntos de sus soluciones coinciden).

Por ejemplo, el sistema

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ ca_{21}x + ca_{22}y + ca_{23}z &= cb_2 \qquad (c \neq 0), \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \end{aligned}$$

es equivalente al sistema (2).

2) Si ambos miembros de cualquier ecuación del sistema, multiplicados por cierto número distinto de cero, los sustraemos de los miembros respectivos de otra ecuación y componemos un sistema, en el cual en vez de una ecuación de las citadas más arriba se encuentra una ecuación, obtenida como resultado de la sustracción, y las demás ecuaciones quedan sin cambiarse, entonces el sistema obtenido es igual al inicial.

Por ejemplo, cada uno de los sistemas

$$\begin{gathered} a_{11}x + a_{12}y + a_{13}z = b_1, \\ (a_{11} - ca_{21})\,x + (a_{12} - ca_{22})\,y + (a_{13} - ca_{23})\,z = b_1 - cb_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3; \end{gathered}$$

$$\begin{gathered} (a_{11} - ca_{21})\,x + (a_{12} - ca_{22})\,y + (a_{13} - ca_{23})\,z = \\ = b_1 - cb_2, \\ a_{21}x + a_{22}y + a_{23}z = b_2, \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{gathered}$$

para $c \neq 0$ será equivalente al sistema inicial (2) y ambos sistemas son equivalentes entre sí.

Observemos que, si después de efectuar estas transformaciones aparece en el sistema una ecuación, la cual tiene todos los coeficientes del miembro primero iguales a cero, y el miembro segundo (el término independiente) también es igual a cero, entonces es evidente que esta ecuación se convierte en identidad para cualesquiera valores de las incógnitas, y después de eliminar esta ecuación se obtiene un sistema de ecuaciones que es equivalente al sistema inicial. Si el término independiente de tal ecuación no es igual a cero, entonces esta ecuación no puede convertirse en identidad para ninguna lista de valores de las incógnitas, y por eso el sistema obtenido de ecuaciones, al igual que el sistema inicial equivalente, es incompatible.

Método de eliminación sucesiva de incógnitas. Sea dado un sistema s de ecuaciones lineales con n incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n &= b_s. \end{aligned} \qquad (3)$$

Supongamos para mayor claridad que el coeficiente a_{11} no es igual a cero (aunque puede resultar ser igual a cero, y entonces es necesario comenzar el proceso descrito más abajo con otra ecuación del sistema, es decir, con la ecuación que tiene un coeficiente para la incógnita x_1 distinto de cero). Transformemos el sistema (3), eliminando la incógnita x_1 de todas las ecuaciones, menos de la primera. Para esto, multiplicamos ambos miembros de la primera ecuación por el número $\frac{a_{21}}{a_{11}}$ y los sustraemos de los respectivos miembros de la segunda ecuación, luego ambos miembros de la primera ecuación, multiplicados por el número $\frac{a_{31}}{a_{11}}$, los sustraemos de los respectivos miembros de la tercera ecuación, etc.

De resultas de las transformaciones indicadas llegamos al sistema de s ecuaciones lineales con n incógnitas

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3+\ldots+a_{1n}x_n&=b_1,\\ a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\ldots+a_{2n}^{(1)}x_n&=b_2^{(1)},\\ a_{32}^{(1)}x_2+a_{33}^{(1)}x_3+\ldots+a_{3n}^{(1)}x_n&=b_3^{(1)},\\ \cdots\cdots\cdots\cdots\cdots\cdots&\cdots\\ a_{s2}^{(1)}x_2+a_{s3}^{(1)}x_3+\ldots+a_{sn}^{(1)}x_n&=b_s^{(1)}, \end{aligned} \tag{4}$$

donde $a_{ij}^{(1)}$ y $b_i^{(1)}$ son los nuevos coeficientes para las incógnitas y los nuevos términos independientes que se expresan a través de los coeficientes y de los términos independientes del sistema inicial (3). El sistema (4) es equivalente al sistema (3). Transformemos ahora el sistema (4). Para esto no vamos a tocar la primera ecuación y transformaremos sólo una parte del sistema (4) que se compone de todas las ecuaciones menos la primera. Además, vamos a considerar que entre estas ecuaciones no hay las que tienen todos los coeficientes de los primeros miembros iguales a cero; tales ecuaciones las eliminaríamos, si sus términos independientes fuesen iguales a cero (en caso contrario ya demostraríamos la incompatibilidad de nuestro sistema). Pues, entre los coeficientes $a_{ij}^{(1)}$ existen coeficientes distintos de cero; para mayor claridad admitamos que $a_{22}^{(1)} \neq 0$. Transformemos ahora el sistema (4), sustrayendo de ambos miembros de la tercera y de cada una de las ecuaciones siguientes ambos miembros de la segunda ecuación, multiplicados respectivamente por los números

$$\frac{a_{32}^{(1)}}{a_{22}^{(1)}}, \quad \frac{a_{42}^{(1)}}{a_{22}^{(1)}}, \quad \ldots, \quad \frac{a_{s2}^{(1)}}{a_{22}^{(1)}}.$$

De resultas de todas las ecuaciones, menos la primera y segunda ecuación, se elimina la incógnita x_2 y se obtiene el siguiente sistema de ecuaciones, que es equivalente al sistema (4), y, por consiguiente, al sistema (3):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= b_3^{(2)}, \\ \cdots\cdots\cdots\cdots\cdots & \\ a_{t3}^{(2)}x_3 + \ldots + a_{tn}^{(2)}x_n &= b_t^{(2)}. \end{aligned}$$

El sistema contiene ahora t ecuaciones, además $t \leqslant s$, puesto que ciertas ecuaciones resultaron ser, posiblemente, eliminadas. Más adelante transformaremos sólo una parte del sistema obtenido, la cual contiene todas las ecuaciones, menos las dos primeras.

Como consecuencia de este proceso de eliminación sucesiva de las incógnitas puede resultar que

1) si se obtiene tal sistema, una de cuyas ecuaciones tiene un término independiente distinto de cero y todos los coeficientes del primer miembro son iguales a cero, entonces el sistema inicial es incompatible;

2) o se obtiene el siguiente sistema compatible de ecuaciones que es equivalente al sistema (3):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= b_3^{(2)}, \\ \cdots\cdots\cdots\cdots\cdots & \\ a_{kk}^{(k-1)}x_k + \ldots + a_{kn}^{(k-1)}x_n &= b_n^{(k-1)} \end{aligned} \tag{5}$$

Aquí $a_{11} \neq 0$, $a_{22}^{(1)} \neq 0$, $a_{33}^{(2)} \neq 0$, ..., $a_{kk}^{(k-1)} \neq 0$, y el índice superior que se encuentra entre paréntesis, indica el número de transformación, con el cual se obtienen los coeficientes para las incógnitas y los términos independientes.

Podemos anotar también que $k \leqslant s$ y, evidentemente, $k \leqslant n$.

Este sistema es determinado para $k=n$ (el número de ecuaciones es igual al número de incógnitas) y es indeterminado para $k<n$ (el número de ecuaciones es menor que el número de incógnitas). En realidad, para $k=n$ el sistema (5) tiene la forma

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3+\ldots+a_{1n}x_n&=b_1,\\ a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\ldots+a_{2n}^{(1)}x_n&=b_2^{(1)},\\ a_{33}^{(2)}x_3+\ldots+a_{3n}^{(2)}x_n&=b_3^{(2)},\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{nn}^{(n-1)}x_n&=b_n^{(n-1)} \end{aligned} \qquad (6)$$

De la última ecuación se obtiene un valor bastante determinado para la incógnita x_n. Sustituyéndolo en la penúltima ecuación, hallamos el único valor para la incógnita x_{n-1}. Continuando más adelante hallamos que el sistema (6), y por lo tanto también el sistema (3), tienen una sola solución, es decir, son compatibles y determinados (en este caso también se dice que el sistema (3) se reduce a la *forma triangular*).

Si $k<n$, entonces para las incógnitas «independientes» $x_{k+1}, \ldots, x_n$ tomamos los valores numéricos arbitrarios, después de lo cual, moviéndonos por el sistema (5) de abajo hacia arriba, hallamos para las incógnitas $x_k, x_{k-1}, \ldots$ $\ldots, x_2, x_1$ los valores bastante determinados. Puesto que los valores de las incógnitas independientes se pueden elegir mediante un número infinito de procedimientos distintos, entonces el sistema (5), y por consiguiente, el sistema (3) son compatibles, pero no indeterminados. (En este caso se dice que el sistema se reduce a la *forma trapezoidal*).

Así pues, el método de Gauss es aplicable a cualquier sistema de ecuaciones lineales. Al mismo tiempo el sistema es incompatible, si en el proceso de transformaciones se obtiene una ecuación, en la cual los coeficientes para todas las incógnitas son iguales a cero, y el término independiente es distinto de cero; si no se encuentra tal ecuación, entonces el sistema es compatible. El sistema compatible de ecuaciones es determinado, si se reduce a la forma triangular

(6), y es indeterminado, si se reduce a la forma trapezoidal (5) para $k < n$.

Ejemplo. Examínese, para qué valores de los parámetros α y β el sistema de tres ecuaciones lineales con tres incógnitas

$$\begin{aligned} x + y + z &= 0, \\ x - y + 2z &= 1, \\ 2x + 4y + \alpha z &= \beta \end{aligned} \tag{7}$$

a) tiene una sola solución (y hallar esta solución), b) es incompatible y c) es indeterminado.

Examinemos el sistema (7) mediante el método de Gauss. Multipliquemos ambos miembros de la segunda ecuación del sistema por (—1) y ambos miembros de la tercera ecuación del sistema por (—1) y ambos miembros de la tercera ecuación por $(-{}^1/_2)$. Escribimos el sistema que es equivalente al sistema (7):

$$\begin{aligned} x + y + z &= 0, \\ -x + y - 2z &= -1 \\ -x - 2y - \frac{\alpha}{2} z &= -\frac{\beta}{2}. \end{aligned} \tag{8}$$

Adicionamos la primera ecuación del sistema (8) con la segunda, y la tercera con la primera; las ecuaciones, obtenidas como resultado de la adición, las tomemos en calidad de segunda y de tercera ecuación del nuevo sistema que es equivalente al sistema (8):

$$\begin{aligned} x + y + {}& z = 0, \\ 2y - {}& z = -1, \\ -y + {}& \left(1 - \frac{\alpha}{2}\right) z = -\frac{\beta}{2}. \end{aligned} \tag{9}$$

Ahora, multiplicamos ambos miembros de la tercera ecuación por 2 y con ello igualamos los coeficientes para la incógnita y en la segunda y tercera ecuaciones del sistema. Adicionamos la segunda ecuación con la tercera ecuación transformada y tomamos la ecuación obtenida como tercera ecuación del sistema. De resultas el sistema (9) obtiene la forma

$$\begin{aligned} x + y + z &= 0, \\ 2y - z &= -1, \\ (1 - \alpha) z &= -(1 + \beta). \end{aligned} \tag{10}$$

En dependencia de los valores de los parámetros α, β son posibles los casos siguientes:

1) Para $\alpha \neq 1$ y cualquier β el sistema tiene una sola solución

$$x = \frac{\alpha - 3\beta - 4}{2(\alpha - 1)}, \quad y = \frac{-\alpha + \beta + 2}{2(\alpha - 1)}, \quad z = \frac{\beta + 1}{\alpha - 1}.$$

2) Para $\alpha = 1$ y $\beta \neq -1$ el primer miembro de la tercera ecuación del sistema (10) se anula y el segundo miembro es distinto de cero. Por consiguiente, para estos valores de dos parámetros α y β el sistema resulta ser incompatible.

3) Para $\alpha = 1$ y $\beta = -1$ el sistema (10) adquiere la forma «trapezoidal». Para estos valores de los parámetros α y β el sistema es compatible e indeterminado. Tiene la forma

$$\begin{aligned} x + y + z &= 0, \\ 2y - z &= -1. \end{aligned}$$

Suponiendo $z = c$, donde c es un número cualquiera, obtenemos un conjunto de soluciones

$$x = \frac{1 - 3c}{2}, \quad y = \frac{c - 1}{2}, \quad z = c.$$

El método de Gauss de solución de sistemas de ecuaciones lineales con coeficientes numéricos, en vigor de su simplicidad y de un mismo tipo de operaciones a cumplir, es aplicable para el cálculo en las computadoras electrónicas. Un defecto esencial de este método es la imposibilidad de formular las condiciones de compatibilidad y de determinación del sistema en dependencia de los valores de los coeficientes y de los términos independientes. Por otra parte, incluso en el caso de un sistema determinado este método no permite hallar las fórmulas generales, que expresan la solución del sistema a través de sus coeficientes y de los términos independientes, los cuales es necesario tener para las investigaciones teóricas. En el curso de matemáticas superiores se dan otros métodos de solución de sistemas de ecuaciones lineales, los cuales no tienen los defectos indicados. Estos métodos se basan en la teoría de matrices y determinantes.

7.4. Sistemas de dos ecuaciones lineales con dos incógnitas. Formulemos las condiciones de compatibilidad y determinación del sistema de dos ecuaciones lineales con dos incógnitas

$$\begin{aligned} a_{11}x + a_{12}y &= b_1, \\ a_{21}x + a_{22}y &= b_2, \end{aligned} \tag{11}$$

además, en cada una de las ecuaciones del sistema por lo menos uno de los coeficientes es distinto de cero.

Designemos a través de Δ, Δ_x, Δ_y los determinantes

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\Delta_x = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1a_{22} - b_2a_{12},$$

$$\Delta_y = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = a_{11}b_2 - b_1a_{21}.$$

Si los coeficientes del sistema (11) son tales que $\Delta \neq 0$, entonces el sistema es compatible y determinado. Las soluciones del sistema son

$$x = \frac{\Delta_x}{\Delta}, \quad y = \frac{\Delta_y}{\Delta}.$$

Sea que ahora los coeficientes del sistema (11) son tales que $\Delta = 0$, entonces:

1) si $\Delta_x \neq 0$ ó $\Delta_y \neq 0$ ó bien $\Delta_x \neq 0$, y $\Delta_y \neq 0$, entonces el sistema (11) es incompatible;

2) si $\Delta_x = \Delta_y = 0$, entonces el sistema (11) es compatible e indeterminado.

7.5. Interpretación geométrica de las soluciones de los sistemas de ecuaciones lineales. Sea dado un sistema de dos ecuaciones algebraicas lineales con dos incógnitas

$$\begin{aligned} a_{11}x + a_{12}y &= b_1, \\ a_{21}x + a_{22}y &= b_2, \end{aligned} \tag{12}$$

además, en cada una de las ecuaciones del sistema por lo menos uno de los coeficientes de las incógnitas es distinto de cero.

Cada una de las ecuaciones del sistema (12) representa una correspondencia lineal entre las variables x e y. Cual-

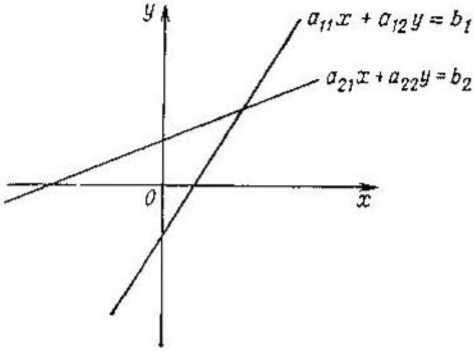


Fig. 4.2.

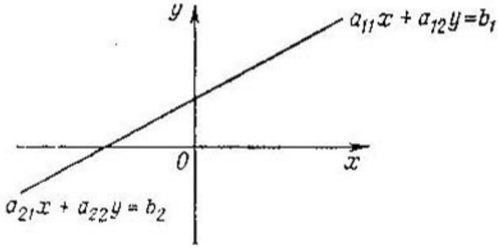


Fig. 4.3.

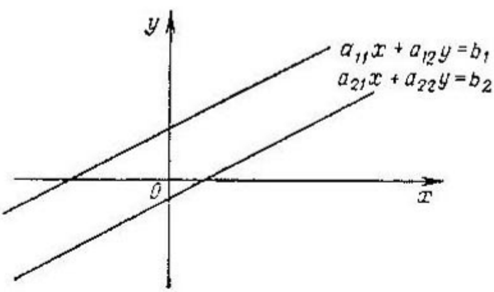


Fig. 4.4.

quier correspondencia lineal entre las variables x e y determina en el sistema de coordenadas rectangulares cierta recta. En el caso, cuando el sistema tiene una sola solución, las rectas definidas por la primera y segunda ecuaciones se intersecan (fig. 4.2). Si el sistema tiene un conjunto infinito de soluciones las rectas coinciden (fig. 4.3); si el sistema es incompatible, las rectas son paralelas (fig. 4.4).

Sea dado un sistema de tres ecuaciones lineales con tres incógnitas

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1, \\ a_{21}x + a_{22}y + a_{23}z &= b_2, \\ a_{31}x + a_{32}y + a_{33}z &= b_3, \end{aligned} \tag{13}$$

al mismo tiempo, en cada ecuación del sistema por lo menos uno de los coeficientes de las incógnitas es distinto de cero. La ecuación lineal con tres variables determina el plano en el espacio y, respectivamente, las tres ecuaciones del sistema (13), es decir, tres planos. Son posibles los casos siguientes de la posición recíproca de estos tres planos α_1, α_2, α_3, dados por la primera, segunda y tercera ecuaciones, respectivamente:

A. *Todos los tres planos son distintos.*

1) Los tres planos se intersecan en un punto (fig. 4.5). El sistema tiene una sola solución.

2) Dos planos α_1, α_2 se intersecan por la recta l, paralela al tercer plano α_3 (fig. 4.6). El sistema no tiene soluciones.

3) Dos planos se intersecan por la recta que pertenece al tercer plano (fig. 4.7). El sistema tiene un conjunto infinito de soluciones.

4) Los tres planos son paralelos (fig. 4.8). El sistema no tiene soluciones.

B. *Dos planos coinciden y el tercer plano es distinto de ellos.*

1) Los planos α_1 y α_2 que coinciden se intersecan con el tercer plano α_3 (fig. 4.9). El sistema tiene un conjunto infinito de soluciones.

2) Los planos que coinciden α_1 y α_2 son paralelos al tercer plano α_3 (fig. 4.10). El sistema no tiene soluciones.

C. *Los tres planos coinciden* (fig. 4.11). El sistema tiene un conjunto infinito de soluciones.

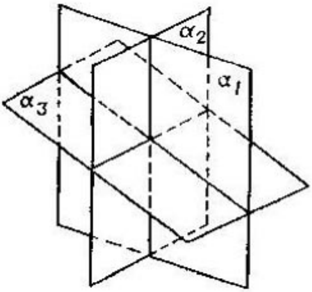


Fig. 4.5.

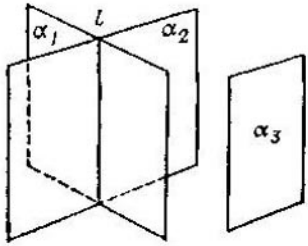


Fig. 4.6.

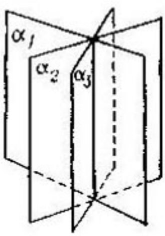


Fig. 4.7.

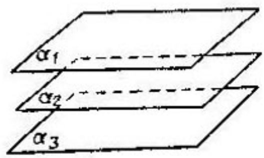


Fig. 4.8.

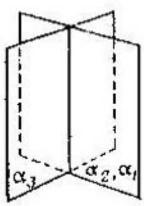


Fig. 4.9.

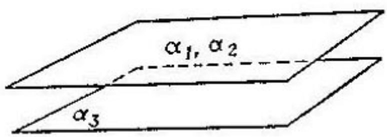


Fig. 4.10.

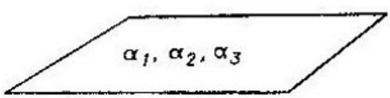


Fig. 4.11.

§ 8. Sistemas de ecuaciones algebraicas no lineales

Para los sistemas de ecuaciones algebraicas no lineales, a diferencia de los sistemas de ecuaciones lineales, no existe ningún método de solución universal que sea prácticamente cómodo y, por eso, para resolver cada sistema concreto de las ecuaciones no lineales es necesario utilizar procedimientos especiales de solución, basados en la utilización de las singularidades de las ecuaciones algebraicas, que constituyen el sistema dado.

Examinemos en ejemplos concretos dos procedimientos de solución de los sistemas de ecuaciones algebraicas no lineales.

Ejemplo 1. Resolver el sistema de ecuaciones

$$\begin{gathered} x^2 - 5xy + 4y^2 = 0, \\ x^2 - 15y^2 - x + 11y = -4. \end{gathered} \tag{1}$$

Observando, que el polinomio de dos variables x e y que se encuentra en el primer miembro de la primera ecuación del sistema (1) es homogéneo, el sistema (1) puede reducirse a dos sistemas siguientes que son equivalentes al sistema (1):

$$\begin{aligned} &x^2 - 5xy + 4y^2 = 0, \\ &y = 0, \\ &x^2 - 15y^2 - x + 11y = -4; \end{aligned} \qquad (2) \qquad \begin{aligned} &\left(\frac{x}{y}\right)^2 - 5\left(\frac{x}{y}\right) + 4 = 0, \\ &x^2 - 15y^2 - x + \\ &\quad + 11y = -4. \end{aligned} \qquad (2')$$

Hallamos las soluciones del sistema (2). Sustituyendo el valor $y = 0$ en la primera y tercera ecuaciones del sistema (2), obtenemos un sistema de dos ecuaciones para hallar el valor de una incógnita x:

$$x = 0, \quad x^2 - x = -4.$$

No existe un número x que satisfaga ambas ecuaciones del sistema, es decir, el sistema (2) es incompatible. Pasamos a la solución del sistema (2′). Resolvemos la primera ecuación del sistema (2′) como una ecuación cuadrática respecto a la variable $u = \frac{x}{y}$. La ecuación

$$u^2 - 5u + 4 = 0$$

tiene las raíces $u_1 = 4$, $u_2 = 1$, es decir, $\frac{x}{y} = 4$, $\frac{x}{y} = 1$.

De esta manera, el sistema (2') es equivalente a dos sistemas:

$$\begin{gathered} x = 4y, \\ x^2 - 15y^2 - x + 11y = -4. \end{gathered} \tag{3}$$

$$\begin{gathered} x = y, \\ x^2 - 15y^2 - x + 11y = -4. \end{gathered} \tag{3'}$$

Sustituyendo las expresiones x a través de y en las segundas ecuaciones de los sistemas obtenemos las respectivas ecuaciones para la obtención de y:

$$y^2 + 7y + 4 = 0, \quad -14y^2 + 10y + 4 = 0.$$

Las soluciones de la primera ecuación son:

$$y_1 = \frac{-7+\sqrt{33}}{2}, \quad y_2 = \frac{-7-\sqrt{33}}{2}.$$

Las soluciones de la segunda ecuación son:

$$y_3 = 1, \quad y_4 = -2/7.$$

Los sistemas (3) y (3') tienen, respectivamente, las soluciones siguientes:

$$x_1 = 2\sqrt{33} - 14, \quad x_2 = -2\sqrt{33} - 14,$$

$$x_3 = 1, \quad x_4 = -2/7,$$

$$y_1 = \frac{\sqrt{33}-7}{2}, \quad y_2 = \frac{-\sqrt{33}-7}{2}, \quad y_3 = 1, \quad y_4 = -\frac{2}{7},$$

las cuales son las soluciones del sistema (1).

Sistemas simétricos de ecuaciones. Un sistema de ecuaciones con dos incógnitas x e y se llama *simétrico*, si el mismo no se cambia con la sustitución de la incógnita x por la incógnita y, y de la incógnita y por la incógnita x.

Muy a menudo la solución de tales sistemas puede ser hallada mediante la introducción de nuevas variables, es decir, de polinomios *simétricos* elementales σ_1 y σ_2:

$$\sigma_1 = x + y, \quad \sigma_2 = xy.$$

Ejemplo 2. Hallar las soluciones del sistema

$$\begin{gathered} x^2 + xy + y^2 = 4, \\ x + xy + y = 2. \end{gathered} \tag{4}$$

tiene las raíces $u_1 = 4$, $u_2 = 1$, es decir, $\frac{x}{y} = 4$, $\frac{x}{y} = 1$.

De esta manera, el sistema (2′) es equivalente a dos sistemas:

$$\begin{aligned} x &= 4y, \\ x^2 - 15y^2 - x + 11y &= -4. \end{aligned} \tag{3}$$

$$\begin{aligned} x &= y, \\ x^2 - 15y^2 - x + 11y &= -4. \end{aligned} \tag{3′}$$

Sustituyendo las expresiones x a través de y en las segundas ecuaciones de los sistemas obtenemos las respectivas ecuaciones para la obtención de y:

$$y^2 + 7y + 4 = 0, \quad -14y^2 + 10y + 4 = 0.$$

Las soluciones de la primera ecuación son:

$$y_1 = \frac{-7 + \sqrt{33}}{2}, \quad y_2 = \frac{-7 - \sqrt{33}}{2}.$$

Las soluciones de la segunda ecuación son:

$$y_3 = 1, \quad y_4 = -2/7.$$

Los sistemas (3) y (3′) tienen, respectivamente, las soluciones siguientes:

$$x_1 = 2\sqrt{33} - 14, \quad x_2 = -2\sqrt{33} - 14,$$

$$x_3 = 1, \quad x_4 = -2/7,$$

$$y_1 = \frac{\sqrt{33} - 7}{2}, \quad y_2 = \frac{-\sqrt{33} - 7}{2}, \quad y_3 = 1, \quad y_4 = -\frac{2}{7},$$

las cuales son las soluciones del sistema (1).

Sistemas simétricos de ecuaciones. Un sistema de ecuaciones con dos incógnitas x e y se llama *simétrico*, si el mismo no se cambia con la sustitución de la incógnita x por la incógnita y, y de la incógnita y por la incógnita x.

Muy a menudo la solución de tales sistemas puede ser hallada mediante la introducción de nuevas variables, es decir, de polinomios *simétricos* elementales σ_1 y σ_2:

$$\sigma_1 = x + y, \quad \sigma_2 = xy.$$

Ejemplo 2. Hallar las soluciones del sistema

$$\begin{aligned} x^2 + xy + y^2 &= 4, \\ x + xy + y &= 2. \end{aligned} \tag{4}$$

Introduzcamos nuevas variables, los polinomios simétricos elementales

$$\sigma_1 = x + y, \quad \sigma_2 = xy.$$

Respecto a las variables σ_1 y σ_2 el sistema (4) se escribe en la forma

$$\begin{aligned} \sigma_1^2 - \sigma_2 = 4, \\ \sigma_1 + \sigma_2 = 2. \end{aligned} \tag{5}$$

De la segunda ecuación del sistema obtenemos

$$\sigma_2 = 2 - \sigma_1. \tag{6}$$

Sustituyendo en la primera ecuación del sistema (5) σ_2 por la expresión $2 - \sigma_1$, obtenemos la ecuación cuadrática

$$\sigma_1^2 + \sigma_1 - 6 = 0 \text{ con las raíces } \sigma_1 = 2, \ \sigma_1 = -3.$$

Sustituyendo los valores hallados σ_1 en la relación (6), obtenemos los valores σ_2. De esta manera, el conjunto de soluciones del sistema (5) tiene la forma

$$\sigma_1 = 2, \quad \sigma_2 = 0; \qquad \sigma_1 = -3, \quad \sigma_2 = 5.$$

Ahora el conjunto de soluciones del sistema inicial (4) puede ser obtenido como la unión de los conjuntos de soluciones de dos sistemas más simples:

$$\begin{aligned} x + y = 2, \qquad & x + y = -3, \\ xy = 0; \qquad & xy = 5. \end{aligned}$$

Ejemplo 3. Resolver el sistema de ecuaciones

$$\begin{aligned} x + y = 5, \\ x^5 + y^5 = 275. \end{aligned}$$

La sustitución de $\sigma_1 = x + y$, $\sigma_2 = xy$ reduce el sistema dado de ecuaciones a la forma

$$\begin{aligned} \sigma_1 = 5, \\ \sigma_1^5 - 5\sigma_1^3\sigma_2 + 5\sigma_1\sigma_2^2 = 275. \end{aligned}$$

Sustituyendo el valor $\sigma_1 = 5$ en la segunda ecuación, obtenemos una ecuación para σ_2:

$$\sigma_2^2 - 25\sigma_2 + 114 = 0.$$

Las raíces de esta ecuación serán

$$\sigma_2 = 6, \quad \sigma_2 = 19.$$

De esta manera, el conjunto de soluciones del sistema inicial se halla como una unión de conjuntos de soluciones de dos sistemas más simples:

$$\begin{aligned} x + y = 5, \qquad & x + y = 5, \\ xy = 6; \qquad & xy = 19. \end{aligned}$$

En el p. 7.3 fueron examinados los sistemas de ecuaciones algebraicas lineales, en las cuales el número de ecuaciones era menor que el número de incógnitas. En este caso resultó ser que unas incógnitas se expresaban a través de otras («independientes») incógnitas, es decir, el sistema tenía un conjunto infinito de soluciones. De otra manera sucede en el caso de solución de sistemas de ecuaciones algebraicas no lineales, en las cuales el número de incógnitas es mayor que el número de ecuaciones. Puede resultar que tal sistema tiene tanto un número infinito, como finito de soluciones o no tiene, incluso, las mismas. Ahora examinemos como ejemplo un sistema de dos ecuaciones algebraicas con tres incógnitas, el cual tiene una sola solución (es decir, el conjunto de soluciones representa la única terna ordenada de números que convierte a ambas ecuaciones en identidad).

Ejemplo 4. Hallar las soluciones reales del sistema

$$x + y = 2$$
$$xy - z^2 = 1.$$

Después de sustituir en la segunda ecuación $y = 2 - x$ la solución del sistema se reduce a la solución de la ecuación

$$(x - 1)^2 + z^2 = 0,$$

la cual es equivalente al sistema

$$x - 1 = 0, \quad z = 0.$$

Este sistema tiene una sola solución, por eso también el sistema (7) de dos ecuaciones con tres incógnitas tiene sólo una solución (1; 1; 0).

§ 9. Desigualdades

9.1. Definiciones y propiedades principales de las desigualdades. Las expresiones que tienen la forma

$$a < b \ (a \leqslant b), \quad a > b \ (a \geqslant b),$$

donde a y b pueden ser números o funciones, se llaman *desigualdades*. Los símbolos $<$ ($\leqslant$), $>$ ($\geqslant$) se llaman *signos de desigualdad* y se leen respectivamente: menor (menor o igual), mayor (mayor o igual).

Las desigualdades que se escriben mediante los signos $>$ y $<$, se llaman *estrictas*, y las desigualdades cuyas expresiones contienen los signos $\geqslant$ y $\leqslant$ se llaman *no estrictas*.

La desigualdad no estricta es equivalente a la desigualdad estricta del mismo signo y a la igualdad.

Se distingue dos tipos de desigualdades: las *aritméticas* (o numéricas), cuya expresión contiene sólo números, y *no aritméticas*, cuya expresión contiene, junto con los números, funciones de una o de varias variables.

Por ejemplo, las desigualdades numéricas son

$$2>1, \quad \sqrt{2}\leqslant 7.$$

Las desigualdades no aritméticas, por ejemplo, son las desigualdades

$$a<1, \ \operatorname{tg}^2 x+\operatorname{tg} x>0, \ x^2+y^2\geqslant R^2, \ \ldots$$

Las funciones que entran en las desigualdades pueden admitir distintos valores numéricos en dependencia de los distintos valores de sus argumentos. Para unos valores de los argumentos la desigualdad puede convertirse en desigualdad estricta, para otros valores, no.

*Las propiedades principales de las desigualdades**) *son:*

1) Si $a>b$, entonces $b<a$; si $a<b$, entonces $b>a$ (antisimetría)

2) Si $a>b$ y $b>c$, entonces $a>c$; si $a<b$ y $b<c$, entonces $a<c$

Operaciones aritméticas con las desigualdades.

1) Dos desigualdades de un mismo signo pueden sumarse; en este caso como resultado de la adición se obtiene una desigualdad del mismo signo:

si $a>b$ y $c>d$, entonces $a+c>b+d$;

si $a<b$ y $c<d$, entonces $a+c<b+d$.

2) Si ambos miembros de la desigualdad se multiplican (se dividen) por un mismo valor positivo, entonces se obtiene una desigualdad del mismo signo; si ambos miembros de la desigualdad se multiplican (se dividen) por un valor

*) Todo lo dicho más abajo para los casos de las desigualdades estrictas es válido también para las desigualdades no estrictas.

negativo, entonces se obtiene una desigualdad de signo opuesto:

si $a > b$ y $c > 0$, entonces $ac > bc$;

si $a < b$ y $c > 0$, entonces $ac < bc$;

si $a > b$ y $c < 0$, entonces $ac < bc$;

si $a < b$ y $c < 0$, entonces $ac > bc$.

3) Como consecuencia de las reglas de adición de las desigualdades y de multiplicación de ambos miembros de las desigualdades por un mismo valor, se obtiene la regla de sustracción de las desigualdades de distinto signo:

De una desigualdad se puede miembro a miembro (del primer miembro sustraer el primero, y del segundo miembro sustraer el segundo) sustraer otra desigualdad de signo opuesto. De resultas se obtiene una desigualdad que tiene el signo de la primera desigualdad:

si $a < b$ y $c > d$, entonces $a - c < b - d$;

si $a > b$ y $c < d$, entonces $a - c > b - d$.

4) Si al primero y segundo miembros de la desigualdad agregamos el mismo valor, entonces de resultas se obtiene una desigualdad del mismo signo:

si $a > b$, entonces $a + c > b + c$.

9.2. Algunas desigualdades importantes.

1) Para cualesquiera números reales a y b se cumple la desigualdad

$$|a + b| \leqslant |a| + |b|,$$

es decir, el valor absoluto de la suma de dos números reales no supera la suma de valores absolutos de estos números.

Mediante el método de inducción matemática se puede demostrar que para cualquier número finito de sumandos a_i es válida la desigualdad

$$\left|\sum_{i=1}^{n} a_i\right| \leqslant \sum_{i=1}^{n} |a_i|,$$

además, la igualdad se obtiene si todos los sumandos son números del mismo signo.

2) Para cualesquiera números reales a y b se cumple la desigualdad

$$|a-b| \geqslant ||a|-|b||,$$

es decir, el valor absoluto de la diferencia de dos números no es menor que el valor absoluto de la diferencia de los valores absolutos de estos números.

3) Para cualesquiera dos números reales a y b se cumple la desigualdad

$$a^2+b^2 \geqslant 2|ab|,$$

al mismo tiempo, la igualdad se obtiene cuando y sólo cuando $|a|=|b|$.

4) Si a y b son números reales del mismo signo ($ab>0$), entonces

$$\frac{a}{b}+\frac{b}{a} \geqslant 2,$$

además, la igualdad se obtiene cuando y sólo cuando $a=b$.

5) *Desigualdad de Cauchy*: si a y b son números reales no negativos, entonces

$$\frac{a+b}{2} \geqslant \sqrt{ab},$$

es decir, la media aritmética de dos números no negativos no es menor que su media geométrica.

Una desigualdad análoga es válida también para cualquier lista finita de números no negativos a_i ($i=1, 2, 3, \ldots$ $\ldots, n$):

$$\frac{a_1+a_2+a_3+\ldots+a_n}{n} \geqslant \sqrt[n]{a_1 a_2 a_3 \ldots a_n},$$

es decir, la media aritmética n de los números no negativos es mayor o igual a su media geométrica; la igualdad se obtiene, cuando todas las n de números son iguales.

6) *Desigualdad de Cauchy—Buniakovski*: para cualesquiera números reales $a_1, a_2, a_3, \ldots, a_n b_1, b_2, \ldots, b_n$ se cumple la desigualdad

$$(a_1b_1+a_2b_2+a_3b_3+\ldots+a_nb_n)^2 \leqslant$$
$$\leqslant (a_1^2+a_2^2+\ldots+a_n^2)(b_1^2+b_2^2+\ldots+b_n^2).$$

La igualdad se cumple cuando y sólo cuando los números a_i y b_i son proporcionales, es decir, cuando existen tales números α y β que $\alpha^2+\beta^2 \neq 0$, y para todos $i=1, 2, 3, \ldots, n$ se cumple la igualdad

$$\alpha a_i+\beta b_i=0.$$

7) Relación entre la media aritmética y cuadrática de varios números: el valor absoluto de la media aritmética de cualquier número finito de números reales no supera a la media cuadrática de estos números es decir,

$$\frac{|a_1+a_2+\ldots+a_n|}{n} \leqslant \sqrt{\frac{a_1^2+a_2^2+\ldots+a_n^2}{n}},$$

además, la igualdad se cumple cuando y sólo cuando todas las n de los números son iguales entre sí.

8) *Desigualdad de Hölder* (generalización de la desigualdad de Cauchy — Buniakovski): para cualesquiera números reales $a_1, a_2, \ldots, a_n, b_1, \ldots, b_n$ con cualquier $p > 1$ se cumple la desigualdad

$$|a_1b_1 + a_2b_2 + \ldots + a_nb_n| \leqslant$$
$$\leqslant (|a_1|^p + |a_2|^p + \ldots + |a_n|^p)^{1/p} (|b_1|^q + |b_2|^q + \ldots + |b_n|)^{1/q},$$

donde $q = \dfrac{p}{p-1}$.

§ 10. Solución de desigualdades y de sistemas de desigualdades

10.1. Definiciones principales. Sea f una función numérica de una o varias variables (argumentos). Resolver la desigualdad

$$f < 0 \tag{1}$$

significa hallar un conjunto de valores del argumento (argumentos) de la función f, para los cuales la función f es negativa, es decir, para los cuales la desigualdad (1) se convierte en cierta desigualdad numérica.

Análogamente, resolver la desigualdad

$$f > 0 \tag{2}$$

significa hallar un conjunto de valores del argumento (argumentos) de la función f, para los cuales los valores de esta función son positivos, es decir, para los cuales la desigualdad (2) se convierte en una desigualdad numérica válida.

Un conjunto de valores de los argumentos de la función f, para los cuales la desigualdad se convierte en una desigualdad numérica válida, se llama conjunto de soluciones de la desigualdad o simplemente solución de la desigualdad. Dos desigualdades se consideran *equivalentes*, si los conjuntos de sus soluciones coinciden.

Resolver un sistema de varias desigualdades significa hallar un conjunto de valores de los argumentos de las funciones que entran en las desigualdades, para los cuales todas las desigualdades del sistema *simultáneamente* se convierten en desigualdades numéricas válidas.

Dos sistemas de desigualdades se consideran *equivalentes*, si los conjuntos de sus soluciones coinciden. El sistema de desigualdades S se considera equivalente a dos sistemas de desigualdades S_1 y S_2, si el conjunto de soluciones del sistema S coincide con la unión de los conjuntos de soluciones de sistemas S_1 y S_2.

10.2. Desigualdades lineales y sistemas de desigualdades lineales con una incógnita. Por desigualdades lineales se entiende una desigualdad que tiene la forma

$$ax + b > 0,\ ax + b < 0,\ ax + b \geqslant 0,\ ax + b \leqslant 0,$$

donde a y b son números reales ($a \neq 0$).

La solución de las desigualdades lineales se efectúa mediante la sustitución de la desigualdad inicial por la desigualdad equivalente. Para esto se usan las siguientes transformaciones de desigualdades que conducen a las desigualdades equivalentes: la adición a ambos miembros de la desigualdad el mismo número y la multiplicación (división) de ambos miembros de la desigualdad por el mismo número. Así, un conjunto de soluciones de la desigualdad

$$ax + b > 0 \tag{3}$$

puede ser hallado de la manera siguiente.

Agregamos a ambos miembros de la desigualdad (3) el número $-b$, de resultas obtenemos una desigualdad equivalente

$$ax > -b, \tag{4}$$

y dividimos ambos miembros de la desigualdad por a.

Si $a > 0$, entonces la desigualdad (4) se convierte en la desigualdad

$$x > -\frac{b}{a},$$

la cual da un conjunto de soluciones de la desigualdad inicial (3). Este conjunto de soluciones puede escribirse también en la forma

$$x \in \left(-\frac{b}{a};\ +\infty\right).$$

Si $a < 0$, entonces la desigualdad (4) es equivalente a la desigualdad

$$x < -\frac{b}{a}.$$

Un conjunto de números reales que satisfacen a esta desigualdad, es el conjunto de soluciones de la desigualdad inicial (3):

$$x \in \left(-\infty;\ -\frac{b}{a}\right).$$

Análogamente pueden ser halladas las soluciones de cualesquiera desigualdades lineales.

Un conjunto de soluciones de un sistema de dos desigualdades lineales

$$ax + b > 0,$$
$$cx + d > 0$$

se halla como la intersección de conjuntos de soluciones de estas desigualdades.

10.3. Desigualdades cuadráticas. Por desigualdad *cuadrática* se entiende una desigualdad, que puede ser reducida a una de las desigualdades siguientes:

$$ax^2 + bx + c > 0,\quad ax^2 + bx + c < 0,$$
$$ax^2 + bx + c \geqslant 0,\quad ax^2 + bx + c \leqslant 0,$$

donde a, b, c son ciertos números reales y $a \neq 0$.

Serán desigualdades cuadráticas simples las desigualdades

$$x^2 < m \text{ y } x^2 > m.$$

El conjunto de soluciones de la desigualdad $x^2 < m$ es:

1) para $m \leqslant 0$ $x = \phi$ (es decir, no hay soluciones);

2) para $m > 0$ $x \in (-\sqrt{m};\ \sqrt{m})$, es decir, $-\sqrt{m} < x < \sqrt{m}$, o $|x| < \sqrt{m}$.

El conjunto de soluciones de la desigualdad $x^2 > m$ es:

1) para $m < 0$ $x \in \mathbf{R}$ (es decir, x es cualquier número real);

2) para $m \geqslant 0$ $x \in (-\infty; -\sqrt{m}) \cup (\sqrt{m}; +\infty)$, es decir, $-\infty < x < -\sqrt{m}$ y $\sqrt{m} < x < +\infty$, o $|x| > \sqrt{m}$.

La desigualdad cuadrática $ax^2 + bx + c > 0$ tiene, en dependencia de los valores de sus coeficientes a, b y c, de un conjunto de soluciones:

1) para $a > 0, D = b^2 - 4ac \geqslant 0$ $x \in \left(-\infty; \frac{-b-\sqrt{D}}{2a}\right) \cup \left(\frac{-b+\sqrt{D}}{2a}; +\infty\right)$;

2) para $a > 0$, $D < 0$ $x \in R$;

3) para $a < 0, D \geqslant 0$ $x \in \left(\frac{-b+\sqrt{D}}{2a}; \frac{-b-\sqrt{D}}{2a}\right)$;

4) para $a < 0, D < 0$ $x = \phi$ (es decir, no hay soluciones).

La solución de la desigualdad $ax^2 + bx + c < 0$ se reduce a la solución de la desigualdad examinada más arriba, si ambos miembros de la desigualdad se multiplican por -1.

Un conjunto de soluciones de las desigualdades no estrictas $ax^2 + bx + c \geqslant 0$ y $ax^2 + bx + c \leqslant 0$ se halla como una unión de los conjuntos de soluciones de las desigualdades estrictas correspondientes y de la ecuación $ax^2 + bx + c = 0$.

Las desigualdades que se reducen a la forma

$$\frac{ax+b}{cx+d} > k, \tag{5}$$

donde a, b, c, d, k son ciertos números reales y $c \neq 0$ (si $c = 0$, entonces la desigualdad lineal fraccional se convierte en lineal), se llaman desigualdades *lineales fraccionales*. También pertenecen a estas desigualdades las del tipo (5), las cuales en vez del signo $>$ tienen los signos $<$, $\geqslant$, $\leqslant$.

La solución de la desigualdad lineal fraccional se reduce a la solución de la desigualdad cuadrática. Para esto, es necesario multiplicar ambos miembros de la desigualdad (5) por la expresión $(cx + d)^2$, la cual es positiva para todos $x \in \mathbf{R}$ y $x \neq -d/c$.

10.4. Método de intervalos. Sea $P(x)$ un polinomio de n-ésimo grado con coeficientes reales, $c_1, c_2, \ldots, c_l$ son todas las raíces reales del polinomio con las multiplicidades $k_1, k_2, \ldots, k_l$, respectivamente, además, $c_1 > c_2 > \ldots$

$\ldots > c_l$. El polinomio $P(x)$ puede representarse en la forma

$$P(x) = (x - c_1)^{k_1} (x - c_2)^{k_2} \ldots (x - c_l)^{k_l} Q(x), \qquad (6)$$

donde el polinomio $Q(x)$ no tiene raíces reales y es positivo o negativo para todos $x \in \mathbf{R}$. Supongamos, para mayor precisión, que $Q(x) > 0$. Entonces para $x > c_1$ todos los factores en el desarrollo (6) son positivos y $P(x) > 0$. Si c_1 es la raíz de la multiplicidad impar (k_1 es impar), entonces para $c_2 < x < c_1$ todos los factores en el desarrollo (6), a excepción del primero, son positivos y $P(x) < 0$. En este caso se dice que el polinomio $P(x)$ cambia su signo con el paso a través de la raíz c_1. Si c_1 es la raíz de la multiplicidad par (k_1 es par), entonces todos los factores (incluyendo el primero) para $c_2 < x < c_1$ son positivos y, por consiguiente, $P(x) > 0$. En este caso se dice que el polinomio $P(x)$ no cambia su signo cuando pasa a través de la raíz c_1.

Análogamente, utilizando el desarrollo (6), es fácil convencerse que al pasar a través de la raíz c_2 el polinomio $P(x)$ cambia el signo, si k_2 es impar y no lo cambia, si k_2 es par.

La propiedad examinada de los polinomios se usa para resolver las desigualdades por el método de intervalos. Para hallar todas las soluciones de la desigualdad

$$P(x) > 0,$$

es suficiente saber las raíces reales del polinomio $P(x)$, sus multiplicidades y el signo del polinomio $P(x)$ en un punto arbitrario x_0 que no coincide con la raíz del polinomio.

Ejemplo 1. Resolver la desigualdad

$$x^2 (x + 2)(x - 1)^3 (x^2 + 1) > 0. \qquad (7)$$

Situemos en el eje numérico Ox las raíces del polinomio que se encuentra en el primer miembro de la desigualdad (fig. 4.12).

Para $x > 1$ el polinomio es positivo, puesto que todos los factores que se encuentran en el primer miembro de la igualdad son positivos. Vamos a movernos por el eje Ox de derecha a izquierda. Al pasar a través del punto $x = 1$ el polinomio cambia su signo y se convierte en negativo, puesto que $x = 1$ es la raíz de la multiplicidad 3; al pasar a través del punto $x = 0$ el polinomio no cambia de signo, puesto que $x = 0$ es la raíz de la multiplicidad 2; al pasar a través del punto $x = -2$ el polinomio de nuevo cambia de signo y llega a ser positivo.

Los intérvalos del polinomio dado de signos constantes se representan esquemáticamente en la fig. 4.12. Usando

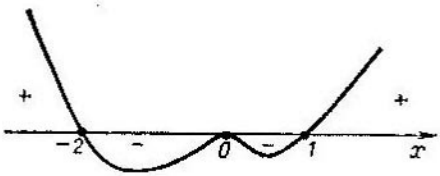


Fig. 4.12.

esta figura, es muy fácil escribir un conjunto de solución de la desigualdad (7):

$$x \in (-\infty; -2) \cup (1; +\infty).$$

La solución de la desigualdad racional que tiene la forma

$$\frac{P(x)}{Q(x)} > 0, \tag{8}$$

donde $P(x)$ y $Q(x)$ son polinomios, se reduce a la solución de la desigualdad algebraica equivalente de la manera siguiente: multiplicando ambos miembros de la desigualdad (8) por el polinomio $[Q(x)]^2$, el cual es positivo para todos los valores admisibles de la desigualdad (8), obtenemos la desigualdad algebraica

$$P(x)\,Q(x) > 0,$$

equivalente a la desigualdad (8).

Ejemplo 2. Resolver la desigualdad racional

$$\frac{x^2 (x-1)^3 (x+2)}{x-3} < 0. \tag{9}$$

Multiplicando ambos miembros de la desigualdad por $(x-3)^2$, obtenemos la desigualdad equivalente a la desigualdad (9):

$$x^2 (x-1)^3 (x+2)(x-3) < 0.$$

El conjunto de soluciones de la última desigualdad se halla mediante el método de intervalos y es el siguiente:

$$x \in (-\infty; -2) \cup (1; 3).$$

10.5. Solución de desigualdades irracionales. Por desigualdad irracional se entiende la desigualdad, en la cual los valores incógnitos (o las funciones racionales de valores incógnitos) se encuentran bajo el signo del radical. Para hallar el conjunto de soluciones de la desigualdad irracional, es necesario, por regla general, elevar ambos miembros de la desigualdad a una potencia natural. A pesar de la apariencia exterior del proceso de solución de la ecuación irracional (véase p. 5.9) y de la desigualdad irracional, entre las mismas existe una gran diferencia. Cuando resolvemos las ecuaciones algebraicas no nos preocupa que después de la elevación a la potencia se obtenga una ecuación equivalente a la inicial: la ecuación algebraica tiene un número finito de raíces y podemos elegir de un conjunto de estas raíces las soluciones de la ecuación irracional.

Un conjunto de soluciones de desigualdades representa, por regla general, un conjunto infinito de números, y por eso la verificación directa del conjunto de soluciones mediante la sustitución de estos números en la desigualdad inicial es, de principio, imposible. El único procedimiento que garantiza una respuesta correcta consiste en que tenemos que observar, que para cualquier transformación de la desigualdad inicial se obtenga una desigualdad equivalente a la inicial.

Para la solución de las desigualdades irracionales es necesario tener en cuenta que cuando elevamos a una potencia impar ambos miembros de la desigualdad siempre se obtiene una desigualdad equivalente a la desigualdad inicial. Si ambos miembros de la desigualdad los elevamos a la potencia par, obtendremos una desigualdad equivalente a la inicial que tiene el mismo signo de desigualdad sólo en el caso, si ambos miembros de la desigualdad inicial no son negativos.

Ejemplo. Resolver la desigualdad

$$\sqrt{x-5}-\sqrt{9-x}>1, \tag{10}$$

El conjunto de los valores admisibles de x es:

$$x \in [5;\ 9].$$

La desigualdad (10) es equivalente a la desigualdad

$$\sqrt{x-5}>1+\sqrt{9-x}, \tag{11}$$

ambos miembros de la cual son no negativos. Elevando ambos miembros de la desigualdad (11) al cuadrado, obtenemos la desigualdad equivalente

$$2x-15>2\sqrt{9-x}. \tag{12}$$

1) Si $2x-15 \leqslant 0$, es decir, $x \leqslant 15/2$, el primer miembro de la desigualdad es negativo o es igual a cero, y el segundo miembro es no negativo. Por eso para ningún $x \in [5;\ 15/2]$ la desigualdad (12) puede convertirse en una desigualdad numérica válida.

2) Si $2x-15>0$, entonces ambos miembros de la desigualdad son no negativos y después de elevarlos al cuadrado se obtiene una desigualdad equivalente a la desigualdad (12):

$$(2x-15)^2>4\,(9-x).$$

De esta manera, el conjunto de soluciones de la desigualdad (10) se obtiene como el conjunto de soluciones del sistema de desigualdades

$$5 \leqslant x \leqslant 9,$$
$$2x-15>0,$$
$$(2x-15)^2>4\,(9-x),$$

de donde obtenemos $x \in \left(\frac{14+\sqrt{7}}{2};\ 9\right]$.

10.6. Desigualdades exponenciales. Un ejemplo de desigualdad *exponencial* es la que tiene la forma

$$a^x>b, \tag{13}$$

donde a y b son ciertos números reales ($a>0$, $a \neq 1$).

En función de los valores de los parámetros a y b el conjunto de soluciones de la desigualdad (13) es el siguiente:

1) para $a > 1$, $b > 0$ $x \in (\log_a b;\ +\infty)$;
2) para $0 < a < 1$, $b > 0$ $x \in (-\infty;\ \log_a b)$;
3) para $a > 0$, $b < 0$, $x \in \mathbf{R}$.

El conjunto de soluciones de la desigualdad

$$a^x < b \tag{14}$$

en dependencia de los valores de los parámetros a y b es el siguiente:

1) para $a > 1$, $b > 0$ $x \in (-\infty;\ \log_a b)$;
2) para $0 < a < 1$, $b > 0$ $x \in (\log_a b;\ +\infty)$;
3) para $a > 0$, $b < 0$ $x = \phi$ (es decir, la desigualdad no tiene soluciones).

Las desigualdades (13) y (14) pueden ser generalizadas para el caso, cuando en el exponente de la potencia se encuentra cierto polinomio de x. Así pues, el conjunto de soluciones de la desigualdad

$$2^{f(x)} > 3 \tag{15}$$

se halla como un conjunto de soluciones de la desigualdad

$$f(x) > \log_2 3$$

que es equivalente a la desigualdad (15).

10.7. Desigualdades logarítmicas. La desigualdad que tiene la forma

$$\log_a f(x) > b, \tag{16}$$

donde a y b son ciertos números reales ($a > 0$ y $a \neq 1$), es una desigualdad logarítmica simple. La desigualdad (16) en dependencia de los valores del parámetro a es equivalente a los sistemas de desigualdades siguientes*):

1) para $a > 1$ $\quad$ 2) para $0 < a < 1$

$$\begin{cases} f(x) > 0, \\ f(x) > a^b, \end{cases} \qquad \begin{cases} f(x) > 0, \\ f(x) < a^b, \end{cases}$$

Respectivamente la desigualdad

$$\log_a f(x) < b \tag{17}$$

---

*) En caso, si la desigualdad (16) es no estricta, las segundas desigualdades de estos sistemas también son no estrictas.

es equivalente a los sistemas siguientes de desigualdades:

1) para $a > 1$ 2) para $0 < a < 1$

$$\begin{cases} f(x) > 0, \\ f(x) < a^b, \end{cases} \qquad \begin{cases} f(x) > 0, \\ f(x) > a^b. \end{cases}$$

El conjunto de soluciones de las desigualdades que tienen la forma

$$R\,(a^x) > 0 \text{ y } R\,(\log_a x) > 0, \tag{18}$$

así como las desigualdades $R < 0$, $R \geqslant 0$, $R \leqslant 0$, donde R es el polinomio de argumentos indicados, se halla de la manera siguiente.

Se introduce una incógnita nueva $y = a^x$ (respectivamente $y = \log_a x$) y las desigualdades (18) se resuelven como algebraicas respecto a la incógnita y. Después, la solución de la desigualdad inicial se reduce a la solución de la desigualdad simple respectiva (13), (14), (16), (17) o de los sistemas de estas desigualdades.

Ejemplo 1. Resolver la desigualdad logarítmica

$$\log_4^2 x - \log_2 x - 15 > 0. \tag{19}$$

Reducimos todos los logaritmos a una base (por ejemplo, a la base 2) y obtenemos una desigualdad que es equivalente a la desigualdad logarítmica (19):

$$\frac{1}{4} \log_2^2 x - \log_2 x - 15 > 0. \tag{20}$$

Resolvemos esta desigualdad como cuadrática respecto a la incógnita nueva

$$y = \log_2 x.$$

De resultas obtenemos que la desigualdad logarítmica (20) es equivalente a dos desigualdades simples:

$$\log_2 x > 10 \text{ y } \log_2 x < -6,$$

cuya unión de los conjuntos de soluciones da un conjunto de soluciones de la desigualdad (19):

$$x \in (0;\ 2^{-6}) \cup (2^{10};\ +\infty).$$

En conclusión indiquemos un tipo más de desigualdades logarítmicas que se encuentra con frecuencia, en las cuales la base del logaritmo y el argumento del logaritmo son funciones de la incógnita x. Tal desigualdad logarítmica es, por ejemplo, la desigualdad

$$\log_{g(x)} f(x) > c, \tag{21}$$

donde $f(x)$ y $g(x)$ son ciertos polinomios, y c es cierto número real. El conjunto de valores admisibles de la incógnita x se halla como un conjunto de soluciones del sistema

$$\begin{aligned} f(x) &> 0, \\ g(x) &> 0, \\ g(x) &\neq 1. \end{aligned}$$

La desigualdad logarítmica (21) es equivalente a dos sistemas de las desigualdades algebraicas:

$$\begin{array}{ll} f(x) > 0, & f(x) > 0, \\ g(x) > 1, & 0 < g(x) < 1, \\ f(x) > [g(x)]^c; & f(x) < [g(x)]^c. \end{array}$$

Ejemplo 2. Resolver la desigualdad

$$\log_x \left(2x - \frac{3}{4}\right) > 2. \tag{22}$$

La desigualdad logarítmica dada es equivalente a dos sistemas de desigualdades:

$$\begin{array}{rr} 2x - \frac{3}{4} > 0, & 2x - \frac{3}{4} > 0 \\ x > 1, & 0 < x < 1, \\ 2x - \frac{3}{4} > x^2; & 2u - \frac{3}{4} < x^2. \end{array}$$

El conjunto de soluciones de la desigualdad logarítmica (22) se obtiene como una unión de conjuntos de soluciones de estos dos sistemas:

$$x \in \left(\frac{3}{8};\ \frac{1}{2}\right) \cup \left(1;\ \frac{3}{2}\right).$$

10.8. Representación geométrica de un conjunto de soluciones de una desigualdad con dos incógnitas. Supongamos que tenemos cierta desigualdad con dos incógnitas. El conjunto de soluciones de tal desigualdad es un conjunto de pares ordenados de números reales $(x; y)$, el cual puede ser representado geométricamente como un conjunto de puntos del plano de coordenadas Oxy.

La unión del conjunto de soluciones de las desigualdades

$$F(x, y) > 0, \quad F(x, y) < 0$$

y de la ecuación

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

da el conjunto de todos los puntos del plano de coordenadas Oxy, a excepción de los puntos, en los cuales la función de dos variables $F(x, y)$ no está determinada. Así por ejemplo, la unión de un conjunto de soluciones de las desigualdades

$$\frac{x}{y} > 0, \quad \frac{x}{y} < 0$$

y de la ecuación

$$\frac{x}{y} = 0$$

da un conjunto de todos los puntos del plano numérico $\mathbf{R}^2$, a excepción del conjunto de los puntos que pertenecen al eje Ox.

La ecuación $F(x, y) = 0$ proporciona un conjunto de puntos de «la frontera de separación» (o una parte de la misma) entre los conjuntos de los puntos definidos por las desigualdades $F(x, y) > 0$ y $F(x, y) < 0$.

Si la ecuación $F(x, y) = 0$ tiene una única solución respecto a la variable y (x se considera parámetro), entonces se dice que la ecuación $F(x, y) = 0$ define la función

$$y = f(x).$$

Si la ecuación $F(x, y) = 0$ tiene una única solución respecto a la variable x (y se considera parámetro), entonces se dice que la ecuación $F(x, y) = 0$ define la función

$$x = g(y),$$

donde y se considera una variable independiente, y x, una variable dependiente.

Puede resultar que la ecuación $F(x, y) = 0$ no tiene una sola solución ni respecto a la variable x ni respecto a la variable y. En este caso se dice que la ecuación $F(x, y) = 0$ representa una correspondencia entre los conjuntos de los valores admisibles de las variables x e y.

En la fig. 4.13—4.28 están representadas ciertas figuras definidas por las desigualdades que tienen la forma $y >$

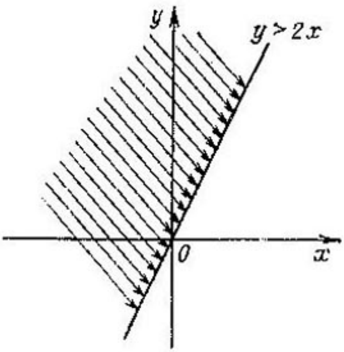


Fig. 4.13.

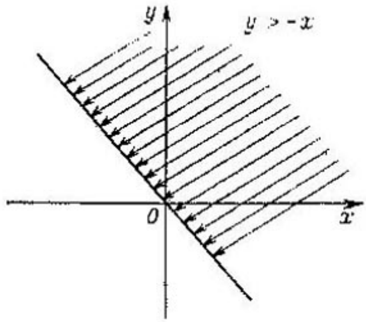


Fig. 4.14.

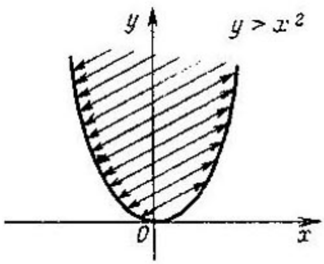


Fig. 4.15.

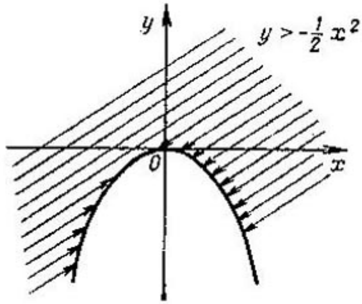


Fig. 4.16.

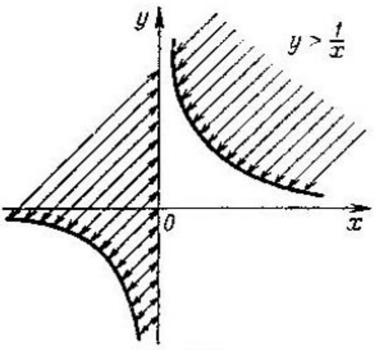


Fig. 4.17.

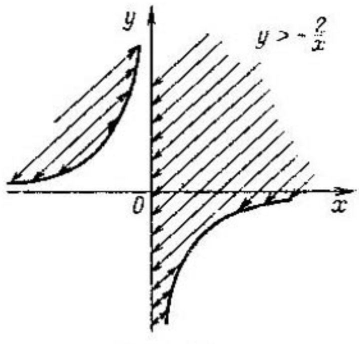


Fig. 4.18,

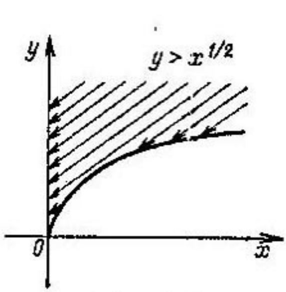


Fig. 4.19.

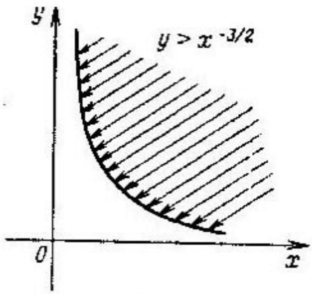


Fig. 4.20.

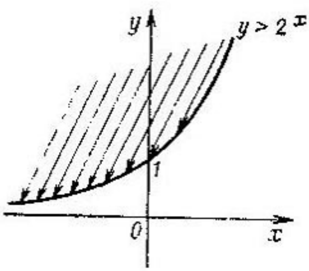


Fig. 4.21.

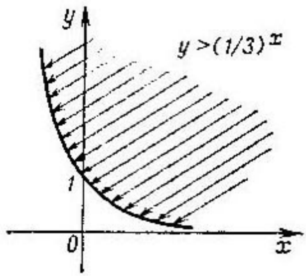


Fig. 4.22.

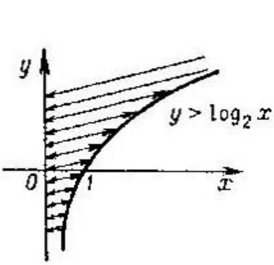


Fig. 4.23.

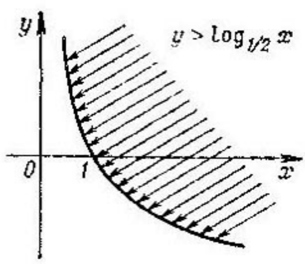


Fig. 4.24.

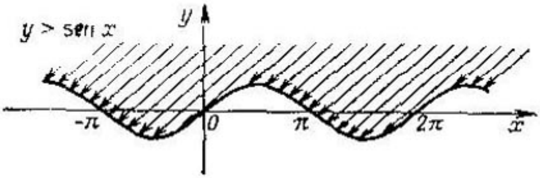


Fig. 4.25.

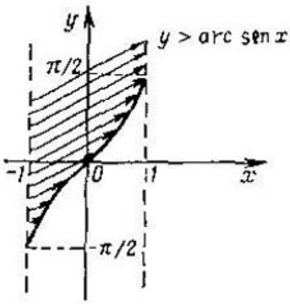


Fig. 4.26.

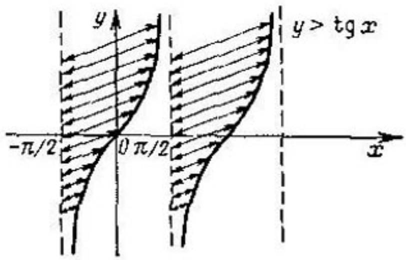


Fig. 4.27.

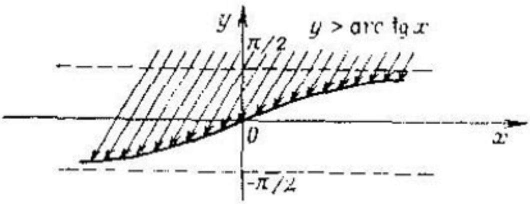


Fig. 4.28.

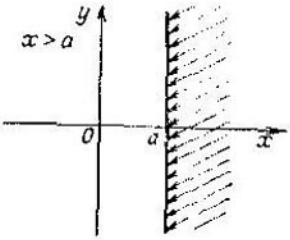


Fig. 4.29.

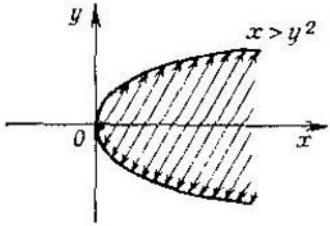


Fig. 4.30.

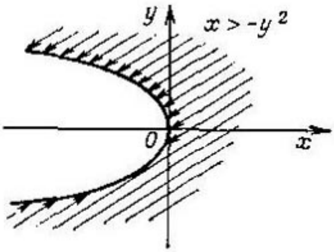


Fig. 4.31.

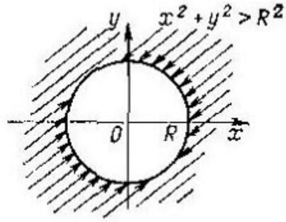


Fig. 4.32.

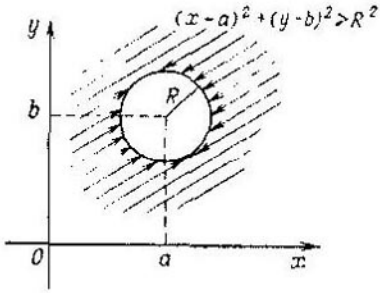


Fig. 4.33.

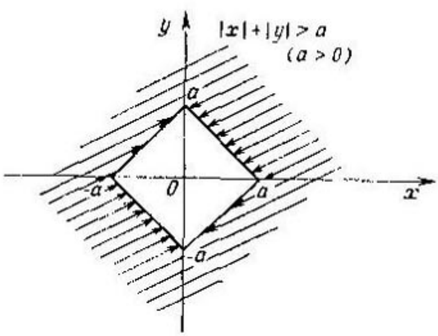


Fig. 4.34.

$> f(x)$. Estas figuras representan una parte del plano de coordenadas Oxy que se encuentra más arriba de la línea $y = f(x)$. Las figuras dadas por las desigualdades $y < f(x)$ representan una parte del plano que se encuentra más abajo de la línea $y = f(x)$.

En las fig. 4.29—4.31 se sombrean las figuras representadas por las desigualdades que tienen la forma $x > g(y)$.

En las fig. 4.32—4.34 están representadas las figuras dadas por las desigualdades que tienen la forma $F(x, y) > 0$, en el caso, cuando la ecuación $F(x, y) = 0$ representa la correspondencia entre las variables x e y:

1) la desigualdad $x^2 + y^2 > R^2$ representa el exterior del círculo del radio R con el centro en el origen de coordenadas (véase fig. 4.32);

2) la desigualdad $(x - a)^2 + (y - b)^2 > R^2$ representa el exterior del círculo de radio R con el centro en el punto con las coordenadas $(a; b)$ (véase la fig. 4.33);

3) la desigualdad que tiene la forma $|x| + |y| > a$ $(a > 0)$ representa el exterior de un cuadrado con los vértices en los puntos $(a; 0)$, $(0; a)$, $(-a; 0)$ y $(0; -a)$ (véase la fig. 4.34).

En las fig. 4.13—4.34 se representan las figuras simples dadas por las desigualdades con dos variables x e y. Sin embargo, destreza para construir estas figuras simples permite construir también figuras más complicadas representadas, por ejemplo, por un sistema de varias desigualdades con dos variables. Una figura geométrica dada por un sistema de desigualdades con dos variables representa la parte general (la intersección) de todas las figuras dadas por las desigualdades del sistema.

§ 11. Procedimientos de demostración de la validez de las desigualdades

11.1. Demostración de la validez de las desigualdades mediante una cadena de desigualdades equivalentes. Uno de los métodos de demostración de la validez de las desigualdades se basa en la construcción de una cadena de desigualdades equivalentes, al final de la cual se encuentra una desigualdad válida evidente.

Ejemplo 1. Demostrar la desigualdad

$$\frac{a^3+b^3}{2}\geqslant\left(\frac{a+b}{2}\right)^3,\quad \text{donde } a>0 \text{ y } b>0.$$

Sustituyamos esta desigualdad por la desigualdad equivalente

$$\frac{a^3+b^3}{2}-\left(\frac{a+b}{2}\right)^3\geqslant 0.$$

Suprimiendo los paréntesis y agrupando los sumandos del primer miembro de la desigualdad, obtenemos una desigualdad equivalente

$$\frac{3}{8}(a+b)(a-b)^2\geqslant 0.$$

Puesto que $a>0$ y $b>0$, la validez de la última desigualdad es evidente, lo que demuestra la validez de la desigualdad inicial equivalente.

Ejemplo 2. Demostrar la desigualdad

$$\frac{1}{\log_2\pi}+\frac{1}{\log_\pi 2}>2.$$

Basándose en la propiedad de los logaritmos

$$\frac{1}{\log_\pi 2}=\log_2\pi$$

la desigualdad inicial es equivalente a la desigualdad

$$\log_2\pi+\frac{1}{\log_2\pi}>2. \tag{1}$$

En vigor de que

$$\log_2\pi>0,$$

en el primer miembro de la desigualdad (1) se encuentra la suma de dos números positivos inversos recíprocamente, distintos de la unidad. En vigor de la desigualdad 4) p. 9.2 la desigualdad (1) es válida y, por consiguiente, es válida también la desigualdad inicial.

Ejemplo 3. Demostrar que para cualesquiera x e y reales se cumple la desigualdad

$$x^2+2xy+3y^2+2x+6y+3\geqslant 0.$$

Efectuemos las transformaciones siguientes del primer miembro de la desigualdad

$$\begin{aligned} x^2 + 2xy + 3y^2 + 2x + 6y + 3 &= \\ &= (x^2 + 2xy + y^2) + 2y^2 + 2x + 6y + 3 = \\ &= (x + y)^2 + 2(x + y) + 2y^2 + 4y + 3 = \\ &= [(x + y)^2 + 2(x + y) + 1] + 2y^2 + 4y + 2 = \\ &= [(x + y) + 1]^2 + 2(y^2 + 2y + 1) = \\ &= [(x + y) + 1]^2 + 2(y + 1)^2. \end{aligned}$$

De esta manera sucede que como resultado de las transformaciones idénticas del primer miembro de la desigualdad inicial ésta se reduce a la desigualdad equivalente

$$(x + y + 1)^2 + 2(y + 1)^2 \geqslant 0,$$

la validez de la cual es evidente.

Análogamente, utilizando las desigualdades del p. 9.2, se puede demostrar, por ejemplo, las desigualdades siguientes:

1) $a^2 + b^2 + c^2 + 3 \geqslant 2(a + b + c)$;

2) $\frac{bc}{a} + \frac{ac}{b} + \frac{ab}{c} \geqslant a + b + c \quad (a > 0,\ b > 0,\ c > 0)$;

3) $(a + b)(b + c)(a + c) \geqslant 8abc \quad (a > 0,\ b > 0,\ c > 0)$;

4) $a^2 + b^2 + c^2 \geqslant ab + bc + ac$ (a, b, c son los números de un mismo signo).

11.2. Demostración de la validez de las desigualdades con la utilización de las propiedades de las funciones que entran en las desigualdades. La validez de ciertas desigualdades puede ser demostrada con la utilización de las propiedades de las funciones que entran en las desigualdades.

Así, por ejemplo, la validez de la desigualdad

$$\operatorname{tg} x + \operatorname{ctg} x \geqslant 2$$

para todos $x \in (0; \pi/2)$ se desprende de la relación

$$\operatorname{ctg} x = \frac{1}{\operatorname{tg} x}.$$

Efectivamente, debido a esta relación la desigualdad inicial se escribe en la forma de la desigualdad

$$\operatorname{tg} x + \frac{1}{\operatorname{tg} x} \geqslant 2,$$

la cual es válida para todos los valores $x \in (0; \pi/2)$, puesto que en este intervalo $\operatorname{tg} x$ es positiva, y la suma de dos números cualesquiera positivos inversos recíprocamente es mayor o igual a dos.

Otro ejemplo de desigualdad, cuya validez se demuestra basándose en las propiedades de las funciones que entran en la misma, es la desigualdad

$$0 \leqslant \operatorname{sen}^8 x + \cos^{20} x \leqslant 1.$$

La validez de esta desigualdad, y, por ejemplo, de la desigualdad

$$\sqrt[4]{\operatorname{sen} x} + \sqrt[4]{\cos x} \geqslant 1$$

se deduce de la acotación de las funciones $y = \operatorname{sen} x$, $y = \cos x$ y de las relaciones del p. 3.2 del capítulo 7.

En el ejemplo siguiente un factor decisivo es la aplicación de la propiedad de la monotonía de la función exponencial.

Demostremos que si $a + b = c$ $(a > 0,\ b > 0)$, entonces

1) $a^\alpha + b^\alpha > c^\alpha$, donde $\alpha < 1$;

2) $a^\beta + b^\beta < c^\beta$, donde $\beta > 1$.

Examinemos los números $\frac{a}{c}$ y $\frac{b}{c}$. Según la condición del problema

$$\frac{a}{c} + \frac{b}{c} = 1 \quad \text{y} \quad \frac{a}{c} > 0, \qquad \frac{b}{c} > 0.$$

Si la suma de dos números positivos es igual a la unidad, entonces cada uno de estos números es menor que la unidad:

$$0 < \frac{a}{c} < 1, \quad 0 < \frac{b}{c} > 1.$$

Entonces, en vigor de la propiedad de la función exponencial monótona decreciente con una base que es menor que la unidad, tenemos

$$\left(\frac{a}{c}\right)^{\alpha} > \frac{a}{c}, \quad \left(\frac{b}{c}\right)^{\alpha} > \frac{b}{c},$$
$$\left(\frac{a}{c}\right)^{\beta} < \frac{a}{c}, \quad \left(\frac{b}{c}\right)^{\beta} < \frac{b}{c}. \tag{2}$$

donde α es un número real cualquiera menor que la unidad, y β, un número real mayor que la unidad. De la desigualdad (2) se desprenden las desigualdades

$$\left(\frac{a}{c}\right)^{\alpha} + \left(\frac{b}{c}\right)^{\alpha} > \frac{a}{c} + \frac{b}{c} = 1 \Leftrightarrow a^{\alpha} + b^{\alpha} > c^{\alpha},$$

$$\left(\frac{a}{c}\right)^{\beta} + \left(\frac{b}{c}\right)^{\beta} < \frac{a}{c} + \frac{b}{c} = 1 \Leftrightarrow a^{\beta} + b^{\beta} < c^{\beta}.$$

11.1. Ciertos procedimientos especiales de demostración de la validez de las desigualdades.

1. Uno de los procedimientos especiales de la demostración de la validez de las desigualdades consiste en lo siguiente. Supongamos, por ejemplo, que es necesario demostrar la validez de la desigualdad $A < B$. Si se consigue elegir tal valor C que $A < C$ y demostrar la validez de la desigualdad $C \leqslant B$, entonces será demostrada también la validez de la desigualdad inicial $A < B$. Precisamente tal procedimiento se usa para la demostración de la acotación de sucesión (véase p. 2.4 del capítulo 8)

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

2. La validez de ciertas desigualdades, en las cuales el primer y segundo miembros son funciones del argumento natural, puede ser demostrada mediante el procedimiento siguiente. La desigualdad inicial se transforma en desigualdad equivalente, cuyo segundo miembro es constante, y el primer miembro es función del argumento natural. Entonces, considerando la función del argumento natural que se encuentra en el primer miembro de la desigualdad como un término común de cierta sucesión, se puede reducir la demostración de la validez de la desigualdad a la examinación de las propiedades de esta sucesión.

Ejemplo. Demostrar que la desigualdad

$$2^{n-1} \leqslant n!$$

es válida para cualquier n natural.

Transformamos la desigualdad dada en desigualdad equivalente

$$\frac{2^{n-1}}{n!} \leqslant 1. \qquad (3)$$

Examinemos la sucesión (x_n) representada por la fórmula del término común

$$x_n = \frac{2^{n-1}}{n!}.$$

Es muy fácil demostrar que esta sucesión es una sucesión monótona decreciente (véase p. 1.5 del capítulo 8) y, por consiguiente, el término mayor de la sucesión dada es el primer término, que es igual a la unidad. Pero, puesto que para $n = 1$ la desigualdad (3) es válida, entonces ella es válida y para todos los demás valores n.

3. Ciertas desigualdades pueden ser demostradas mediante el método de inducción matemática. Precisamente este método se usa para la demostración de la acotación de la sucesión (véase el ejemplo 3 p. 2.4 del capítulo 8)

$$x_n = \underbrace{\sqrt{c + \sqrt{c + \ldots + \sqrt{c}}}}_{n \text{ raíces}}.$$

Mediante el método de inducción matemática pueden ser demostradas tambien las desigualdades siguientes:

1) $n! < \left(\frac{n+1}{n}\right)^n$;

2) $|\operatorname{sen} nx| \leqslant n\,|\operatorname{sen} x|$;

3) $n! > 2^{n-1}$, si $n > 2$;

4) $2^n n! < n^n$, si $n > 2$.

11.4. Ciertos procedimientos de verificación de la validez de las desigualdades numéricas. Cuando resolvemos ecuaciones y desigualdades puede surgir la necesidad de aclarar, cuál de dos (o varios) números es mayor y cuál es menor. A continuación, en ejemplos, se examinan ciertos procedimientos simples de verificación de la validez de las desigualdades numéricas.

1) Para aclarar, cuál de dos números racionales $\frac{m}{n}$ y $\frac{p}{q}$ es mayor, calculemos la diferencia de estos números

$$\frac{m}{n}-\frac{p}{q}=\frac{mq-np}{nq}.$$

Si la diferencia es mayor que cero, entonces $\frac{m}{n}>\frac{p}{q}$; si la diferencia es menor que cero, entonces $\frac{m}{n}<\frac{p}{q}$.

2) Para comparar dos números irracionales a y b se usa muy a menudo el procedimiento siguiente: se supone que un número es mayor que el otro (por ejemplo, $a>b$) y se construye una cadena de desigualdades equivalentes que da bien una desigualdad numérica válida (y entonces la suposición de que $a>b$ es válida), o bien una desigualdad numérica que no es válida (en este caso la suposición de que $a>b$ no es válida y por consiguiente, $a<b$). Por ejemplo, sea que es necesario comparar, cuál de dos números $\sqrt{14}-\sqrt{3}$ y 2 es mayor. Supongamos que $\sqrt{14}-\sqrt{3}$ es menor que 2:

$$\sqrt{14}-\sqrt{3}<2; \tag{4}$$

ambos miembros de la desigualdad (4) son positivos. Elevemos ambos miembros de la desigualdad al cuadrado:

$$(\sqrt{14}-\sqrt{3})^2<4,$$

$$14-2\sqrt{42}+3<4.$$

La última desigualdad se puede reescribir en la forma

$$13<2\sqrt{42}.$$

Elevando ambos miembros de la desigualdad al cuadrado, obtenemos la desigualdad $169<168$.

Es evidente que la última desigualdad no es válida. Por consiguiente, la suposición de que el número $\sqrt{14}-\sqrt{3}$ es menor que 2 no es válida. El número $\sqrt{14}-\sqrt{3}$ es mayor que 2.

3) Examinemos en dos ejemplos, como se puede verificar la validez de una desigualdad mediante cálculos aproximados.

Ejemplo 1. Demostrar que

$$\operatorname{sen} 39^\circ > 1/\sqrt{3}. \tag{5}$$

Tratamos de elegir el número a de tal manera que sen $39^\circ > a$ y $a > 1/\sqrt{3}$. El valor más próximo en la «tabla» de la función sen x es sen $30^\circ = 1/2$:

$$\operatorname{sen} 30^\circ < \operatorname{sen} 39^\circ. \tag{6}$$

Sin embargo, tomando como número a el número $1/2$, no se puede demostrar nada, puesto que $1/2 < 1/\sqrt{3}$. Tratamos de mejorar la estimación (6). Esta estimación del sen 39° puede ser mejorada mediante dos procedimientos. El primer procedimiento consiste en lo siguiente: mediante las fórmulas de la mitad del ángulo se puede calcular sucesivamente el sen 15°, cos 15°, sen $7^\circ 30'$, cos $7^\circ 30'$ y luego podemos calcular

$$\operatorname{sen} 37^\circ 30' = \operatorname{sen} 30^\circ \cos 7^\circ 30' + \operatorname{sen} 7^\circ 30' \cos 30^\circ$$

y, teniendo en cuenta la desigualdad

$$\operatorname{sen} 37^\circ 30' < \operatorname{sen} 39^\circ, \tag{7}$$

tomando en calidad de a el número $37^\circ 30'$, se puede demostrar que sen $37^\circ 30' > 1/\sqrt{3}$.

En el segundo procedimiento usamos una estimación más grosera en comparación con la estimación (7). En el capítulo 7 fue calculado el sen 36°:

$$\operatorname{sen} 36^\circ = \frac{\sqrt{10-2\sqrt{5}}}{4}.$$

Puesto que en el primer cuarto sen x es una función monótona creciente, de la desigualdad

$$36^\circ < 39^\circ$$

sigue la desigualdad

$$\operatorname{sen} 36^\circ < \operatorname{sen} 39^\circ,$$

es decir,

$$\operatorname{sen} 39^\circ > \frac{\sqrt{10-2\sqrt{5}}}{4}.$$

Ahora mediante el procedimiento descrito más arriba, demostramos que

$$\frac{\sqrt{10-2\sqrt{5}}}{4} > \frac{1}{\sqrt{3}}.$$

La validez de la última desigualdad se desprende de los cálculos citados más abajo:

$$\left(\frac{\sqrt{10-2\sqrt{5}}}{4}\right)^2 > \frac{1}{3},$$

$$\frac{5-\sqrt{5}}{8} > \frac{1}{3},$$

$$15-3\sqrt{5} > 8,$$

$$7 > 3\sqrt{5},$$

$$49 > 45.$$

De esta manera fue demostrado que

$$\operatorname{sen} 39^\circ > \operatorname{sen} 36^\circ > 1/\sqrt{3}.$$

Ejemplo 2. Comparar, cuál de los números es mayor: $\log_2 3$ ó $\log_3 5$.

Estimemos los números, cuyos logaritmos se calculan por las potencias de las bases de los logaritmos:

$$2^1 < 3 < 2^2, \quad 3^1 < 5 < 3^2. \tag{8}$$

Puesto que la función $\log_a x$ para $a > 1$ es una función monótona creciente, de las desigualdades (8) se deducen las desigualdades

$$\log_2 (2^1) < \log_2 3 < \log_2 (2^2),$$

$$1 < \log_2 3 < 2,$$

$$\log_3 (3^1) < \log_3 5 < \log_3 (3^2),$$

$$1 < \log_3 5 < 2;$$

de las últimas desigualdades es imposible deducir, cuál de los logaritmos es mayor, puesto que ellos se encuentran entre los mismos números.

Mejoremos las estimaciones (8). Vamos a estimar los números 3 y 5 por las potencias fraccionarias de los números 2 y 3 respectivamente. Tomemos los números $2^{1+\frac{1}{2}}=2\sqrt{2}$ y $3^{1+\frac{1}{2}}=3\sqrt{3}$ y comparemos con estos números los números 3 y 5 respectivamente. Mediante el procedimiento descrito en 2) se puede mostrar que

$$3>2\sqrt{2} \text{ y } 5<3\sqrt{3}.$$

De esta manera, tenemos

$$2^{3/2}<3<2^2, \quad 3^1<5<3^{3/2},$$

de donde respectivamente

$$\frac{3}{2}<\log_2 3<2, \quad 1<\log_3 5<\frac{3}{2}. \tag{9}$$

Comparando las últimas desigualdades, obtenemos que $\log_2 3$ es mayor que $\log_3 5$.

Observemos que si surge la necesidad, se puede seguir mejorando las estimaciones (9). Pues bien, demostrando la validez de las desigualdades

$$2^{\frac{3}{2}+\frac{1}{4}}=2\cdot\sqrt{2}\cdot\sqrt[4]{2}>3, \quad 3^{1+\frac{1}{4}}=3\sqrt[4]{3}<5,$$

obtenemos las estimaciones siguientes para los $\log_2 3$ y $\log_3 5$:

$$\frac{3}{2}<\log_2 3<\frac{7}{4}, \quad \frac{5}{4}<\log_3 5<\frac{3}{2}$$

CAPÍTULO 5

Método de coordenadas

El concepto de sistema de coordenadas rectangular en el plano apareció primeramente en geometría antes del comienzo de nuestra era. Mediante este sistema un matemático de la escuela de Alejandría, Apolonio, definía y estudiaba las curvas de segundo orden; la elipse, hipérbola y parábola. En el siglo XVIII el matemático francés R. Descartes (y al mismo tiempo Fermat) introdujo la regla de elección de signos en el sistema de coordenadas rectangular y créó las bases de la geometría analítica en el plano, que es la parte de las matemáticas que establece la relación entre el álgebra y la geometría. Las obras de Descartes fueron preparadas por las obras de otro matemático francés, Viète, el cual fue el primero en introducir en el álgebra las designaciones literales (tanto de los valores conocidos, como de los desconocidos). La geometría analítica jugó un papel muy importante en el desarrollo del concepto de número: gracias a la regla de elección de los signos de las coordenadas los números negativos (los cuales no eran reconocidos por la mayoría de los matemáticos de la Edad Media) obtuvieron una representación clara y fueron aceptados en las matemáticas definitivamente. Muchos años más tarde, la aplicación del sistema de coordenadas cartesianas rectangulares jugó un papel decisivo en la afirmación de los números complejos en las matemáticas.

§ 1. Sistemas de coordenadas

1.1. Eje de coordenadas. La recta, en la cual se fijan dos puntos distintos: el punto O, que se llama *origen de coordenadas*, y el punto E, que se llama *punto de unidad* (fig. 5.1), se denomina *eje de coordenadas.*

El punto E se suele situar a la derecha del punto O. Se considera dirección positiva del eje de coordenadas la dirección del rayo que sale del punto O y contiene el punto E. La dirección contraria se considera dirección negativa

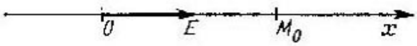


Fig. 5.1.

del eje de coordenadas. El segmento OE se llama de escala o segmento unidad. El eje de coordenadas se suele designar por Ox.

El vector $\overrightarrow{OE}$ se denomina *vector unitario* (o *versor*) y se suele designar con $\boldsymbol{e}$.

Se llama *coordenada* del punto M_0, situado en el eje de coordenadas, el número x_0 que se define por la igualdad $x_0 = \pm \frac{|OM_0|}{|OE|}$; además delante de la fracción se toma el signo más, cuando los puntos E y M_0 están situados a un lado del punto O, y el signo menos, cuando los puntos E y M_0 están situados a distintos lados respecto al punto O; si el punto M_0 coincide con el punto O, entonces $x_0 = 0$.

La distancia entre los puntos M_1 y M_2 con las coordenadas x_1 y x_2 se calcula según la fórmula

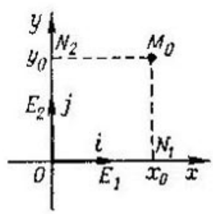


Fig. 5.2.

$$d = |x_2 - x_1|.$$

1.2. Sistema de coordenadas cartesiano rectangular en el plano. Un par ordenado de dos ejes de coordenadas Ox y Oy, perpendiculares entre sí, se llama *sistema de coordenadas cartesiano rectangular* en el plano, además, el origen de coordenadas para cada uno de los ejes es su punto común O (origen de coordenadas) (fig. 5.2). Los ejes Ox y Oy se ordenan de la manera siguiente: si giramos el eje Ox alrededor del punto O en un ángulo $\pi/2$ en sentido antihorario, entonces coincidirá con el eje Oy. Al mismo tiempo los segmentos de escala OE_1 y OE_2 de los ejes de coordenadas Ox y Oy se eligen de tal manera que sus longitudes sean iguales:

$$|OE_1| = |OE_2|.$$

Entonces al girar el eje Ox en un ángulo $\pi/2$ en sentido antihorario alrededor del punto O los puntos E_1 y E_2 coinciden. El eje Ox se llama *eje de abscisas* y el eje Oy, *eje de ordenadas*.

Los vectores $\boldsymbol{e}_1 = \overrightarrow{OE_1}$ y $\boldsymbol{e}_2 = \overrightarrow{OE_2}$ se llaman *vectores básicos* del sistema de coordenadas cartesiano rectangular (o *versores*) y se suelen designar con letras $\boldsymbol{i}$ y $\boldsymbol{j}$:

$$\boldsymbol{e}_1 = \boldsymbol{i}, \quad \boldsymbol{e}_2 = \boldsymbol{j}.$$

De esta manera, se puede considerar que el sistema de coordenadas cartesiano rectangular se representa por cierto punto O (por el origen de coordenadas) y por un par ordenado de vectores unitarios perpendiculares entre sí $(\boldsymbol{i}; \boldsymbol{j})$.

Los ejes coordenados Ox y Oy dividen, el plano en cuatro *cuadrantes*. La parte del plano que está situada más arriba del eje Ox y a la derecha del eje Oy se considera primer cuadrante; la parte del plano que se encuentra más arriba del eje Ox y a la izquierda del eje Oy es el segundo cuadrante; la parte del plano situada más abajo del eje Ox y a la izquierda del eje Oy es el tercer cuadrante, y la parte del plano que se encuentra más abajo del eje Ox y a la derecha del eje Oy, el cuarto cuadrante. El plano con el sistema de coordenadas construidos se llama *plano de coordenadas*.

Para el procedimiento expuesto más arriba de ordenación de los ejes de coordenadas el sistema de coordenadas se llama sistema de coordenadas cartesiano rectangular *derecho* a diferencia del sistema de coordenadas cartesiano rectangular *izquierdo*, en el cual los ejes se ordenan de la manera siguiente: el primer eje (eje Ox) coincide con el segundo (eje Oy) mediante el giro en un ángulo $\pi/2$ en sentido horario. En lo sucesivo bajo el término «sistema de coordenadas cartesiano rectangular» se entenderá el sistema de coordenadas cartesiano rectangular derecho.

En los suplementos se usan también los sistemas de coordenadas *oblicuángulos* (derechos e izquierdos), en los cuales la coincidencia de los ejes se realiza al girar en un ángulo, distinto del recto.

Coordenadas de un punto en el plano. Sea que Oxy es un sistema de coordenadas cartesiano rectangular en el plano con el origen O (véase fig. 5.2) y M_0, cierto punto del plano. Bajemos del punto M_0 sobre los ejes Ox y Oy las perpendiculares que intersecan los ejes coordenados indicados en los puntos N_1 y N_2 respectivamente. Designemos la coordenada del punto N_1, situada en el eje Ox, a través de x_0, y la coordenada del punto N_2 que se encuentra en el eje Oy, mediante y_0.

Se llaman *coordenadas del punto* M_0 en el sistema de coordenadas cartesiano rectangular a un par ordenado de números $(x_0;\ y_0)$.

El número x_0 se llama *abscisa* del punto M_0, y el número y_0, *ordenada* de este mismo punto y se escribe M_0 $(x_0;\ y_0)$.

La *distancia* entre los puntos A y B que tienen las coordenadas $(x_1;\ y_1)$ y $(x_2;\ y_2)$, respectivamente, se calcula mediante la fórmula

$$|AB| = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}.$$

Transformaciones de las coordenadas rectangulares por traslación paralela de los ejes. Sea que Oxy y $O'x'y'$ son dos sistemas de coordena-

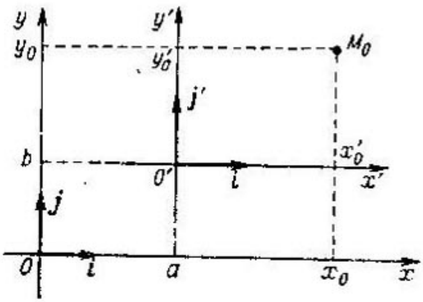


Fig. 5.3.

das cartesianos rectangulares en el plano con el origen de coordenadas O y O' $(O \neq O')$ respectivamente, los ejes de una misma dirección (fig. 5.3) y los mismos segmentos de escala. Y sea que el punto O' tiene las coordenadas $(a;\ b)$ respecto al sistema de coordenadas Oxy. El número a se llama magnitud de desplazamiento del sistema de coordenadas $O'x'y'$ respecto al sistema Oxy en dirección al eje Ox, y el número b, magnitud de desplazamiento en dirección del eje Oy.

Elijamos en el plano un punto M_0 con las coordenadas $(x_0;\ y_0)$ respecto al sistema Oxy. Entonces sus coordenadas $(x'_0;\ y'_0)$ respecto al sistema $O'x'y'$ están relacionadas con las coordenadas respecto al sistema Oxy mediante las fórmulas

$$x'_0=x_0-a, \qquad y'_0=y_0-b.$$

A estas fórmulas se las llama *fórmulas de transformación* de las coordenadas rectangulares del punto M_0.

Formulemos la *regla* de transformación de coordenadas del punto M_0 para el paso de un sistema de coordenadas rectangular al otro:

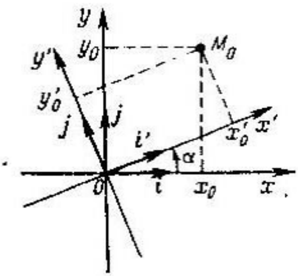


Fig. 5.4.

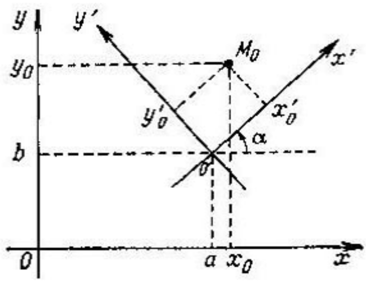


Fig. 5.5.

En la traslación paralela del sistema de coordenadas rectangular en una magnitud a en dirección del eje Ox y en una magnitud b en dirección del eje Oy las abscisas de todos los puntos disminuyen o aumentan en función del signo de la magnitud a en una magnitud a y las ordenadas, en una magnitud b.

Fórmulas de transformación de las coordenadas rectangulares por giro de los ejes. Sea Oxy y $Ox'y'$ dos sistemas de coordenadas cartesianas rectangulares, además, el sistema de coordenadas $Ox'y'$ se obtuvo del sistema Oxy mediante un giro en un ángulo α en torno al origen común de coordenadas O (fig. 5.4).

Elijamos en el plano un punto M_0 que tenga unas coordenadas $(x_0;\ y_0)$ respecto al sistema de coordenadas Oxy, y unas coordenadas $(x'_0;\ y'_0)$ respecto al sistema de coordenadas $Ox'y'$. Las coordenadas del punto M_0 en los sistemas de coordenadas Oxy y $Ox'y'$ se relacionan por las fórmulas

$$\begin{aligned} x_0&=x'_0\cos\alpha-y'_0\operatorname{sen}\alpha,\\ y_0&=x'_0\operatorname{sen}\alpha+y'_0\cos\alpha \end{aligned} \tag{1}$$

o mediante las fórmulas equivalentes

$$\begin{aligned} x'_0&=x_0\cos\alpha+y_0\operatorname{sen}\alpha,\\ y'_0&=-x_0\operatorname{sen}\alpha+y_0\cos\alpha. \end{aligned}$$

Fórmulas de transformación de las coordenadas rectangulares por traslación paralela, y por un giro de los ejes. Sean Oxy y $O'x'y'$ dos sistemas de coordenadas cartesianas rectangulares y sea que el sistema de coordenadas $O'x'y'$ se obtuvo del sistema Oxy por traslación paralela con una magnitud de desplazamiento, respectivamente, a y b, y por un giro posterior alrededor del punto O' (imagen del punto O) en un ángulo α (fig. 5.5).

Sea que M_0 es un punto del plano que tiene las coordenadas $(x_0; y_0)$ respecto al sistema de coordenadas Oxy y las coordenadas

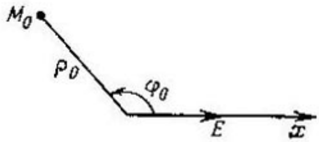


Fig. 5.6.

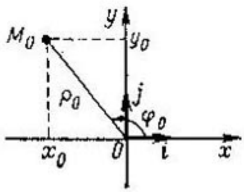


Fig. 5.7.

$(x_0'; y_0')$ respecto al sistema de coordenadas $O'x'y'$. Las coordenadas del punto M_0 respecto a los sistemas de coordenadas Oxy y $O'x'y'$ se relacionan por las fórmulas

$$\begin{aligned} x_0 &= x_0' \cos\alpha - y_0' \operatorname{sen}\alpha + a, \\ y_0 &= x_0' \operatorname{sen}\alpha + y_0' \cos\alpha + b, \end{aligned} \tag{2}$$

las cuales expresan las coordenadas del punto M_0 respecto al sistema de coordenadas Oxy a través de las coordenadas del mismo punto M_0 respecto al sistema $O'x'y'$.

Resolviendo las relaciones (2) respecto a las magnitudes x_0' e y_0', obtenemos las expresiones para las coordenadas del punto M_0 respecto al sistema $O'x'y'$ a través de las coordenadas del punto M_0 respecto al sistema Oxy:

$$\begin{aligned} x_0' &= (x_0 - a)\cos\alpha + (y_0 - b)\operatorname{sen}\alpha, \\ y_0' &= -(x_0 - a)\operatorname{sen}\alpha + (y_0 - b)\cos\alpha. \end{aligned}$$

1.3. Sistema de coordenadas polares. Relación entre las coordenadas rectangulares y polares. El *sistema de coordenadas polares* en el plano se determina por la elección de un punto O, llamado *polo* y por un rayo dirigido Ox con el origen en el punto O (*eje polar*) con un segmento OE elegido en este rayo, cuya longitud se toma por unidad de escala (fig. 5.6).

Al punto M_0 que pertenece al plano y es distinto del punto O se le asigna respectivamente un par ordenado de números $(\rho_0; \varphi_0)$, llamado *coordenadas polares* del punto M_0. El número ρ_0 que se llama *radio polar* del punto M_0 representa la longitud del segmento OM_0. El número φ_0 que se llama *ángulo polar* es el valor del ángulo xOM_0, medido en radianes (véase fig. 5.6).

Entre el conjunto de todos los puntos del plano (a excepción del punto O) y el conjunto de los pares ordenados de números $(\rho_0;\ \varphi_0)$ donde φ_0 $[0;\ 2\pi]$, existe una correspondencia biunívoca. Para el punto O (del polo) el valor del ángulo polar no está definido.

Relación entre las coordenadas rectangulares y polares. Si tomamos el polo por el origen del sistema de coordenadas cartesiano rectangular, la dirección del eje polar, por la dirección positiva del eje Ox (fig. 5.7), entonces entre las coordenadas polares del punto M_0 $(\rho_0;\ \varphi_0)$, el cual no coincide con el polo, y las coordenadas rectangulares $(x_0;\ y_0)$ del mismo punto M_0 existe la dependencia siguiente:

$$x_0 = \rho_0 \cos \varphi_0, \qquad y_0 = \rho_0 \operatorname{sen} \varphi_0.$$

Estas fórmulas se llaman *fórmulas del paso* de las coordenadas polares de un punto M_0 a las coordenadas rectangulares del mismo punto.

A la inversa, representando las coordenadas rectangulares $(x_0;\ y_0)$ de cualquier punto M_0 que no coincide con el origen de coordenadas, se calculan sus coordenadas polares mediante las fórmulas

$$\rho_0 = \sqrt{x_0^2 + y_0^2},$$

$$\operatorname{sen} \varphi_0 \frac{y_0}{\sqrt{x_0^2+y_0^2}}, \qquad \cos \varphi_0 = \frac{x_0}{\sqrt{x_0^2+y_0^2}}.$$

1.4. Sistema de coordenadas cartesianas rectangulares en el espacio. Sea que α es cierto plano en el espacio con

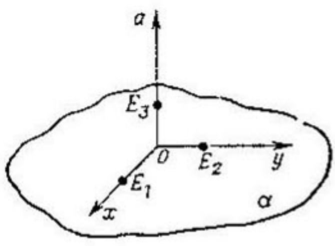


Fig. 5.8.

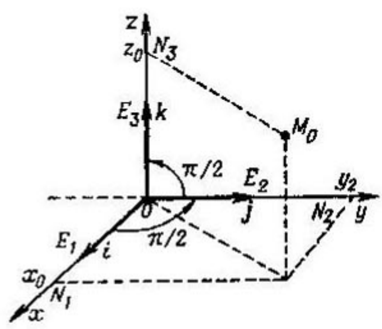


Fig. 5.9.

el sistema de coordenadas cartesiano rectangular Oxy eligido en él. Tracemos por el punto O una recta a perpendicular al plano α (fig. 5.8). Elijamos en la recta a una dirección y un segmento de escala OE_3, el cual es igual a los segmentos de escala OE_1 y OE_2 (en la fig. 5.8 la dirección positiva en la recta a se indica con una flecha). Este es el eje de coordenadas que designamos con Oz.

Una terna ordenada de los ejes de coordenadas perpendiculares entre sí Ox, Oy y Oz se llama *sistema cartesiano de coordenadas rectangulares* en el espacio. El sistema cartesiano de coordenadas rectangulares $Oxyz$ se llama *derecho* (fig. 5.9), si Oxy es un sistema de coordenadas rectangulares derecho, y el eje Oy coincide con el eje Oz mediante un giro en un ángulo $\pi/2$ en sentido horario. El sistema cartesiano de coordenadas rectangulares $Oxyz$ es izquierdo, si el sistema de coordenadas Oxy es izquierdo o, si Oxy es derecho el eje Oy coincide con el eje Oz mediante un giro en un ángulo $\frac{\pi}{2}$ en sentido horario.

El eje Ox se llama *eje de abscisas*, el eje Oy, *eje de ordenadas* y el eje Oz, *eje de z-coordenadas*. Los planos que pasan a través de cada dos ejes de coordenadas se llaman *planos de coordenadas*; el espacio, en el cual está eligido un sistema de coordenadas, se llama *espacio de coordenadas*.

Los vectores $\vec{OE_1}$, $\vec{OE_2}$, $\vec{OE_3}$ se llaman *vectores básicos* del sistema de coordenadas cartesianas rectangulares (o *versores*) y se suelen designar respectivamente por $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$.

Coordenadas de un punto en el espacio. Sea M_0 cierto punto del espacio y $Oxyz$ cierto sistema espacial de coordenadas cartesianas rectangulares con el origen en el punto O (véase fig. 5.9). Construyamos un plano que contiene un punto M_0, paralelo al plano de coordenadas Oyz. El plano construido interseca el eje de coordenadas Ox en cierto punto N_1. Designemos a la coordenada del punto N_1 con x_0. Al construir los planos, paralelos a los planos de coordenadas Oxz y Oxy, que pasan por el punto M_0, obtenemos los puntos N_2 y N_3, los cuales son puntos de intersección de los planos construidos con los ejes Oy y Oz, respectivamente. Designemos las coordenadas de los puntos N_2 y N_3 que pertenecen respectivamente a los ejes Oy y Oz, con y_0 y z_0.

Se llaman *coordenadas del punto* M_0 respecto al sistema cartesiano de coordenadas rectangulares $Oxyz$ a una terna ordenada de números $(x_0; y_0; z_0)$ y se escribe

$$M_0\,(x_0;\ y_0;\ z_0).$$

El número x_0 se llama *abscisa* del punto M_0, y_0, su *ordenada* y z_0, su *z-coordenada*.

La *distancia* entre los puntos $A(x_1;\ y_1;\ z_1)$ y $B(x_2;\ y_2;\ z_2)$ se calcula por la fórmula

$$|AB| = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}.$$

1.5. Ecuación del plano. Cada ecuación

$$Ax + By + Cz + D = 0, \tag{3}$$

que es lineal respecto a las variables x, y y z, define en el sistema espacial de coordenadas rectangulares $Oxyz$ un plano (A, B, C son números, de los cuales por lo menos uno no es igual a cero). Y a la inversa, cualquier plano en el sistema de coordenadas rectangulares $Oxyz$ puede ser representado por una ecuación que tiene la forma (3).

Las coordenadas de todos los puntos del espacio situados a un lado del plano determinado por la ecuación (3) satisfacen la desigualdad

$$Ax + By + Cz + D > 0,$$

y las coordenadas de todos los puntos que se encuentran en el otro lado satisfacen la desigualdad

$$Ax + By + Cz + D < 0.$$

En el sistema cartesiano de coordenadas rectangulares el vector $N = (A;\ B;\ C)$ es perpendicular al plano

$$Ax + By + Cz + D = 0.$$

La *distancia* d del punto $M_0(x_0;\ y_0;\ z_0)$ al plano

$$Ax + By + Cz + D = 0$$

se calcula mediante la fórmula

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2+B^2+C^2}}.$$

§ 2. Vectores

2.1. Vectores. Conceptos principales. Sea que A y B son dos puntos distintos del plano. El segmento AB, cuyo punto A se considera el origen, y el punto B, el extremo, se llama *vector* AB, y se designa $\overrightarrow{AB}$. También se dice que el vector $\overrightarrow{AB}$ está aplicado al punto A. La dirección que se define por un rayo AB se llama *dirección* del vector $\overrightarrow{AB}$, y la longitud del segmento AB se llama *longitud* (o *módulo*) del vector $\overrightarrow{AB}$. La longitud (el módulo) del vector $\overrightarrow{AB}$ se designa $|\overrightarrow{AB}|$. En los diagramas el vector $\overrightarrow{AB}$ se suele representar con una flecha rectangular con el origen en el punto A y el extremo en el punto B (fig. 5.10).

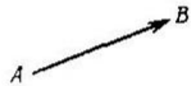


Fig. 5.10.

Entre todos los vectores con el origen en el punto A existe un vector, cuya longitud es igual a cero. Se considera que este vector comienza y termina en el punto A. El vector que comienza y termina en el punto A se llama vector *nulo* y se designa $\overrightarrow{AA}$. El concepto de dirección para el vector nulo $\overrightarrow{AA}$ no se introduce.

La definición de un vector espacial coincide exactamente con la definición de un vector en el plano, al mismo tiempo, sólo se considera que los puntos A y B son dos puntos del espacio.

Igualdad de vectores. Sea $\overrightarrow{AB}$ y $\overrightarrow{CD}$ dos vectores del plano (o del espacio). Se dice que el vector $\overrightarrow{AB}$ es igual al vector $\overrightarrow{CD}$, si

1) la longitud del segmento AB es igual a la longitud del segmento CD;

2) los rayos AB y CD están dirigidos en la misma dirección.

La igualdad de vectores se escribe mediante el signo de igualdad:

$$\overrightarrow{AB} = \overrightarrow{CD}.$$

Para tal definición de la igualdad de vectores el conjunto de todos los vectores, iguales al vector $\overrightarrow{AB}$, se llama *vector libre*. Los vectores libres se suelen designar mediante las letras latinas minúsculas $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{x}$, y sus longitudes se designan respectivamente $|\boldsymbol{a}|$, $|\boldsymbol{b}|$, $|\boldsymbol{c}|$, $|\boldsymbol{x}|$. El vector nulo se suele designar 0. Cada vector libre no nulo $\boldsymbol{a}$ tiene una infinidad de representaciones en forma de segmentos dirigidos, trazados desde distintos puntos del plano (espacio) A_1, A_2, . . . A_n, . . ., tales que las longitudes de los segmentos A_1B_1, A_2B_2, . . ., A_nB_n, . . . son iguales a la longitud del vector $\boldsymbol{a}$, y los rayos A_1B_1, A_2B_2, . . ., A_nB_n, .. son codirigidos y su dirección coincide con la dirección del vector $\boldsymbol{a}$ (fig. 5.11). El vector libre nulo 0 tiene también una infinidad de muchas representaciones en forma de puntos de un plano.

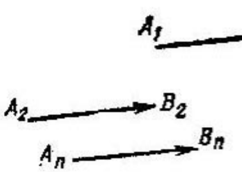


Fig. 5.11.

El vector libre se suele llamar simplemente *vector*. Un conjunto de vectores iguales entre sí, pertenecientes a la misma recta, se llama vector *deslizante*. Junto con los vectores libres y deslizantes se examinan también los vectores *fijos*, los cuales se caracterizan por un módulo, dirección y por la posición del punto inicial, llamado *punto de aplicación*. Dos vectores fijos se consideran *iguales*, si tienen no sólo los módulos y las direcciones iguales, sino también un punto inicial común, es decir, si los vectores iguales fijos son vectores coincidentes.

Un vector en el plano puede ser introducido mediante la utilización del concepto de isometría o, mejor decir, de un caso particular de isometría, del concepto de traslación paralela. Recordamos las definiciones de estos conceptos.

La aplicación de un plano sobre sí, para la cual se conservan las distancias entre los puntos, se llama *isometría*.

La aplicación del plano sobre sí, para la cual todos los puntos del plano se desplazan en la misma dirección a la misma distancia, se llama *traslación paralela*.

En estos términos la transformación del plano, para la cual cada punto M del plano se aplica sobre tal punto M_1 de este plano, que el rayo MM_1 está codirigido con el rayo AB y la distancia $|MM_1|$ es igual a la distancia $|AB|$, se llama *vector* (traslación paralela) que se define por un par ordenado de no coincidentes $(A; B)$ del plano.

Cualquier par de puntos coincidentes representa una transformación idéntica (es decir, una transformación, que aplica sobre sí cualquier punto del plano). La transformación idéntica se llama vector nulo y se designa mediante el símbolo $\mathbf{0}$.

Adición de vectores. Sean $\boldsymbol{a}$ y $\boldsymbol{b}$ dos vectores no nulos. Trazemos el vector $\boldsymbol{a}$ desde el punto O y designemos su extremo con la letra A (fig. 5.12): $\overrightarrow{OA} = \boldsymbol{a}$. Trazemos desde punto A el vector $\boldsymbol{b}$ y designemos su extremo con la letra B: $\overrightarrow{AB} = \boldsymbol{b}$. Se llama *suma* de dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ a un vector que tiene el origen en el punto O y el extremo en el punto B ($\overrightarrow{OB} = \boldsymbol{c}$) y se escribe

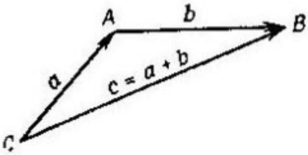


Fig. 5.12.

$$\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}.$$

Del vector $\boldsymbol{c}$ se dice que está obtenido como resultado de la suma de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$.

Las propiedades de la operación de adición de los vectores son:

1) la propiedad del vector nulo:

$$\boldsymbol{a} + \mathbf{0} = \boldsymbol{a};$$

2) la asociatividad de adición:

$$(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c});$$

3) la conmutatividad de adición:

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}.$$

La suma de dos vectores no colineales $\boldsymbol{a}$ y $\boldsymbol{b}$ se puede construir mediante la regla llamada *regla del paralelógramo*. Tracemos desde el punto A los vectores $\overrightarrow{AB} = \boldsymbol{b}$ y $\overrightarrow{AD} = \boldsymbol{a}$ (fig. 5.13). Tracemos por el extremo del vector $\boldsymbol{a}$ una recta paralela al vector $\boldsymbol{b}$ y por el extremo del vector $\boldsymbol{b}$, una recta paralela al vector $\boldsymbol{a}$. Sea C un punto de intersección de las rectas construidas. El vector $\overrightarrow{AC} = \boldsymbol{c}$ es la suma de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$:

$$\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}.$$

El vector que se designa $-\boldsymbol{a}$, tal, que la suma del vector $\boldsymbol{a}$ y del vector $-\boldsymbol{a}$ es igual al vector nulo:

$$\boldsymbol{a} + (-\boldsymbol{a}) = 0,$$

se llama vector *opuesto* al vector $\boldsymbol{a}$.

El vector opuesto al vector $\overrightarrow{AB}$ se designan también $\overrightarrow{BA}$:

$$\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \mathbf{0}.$$

Los vectores opuestos no nulos tienen las longitudes iguales ($|\boldsymbol{a}| = |-\boldsymbol{a}|$) y las direcciones contrarias.

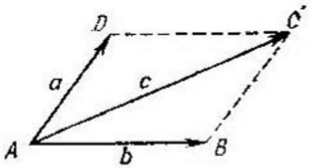


Fig. 5.13.

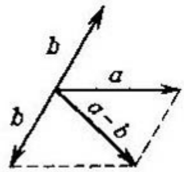


Fig. 5.14.

La diferencia de dos vectores $\boldsymbol{a} - \boldsymbol{b}$ es la suma del vector $\boldsymbol{a}$ y del vector opuesto al vector $\boldsymbol{b}$ (fig. 5.14), es decir, es el vector

$$\boldsymbol{a} + (-\boldsymbol{b}).$$

La diferencia de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ se designa $\boldsymbol{a} - \boldsymbol{b}$.

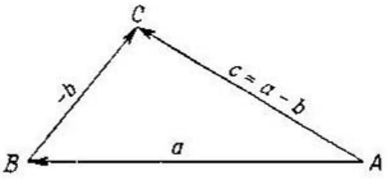


Fig. 5.15.

La diferencia de dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ puede ser obtenida mediante una *regla del triángulo*. Tracemos del punto A un vector $\boldsymbol{a}$ (fig. 5.15): $\overrightarrow{AB} = \boldsymbol{a}$; del extremo del vector $\overrightarrow{AB}$ tracemos el vector $\overrightarrow{BC} = -\boldsymbol{b}$; el vector $\overrightarrow{AC} = \boldsymbol{c}$ es la diferencia de vectores $\boldsymbol{a}$ y $\boldsymbol{b}$:

$$\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}.$$

Un vector que tiene la dirección del vector $\boldsymbol{a}$, si λ es positivo, y la dirección contraria, si λ es negativo, se llama *producto* de un vector no nulo $\boldsymbol{a}$ por el número λ. La longitud de este vector es igual al producto de la longitud del vector $\boldsymbol{a}$ por el valor absoluto (módulo) del número λ.

El producto del vector $\boldsymbol{a}$ por el número λ se designa por $\lambda\boldsymbol{a}$.

En el caso, si $\boldsymbol{a} = \boldsymbol{0}$ ó $\lambda = O$, respectivamente se consideran

$$\lambda \cdot \boldsymbol{0} = \boldsymbol{0} \text{ para cualquier } \lambda,$$

$$O \cdot \boldsymbol{a} = O \text{ para cualquier } \boldsymbol{a}.$$

Las propiedades de la operación de multiplicación del vector por el número son:

1) la conmutatividad:

$$\lambda\boldsymbol{a} = \boldsymbol{a}\lambda;$$

2) la asociatividad:

$$\lambda(\mu\boldsymbol{a}) = (\lambda\mu)\boldsymbol{a};$$

3) la distributividad respecto a la adición de vectores:

$$\lambda(\boldsymbol{a} + \boldsymbol{b}) = \lambda\boldsymbol{a} + \lambda\boldsymbol{b};$$

4) la distributividad respecto a la adición de números:

$$(\lambda + \mu)\boldsymbol{a} = \lambda\boldsymbol{a} + \mu\boldsymbol{a}.$$

Dos vectores no nulos se llaman *colineales*, si sus direcciones coinciden o son opuestas.

Un vector nulo se considera, según la definición, colineal a cualquier vector.

La condición necesaria y suficiente de colinealidad de los vectores no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$ es la existencia del número λ que satisface a la igualdad $\boldsymbol{b} = \lambda\boldsymbol{a}$.

La condición necesaria y suficiente de colinealidad de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ en el plano, dados por sus coordenadas respecto al sistema de coordenadas cartesianas rectangulares Oxy:

$$\boldsymbol{a} = (x_1;\ y_1), \qquad \boldsymbol{b} = (x_2;\ y_2),$$

es la igualdad igual a cero del determinante de segundo orden:

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = 0.$$

La condición necesaria y suficiente de colinealidad de dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ en el espacio, dados por sus coordenadas respecto al sistema de coordenadas cartesiano rectangular $Oxyz$:

$$\boldsymbol{a} = (x_1;\ y_1;\ z_1), \qquad \boldsymbol{b} = (x_2;\ y_2;\ z_2),$$

es la igualdad igual a cero de todos los determinantes sucesivos de segundo orden:

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}, \qquad \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}, \qquad \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}.$$

Tres vectores no nulos se llaman *coplanares*, si los rayos que representan sus direcciones están situados en las rectas paralelas a cierto plano. Si entre los tres vectores existe por lo menos un vector nulo, entonces tales vectores también se consideran coplanares.

La condición necesaria y suficiente para el carácter coplanar de tres vectores $\boldsymbol{a}$, $\boldsymbol{b}$ y $\boldsymbol{c}$ dados por sus coordenadas en el sistema de coordenadas cartesianas rectangulares $Oxyz$:

$$\boldsymbol{a} = (x_1;\ y_1;\ z_1), \qquad \boldsymbol{b} = (x_2;\ y_2;\ z_2),$$
$$\boldsymbol{c} = (x_3;\ y_3;\ z_3),$$

es la igualdad igual a cero del determinante de tercer orden:

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = 0.$$

Si los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ no son colineales, entonces cualquier vector $\boldsymbol{c}$, coplanar con los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$, puede representarse de una sola forma

$$\boldsymbol{c} = x\boldsymbol{a} + y\boldsymbol{b}. \tag{1}$$

(Por unicidad de representación se entiende que existe un solo par ordenado de números $(x;\ y)$, para el cual se cumple la igualdad (1).)

La representación del vector $\boldsymbol{c}$, coplanar con los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$, en forma de la suma $\boldsymbol{c} = x\boldsymbol{a} + y\boldsymbol{b}$ se llama *desarrollo* del vector $\boldsymbol{c}$ en vectores $\boldsymbol{a}$ y $\boldsymbol{b}$.

Regla del paralelepípedo. Para hallar un vector, que es la suma de tres vectores no coplanares $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, se suele utilizar muy a menudo una regla que se llama *regla del paralelepípedo*.

Desde un punto arbitrario del espacio O se trazan los vectores $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$ y $\overrightarrow{OC} = \boldsymbol{c}$. Se construye un paralelepípedo de tal manera que los segmentos OA, OB, OC sean sus aristas (fig. 5.16). El vector $\overrightarrow{OS}$, donde OS es la diagonal del paralelepípedo, es la suma de tres vectores dados:

$$\overrightarrow{OS} = \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c}.$$

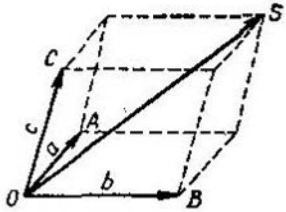


Fig. 5.16.

Desarrollo del vector en tres vectores no coplanares. Sea $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ tres vectores no coplanares (no nulos). Entonces cualquier vector espacial $\boldsymbol{d}$ puede ser representado de una sola manera en forma de suma:

$$\boldsymbol{d} = x\boldsymbol{a} + y\boldsymbol{b} + z\boldsymbol{c}. \qquad (2)$$

(Aquí por unicidad de representación se entiende que existe una sola terna ordenada de números $(x; y; z)$, para la cual se cumple la igualdad (2).)

2.2. Ángulo entre vectores. Producto escalar de vectores. Un ángulo entre dos vectores no nulos se llama *ángulo* entre las direcciones de estos vectores. El ángulo entre los dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ se designa $(\widehat{\boldsymbol{a}, \boldsymbol{b}})$.

Si el ángulo entre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es igual a 90°, entonces estos vectores se llaman *perpendiculares* y se escribe $\boldsymbol{a} \perp \boldsymbol{b}$.

Un número igual al producto de las longitudes de los vectores, que se designan $(\boldsymbol{a}, \boldsymbol{b})$ (o $\boldsymbol{a} \cdot \boldsymbol{b}$), por el coseno del ángulo entre estos vectores se llama *producto escalar* de dos vectores no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$:

$$(\widehat{\boldsymbol{a}, \boldsymbol{b}}) = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cos \varphi,$$

donde $\varphi = (\widehat{\boldsymbol{a}, \boldsymbol{b}})$. Si, por lo menos, uno de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es un vector nulo, entonces el producto escalar de estos dos vectores se considera igual a cero. Del producto escalar $(\boldsymbol{a}, \boldsymbol{b})$ se dice que está obtenido como resultado de la multiplicación escalar del vector $\boldsymbol{a}$ por el vector $\boldsymbol{b}$.

Las propiedades de la multiplicación escalar de vectores es:

1) la conmutatividad:

$$(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a});$$

2) la asociatividad respecto a la multiplicación por un número:

$$((k\boldsymbol{a}), \boldsymbol{b}) = k(\boldsymbol{a}, \boldsymbol{b}).$$

Las operaciones de adición y multiplicación escalar de vectores se relacionan por la ley distributiva de multiplicación respecto a la adición:

$$(\boldsymbol{a}, (\boldsymbol{b} + \boldsymbol{c})) = (\boldsymbol{a}, \boldsymbol{b}) + (\boldsymbol{a}, \boldsymbol{c}).$$

La condición necesaria y suficiente de perpendicularidad de dos vectores no nulos es la igualdad igual a cero de su producto escalar.

El producto escalar del vector por este mismo vector es igual al cuadrado del valor numérico de su longitud. El producto $\boldsymbol{a}\cdot\boldsymbol{a}$ se escribe como $(\boldsymbol{a})^2$ y se llama *cuadrado escalar* del vector $\boldsymbol{a}$.

2.3. Coordenadas de un vector en el plano. Sea Oxy un sistema de coordenadas cartesianas rectangulares con los vectores básicos $(\boldsymbol{i}; \boldsymbol{j})$. Cualquier vector $\boldsymbol{a}$, perteneciente al plano, puede ser descompuesto por los vectores de base $\boldsymbol{i}, \boldsymbol{j}$, es decir, para cualquier vector $\boldsymbol{a}$ existe, y además sólo uno, un par ordenado de números $(x_0; y_0)$ tal que

$$\boldsymbol{a} = x_0\boldsymbol{i} + y_0\boldsymbol{j}.$$

Los números x_0, y_0 se llaman *coordenadas* del vector $\boldsymbol{a}$ en la base $(\boldsymbol{i}; \boldsymbol{j})$ y se escribe

$$\boldsymbol{a} = (x_0; y_0).$$

Si en el sistema de coordenadas rectangulares los puntos A y B tienen las coordenadas $(x_1; y_1)$ y $(x_2; y_2)$, entonces las coordenadas del vector $\overrightarrow{AB}$ serán un par ordenado de números $(x_2 - x_1; y_2 - y_1)$, es decir,

$$\overrightarrow{AB} = (x_2 - x_1; y_2 - y_1).$$

Reglas de operaciones con los vectores, determinados por sus coordenadas. Supongamos que en la base ($\boldsymbol{i}$; $\boldsymbol{j}$) están dados los vectores

$$\boldsymbol{a} = (x_1;\ y_1), \quad \boldsymbol{b} = (x_2;\ y_2).$$

1) Las coordenadas de la suma de dos vectores son iguales a la suma de las coordenadas respectivas de los sumandos:

$$\boldsymbol{a} + \boldsymbol{b} = (x_1 + x_2;\ y_1 + y_2).$$

2) Las coordenadas de la diferencia de dos vectores son iguales a la diferencia de las coordenadas respectivas de estos vectores:

$$\boldsymbol{a} - \boldsymbol{b} = (x_1 - x_2;\ y_1 - y_2).$$

3) Las coordenadas del producto del vector por un número son iguales al producto de las coordenadas respectivas del vector dado por este número:

$$p\boldsymbol{a} = (px_1;\ py_1).$$

4) El producto escalar de dos vectores es igual a la suma de los productos de las coordenadas respectivas de estos vectores:

$$(\boldsymbol{a},\ \boldsymbol{b}) = x_1x_2 + y_1y_2.$$

El módulo del vector $\boldsymbol{a}$, definido por sus coordenadas (x_1; y_1) se calcula mediante la fórmula

$$|\boldsymbol{a}| = \sqrt{x_1^2 + y_1^2}.$$

El ángulo φ entre los dos vectores no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$ dados por sus coordenadas (x_1; y_1) y (x_2; y_2), respectivamente, se calcula de la relación

$$\cos\varphi = \cos\widehat{(\boldsymbol{a},\ \boldsymbol{b})} = \frac{x_1x_2 + y_1y_2}{\sqrt{x_1^2 + y_1^2}\ \sqrt{x_2^2 + y_2^2}}.$$

Los cosenos de los ángulos entre el vector $\boldsymbol{a} = (x_1;\ x_2)$ y los vectores de base $\boldsymbol{i}$, $\boldsymbol{j}$ se llaman *cosenos directores* del vector $\boldsymbol{a}$ y se calculan según las fórmulas

$$\cos\alpha = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad \cos\beta = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}.$$

Los cosenos directores de cualquier vector $\boldsymbol{a}$ se relacionan por la igualdad

$$\cos^2\alpha + \cos^2\beta = 1.$$

2.4. Coordenadas de un vector en el espacio. Sea $Oxyz$ un sistema de coordenadas cartesianas rectangulares con una terna de los vectores básicos $(\boldsymbol{i};\ \boldsymbol{j};\ \boldsymbol{k})$. Cualquier vector espacial $\boldsymbol{a}$ puede ser descompuesto en vectores de base $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, es decir, para cualquier vector $\boldsymbol{a}$ existe, y además, sólo una terna de números $(x_0;\ y_0;\ z_0)$ tal que

$$\boldsymbol{a} = x_0\boldsymbol{i} + y_0\boldsymbol{j} + z_0\boldsymbol{k}.$$

Los números x_0, y_0, z_0 se llaman *coordenadas* del vector $\boldsymbol{a}$ en la base $(\boldsymbol{i};\ \boldsymbol{j};\ \boldsymbol{k})$ y se escribe

$$\boldsymbol{a} = (x_0;\ y_0;\ z_0).$$

Si en el sistema de coordenadas rectangulares los puntos A y B tienen las coordenadas $(x_1;\ y_1;\ z_1)$ y $(x_2;\ y_2;\ z_2)$, entonces las coordenadas del vector $\overrightarrow{AB}$ serán una terna ordenada de números $(x_2 - x_1;\ y_2 - y_1;\ z_2 - z_1)$, es decir,

$$\overrightarrow{AB} = (x_2 - x_1;\ y_2 - y_1;\ z_2 - z_1).$$

Reglas de operaciones con los vectores definidos por sus coordenadas. Supongamos que en la base $(\boldsymbol{i},\ \boldsymbol{j},\ \boldsymbol{k})$ están dados los vectores

$$\boldsymbol{a} = (x_1;\ y_1;\ z_1) \quad \text{y} \quad \boldsymbol{b} = (x_2;\ y_2;\ z_2).$$

1) Las coordenadas de la suma de dos vectores son iguales a la suma de las coordenadas respectivas de los sumandos:

$$\boldsymbol{a} + \boldsymbol{b} = (x_1 + x_2;\ y_1 + y_2;\ z_1 + z_2).$$

2) Las coordenadas de la diferencia de dos vectores son iguales a la diferencia de las coordenadas respectivas de estos vectores:

$$\boldsymbol{a} - \boldsymbol{b} = (x_1 - x_2;\ y_1 - y_2;\ z_1 - z_2).$$

3) Las coordenadas del producto de un vector por un número son iguales al producto de las coordenadas respectivas del vector dado por este número:

$$p\boldsymbol{a} = (px_1;\ py_1;\ pz_1).$$

4) El producto escalar de dos vectores es igual a la suma de los productos de las coordenadas respectivas de estos vectores:

$$(\boldsymbol{a},\ \boldsymbol{b}) = x_1x_2 + y_1y_2 + z_1z_2.$$

El módulo del vector $\boldsymbol{a}$ determinado por sus coordenadas $(x_1;\ y_1;\ z_1)$ respecto a cierta base rectangular $(\boldsymbol{i};\ \boldsymbol{j};\ \boldsymbol{k})$ se calcula mediante la fórmula

$$|\boldsymbol{a}| = \sqrt{x_1^2 + y_1^2 + z_1^2}.$$

El angulo φ entre los dos vectores espaciales no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$ dados por sus coordenadas $(x_1;\ y_1;\ z_1)$ y $(x_2;\ y_2;\ z_2)$ respecto a cierta base rectangular $(\boldsymbol{i};\ \boldsymbol{j};\ \boldsymbol{k})$ se calcula de la relación

$$\cos\varphi = \cos\widehat{(\boldsymbol{a},\ \boldsymbol{b})} = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2}\,\sqrt{x_2^2 + y_2^2 + z_2^2}}.$$

Los cosenos de los ángulos entre el vector $\boldsymbol{a} = (x_1;\ y_1;\ z_1)$ y los vectores de base $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ se llaman *cosenos directores* del vector $\boldsymbol{a}$ y se calculan según las fórmulas

$$\cos\alpha = \frac{x_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}},$$

$$\cos\beta = \frac{y_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}},$$

$$\cos\gamma = \frac{z_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}}.$$

Los cosenos directores de cualquier vector a se relacionan por la igualdad

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

2.5. Producto vectorial. Se llama *producto vectorial* de dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$, que se designan $[\boldsymbol{a},\ \boldsymbol{b}]$ (o $\boldsymbol{a} \times \boldsymbol{b}$), al vector que se define por las tres condiciones siguientes:

1) el módulo del vector $[\boldsymbol{a},\ \boldsymbol{b}]$ es igual a $|\boldsymbol{a}|\cdot|\boldsymbol{b}| \times$ sen φ, donde φ es el ángulo entre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$;

2) el vector $[\boldsymbol{a},\ \boldsymbol{b}]$ es perpendicular tanto al vector $\boldsymbol{a}$, como también al vector $\boldsymbol{b}$;

3) la terna ordenada de los vectores $(\boldsymbol{a};\ \boldsymbol{b};\ [\boldsymbol{a},\ \boldsymbol{b}])$ trazados desde un punto forma la base derecha (en el caso general, una base oblicuángula) (fig. 5.17).

Si por lo menos uno de los vectores $\boldsymbol{a}$ o $\boldsymbol{b}$ es un vector nulo, entonces el producto vectorial de dos vectores se considera igual al vector nulo:

$$[\boldsymbol{a}, \boldsymbol{b}] = 0.$$

El producto vectorial de dos vectores no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$ representados por sus coordenadas: $\boldsymbol{a} = (x_1;\ y_1;\ z_1)$ y $\boldsymbol{b} = (x_2;\ y_2;\ z_2)$ es el vector

$$[\boldsymbol{a}, \boldsymbol{b}] = \left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix};\ -\begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix};\ \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right);$$

donde $\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}$, $\begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}$, $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$ son los determinantes de segundo orden.

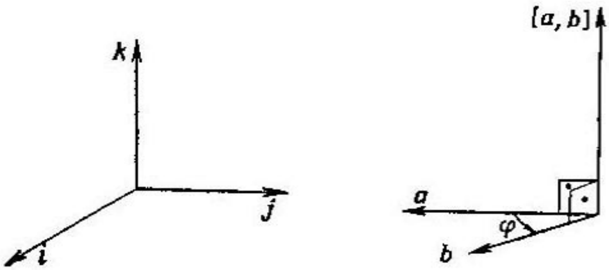


Fig. 5.17.

Las propiedades del producto vectorial:

$$[\boldsymbol{a}, \boldsymbol{b}] = -[\boldsymbol{b}, \boldsymbol{a}];$$

$$[(p\boldsymbol{a}), \boldsymbol{b}] = p\,[\boldsymbol{a}, \boldsymbol{b}] \quad (p \text{ es un número});$$

$$[(\boldsymbol{a} + \boldsymbol{b}), \boldsymbol{c}] = [\boldsymbol{a}, \boldsymbol{c}] + [\boldsymbol{b}, \boldsymbol{c}];$$

$$[\boldsymbol{a}, [\boldsymbol{b}, \boldsymbol{c}]] = \boldsymbol{b}\,(\boldsymbol{a}, \boldsymbol{c}) - \boldsymbol{c}\,(\boldsymbol{a}, \boldsymbol{b});$$

$$[[\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{c}] = \boldsymbol{b}\,(\boldsymbol{a}, \boldsymbol{c}) - \boldsymbol{a}\,(\boldsymbol{b}, \boldsymbol{c}).$$

2.6. Producto mixto (escalar-vectorial) de vectores. Sea $Oxyz$ un sistema de coordenadas cartesianas rectangulares con una base derecha $(\boldsymbol{i},\ \boldsymbol{j};\ \boldsymbol{k})$. Se llama *producto mixto* de tres vectores no nulos y no coplanares, representados por sus coordenadas respecto a la base $(\boldsymbol{i};\ \boldsymbol{j};\ \boldsymbol{k})$,

$$\boldsymbol{a} = (x_1;\ y_1;\ z_1), \qquad \boldsymbol{b} = (x_2;\ y_2;\ z_2), \qquad \boldsymbol{c} = (x_3;\ y_3;\ z_3),$$

al número, cuyo valor absoluto es igual al volumen de un paralelepípedo, las aristas del cual son los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, trazados desde un mismo punto; este número es positivo, si la terna ordenada de los vectores ($\boldsymbol{a}$; $\boldsymbol{b}$; $\boldsymbol{c}$) forma una base de mano derecha y es negativo, si ($\boldsymbol{a}$; $\boldsymbol{b}$; $\boldsymbol{c}$) forma una base de mano izquierda*) (fig. 5.18).

El producto mixto de los vectores $\boldsymbol{a}$; $\boldsymbol{b}$; $\boldsymbol{c}$ se designa $\langle \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\rangle$.

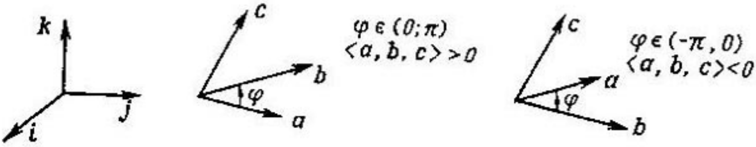


Fig. 5.18.

Si los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ son coplanares, entonces el producto mixto se considera igual a cero.

$$\langle \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\rangle = 0.$$

El producto mixto de los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, representados por sus coordenadas se calcula como un determinante de tercer orden:

$$\langle \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\rangle = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Las propiedades del producto mixto:

$$\langle \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\rangle = ([\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{c});$$

$$\langle \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\rangle = ([\boldsymbol{b}, \boldsymbol{c}], \boldsymbol{a}) = (\boldsymbol{c}, [\boldsymbol{a}, \boldsymbol{b}]) = -(\boldsymbol{b}, [\boldsymbol{a}, \boldsymbol{c}]).$$

§ 3. Principios de geometría analítica

La geometría analítica del plano es una parte de la geometría, en la cual las figuras geométricas se trazan según la determinación de la ecuación de la figura. Supongamos que en un plano se elige un sistema de coordenadas cartesianas rectangulares Oxy. La figura geométrica (llamada corrientemente *curva*) se representa por la ecuación que tiene la forma

$$F(x, y) = 0, \tag{1}$$

la cual se satisface por las coordenadas de los puntos pertenecientes a la figura dada y no se satisface por las coordenadas de ningún punto, que no pertenecen a esta figura.

*) En caso de que la base ($\mathbf{i}$; $\mathbf{j}$; $\mathbf{k}$) sea de mano izquierda, el producto mixto de los vectores $\boldsymbol{a}$; $\boldsymbol{b}$; $\boldsymbol{c}$ es positivo, si la terna de vectores ($\boldsymbol{a}$; $\boldsymbol{b}$; $\boldsymbol{c}$) forma una base de mano izquierda, y negativo, en caso contrario.

Las propiedades de una figura geométrica se aclaran mediante el estudio de las propiedades de la ecuación (1).

Las curvas simples que se estudian en la geometría analítica del plano son la recta, hipérbola, parábola y elipse (cuyo caso particular es la circunferencia).

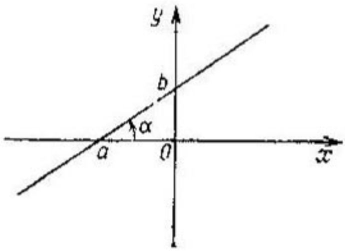


Fig. 5.19.

3.1. La recta. Cada ecuación, lineal respecto a las variables x e y, es decir, la ecuación que tiene la forma

$$Ax + By + C = 0, \quad (2)$$

donde A y B no son iguales a cero simultáneamente, representa una recta en el sistema de coordenadas cartesianas rectangulares Oxy. Y, recíprocamente, cada recta puede ser representada por una ecuación lineal que tiene la forma (2).

La ecuación (2) se llama *ecuación general de la recta*.

Ciertas formas particulares de expresión de la ecuación de la recta:

1) La recta que forma un ángulo α con dirección positiva del eje Ox e interseca el eje Oy en el punto $(0;\ b)$ se representa por la ecuación (fig. 5.19)

$$y = kx + b,$$

donde $k = \operatorname{tg} \alpha$, $\alpha \in [0;\ \pi/2) \cup (\pi/2;\ \pi)$. El coeficiente k se llama *coeficiente angular* de la recta.

La recta que forma el ángulo α con dirección positiva del eje Ox y que pasa por el punto $(x_0;\ y_0)$ se representa por la ecuación $y - y_0 = k(x - x_0)$ $(k = \operatorname{tg} \alpha)$.

2) *Ecuación segmentaria de una recta*. La recta que interseca el eje Ox en el punto $(a;\ 0)$ y el eje Oy en el punto $(0;\ b)$ (véase fig. 5.19) se representa por la ecuación

$$\frac{x}{a} + \frac{y}{b} = 1 \quad (a \neq 0,\ b \neq 0).$$

3) La ecuación de la recta que pasa por dos puntos dados (no coincidentes) $M_1(x_1;\ y_1)$ y $M_2(x_2;\ y_2)$ tiene la forma

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

Posición recíproca de tres puntos. La condición necesaria y suficiente de pertenencia de tres puntos $M_1(x_1;\ y_1)$, $M_2(x_2;\ y_2)$, $M_3(x_3;\ y_3)$ a una recta es la igualdad a cero del determinante de tercer orden:

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

Las coordenadas del punto M_0 que dividen el segmento M_1M_2 con los extremos $M_1(x_1;\ y_1)$, $M_2(x_2;\ y_2)$ respecto a λ se definen por las relaciones

$$x_0 = \frac{x_1 + \lambda x_2}{1+\lambda}, \qquad y_0 = \frac{y_1 + \lambda y_2}{1+\lambda}.$$

Distancia de un punto a una recta. La distancia d del punto $M_0(x_0;\ y_0)$ a la recta representada respecto al sistema de coordenadas rectangulares por la ecuación $Ax + By + C = 0$ se calcula mediante la fórmula

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}.$$

Posición recíproca de dos rectas en el plano. Dos rectas, representadas por sus ecuaciones generales

$$A_1x + B_1y + C_1 = 0, \qquad A_2x + B_2y + C_2 = 0,$$

se intersecan cuando y sólo cuando se diferencia del cero el determinante de segundo orden, compuesto de los coeficientes para las variables x e y:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0.$$

Las coordenadas $(x_0;\ y_0)$ del punto de intersección de dos rectas dadas se definen mediante las *fórmulas de Cramer*:

$$x_0 = \frac{\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}, \qquad y_0 = \frac{\begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}.$$

Para que dos rectas, representadas por las ecuaciones generales

$$A_1x + B_1y + C_1 = 0, \quad A_2x + B_2y + C_2 = 0,$$

sean perpendiculares entre sí, es necesario y suficiente que

$$A_1A_2 + B_1B_2 = 0.$$

Para que dos rectas, representadas en forma de ecuaciones con coeficientes angulares k_1 y k_2:

$$y = k_1x + b_1, \quad y = k_2x + b_2,$$

sean perpendiculares entre sí, es necesario y suficiente que los coeficientes angulares k_1 y k_2 estén relacionados entre sí por la igualdad

$$k_1k_2 = -1.$$

3.2. La circunferencia. Se llama *circunferencia* al conjunto de todos los puntos del plano que se encuentran a la distancia dada de un punto dado, situado en este plano.

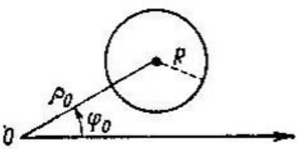


Fig. 5.20.

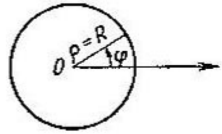


Fig. 5.21.

La ecuación de la circunferencia de radio R con centro en el origen O del sistema de coordenadas cartesianas rectangulares Oxy tiene la forma

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

La ecuación de la circunferencia de radio R con centro en el punto con las coordenadas $(a;\ b)$ en el sistema de coordenadas cartesianas rectangulares tiene la forma

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

La ecuación de la circunferencia de radio R con centro en el punto $(\rho_0;\ \varphi_0)$ en el sistema polar de coordenadas tiene la forma (fig. 5.20)

$$\rho^2 - 2\rho\rho_0 \cos(\varphi - \varphi_0) + \rho_0^2 = R^2.$$

Si el centro de la circunferencia coincide con el polo (fig. 5.21), entonces la ecuación de la circunferencia de radio R en el sistema polar de coordenadas obtiene la forma

$$\rho = R.$$

3.3. La elipse. Se llama *elipse* al conjunto de puntos del plano, para los cuales la suma de las distancias a dos puntos dados del plano F_1 y F_2 llamados *focos* es un valor constante positivo.

Para que el conjunto indicado de los puntos del plano no sea un conjunto vacío, es necesario que el valor constante indicado sea mayor que la distancia entre los focos.

Ecuación canónica de la elipse. Sea que en el plano se elige un sistema de coordenadas cartesianas rectangulares Oxy (fig. 5.22) y sea

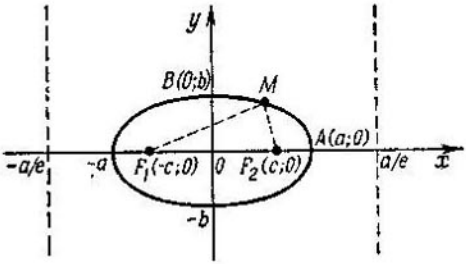


Fig. 5.22.

que los puntos F_1 y F_2 con las coordenadas $(-c;\ 0)$ y $(c;\ 0)$ son respectivamente los focos de la elipse. Si M es un punto arbitrario de la elipse, entonces, en vigor de la definición de la elipse, existe la igualdad*).

$$|F_1M| + |F_2M| = 2a.$$

Los segmentos F_1M y F_2M se llaman *radios focales* de la elipse:

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

Sustituyendo las expresiones para los radios focales en la igualdad (3), después de unas transformaciones no complicadas obtenemos la ecuación equivalente a la ecuación (3):

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

Introduciendo un valor nuevo $b = \sqrt{a^2 - c^2}$, reducimos la ecuación (4) a la forma

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{5}$$

La ecuación (5) se llama *ecuación canónica* de la elipse.

Los puntos de intersección de la elipse con sus ejes de simetría (en nuestro caso los ejes de simetría son los ejes de coordenadas Ox

*) Aquí es cómodo designar la distancia constante por $2a$.

y Oy) se llaman *vértices* de la elipse. En la fig. 5.22 los vértices de la elipse son los puntos con las coordenadas $(a; 0)$, $(-a; 0)$, $(0, b)$, $(0; -b)$. El segmento OA ($|OA| = a$) se llama *semieje* mayor de la elipse y el segmento OB ($|OB| = b$), semieje menor de la elipse. El punto O (en nuestro caso el origen del sistema de coordenadas) se llama *centro* de la elipse. La elipse es una figura simétrica central respecto al punto O.

La relación de la distancia entre los focos de la elipse a la longitud de su eje mayor se llama *excentricidad* de la elipse:

$$e = \frac{c}{a} < 1, \quad \text{o} \quad e = \sqrt{1 - \left(\frac{b}{a}\right)^2}. \tag{6}$$

La excentricidad caracteriza la forma de la elipse: cuanto más cerca es la excentricidad a la unidad, tanto menor es la relación b/a y tanto más alargada es la elipse; cuanto menor es la excentricidad, tanto más cerca a la unidad es la relación b/a y tanto más parecida es la elipse a la circunferencia.

Utilizando la definición de excentricidad de la elipse (6), la ecuación (5) y la relación $b^2 = a^2 - c^2$ se pueden obtener las expresiones de los radios focales de la elipse a través de su excentricidad

$$|F_1M| = a + ex; \qquad |F_2M| = a - ex.$$

Directrices de la elipse. Supongamos que la elipse que se representa por la ecuación canónica

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

está alargada en dirección del eje Ox, es decir, $a > b$. Dos rectas perpendiculares al eje mayor de la elipse (en nuestro caso al eje Ox) y situadas a la distancia a/e del centro de la elipse (en nuestro caso del punto O) se llaman *directrices* de la elipse (véase fig. 5.22). Las ecuaciones de las directrices de la elipse tienen la forma

$$x = -\frac{a}{e}, \qquad x = \frac{a}{e}.$$

Puesto que para la elipse $e < 1$, entonces la directriz de mano derecha se sitúa más a la derecha del vértice derecho de la elipse; análogamente, la directriz de mano izquierda se sitúa más a la izquierda de su vértice izquierdo.

Si r es la distancia de un punto arbitrario de la elipse a cierto foco, d es la distancia del mismo punto a la directriz que corresponde a este foco, entonces la relación r/d es un valor constante, igual a la excentricidad de la elipse:

$$\frac{r}{d} = e.$$

La ecuación de la recta tangente a la elipse, dada por la ecuación canónica que pasa por el punto $M_0\,(x_0; y_0)$ y que pertenece a la elipse tiene la forma

$$\frac{x_0 x}{a^2}+\frac{y_0 y}{b^2}=1.$$

3.4. La hipérbola. Se llama *hipérbola* al conjunto de puntos, para los cuales el valor absoluto de la diferencia de las distancias a dos puntos dados del plano F_1 y F_2, llamados *focos*, es un valor positivo constante.

El conjunto de puntos M que forman la hipérbola, en vigor de la definición dada de la hipérbola, satisface a la ecuación*)

$$\|F_1M| - |F_2M\| = 2a.$$

Los segmentos F_1M y F_2M se llaman *radios focales* de la hipérbola.

Ecuación canónica de la hipérbola. Sea que en un plano está elegido el sistema de coordenadas cartesianas rectangulares Oxy (fig. 5.23)

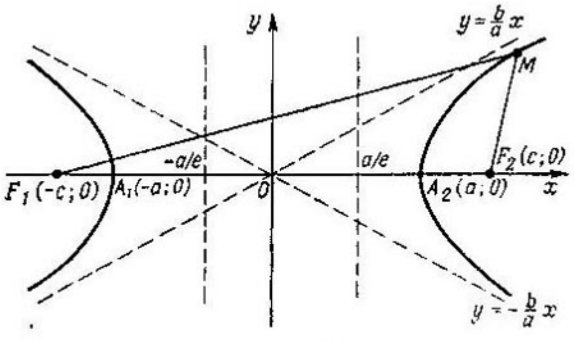


Fig. 5.23.

y sea que los puntos con las coordenadas $(-c; 0)$ y $(c; 0)$ son focos de la hipérbola. Las coordenadas de cualquier punto M, perteneciente a la hipérbola, satisfacen la ecuación

$$\left|\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}\right|=2a. \tag{7}$$

Suprimiendo la irracionalidad, la ecuación (7) se puede reducir a la forma

$$(c^2-a^2)\,x^2-a^2y^2=a^2\,(c^2-a^2). \tag{8}$$

Introduciendo en la examinación un valor nuevo $b=\sqrt{c^2-a^2}$ $(c>a)$, la ecuación (8) se puede reducir a la forma

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{9}$$

La ecuación (9) se llama *ecuación canónica* de la hipérbola.

*) Es cómodo designar el valor constante por $2a$.

Los puntos de intersección de la hipérbola con sus ejes de simetría (en nuestro caso los ejes de simetría son los ejes de coordenadas Ox y Oy) se llaman *vértices* de la hipérbola (en la fig. 5.23 los puntos A_1 y A_2 con las coordenadas $(-a;\ 0)$ y $(a;\ 0)$ respectivamente). El punto O (en nuestro caso el origen de coordenadas) se llama *centro* de la hipérbola. La hipérbola es tanto una figura central simétrica, como también una figura simétrica respecto a los ejes Ox y Oy. Las ramas de una hipérbola que se encuentran en el primer y tercer cuadrante de coordenadas (véase fig. 5.23) para $|x| \to \infty$ se aproximan a la recta $y = \frac{b}{a}x$. Las ramas de la hipérbola que se encuentran en el segundo y cuarto cuadrante de coordenadas para $|x| \to \infty$ se aproximan a la recta $y = -\frac{b}{a}x$.

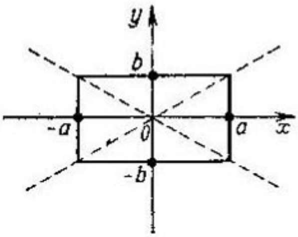


Fig. 5.24.

Un rectángulo con los lados $2a$ y $2b$, situado simétricamente respecto a los ejes de la hipérbola, se llama *rectángulo principal* de la hipérbola (fig. 5.24). En la literatura matemática también se suelen llamar *ejes* de la hiperbola a los segmentos de longitud $2a$ y $2b$ que unen los puntos medios de los lados opuestos del rectángulo principal. En vigor de esta terminología se dice que la ecuación

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

define la hipérbola con los semiejes a y b.

La relación de la distancia entre los focos de esta hipérbola y la distancia entre sus vértices se llama *excentricidad* de una hipérbola

$$e = \frac{c}{a} > 1,$$

$$e = \sqrt{1 + \left(\frac{b}{a}\right)^2}.$$

La excentricidad de una hipérbola caracteriza la forma de su rectángulo principal y, por consiguiente, también la forma de la misma hipérbola: cuanto menor es la excentricidad de la hipérbola, tanto más alargado es su rectángulo principal en dirección del eje que une los vértices.

Una hipérbola se llama *equilateral*, si los lados del rectángulo principal son iguales.

Utilizando la definición de excentricidad de una hipérbola (10), la ecuación de la hipérbola (9) y la relación $b^2 = c^2 - a^2$, se pueden obtener las expresiones de los radios focales mediante la excentricidad. Para los puntos de la rama derecha de la hipérbola $|F_1M| = a + ex$,

$|F_2M| = a - ex$. Para los puntos de la rama izquierda de la hipérbola $|F_1M| = -a - ex$, $|F_2M| = a - ex$.

Directrices de una hipérbola. Dos rectas que son perpendiculares al eje de la hipérbola que la interseca, y que se encuentran a la distancia $\frac{a}{e}$ del centro de la hipérbola, se llaman *directrices* de una hipérbola.

Las ecuaciones de las directrices en el sistema de coordenadas rectangulares, respecto al cual la hipérbola se da por la ecuación canónica

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

tienen la forma $x = -a/e$, $x = a/e$. Puesto que para la hipérbola $e > 1$, entonces la directriz de mano derecha se encuentra entre el centro y el vértice derecho de la hipérbola; análogamente, la directriz de mano izquierda se encuentra entre el centro y el vértice izquierdo (véase fig. 5.23). Las directrices de la hipérbola tienen la propiedad siguiente: si r es la distancia de un punto arbitrario de la hipérbola a cualquier foco, d, la distancia del mismo punto a la directriz, correspondiente a este foco, entonces la relación r/d es un valor constante, igual a la excentricidad de la hipérbola: $r/d = e$.

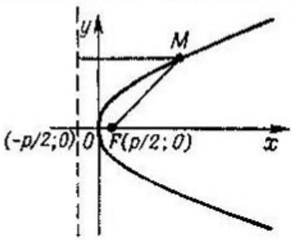


Fig. 5.25.

La ecuación de una recta tangente a la hipérbola, dada por la ecuación canónica (9) y que pasa por el punto M_0 $(x_0; y_0)$ perteneciente a la hipérbola, tiene la forma

$$\frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1.$$

3.5. La parábola. Se llama *parábola* al conjunto de los puntos de un plano, para los cuales la distancia hasta cierto punto dado F que se llama *foco* es igual a la distancia hasta cierta recta dada llamada *directriz**).

La distancia p del foco F a la directriz de la parábola se llama *parámetro* de la parábola.

El conjunto de los puntos M que forman la parábola, en vigor de la definición dada de la parábola, satisface la ecuación

$$|FM| = d.$$

Ecuación canónica de la parábola. Sea que en el plano está elegido un sistema de coordenadas cartesianas rectangulares Oxy y sea que el punto con las coordenadas $(p/2; 0)$ es el foco de la parábola, y la recta, dada por la ecuación $x = -p/2$, es la directriz de la parábola (fig. 5.25).

---

*) Al mismo tiempo se supone que el punto F no pertenece a la directriz.

Las coordenadas $(x; y)$ de cualquier punto M, perteneciente a la parábola, satisfacen la ecuación

$$\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}=x+\frac{p}{2}. \qquad (11)$$

Suprimiendo la irracionalidad, se puede reducir la ecuación (11) a la forma

$$y^2 = 2px. \qquad (12)$$

La ecuación (12) se llama *ecuación canónica* de la parábola.

El eje de simetría de la parábola (en nuestro caso el eje de simetría es el eje Ox) se llama *eje* de la parábola. El punto de intersección de la parábola con el eje de simetría se llama *vértice* de la parábola (en nuestro caso el vértice de la parábola coincide con el origen de coordenadas).

Observemos que en vigor de la definición de la parábola y de las propiedades de las directrices de la elipse e hipérbola se puede suponer condicionalmente que la parábola tiene una excentricidad igual a la unidad.

La ecuación de la recta tangente a la parábola, dada por la ecuación canónica (12) y que pasa por el punto M_0 $(x_0; y_0)$, perteneciente a la parábola, tiene la forma

$$y_0y = p\,(x + x_0).$$

CAPÍTULO 6

Geometría

El surgimiento de la geometría se refiere a la antigüedad profunda y fué acondicionado por las necesidades prácticas de la actividad humana (por la necesidad de medición de terrenos, medición de volúmenes de distintos objetos, etc.). Los conocimientos y conceptos geométricos simples ya eran conocidos en Egipto Antiguo. En este período las afirmaciones geométricas se formulaban en forma de reglas sin demostraciones. Del siglo VII antes de nuestra era hasta el I siglo de nuestra era la geometría se desarrolló aceleradamente en Grecia Antigua. En este período no sólo se acumularon los distintos conocimientos, sino que se elaboró la metodología de las demostraciones de las afirmaciones geométricas. También se realizaron los primeros intentos de formular los principales conceptos primarios (axiomas) de la geometría, de los cuales mediante razonamientos puramente lógicos se deduce un conjunto de distintas afirmaciones geométricas. El nivel de desarrollo de la geometría en la Grecia Antigua se refleja en la obra de Euclides los *Elementos*. En esta obra por primera vez fué hecho un intento de dar una construcción sistemática de la planimetría sobre la base de los conceptos geométricos indefinibles fundamentales y de axiomas (postulados). Un lugar singular en la historia de las matemáticas lo ocupa el quinto postulado de Euclides (axioma de las rectas paralelas). Durante mucho tiempo los matemáticos trataron sin éxito de deducir el quinto postulado de los demás postulados de Euclides, y sólo a mediados del siglo XIX, gracias a las investigaciones de N. I. Lobachevski, B. Riemann y Bolyai, se hizo claro que el quinto postulado no puede ser deducido de los demás, y el sistema de axiomas, propuesto por Euclides, no es el único posible.

Los *Elementos* de Euclides ejercieron una enorme influencia en el desarrollo de las matemáticas. Esta obra durante más de dos mil años ha sido no sólo un manual de geometría, sino también un punto de partida para muchísimas investigaciones matemáticas, de resultas de las cuales surgieron nuevas partes independientes de las matemáticas.

La construcción sistemática de la geometría suele realizarse según el esquema siguiente:

1) Se enumeran los conceptos geométricos fundamentales, los cuales se introducen sin definiciones.

2) Se da la formulación de los axiomas de geometría.

3) Sobre la base de los axiomas y de los conceptos geométricos fundamentales se formulan los demás conceptos geométricos y teoremas.

En calidad de conceptos fundamentales (indefinibles) de la geometría se suele tomar los objetos de los tres tipos siguientes:

1) los puntos que suelen designarse con las letras A, B, C, . . .;

2) las rectas que suelen designarse con letras a, b, c, ...;

3) los planos que suelen designarse con letras α, β, γ, ...

se llama *figura* geométrica a un conjunto cualquiera de puntos del espacio (plano). Una figura geométrica se llama plana, si todos sus puntos le pertenecen a un plano.

La intersección de dos (o varias) figuras geométricas es una figura compuesta de todos aquellos y sólo aquellos puntos, los cuales pertenecen a cada una de las figuras dadas. En la teoría de los conjuntos dos conjuntos cualesquiera tienen una intersección (la cual puede ser también un conjunto vacío). En geometría en vez de las palabras «la intersección de dos figuras es vacía» se suele decir que las figuras no se intersecan.

La unión de dos (o varias) figuras geométricas es una figura, compuesta de todos aquellos y sólo aquellos puntos, los cuales pertenecen por lo menos a una de las figuras dadas.

§ 1. Rayo. Segmento

1.1. Rayo. Sea a cierta recta, y O cierto punto de la recta a. El punto O parte el conjunto de puntos de una recta en dos conjuntos: uno que se encuentra más a la izquierda del punto O y otro que se encuentra más a la derecha del mismo punto (fig. 6.1). Estos conjuntos se llaman rayos abiertos, que salen del punto O (o con el origen en el punto O).

El conjunto de todos los puntos de la recta a que se encuentran más a la derecha (a la izquierda) del punto O, incluyendo el punto O, se llama *rayo* y se designa Oa (indican-

Fig. 6.1.

do la recta y el origen del rayo) *). Sobre el rayo Oa se dice que pertenece a la recta a.

Sea O_1h_1 y O_2h_2 dos rayos. Son posibles los siguientes casos de su posición recíproca:

1) Los rayos O_1h_1 y O_2h_2 están situados en una recta. Se llaman *codirigidos*, si uno de los rayos se contiene en el

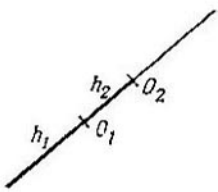


Fig. 6.2.

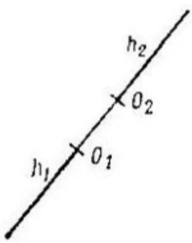


Fig. 6.3.

otro (fig. 6.2), es decir, si la intersección de ellos es un rayo y se llaman *de direcciones opuestas*, si ninguno de los rayos se contiene en el otro (fig. 6.3), es decir, si la intersección de ellos no es un rayo.

2) O_1h_1 y O_2h_2 están situados en rectas paralelas. Estas dos rectas pertenecen a cierto plano. Tracemos por los puntos O_1 y O_2 una recta que parte el plano en dos semiplanos con la frontera O_1O_2. Si ambos rayos están situados en uno de

*) Se usan también las designaciones mediante una letra latina (cuando se conoce el punto de origen) y la designación OA, donde A es un punto arbitrario de un rayo (abierto).

estos semiplanos (fig. 6.4), entonces tales rayos se llaman codirigidos y se escribe $O_1h_1 \uparrow\uparrow O_2h_2$. Si los rayos O_1h_1 y O_2h_2 están situados en semiplanos distintos, entonces se llama de direcciones opuestas (fig. 6.5) y se escribe $O_1h_1 \uparrow\downarrow O_2h_2$.

Para los rayos, pertenecientes a las rectas no paralelas, no se introduce el concepto de codirección (o de la dirección contraria).

Propiedades de los rayos codirigidos.

1) Cualquier rayo está codirigido a sí mismo (reflexividad).

2) Si el rayo O_1h_1 está codirigido al rayo O_2h_2, entonces también el rayo O_2h_2 está codirigido al rayo O_1h_1 (simetría).

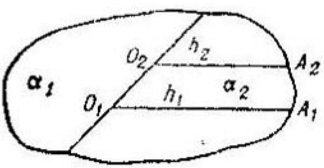


Fig. 6.4.

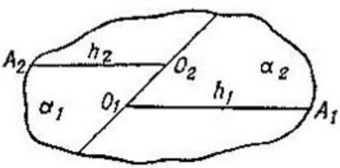


Fig. 6.5.

3) Si el rayo O_1h_1 está codirigido al rayo O_2h_2, y el rayo O_2h_2 está codirigido al rayo O_3h_3, entonces el rayo O_1h_1 está codirigido al rayo O_3h_3 (transitividad).

El conjunto de todos los rayos del plano, cada uno de los cuales está codirigido con un mismo rayo, se llama *dirección* en el plano.

El conjunto de todos los rayos del espacio, cada uno de los cuales está codirigido con un mismo rayo, se llama *dirección* en el espacio.

1.2. Segmento. Sean A y B dos puntos distintos de la recta a. Se llama *segmento* al conjunto de todos los puntos de una recta a, situados entre los puntos A y B, incluyendo los puntos A, B. Los puntos A y B se llaman *extremos* de un segmento, y todos los demás puntos, *puntos interiores* del segmento. Un segmento con los extremos A, B se designa AB.

La longitud del segmento AB se designa frecuentemente $|AB|$. Los segmentos se consideran iguales, si sus longitudes son iguales.

§ 2. Ángulos del plano

2.1. Noción de ángulo. Se llama ángulo a un par de rayos distintos Oa y Ob que salen del mismo punto O, y se designa con el símbolo $\angle(a, b)$. El punto O se llama *vértice* del ángulo, y los rayos Oa y Ob, *lados* del ángulo. Si A y B son dos puntos de los rayos Oa y Ob, entonces $\angle(a, b)$ se designa también con el símbolo $\angle AOB$ (fig. 6.6).

Al ángulo $\angle(a, b)$ se llama *llano*, si los rayos Oa y Ob que salen del mismo punto están situados en la misma recta y no coinciden (es decir, se dirigen en direcciones opuestas).

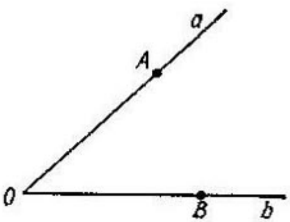


Fig. 6.6.

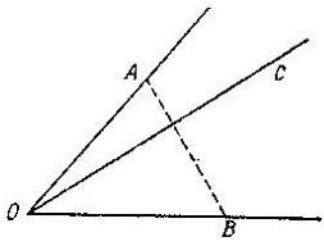


Fig. 6.7.

Junto con la definición de ángulo citada más arriba muy a menudo se usa tal definición de ángulo: se llama *ángulo* a la figura, formada por dos rayos con un origen común y limitada por los mismos por una parte del plano.

Dos ángulos se consideran *iguales*, si uno de ellos se puede superponer sobre el otro de tal manera que los lados de los ángulos coincidan. O, mejor dicho, dos ángulos se consideren iguales, si existe una isometría que traslade un ángulo en otro (sobre la isometría del plano véase p. 25.3).

Se dice que el rayo OC que sale del vértice del ángulo $\angle AOB$ está situado entre sus lados, si interseca el segmento AB (fig. 6.7). Se dice que el punto C se encuentra entre los lados del ángulo, si por este punto se puede trazar un rayo con el origen en el vértice del ángulo, situado entre los lados del ángulo. El conjunto de todos los puntos del plano, situados entre los lados del ángulo, forma una *región interior* del ángulo (fig. 6.8). La parte restante del plano se llama *región exterior* del ángulo.

El ángulo $\angle(a, b)$ se considera mayor que el ángulo $\angle(c, d)$, si el ángulo $\angle(c, d)$ se puede superponer sobre el

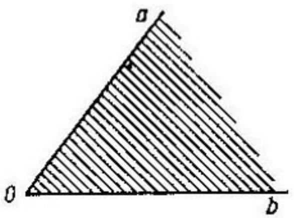


Fig. 6.8.

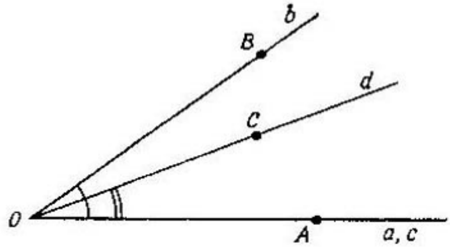


Fig. 6.9.

ángulo $\angle(a, b)$ de tal manera, que después de coincidir un par de lados, el otro lado del ángulo $\angle(c, d)$ estará situado entre los lados del ángulo $\angle(a, b)$. En la fig. 6.9 $\angle AOB$ es mayor que $\angle AOC$.

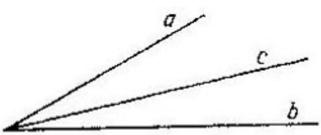


Fig. 6.10.

Sea que un rayo c está situado entre los lados del ángulo $\angle(a, b)$ (fig. 6.10). Los pares de los rayos a, c y c, b forman dos ángulos. Del ángulo $\angle(a, b)$ se dice que es la suma de dos ángulos $\angle(a, c)$ y $\angle(c, b)$ y se escribe

$$\angle(a, b) = \angle(a, c) + \angle(c, b).$$

Se llama *bisectriz* del ángulo al rayo que tiene el origen en el vértice del ángulo, el cual divide este ángulo en dos ángulos iguales.

2.2. Medición de los ángulos en grados. Para la medición por grados de los ángulos como unidad principal de medición de los ángulos («patrón de referencia» de un ángulo, con el cual se comparan los distintos ángulos) se toma el ángulo de un grado (se designa 1°). El ángulo de un grado es el que es igual a 1/180 parte del ángulo llano. El ángulo igual a 1/60 parte del ángulo de 1° es el ángulo de un minuto (se designa 1′). El ángulo igual a 1/60 parte del ángulo de un minuto es el ángulo de un segundo (se designa 1″).

2.3. Medición de los ángulos en radianes. Junto con la medición de los ángulos en grados en geometría y trigonometría se practica también la medición de los ángulos en *radianes*. Examinemos una circunferencia de radio R con centro O. Trazemos dos radios OA y OB de tal manera que la longitud del arco AB sea igual al radio de la circunferencia (fig. 6.11). El ángulo central obtenido $\angle AOB$ es un ángulo de un *radián*. El ángulo igual a un radián se toma por unidad de medición angular en radianes. En la medición angular en radianes el ángulo llano es igual a π radianes.

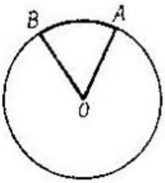


Fig. 6.11.

Las unidades de medición en grados y radianes de los ángulos están relacionadas por las igualdades:

$$1 \text{ radián} = \frac{180^\circ}{\pi} \approx 57^\circ 17' 45'';$$

$$1^\circ = \frac{\pi}{180} \text{ del radián} \approx 0{,}017453 \text{ del radián};$$

$$1' = \frac{\pi}{180 \cdot 60} \text{ del radián} \approx 0{,}000291 \text{ del radián};$$

$$1'' = \frac{\pi}{180 \cdot 60 \cdot 60} \text{ del radián} \approx 0{,}000005 \text{ del radián}.$$

La medida en grados (o en radianes) se llama también *magnitud* del ángulo. La magnitud del ángulo $\angle AOB$ a veces se designa $\widehat{AOB}$.

2.4. Clasificación de los ángulos. Un ángulo, igual a 90°, o en radianes $\pi/2$, se llama ángulo *recto*; frecuentemente se designa con la letra d. Un ángulo, menor de 90°, se llama *agudo*; un ángulo mayor de 90°, pero menor de 180°, se llama *obtuso*.

Dos ángulos que tienen un lado común, cuya suma es igual a 180°, se llaman *adyacentes*. Dos ángulos que tienen un lado común, cuya suma es igual a 90°, se llaman *complementarios*.

2.5. El ángulo entre direcciones. Si en el plano están elegidas dos direcciones, entonces en cada punto O del plano comienza un rayo OA de la primera dirección y un rayo OB de la segunda dirección. El ángulo entre los rayos OA y OB se llama ángulo entre dos direcciones en el plano. El ángulo entre las direcciones no depende de la elección de un punto inicial de los rayos.

Se llama *ángulo* entre dos direcciones en el espacio al ángulo entre dos rayos cualesquiera de estas direcciones que tienen un comienzo común.

§ 3. Paralelismo y perpendicularidad en el plano

3.1. Paralelismo en el plano. Dos rectas distintas a y b que se encuentran en un mismo plano se llaman *paralelas*, si no tienen ningún punto común.

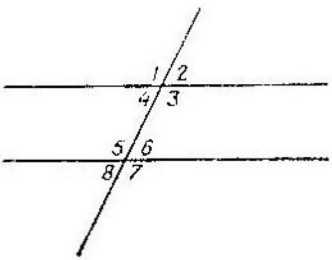


Fig. 6.12.

Cualquier recta se considera paralela a sí misma. Para la designación del paralelismo de las rectas se usa el símbolo $\parallel$.

Las propiedades de las rectas paralelas son:

1) Cualquier recta es paralela a sí misma (reflexividad).

2) Si la recta a es paralela a la recta b, entonces también la recta b es paralela a la recta a (simetría).

3) Si la recta a es paralela a la recta b, y la recta b es paralela a la recta c, entonces también la recta a es paralela a la recta c (transitividad).

El conjunto de todas las rectas paralelas en el plano se llama *haz* de rectas paralelas.

De resultas de la intersección de dos rectas por una tercera se forman ocho ángulos (fig. 6.12). Los pares de ángulos $\angle 1$, $\angle 5$; $\angle 2$, $\angle 6$; $\angle 3$, $\angle 7$; $\angle 4$, $\angle 8$ se llaman ángulos *correspondientes*; los pares $\angle 3$, $\angle 5$; $\angle 4$, $\angle 6$ se llaman ángulos *alternos internos*; los pares $\angle 1$, $\angle 7$; $\angle 2$, $\angle 8$ se llaman ángulos *alternos externos*; los pares $\angle 4$, $\angle 5$; $\angle 3$, $\angle 6$ se llaman *ángulos adyacentes*.

Criterios de paralelismo de las rectas:

1) Si al intersecarse dos rectas a y b por una tercera los ángulos correspondientes son iguales, entonces las rectas a y b son paralelas.

2) Si al intersecarse dos rectas a y b por una tercera los ángulos alternos internos (o externos) son iguales, entonces las rectas a y b son paralelas.

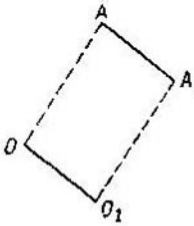


Fig. 6.13.

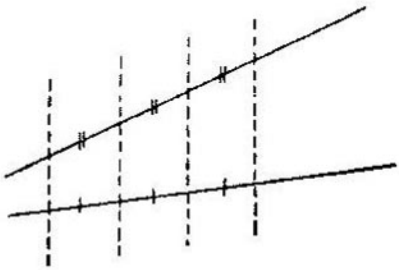


Fig. 6.14.

3) Si al intersecarse dos rectas a y b por una tercera los ángulos adyacentes suman 180°, entonces las rectas a y b son paralelas.

Teoremas sobre los segmentos iguales:

1) Si desde los puntos O y O_1 trazamos en la misma dirección segmentos iguales OA y O_1A_1, entonces los segmentos OO_1 y AA_1 son iguales y paralelos*) (fig. 6.13).

2) (*Teorema de Tales*). Si en una recta trazamos varios segmentos iguales y a través de sus extremos trazamos rectas paralelas que intersecan la segunda recta, entonces ellas cortarán en la segunda recta segmentos iguales (fig. 6.14).

Los segmentos se llaman *proporcionales*, si sus longitudes son proporcionales.

Teoremas sobre los segmentos proporcionales:

1) Las rectas paralelas que intersecan los lados del ángulo cortan en ellos segmentos proporcionales.

---

*) Los segmentos pertenecientes a las rectas paralelas, se llaman *segmentos paralelos*.

En la fig. 6.15 las rectas AA_1 y BB_1 son paralelas. En vigor del teorema formulado arriba

$$\frac{|OA_1|}{|OB_1|}=\frac{|OA|}{|OB|}.$$

2) Si los segmentos OA_1 y OB_1 son proporcionales a los segmentos OA y OB y están situados respectivamente en

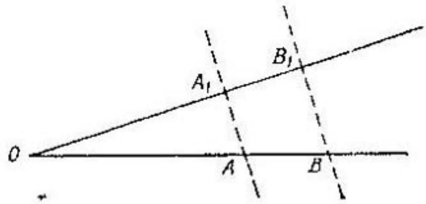


Fig. 6.15.

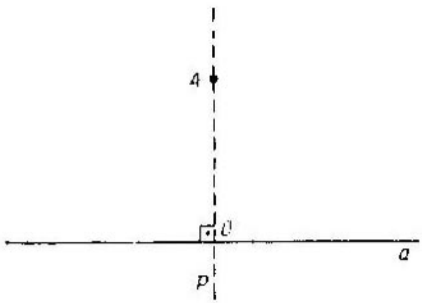


Fig. 6.16.

los rayos OA_1 y OA, entonces las rectas AA_1 y BB_1 son paralelas (véase fig. 6.15).

3.2. Perpendicularidad en el plano. Dos rectas, en cuya intersección se forman ángulos rectos, se llaman *perpendiculares* entre sí. La perpendicularidad mutua de las rectas se anota con el símbolo $\perp$:

$$a \perp b.$$

Teoremas sobre las rectas perpendiculares.

1) Si en un plano se dan un punto y una recta, entonces existe una sola recta que contiene el punto dado y es perpendicular a la recta dada.

2) Si una recta es perpendicular a una de las rectas del haz de las rectas paralelas, entonces ella es perpendicular a cualquiera otra recta de este haz.

3.3. Distancia de un punto a una recta. Sea que el punto A está situado fuera de la recta a. Construyamos una recta p perpendicular a la recta dada a y que pasa por el punto A (fig. 6.16). Designemos a través de O el punto de intersección de las rectas a y p. El punto O se llama *base* de la perpendicular p.

La *distancia del punto A a la recta a* se llama longitud del segmento OA. La distancia del punto A a cualquier punto de la recta a, distinta del punto O, es mayor que la distancia del punto A a la recta a.

§ 4. Paralelismo y perpendicularidad en el espacio

4.1. Paralelismo de la recta y el plano. Un plano α y una recta a que no pertenece por completo al plano α se llaman *paralelos*, si no tienen ningún punto común. Cualquier recta a que pertenece al plano α se considera paralela al plano α. El paralelismo del plano α y la recta a se escribe mediante el símbolo $\parallel$:

$$a \parallel \alpha \text{ o } \alpha \parallel a.$$

Criterio de paralelismo de una recta y un plano:

Si una recta es paralela a cualquiera recta situada en un plano, entonces la recta y el plano dados son paralelos.

Teoremas sobre el plano y la recta, paralela al plano:

1) Si un plano contiene una recta paralela a otro plano e interseca este plano, entonces la línea de intersección de los planos es paralela a la recta dada.

2) Si a través de cada una de dos rectas paralelas está trazado un plano arbitrario y estos planos se intersecan, entonces la línea de intersección es paralela a cada una de las rectas dadas.

4.2. Paralelismo de los planos. Dos planos α y β se llaman *paralelos*, si no tienen un punto común o coinciden.

Para la designación del paralelismo de los planos se usa el símbolo $\|$:

$$\alpha \parallel \beta.$$

Propiedades de planos paralelos:

1) Cualquier plano es paralelo a sí mismo (reflexividad).

2) Si el plano α es paralelo al plano β, entonces también el plano β es paralelo al plano α (simetría).

3) Si el plano α es paralelo al plano β, y el plano β es paralelo al plano γ, entonces el plano α es paralelo al plano γ (transitividad).

Criterio de paralelismo de dos planos:

Si dos rectas que se intersecan de un mismo plano son paralelas respectivamente a dos rectas de otro plano, entonces estos planos son paralelos.

Teoremas sobre los planos paralelos:

1) Si dos planos paralelos se intersecan por un tercero, entonces las líneas de intersección son paralelas.

2) A través de un punto dado que no pertenece al plano dado se puede trazar uno y sólo un plano paralelo al plano dado.

3) Si cada uno de dos planos dados es paralelo al tercero, entonces dos planos dados son paralelos entre sí.

4.3. Perpendicularidad de una recta y un plano. Una recta y un plano se llaman mutuamente *perpendiculares*, si la recta es perpendicular a cada recta perteneciente al plano.

Para la notación de la perpendicularidad de la recta y del plano se usa el símbolo $\perp$:

$$a \perp \alpha \text{ o } \alpha \perp a.$$

Una recta que es perpendicular al plano se llama *perpendicular* a este plano. Por cada punto dado del espacio se puede trazar una y sólo una recta perpendicular al plano dado.

Criterio de perpendicularidad de una recta y un plano:

Si una recta es perpendicular a cada una de las dos rectas que se intersecan y que se encuentran en el plano, entonces esta recta y el plano son perpendiculares entre sí.

Teoremas sobre la perpendicularidad de la recta y el plano:

1) Dos perpendiculares al plano distintas son paralelas,

2) Si una de las dos rectas paralelas es perpendicular al plano, entonces también la otra recta es perpendicular a este plano.

3) Una recta que es perpendicular a uno de las dos planos paralelos es también perpendicular a otro plano.

4) Dos planos, perpendiculares a una misma recta, son paralelos.

4.4. Distancia de un punto a un plano. Sea que el punto A_1 no pertenece al plano α. Tracemos por el punto A_1 una perpendicular al plano. Designemos el punto de intersección de la perpendicular con el plano con la letra A (fig. 6.17). El punto A es la base de la perpendicular y la longitud del segmento AA_1, la *distancia* del punto A_1 al plano α.

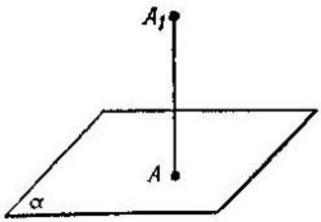


Fig. 6.17.

La distancia del punto A_1 al plano α es menor que la distancia del punto A_1 a cualquier punto del plano α, distinto del punto A.

4.5. Perpendicularidad de los planos. Dos planos se llaman *perpendiculares* entre sí, si el ángulo entre ellos es igual a 90°. Para designar la perpendicularidad mutua de dos planos se usa el símbolo $\perp$.

Criterio de perpendicularidad de los planos:

Si un plano contiene una perpendicular a otro plano, entonces aquél es perpendicular a éste.

Teorema sobre los planos perpendiculares entre sí:

Si dos planos son perpendiculares entre sí, entonces la recta, perteneciente a uno de los planos y perpendicular a la línea de intersección de los planos, es perpendicular al otro plano.

4.6. La oblicua. Una recta que interseca un plano, pero que no es perpendicular a él, se llama *oblicua* al plano.

Teorema sobre las tres perpendiculares:

Para que una recta que está situada en el plano sea perpendicular a una oblicua, es necesario y suficiente que esta recta sea perpendicular a la proyección de la oblicua sobre este plano.

4.7. Rectas cruzadas. Dos rectas se llaman cruzadas, si no se intersecan y no son paralelas. Para designar las rectas cruzadas a y b se usa la notación $a \dot{-} b$.

Criterio de rectas cruzadas:

Si una de dos rectas se encuentra en cierto plano, y la otra interseca este plano en un punto que no pertenece a la primera recta, entonces las rectas dadas se cruzan.

Se llama *perpendicular común* de dos rectas cruzadas al segmento que satisface las condiciones:

1) un extremo del segmento pertenece a una recta, y el otro, a otra;

2) la recta que contiene el segmento es perpendicular a ambas rectas cruzadas;

Se llama *distancia* entre dos rectas cruzadas a la longitud de un segmento de una recta, que es perpendicular tanto a una, como a otra recta, con los extremos en las réctas cruzadas.

§ 5. Proyección sobre un plano

5.1. Proyección paralela. Sean dados un plano α y una recta l que interseca el plano α (fig. 6.18). Tomemos un punto arbitrario del espacio A_1 y tracemos por él una recta l_1, paralela a l. La recta l_1 intersecará el plano α en cierto punto A. El punto A obtenido de tal manera se llama *proyección* del punto A_1 sobre el plano α en la proyección paralela de la recta l. En breve se suele decir que el punto A es la proyección paralela del punto A_1.

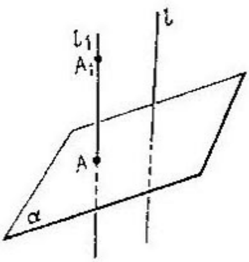


Fig. 6.18.

Se llama *proyección paralela* de la figura espacial Φ_1 al conjunto Φ de proyecciones paralelas de todos los puntos de la figura dada.

Las propiedades de la proyección paralela *) *son:*

1) La proyección de una recta es la recta.

2) Las proyecciones de las rectas paralelas son paralelas.

---

*) Aquí se supone que la proyección se efectúa paralelamente a la recta l, la cual no es paralela a las rectas o a los segmentos a proyectar.

3) La relación de las proyecciones de dos segmentos paralelos es igual a la relación de dos segmentos a proyectar.

5.2. Proyección ortogonal. La proyección ortogonal es un caso particular de la proyección paralela.

Sean dados un plano α y una recta l, perpendicular a α. Tomemos un punto arbitrario del espacio A_1 y tracemos por este punto una recta l_1, paralela a l (por consiguiente, perpendicular al plano α). La recta l_1 interseca el plano α en cierto punto A (fig. 6.19). El punto A obtenido se llama *proyección ortogonal* del punto A_1 sobre el plano α.

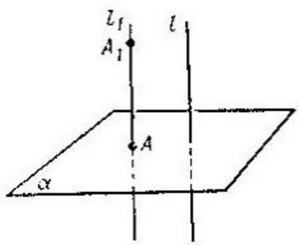


Fig. 6.19.

El conjunto Φ de proyecciones ortogonales de todos los puntos de una figura dada Φ_1 se llama *proyección ortogonal* de la figura Φ_1 sobre el plano α. Como un caso particular de la proyección paralela, la proyección ortogonal tiene todas las propiedades de la proyección paralela.

La propiedad de la proyección ortogonal de un polígono plano es:

El área de la proyección ortogonal de un polígono plano sobre el plano α es igual al área del polígono a proyectar, multiplicado por el coseno del ángulo entre el plano del polígono y el plano α.

§ 6. Angulos en el espacio

6.1. Angulo entre la oblicua y el plano. El ángulo entre la oblicua y su proyección ortogonal sobre el plano se llama *ángulo* entre la oblicua y el plano.

El ángulo entre la oblicua a y el plano α suele designarse

$$\angle (a, \alpha).$$

6.2. Angulo diedro. Sean dados dos planos no paralelos α y β. El plano α sirve de frontera común de dos semiespacios P_1 y P_2, y el plano β, de frontera para los semiespacios Q_1 y Q_2. Elijamos un semiespacio de cada par, por ejemplo, P_2 y Q_2 (fig. 6.20) y examinemos su intersección $P_2 \cap Q_2$.

Se llama *ángulo diedro* *) a la intersección de dos semiespacios, cuyas fronteras son planas no paralelas.

Cada ángulo diedro está limitado por dos semiplanos que se llaman sus *caras*. Una recta que es frontera común de dos caras del ángulo diedro se llama *arista* del ángulo diedro. Todos los puntos del ángulo diedro que no pertenecen a sus caras forman su región interior.

El ángulo diedro se designa mediante el signo $\angle$ y con las letras que indican sus caras y su arista. Además, la letra

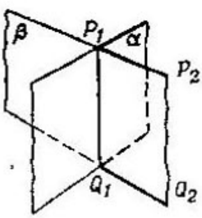


Fig. 6.20.

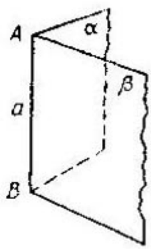


Fig. 6.21.

que designa el arista se pone entre las letras que designan sus caras. Por ejemplo, el ángulo que se muestra en la fig. 6.21 se designa $\angle \alpha a \beta$. Se usa también la designación breve del ángulo diedro por su arista, por ejemplo: $\angle a$; $\angle AB$ (fig. 6.21).

La intersección del ángulo diedro y un plano perpendicular a su arista se llama ángulo *lineal* del ángulo diedro.

El ángulo diedro se mide por su ángulo lineal. Cuantos grados, minutos y segundos tiene el ángulo lineal, tantos grados, minutos y segundos tiene el ángulo diedro.

El ángulo diedro será recto, agudo u obtuso, en función de si el ángulo lineal del ángulo diedro es recto, agudo u obtuso.

Dos ángulos diedros son iguales, si son iguales sus ángulos lineales.

6.3. Ángulo entre dos planos. Dos planos que se intersecan definen en el espacio cuatro ángulos diedros. Estos

*) El ángulo diedro se puede definir también como dos semiplanos que tienen una frontera común.

ángulos diedros son por pares iguales y la suma de dos ángulos que tienen una cara común es igual a 180°. El menor de los dos ángulos se llama *ángulo* entre estos planos. Si dos planos son paralelos, entonces el ángulo entre ellos se considera igual a 0°.

El ángulo entre los planos α, β se designa $\angle(\alpha, \beta)$.

§ 7. Quebrada. Polígono

Se llama *línea quebrada* (o simplemente *quebrada*) la unión de segmentos, en la cual el extremo de cada segmento (menos, puede ser, el último) es el comienzo del siguiente,

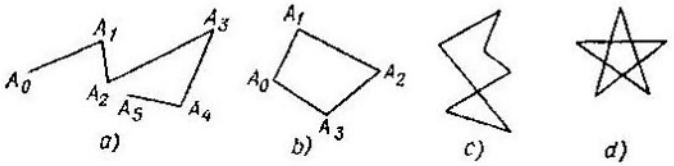


Fig. 6.22.

además, los segmentos que tienen un extremo común no están situados en una misma recta. Los segmentos que componen la quebrada se llaman *lados* de una quebrada, y los segmentos que tienen un extremo común se llaman *lados adyacentes* de una quebrada.

Una quebrada se llama *cerrada*, si el extremo de su último lado coincide con el origen del primer lado.

Una quebrada se llama *simple*, si cada uno de sus lados tiene sólo un punto común con el otro lado y este punto es el extremo del lado.

En la fig. 6.22 se muestran dos tipos de quebradas. Las quebradas que se muestran en la fig. *a*) y *b*) son simples. Sólo tales quebradas examinaremos posteriormente.

Una quebrada cerrada simple parte el plano en dos regiones, interior y exterior, respecto a la quebrada dada. La región exterior tiene la propiedad de que en ella se puede trazar una recta que pertenece por completo a esta región.

Se llama *polígono* a una quebrada cerrada simple junto con su región interior. Al mismo tiempo, la misma quebrada se llama *frontera* del polígono, y su región interior se llama *región interior* del polígono. Los lados de la frontera del polígono se llaman *lados* del mismo, y los puntos de intersección de los lados, *vértices* del polígono. El número de vértices de un polígono es igual al número de sus lados. Por lo general el polígono se designa enumerando sus vérti-

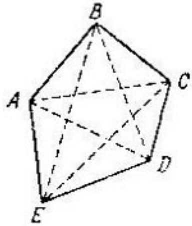


Fig. 6.23.

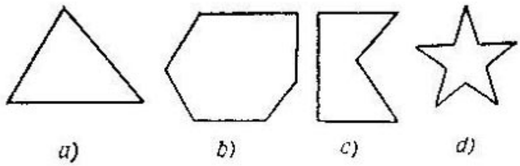


Fig. 6.24.

ces. En la fig. 6.23 en el polígono $ABCDE$ AB, BC, CD, DE y AE son los lados del polígono, A, B, C, D, E, los vértices.

Un polígono se llama *convexo*, si contiene completamente el segmento que une dos de sus puntos cualesquiera, o, con otras palabras, un polígono se llama *convexo*, si al prolongar cualquiera de sus lados todo el polígono se encuentra a un lado de esta recta. En la fig. 6.24, a, b están representados los polígonos convexos; en la fig. 6.24, c, d, los no convexos. En lo sucesivo vamos a examinar sólo los polígonos convexos.

Tracemos del vértice de un polígono dos rayos que contienen dos lados adyacentes del polígono convexo (fig. 6.25). Estos dos rayos parten el plano en dos regiones. El polígono convexo se encuentra completamente en una de las regiones. Esta región se llama *ángulo interior* (o simple *ángulo*) del polígono. El ángulo adyacente a su ángulo interior se llama *ángulo exterior* de un polígono convexo. En la fig. 6.26 se indica con un rayado uno de los ángulos exteriores del hexágono $ABCDEF$.

El polígono se denomina por el número de sus ángulos. Así, por ejemplo, un polígono con tres ángulos se llama *triángulo*, un polígono con cuatro ángulos se llama *cuadrilátero*, etc.

La suma de los ángulos interiores de un n-polígono convexo es igual a $2d\,(n-2)$ (d es el valor del ángulo recto).

La suma de los ángulos exteriores de un polígono convexo, tomados de uno en uno para cada vértice, es igual a $4d$.

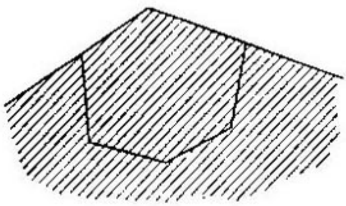


Fig. 6.25.

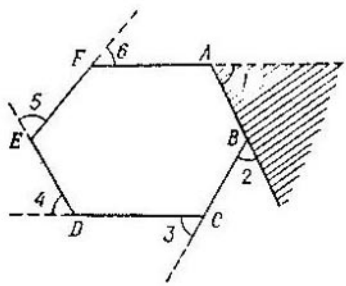


Fig. 6.26.

La suma de todos los lados de un polígono se llama *perímetro* del polígono. El perímetro suele designarse por el símbolo P_n, donde el índice n indica el número de lados del polígono.

Se llama *diagonal* del polígono al segmento que une dos vértices no vecinos del mismo. En la fig. 6.23 AC, AD, BD, BE, CE son las diagonales del pentágono $ABCDE$. El número de diagonales de un n-polígono convexo es igual a $\frac{1}{2}\,n\,(n-3)$.

Un polígono se llama *regular*, si todos sus lados son iguales y todos los ángulos interiores también son iguales entre sí.

§ 8. Triángulos

8.1. Principales propiedades. Un polígono con tres ángulos (y con tres lados) se llama *triángulo*.

Los principales elementos del triángulo son sus lados y sus ángulos.

Se dice que en el triángulo ABC (fig. 6.27) el lado a está situado frente al ángulo α y, inversa, frente al lado a se encuentra el ángulo α. Análogamente, b está situado frente a β, c, frente a γ. Un triángulo se define completamente *) por cualquier terna siguiente de sus elementos principales: o por sus tres lados, o por un lado y dos ángulos, o por dos lados y el ángulo entre ellos.

Para la existencia de un triángulo, representado por tres lados a, b, c, es necesario y suficiente el cumplimiento de las desigualdades llamadas desigualdades del triángulo:

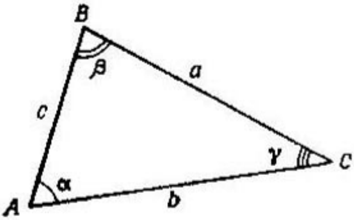


Fig. 6.27.

$$a + b > c,$$
$$a + c > b,$$
$$b + c > a.$$

Para la existencia de un triángulo, representado por el lado a y los ángulos α, β, es necesario y suficiente el cumplimiento de la desigualdad

$$\alpha + \beta < 180°,$$

y para la existencia del triángulo, representado por los lados b, c y el ángulo γ entre ellos, es necesario y suficiente el cumplimiento de la desigualdad

$$\gamma < 180°.$$

En calidad de terna de los elementos que definen unívocamente el triángulo, se pueden elegir también otras combinaciones de elementos. Algunas de tales combinaciones se dan en el § 13.

No cualquiera terna de los elementos principales del triángulo representa unívocamente el triángulo. Así, por ejemplo, dando tres ángulos del triángulo α, β, γ (ellos no son independientes y se relacionan entre sí por la igualdad $\alpha + \beta + \gamma = 180°$) , se puede construir tanto como se quiera muchos triángulos desiguales con los mismos ángulos α, β, γ (estos triángulos son semejantes).

Relaciones entre los lados y los ángulos del triángulo:

1) Frente al lado mayor se encuentra el ángulo mayor.

2) Frente al ángulo mayor se encuentra el lado mayor.

---

*) Más exactamente, la representación de estas ternas de elementos da un conjunto de triángulos iguales.

3) Frente a los lados iguales se encuentran los ángulos iguales, y, a la inversa, frente a los ángulos iguales se encuentran los lados iguales.

Relación entre los ángulos interiores y exteriores de un triángulo:

La suma de dos ángulos interiores cualesquiera del triángulo es igual a su ángulo exterior, adyacente con el tercer ángulo.

Los lados y ángulos del triángulo se relacionan entre sí mediante proporciones que se llaman teorema de los senos y teorema de los cosenos (véase capítulo 7).

Un triángulo se llama *obtusángulo*, *rectángulo* o *acutángulo*, si su ángulo interior mayor es mayor, igual o menor que 90° respectivamente.

El *área del triángulo* S puede ser calculada mediante las fórmulas:

$$S = \frac{1}{2} ah_a = \frac{1}{2} bh_b = \frac{1}{2} ch_c,$$

$$S = \sqrt{p(p-a)(p-b)(p-c)} \quad \text{(fórmula de Geron)},$$

$$S = \frac{1}{2} ab \operatorname{sen} \gamma,$$

$$S = \frac{abc}{4R},$$

$$S = pr.$$

Aquí a, b, c son los lados del triángulo, h_a, h_b, h_c, las alturas, $p = \frac{1}{2}(a+b+c)$, es el semiperímetro, R, el radio de la circunferencia circunscrita alrededor del triángulo y r, el radio de la circunferencia inscrita en el triángulo.

8.2. Medianas del triángulo. Se llama mediana de un triángulo al segmento que une un vértice de éste con el punto medio del lado opuesto (en la fig. 6.28 los segmentos AD, BE, CF son medianas del triángulo ABC). Las medianas se intersecan en un punto situado dentro del triángulo.

Propiedades principales de las medianas del triángulo:

1) Las medianas del triángulo se dividen por el punto de su intersección en relación 2 : 1 (contando desde los vértices del triángulo).

2) La mediana divide el triángulo en dos triángulos equivalentes. (Dos triángulos son equivalentes, si sus áreas son iguales.)

La *mediana del triángulo* m_a, trazada al lado a, se expresa a través de los lados del triángulo mediante la fórmula

$$m_a = \frac{1}{2}\sqrt{2b^2 + 2c^2 - a^2}.$$

8.3. Alturas del triángulo. Sea ABC cierto triángulo arbitrario. Tracemos por el vértice A una perpendicular a la

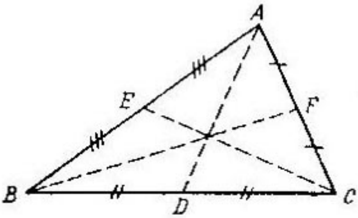


Fig. 6.28.

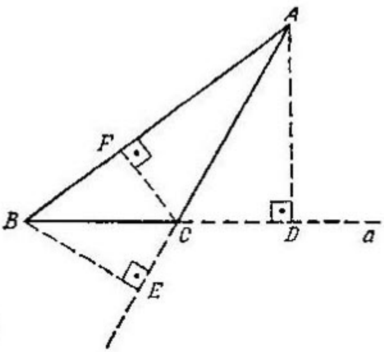


Fig. 6.29.

recta a que contiene al lado BC (fig. 6.29). Designemos la base de la perpendicular con la letra D. El segmento de la perpendicular AD se llama *altura* del triángulo ABC, bajada del vértice A al lado BC. El lado BC se llama, en este caso, *base* del triángulo ABC.

En el triángulo obtusángulo ΔABC (véase fig. 6.29) dos alturas (AD y BE) intersecan la prolongación de los lados y se encuentran fuera del triángulo; la tercer altura (CF) interseca el lado del triángulo. En el triángulo acutángulo (fig. 6.30) todas las tres alturas están situadas dentro del triángulo. En el triángulo rectángulo los catetos son también alturas. Tres rectas que contienen distintas alturas del triángulo siempre se intersecan en el mismo punto, llamado *ortocentro* del triángulo. En el triángulo obtusángulo el ortocentro se encuentra fuera de éste; en el acutángulo, se encuentra dentro; en el triángulo rectángulo el ortocentro coincide con el vértice del ángulo recto. La altura del triángulo bajada al lado del triángulo a, suele designarse h_a.

La *altura del triángulo* h_a se expresa a través de los lados mediante la fórmula

$$h_a = \frac{2\sqrt{p(p-a)(p-b)(p-c)}}{a},$$

donde $p = \frac{1}{2}(a+b+c)$.

8.4. Bisectrices del triángulo. El segmento de la bisectriz del ángulo interior del triángulo desde su vértice hasta

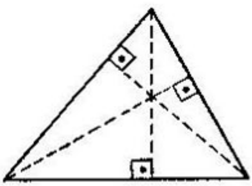


Fig. 6.30.

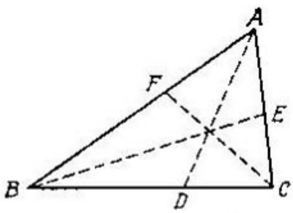


Fig. 6.31.

el punto de intersección con el lado opuesto se llama *bisectriz del triángulo*.

Tres bisectrices del triángulo (AD, BE, CF en la figura 6.31) se intersecan en un mismo punto situado dentro del triángulo y es el centro de la circunferencia inscrita en el triángulo.

Propiedades de la bisectriz del ángulo de un triángulo:

1) La bisectriz divide el lado opuesto en partes proporcionales a los lados adyacentes a ella. Así, por ejemplo, para el triángulo ABC que se muestra en la fig. 6.31

$$\frac{|AE|}{|EC|} = \frac{|AB|}{|BC|}.$$

2) La bisectriz del triángulo divide el área de éste en una relación que es proporcional a los lados adyacentes:

$$\frac{S_{\triangle ABE}}{S_{\triangle BCE}} = \frac{|AB|}{|BC|}.$$

8.5. Línea media del triángulo. Se llama línea media de un triángulo al segmento que une los puntos medios de dos lados de un triángulo (en la fig. 6.32 DE, DF, FE son las líneas medias).

Propiedades de la línea media del triángulo:

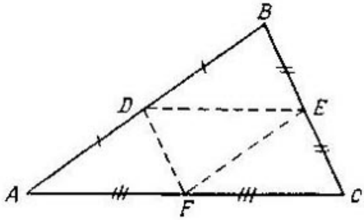


Fig. 6.32.

1) La recta que contiene la línea media de un triángulo es paralela a la recta que contiene el tercer lado del triángulo.

2) La línea media de un triángulo es igual a la mitad del tercer lado.

3) La línea media del triángulo corta de éste un triángulo semejante. El área del triángulo cortado se relaciona al área del triángulo principal en proporción de 1 : 4.

8.6. Triángulo isósceles. El triángulo con dos lados iguales se llama *isósceles*.

En el triángulo isósceles se suele tomar por base el lado que no es igual a ninguno de los otros dos lados.

Propiedades del triángulo isósceles:

1) En el triángulo isósceles los ángulos de la base del triángulo son iguales.

2) La altura, trazada del vértice, es también bisectriz y mediana.

8.7. Triángulo equilátero. Se llama *triángulo equilátero* (o *regular*) el que tiene todos los lados iguales.

Propiedades del triángulo regular:

1) En el triángulo equilátero todos los ángulos son iguales (cada ángulo es igual a 60°).

2) Cada una de las tres alturas del triángulo equilátero es también bisectriz y mediana.

Además, el triángulo equilátero, como caso particular del polígono regular, tiene todas sus propiedades.

8.8. Triángulo rectángulo. Se llama *triángulo rectángulo* el que tiene un ángulo recto.

El lado del triángulo rectángulo opuesto al ángulo recto se llama *hipotenusa*, y los otros dos lados se llaman *catetos*.

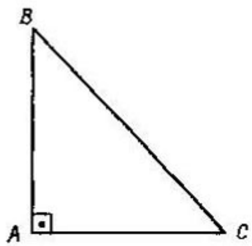


Fig. 6.33.

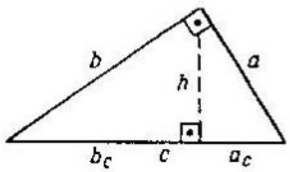


Fig. 6.34.

En la fig. 6.33 está representado un triángulo rectángulo: $\widehat{BAC} = 90°$, BC es la hipotenusa, AB y AC, los catetos.

El triángulo rectángulo que tiene los catetos iguales se llama triángulo rectángulo *isósceles*. En el triángulo rectángulo isósceles los ángulos agudos son iguales (cada uno de ellos es igual a 45°).

Los lados de cualquier triángulo rectángulo a, b y c (c es la hipotenusa) se relacionan entre sí por la proporción denominada *teorema de Pitágoras*:

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

el cual se lee del modo siguiente: el cuadrado de la hipotenusa es igual a la suma de los cuadrados de los catetos.

Propiedades del triángulo rectángulo:

1) El cateto del triángulo rectángulo es la media proporcional entre la hipotenusa y la proyección de este cateto sobre la hipotenusa (fig. 6.34):

$$b_c : b = b : c, \; a_c : a = a : c, \text{ o } b^2 = b_c c \text{ y } a^2 = a_c c.$$

2) La altura del triángulo rectángulo, trazada del vértice del ángulo recto, es la media proporcional entre las

proyecciones de los catetos sobre la hipotenusa:

$$b_c : h = h : a_c, \text{ o } h^2 = a_c b_c.$$

El área del triángulo rectángulo puede ser calculada por las fórmulas generales de cálculo del área de un triángulo. Además, el *área del triángulo rectángulo* puede ser calculada mediante la fórmula

$$S = \frac{1}{2} ab,$$

es decir, el área del triángulo rectángulo es igual a la mitad del producto de sus catetos.

§ 9. Cuadriláteros

9.1. Paralelogramo. Se llama *paralelogramo* (véase fig. 6.35) al cuadrilátero de lados opuestos por pares paralelos.

Las propiedades del paralelogramo son:

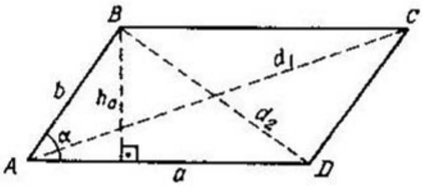


Fig. 6.35.

1) El punto medio de la diagonal del paralelogramo es su centro de simetría.

2) Los lados opuestos del paralelogramo son iguales.

3) Los ángulos opuestos del paralelogramo son iguales.

4) Cada diagonal del paralelogramo lo divide en dos triángulos iguales.

5) Las diagonales del paralelogramo se dividen por el punto de intersección por la mitad.

6) La suma de los cuadrados de las diagonales del paralelogramo (d_1 y d_2) son iguales a la suma de los cuadrados de todos sus lados:

$$d_1^2 + d_2^2 = 2(a^2 + b^2).$$

Criterios de paralelogramo:

1) Si en un cuadrilátero los lados opuestos son iguales por pares, entonces este cuadrilátero es un paralelogramo.

2) Si en un cuadrilátero dos lados opuestos son iguales y paralelos, entonces este cuadrilátero es un paralelogramo.

Cada uno de los criterios de paralelogramo puede ser tomado como definición del paralelogramo. Así, por ejemplo, del primer criterio se obtiene la definición siguiente del paralelogramo:

Un cuadrilátero, que tiene los lados opuestos iguales por pares, se llama *paralelogramo*.

Puesto que el paralelogramo representa un cuadrilátero con propiedades especiales, entonces tiene todas las propiedades de un cuadrilátero arbitrario. En particular, la suma de los ángulos interiores del paralelogramo es igual a $4d$ (360°).

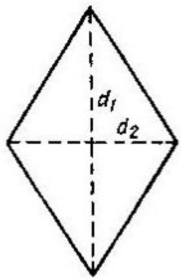


Fig. 6.36.

Se llama *altura* del paralelogramo al segmento de la perpendicular a los lados del paralelogramo, comprendido entre ellos.

Área del paralelogramo:

1) El área del paralelogramo es igual al producto de su base por la altura (véase fig. 6.35):

$$S = ah_a.$$

2) El área del paralelogramo es igual al producto de los lados adyacentes del paralelogramo por el seno del ángulo entre ellos (véase fig. 6.35):

$$S = ab \operatorname{sen} \alpha.$$

9.2. Rombo. Se llama *rombo* (fig. 6.36) al paralelogramo que tiene todos sus lados iguales. Puesto que el rombo es un caso particular del paralelogramo, entonces tendrá todas sus propiedades. Además, el rombo tiene las siguientes *propiedades* especiales:

1) La recta que contiene la diagonal del rombo en su eje de simetría.

2) Las diagonales del rombo son perpendiculares entre sí.

3) Las diagonales del rombo son bisectrices de sus ángulos interiores.

Además de las fórmulas generales de cálculo del área del rombo, como es el área del paralelogramo, el *área del rombo* puede ser calculada mediante la fórmula

$$S = \frac{1}{2} d_1 d_2,$$

donde d_1 y d_2 son las diagonales del rombo.

9.3. Rectángulo. Se llama *rectángulo* (fig. 6.37) al paralelogramo que tiene todos los ángulos rectos.

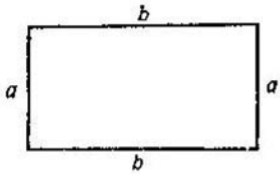


Fig. 6.37.

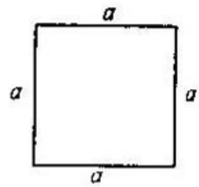


Fig. 6.38.

Puesto que el rectángulo es un paralelogramo, entonces tiene todas sus propiedades. Además, el rectángulo tiene las siguientes *propiedades* especiales:

1) La perpendicular que pasa por los puntos medios de los lados opuestos del rectángulo es su eje de simetría.

2) El rectángulo tiene dos ejes de simetría.

3) Las diagonales del rectángulo son iguales.

El *área del rectángulo* se calcula por la fórmula

$$S = ab,$$

donde a y b son los lados adyacentes del rectángulo.

9.4. Cuadrado. Se llama *cuadrado* al rectángulo que tiene todos sus lados iguales.

De las definiciones del cuadrado y del rombo se desprende que el cuadrado (fig. 6.38) es un rombo que tiene todos sus ángulos rectos. Puesto que el cuadrado es también un paralelogramo, así como rectángulo y rombo, entonces tendrá todas sus propiedades.

El área del cuadrado se calcula por la fórmula

$$S = a^2,$$

donde a es un lado del cuadrado.

9.5. Trapecio. Se llama *trapecio* al cuadrilátero que tiene paralelos solamente dos de sus lados.

Los lados paralelos se llaman *bases* del trapecio, y los no paralelos, sus *lados laterales* (en la fig. 6.39 BC y AD son las bases, AB y CD, los lados laterales).

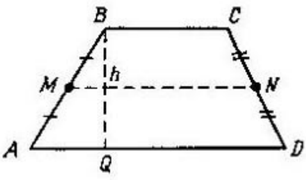


Fig. 6.39.

Se llama *altura* del trapecio (en la fig. 6.39 el segmento BQ es la altura) al segmento de la perpendicular a las bases del trapecio, comprendido entre las bases. El trapecio, cuyos lados laterales son iguales ($|AB| = |CD|$), se llama *isósceles*. En el trapecio isósceles los ángulos de la base son iguales por pares:

$$\angle BAD = \angle ADC,$$
$$\angle ABC = \angle BCD.$$

El segmento que une los puntos medios de los lados no paralelos del trapecio se llama línea media del trapecio. La línea media del trapecio es paralela a sus bases y es igual a la semisuma de la longitud de las dos bases:

$$|MN| = \frac{|BC| + |AD|}{2}.$$

La línea media del trapecio divide la altura del trapecio en dos segmentos iguales.

El *área del trapecio* se calcula mediante la fórmula

$$S = \frac{a+b}{2}\, h,$$

donde a y b son las bases del trapecio y h, la altura.

§ 10. Polígonos semejantes

10.1. Criterio de semejanza de los polígonos. Si los lados de un polígono son proporcionales a los lados de otro polígono y los ángulos correspondientes (es decir, los ángulos que están situados entre los lados proporcionales) de estos polígonos son iguales, entonces tales polígonos son semejantes.

En la fig. 6.40 está representado un pentágono $ABCDE$, semejante al pentágono $A_1B_1C_1D_1E_1$, con el coeficiente de semejanza $k = 2$:

$$\frac{|AB|}{|A_1B_1|} = \frac{|BC|}{|B_1C_1|} = \frac{|CD|}{|C_1D_1|} = \frac{|DE|}{|D_1E_1|} = \frac{|AE|}{|A_1E_1|} = k = 2.$$

Dos cuadriláteros son semejantes, si tres pares de sus lados correspondientes son proporcionales y los pares de los

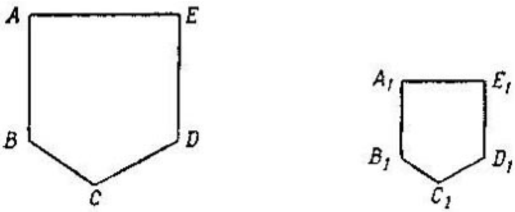


Fig. 6.40.

ángulos correspondientes, comprendidos entre los lados correspondientes, son iguales (fig. 6.41):

$$\frac{|AB|}{|A_1B_1|} = \frac{|BC|}{|B_1C_1|} = \frac{|CD|}{|C_1D_1|}, \qquad \angle B = \angle B_1, \qquad \angle C = \angle C_1.$$

10.2. Criterios de semejanza de los triángulos.

1) Si tres lados de un triángulo son proporcionales a los tres lados de otro triángulo, entonces tales triángulos son semejantes (fig. 6.42):

$$\frac{|AB|}{|A_1B_1|} = \frac{|BC|}{|B_1C_1|} = \frac{|AC|}{|A_1C_1|}.$$

2) Si dos ángulos de un triángulo son iguales a dos ángulos de otro triángulo, entonces tales triángulos son semejantes (fig. 6.43):

$$\angle A = \angle A_1, \quad \angle B = \angle B_1.$$

3) Si dos lados cualesquiera de un triángulo son proporcionales a dos lados de otro triángulo, y los ángulos entre estos lados son iguales, entonces tales triángulos son semejantes (fig. 6.44):

$$\frac{|AB|}{|A_1B_1|} = \frac{|BC|}{|B_1C_1|}, \qquad \angle B = \angle B_1.$$

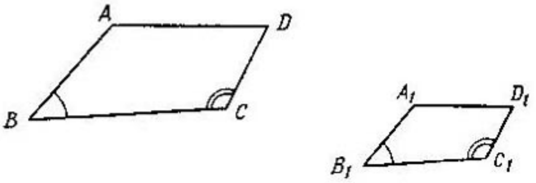


Fig. 6.41.

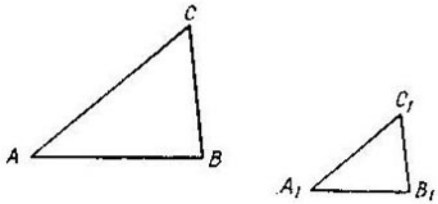


Fig. 6.42.

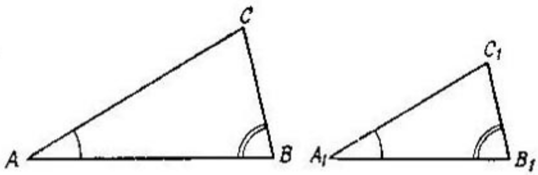


Fig. 6.43.

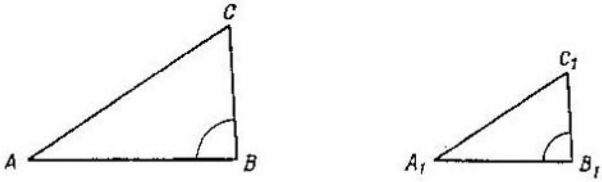


Fig. 6.44.

Para la semejanza de los triángulos de tipo especial (rectangulares, isósceles, equilaterales) es necesario cumplir menor cantidad de condiciones que para los triángulos arbitrarios. Así, a) los triángulos rectangulares son semejantes, si la hipotenusa y el cateto de un triángulo son proporcionales a la hipotenusa y al cateto de otro triángulo; b) los triángulos rectangulares son semejantes, si el ángulo agudo de un triángulo es igual al ángulo agudo de otro.

Una afirmación análoga es válida también para los cuadriláteros (polígonos) de tipo especial. Así, por ejemplo, dos n-polígonos regulares cualesquiera siempre son semejantes.

Propiedades de los polígonos semejantes:

1) La relación de los perímetros de los polígonos semejantes es igual a la razón de sus lados correspondientes (razón de semejanza).

2) La relación de las áreas de los polígonos semejantes es igual al cuadrado de la razón de semejanza.

§ 11. La circunferencia y el círculo

***11.1. La circunferencia y el círculo.** Se llama *circunferencia* al conjunto de todos los puntos del plano que se encuentran a una distancia positiva dada de cierto punto dado del plano que se llama *centro* de la circunferencia.

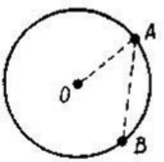


Fig. 6.45.

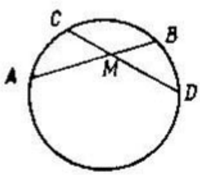


Fig. 6.46.

Se llama *radio* de la circunferencia al segmento que une el centro de ésta con cualquier punto de la circunferencia (en la fig. 6.45 el segmento OA es el radio). El radio de la circunferencia suele designarse con las letras r, R.

El segmento que une dos puntos de una circunferencia se llama *cuerda* (en la fig. 6.45 AB es la cuerda).

Una cuerda que pasa por el centro de la circunferencia se llama *diámetro*. El diámetro es igual al radio doble de la circunferencia. El diámetro suele designarse con las letras d, D.

Propiedades de las cuerdas de una circunferencia:

1) El diámetro que divide la cuerda por la mitad es perpendicular a esta cuerda.

2) En la circunferencia las cuerdas iguales se encuentran a la misma distancia del centro de la circunferencia y, a la inversa, las cuerdas son iguales, si se encuentran a la misma distancia del centro de la circunferencia.

3) De dos cuerdas no iguales de la circunferencia la que más cerca está del centro de la circunferencia es la mayor, y, a la inversa, de dos cuerdas no iguales la mayor será aquella que está más cerca del centro.

4) Entre los segmentos de las cuerdas que se intersecan AM, MB, CM, y MD (fig. 6.46) existe la relación siguiente:

$$|AM|\cdot|MB|=|CM|\cdot|MD|.$$

Se llama *círculo* al conjunto de todos los puntos del plano, cuya distancia de cierto punto dado del plano (llamado *centro* del círculo) no es mayor que la distancia dada.

El radio, la cuerda y el diámetro de una circunferencia son radio, cuerda y diámetro del círculo correspondiente.

11.2. Tangente y secante. Una recta que tiene con la circunferencia sólo un punto común y que pertenece al plano de la circunferencia se llama *tangente* a esta circunferencia.

Para que la recta sea tangente a la circunferencia, es necesario y suficiente que esta recta sea perpendicular al diametro de la circunferencia y pase por su extremo.

Por cualquier punto que se encuentra fuera de la circunferencia y que pertenece al plano de la circunferencia se pueden trazar dos tangentes distantes (fig. 6.47).

La recta que tiene con la circunferencia dos puntos comunes se llama *secante*. En la fig. 6.48 son secantes AD y AD_1.

Si a través del punto A que se encuentra fuera del círculo trazamos una tangente y una secante (véase fig. 6.48), entonces los segmentos de la tangente y de la secante se relacionan por las proporciones

$$|AB|^2 = |AD| \cdot |AC| = |AD_1| \cdot |AC_1|.$$

11.3. Posición recíproca de dos circunferencias. Sea que en un plano se dan dos puntos no coincidentes O_1 y O_2,

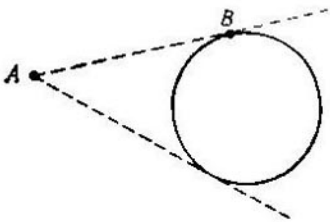


Fig. 6.47.

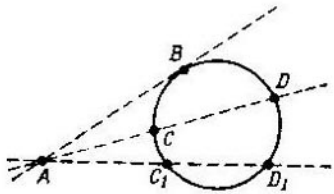


Fig. 6.48.

la distancia entre los cuales $h = |O_1O_2|$. Construyamos dos circunferencias con radios R_1 y R_2 y centros en los puntos O_1 y O_2 respectivamente. Para mayor exactitud supongamos que $R_1 \geqslant R_2$. Para la distancia dada $|O_1O_2| = h$

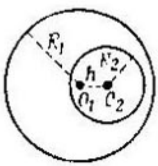


Fig. 6.49.

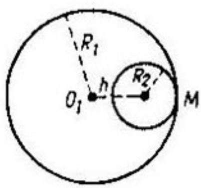


Fig. 6.50.

y los radios dados R_1 y R_2 pueden ser los siguientes casos de una posición recíproca de las circunferencias:

1) Si $h < R_1 - R_2$, entonces las circunferencias no se intersecan y el círculo de radio R_2 pertenece enteramente al círculo de radio R_1 (fig. 6.49).

2) Si $h = R_1 - R_2$, entonces el círculo de radio R_2 pertenece enteramente al círculo de radio R_1, y las circunferencias tienen un punto común M (fig. 6.50). Sobre tal posición de las circunferencias se dice que son tangentes por dentro en el punto M.

3) Si $h > R_1 + R_2$, entonces las circunferencias no se intersecan, y los círculos no tienen ningún punto común (fig. 6.51).

4) Si $h = R_1 + R_2$, entonces las circunferencias (y los círculos) tienen un punto común M (fig. 6.52). Sobre tal

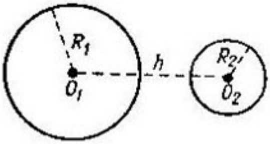


Fig. 6.51.

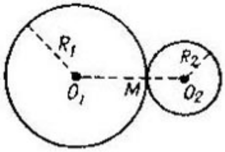


Fig. 6.52.

posición de las circunferencias se dice que son tangentes entre sí por fuera en el punto M.

5) Si $R_1 - R_2 < h < R_1 + R_2$, entonces las circunferencias tienen dos puntos comunes. En este caso se dice que las circunferencias se intersecan en los puntos M_1 y M_2 (fig. 6.53).

El segmento O_1O_2 se llama *línea de los centros* de las circunferencias. Se llama *tangente exterior* de dos circunferencias dadas a la recta que es tangente a ambas circunferencias y no interseca las líneas de los centros. En los cinco casos enumerados más arriba la tangente exterior a dos circunferencias puede ser construida en todos los casos, menos en 1).

Se llama *tangente interior* de dos circunferencias dadas a la recta que es tangente a ambas circunferencias y que interseca la línea de sus centros. La tangente interior puede ser construida en los casos 3) y 4).

Si las dos circunferencias se intersecan (véase el caso 5) fig. 6.53), entonces el segmento M_1M_2 se llama cuerda común de dos circunferencias que se intersecan. La cuerda común de dos circunferencias que se intersecan es recíprocamente perpendicular con la línea de los centros y se divide por el punto de intersección en dos segmentos iguales:

$$|M_1K| = |KM_2|.$$

11.4. Ángulos centrales y arcos de la circunferencia. Se llama *ángulo central* de una circunferencia al ángulo con el vértice en el centro de la circunferencia.

Dos rayos distintos que salen del centro de la circunferencia definen dos ángulos centrales. Los puntos de intersección de los rayos con la circunferencia (en la fig. 6.54,

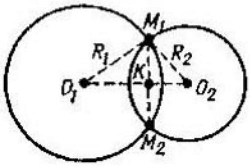


Fig. 6.53.

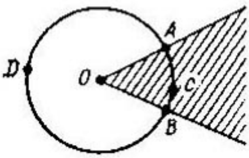


Fig. 6.54.

los puntos A y B) dividen la circunferencia en dos partes. Estas partes se llaman *arcos* de la circunferencia. Para formar uno de los arcos indicados, se escogen en los arcos puntos arbitrarios (en la fig. 6.54 son los puntos C y D) y se dice sobre los arcos ACB y ADB. Para designar los arcos se usa el símbolo $\smile$. Así, por ejemplo, los arcos ACB y ADB se designan:

$$\smile ACB \text{ y } \smile ADB.$$

Se llaman centros de los arcos el centro de la circunferencia.

Los arcos de una circunferencia se miden en grados, minutos y segundos. Cuantos grados, minutos y segundos tiene el ángulo central dado, tantos grados, minutos y segundos tiene el arco respectivo. El valor angular del arco ACB ($\smile ACB$) a veces se designa con el símbolo $\overset{\frown}{ACB}$.

El arco que corresponde al ángulo central de 180° se llama *semicircunferencia*.

Entre los arcos y los ángulos centrales de una circunferencia que les corresponden existen las relaciones siguientes:

1) Dos arcos, pertenecientes a las circunferencias de un mismo radio, son iguales cuando y sólo cuando sus valores angulares son iguales *).

---

*) Los valores angulares de los arcos de las circunferencias de radio distinto pueden ser iguales, pero los arcos, sin embargo, no serán iguales.

2) En una misma circunferencia al ángulo central le corresponde el arco mayor.

11.5. Arcos y cuerdas de la circunferencia. Sea que en la circunferencia está trazada la cuerda AB que no pasa por el centro de la circunferencia (fig. 6.55). De la cuerda AB se dice que tensa el arco AB (en este caso se supone que de dos arcos, en los cuales los puntos A y B parten la circunferencia, se elige el arco menor, es decir, el valor angular del

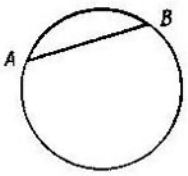


Fig. 6.55.

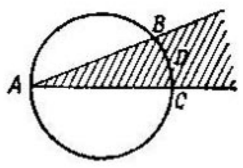


Fig. 6.56.

arco AB está comprendido en el intervalo (0°; 180°). Entre las cuerdas de la circunferencia y los arcos unidos por ellas existen las relaciones siguientes:

1) Los arcos iguales están unidos por cuerdas iguales.

2) Las cuerdas iguales unen arcos iguales.

3) El diámetro que es perpendicular a la cuerda divide por la mitad el arco que une esta cuerda.

11.6. Ángulos en la circunferencia. Un ángulo, cuyo vértice pertenece a la circunferencia, y los lados intersecan la circunferencia se llama ángulo *inscrito* en esta circunferencia (fig. 6.56). Se dice también que el ángulo BAC descansa en el arco BDC. El valor del ángulo inscrito es igual a la mitad del valor angular del arco, en el cual este ángulo descansa (o a la mitad del valor del ángulo central que corresponde al arco dado):

$$\widehat{BAC} = \frac{1}{2}\overset{\frown}{BDC}.$$

El ángulo formado por dos tangentes a la circunferencia, que pasan por el mismo punto, se llama ángulo circunscrito (en la fig. 6.57 son tangentes CA y CB).

El valor del ángulo circunscrito es igual a la semidiferencia de los valores angulares de los arcos, comprendidos entre sus lados:

$$\widehat{ACB} = \frac{1}{2}(\smile{ALB} - \smile{ADB}).$$

El valor de un ángulo, formado por dos secantes que tienen un punto común, el cual se encuentra fuera de la circunferencia, es igual a la semidiferencia de los valores angulares

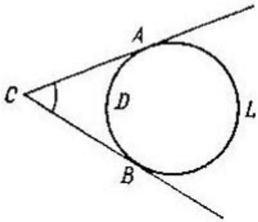


Fig. 6.57.

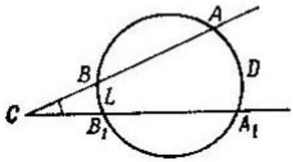


Fig. 6.58.

de los arcos, comprendidos entre sus lados (fig. 6.58):

$$\widehat{ACA_1} = \frac{1}{2}(\smile{ADA_1} - \smile{BLB_1}).$$

El valor de un ángulo, formado por dos secantes que tienen un punto común situado en el interior de la circunferencia, es igual a la semisuma de valores angulares de los arcos, comprendidos entre sus lados (fig. 6.59):

$$\widehat{BCB_1} = \frac{1}{2}(\smile{BLB_1} + \smile{ADA_1}).$$

11.7. Longitudes y áreas en la circunferencia y en el círculo. Se llama *longitud de una circunferencia* al límite de sucesión de los perímetros de los polígonos regulares, inscritos en la circunferencia dada para un aumento no limitado del número de lados. La *longitud de la circunferencia* L se calcula mediante la fórmula

$$L = \pi d,$$

donde d es el diámetro de la circunferencia, o por la fórmula

$$L = 2\pi r,$$

donde r es el radio de la circunferencia.

La *longitud del arco* de la circunferencia con un valor angular de $\alpha°$ se calcula por la fórmula

$$l = \frac{\pi r \alpha}{180}.$$

Se llama *área de un círculo* al límite de sucesión de las áreas de los polígonos regulares, inscritos en la circunferen-

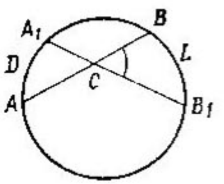


Fig. 6.59.

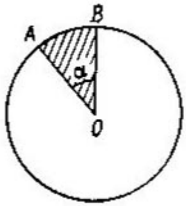


Fig. 6.60.

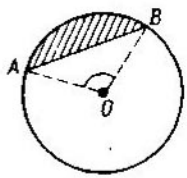


Fig. 6.61.

cia dada, para un aumento no limitado del número de los lados.

El *área del círculo* de radio r se calcula mediante la fórmula

$$S = \pi r^2.$$

Se llama *sector* a la parte del círculo, limitada por sus dos radios (fig. 6.60).

El *área de un sector* con un valor angular del arco de α se calcula mediante la fórmula

$$S_{\text{sect}} = \frac{\pi r^2 \alpha}{360}.$$

Se llama *segmento* (fig. 6.61) a la parte del círculo, limitada por una cuerda y el arco que la une.

El *área de un segmento* se calcula como la diferencia del área del sector, limitado por los radios OA y OB y el área del triángulo AOB (véase fig. 6.61).

§ 12. Los polígonos y la circunferencia

12.1. Polígonos inscritos y circunscritos. Un polígono, todos los vértices del cual pertenecen a la circunferencia, se llama *inscrito* en esta circunferencia, y la circunferencia se llama, circunferencia *circunscrita* alrededor del polígono.

19—01477

Un polígono, todos los lados del cual son tangentes a la circunferencia, se llama *circunscrito* alrededor de esta circunferencia, y la circunferencia se llama *inscrita* en este polígono.

Alrededor de cualquier polígono regular se puede circunscribir una circunferencia, y en cualquier polígono regular se puede inscribir una circunferencia.

El centro de una circunferencia inscrita en un polígono regular coincide con el centro de la circunferencia circunscrita alrededor del polígono regular; este punto se llama *centro* de un polígono regular.

El segmento de una perpendicular, bajada del centro de un polígono regular a su lado, se llama *apotema* de este polígono regular.

El *lado de un n-polígono regular* a_n se expresa a través del radio R de la circunferencia circunscrita alrededor de él, mediante la fórmula

$$a_n = 2R \operatorname{sen} \frac{180^\circ}{n}.$$

El *área de un n-polígono regular* es igual a la mitad del producto de su perímetro por el radio de la circunferencia inscrita:

$$S_n = \frac{1}{2} P_n r.$$

El *área de un n-polígono regular* se expresa a través del radio de la circunferencia circunscrita R mediante la fórmula

$$S = \frac{1}{2} R^2 n \operatorname{sen} \frac{360^\circ}{n}.$$

12.2. Triángulos inscritos. El triángulo, todos los vértices del cual pertenecen a la circunferencia, se llama triángulo *inscrito* en esta circunferencia, y la circunferencia se llama *circunscrita* alrededor de este triángulo. Alrededor de cualquier triángulo se puede circunscribir solamente una circunferencia. El centro de la circunferencia circunscrita alrededor de un triángulo es el punto de intersección de las perpendiculares medias a los lados de este triángulo: este punto se encuentra dentro del triángulo, si el triángulo es acutángulo; fuera del triángulo, si es obtusángulo y en el punto medio de la hipotenusa, si el triángulo es rectángulo.

El *radio de la circunferencia*, circunscrita alrededor de un triángulo arbitrario, se calcula mediante las fórmulas

$$R=\frac{1}{2}\cdot\frac{a}{\operatorname{sen}\alpha}=\frac{1}{2}\cdot\frac{b}{\operatorname{sen}\beta}=\frac{1}{2}\cdot\frac{c}{\operatorname{sen}\gamma},$$

$$R=\frac{abc}{4S_\Delta}=\frac{abc}{4\sqrt{p(p-a)(p-b)(p-c)}};$$

a, b, c son los lados del triángulo, $p=\frac{1}{2}(a+b+c)$ es el semiperímetro, S_Δ, el área del triángulo. α, β, γ son los ángulos del triángulo opuestos a los lados a, b, c respectivamente.

12.3. Triángulos circunscritos. Un triángulo, todos los lados del cual son tangentes a la circunferencia, se llama *circunscrito* alrededor de esta circunferencia, y la circunferencia se llama *inscrita* en este triángulo.

En cualquier triángulo se puede inscribir una circunferencia y sólo una. El centro O de la circunferencia inscrita en el triángulo es el punto de intersección de las bisectrices de los ángulos interiores del triángulo.

El *radio de la circunferencia*, inscrita en un triángulo arbitrario, se calcula según la fórmula

$$r=\frac{S_\Delta}{p}=\sqrt{\frac{(p-a)(p-b)(p-c)}{p}}.$$

12.4. Circunferencia inscrita por fuera. Se llama circunferencia *inscrita por fuera* a la que es tangente a un lado del triángulo y a la prolongación de los otros dos.

Las bisectrices de un par de ángulos exteriores del triángulo, adyacentes con los ángulos β y γ (fig. 6.62) se intersecan en el punto O_a. Por este mismo punto pasa la bisectriz del ángulo interior α. El punto O_a es el centro de la circunferencia inscrita por fuera, tangente al lado a y a la prolongación de los lados b y c. Análogamente se hallan los puntos O_b y O_c que son los centros de las circunferencias inscritas por fuera, tangentes a los lados b y c respectivamente (véase fig. 6.62).

Los *radios de las circunferencias inscritas por fuera* r_a, r_b, r_c, tangentes a los lados a, b, c respectivamente, se calculan por las fórmulas

$$r_a = \frac{S_\triangle}{p-a} = \frac{2S_\triangle}{b+c-a},$$

$$r_b = \frac{S_\triangle}{p-b} = \frac{2S_\triangle}{a+c-b},$$

$$r_c = \frac{S_\triangle}{p-c} = \frac{2S_\triangle}{a+b-c}.$$

12.5. Relaciones entre los lados de los triángulos regular y rectángulo y los radios de las circunferencias inscrita

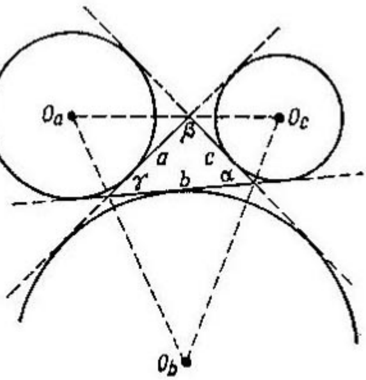


Fig. 6.62.

y circunscrita. El lado del triángulo regular a se relaciona con el radio de la circunferencia circunscrita R y el radio de la circunferencia inscrita r mediante las proporciones

$$R = \frac{a\sqrt{3}}{3}, \quad r = \frac{a\sqrt{3}}{6}.$$

El centro de la circunferencia inscrita en el triángulo regular coincide con el centro de la circunferencia, circunscrita alrededor de él y se llama *centro* del triángulo regular.

Circunferencia, circunscrita alrededor de un triángulo rectángulo. El centro de la circunferencia, circunscrita alrededor de un triángulo rectángulo, está situado en el punto medio de la hipotenusa y, por consiguiente, la hipotenusa del triángulo rectángulo es el diámetro de la circunferencia, circunscrita alrededor del triángulo rectangular.

Todos los puntos de la circunferencia con diámetro AB (a excepción de los puntos A y B) son vértices de los triángulos rectángulos con hipotenusa AB (fig. 6.63).

12.6. Cuadriláteros inscritos. Un cuadrilátero, todos los vértices del cual pertenecen a la circunferencia, se llama *inscrito* en esta circunferencia, y la circunferencia se llama

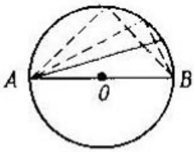


Fig. 6.63.

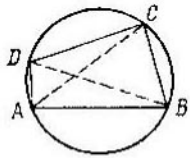


Fig. 6.64.

circunscrita alrededor de este cuadrilátero. No todo cuadrilátero tiene la propiedad de que alrededor de él se puede circunscribir una circunferencia. Para poder circunscribir una circunferencia alrededor de un cuadrilátero, es necesario y suficiente que la suma de los ángulos opuestos de éste sea igual a 180°.

En particular, de todos los paralelogramos solamente alrededor de un rectángulo (cuadrado) se puede circunscribir una circunferencia.

Alrededor de un trapecio se puede circunscribir una circunferencia cuando y sólo cuando este trapecio es isósceles.

Si el cuadrilátero $ABCD$ se puede inscribir en una circunferencia, entonces el producto de las diagonales de este cuadrilátero es igual a la suma de los productos de los lados opuestos (fig. 6.64):

$$|AC|\cdot|BD| = |AB|\cdot|CD| + |AD|\cdot|BC|.$$

12.7. Cuadriláteros circunscritos. Un cuadrilátero, todos los lados del cual son tangentes a la circunferencia, se llama *circunscrito* alrededor de la circunferencia, y la circunferencia se llama *inscrita* en este cuadrilátero.

Para poder inscribir en un cuadrilátero la circunferencia, es necesario y suficiente que las sumas de los lados opuestos de este cuadrilátero sean iguales.

De todos los paralelogramos solamente en el rombo (en particular, en el cuadrado) se puede inscribir una circunferencia. El centro de la circunferencia inscrita se encuentra en la intersección de diagonales de los cuadriláteros indicados.

§ 13. Construcciones geométricas

Los problemas con construcciones geométricas representan una de las partes tradicionales de geometría. Las construcciones geométricas se efectúan mediante una regla «simple» y un compás. Por regla «simple» se comprende un instrumento, mediante el cual se puede realizar una sola acción, trazar una recta (un rayo, segmento) por dos puntos dados. Por compás se entiende un instrumento, mediante el cual se pueden construir circunferencias y trazar en una recta un segmento geométricamente dado. En los problemas con construcciones citados más abajo las palabras «se da un segmento» y «se da un ángulo» significan que se da una representación geométrica de un segmento (respectivamente de un ángulo), y no su valor numérico. Para resolver los problemas con construcciones, es necesario tener en cuenta también que el problema de construir una figura geométrica consiste no en trazar prácticamente una figura con un grado conocido de precisión, sino en como mediante una regla y un compás puede ser realizada teóricamente la construcción requerida en la suposición de que nuestros instrumentos dan una precisión de construcción absoluta.

En conclusión formulemos tres problemas clásicos de construcción de figuras, cuya solución fue objeto de búsqueda durante varios siglos, hasta que fue demostrado que estos problemas no pueden ser resueltos mediante una regla y un compás.

1) *Duplicación del cubo.* Si un cubo dado tiene una arista igual a la unidad de longitud, entonces su volumen es igual a la unidad cúbica. Construir la arista de un cubo, cuyo volumen es dos veces mayor.

2) *Trisección del ángulo.* Dividir un ángulo arbitrario en tres ángulos iguales.

3) *Cuadratura del círculo.* Construir un cuadrado (un lado del cuadrado), cuya área es igual al área de un círculo, el radio del cual se toma por la unidad de longitud.

13.1. Construcción de rectas, paralelas y perpendiculares a una recta dada (a un segmento dado).

1) *Construir una recta paralela a una recta dada AB que pasa por un punto dado C.*

Construimos una circunferencia con centro en el punto dado C de tal manera que interseque la recta dada AB mediante la abertura arbitraria del compás (fig. 6.65). Mediante la misma abertura del compás desde un punto de

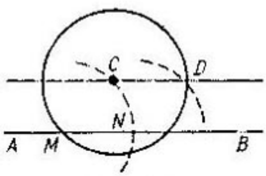


Fig. 6.65.

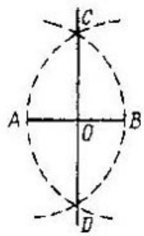


Fig. 6.66.

intersección de la recta y la circunferencia (en la fig. 6.65 el punto M) trazamos en la recta AB en cualquiera dirección el segmento MN. De nuevo, mediante la misma abertura del compás cortamos desde el punto N un punto de la circunferencia D. Trazamos una recta por los puntos C y D. La recta CD, es la recta buscada.

2) *Dividir el segmento dado por la mitad y construir una perpendicular al segmento en su punto medio.*

De los extremos de un segmento dado AB como de un centro (fig. 6.66), mediante el mismo radio arbitrario (mayor $\frac{1}{2} \mid AB \mid$), construimos dos arcos que se intersecan. La recta que pasa por los puntos de intersección de los arcos C y D es la perpendicular buscada. El punto O de intersección de las rectas AB y CD es el punto medio del segmento AB.

3) *Levantar una perpendicular a la recta dada MN en un punto dado A (construcción de un ángulo recto).*

Tomemos un punto arbitrario O, el cual no pertenece a la recta dada MN (fig. 6.67), y construyamos una circunferencia con el centro en el punto O de radio OA (A es el punto dado). Trazamos por el segundo punto de intersección de la circunferencia construida con la recta MN (en la fig. 6.67 por el punto B), el diámetro BC. La recta que pasa por los puntos C y A, es la perpendicular buscada. El ángulo, formado por los rayos AC y AB, es recto.

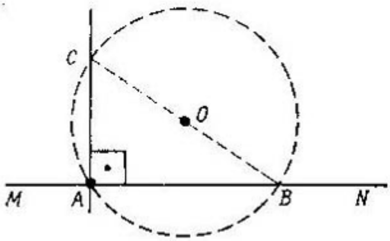


Fig. 6.67.

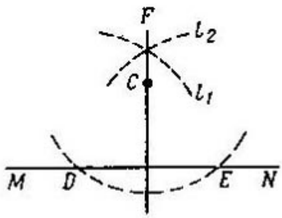


Fig. 6.68.

4) *Bajar una perpendicular del punto dado C a la recta dada MN.*

Del punto dado C como de un centro (fig. 6.68) mediante un radio arbitrario trazamos el arco DE que interseca la rec-

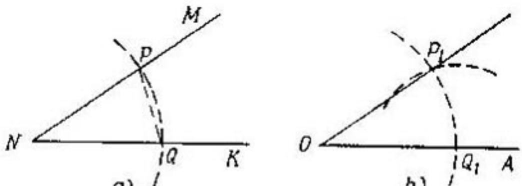


Fig. 6.69.

ta dada MN en los puntos D y E. De los puntos D y E como de un centro trazamos mediante el mismo radio arbitrario dos arcos l_1, l_2 que se intersecan en el punto F. La recta que pasa por los puntos C y F, es la perpendicular buscada.

13.2. Construcción de ángulos.

1) *Construir un ángulo, igual al ángulo dado MNK* (fig. 6.69, *a*).

Elegimos en el plano un punto arbitrario O y trazamos un rayo OA (fig. 6.69, *b*). Del vértice N del ángulo dado como de un centro circunscribimos un arco PQ de radio arbitrario. Por la misma abertura del compás circunscribimos del centro O un arco P_1Q_1. Del punto Q_1 como de un centro mediante un radio, igual a la longitud de la cuerda PQ, cortamos el punto P_1. Trazando el rayo OP_1, obtenemos el ángulo P_1OQ_1, igual al ángulo dado MNK.

2) *Construir los ángulos, iguales a* $60°$ *y* $30°$.

De los extremos A y B de un segmento arbitrario AB como de centros, mediente un radio igual a $|AB|$, circunscribimos dos arcos intersecantes (fig. 6.70). Construimos

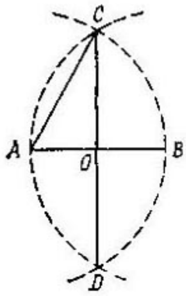


Fig. 6.70.

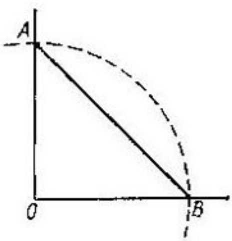


Fig. 6.71.

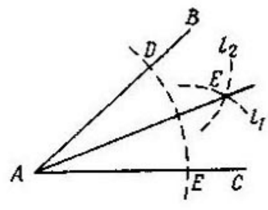


Fig. 6.72.

unos segmentos CD y AC por los puntos de su intersección C y D. El ángulo ACO es igual a $30°$, y el ángulo CAO, igual a $60°$.

3) *Construir un ángulo, igual a* $45°$.

En los lados del ángulo recto AOB (fig. 6.71) trazamos segmentos iguales OA y OB. Por los puntos A y B trazamos la recta AB. El ángulo, formado por los rayos BA y BO, es igual a $45°$.

4) *Dividir el ángulo dado* BAC *en dos ángulos iguales (construir la bisectriz del ángulo).*

Del vértice del ángulo BAC como de un centro trazamos el arco DE de una circunferencia de radio arbitrario (fig. 6.72). De los puntos D y E de su intersección con los rayos AB y AC circunsribimos mediante radios iguales arbitrarios los arcos l_1 y l_2. Por el punto de intersección de éstos F trazamos el rayo AF. Los ángulos obtenidos BAF y FAC son iguales, y el rayo AF es la bisectriz del ángulo BAC,

13.3. Construcción de segmentos.

1) *Dividir el segmento dado AB por el número dado de segmentos iguales.*

Construimos una recta, paralela (pero no coincidente) con la recta que contiene el segmento dado AB. Tomamos un punto arbitrario A_1, perteneciente a la recta construida (fig. 6.73). Del punto A_1 trazamos tantos segmentos iguales

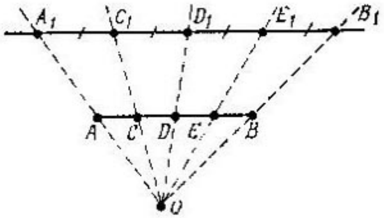


Fig. 6.73.

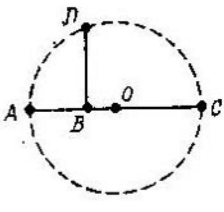


Fig. 6.74.

A_1C_1, C_1D_1, D_1E_1, ... de longitud arbitraria, en cuantas partes es necesario dividir el segmento AB. Designamos con la letra B_1 el extremo del último segmento. Trazamos las rectas AA_1 y BB_1 hasta intersecarse en el punto O. A través de los pares de puntos $(C_1; O)$, $(D_1; O)$, $(E_1; O)$, ... trazamos los rayos que intersecan la recta AB en los puntos C, D, E, ..., los cuales dividen el segmento AB en un número necesario de segmentos iguales.

2) *Dividir un segmento dado en segmentos, proporcionales a los valores dados.*

El problema se resuelve de la misma manera que el anterior, sólo que del punto A_1 se trazan segmentos, proporcionales a los valores dados.

3) *Construir un segmento que sea la media proporcional a dos segmentos dados*).*

En una recta arbitraria tracemos los segmentos dados de tal manera que un extremo de un segmento coincida con el origen del otro (en la fig. 6.74) los segmentos AB y BC). Dividamos el segmento AC por la mitad y mediante un ra-

*) El segmento x que satisface a la igualdad $a/x = x/b$, donde a y b, son los segmentos dados, se llama *segmento de media proporcional* a dos segmentos dados.

dio, igual a la mitad del segmento AC, construimos una circunferencia con centro en el punto medio del segmento AC (en la fig. 6.74 el punto O es el punto medio del segmento AC). Del punto B levantamos una perpendicular al segmento AC. Designemos el punto de intersección de la perpendicular y la circunferencia con la letra D. El segmento BD es la media proporcional a los segmentos AB y BC:

$$\frac{|AB|}{|BD|}=\frac{|BD|}{|BC|}.$$

4) *Construir el cuarto segmento proporcional (por los segmentos dados, a, b y c construir el segmento x tal, que $a : b = c : x$).*

En un lado de un ángulo arbitrario α (fig. 6.75) trazamos sucesivamente desde su vértice los segmentos a y c, y en

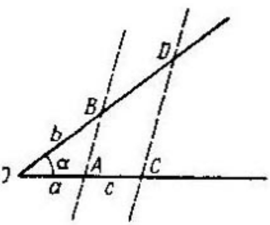


Fig. 6.75.

Fig. 6.76.

el otro lado, el segmento b. Por el punto C trazamos una recta, paralela a la recta AB. El segmento BD, cortado en el lado del ángulo OB, es el buscado.

5) *Construir un segmento, conmensurable con el dado.*

Un segmento AB se llama *conmensurable* con el segmento CD, si la relación de las longitudes de estos segmentos es un número racional:

$$\frac{|AB|}{|CD|}=\frac{m}{n}\ (m,\ n \text{ son números naturales}).$$

Sea dado cierto segmento AB y se requiere construir un segmento conmensurable con el dado. Dividamos el segmento AB en m partes iguales. Tomemos una de estas partes iguales y tracémosla n veces en el rayo CN (fig. 6.76) de tal manera que el extremo del segmento anterior (a excepción del primero) sea el origen del segmento siguiente. Entonces el extremo del último segmento da el punto D, que es el extremo del segmento buscado CD.

6) *Ejemplos de construcción de segmentos, inconmensurables al segmento dado.*

a) *Por el segmento dado de longitud a construir un segmento de longitud* $a\sqrt{2}$.

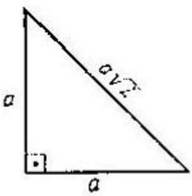


Fig. 6.77.

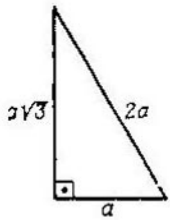


Fig. 6.78.

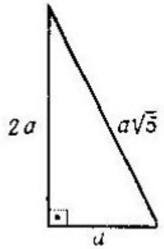


Fig. 6.79.

Construimos un triángulo rectángulo isósceles con catetos de longitud a (fig. 6.77). La hipotenusa de tal triángulo será el segmento $a\sqrt{2}$.

b) *Por un segmento dado a construir el segmento* $a\sqrt{3}$.

Construimos un triángulo rectángulo con un cateto a y una hipotenusa $2a$ (fig. 6.78).

El segmento $a\sqrt{3}$ será el segundo cateto del triángulo rectangular construido.

c) *Por un segmento dado a construir el segmento* $a\sqrt{5}$.

Construimos un triángulo rectangular con los catetos a y $2a$ (fig. 6.79). El segmento $a\sqrt{5}$ será la hipotenusa del triángulo construido.

7) *Ejemplos de construcción de segmentos, representados por expresiones algebraicas.* Citemos aquí tres ejemplos de construcción de segmentos, representados por expresiones algebraicas (por las palabras «el segmento se da por una expresión algebraica» se entiende que el segmento buscado se representa como una función algebraica de los segmentos dados).

Para la construcción de segmentos, representados por expresiones algebraicas, se usan las construcciones elementales siguientes:

construcción de un segmento que sea la media proporcional a dos segmentos dados;

construcción del cuarto segmento proporcional;

construcción de un segmento, inconmensurable al dado.

Ejemplo 1. Construir un segmento, representado por la expresión

$$\sqrt{a^2+ab+c^2},$$

donde a, b, y c son los segmentos dados.

Hagamos la siguiente transformación de la expresión que se encuentra bajo el signo del radical:

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

Es muy fácil observar que si construimos un segmento x tal, que

$$x^2=a(a+b),$$

es decir, un segmento que es la media proporcional a los segmentos a y $a+b$, entonces el segmento buscado se obtendrá como la hipotenusa de un triángulo rectangular con catetos x y c. De tal manera, la construcción del segmento buscado se reduce a la realización de dos construcciones sucesivas, es decir, a la construcción de la media proporcional y a la construcción de un triángulo rectangular por dos catetos.

Ejemplo 2. Construir un segmento, representado por la expresión

$$\sqrt{3a^2+5b^2},$$

donde a y b son los segmentos dados.

Construimos el segmento $a\sqrt{3}$ como un cateto del triángulo rectangular con la hipotenusa $2a$ y el cateto a dados.

Construimos el segmento $b\sqrt{5}$ como la hipotenusa del triángulo rectangular con los catetos dados $2b$ y b.

El segmento buscado se construye como la hipotenusa de un triángulo rectangular con los catetos $a\sqrt{3}$ y $b\sqrt{5}$.

Ejemplo 3. Construir un segmento x, representado por la relación

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

donde a y b son los segmentos dados.

Realizando las transformaciones evidentes de la igualdad dada:

$$\frac{1}{x}=\frac{1}{a}+\frac{1}{b}=\frac{a+b}{ab} \Longleftrightarrow \frac{b}{x}=\frac{a+b}{a},$$

observamos que la construcción del segmento x se reduce a la construcción del cuarto segmento proporcional.

13.4. Construcción de circunferencias y de arcos de circunferencias.

1) *Trazar una circunferencia de radio dado por dos puntos dados.*

De los puntos dados A y B como de un centro mediante el radio r dado circunscribimos dos arcos que se intersecan

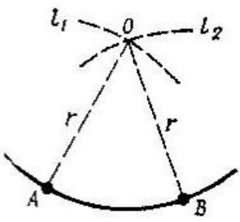


Fig. 6.80.

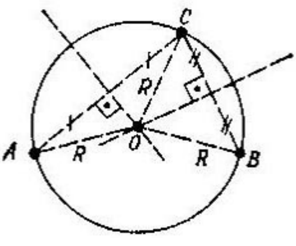


Fig. 6.81.

l_1 y l_2. El punto de su intersección O (fig. 6.80) es el centro de la circunferencia buscada *).

2) *Circunscribir una circunferencia por tres puntos dados A, B, C que no están situados en una recta.*

Construimos unos segmentos BC y AC (fig. 6.81), cuyos extremos son los puntos dados A, B, C. Trazamos unas perpendiculares por los puntos medios de estos segmentos. El

*) Aquí son posibles tres casos: existen dos circunferencias que pasan por los puntos A y B (si $|AB| < 2r$); si $|AB| = 2r$, tal circunferencia sólo hay una; para $|AB| > 2r$ no existe la circunferencia buscada.

punto de intersección de estas perpendiculares (en la fig. 6.81, el punto O) es el centro de la circunferencia buscada. El radio de la circunferencia dada es igual a la distancia del punto O a cualquiera de los tres puntos equidistantes de O.

3) *Por dos puntos dados A y B trazar una circunferencia, tangentes a la recta dada a, la cual no pasa por A y B.*

Trazamos una recta por los puntos dados A y B. Son posibles dos casos:

a) la recta AB se interseca con la recta dada a;

b) la recta AB es paralela a la recta dada a.

Examinemos el *primer caso*. Designemos el punto de intersección de las rectas AB y a con la letra C (fig. 6.82).

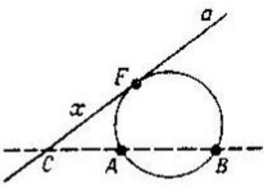


Fig. 6.82.

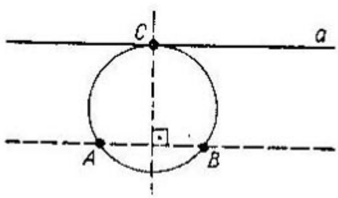


Fig. 6.83.

Construimos el segmento x, que es la media proporcional a los segmentos CA y CB. Trazamos el segmento x en la recta a del punto C (en la fig. 6.82 $x = CF$). El punto F es el punto de tangencia de la circunferencia buscada con la recta dada. De esta manera, el problema se reduce a la construcción de una circunferencia que pasa por tres puntos dados A, B y F no situados en una misma recta. El problema tiene una segunda solución: se puede construir una segunda circunferencia que pasa por los puntos dados A y B y que es tangente a la recta dada a. Para esto es necesario trazar en la recta a el segmento x por el otro lado del punto C.

Segundo caso. Si las rectas AB y a son paralelas, entonces trazamos una perpendicular por el punto medio del segmento AB hasta la intersección con la recta a en el punto C (fig. 6.83). El punto C es el punto de tangencia de la circunferencia buscada y la recta dada a. La circunferencia que pasa por los puntos A, B y C es la buscada.

4) *Hallar el centro del arco dado de una circunferencia.*

En un arco dado elegimos tres puntos cualesquiera A, B, C y construimos el centro de la circunferencia que pasa por los tres puntos dados.

5) *Dividir el arco dado de una circunferencia en dos partes iguales.* Mediante una cuerda unimos los extremos del arco dado. Trazamos una perpendicular por el punto medio de la cuerda. El punto de intersección de la perpendicular y el arco divide el arco dado en dos arcos iguales.

6) *Hallar un conjunto de puntos del plano, de los cuales el segmento dado AB se ve bajo el ángulo dado α.*

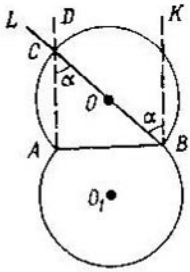


Fig. 6.84.

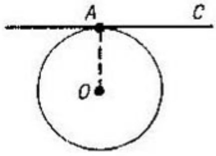


Fig. 6.85.

Un conjunto buscado de puntos representa arcos iguales de circunferencias iguales con una cuerda común AB (fig. 6.84). Los centros y radios de estas circunferencias se hallan de la manera siguiente. De los puntos A y B levantamos perpendiculares al segmento AB. Designemos estas perpendiculares AD y BK respectivamente. Trazamos el ángulo $KBL = \alpha$ partiendo del rayo BK de tal manera que el rayo BL interseque el rayo AD en cierto punto C. El punto medio O del segmento BC es el centro de una de las circunferencias buscadas. El radio de esta circunferencia es igual a $\frac{1}{2} \mid BC \mid$. La otra circunferencia se construye análogamente, más bajo que la recta que pasa por los puntos A, B.

13.5. Construcción de tangentes a las circunferencias.

1) *Trazar por un punto dado una tangente a la circunferencia dada.*

Si un punto dado A pertenece a la circunferencia dada con centro O (fig. 6.85), entonces la perpendicular a AO, trazada por el extremo A del segmento AO, es la tangente buscada.

Si el punto A se encuentra fuera del círculo (fig. 6.86), entonces es necesario dividir el segmento AO por la mitad y de su punto medio B, mediante el radio OB, trazar el arco DOC (D y C son los puntos de intersección del arco y la circunferencia dada). Por los pares de los puntos (D; A) y (C; A) trazamos las rectas, que serán las tangentes buscadas.

2) *Trazar a las dos circunferencias dadas una tangente exterior común.*

El problema no tiene soluciones, si un círculo se encuentra por completo en el otro y las circunferencias no son tangentes. Si un círculo se contiene por completo en el otro y las

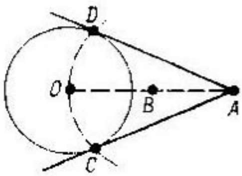


Fig. 6.86.

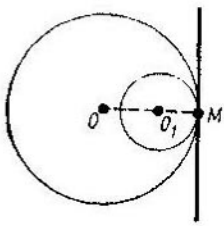


Fig. 6.87.

circunferencias son tangentes por dentro, entonces el problema tiene una sola solución (existe sólo una recta, tangente a las dos circunferencias). En todos los demás casos de posición mutua de dos circunferencias el problema tiene dos soluciones (es decir, existen dos tangentes exteriores comunes diferentes a las circunferencias dadas).

Examinemos un caso cuando dos circunferencias son tangentes por el interior. Trazamos la línea de los centros OO_1 de estas circunferencias (fig. 6.87). En el punto de tangencia de dos circunferencias (en la fig. 6.87, el punto M) construimos una recta que es perpendicular a la línea de los centros. Esta es la tangente buscada.

Examinemos un caso, cuando el problema tiene dos soluciones.

a) Si los radios de las circunferencias dadas son iguales, entonces trazamos por los centros de las circunferencias O

y O_1 (fig. 6.88) los diámetros AB y A_1B_1 que son perpendiculares a la línea de los centros OO_1.

Las rectas que pasan por los pares de los puntos (A; A_1) y (B; B_1) son las tangentes buscadas.

b) Si los radios R y R_1 de las circunferencias dadas no son iguales ($R > R_1$), entonces del centro del círculo mayor trazamos una circunferencia de radio $OC = R - R_1$ (fig. 6.89).

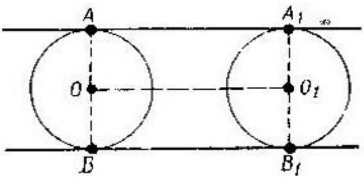


Fig. 6.88.

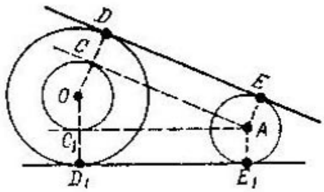


Fig. 6.89.

A la circunferencia construida trazamos las tangentes AC y AC_1, que pasan por el centro del círculo menor A. Por el centro de la circunferencia mayor O y por los puntos de tangencia C y C_1 trazamos los rayos OC y OC_1, los cuales intersecan la circunferencia mayor en los puntos D y D_1. Trazamos los radios AE y AE_1, perpendiculares a las rectas AC y AC_1 respectivamente. Las rectas que pasan por los pares de los puntos (D; E), (D_1; E_1), son las tangentes buscadas.

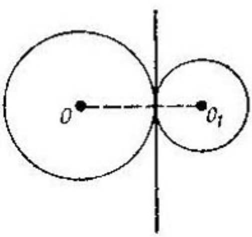


Fig. 6.90.

3) *Trazar a dos circunferencias dadas una tangente común interior.*

El problema no tiene solución, si los círculos (circunferencias) se intersecan. Si las circunferencias son tangentes por fuera, el problema tiene una sola solución, por el punto de tangencia se traza una perpendicular a la línea de los centros (fig. 6.90).

En los demás casos el problema tiene dos soluciones (es decir, existen dos rectas diferentes, cada una de las cuales será tangente tanto a la una, como a la otra circunferencia). Las tangentes se construyen de la manera siguiente. Del centro de cualquiera circunferencia (por ejemplo, de una circunferencia con centro A (fig. 6.91)) trazamos una circunferencia de radio igual a la suma de los radios de las circunferencias dadas. Del centro B de la segunda circunferencia trazamos las tangentes BC y BC_1 a la circunferencia construida. Trazamos los radios AC y AC_1 en los puntos de tangencia. Estos radios intersecan la circunferencia dada con el

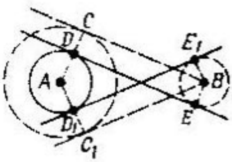


Fig. 6.91.

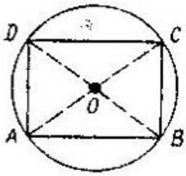


Fig. 6.92.

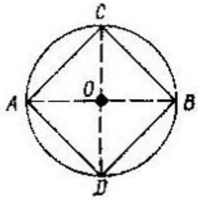


Fig. 6..93

centro A en los puntos D y D_1. Por el centro de la segunda circunferencia B trazamos los redios BE y BE_1 que son perpendiculares a las rectas BC y BC_1 respectivamente. Las rectas que pasan por los pares de los puntos (D; E) y (D_1; E_1) son las tangentes buscadas.

13.6. Construcción de una circunferencia, circunscrita alrededor de un polígono, y de un polígono, inscrito en una circunferencia.

1) *Circunscribir una circunferencia alrededor de un triángulo dado.*

Por los vértices del triángulo A, B, C se circunscribe una circunferencia.

2) *Circunscribir una circunferencia alrededor de un rectángulo dado (o de un cuadrado).*

Trazamos las diagonales AC y BD (fig. 6.92). Del punto O de su intersección como de un centro circunscribimos una circunferencia de radio $|OA|$. La circunferencia construida es la buscada.

3) *Inscribir un cuadrado en un círculo dado.*

Trazamos dos diámetros AB y CD perpendiculares entre sí. El cuadrilátero con los vértices A, B, C, D es el cuadrado buscado (fig. 6.93).

4) *Inscribir en una circunferencia dada un hexágono regular y un triángulo regular.*

Eligiendo un punto arbitrario (por ejemplo, el punto A) que pertenece a la circunferencia, con la abertura del compás igual al radio de la circunferencia hacemos en ésta unas marcas obteniendo sucesivamente los puntos B, C, D, E, F (fig. 6.94). Uniendo sucesivamente los puntos indicados, obtenemos un hexágono regular, inscrito en la circunferencia

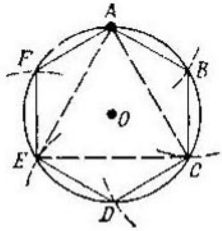


Fig. 6.94.

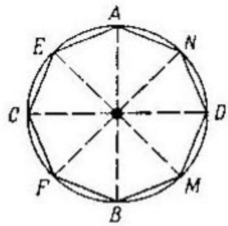


Fig. 6.95.

dada. Uniendo un punto sí y otro no, obtenemos un triángulo regular (equilátero), inscrito en la circunferencia dada.

5) *Inscribir en la circunferencia dada un octágono regular.* Trazamos dos diámetros AB y CD perpendiculares entre sí (fig. 6.95). Trazando las bisectrices de cuatro ángulos rectos, obtenemos los puntos de intersección de las bisectrices con la circunferencia (en la fig. 6.95, los puntos E, F, M, N). Uniendo sucesivamente los ocho puntos obtenidos A, E, C, F, B, M, D, N, obtenemos el octágono buscado.

13.7. Construcción de una circunferencia, inscrita en un polígono, y de un polígono, circunscrito alrededor de una circunferencia.

1) *Inscribir una circunferencia en un triángulo dado.* Dividimos dos ángulos interiores cualesquiera de un triángulo en dos pares de ángulos iguales (es decir, construimos dos bisectrices de los ángulos interiores del triángulo). Del punto O de intersección de bisectrices (fig. 6.96) trazamos una perpendicular a cualquier lado del triángulo (en la fig. 6.96, hacia el lado AB), la cual interseca el lado en el punto D. Mediante el radio $|OD|$ con centro en el punto O circunscribimos la circunferencia buscada.

2) *Inscribir una circunferencia en un rombo (o en un cuadrado).*

Trazamos las diagonales del rombo (en la fig. 6.97, los segmentos AC y BD). Por el punto O de su intersección trazamos una perpendicular a cualquier lado del rombo (en la fig. 6.97, al lado BC; E es el punto de intersección de la perpendicular con el lado). La circunferencia con centro O y de radio igual a $|OE|$ es la circunferencia buscada.

3) *Circunscribir un cuadrado alrededor de una circunferencia dada.*

Trazamos dos diámetros perpendiculares entre sí de la circunferencia dada (en la fig. 6.98 AB y CD, los diámetros).

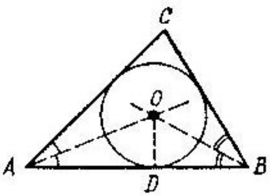


Fig. 6.96.

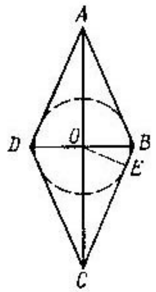


Fig. 6.97.

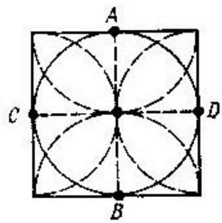


Fig. 6.98.

De sus extremos (de los puntos A, B, C, D) como de un centro circunscribimos cuatro semicircunferencias mediante radios iguales al radio de la circunferencia dada. Los puntos de intersección de las semicircunferencias son los vértices del cuadrado buscado.

4) *Circunscribir alrededor de una circunferencia dada un hexágono y un octágono regulares.*

Marcando en la circunferencia seis (ocho) puntos de la misma manera, que en los problemas 4), 5) p. 13,6, construimos las tangentes a la circunferencia dada en los puntos marcados. Los puntos de intersección de las tangentes vecinas darán los vértices del hexágono regular (del octágono).

13.8. Construcción de triángulos.

1) *Construir un triángulo por tres lados a, b, c.*

De los extremos del segmento $AB = a$ como de los centros circunscribimos dos arcos de circunferencias mediante los radios b y c (fig. 6.99). El punto de su interseccion C es el tercer vértice del triángulo buscado ABC. El problema tiene solución, si los segmentos a, b y c satisfacen las desigualdades del triángulo.

2) *Construir un triángulo por dos lados a, b y el ángulo entre ellos.*

Construimos un ángulo que es igual al ángulo γ. Del vértice del ángulo trazamos en sus lados los segmentos CA

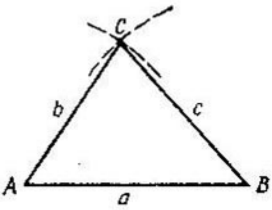


Fig. 6.99.

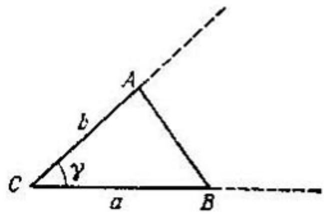


Fig. 6.100.

y CB, respectivamente iguales a los segmentos b, a (fig. 6.100). Juntamos los puntos A y B. El triángulo ABC es el triángulo buscado.

3) *Construir un triángulo por un lado a y los ángulos adyacentes a él β y γ ($\beta + \gamma < 180°$).*

Para cada uno de los extremos del segmento BC ($BC = a$) trazamos los ángulo β y γ (fig. 6.101). El punto A de intersección de los rayos BA y CA es el tercer vértice del triángulo ABC.

4) *Construir un triángulo por la base a, la altura h_a y el ángulo del vértice α.*

Construimos un arco de una circunferencia, del cual el segmento $AB = a$ se ve bajo el ángulo dado α (véase el problema 6) p. 13.4). Trazamos una perpendicular arbitraria al segmento AB y trazamos en él un segmento $AN = h_a$. Por el punto N trazamos una recta NL que es paralela a la recta que pasa por los puntos A, B. El punto obtenido de intersección de la recta NL con el arco de la circunferencia (el punto C en la fig. 6.102) es el vértice del triángulo buscado ABC (el problema tiene dos soluciones).

5) *Construir un triángulo por dos lados a, b y el ángulo que es opuesto al lado a.*

Construimos un arco de la circunferencia, de la cual el segmento $AB = a$ se ve bajo el ángulo dado α. Del extremo A del segmento AB como centro mediante un radio igual a b trazamos un arco que se interseca con el primer arco. El

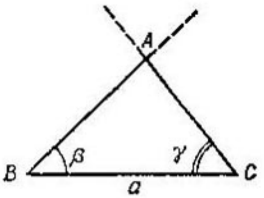


Fig. 6.101.

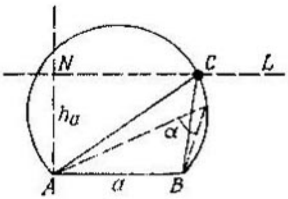


Fig. 6.102.

punto de intersección de estos arcos (en la fig. 6.103 son los puntos C y C_1) será el tercer vértice del triángulo buscado ABC (C_1). En dependencia de los valores a, b y α el problema tiene una o dos soluciones o no tiene soluciones.

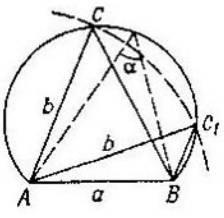


Fig. 6.103.

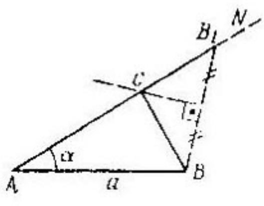


Fig. 6.104.

6) *Construir un triángulo por el lado a, el ángulo de la base α y la suma de los otros dos lados $b + c$.*

Del extremo A del segmento AB ($AB = a$) trazamos el ángulo α. En el rayo AN trazamos la suma dada de dos lados del triángulo buscado (en la fig. 6.104 $AB_1 = b + c$). Del punto medio del segmento BB_1 levantamos una perpendicular hasta la intersección con el rayo AB_1. El punto de intersección del rayo y la perpendicular (en la fig. 6.104 es el punto C) es el tercer vértice del triángulo buscado ABC.

7) *Construir un triángulo por un lado a, un ángulo de la base α y la diferencia de los otros dos lados $b - c$.*

Del extremo A del segmento AB ($AB = a$) trazamos un ángulo α. En el rayo AN trazamos la diferencia dada $b - c$ de dos lados del triángulo buscado (en la fig. 6.105, el segmento AB_1, $AB = b - c$). Del punto medio del segmento BB_1 levantamos una perpendicular hasta la intersección con el rayo AN. El punto de intersección del rayo AN y la per-

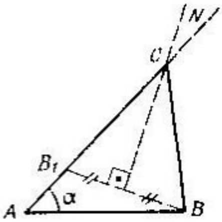


Fig. 6.105.

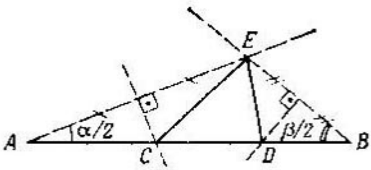


Fig. 6.106.

pendicular (en la fig. 6.105, el punto C) es el tercer vértice del triángulo buscado ABC.

8) *Construir un triángulo por dos ángulos dados α, β y el perímetro P.*

Construimos los ángulos $\alpha/2$, $\beta/2$. Construimos el segmento $AB = P$. De los extremos A y B trazamos los ángulos $\alpha/2$ y $\beta/2$ de tal manera que los rayos que son los segundos lados de estos ángulos, pertenezcan a un semiplano respecto a la recta AB (fig. 6.106). De los puntos medios de los segmentos AE y BE (E es el punto de intersección de los lados de los ángulos) levantamos las perpendiculares hasta la intersección con el segmento AB en los puntos C y D. El triángulo CED es el buscado.

9) *Construir un triángulo por la altura h_a, mediana m_a y bisectriz l_a, trazadas desde el mismo vértice.*

Construimos un triángulo rectangular ABC con la hipotenusa $AC = m_a$ y el cateto $AB = h_a$ (fig. 6.107). Mediante un radio que es igual a l_a, con centro en el punto A construimos un arco que interseca el cateto BC en el punto D. Construimos una recta MN que es perpendicular a la recta BC y que pasa por el punto C. Designamos el punto de intersección del rayo AD y de la recta MN con la letra L. Levantamos una perpendicular al segmento AL en su punto medio. El punto de intersección O de la perpendicular dada con la recta MN es el centro de la circunferencia circunscrita alrededor del triángulo buscado. Mediante un radio que es igual a AO, con centro en el punto O trazamos una circunferencia que interseca la recta BC en los puntos R y S. El triángulo con los vértices A, R, S es el buscado.

10) *Construir un triángulo por tres medianas* m_a, m_b, m_c.

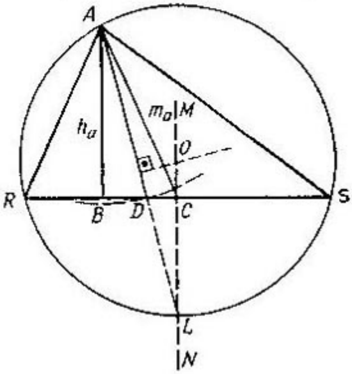


Fig. 6.107.

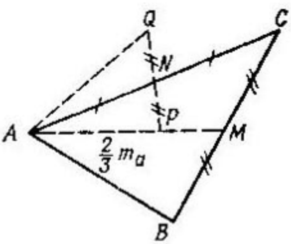


Fig. 6.108.

Construimos los segmentos $\frac{2}{3}m_a$, $\frac{2}{3}m_b$, $\frac{2}{3}m_c$.

Construimos un triángulo APQ (fig. 6.108) con los lados $\frac{2}{3}m_a$, $\frac{2}{3}m_b$, $\frac{2}{3}m_c$ (véase el problema 1) del mismo punto). En el rayo AP del punto P trazamos el segmento $PM = \frac{1}{3}m_a$. Hallamos el punto medio del segmento PQ y en el rayo AN trazamos el segmento $NC = AN$. Construimos un rayo CM y trazamos en él un segmento $BM = CM$ de tal manera que el punto M pertenezca al segmento BC. Los puntos A, B y C son los vértices del triángulo buscado.

11) *Construcción de triángulos rectangulares.*

La construcción del triángulo rectangular por dos catetos se efectúa de la misma manera que en el problema 2) del presente punto, para el caso $\beta = 90°$.

La construcción de un triángulo rectangular por un cateto y un ángulo agudo α se efectúa de la misma manera que en el problema 3), si el ángulo β lo hacemos igual a $90°$.

La construcción de un triángulo rectangular por la hipotenusa y el ángulo agudo α se reduce al problema 3) para el caso de los ángulos agudos α y $\beta = 90° - \alpha$.

12) *Construir un paralelogramo por los lados dados a, b y el ángulo α entre ellos.*

Construimos el ángulo MAN que es igual al ángulo dado α (fig. 6.109). En el rayo AM trazamos un segmento $AC = a$, y en el rayo AN trazamos el segmento $AB = b$. Trazamos del punto B como de un centro el arco de radio AC,

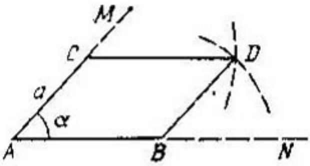


Fig. 6.109.

y del punto C, el arco de radio AB. Juntamos el punto de intersección de estos arcos (en la fig. 6.109, el punto D) con los puntos B y C. El cuadrilátero $ABDC$ es el paralelogramo buscado.

§ 14. Ángulo poliedro

Sean dados un polígono plano Φ y cualquier punto S que no pertenece al plano del polígono dado (en la fig. 6.110 el polígono Φ es un hexágono $ABCDEF$). La unión de todos los rayos que tienen un origen común S e intersecan un polígono dado Φ (fig. 6.110) se llama *ángulo poliedro**).

El punto S se llama *vértice* del ángulo poliedro; los rayos SA, SB, SC, SD, SE, SF que contienen los vértices del polígono se llaman sus *aristas*; los planos que contienen los triángulos SAB, SBC, etc., sus *caras*.

Según el número de caras se distinguen los ángulos triedros, tetraedros, pentaedros, etc. El ángulo poliedro se designa mediante letras, anotando primero el vértice, y después sucesivamente un punto en cada una de sus aristas.

*) A veces también se llama *ángulo poliedro* al conjunto de todos los rayos que tienen un origen común S e intersecan una quebrada cerrada $ABCDEF$, es decir, a la unión de todas las caras del ángulo poliedro.

Así, por ejemplo, el ángulo de seis caras representado en la fig. 6.110 se denota

$$SABCDEF.$$

Cada dos caras de un ángulo poliedro que tienen una cara común forman un ángulo diedro.

Se llama *región interior del* ángulo poliedro al conjunto de todos sus puntos que no pertenecen a las caras. El ángulo poliedro, cuya región interior está situada a un lado del plano de cada una de sus caras, se llama ángulo poliedro *convexo.* En caso contrario el ángulo poliedro se llama *no convexo.*

Los ángulos ASB, BCS, etc. se llaman *ángulos planos* del ángulo poliedro $SABCDEF$ (véase fig. 6.110).

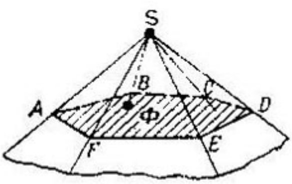


Fig. 6.110.

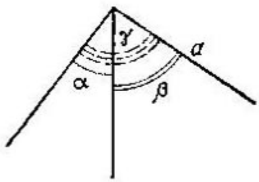


Fig. 6.111.

Las propiedades de los ángulos planos de un ángulo poliedro son:

1) Cada ángulo plano del ángulo poliedro es menor que la suma de los demás ángulos planos suyos.

2) En un ángulo poliedro convexo la suma de los ángulos planos es menor que 360°.

Un ángulo poliedro simple es un ángulo triedro. El cumplimiento de las desigualdades

$$|\beta-\gamma|<\alpha<\beta+\gamma, \quad \alpha+\beta+\gamma<360^\circ$$

es una condición necesaria y suficiente para la existencia de un ángulo triedro con los ángulos planos α, β, γ.

Teorema de los cosenos para el ángulo triedro. El coseno de un ángulo plano del ángulo triedro es igual al producto de los cosenos de los otros dos ángulos planos, sumado con el producto de los senos de los mismos ángulos y el coseno del ángulo diedro definido por estos ángulos planos (fig. 6.111):

$$\cos\alpha=\cos\beta\cos\gamma+\operatorname{sen}\beta\operatorname{sen}\gamma\cos a.$$

§ 15. Superficie poliédrica. Poliedro

Se llama *superficie poliédrica* al conjunto de un número finito de polígonos planos tal, que cada lado de cualquiera de los polígonos es al mismo tiempo el lado del otro (pero sólo de uno) polígono, que se llama *adyacente* con el primero.

Desde cualquiera de los polígonos que forman una superficie poliédrica se puede llegar hasta cualquier otro, moviéndose por los polígonos adyacentes.

Los polígonos que forman una superficie poliédrica se llaman *caras* de ésta; los lados de los polígonos se llaman *aristas*, y los vértices, *vértices* de la superficie poliédrica.

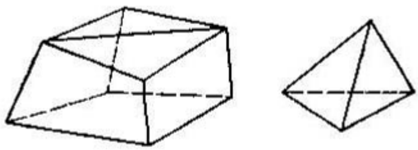


Fig. 6.112.

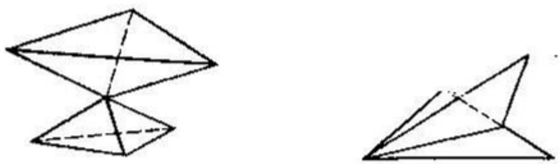


Fig. 6.113.

En la fig. 6.112 están representadas las uniones de los polígonos que satisfacen las exigencias indicadas y que son superficies poliédricas. En la fig. 6.113 se muestran figuras que no son superficies poliédricas.

La superficie poliédrica divide el espacio en dos partes, la región interior de la superficie poliédrica y la región exterior. De las dos regiones la exterior será aquella, en la cual existen rectas que pertenecen por completo a la región.

La unión de la superficie poliédrica y su región interior se llama *poliedro*. Al mismo tiempo la superficie poliédrica y su región interior se llaman respectivamente *superficie* y *región interior* del poliedro. Las caras, aristas y vértices de la superficie del poliedro se llaman respectivamente *caras*, *aristas* y *vértices* del poliedro.

Un poliedro se llama *convexo*, si todo él está situado a un lado del plano de cualquiera de sus caras.

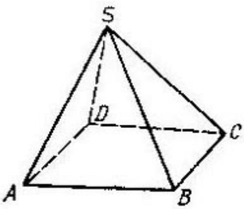


Fig. 6.114.

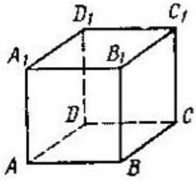


Fig. 6.115.

Se llama *diagonal* de un poliedro al segmento que une dos vértices de éste, no situados en la misma cara.

El poliedro se suele designar enumerando sus vértices e indicando sus propiedades especiales. Por ejemplo, el poliedro $SABCD$ representado en la fig. 6.114 es una pirámide, el poliedro $ABCDA_1B_1C_1D_1$ (fig. 6.115), un paralelepípedo.

§ 16. Prisma

Se llama *prisma n-angular* al polígono, dos caras del cual son n-polígonos iguales situados en los planos paralelos, y las demás n caras son paralelogramos.

Se llaman *bases* del prisma a un par de n-polígonos iguales. Las demás caras del prisma se llaman *laterales*, y su unión se llama *superficie lateral* del prisma. En la fig. 6.116 se muestra un prisma pentagonal.

Los lados de las caras del prisma se llaman *aristas*, y los extremos de sus aristas, *vértices* del prisma. Las aristas que no están situadas en la base del prisma se llaman *aristas laterales*.

Un prisma es *recto* si las aristas son perpendiculares a los planos de las bases. En caso contrario el prisma se llama *inclinado*.

Se llama *altura* del prisma al segmento de la perpendicular a los planos de las bases del prisma, cuyos extremos pertenecen a estos planos.

Se llama prisma *regular* al prisma recto, cuyas bases son polígonos regulares.

Área de la superficie lateral del prisma. Sea dado un prisma arbitrario (en la fig. 6.117, el prisma pentagonal).

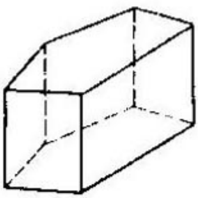


Fig. 6.116.

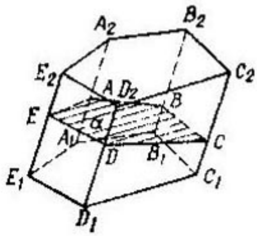


Fig. 6.117.

Por el punto A que pertenece a una de sus aristas laterales trazamos el plano α, que es perpendicular a esta arista (y, por consiguiente, perpendicular a todas las demás aristas laterales). Si el plano α interseca todas las aristas laterales del prisma, entonces el polígono obtenido (en la fig. 6.117, el pentágono $ABCDE$) se llama *sección* perpendicular del prisma (si no existe tal polígono, entonces por sección perpendicular del prisma se toma el polígono que tiene los vértices en los puntos de intersección del plano α con las prolongaciones de las aristas laterales).

El *área de la superficie lateral* del prisma se calcula según la fórmula

$$S_{\text{lat}} = P_n \cdot |A_1A_2|,$$

donde P_n, es el perímetro de la sección perpendicular del prisma, $|A_1A_2|$, la longitud de la arista lateral.

El *área de la superficie total* del prisma es igual a la suma de las áreas de la superficie lateral del prisma y de sus dos bases.

El *volumen del prisma inclinado* se calcula por la fórmula

$$V = S_n \cdot |A_1A_2|,$$

donde S_n es el área de la sección perpendicular del prisma, $|A_1A_2|$, la longitud de la arista lateral, o según la fórmula

$$V = S_{base} \cdot H,$$

donde S_{base} es el área de la base del prisma, H es la altura.

§ 17. Paralelepípedo. Cubo

Se llama *paralelepípedo* al prisma cuyas bases son paralelogramos.

Todas las seis caras del paralelepípedo (fig. 6.118) son paralelogramos.

Se llaman *diagonales* del paralelepípedo a los segmentos que unen los vértices de éste, no situados en la misma cara.

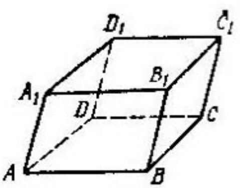


Fig. 6.118.

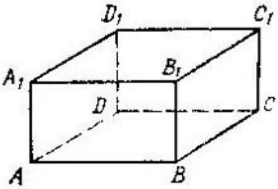


Fig. 6.119.

Propiedades del paralelepipedo:

1) El punto medio de la diagonal del paralelepípedo es su centro de simetría.

2) Las caras opuestas del paralelepípedo son iguales dos a dos y paralelas.

3) Todas las cuatro diagonales del paralelepípedo se intersecan en un mismo punto, en el cual se dividen por la mitad.

Se llama paralelepípedo *recto* al que tiene las aristas laterales perpendiculares al plano de la base del paralelepípedo, (en la fig. 6.119 $ABCDA_1B_1C_1D_1$ es un paralelepípedo recto).

Se llama *rectangular* al paralelepípedo recto, cuya base es un rectángulo. Todas las caras del paralelepípedo rectangular son rectángulos. Las longitudes de tres aristas de un paralelepípedo rectangular que salen de un mismo vértice se llaman *dimensiones* del paralelepípedo rectangular.

Propiedades del paralelepípedo rectangular:

1) El cuadrado de la diagonal del paralelepípedo rectangular es igual a la suma de los cuadrados de sus tres dimensiones:

$$d^2 = a^2 + b^2 + c^2.$$

2) Todas las diagonales del paralelepípedo rectangular son iguales.

Puesto que el paralelepípedo es un caso particular del prisma, el área de la superficie y el volumen del paralelepípedo se calculan mediante las fórmulas del área de la superficie y del volumen del prisma. Además, el *volumen del paralelepípedo rectangular* se puede calcular según la fórmula

$$V = abc,$$

donde a, b, c son las tres dimensiones del paralelepípedo rectangular.

Cubo. Se llama *cubo* al paralelepípedo rectangular que tiene dimensiones iguales. Todas las caras del cubo son cuadrados iguales.

El *volumen del cubo* se calcula por la fórmula

$$V = a^3,$$

donde a es la dimensión del cubo.

§ 18. Pirámide. Pirámide truncada

Se llama *pirámide* al poliedro una de cuyas caras es un polígono cualquiera y las demás son triángulos que tienen un vértice común.

El polígono se llama *base* de la pirámide, y las demás caras (triángulos) se llaman *caras laterales* de la pirámide.

Existen pirámides triangulares, cuadrangulares, pentagonales, etc., según el tipo de polígono que se encuentra en la base. La pirámide triangular se llama también *tetraedro*. En la fig. 6.120 se muestra una pirámide cuadrangular $SABCD$, cuya base es $ABCD$ y las caras laterales son SAB, SBC, SCD, SAD.

Los lados de las caras de la pirámide se llaman *aristas*. Las aristas que pertenecen a la base de la pirámide se llaman *aristas de la base*, y todas las demás aristas se llaman *laterales*. El vértice común de todos los triángulos (caras laterales) se llama *vértice* de la pirámide (en la fig. 6.120 el punto S es el vértice de la pirámide, los segmentos SA, SB, SC, SD son las aristas laterales y los segmentos AB, BC, CD, AD, las aristas de la base).

Se llama *altura* de una pirámide al segmento de la perpendicular, trazada del vértice de la pirámide S al plano de

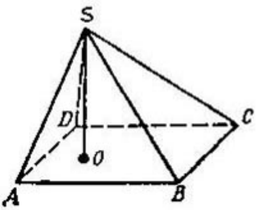


Fig. 6.120.

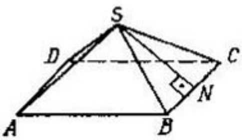


Fig. 6.121.

la base (los extremos de este segmento son el vértice de la pirámide y la base de la perpendicular). En la fig. 6.120 SO es la altura de la pirámide.

Pirámide regular. Una pirámide es *regular*, si la base es un polígono regular, y la proyección ortogonal del vértice sobre el plano de la base coincide con el centro del polígono, situado en la base de la pirámide.

Todas las aristas laterales de la pirámide regular son iguales entre sí; todas las caras laterales, son triángulos isósceles iguales.

La altura de cualquiera de las caras laterales de la pirámide regular, trazada desde su vértice, se llama *apotema* de esta pirámide. En la fig. 6.121 SN es la apotema. Todas las apotemas de una pirámide regular son iguales entre sí.

El *área de la superficie lateral de una pirámide* es igual a la suma de las áreas de los triángulos de las caras laterales, y el *área de la superficie total* es igual a la suma del área de la superficie lateral y del área de la base.

El *área de la superficie lateral de una pirámide regular* se calcula mediante la fórmula

$$S=\frac{1}{2}Ph,$$

donde P es el perímetro de la base de la pirámide, h es la apotema.

El *volumen de una pirámide* se calcula por la fórmula

$$V=\frac{1}{3}SH,$$

donde S es el área de la base de la pirámide y H, su altura.

Pirámide truncada. Tomemos una pirámide arbitraria. Por un punto perteneciente a cualquier arista lateral y que no coincide con sus extremos trazamos un plano, paralelo al plano de la base (fig. 6.122). El plano trazado cortará la pirámide en dos partes: en la pirámide $SA_1B_1C_1D_1$ (en la fig. 6.122 la que está situada más arriba del plano de la sección) y en la pirámide truncada (en la fig. 6.122 la pirámide truncada está situada más abajo del plano de la sección).

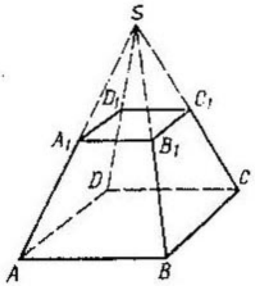


Fig. 6.122.

Se llama *poliedro truncado* al que tiene por vértices los vértices de la base de la pirámide y los vértices de la base de la pirámide cortada.

Las bases de la pirámide truncada son polígonos homotéticos (en la fig. 6.122 los cuadriláteros $ABCD$ y $A_1B_1C_1D_1$ son bases de la pirámide truncada). El centro de homotecia es el vértice de la pirámide. Se llama *altura* de la pirámide truncada a la perpendicular al plano de las bases que tiene los extremos en los planos de las bases de la pirámide. Las caras laterales de una pirámide truncada son trapecios.

Una pirámide truncada se llama *regular*, si es una parte de la pirámide regular. Las caras laterales de una pirámide truncada regular son trapecios isósceles iguales. La altura de cada una de estos trapecios se llama *apotema* de la pirámide truncada regular.

El *área de la superficie lateral* de una pirámide truncada regular se calcula por la fórmula

$$S=\frac{1}{2}(P+p)h,$$

donde P y p son los perímetros de las bases de la pirámide, h es la apotema.

El *área de la superficie total* de una pirámide truncada es igual a la suma de las áreas de las bases y del área de la superficie lateral de la pirámide.

El *volumen de una pirámide truncada* se calcula por la fórmula

$$V=\frac{1}{3}H\left(S_1+\sqrt{S_1S_2}+S_2\right),$$

donde H es la altura de la pirámide truncada, S_1 y S_2 son las áreas de las bases.

§ 19. Poliedros regulares

Un poliedro se llama *regular*, si todas sus caras son polígonos regulares, y todos sus ángulos poliedros tienen el mismo número de caras.

Todas las aristas de un poliedro regular son segmentos iguales, todos los ángulos planos de un poliedro regular también son iguales.

Existen cinco tipos distintos de poliedros regulares (convexos): el cubo, tetraedro regular, octaedro regular, dodecaedro regular, icosaedro regular. En la fig. 6.123—6.127 se muestran los poliedros regulares enumerados, así como las desarrollantes de sus superficies.

En lo posterior vamos a usar las denotaciones siguientes: V es el volumen; S es el área de la superficie; R es el radio de la esfera circunscrita; r es el radio de la esfera inscrita; H es la altura (para aquellos poliedros, para los cuales este concepto tiene sentido); con a denotamos cada una de las aristas iguales.

El cubo. Todas las seis caras son cuadrados iguales. El cubo tiene ocho vértices y doce aristas.

$$V=a^3,$$

$$S=6a^2,$$

$$R = \frac{a\sqrt{3}}{2},$$

$$r = \frac{a}{2}.$$

El tetraedro. Todas las cuatro caras son triángulos equiláteros iguales. El tetraedro tiene cuatro vértices y seis aristas.

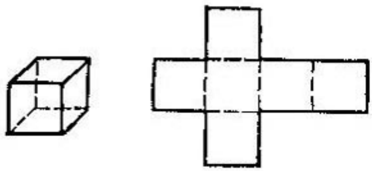


Fig. 6.123.

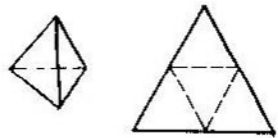


Fig. 6.124.

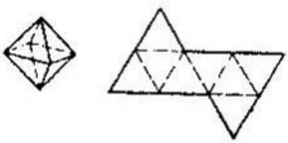


Fig. 6.125.

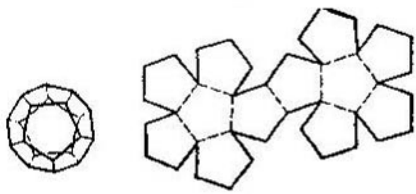


Fig. 6.126.

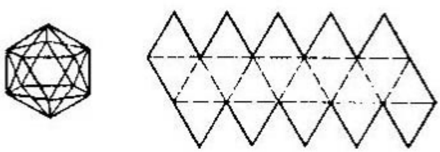


Fig. 6.127.

$$V = \frac{a^3\sqrt{2}}{12},$$

$$S = a^2\sqrt{3},$$

$$R = \frac{a\sqrt{6}}{4},$$

$$r = \frac{a\sqrt{6}}{12},$$

$$H = \frac{a\sqrt{6}}{3}.$$

El octaedro. Todas las ocho caras son triángulos equiláteros iguales. El octaedro tiene seis vértices y doce aristas.

$$V = \frac{a^3\sqrt{2}}{3},$$

$$S = 2a^2\sqrt{3},$$

$$R = \frac{a\sqrt{2}}{2},$$

$$r = \frac{a\sqrt{6}}{6}.$$

El dodecaedro. Todas las doce caras son pentágonos regulares iguales. Un dodecaedro tiene veinte vértices y treinta aristas.

$$V = \frac{a^3(15+7\sqrt{5})}{4},$$

$$S = 3a^2\sqrt{5(5+2\sqrt{5})},$$

$$R = \frac{a\sqrt{3}(1+\sqrt{5})}{4},$$

$$r = \frac{a\sqrt{10(25+11\sqrt{5})}}{20}.$$

El icosaedro. Todas las veinte caras son triángulos equiláteros iguales. El icosaedro tiene doce vértices y treinta aristas.

$$V = \frac{5a^3(3+\sqrt{5})}{12},$$

$$S = 5a^2\sqrt{3},$$

$$R = \frac{a\sqrt{2(5+\sqrt{5})}}{4},$$

$$r = \frac{a\sqrt{3}(3+\sqrt{5})}{12}.$$

El número de aristas, el número de vértices y el número de caras de los poliedros regulares se relacionan entre sí por la fórmula de Euler*)

$$N - L + F = 2,$$

donde N es el número de vértices, F es el número de las caras y L, el número de aristas.

§ 20. Figuras de rotación

Supongamos que en el espacio se da una recta l y se elige cierto punto arbitrario M que no pertenece a la recta l. Construimos un plano α perpendicular a la recta l y que contiene el punto M (fig. 6.128). El pleno α interseca la recta l

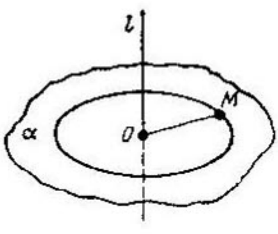


Fig. 6.128.

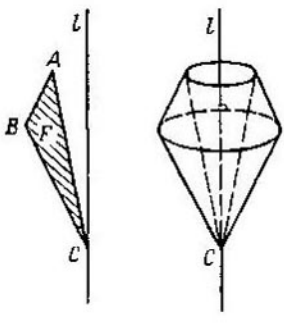


Fig. 6.129.

en cierto punto O. Examinemos una circunferencia de radio OM con centro en el punto O, perteneciente al plano α. De tal circunferencia se dice que se obtiene por rotación del punto M alrededor del eje l.

Examinemos ahora en el espacio la recta l y cierta figura F, perteneciente al plano que pasa por la recta dada l (fig. 6.129). A cada punto de la figura F, no perteneciente a la recta l, le asignamos una circunferencia que se obtiene

*) La fórmula de Euler es válida no sólo para los poliedros regulares, sino también en general para todos los poliedros convexos (por ejemplo, para el prisma, la pirámide, etc.).

como resultado de la rotación del punto dado alrededor de la recta l.

La unión de todas tales circunferencias y puntos de la figura F, pertenecientes a la recta l, se llama *figura de rotación*, obtenida como resultado de la rotación de la figura F alrededor de la recta l, la cual se llama *eje de rotación*.

Se llama *sección* de una figura de rotación al corte no vacío de esta figura y el plano.

La sección de una figura de rotación se llama sección *axial*, si representa un corte de la figura de rotación por el plano que pasa por el eje de rotación de esta figura.

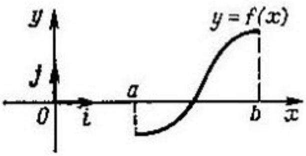


Fig. 6.130.

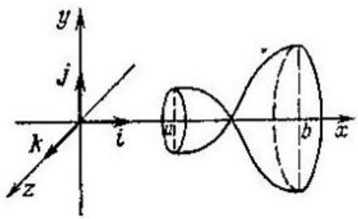


Fig. 6.131.

Examinemos en un plano de coordenadas Oxy un trapecio curvilíneo, limitado por una gráfica de una función continua $y = f(x)$, el eje Ox y las rectas $x = a$ y $x = b$ (fig. 6.130). Girando el trapecio curvilíneo alrededor del eje Ox, obtenemos una figura de rotación (fig. 6.131) cuyo volumen se calcula por la fórmula

$$V = \pi \int_a^b f^2(x)\, dx.$$

§ 21. Cilindro

Se llama *cilindro circular recto* (o simplemente *cilindro*) a la figura engendrada por un rectángulo que gira alrededor del eje que pasa por uno de sus lados.

Se llama *superficie* cilíndrica (fig. 6.132) a la engendrada por rotación alrededor del mismo eje de una quebrada, compuesta por lados de un rectángulo que no están situados en el eje de rotación.

Se llaman *bases* del cilindro a los círculos obtenidos por rotación de los lados, adyacentes al lado que pertenece al eje de rotación (en la fig. 6.132 las bases del cilindro se obtienen como resultado de la rotación de los lados del rectángulo

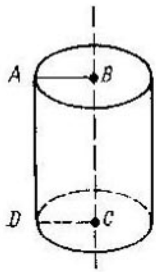


Fig. 6.132.

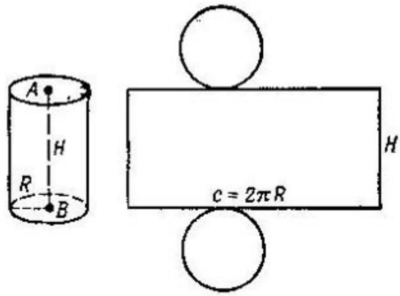


Fig. 6.133.

AB y DC alrededor del eje BC). El radio de estos dos círculos iguales se llama *radio* de la base del cilindro. En la fig. 6.132 el segmento AB es el radio de la base.

Se llama *superficie lateral* de un cilindro a la figura, generada por rotación de un lado del rectángulo que no es adyacente al lado perteneciente al eje de rotación. En la fig. 6.132 la superficie lateral del cilindro se obtiene por rotación del lado AD alrededor del eje BC. El segmento AD se llama *generatriz* del cilindro.

Se llama *altura* de un cilindro a la perpendicular a los planos de las bases del cilindro, cuyos extremos coinciden con los centros de las bases de este cilindro. En la fig. 6.132 el segmento BC es la altura del cilindro.

La *desarrollante* de un cilindro representa un rectángulo, un lado del cual es igual a la longitud de la circunferencia de la base del cilindro, y el otro a la altura del cilindro, y dos círculos con radios, son iguales al radio del cilindro (fig. 6.133).

El *volumen del cilindro* se calcula por la fórmula

$$V = \pi R^2 H,$$

donde R es el radio de la base del cilindro, H es la altura del cilindro.

Por *área de la superficie lateral* de un cilindro se toma el límite de la relación entre el incremento del volumen de un cilindro y el incremento del radio de la base del cilindro, cuando el incremento del radio tiende a cero:

$$S_{\text{lat. cil}} = \lim_{\Delta R \to 0} \frac{\Delta V_{\text{cil}}}{\Delta R}$$

El *área de la superficie lateral y total* de un cilindro se calcula mediante las fórmulas

$$S_{\text{lat.}} = 2\pi RH,$$
$$S_{\text{cil.}} = 2\pi RH + 2\pi R^2,$$

donde R es el radio de la base del cilindro y H es la altura del cilindro.

Las fórmulas para calcular el volumen y el área de la superficie lateral de un cilindro pueden ser obtenidas también sin utilizar los conceptos de la integral definida y de la derivada. Tomemos un cilindro circular recto e inscribamos en las bases del cilindro dos n-polígonos regulares de tal manera, que para la proyección ortogonal sobre el plano de la base inferior del cilindro la proyección del n-polígono superior coincida con el n-polígono inferior. Construyamos un prisma, cuyas bases son n-polígonos, que se encuentran en las bases del cilindro (de tal prisma se dice que está inscrito en el cilindro).

Por *volumen del cilindro* se toma el límite de una sucesión de volúmenes de los prismas regulares, inscritos en el cilindro, para el incremento infinito del número de lados de los polígonos regulares situados en la base del prisma.

El *área de un n-polígono regular* s_n se expresa por el radio de la circunferencia circunscrita R, según la fórmula

$$s_n = R^2 \cdot \frac{n}{2} \cdot \operatorname{sen} \frac{2\pi}{n}.$$

El *volumen de un prisma regular* V_n, en las bases del cual se encuentran unos n-polígonos regulares, es igual

$$V_n = R^2 \cdot H \cdot \frac{n}{2} \cdot \operatorname{sen} \frac{2\pi}{n}$$

(aquí H es la altura del prisma).

El *volumen del cilindro* es:

$$V_{\text{cil}} = \lim_{n\to\infty} V_n = \pi R^2 H.$$

Por *área de la superficie lateral* de un cilindro se toma el límite de la sucesión de las áreas de las superficies laterales de los prismas regulares, inscritos en el cilindro, para un aumento infinito del número de los lados de los polígonos regulares, situados en la base de los prismas.

El área de la superficie lateral de un prisma de altura H, en cuya base se encuentra un n-polígono regular es:

$$S_n = 2RH\cdot n\cdot \operatorname{sen}\frac{\pi}{n},$$

$$\lim_{n\to\infty} S_n = 2\pi RH = S_{\text{lat. cil.}}.$$

§ 22. Cono. Cono truncado

Se llama *cono circular recto* (o simplemente *cono*) a la figura engendrada por rotación de un triángulo rectangular alrededor de un eje que contiene su cateto.

Se llama *superficie* de un cono a la figura engendrada por rotación alrededor del mismo eje de una quebrada, compuesta por una hipotenusa y un cateto que no pertenece al eje de rotación. La figura, obtenida por rotación de la hipotenusa, se llama *superficie lateral* del cono, y la figura (el círculo), obtenida por rotación de un cateto, se llama *base* del cono (fig. 6.134). El radio de este círculo se llama *radio* de la base del cilindro (en la fig. 6.134, el segmento OA).

Se llama *altura* del cono al cateto del triángulo, perteneciente al eje (en la fig. 6.134, el segmento SO es la altura del cono). La *hipotenusa* de un triángulo rectangular se llama generatriz del cono (en la fig. 6.134, segmento AS es la generatriz del cono).

La *desarrollante* de la superficie lateral de un cono es un sector circular, y la *desarrolante total* de la superficie del cono representa un sector circular y un círculo (fig. 6.135).

El *volumen del cono* se calcula mediante la fórmula

$$V_{\text{con}} = \frac{1}{3}\pi R^2 H,$$

donde R es el radio de la base del cono y H, la altura del cono.

Por área de la superficie lateral del cono se toma el límite de la relación entre el incremento del volumen del cono y el incremento de la normal mayor de la generatriz del

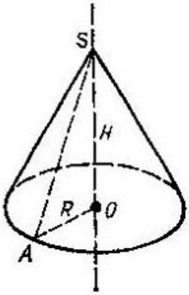


Fig. 6.134.

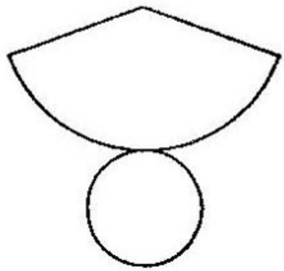


Fig. 6.135.

cono, cuando el incremento de esta normal tiende a cero:

$$S_{\text{lat. con.}} = \lim_{\Delta p \to 0} \frac{\Delta V_k}{\Delta p}$$

(La longitud del mayor de los segmentos, perpendiculares a la generatriz del cono, que tiene los extremos en la generatriz y en el eje del cono, se llama *normal* mayor de la generatriz de un cono).

El *área de la superficie lateral* de un cono se calcula por la fórmula

$$S_{\text{lat}} = \pi R L,$$

donde R es el radio de la base del cono y L, la generatriz del cono ($L = |AS|$ en la fig. 6.134).

Las fórmulas para calcular el volumen y el área de la superficie lateral de un cono pueden ser deducidas también sin aplicar los conceptos de la integral definida y de la derivada. Inscribamos en la base del cono un n-polígono y construyamos una pirámide regular, cuya base es un n-polígono regular, y el vértice coincide con el vértice del cono (sobre tal pirámide se dice que está inscrita en el cono).

Por *volumen de un cono* se toma el límite de la sucesión de las pirámides regulares, inscritas en el cono, para un incremento infinito del número de los lados de un polígono regular, que es la base de la pirámide.

Por *área de la superficie* lateral de un cono se toma el límite de la sucesión de las áreas de las superficies laterales de las

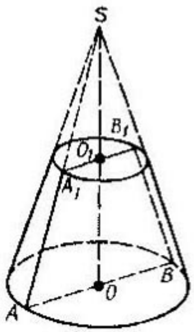


Fig. 6.136.

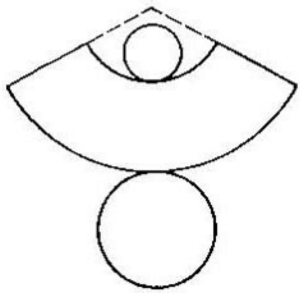


Fig. 6.137.

pirámides regulares, inscritas en el cono, para un incremento infinito del número de lados de un polígono regular situado en las bases de las pirámides.

La deducción de las fórmulas para calcular el volumen y el área de la superficie lateral de un cono es análoga a la deducción de las respectivas fórmulas para el cilindro (véase § 21).

El *área de la superficie total de un cono* se calcula por la fórmula

$$S_{\text{con}} = \pi RL + \pi R^2.$$

Cono truncado. Se llama *cono truncado* a la parte del cono que está limitada por su base y sección, paralelas al plano de la base.

Las bases del cono truncado son círculos homotéticos con centro de homotecia en el vértice del cono (en la fig. 6.136 en el punto S).

Un cono truncado se puede obtener como resultado de la rotación de un trapecio isósceles alrededor de su eje de simetría (en la fig. 6.136, del trapecio AA_1B_1B). Al girar la frontera de este trapecio se obtiene la superficie de un cono truncado.

Se llama *generatriz* del cono truncado al lado lateral del trapecio; los círculos, obtenidos por rotación de las bases del trapecio se llaman *bases* del cono truncado.

La *desarrollante* de un cono truncado representa la unión de una parte del anillo circular y de dos círculos (fig. 6.137).

El *volumen del cono truncado* se calcula por la fórmula

$$V=\frac{1}{3}\pi H\left(R_1^2+R_1R_2+R_2^2\right),$$

donde H es la altura del cono truncado (en la fig. 6.136 $H=|OO_1|$), R_1 y R_2 son los radios de las bases superior e inferior del cono truncado (en la fig. 6.136 $R_1=|OA|$, $R_2=|O_1A_1|$).

El área de la superficie lateral de un cono truncado es igual a la diferencia de las superficies laterales de un cono perfecto y un cono cortado por un plano, paralelo a la base del cono.

El *área de la superficie lateral de un cono truncado se* calcula según la fórmula

$$S_{\text{lat}}=\pi\left(R_1+R_2\right)L,$$

donde L es la generatriz del cono truncado (en la fig. 6.136 $L=|AA_1|$).

§ 23. La esfera y el cuerpo esférico

Se llama *esfera* al conjunto de todos los puntos del espacio que se encuentran a una distancia positiva dada R del punto dado del espacio O.

El punto dado O se llama *centro* de la esfera (fig. 6.138).

El segmento OM (M es un punto arbitrario de la esfera) se llama *radio* de la esfera ($OM=R$). El segmento que une dos puntos cualesquiera de la esfera se llama su *cuerda*. La cuerda que pasa por el centro de la esfera se llama *diámetro* de la esfera. El diámetro de la esfera es igual a su radio doble.

El conjunto de todos los puntos del espacio que se encuentran del punto dado O a una distancia no mayor que la distancia dada R, se llama cuerpo esférico (esfera).

El cuerpo esférico se puede obtener por rotación de un semicírculo alrededor de un eje que contiene el diámetro del semicírculo (véase fig. 6.138). La figura, obtenida por rotación de una semicircunferencia, es una esfera, es decir, es la superficie del cuerpo esférico. El centro, el radio y las

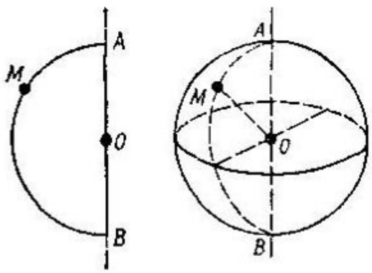


Fig. 6.138.

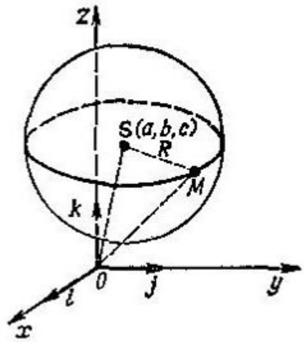


Fig. 6.139.

cuerdas de esta esfera se llaman respectivamente el *centro*, *radio* y *cuerdas* del cuerpo esférico. Todos los puntos de éste, que no pertenecen a su superficie, se llaman *puntos interiores* del cuerpo esférico.

En el sistema de coordenadas rectangular espacial $Oxyz$ la *ecuación de la esfera* de radio R con centro en el punto S $(a; b; c)$ tiene la forma (fig. 6.139)

$$(x-a)^2+(y-b)^2+(z-c)^2=R^2.$$

El *volumen del cuerpo esférico de* radio R se calcula por la fórmula

$$V=\frac{4}{3}\pi R^3.$$

Por *área de una esfera* (área de la superficie del cuerpo esférico) se toma el límite de la relación entre el incremento del volumen del globo, limitado por esta esfera y el incremento del radio, cuando este último tiende a cero:

$$S=\lim_{\Delta R\to 0}\frac{\Delta V}{\Delta R}.$$

El cuerpo esférico se puede obtener por rotación de un semicírculo alrededor de un eje que contiene el diámetro del semicírculo (véase fig. 6.138). La figura, obtenida por rotación de una semicircunferencia, es una esfera, es decir, es la superficie del cuerpo esférico. El centro, el radio y las

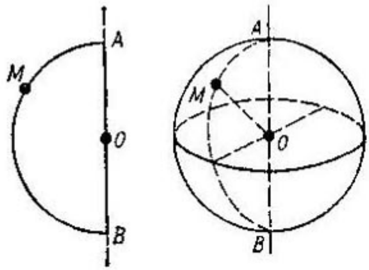


Fig. 6.138.

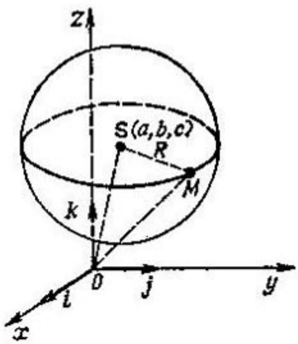


Fig. 6.139.

cuerdas de esta esfera se llaman respectivamente el *centro*, *radio* y *cuerdas* del cuerpo esférico. Todos los puntos de éste, que no pertenecen a su superficie, se llaman *puntos interiores* del cuerpo esférico.

En el sistema de coordenadas rectangular espacial $Oxyz$ la *ecuación de la esfera* de radio R con centro en el punto S $(a; b; c)$ tiene la forma (fig. 6.139)

$$(x-a)^2+(y-b)^2+(z-c)^2=R^2.$$

El *volumen del cuerpo esférico de* radio R se calcula por la fórmula

$$V=\frac{4}{3}\pi R^3.$$

Por *área de una esfera* (área de la superficie del cuerpo esférico) se toma el límite de la relación entre el incremento del volumen del globo, limitado por esta esfera y el incremento del radio, cuando este último tiende a cero:

$$S=\lim_{\Delta R\to 0}\frac{\Delta V}{\Delta R}.$$

El *área de una esfera* de radio R se calcula según la fórmula

$$S = 4\pi R^2.$$

Las fórmulas para calcular el volumen y el área de la superficie esférica, también pueden ser obtenidas, en general, como límites de las sucesiones de los volúmenes y áreas de figuras, inscritas en el cuerpo esférico. Sin embargo, (en comparación con la deducción de las fórmulas respectivas para el cilindro y cuerpo esférico) esta deducción es mucho más complicada y no la citaremos.

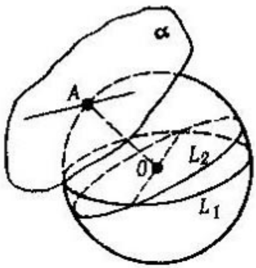


Fig. 6.140.

La *sección* de una esfera por un plano es:

1) una circunferencia, si la distancia del centro de la esfera al plano de sección es menor que el radio de la esfera;

2) un punto, si la distancia del centro de la esfera al plano de sección es igual al radio.

La intersección de la esfera por un plano que pasa por su centro se llama *círculo máximo* (mayor). El radio del círculo mayor es igual al radio de la esfera. En la fig. 6.140, L_1 y L_2 son las circunferencias de los círculos máximos.

Cualquier par de círculos máximos se interseca por el diámetro de la esfera, que es también diámetro de cada uno de los círculos que se intersecan. Por dos puntos de la esfera que se encuentran en los extremos de un mismo diámetro se puede trazar un conjunto infinito de círculos mayores. Por dos puntos que no están situados en los extremos de un diámetro de la esfera se puede trazar uno y solamente un círculo mayor.

El plano que tiene con la esfera un solo punto común se llama *plano tangente* a la esfera. Este punto (el punto A en la fig. 6.140) se llama *punto de tangencia* de la esfera y del plano. Para que el plano sea tangente a la esfera, es necesario y suficiente que este plano sea perpendicular al radio de la esfera y pase por su extremo. En la fig. 6.140 el plano tangente α es perpendicular al radio OA.

Una recta, perteneciente al plano tangente a la esfera y que pasa por el punto de tangencia (por el punto A en la fig. 6.140) se llama recta *tangente* a la esfera (cuerpo esférico). Por cada punto, perteneciente a la esfera, se puede trazar tantas como se quieran tangentes.

§ 24. Partes de la esfera

24.1. Segmento esférico. Se llama *segmento esférico* a la figura, obtenida por rotación de un segmento circular alrededor de un diámetro perpendicular a su cuerda (en la fig. 6.141 AB es la cuerda, CD, el diámetro).

La figura, obtenida por rotación del arco de un segmento circular, se llama *superficie de segmento*, y la figura, obtenida como resultado de la rotación de una cuerda, *base* del

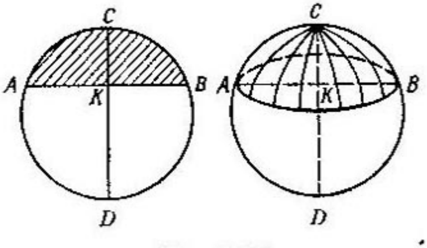


Fig. 6.141.

segmento esférico. El segmento del diametro, perteneciente al mismo tiempo al eje de rotación y al segmento circular (en la fig. 6.141, el segmento KC), se llama *altura* del segmento circular (así como, altura de la superficie de segmento).

El *volumen de un segmento esférico* de una esfera de radio R y altura H se calcula mediante la fórmula

$$V = \frac{1}{3}\pi H^2 (3R - H).$$

El *área de la superficie de segmento* se calcula por la fórmula

$$S = 2\pi RH.$$

24.2. Sector esférico. La figura, obtenida por rotación de un sector circular alrededor de un eje que pasa por uno de sus radios (fig. 6.142), se llama *sector esférico*. El arco del sector circular forma para esta rotación una superficie de segmento.

El *volumen de un sector esférico* se calcula por la fórmula

$$V=\frac{2}{3}\pi R^2H,$$

donde R es el radio de la esfera y H, la altura del segmento esférico (en la fig. 6.142 $H = |KC|$).

El área de la superficie total de un sector esférico puede ser obtenido como la suma del área de la superficie del segmento esférico y del área de la superficie lateral del cono,

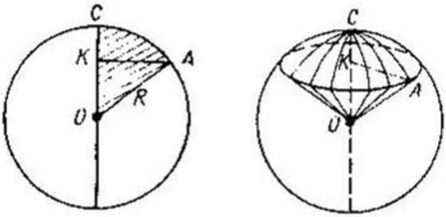


Fig. 6.142.

obtenido por rotación de un triángulo rectangular OKA alrededor del eje CO que contiene su cateto KO (véase fig. 6.142).

El *área de la superficie total* de un sector esférico se calcula según la fórmula

$$S_{\text{sect. esf.}} = S_{\text{segm}} + S_{\text{con. lat.}} =$$

$$= 2\pi RH + \pi R\sqrt{2RH-H^2} = \pi R\left(2H+\sqrt{2RH-H^2}\right).$$

24.3. Capa esférica. La figura, obtenida por rotación de una parte del círculo, comprendida entre dos cuerdas paralelas, alrededor de un eje que pasa por el diámetro, perpendicular a las cuerdas, se llama *capa esférica* (fig. 6.143).

Los círculos, obtenidos por rotación de las cuerdas, se llaman *bases* de la capa esférica. Los radios de estos círculos (en la fig. 6.143, los segmentos AD y BC) se llaman *radios* de la capa esférica. El segmento del diámetro, perteneciente al

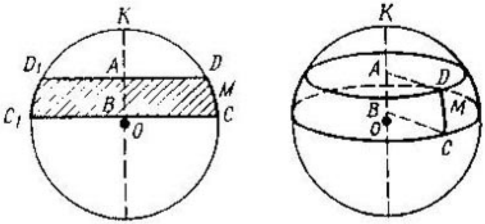


Fig. 6.143.

mismo tiempo al eje de rotación y a la capa esférica (en la fig. 6.143, el segmento AB), se llama *altura* de la capa esférica.

El volumen de la capa esférica se calcula mediante la fórmula

$$V = \frac{1}{6}\pi H^3 + \frac{1}{2}\pi (r_1^2 + r_2^2) H,$$

donde r_1 y r_2 son los radios de la capa esférica (en la fig. 6.143 $|AD| = r_1$, $|BC| = r_2$) y H es la altura de la capa esférica (en la fig. 6.143 $H = |AB|$).

24.4. Zona esférica. Se llama *zona esférica* a la figura, obtenida por rotación del arco de una circunferencia, comprendida entre las dos cuerdas paralelas, alrededor del eje que pasa por el diámetro, perpendicular a las cuerdas. En la fig. 6.143 la zona esférica está obtenida como resultado de la rotación del arco DMC alrededor del eje AB.

La altura de la capa esférica es al mismo tiempo la altura de la zona esférica respectiva.

El área de la zona esférica representa la diferencia de las áreas de dos superficies de segmento (en la fig. 6.143, la diferencia de las superficies de segmento, formadas como resultado de la rotación de los arcos C_1KC y D_1KD).

El *área de la zona esférica* se calcula mediante la fórmula

$$S_{z.esf} = 2\pi RH,$$

donde H es la altura de la zona esférica, R es el radio del arco, de cuya rotación se obtiene la zona esférica (el radio de la circunferencia que contiene un arco dado se llama radio del arco).

§ 25. Transformaciones del plano y del espacio

25.1. Aplicación de una figura en una figura, y aplicación de una figura sobre una figura. Sea Φ y Φ_1 dos figuras geométricas. Si a cada punto M de la figura Φ se le asigna de cierto modo un solo punto M_1 de la figura Φ_1, entonces tal correspondencia se llama *aplicación de la figura Φ en la figura* Φ_1. En este caso el punto M_1 se llama *imagen* del punto M, y el punto M, *preimagen* del punto M_1. Si para la aplicación elegida de la figura Φ en la figura Φ_1 ciertos puntos de la figura Φ_1 tienen varias preimágenes, entonces el conjunto de todas las preimágenes del punto $M_1 \in \Phi_1$ se llama *preimagen completa* del punto M_1.

El conjunto Φ' de todas las imágenes de los puntos de la figura Φ se llama imagen de la figura Φ.

Si para la aplicación elegida f la imagen de la figura Φ coincide con la figura Φ_1, es decir, $\Phi' = \Phi_1$, entonces se dice que la figura Φ se aplica sobre la figura Φ_1, y se escribe

$$\Phi \xrightarrow{f} \Phi_1$$

La aplicación de la figura Φ sobre la figura Φ_1 se llama *invertible*, si cada punto M_1 de la figura Φ_1 tiene una sola preimagen. Para una aplicación invertible existe una aplicación inversa, la cual aplica la figura Φ_1 sobre la figura Φ.

Si f es la aplicación invertible de la figura Φ sobre la figura Φ_1, entonces la aplicación inversa a ella se denota por el símbolo f^{-1} y se escribe:

$$\Phi_1 \xrightarrow{f^{-1}} \Phi.$$

La aplicación invertible de una figura sobre otra también se llama aplicación *biyectiva*.

25.2. Transformación del plano y del espacio. La aplicación biyectiva del espacio (plano) sobre sí se llama *transformación* del espacio (plano).

Las transformaciones más simples del espacio (plano) son:

1) Las transformaciones del espacio (plano) que conservan las distancias. Tales transformaciones se llaman *isometría**).

2) Las transformaciones del espacio (plano), para las cuales las distancias entre dos puntos cualesquiera se cambian en una misma relación $k > 0$. Tales transformaciones del espacio (plano) se llaman transformaciones *de semejanza.*

La partición dada de las transformaciones en dos tipos es muy condicional, pues la isometría es un caso particular de la transformación de semejanza; la transformación de semejanza es una isometría, si $k = 1$.

Designemos la transformación arbitraria del espacio (plano) por f. Si la transformación f aplica el punto M del espacio (plano) sobre el punto M_1 del espacio (plano), entonces se escribe

$$f(M) = M_1.$$

El punto M_1 se llama *imagen* del punto M para la transformación del espacio (plano) f, y el punto M es la *preimagen* del punto M_1.

La transformación del espacio (plano) f se puede considerar como un conjunto de todos los pares ordenados de los puntos $(M;\ M_1)$ del espacio (plano) donde $M_1 = f(M)$. El conjunto de todos los pares ordenados de los puntos del espacio (plano) $(M_1;\ M)$ define la transformación inversa que aplica M_1 en M. Tal transformación se llama transformación *inversa* respecto a la transformación f y se designa con el símbolo f^{-1}.

Composición de las transformaciones. Sea f_1 y f_2 dos transformaciones sucesivas del espacio (plano):

$$M_1 = f_1(M) \text{ y } M_2 = f_2(M_1).$$

Una transformación del espacio (plano) que aplica el punto M del espacio (plano) en el punto M_2 del espacio (plano) se llama *composición* de las transformaciones f_1 y f_2

---

*) En el curso del programa escolar en vez del término «isometría» se usa el término «desplazamiento».

y se denota

$$f_2 \circ f_1.$$

Transformación idéntica. Una transformación del espacio (plano) que aplica cada punto del espacio (plano) sobre sí se llama transformación *idéntica* del espacio (plano). La transformación idéntica se designa con la letra E. Según la definición de la transformación idéntica del espacio (plano) tenemos que

$$E(M) = M.$$

La transformación E puede considerarse como un conjunto de todos los pares de los puntos coincidentes del espacio (plano).

25.3. Isometría del espacio y del plano. Igualdad de figuras. De todas las transformaciones del espacio (plano) se pueden destacar las transformaciones, para las cuales la distancia entre dos puntos cualesquiera del espacio (plano) es igual a la distancia entre sus imágenes; es decir, si $f(A) = A_1$ y $f(B) = B_1$, donde A y B son dos puntos cualesquiera del espacio (plano), entonces $|AB| = |A_1B_1|$. Sobre tales transformaciones del espacio (plano) se dice que conservan la distancia.

La transformación del espacio que conserva la distancia se llama *isometría*. De la definición de la isometría se desprende:

1) La transformación idéntica es una isometría.

2) La transformación inversa a la isometría es una isometría.

3) La composición de dos isometrías es una isometría.

Cualquiera isometría del plano puede ser representada como el cumplimiento sucesivo de tres transformaciones: la traslación paralela, el giro alrededor de un punto y la simetría axial.

Dos figuras geométricas Φ y Φ_1 son iguales, si existe isometría que aplica la figura Φ sobre la figura Φ_1*).

25.4. Rotación de un plano alrededor de un punto. Se llama *rotación* de un plano alrededor del centro O tal isometría del plano, para la cual:

*) En el curso de geometría de secundaria básica en vez del término «igualdad» se usa el término «congruencia».

1) el punto O se aplica sobre sí;

2) el ángulo entre cualquier rayo OX y el rayo OX_1 que le corresponde (la imagen del rayo OX) tiene la misma magnitud α, llamada *ángulo de rotación*.

Si el rayo OX coincide con el rayo OX_1 mediante un giro en sentido antihorario, entonces la dirección de rotación se considera *positiva*; si el rayo OX coincide con el rayo OX_1 mediante un giro en sentido horario, entonces la dirección

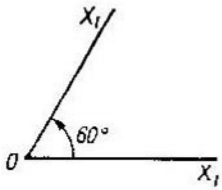


Fig. 6.144.

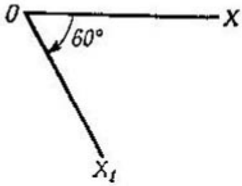


Fig. 6.145.

de rotación se considera *negativa*. Así, por ejemplo, en la fig. 6.144 los rayos OX y OX_1 dan un giro en 60°, y en la fig. 6.145 los rayos OX y OX_1 dan un giro en −60°.

El giro alrededor del centro O con un ángulo de rotación α se designa R_0^α, donde $\alpha \in (-180°; 180°]$.

El giro en un ángulo $\beta = \alpha + 360° \cdot n$, donde $n \in Z$ y $\alpha \in (-180°; 180°]$, se identifica con el giro R_0^α.

Puesto que el giro de un plano alrededor de un punto es un caso particular de isometría, entonces para él son válidas todas las propiedades generales de isometría:

1) La aplicación, inversa a un giro alrededor del centro O, también es un giro alrededor del centro O. Así, por ejemplo, para un giro en sentido antihorario en un ángulo $\alpha = 30°$ es inverso el giro en 30° también, pero en sentido horario.

2) Una composición de dos giros alrededor del centro O en los ángulos α_1 y α_2 es un giro en un ángulo $\alpha_1 + \alpha_2$: $R_0^{\alpha_1} \circ R_0^{\alpha_2} = R_0^{\alpha_1+\alpha_2}$, además, la composición de dos giros tiene propiedad de conmutatividad:

$$R_0^{\alpha_1} \circ R_0^{\alpha_2} = R_0^{\alpha_2} \circ R_0^{\alpha_1}.$$

25.5. Simetría central y figuras simétricas centrales. Se llama *simetría central* con centro O a un caso particular de rotación de un plano alrededor del centro O, precisamente al giro en 180°.

La simetría central con centro O se denota con el símbolo Z_0. En vigor de la definición de la simetría central tenemos que

$$Z_0 = R_0^{180°}.$$

Se llama *simétrica central* a la figura geométrica que para una simetría central con centro O se aplica sobre sí misma

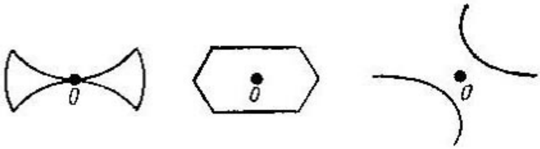


Fig. 6.146.

(se dice también que tal figura tiene un *centro de simetría*, O). En la fig. 6.146 se muestran varias figuras simétricas centrales.

25.6. Simetría axial de un plano. Se dice que los puntos M_1 y M_2 del plano α son simétricos respecto a la recta l perteneciente al plano α y que no contiene los puntos M_1 y M_2, si el segmento M_1M_2 es perpendicular a la recta l y se divide por esta recta por la mitad ($|M_1O| = |OM_2|$), (fig. 6.147).

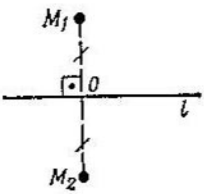


Fig. 6.147.

Cualquier punto de la recta l se considera simétrico a sí mismo.

La isometría de un plano, para la cual cada punto del plano se aplica sobre un punto simétrico a ella (respecto a la recta dada) se llama *simetría axial* del plano.

La recta dada l se llama *eje de simetría*, y la simetría axial respecto a la recta dada l se denota con el símbolo S_l.

Puesto que la simetría axial es un caso particular de isometría, para ella son válidas todas las propiedades generales de isometría. En particular, la aplicación, inversa a la simetría axial, es una simetría axial.

Para la simetría axial es válida también una afirmación más fuerte:

La aplicación, inversa a la simetría axial, es la misma simetría axial.

Una recta l se llama *eje de simetría* de la figura Φ, si la figura Φ para una simetría axial que tiene un eje l se aplica

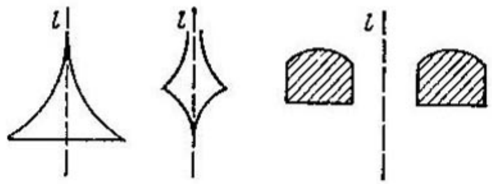


Fig. 6.148.

sobre sí misma. En este caso de la figura Φ se dice que es *simétrica* respecto al eje l.

En la fig. 6.148 se muestran varias figuras geométricas simples, simétricas respecto al eje l.

25.7. Simetría axial del espacio. Se dice que los puntos M_1 y M_2 *son simétricos* respecto a la recta l, que no contiene los puntos M_1 y M_2, si el segmento M_1M_2 es perpendicular a l y se divide por esta recta por la mitad ($|M_1O| = |OM_2|$).

Cualquier punto de la recta l se considera simétrico a sí mismo. Se llama *simetría axial* del espacio a la isometría del espacio, para la cual cada punto del espacio se aplica sobre un punto simétrico a ella (respecto a la recta dada l).

La recta dada l se llama *eje de simetría*. La simetría axial del espacio se disigna con el símbolo S_l.

La simetría axial del espacio es un caso particular de la isometría del espacio, por eso para aquélla son válidas todas las propiedades generales de ésta.

Si la figura Φ para una simetría axial con eje l se aplica sobre sí misma, se dice que la figura Φ *es simétrica* respecto al eje l, y el eje l se llama *eje de simetría* de la figura Φ.

En la fig. 6.149 están representadas dos figuras geométricas espaciales que tienen un eje de simetría l.

25.8. Simetría respecto al plano. Se dice que los puntos M y M_1 *son simétricos* respecto al plano α (fig. 6.150), si el segmento MM_1 es perpendicular a este plano y se divide

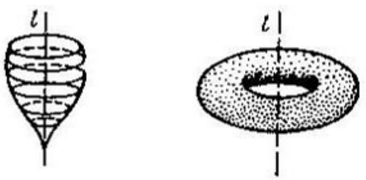


Fig. 6.149.

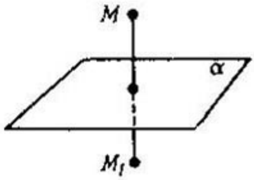


Fig. 6.150.

por él por la mitad. Cualquier punto del plano α se considera simétrico a sí mismo.

La transformación del espacio, para la cual cada punto se aplica sobre un punto simétrico a ella respecto al plano dado, se llama *simetría* respecto a este plano.

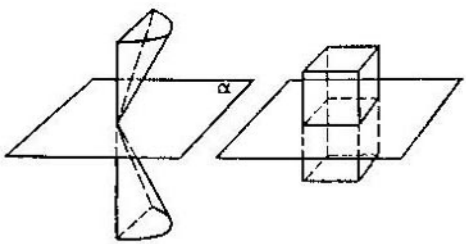


Fig. 6.151.

La simetría respecto al plano α se designa con el símbolo S_α.

El plano α se llama *plano de simetría* de la figura Φ, si para la simetría respecto al plano α la figura Φ se aplica sobre sí misma. En este caso de la figura Φ se dice que es simétrica respecto al plano α.

En la fig. 6.151 se muestran dos figuras geométricas, simétricas respecto al plano α.

25.9. Homotecia del plano. Se llama homotecia del plano con *centro* O y un *coeficiente* $k \neq 0$ a la aplicación del plano sobre sí, para la cual el punto X_1 es tal que

$$\overrightarrow{OX_1} = k\overrightarrow{OX},$$

es la imagen de un punto arbitrario X.

La homotecia con centro O y el coeficiente k se disigna por el símbolo H_0^k. Utilizando esta designación, se puede escribir la homotecia del plano como sigue:

$$H_0^k(X) = X_1, \text{ si } \overrightarrow{OX_1} = k\overrightarrow{OX}.$$

Hablando de un conjunto de homotecias con cualquier centro definido, se omiten en la designación de la homotecia la letra O, es decir, en vez de H_0^k se escribe H^k. Si hablamos de una homotecia concreta, es decir, de una homotecia con centros y coeficientes concretos, entonces se disignan simplemente con H.

Propiedades principales de homotecia:

1) El centro de homotecia se aplica sobre sí.

2) Si el coeficiente $k > 0$, entonces los puntos X y $X_1 = H(X)$ están situados en la recta OX a un lado del centro de homotecia, es decir, el punto X_1 pertenece al rayo OX.

Si $k < 0$, entonces los puntos X y $X_1 = H(X)$ se encuentran en la recta OX a distintos lados del centro de homotecia, es decir, el punto X_1 pertenece al rayo que tiene el mismo origen O que el rayo OX y dirección contraria al rayo OX.

3) Si para la homotecia con un coeficiente k los puntos X e Y se aplican respectivamente sobre los puntos X_1 e Y_1, entonces

$$\overrightarrow{X_1Y_1} = k\overrightarrow{XY}.$$

4) Para una homotecia con coeficiente 1 cada punto pasa a sí mismo. Por eso la aplicación idéntica del plano sobre sí es una homotecia con cualquier centro y un coeficiente $k = 1$.

5) Para la homotecia con centro O y coeficiente -1 cada punto X pasa al punto X_1, para el cual $\overrightarrow{OX_1} = -\overrightarrow{OX}$, es decir, pasa a un punto que es centralmente simétrico al

punto X. Por eso la homotecia con un coeficiente $k = -1$ es una simetría central:

$$H_0^{-1} = Z_0.$$

6) La homotecia con el mismo centro y coeficiente $k' = 1/k$ es una aplicación inversa a la homotecia con el coeficiente k.

7) Para una homotecia cuyo coeficiente es k todas las distancias entre los puntos se multiplican por $|k|$:

$$|X_1 Y_1| = |k| \cdot |XY|,$$

donde $X_1 = H^k(X)$, $Y_1 = H^k(Y)$.

Dos figuras planas Φ y Φ_1 se llaman *homotéticas*, si existe una homotecia que aplica la figura Φ sobre la figura Φ_1.

Las propiedades de las figuras homotéticas son:

1) Las figuras homotéticas son semejantes.

2) Para una homotecia con un coeficiente positivo cada rayo pasa a un rayo codirigido con él. Para una homotecia con un coeficiente negativo cada rayo pasa a un rayo que tiene una dirección contraria a él. En particular, para una homotecia, cualquier recta que pasa por el centro de homotecia pasa a sí misma; una recta que no pasa por el centro de homotecia (para $k \neq 1$) pasa a una recta paralela; un ángulo pasa a un ángulo igual, etc.

25.10. Homotecia del espacio. La homotecia del espacio se define de la misma manera que la homotecia del plano.

La transformación del espacio, para la cual la imagen de un punto arbitrario X es tal punto X_1, que $\overrightarrow{OX_1} = k\overrightarrow{OX}$, se llama *homotecia* del espacio con *centro* O y coeficiente $k \neq 0$.

La homotecia del espacio se designa con el mismo símbolo H_0^k que la homotecia del plano.

Todas las propiedades de homotecia del plano son también propias de la homotecia del espacio. A las propiedades de homotecia del plano citadas más arriba es necesario añadir dos *propiedades* más *de homotecia del espacio*:

1) Para la homotecia del espacio el plano se aplica o sobre sí mismo o sobre el plano paralelo.

2) Cualquier esfera para la homotecia del espacio con centro que coincide con el centro de la esfera, se aplica sobre la esfera.

La homotecia tanto del espacio como del plano se da bien por el centro de homotecia y un par ordenado de puntos respectivos X y $X_1 = H(X)$ o bien por dos pares ordenados de puntos respectivos $(X;\ X_1)$ e $(Y;\ Y_1)$, donde $X_1 = H(X)$ e $Y_1 = H(Y)$.

Además de los procedimientos de representación enumerados más arriba la homotecia puede ser representada, por ejemplo, por el centro de homotecia O y el coeficiente de homotecia k.

25.11. Transformación de la semejanza de un plano. La aplicación del plano sobre sí, para la cual las distancias entre dos puntos cualesquiera se cambian en la misma relación $k > 0$, se llama *transformación de semejanza* o simplemente *semejanza*.

El número k se llama *coeficiente* de semejanza.

Así, si f es la transformación de semejanza que cambia la distancia en k veces, y A y B son dos puntos cualesquiera del plano, y $f(A) = A_1$, $f(B) = B_1$, entonces $|A_1B_1| = k \cdot |AB|$.

Propiedades de las transformaciones de semejanza:

1) La composición $F_1 \circ F_2$ de dos transformaciones de semejanza F_1 y F_2 con los coeficientes de semejanza k_1 y k_2 respectivamente, es una transformación de semejanza con los coeficientes $k_1 \cdot k_2$. En particular, la composición de homotecia e isometría es una transformación de semejanza.

2) Cada transformación de semejanza es una composición de homotecia e isometría.

3) Cada homotecia es una transformación de semejanza.

4) Cada isometría es una transformación de semejanza con un coeficiente de semejanza igual a la unidad.

25.12. Figuras semejantes. Si la figura Φ se puede aplicar sobre la figura Φ_1 de tal manera, que para cualesquiera puntos A y B de la figura Φ la relación de la distancia $|A_1B_1|$ entre sus imágenes (los puntos de la figura Φ_1) a la distancia $|AB|$ entre los mismos puntos es igual a un mismo número k: $|A_1B_1| = k \cdot |AB|$, entonces se dice que la figura Φ_1 es *semejante* a la figura Φ con un coeficiente de semejanza k y se escribe

$$\Phi_1 \overset{k}{\sim} \Phi.$$

Propiedades de las figuras semejantes:

1) Cada figura es semejante a sí misma con un coeficiente de semejanza 1 (reflexividad):

$$\Phi \overset{1}{\sim} \Phi$$

2) Si la figura Φ_1 es semejante a la figura Φ con un coeficiente k, entonces la figura Φ es semejante a la figura Φ_1 con un coeficiente $k' = 1/k$ (simetría):

$$\Phi_1 \overset{k}{\sim} \Phi \Leftrightarrow \Phi \overset{1/k}{\sim} \Phi_1.$$

Teniendo en cuenta la propiedad de simetría, se suele decir simplemente que dos figuras Φ y Φ_1 son semejantes.

3) Si la figura Φ_1 es semejante a la figura Φ con un coeficiente k_1, y la figura Φ_2 es semejante a la figura Φ_1 con un coeficiente de semejanza k_2, entonces la figura Φ_2 es semejante a la figura Φ con un coeficiente $k = k_1 \cdot k_2$ (transitividad):

$$\Phi_1 \overset{k_1}{\sim} \Phi \ \text{y}\ \Phi_2 \overset{k_2}{\sim} \Phi_1 \Rightarrow \Phi_2 \overset{k_1 \cdot k_2}{\sim} \Phi.$$

§ 26. Sistema de axiomas y de conceptos indefinibles de geometría (según Hilbert)

Son conceptos principales indefinibles de geometría, según Hilbert, tres tipos de objetos:

1) los puntos, que designaremos $A, B, C, \ldots$;

2) las rectas, que designaremos $a, b, c, \ldots$;

3) los planos, que designaremos $\alpha, \beta, \gamma, \ldots$.

Axiomas de geometría según Hilbert

1. *Axiomas de pertenencia.*

Los axiomas de este grupo establecen las relaciones de pertenencia entre los conceptos principales de geometría.

1-1. Para dos puntos cualesquiera A y B existe una recta a que pasa por estos puntos.

1-2. Para dos puntos distintos cualesquiera A y B existe no más de una recta que pasa por estos puntos.

1-3. En una recta existen por lo menos dos puntos. Existen por lo menos tres puntos que no están situados en una recta.

1-4. Para tres puntos cualesquiera A, B y C que no están situados en una misma recta existe un plano α que contiene tres puntos dados A, B y C. Para cualquier plano siempre existe un punto, perteneciente a él.

1-5. Para tres puntos cualesquiera A, B y C que no están situados en una misma recta existe no más que un plano que pasa por estos puntos. Este plano se designa ABC, es decir, por la enumeración de los puntos, por los cuales pasa el plano.

1-6. Si dos puntos A, B de la recta a están situados en el plano α, entonces cualquier punto de la recta a se encuentra en el plano α.

1-7. Si los planos α y β tienen un punto común A, entonces ellos tienen por lo menos un punto común más B.

1-8. Existen por lo menos cuatro puntos que no están situados en un plano.

Los axiomas 1-1—1-3 se llaman axiomas de plano, y los demás axiomas, espaciales.

II. *Axiomas de orden.*

Los axiomas de este grupo definen el concepto «entre», que sirve para describir la relación del orden de los puntos en la recta, en el plano y en el espacio.

Los puntos situados en la recta se encuentran entre sí en relaciones determinadas. Para la descripción de estas relaciones se usa la palabra «entre».

II-1. Si el punto B está situado entre el punto A y el punto C, entonces A, B y C son tres puntos distintos de la recta y el punto B se encuentra también entre C y A.

II-2. Para dos puntos distintos cualesquiera A y C situados en la recta AC existe por lo menos un punto B tal, que el punto C se encuentra entre A y B.

II-3. Entre tres puntos distintos cualesquiera situados en una recta existe no más que un punto situado entre los otros dos.

Además de estos axiomas de orden, en la recta se introduce un axioma más que define la relación de orden en el plano. Antes de formular este axioma, introduzcamos un concepto nuevo, el concepto de segmento, el cual se usa en la formulación de este axioma.

Sea A y B dos puntos distintos de la recta a. El conjunto de todos los puntos de la recta a situados entre los puntos A y B, incluyendo A y B, se llama *segmento*, y los puntos A y B, extremos del segmento. El segmento con los extremos A, B se designa AB (o BA). Cualquier punto del segmento que se encuentra entre sus extremos se llama *punto interior* del segmento.

II-4. Sea A, B y C tres puntos no situados en una recta, y a, una recta en el plano ABC, que no pasa por ninguno de los puntos A, B y C (fig. 6.152). Si en este caso, la recta a pasa por uno de los puntos del segmento AB, entonces ella tiene que pasar por uno de los puntos del segmento AC o por uno de los puntos del segmento BC.

Usando el concepto de «el punto A se encuentra entre los puntos B y C», definido más arriba, se puede introducir el concepto de rayo.

Sea que en la recta a se representan cuatro puntos distintos A, B, A' y O (fig. 6.153) de tal manera que el punto O se encuentra entre A y B, pero no se encuentra entre A y A'. En este caso se dice que los puntos A y A' de la recta a están situados al mismo lado del punto O, y los puntos A, B se encuentran en la recta a distintos lados del punto O. El conjunto de todos los puntos de la recta a situados a un lado del punto O, incluyendo el punto O, se llama *rayo* que sale del punto O.

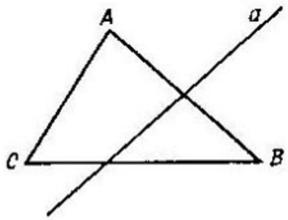


Fig. 6.152.

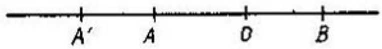


Fig. 6.153.

III. *Axiomas de congruencia.*

Los segmentos se encuentran en relación definida uno con el otro; para la descripción de esta relación sirven los conceptos de «congruencia» o «igualdad» de segmentos*).

III-1. Sea A y B, dos puntos de la recta a y A', cierto punto de la recta a' (el cual, en particular, puede coincidir con la recta a). Entonces en la recta a' existe un punto B', que se encuentra a un lado dado del

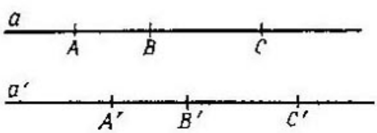


Fig. 6.154.

punto A' tal, que el segmento AB es igual al segmento $A'B'$ (fig. 6.154). Para designar la igualdad de los segmentos se usa el signo $=$.

Este axioma permite trazar los segmentos.

III-2. Si el segmento $A'B'$ y el segmento $A''B''$ son iguales a un mismo segmento AB, entonces los segmentos $A'B'$ y $A''B''$ son iguales.

Del axioma III-2, en particular, se desprende que cualquier segmento es igual a sí mismo, y también que la relación de igualdad de segmentos es simétrica y transitiva.

III-3. Sea AB y BC dos segmentos de la recta a, que no tienen ningún punto interior común, y $A'B'$ y $B'C'$, dos segmentos de la

*) En la axiomática de Hilbert, a diferencia de la axiomática que se propone en el curso de geometría de la escuela media, los términos «congruencia» e «igualdad» son sinónimos.

misma o de otra recta a', que también no tienen puntos interiores comunes (véase la fig. 6.154). Si para esto

$$AB = A'B' \quad \text{y} \quad BC = B'C',$$

entonces también $AC = A'C'$.

Este axioma da la posibilidad de sumar los segmentos.

Los axiomas III-1 y III-3 contienen las afirmaciones que se refieren sólo a la congruencia de segmentos: se llaman *axiomas lineales de igualdad*.

Para formular los dos axiomas siguientes de igualdad se usa el concepto de «ángulo».

Los ángulos se encuentran en una relación definida, para cuya designación se usa también el concepto «igualdad».

III-4. Sean dados un ángulo $\sphericalangle (h_1, h_2)$ en el plano α y una recta a' en el plano α', así como una parte del plano α', bien definida respecto a la recta a una parte del plano α' y separada por la recta a' de todo el plano. Sea h_1' un rayo de la recta a', que sale del punto O'. Entonces en el plano α' existe uno y sólo un rayo h_2' tal, que el ángulo $\sphericalangle (h_1', h_2')$ es igual al ángulo $\sphericalangle (h_1, h_2)$ y todos los puntos interiores del ángulo $\sphericalangle (h_1', h_2')$ se encuentran en el plano α' al lado dado de la recta a':

$$\sphericalangle (h_1, h_2) = \sphericalangle (h_1', h_1').$$

III-5. Si para dos triángulos ABC y $A'B'C'$ tienen lugar las igualdades

$$AB = A'B', \qquad AC = A'C', \qquad \sphericalangle BAC = \sphericalangle B'A'C',$$

entonces tiene lugar también la igualdad

$$\sphericalangle ABC = \sphericalangle A'B'C'.$$

Los axiomas III-4 y III-5 se llaman *axiomas del plano de una igualdad*.

IV. *Axiomas de las paralelas.*

El axioma de las rectas paralelas que se llama también *axioma de Euclides* se formula de la manera siguiente:

Sea a una recta arbitraria, y A, un punto que no pertenece a la recta a. Entonces en el plano que se define por la recta a y el punto A existe no más que una recta que pasa por el punto A y que no interseca la recta a.

Esta recta se llama recta paralela a a, que pasa por el punto A.

A veces para formular los axiomas de geometría en vez de axioma de las paralelas se proponen otras definiciones equivalentes al axioma de las paralelas. Tales definiciones, por ejemplo, son las siguientes:

1) Si dos rectas a y b situadas en un plano no intersecan a una tercera recta, que se encuentra en el mismo plano, entonces ellas tampoco se intersecan entre sí.

2) Si se acepta el axioma de Arquímedes (véase más abajo), entonces el axioma de las paralelas puede ser sustituido por la exigencia de igualdad de la suma de los ángulos interiores de un triángulo con los ángulos rectos.

V. *Axiomas de continuidad.*

V-1 (*axioma de medición, o axioma de Arquímedes*). Sea AB y CD dos segmentos arbitrarios. Entonces en la recta AB existe un número finito de puntos $A_1, A_2, A_3, \ldots, A_n$ tales, que todos los segmentos $AA_1, A_1A_2, A_2A_3, \ldots, A_{n-1}A_n$ son iguales al segmento CD y el punto B está situado entre los puntos A y A_n (fig. 6.155).

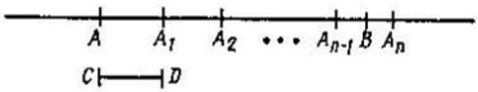


Fig. 6.155.

V-2 (*axioma de la completitud lineal*). Los puntos de la recta forman un sistema, el cual (para la condición de aceptación de los axiomas enumerados) no admite ninguna extensión, es decir, a un conjunto de puntos de la recta es imposible añadir ciertos objetos más de tal manera, que

1) todos los elementos de un conjunto extendido obtenido de nuevo se sometan a los axiomas enumerados;

2) todos los axiomas conserven su sentido inicial, es decir, que la relación de orden entre los puntos, y la igualdad de segmentos, que existían antes de la extensión, se conserven también en el conjunto extendido.

CAPÍTULO 7

Trigonometría

El objeto principal de la trigonometría es la obtención de los parámetros incógnitos del triángulo por medio de los valores dados de otros parámetros suyos. Así, aplicando los métodos trigonométricos, por los lados dados de un triángulo se pueden calcular sus ángulos, por el área y sus dos ángulos, los lados, etc. La necesidad de obtener los parámetros incógnitos de un triángulo dado surgió ante todo en la astronomía, y durante mucho tiempo la trigonometría ha sido una parte de la astronomía.

Los primeros métodos de obtención de los parámetros incógnitos de un triángulo dado fueron desarrollados por los sabios de la Grecia Antigua hace varios siglos antes de nuestra era. Los astrónomos griegos no analizaban los senos, cosenos y tangentes. En vez de las tablas de los valores de estas funciones trigonométricas, ellos hicieron y usaron las tablas que permiten buscar la cuerda de una circunferencia por su arco. La trigonometría tuvo su desarrollo ulterior en La Edad Media en las obras de los sabios árabes e indios (fueron introducidos los conceptos de seno, coseno, tangente, etc.). Las denotaciones contemporáneas de las letras fueron introducidas en la trigonometría a mediados del siglo XVIII. Aproximadamente al mismo tiempo, comenzaron a estudiarse en trigonometría las medidas de ángulos en radianes, fueron introducidas las funciones trigonométricas y trigonométricas inversas del argumento numérico, después de lo cual la trigonometría adquirió su aspecto contemporáneo.

§ 1. Funciones trigonométricas

1.1. Generalización del concepto de ángulo. En geometría fue dada la definición del ángulo llano y del ángulo menor que el llano (véase § 2 del capítulo 6). Tomemos dos ángulos obtusos (a, b) y (c, d). Hagamos coincidir los rayos b y d de tal manera, que los rayos a y c estén situados a distintos lados de la recta que contiene los rayos b y d (fig. 7.1).

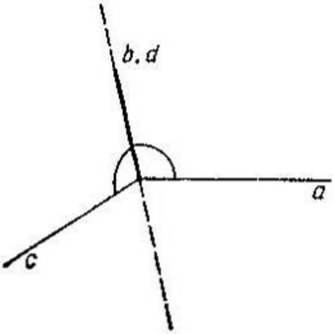


Fig. 7.1.

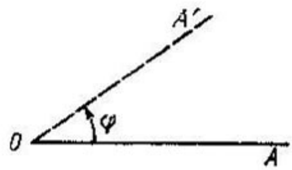


Fig. 7.2.

Como resultado de la adición de estos dos ángulos se obtiene un ángulo, mayor que el llano. Los rayos a y c son los lados de este ángulo. Para poder distinguir el ángulo obtenido del ángulo (a, c) que tiene los mismos lados, marquemos en la figura el ángulo obtenido con un arco. La región interior del ángulo marcado es la región exterior del ángulo $\angle (a, c)$. y la región exterior es la región interior del ángulo $\angle (a, c)$. Análogamente se pueden obtener los ángulos, mayores que dos ángulos llanos (es decir, mayores que el ángulo pleno), etc.

Para poder respresentar mejor un ángulo que es mayor que el ángulo pleno, es cómodo considerar el ángulo como una figura geométrica, obtenida por rotación de un rayo alrededor de su origen.

Tomemos en un plano un rayo OA y hagámoslo girar alrededor del punto O en sentido antihorario (fig. 7.2). Al girar en una cuarta parte de la revolución completa alrededor del punto O, entre el rayo OA y el rayo OA' (la imagen del rayo OA) se obtiene un ángulo igual a 90° ($\pi/2$). Análogamente se obtiene un ángulo llano, cuando el rayo OA realiza la mitad de la revolución completa alrededor del punto O; un ángulo pleno se obtiene, si el rayo OA realiza una revolución completa alrededor del punto O. Los ángulos entre el rayo OA y el rayo OA', mayores que el ángulo pleno,

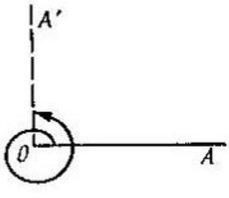


Fig. 7.3.

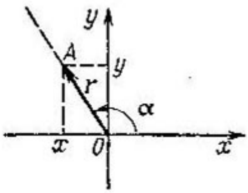


Fig. 7.4.

se obtienen, cuando el rayo OA comienza a realizar la segunda revolución alrededor del punto O. Así, por ejemplo, el ángulo 450° ($5\pi/4$) entre el rayo OA y el rayo OA' está representado en la fig. 7.3.

Podemos girar un rayo alrededor de un punto inicial en dos direcciones: en sentido horario y en sentido antihorario. La dirección de rotación en sentido antihorario se llama *positiva* y en sentido horario, *negativa*. Respectivamente, los ángulos, obtenidos mediante la rotación de un rayo en sentido antihorario se llaman *positivos*, y en sentido horario, *negativos*.

1.2. Funciones trigonométricas. Sea Oxy cierto sistema de coordenadas cartesianas rectangulares; OA es un rayo, obtenido mediante la rotación de un semieje positivo Ox alrededor del origen de coordenadas O en un ángulo α (fig. 7.4) (A es un punto arbitrario del rayo). Designemos las coordenadas del punto A con x e y: $A(x, y)$, y examinemos un vector $\overrightarrow{OA} = (x;\ y)$. El módulo de este vector es igual

$$r = \sqrt{x^2 + y^2}.$$

El ángulo entre la dirección positiva del eje Ox y el rayo OA se llama *ángulo* entre el eje Ox y el vector $\overrightarrow{OA}$.

Las razones

$$\frac{x}{r} \text{ y } \frac{y}{r}$$

de las coordenadas del vector $\overrightarrow{OA}$ a la longitud del vector $\overrightarrow{OA}$ no dependen de la longitud del vector OA (es decir, de la posición del punto A en el rayo $\overrightarrow{OA}$), sino que dependen de su dirección*).

La razón de la ordenada de un vector, que forma con el eje Ox un ángulo α, a la longitud de este vector, se llama *seno* del ángulo α y se designa sen α:

$$\operatorname{sen} \alpha = \frac{y}{r}.$$

La razón de la abscisa de un vector, que forma con el eje Ox un ángulo α, a la longitud de este vector se llama *coseno* del ángulo α y se designa cos β:

$$\cos \alpha = \frac{x}{r}.$$

La razón del seno del ángulo α a su coseno se llama *tangente* del ángulo α y se designa tg α (o tang α):

$$\operatorname{tg} \alpha = \frac{\operatorname{sen} \alpha}{\cos \alpha},$$

o, según la definición del sen α y cos α,

$$\operatorname{tg} \alpha = \frac{y}{x}.$$

*) De un conjunto de todos los vectores con el origen en el punto O se puede formar un subconjunto de vectores con el origen en el punto O y un módulo igual a la unidad. Los extremos de todos tales vectores pertenecen a la circunferencia con centro en el origen de coordenadas y un radio igual a la unidad, es decir, a la circunferencia de unidad. Las funciones trigonométricas se introducen muy a menudo con la utilización de la circunferencia de unidad (o del conjunto de los vectores $\overrightarrow{OA}$ con los extremos, pertenecientes a la circunferencia de unidad). Así, por ejemplo, si el punto A pertenece a la circunferencia de unidad, entonces la ordenada del vector $\overrightarrow{OA}$ que forma con el eje Ox un ángulo α, se llama seno del ángulo α.

La razón del coseno del ángulo α a su seno se llama *cotangente* del ángulo α y se designa ctg α:

$$\operatorname{ctg}\alpha=\frac{\cos\alpha}{\operatorname{sen}\alpha}$$

o, según la definición del sen α y cos α,

$$\operatorname{ctg}\alpha=\frac{x}{y}.$$

El seno y el coseno están determinados para cualquier ángulo α, la tangente está determinada para todos los valores del ángulo α, a excepción de $\alpha=90^\circ+180^\circ\cdot n$ $\left(\alpha=\frac{\pi}{2}+\pi n\right)$, $n\in Z$, y la cotangente, para todos los valores α, a excepción de $\alpha=180^\circ\cdot n$ $(\alpha=\pi n)$, $n\in Z$.

El seno y coseno del mismo ángulo se relacionan por la igualdad

$$\operatorname{sen}^2\alpha+\cos^2\alpha=1,$$

la tangente y el coseno del mismo ángulo, por la igualdad

$$1+\operatorname{tg}^2\alpha=\frac{1}{\cos^2\alpha},$$

la cotangente y el seno, por la igualdad

$$1+\operatorname{ctg}^2\alpha=\frac{1}{\operatorname{sen}^2\alpha},$$

y la tangente y cotangente del mismo ángulo, por la igualdad

$$\operatorname{tg}\alpha\cdot\operatorname{ctg}\alpha=1.$$

A cada ángulo α le corresponden unos valores bastante determinados del sen α y cos α. A cada valor $\alpha\neq 90^\circ+180^\circ\cdot n$ $\left(\alpha\neq\frac{\pi}{2}+\pi n\right)$, $n\in Z$, le corresponde un valor bastante determinado del tg α, y a cada valor $\alpha\neq 180^\circ\cdot n$ $(\alpha\neq\pi n)$, $n\in Z$, un valor determinado del ctg α.

Las funciones sen α, cos α, tg α y ctg α se llaman *funciones trigonométricas* del ángulo α. Además de las funciones trigonométricas básicas sen α, cos α, tg α y ctg α se introducen aún dos funciones trigonométricas más del ángulo α, secante y cosecante, que se designan sec α y cosec α respectivamente. Estas funciones se definen por las igualdades

$$\sec\alpha = \frac{1}{\cos\alpha}, \quad \operatorname{cosec}\alpha = \frac{1}{\operatorname{sen}\alpha}.$$

La secante del ángulo α está definida para todos los valores α, a excepcion de $\alpha = 90° + 180°\cdot n$ $\left(\alpha = \frac{\pi}{2} + \pi n\right)$, $n \in \mathbf{Z}$, y la cosecante, para todos los valores del ángulo α, a excepcion de $\alpha = 180°\cdot n$ $(\alpha = \pi n)$, $n \in \mathbf{Z}$.

1.3. Cuadrantes de la circunferencias de unidad. Signos de valores de las funciones trigonométricas. Los ejes de

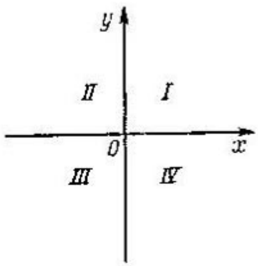


Fig. 7.5.

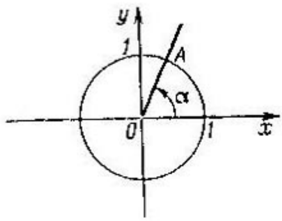


Fig. 7.6.

coordenadas Ox, Oy del sistema de coordenadas rectangulares dividen el plano de coordenadas en cuatro partes que se llaman *cuartos* o *cuadrantes* (fig. 7.5). El primer cuadrante es la parte del plano que está situada más arriba del eje Ox y a la derecha del eje Oy; el segundo cuadrante está más arriba del eje Ox y a la izquierda del eje Oy;más abajo del eje Ox y a la izquierda del eje Oy está el tercer cuadrante; más abajo del eje Oxy a la derecha del eje Oy, el cuarto cuadrante. La circunferencia de unidad (la circunferencia de radio de unidad con centro en el origen de coordenadas) se divide por los ejes de coordenadas en cuatro partes, las cuales también se llaman cuartas (o cuadrantes) y se numeran de la misma manera como las cuartas del plano.

De los puntos de la circunferencia de unidad, pertenecientes a los ejes de coordenadas, se dice que están situados en el semieje positivo (o negativo) de abscisas u ordenadas.

Sea que el rayo OA, obtenido mediante la rotación del semieje positivo Ox en un ángulo α, interseca la circunferencia de unidad en el punto A (fig. 7.6). El ángulo α se considera perteneciente al primero, segundo, tercero o cuarto cuadrante, si el punto A pertenece respectivamente al primero, segundo, tercero o cuarto cuadrante de la circunferencia de unidad.

Las abscisas de los puntos, pertenecientes al I y IV cuadrantes, son positivas, y de los puntos, pertenecientes al II y III cuadrantes son negativas. Las ordenadas de los puntos, pertenecientes al I y II cuadrantes, son positivas, y al III y IV cuadrantes son negativas. Los signos de los valores de las funciones trigonométricas se establecen sobre la base de la definición de las funciones trigonométricas y dependen del cuadrante, al cual pertenece el ángulo α.

Tabla 1

Función / Cuartos	sen α	cos α	tg α	ctg α
I	+	+	+	+
II	+	−	−	−
III	−	−	+	+
IV	−	+	−	−

En la tabla 1 se muestran los signos de los valores de las funciones trigonométricas principales.

1.4. Funciones trigonométricas de un argumento numérico. Los ángulos se pueden medir en grados y radianes. La aplicación del radián para medir los ángulos permite introducir las funciones trigonométricas de un argumento numérico. Esto se puede hacer de la manera siguiente. Entre un conjunto de todos los números reales y un conjunto de todos los ángulos (los cuales se consideran como un giro del rayo OA alrededor del punto O), medidos en radianes, existe una correspondencia biunívoca. Por ejemplo, al número 1 le corresponde un giro del rayo OA alrededor de su origen en un radián en sentido antihorario; al número 3/2 le corresponde un giro del rayo OA alrededor de su origen en 3/2 del radián en sentido horario, etc.

Se llama *seno* del número x al número que es igual al seno del ángulo en x radianes. El número que es igual al coseno del ángulo en x radianes se llama *coseno* del número x. Análogamente se definen y otras funciones trigonométricas del argumento numérico.

Propiedades principales de las funciones trigonométricas.

Propiedades de la función sen x.

1) El campo de definición es un conjunto de todos los números reales.

2) La región de cambio (conjunto de valores) es el intervalo $[-1; 1]$.

3) La función sen x es impar: $\operatorname{sen}(-x) = -\operatorname{sen} x$.

4) La función sen x es periódica. El período menor positivo es igual a 2π:

$$\operatorname{sen}(x + 2\pi) = \operatorname{sen} x.$$

5) Los ceros de la función son: $\operatorname{sen} x = 0$ para $x = \pi n$, $n \in \mathbf{Z}$.

6) Los intervalos de signo constante son:

$$\operatorname{sen} x > 0 \text{ para } x \in (2\pi n; \pi + 2\pi n),\ n \in \mathbf{Z},$$

$$\operatorname{sen} x < 0 \text{ para } x \in (\pi + 2\pi n; 2\pi + 2\pi n),\ n \in \mathbf{Z}.$$

7) La función sen x es continua y tiene una derivada para cualquier valor del argumento:

$$(\operatorname{sen} x)' = \cos x.$$

8) La función sen x crece para $x \in \left(-\frac{\pi}{2} + 2\pi; \frac{\pi}{2} + 2\pi n\right)$, $n \in \mathbf{Z}$, y decrece para $x \in \left(\frac{\pi}{2} + 2\pi n; \frac{3\pi}{2} + 2\pi n\right)$, $n \in \mathbf{Z}$.

9) La función sen x tiene unos valores mínimos que son iguales a -1, para $x = -\frac{\pi}{2} + 2\pi n$, $n \in \mathbf{Z}$, y los valores máximos que son iguales a 1, para $x = \frac{\pi}{2} + 2\pi n$, $n \in \mathbf{Z}$.

La gráfica de la función $y = \operatorname{sen} x$ está representada en la fig. 7.7. La gráfica de la función sex x se llama *sinusoide.*

Propiedades de la función cos x.

1) El campo de definición es un conjunto de todos los números reales.

2) La región de cambio (conjunto de valores) es el intervalo $[-1; 1]$.

3) La función $\cos x$ es par: $\cos(-x) = \cos x$.

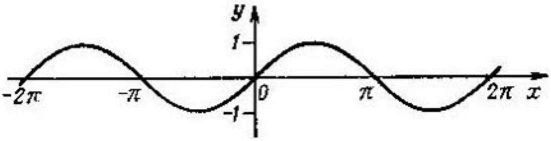


Fig. 7.7.

4) La función $\cos x$ es periódica. El período positivo mínimo de la función es igual a 2π:

$$\cos(x + 2\pi) = \cos x.$$

5) Los ceros de la función son: $\cos x = 0$ para $x = \frac{\pi}{2} + \pi n$, $n \in \mathbf{Z}$.

6) Los intervalos de los signos constantes son:

$$\cos x > 0 \text{ para } x \in \left(-\frac{\pi}{2} + 2\pi n; \frac{\pi}{2} + 2\pi n\right), \ n \in \mathbf{Z},$$

$$\cos x < 0 \text{ para } x \in \left(\frac{\pi}{2} + 2\pi n; \frac{3\pi}{2} + 2\pi n\right), \ n \in \mathbf{Z}.$$

7) La función $\cos x$ es continua y diferenciable para cualquier valor del argumento:

$$(\cos x)' = -\operatorname{sen} x.$$

8) La función $\cos x$ crece para $x \in (-\pi + 2\pi n; 2\pi n)$, $n \in \mathbf{Z}$, y decrece para $x \in (2\pi n; \pi + 2\pi n)$, $n \in \mathbf{Z}$.

9) La función $\cos x$ tiene unos valores mínimos que son iguales a -1, para $x = \pi + 2\pi$, $n \in \mathbf{Z}$, y unos valores máximos que son iguales a 1, para $x = 2\pi n$, $n \in \mathbf{Z}$.

La gráfica de la función $y = \cos x$ se muestra en la fig. 7.8.

Propiedades de la función tg x.

1) El campo de definición de la función es un conjunto de todos los números reales, menos los números $x = \frac{\pi}{2} + \pi n$, $n \in \mathbf{Z}$.

2) La región de cambio (un conjunto de valores) es un conjunto de todos los números reales.

3) La función tg x es impar: $\operatorname{tg}(-x) = -\operatorname{tg} x$.

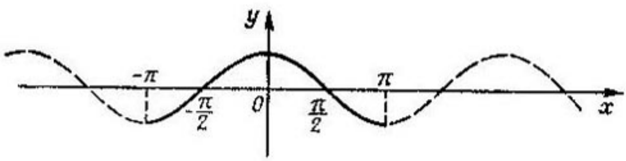


Fig. 7.8.

4) La función tg x periódica. El mínimo período positivo de la función es igual a π:

$$\operatorname{tg}(x + \pi) = \operatorname{tg} x.$$

5) Los ceros de la función son: $\operatorname{tg} x = 0$ para $x = \pi n$, $n \in \mathbf{Z}$.

6) Los intervalos de signo constante son:

$$\operatorname{tg} x > 0 \text{ para } x \in \left(\pi n; \frac{\pi}{2} + \pi n\right), \quad n \in \mathbf{Z},$$

$$\operatorname{tg} x < 0 \text{ para } x \in \left(-\frac{\pi}{2} + \pi n;\ \pi n\right), \quad n \in \mathbf{Z}.$$

7) La función tg x es continua y diferenciable para cualquier valor del argumento del campo de definición de la función:

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x}.$$

8) La función tg x crece en cada uno de los intervalos

$$\left(-\frac{\pi}{2} + \pi n;\ \frac{\pi}{2} + \pi n\right), \quad n \in \mathbf{Z}.$$

La gráfica de la función $y = \operatorname{tg} x$ está representada en la fig. 7.9. La gráfica de la función tg x se llama *tangensoide*.

Propiedades de la función ctg x.

1) El campo de definición de la función es un conjunto de todos los números reales, menos los números de tipo $x = \pi n$, $n \in \mathbf{Z}$.

2) La región de cambio (conjunto de valores) es un conjunto de todos los números reales.

3) La función ctg x es impar: $\operatorname{ctg}(-x) = -\operatorname{ctg} x$.

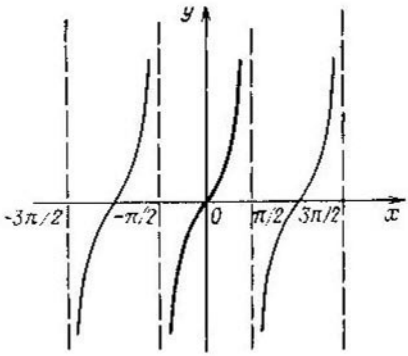


Fig. 7.9.

4) La función ctg x es periódica. El mínimo período positivo de la función es igual a π:

$$\operatorname{ctg}(x+\pi) = \operatorname{ctg}(x).$$

5) Los ceros de la función son: $\operatorname{ctg} x = 0$ para $x = \pi/2 + \pi n$, $n \in \mathbf{Z}$.

6) Los intervalos de signo constante son:

$$\operatorname{ctg} x > 0 \text{ para } x \in \left(\pi n; \frac{\pi}{2} + \pi n\right), \quad n \in \mathbf{Z},$$

$$\operatorname{ctg} x < 0 \text{ para } x \in \left(\frac{\pi}{2} + \pi n; \ \pi(n+1)\right), \quad n \in \mathbf{Z}.$$

7) La función ctg x es continua y diferenciable para cualquier valor del argumento del campo de definición de la función:

$$(\operatorname{ctg} x)' = -\frac{1}{\operatorname{sen}^2 x}.$$

8) La función ctg x decrece en cada uno de los intervalos $(\pi n; \ \pi(n+1))$, $n \in \mathbf{Z}$.

La gráfica de la función $y = \operatorname{ctg} x$ se muestra en la fig. 7.10.

Propiedades de la función sec x.

1) El campo de definición de la función es un conjunto de todos los números reales, menos los números de tipo $x = \pi/2 + \pi n$, $n \in \mathbf{Z}$.

2) La región de cambio (conjunto de valores) es:

$$(-\infty; -1] \cup [1; +\infty).$$

3) La función sec x es par: $\sec(-x) = \sec x$.

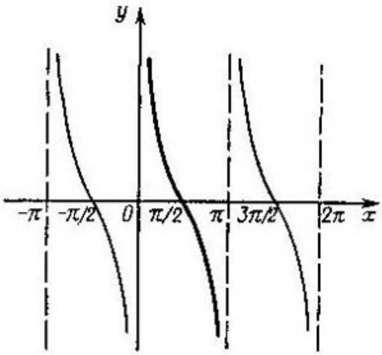


Fig. 7.10.

4) La función sec x es periódica. El mínimo período positivo de la función es igual a 2π: $\sec(x + 2\pi) = \sec x$.

5) La función sec x no se anula para ningún valor del argumento.

6) Los intervalos de signo constante son:

$$\sec x > 0 \text{ para } x \in \left(-\frac{\pi}{2} + 2\pi n; \frac{\pi}{2} + 2\pi n\right), \quad n \in \mathbf{Z},$$

$$\sec x < 0 \text{ para } x \in \left(\frac{\pi}{2} + 2\pi n; \frac{3\pi}{2} + 2\pi n\right), \quad n \in \mathbf{Z}.$$

7) La función sec x es continua y diferenciable para cualquier valor del argumento del campo de definición de la función:

$$(\sec x)' = \frac{\operatorname{sen} x}{\cos^2 x}.$$

8) La función sec x crece en los intervalos $\left[2\pi n; \frac{\pi}{2} + 2\pi n\right)$, $\left(\frac{\pi}{2} + 2\pi n; \pi + 2\pi n\right]$, $n \in \mathbf{Z}$, y decrece en los intervalos $\left[\pi + 2\pi n; \frac{3\pi}{2} + 2\pi n\right)$, $\left(\frac{3\pi}{2} + 2\pi n; 2\pi(n+1)\right]$, $n \in \mathbf{Z}$.

La gráfica de la función $y = \sec x$ está representada en la fig. 7.11

Propiedades de la función cosec x.

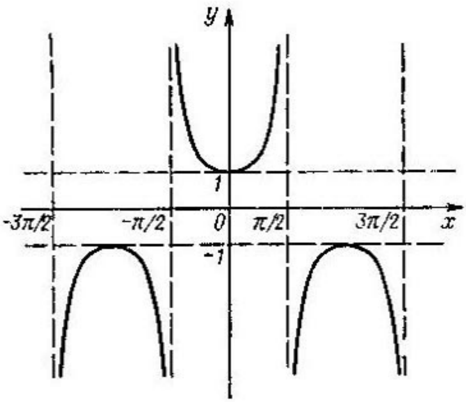


Fig. 7.11.

1) El campo de definición de la función es un conjunto de todos los números reales, menos los números de tipo $x = \pi n$, $n \in \mathbf{Z}$.

2) La región de cambio (conjunto de valores) de la función es:

$$(-\infty; -1] \cup [1; +\infty).$$

3) La función cosec x es impar: $\operatorname{cosec}(-x) = -\operatorname{cosec} x$.

4) La función cosec x es periódica. El mínimo período positivo de la función es igual a 2π:

$$\operatorname{cosec}(x + 2\pi) = \operatorname{cosec} x.$$

5) La función cosec x no se anula para ningún valor del argumento.

6) Los intervalos de signo constante son:

$$\operatorname{cosec} x > 0 \text{ para } x \in (2\pi n; \pi + 2\pi n),\ n \in \mathbf{Z},$$
$$\operatorname{cosec} x < 0 \text{ para } x \in (\pi + 2\pi n; 2\pi (n + 1)),\ n \in \mathbf{Z}.$$

7) La función cosec x es continua y diferenciable para cualquier valor del argumento del campo de definición de

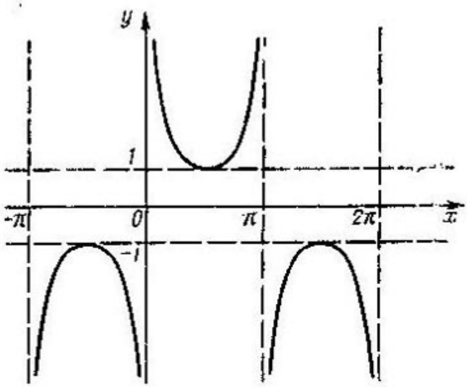


Fig. 7.12.

la función:

$$(\operatorname{cosec} x)' = -\frac{\cos x}{\operatorname{sen}^2 x}.$$

8) La función cosec x crece en los intervalos $\left[\frac{\pi}{2} + 2\pi n; \pi + 2\pi n\right)$ y $\left(\pi + 2\pi n; \frac{3\pi}{2} + 2\pi n\right]$, $n \in \mathbf{Z}$, y decrece en los intervalos $\left(2\pi n; \frac{\pi}{2} + 2\pi n\right]$ y $\left[\frac{3\pi}{2} + 2\pi n; 2\pi + 2\pi n\right)$, $n \in \mathbf{Z}$.

La gráfica de la función $y = \operatorname{cosec} x$ se muestra en la fig. 7.12.

1.5. Funciones trigonométricas inversas.

La *función* arcsen x. La función $y = \operatorname{sen} x$ es definida y continua para cualquier $x \in \mathbf{Z}$. Separemos en el eje numérico Ox un intervalo $[-\pi/2; \pi/2]$. En este intervalo la función $y = \text{sen } x$ crece; en el extremo izquierdo del intervalo (para $x = -\pi/2$) la función $y = \text{sen } x$ adquiere su valor mínimo, igual a -1, y en el extremo derecho (para $x = \pi/2$) adquiere su valor máximo, igual a 1. La función $y = \text{sen } x$, la cual se examina en el intervalo $[-\pi/2; \pi/2]$, tiene una función inversa que se llama *arco seno* y se escribe

$$x = \text{arcsen } y,$$

donde y es una variable independiente, y x, una variable dependiente. Designando, como se suele hacer, a la variable independiente con la letra x, y a la variable dependiente con la letra y, a continuación escribiremos $y = \text{arcsen } x$.

Propiedades de la función arcsen x.

1) El campo de definición es el intervalo $[-1; 1]$

2) La región de cambio (conjunto de valores) es el intervalo $[-\pi/2; \pi/2]$.

3) La función arcsen x es impar: $\text{arcsen } (-x) = -\text{arcsen } x$.

4) Los ceros de la función son: $\text{arcsen } x = 0$ para $x = 0$.

5) Los intervalos de signo constante son:

$$\text{arcsen } x > 0 \text{ para } x \in (0; 1],$$
$$\text{arcsen } x < 0 \text{ para } x \in [-1; 0).$$

6) La función arcsen x es continua y diferenciable en cada punto del intervalo $(-1; 1)$:

$$(\text{arcsen } x)' = \frac{1}{\sqrt{1-x^2}}.$$

7) La función arcsen x crece en el intervalo $[-1; 1]$, obteniendo su valor mínimo, igual a $-\pi/2$— en el extremo izquierdo del intervalo y el valor máximo, igual a $\pi/2$, en el extremo derecho del intervalo $[-1; 1]$.

La gráfica de la función $y = \text{arcsen } x$ se muestra en la fig. 7.13.

La *función* arccos x. La función $y = \cos x$ es definida y continua para cualquier $x \in \mathbf{R}$. Separamos en el eje numérico Ox un intervalo $[0; \pi]$. En este intervalo la función $y = \cos x$ decrece; en el extremo izquierdo del intervalo (para $x = 0$) la función $y = \cos x$ adquiere su valor máximo, igual a 1, y en el extremo derecho (para $x = \pi$) adquiere su valor mínimo, igual a -1. La función $y = \cos x$, la cual examinamos en el intervalo $[0; \pi]$, tiene una función inversa, la cual se llama *arco coseno* y se escribe

$$x = \arccos y,$$

donde y es la variable independiente, y x, la dependiente. Designando, como se suele designar, a la variable independiente con la letra x, y a la variable dependiente con la letra y, a continuación escribiremos $y = \arccos x$.

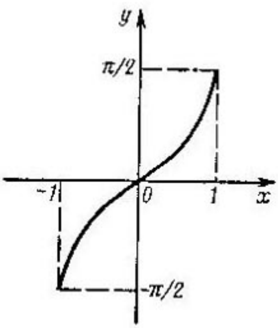


Fig. 7.13.

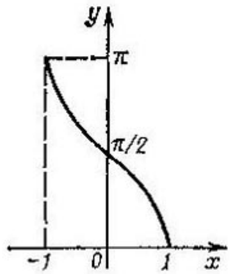


Fig. 7.14.

Propiedades de la función $\arccos x$.

1) El campo de definición es el intervalo $[-1; 1]$.

2) La región de cambio (conjunto de valores) es el intervalo $[0; \pi]$.

3) La función $\arccos x$ no es ni par, ni impar.

4) Los ceros de la función son: $\arccos x = 0$ para $x = 1$.

5) Los intervalos de signo constante son: $\arccos x > 0$ para $x \in [-1; 1)$.

6) La función $\arccos x$ es continua y diferenciable en cada punto del intervalo $(-1; 1)$:

$$(\arccos x) = -\frac{1}{\sqrt{1-x^2}}.$$

7) La función $\arccos x$ decrece en el intervalo $[-1; 1]$, obteniendo el valor máximo, igual a π, en el extremo iz-

quierdo del intervalo y el valor mínimo, igual a 0, en el extremo derecho del intervalo.

La gráfica de la función $y = \arccos x$ está representada en la fig. 7.14.

Función arctg x. La función $y = \operatorname{tg} x$ es continua en su campo de definición (es decir, para todos $x \in \mathbf{R}$, no iguales a $\frac{\pi}{2} + \pi n$, $n \in \mathbf{Z}$). Separamos en el eje numérico Ox el intervalo $(-\pi/2; \pi/2)$. En este intervalo la función $y = \operatorname{tg} x$ crece y obtiene todos sus valores. La función $y = \operatorname{tg} x$, que examinamos en el intervalo $(-\pi/2; \pi/2)$, tiene una función inversa que se llama *arco tangente* y se escribe

$$x = \operatorname{arctg} y,$$

donde y es una variable independiente, y x, una variable dependiente. Designando, como siempre, a la variable independiente por la letra x, y a la dependiente por la letra y, a continuación escribiremos $y = \operatorname{arctg} x$.

Propiedades de la función arctg x.

1) El campo de definición es toda la recta numérica.

2) La región de cambio (conjunto de valores) es el intervalo $(-\pi/2; \pi/2)$.

3) La función arctg x es impar: $\operatorname{arctg}(-x) = -\operatorname{arctg} x$.

4) Los ceros de la función son: $\operatorname{arctg} x = 0$ para $x = 0$.

5) Los intervalos de signo constante:

$$\operatorname{arctg} x > 0 \text{ para } x \in (0; +\infty),$$
$$\operatorname{arctg} x < 0 \text{ para } x \in (-\infty; 0).$$

6) La función arctg x es continua y diferenciable para todos $x \in \mathbf{R}$:

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2}.$$

7) La función arctg x es creciente.

La gráfica de la función $y = \operatorname{arctg} x$ se muestra en la fig. 7.15.

La *función* arcctg x. La función $y = \operatorname{ctg} x$ es continua en su campo de definición (es decir, para todos $x \in \mathbf{R}$, no iguales a πn, $n \in \mathbf{Z}$). Separemos en el eje numérico Ox el intervalo $(0; \pi)$. En este intervalo la función $y = \operatorname{ctg} x$ decrece y adquiere todos sus valores. La función $y = \operatorname{ctg} x$ que examinamos en el intervalo $(0; \pi)$ tiene una función inversa, la cual se llama *arco cotangente* y se escribe

$$x = \operatorname{arcctg} y,$$

donde y es una variable independiente, y x, una variable dependiente. Designando, como siempre, la variable indepen-

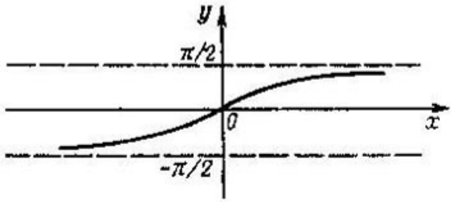


Fig. 7.15.

diente con la letra x, y la dependiente con la letra y, más adelante escribiremos $y = \operatorname{arcctg} x$.

Propiedades de la función $\operatorname{arcctg} x$.

1) El campo de definición es toda la recta numérica.

2) La región de cambio (conjunto de valores) es el intervalo $(0; \pi)$.

3) La función arrcctg x no es ni par, ni impar.

4) La función $\operatorname{arcctg} x$ es positiva para todos $x \in \mathbf{R}$.

5) La función $\operatorname{arcctg} x$ es continua y diferenciable para todos $x \in \mathbf{R}$:

$$(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}.$$

6) La función $\operatorname{arcctg} x$ es decreciente.

La gráfica de la función $y = \operatorname{arcctg} x$ se muestra en la fig. 7.16.

1.6. Valores de las funciones trigonométricas de algunos ángulos. Se pueden calcular los valores de las funciones trigonométricas de los ángulos de 0°, 30°, 45°, 60° y 90°, utilizando las definiciones de las funciones trigonométricas respectivas. Los valores de las funciones trigonométricas de los ángulos de 15°, $\frac{15°}{2}$, $\frac{15°}{4}$, ... se pueden hallar por los valores conocidos de las funciones trigonométricas de 30°, utilizando sucesivamente las fórmulas del ángulo mitad (véase el ejemplo 1). Los valores de las funciones trigonométricas

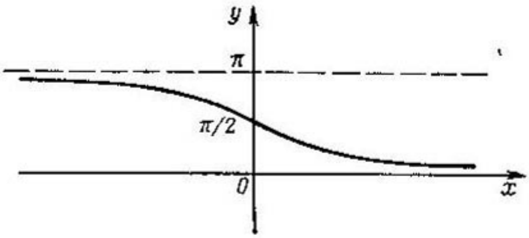


Fig. 7.16.

de los ángulos, múltiplos de 18°, se pueden hallar, conociendo por lo menos el valor de una función trigonométrica de 18°, por ejemplo, el sen 18° (véase el ejemplo 2).

Algunos valores de las funciones trigonométricas se dan en la tabla 2.

Ejemplo 1. Calcular cos 15°.

Los cosenos de los ángulos α y $\alpha/2$ se relacionan por la fórmula

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

Suponiendo $\alpha = 30°$, obtenemos

$$\cos^2 15° = \frac{1+\cos 30°}{2}.$$

$$\cos^2 15° = \frac{1+\frac{\sqrt{3}}{2}}{2},$$

$$\cos^2 15° = \frac{2+\sqrt{3}}{4},$$

$$\cos 15° = \frac{\sqrt{2+\sqrt{3}}}{2} = \frac{\sqrt{3}+1}{2\sqrt{2}}.$$

El sen 15°se puede calcular, utilizando la relación entre las funciones sen α y cos α:

$$\operatorname{sen}^2\alpha + \cos^2\alpha = 1,$$

Tabla 2

Argumento \ Función	sen α	cos α	tg α	ctg α
$0°\quad(0)$	0	1	0	no está determinado
$15°\ \left(\frac{\pi}{12}\right)$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$2-\sqrt{3}$	$2+\sqrt{3}$
$18°\ \left(\frac{\pi}{10}\right)$	$\frac{\sqrt{5}-1}{4}$	$\frac{\sqrt{5+\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}-1}{\sqrt{10+2\sqrt{5}}}$	$\frac{\sqrt{10+2\sqrt{5}}}{\sqrt{5}-1}$
$30°\ \left(\frac{\pi}{6}\right)$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$	$\sqrt{3}$
$36°\ \left(\frac{\pi}{5}\right)$	$\frac{\sqrt{5-\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}+1}{4}$	$\frac{\sqrt{10-2\sqrt{5}}}{\sqrt{5}+1}$	$\frac{\sqrt{5}+1}{\sqrt{10-2\sqrt{5}}}$
$45°\ \left(\frac{\pi}{4}\right)$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	1	1
$54°\ \left(\frac{3\pi}{10}\right)$	$\frac{\sqrt{5}+1}{4}$	$\frac{\sqrt{5-\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}+1}{\sqrt{10-2\sqrt{5}}}$	$\frac{\sqrt{10-2\sqrt{5}}}{\sqrt{5}+1}$
$60°\ \left(\frac{\pi}{3}\right)$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$	$\frac{1}{\sqrt{3}}$
$72°\ \left(\frac{2\pi}{5}\right)$	$\frac{\sqrt{5+\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}-1}{4}$	$\frac{\sqrt{10+2\sqrt{5}}}{\sqrt{5}-1}$	$\frac{\sqrt{5}-1}{\sqrt{10+2\sqrt{5}}}$
$75°\ \left(\frac{5\pi}{12}\right)$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$2+\sqrt{3}$	$2-\sqrt{3}$
$90°\ \left(\frac{\pi}{2}\right)$	1	0	no está determinado	0

en vigor de la cual

$$\operatorname{sen} 15° = \sqrt{1-\cos^2 15°} = \frac{\sqrt{2-\sqrt{3}}}{2} = \frac{\sqrt{3}-1}{2\sqrt{2}}.$$

Ejemplo 2. Calcular los valores de las funciones trigonométricas de un ángulo de 18°.

Examinemos un triángulo isósceles ABC ($|AB| = |BC|$) con el ángulo del vértice de 36° (fig. 7.17). Los ángulos de la base del triángulo ABC son iguales a 72°. Tracemos las bisectrices de los ángulos BAC y ABC. A los segmentos de las bisectrices que se encuentran dentro del triángulo las denotamos respectivamente con AM y BK.

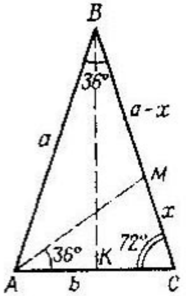


Fig. 7.17.

Examinemos el triángulo AMC. En el triángulo AMC el $\angle MAC$ es igual a 36°, $\angle AMC = \angle MCA$. Por consiguiente, el triángulo AMC es semejante al triángulo ABC.

Designando $|BC|$ por a, $|AC|$ por b, y $|MC|$ por x, la condición de semejanza de los triángulos ABC y AMC se puede escribir en la forma

$$\frac{|AC|}{|BC|}=\frac{|MC|}{|AC|} \quad \text{o} \quad \frac{b}{a}=\frac{x}{b}. \tag{1}$$

Puesto que la bisectriz del ángulo interior del triángulo divide el lado que ella interseca en segmentos, proporcionales a los lados adyacentes, entonces para el triángulo ABC tenemos

$$\frac{|AB|}{|AC|}=\frac{|MB|}{|MC|}, \quad \text{o} \quad \frac{a}{b}=\frac{a-x}{x}. \tag{2}$$

Tomemos en el triángulo ABC la altura BK (véase fig. 7.17) y examinemos un triángulo rectángulo BKC con un ángulo agudo KBC de 18°:

$$\frac{|KC|}{|BC|}=\frac{b}{2a}=\operatorname{sen} 18°.$$

De esta manera, para obtener el valor sen 18° es necesario hallar en el sistema de ecuaciones (1) y (2) el valor $b/(2a)$. De la ecuación (1) obtenemos $x = b^2/a$.

Sustituyendo la expresión de x en la ecuación (2), obtenemos

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

Introduciendo una variable nueva $z = b/a$, la última ecuación se puede escribir en la forma

$$z^2+z-1=0.$$

Hallamos las raíces de la ecuación cuadrática:

$$z_{1,2}=\frac{-1\pm\sqrt{5}}{2}.$$

De esta manera, el valor $b/(2a)$ es igual a $(\sqrt{5}-1)/4$. Este número es, precisamente, el valor sen 18°:

$$\operatorname{sen} 18^\circ = \frac{\sqrt{5}-1}{4}.$$

Luego, usando la fórmula $\operatorname{sen}^2 \alpha + \cos^2 \alpha = 1$ y las definiciones de las funciones tg x, ctg x, se pueden hallar los valores cos 18°, tg 18°, ctg 18°, y mediante las fórmulas del ángulo doble se puede hallar sen 36°, cos 36°. Así, por ejemplo, sen 36° se puede calcular de la manera siguiente:

$$\cos 18^\circ = \sqrt{1-\operatorname{sen}^2 18^\circ} = \frac{\sqrt{10+2\sqrt{5}}}{4},$$

$$\operatorname{sen} 36^\circ = 2 \operatorname{sen} 18^\circ \cos 18^\circ = \frac{1}{8}\,(\sqrt{5}-1)\sqrt{10+2\sqrt{5}} = \frac{1}{4}\sqrt{10-2\sqrt{5}}.$$

§ 2. Fórmulas trigonométricas

2.1. Fórmulas de reducción. El cálculo de valores de las funciones trigonométricas de cualquier ángulo se reduce al cálculo de los valores de las funciones trigonométricas de un ángulo agudo mediante las reglas siguientes:

1) Si el ángulo es positivo y mayor que 2π, entonces las funciones seno y coseno del ángulo dado se reducen a las funciones del ángulo, mayor que 0 y menor que 2π, mediante las fórmulas

$$\operatorname{sen}(\alpha + 2\pi n) = \operatorname{sen}\alpha;$$
$$\cos(\alpha + 2\pi n) = \cos\alpha$$
$$(\alpha \in (0, 2\pi),\ n \in \mathbf{Z}),$$

y las funciones tangente y cotangente del ángulo dado se reducen a las funciones del ángulo que es mayor que 0 y menor que π, mediante las fórmulas

$$\operatorname{tg}(\alpha + \pi n) = \operatorname{tg}\alpha;$$
$$\operatorname{ctg}(\alpha + \pi n) = \operatorname{ctg}\alpha$$
$$(\alpha \in (0; \pi),\ n \in \mathbf{Z}).$$

2) Si el ángulo es negativo, entonces las funciones trigonométricas del ángulo dado se reducen a las funciones trigonométricas del ángulo positivo mediante las fórmulas

$$\operatorname{sen}(-\alpha) = -\operatorname{sen}\alpha; \qquad \cos(-\alpha) = \cos\alpha;$$

$$\operatorname{tg}(-\alpha) = -\operatorname{tg}\alpha; \qquad \operatorname{ctg}(-\alpha) = -\operatorname{ctg}\alpha.$$

3) Las funciones trigonométricas del ángulo mayor que $\pi/2$ y menor que 2π, se reducen a las funciones trigonométricas del ángulo agudo mediante las fórmulas de reducción

Tabla 3

Función \ Argumento	$\beta = \frac{\pi}{2} \pm \alpha$	$\beta = \pi \pm \alpha$	$\beta = \frac{3\pi}{2} \pm \alpha$	$\beta = 2\pi - \alpha$
$\operatorname{sen}\beta$	$\cos\alpha$	$\mp\operatorname{sen}\alpha$	$-\cos\alpha$	$-\operatorname{sen}\alpha$
$\cos\beta$	$\mp\operatorname{sen}\alpha$	$-\cos\alpha$	$\pm\operatorname{sen}\alpha$	$\cos\alpha$
$\operatorname{tg}\beta$	$\mp\operatorname{ctg}\alpha$	$\pm\operatorname{tg}\alpha$	$\mp\operatorname{ctg}\alpha$	$-\operatorname{tg}\alpha$
$\operatorname{ctg}\beta$	$\mp\operatorname{tg}\alpha$	$\pm\operatorname{ctg}\alpha$	$\mp\operatorname{tg}\alpha$	$-\operatorname{ctg}\alpha$

(véase la tabla 3), las cuales se puede formular en forma de la regla siguiente:

Si en la fórmula de reducción el ángulo α se sustrae de $\pi/2$ o se suma a $\pi/2$, tomado un número impar de veces, entonces la función que reducimos se cambia por la cofunción *); si $\pi/2$ está tomado un número par de veces, entonces el nombre de la función se conserva. Al mismo tiempo, delante de la función reducida se pone el mismo signo que tiene la función reducida en el cuadrante respectivo, si consideramos el ángulo α como agudo.

2.2. Relación entre las funciones trigonométricas de un mismo argumento. En la tabla 4 se muestran las fórmulas que expresan las dependencias entre las funciones trigonométricas de un mismo argumento. En las fórmulas citadas aquí, delante del signo del radical debe ser elegido el signo «más» o «menos» en dependencia del cuadrante, en el cual

*) El coseno es cofunción respecto al seno, y, a la inversa. Otro par de cofunciones es la tangente y cotangente.

Tabla 4

	$\operatorname{sen}\alpha$	$\cos\alpha$	$\operatorname{tg}\alpha$	$\operatorname{ctg}\alpha$	$\sec\alpha$	$\operatorname{cosec}\alpha$
$\operatorname{sen}\alpha$		$=\pm\sqrt{1-\cos^2\alpha}$	$=\dfrac{\operatorname{tg}\alpha}{\pm\sqrt{1+\operatorname{tg}^2\alpha}}$	$=\dfrac{1}{\pm\sqrt{1+\operatorname{ctg}^2\alpha}}$	$=\dfrac{\pm\sqrt{\sec^2\alpha-1}}{\sec\alpha}$	$=\dfrac{1}{\operatorname{cosec}\alpha}$
$\cos\alpha$	$=\pm\sqrt{1-\operatorname{sen}^2\alpha}$		$=\dfrac{1}{\pm\sqrt{1+\operatorname{tg}^2\alpha}}$	$=\dfrac{\operatorname{ctg}\alpha}{\pm\sqrt{1+\operatorname{ctg}^2\alpha}}$	$=\dfrac{1}{\sec\alpha}$	$=\dfrac{\pm\sqrt{\operatorname{cosec}^2\alpha-1}}{\operatorname{cosec}\alpha}$
$\operatorname{tg}\alpha$	$=\dfrac{\operatorname{sen}\alpha}{\pm\sqrt{1-\operatorname{sen}^2\alpha}}$	$=\dfrac{\pm\sqrt{1-\cos^2\alpha}}{\cos\alpha}$		$=\dfrac{1}{\operatorname{ctg}\alpha}$	$=\pm\sqrt{\sec^2\alpha-1}$	$=\dfrac{1}{\pm\sqrt{\operatorname{cosec}^2\alpha-1}}$
$\operatorname{ctg}\alpha$	$=\dfrac{\pm\sqrt{1-\operatorname{sen}^2\alpha}}{\operatorname{sen}\alpha}$	$=\dfrac{\cos\alpha}{\pm\sqrt{1-\cos^2\alpha}}$	$=\dfrac{1}{\operatorname{tg}\alpha}$		$=\dfrac{1}{\pm\sqrt{\sec^2\alpha-1}}$	$=\pm\sqrt{\operatorname{cosec}^2\alpha-1}$
$\sec\alpha$	$=\dfrac{1}{\pm\sqrt{1-\operatorname{sen}^2\alpha}}$	$=\dfrac{1}{\cos\alpha}$	$=\pm\sqrt{1+\operatorname{tg}^2\alpha}$	$=\dfrac{\pm\sqrt{1+\operatorname{ctg}^2\alpha}}{\operatorname{ctg}\alpha}$		$=\dfrac{\operatorname{cosec}\alpha}{\pm\sqrt{\operatorname{cosec}^2\alpha-1}}$
$\operatorname{cosec}\alpha$	$=\dfrac{1}{\operatorname{sen}\alpha}$	$=\dfrac{1}{\pm\sqrt{1-\cos^2\alpha}}$	$=\dfrac{\pm\sqrt{1+\operatorname{tg}^2\alpha}}{\operatorname{tg}\alpha}$	$=\pm\sqrt{1+\operatorname{ctg}^2\alpha}$	$=\dfrac{\sec\alpha}{\pm\sqrt{\sec^2\alpha-1}}$	

se encuentra el ángulo α, precisamente de tal manera que el signo de la función trigonométrica que se encuentra en el primer miembro coincida con el signo del valor que se encuentra en el segundo miembro de la igualdad.

2.3. Funciones trigonométricas de la suma y de la diferencia de los ángulos

$$\operatorname{sen}(\alpha \pm \beta) = \operatorname{sen}\alpha \cos\beta \pm \cos\alpha \operatorname{sen}\beta;$$

$$\cos(\alpha \pm \beta) = \cos\alpha \cos\beta \mp \operatorname{sen}\alpha \operatorname{sen}\beta;$$

$$\operatorname{tg}(\alpha \pm \beta) = \frac{\operatorname{tg}\alpha \pm \operatorname{tg}\beta}{1 \mp \operatorname{tg}\alpha \operatorname{tg}\beta};$$

$$\operatorname{ctg}(\alpha \pm \beta) = \frac{\operatorname{ctg}\alpha \operatorname{ctg}\beta \mp 1}{\operatorname{ctg}\alpha \pm \operatorname{ctg}\beta}.$$

2.4. Funciones trigonométricas de los ángulos dobles, triples y de los semiángulos.

$$\operatorname{sen} 2\alpha = 2 \operatorname{sen}\alpha \cos\alpha;$$

$$\cos 2\alpha = \cos^2\alpha - \operatorname{sen}^2\alpha = 1 - 2\operatorname{sen}^2\alpha = 2\cos^2\alpha - 1;$$

$$\operatorname{tg} 2\alpha = \frac{2\operatorname{tg}\alpha}{1-\operatorname{tg}^2\alpha}; \quad \operatorname{ctg} 2\alpha = \frac{\operatorname{ctg}^2\alpha - 1}{2\operatorname{ctg}\alpha};$$

$$\operatorname{sen} 3\alpha = 3\operatorname{sen}\alpha - 4\operatorname{sen}^3\alpha;$$

$$\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha;$$

$$\operatorname{tg} 3\alpha = \frac{3\operatorname{tg}\alpha - \operatorname{tg}^3\alpha}{1 - 3\operatorname{tg}^2\alpha}; \quad \operatorname{ctg} 3\alpha = \frac{\operatorname{ctg}^3\alpha - 3\operatorname{ctg}\alpha}{3\operatorname{ctg}^2\alpha - 1};$$

$$\operatorname{sen}\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}; \quad \cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}};$$

$$\operatorname{tg}\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}} = \frac{\operatorname{sen}\alpha}{1+\cos\alpha} = \frac{1-\cos\alpha}{\operatorname{sen}\alpha};$$

$$\operatorname{ctg}\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{1-\cos\alpha}} = \frac{\operatorname{sen}\alpha}{1-\cos\alpha} = \frac{1+\cos\alpha}{\operatorname{sen}\alpha}.$$

En las fórmulas del ángulo mitad los signos delante de los radicales se toman en dependencia del signo de la función trigonométrica que se encuentra en el primer miembro de la igualdad.

2.5. Transformación de la suma (diferencia) de las funciones trigonométricas en producto (transformación de las expresiones trigonométricas en una forma, que sea cómoda

para la determinación por logaritmos).

$$\operatorname{sen}\alpha+\operatorname{sen}\beta=2\operatorname{sen}\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2};$$

$$\operatorname{sen}\alpha-\operatorname{sen}\beta=2\cos\frac{\alpha+\beta}{2}\operatorname{sen}\frac{\alpha-\beta}{2};$$

$$\cos\alpha+\cos\beta=2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2};$$

$$\cos\alpha-\cos\beta=-2\operatorname{sen}\frac{\alpha+\beta}{2}\operatorname{sen}\frac{\alpha-\beta}{2}=$$

$$=2\operatorname{sen}\frac{\alpha+\beta}{2}\operatorname{sen}\frac{\beta-\alpha}{2};$$

$$\cos\alpha+\operatorname{sen}\alpha=\sqrt{2}\cos(45^\circ-\alpha);$$

$$\cos\alpha-\operatorname{sen}\alpha=\sqrt{2}\operatorname{sen}(45^\circ-\alpha);$$

$$\operatorname{tg}\alpha\pm\operatorname{tg}\beta=\frac{\operatorname{sen}(\alpha\pm\beta)}{\cos\alpha\cos\beta};$$

$$\operatorname{ctg}\alpha\pm\operatorname{ctg}\beta=\frac{\operatorname{sen}(\beta\pm\alpha)}{\operatorname{sen}\alpha\operatorname{sen}\beta};$$

$$\operatorname{tg}\alpha+\operatorname{ctg}\beta=\frac{\cos(\alpha-\beta)}{\cos\alpha\operatorname{sen}\beta};\quad \operatorname{tg}\alpha-\operatorname{ctg}\beta=-\frac{\cos(\alpha+\beta)}{\cos\alpha\operatorname{sen}\beta};$$

$$\operatorname{tg}\alpha+\operatorname{ctg}\alpha=2\operatorname{cosec}2\alpha;\quad \operatorname{tg}\alpha-\operatorname{ctg}\alpha=-2\operatorname{ctg}2\alpha;$$

$$1+\cos\alpha=2\cos^2\frac{\alpha}{2};\quad 1-\cos\alpha=2\operatorname{sen}^2\frac{\alpha}{2};$$

$$1+\operatorname{sen}\alpha=2\cos^2\left(45^\circ-\frac{\alpha}{2}\right);$$

$$1-\operatorname{sen}\alpha=2\operatorname{sen}^2\left(45^\circ-\frac{\alpha}{2}\right);$$

$$1\pm\operatorname{tg}\alpha=\frac{\operatorname{sen}(45^\circ\pm\alpha)}{\cos45^\circ\cos\alpha}=\frac{\sqrt{2}\operatorname{sen}(45^\circ\pm\alpha)}{\cos\alpha};$$

$$1\pm\operatorname{tg}\alpha\operatorname{tg}\beta=\frac{\cos(\alpha\mp\beta)}{\cos\alpha\cos\beta};\quad \operatorname{ctg}\alpha\operatorname{ctg}\beta\pm1=\frac{\cos(\alpha\mp\beta)}{\operatorname{sen}\alpha\operatorname{sen}\beta};$$

$$1-\operatorname{tg}^2\alpha=\frac{\cos2\alpha}{\cos^2\alpha};\quad 1-\operatorname{ctg}^2\alpha=-\frac{\cos2\alpha}{\operatorname{sen}^2\alpha};$$

$$\operatorname{tg}^2\alpha-\operatorname{tg}^2\beta=\frac{\operatorname{sen}(\alpha+\beta)\operatorname{sen}(\alpha-\beta)}{\cos^2\alpha\cos^2\beta};$$

$$\operatorname{ctg}^2\alpha-\operatorname{ctg}^2\beta=\frac{\operatorname{sen}(\alpha+\beta)\operatorname{sen}(\beta-\alpha)}{\operatorname{sen}^2\alpha\operatorname{sen}^2\beta};$$

$$\operatorname{tg}^2\alpha-\operatorname{sen}^2\alpha=\operatorname{tg}^2\alpha\operatorname{sen}^2\alpha;\quad \operatorname{ctg}^2\alpha-\cos^2\alpha=\operatorname{ctg}^2\alpha\cos^2\alpha.$$

2.6. Transformación del producto de las funciones trigonométricas en suma.

$$\operatorname{sen}\alpha\operatorname{sen}\beta=\frac{1}{2}[\cos(\alpha-\beta)-\cos(\alpha+\beta)];$$

$$\cos\alpha\cos\beta=\frac{1}{2}[\cos(\alpha-\beta)+\cos(\alpha+\beta)];$$

$$\operatorname{sen}\alpha\cos\beta=\frac{1}{2}[\operatorname{sen}(\alpha-\beta)+\operatorname{sen}(\alpha+\beta)];$$

$$\operatorname{sen}\alpha\operatorname{sen}\beta\operatorname{sen}\gamma=\frac{1}{4}[\operatorname{sen}(\alpha+\beta-\gamma)+\operatorname{sen}(\beta+\gamma-\alpha)+$$
$$+\operatorname{sen}(\gamma+\alpha-\beta)-\operatorname{sen}(\alpha+\beta+\gamma)];$$

$$\operatorname{sen}\alpha\cos\beta\cos\gamma=\frac{1}{4}[\operatorname{sen}(\alpha+\beta-\gamma)-\operatorname{sen}(\beta+\gamma-\alpha)+$$
$$+\operatorname{sen}(\gamma+\alpha-\beta)+\operatorname{sen}(\alpha+\beta+\gamma)];$$

$$\operatorname{sen}\alpha\operatorname{sen}\beta\cos\gamma=\frac{1}{4}[-\cos(\alpha+\beta-\gamma)+\cos(\beta+\gamma-\alpha)+$$
$$+\cos(\gamma+\alpha-\beta)-\cos(\alpha+\beta+\gamma)];$$

$$\cos\alpha\cos\beta\cos\gamma=\frac{1}{4}[\cos(\alpha+\beta-\gamma)+\cos(\beta+\gamma-\alpha)+$$
$$+\cos(\gamma+\alpha-\beta)+\cos(\alpha+\beta+\gamma)].$$

2.7. Relaciones simples entre las funciones trigonométricas inversas.

$$\operatorname{arcsen}\alpha=-\operatorname{arcsen}(-\alpha)=\frac{\pi}{2}-\arccos\alpha=\operatorname{arctg}\frac{\alpha}{\sqrt{1-\alpha^2}};$$

$$\arccos\alpha=\pi-\arccos(-\alpha)=\frac{\pi}{2}-\operatorname{arcsen}\alpha=\operatorname{arcctg}\frac{\alpha}{\sqrt{1-\alpha^2}};$$

$$\operatorname{arctg}\alpha=-\operatorname{arctg}(-\alpha)=\frac{\pi}{2}-\operatorname{arcctg}\alpha=\operatorname{arcsen}\frac{\alpha}{\sqrt{1+\alpha^2}};$$

$$\operatorname{arcctg}\alpha=\pi-\operatorname{arcctg}(-\alpha)=\frac{\pi}{2}-\operatorname{arctg}\alpha=\arccos\frac{\alpha}{\sqrt{1+\alpha^2}}.$$

Unas dependencias mucho más complicadas entre las funciones trigonométricas inversas pueden ser establecidas mediante los métodos que aplican para la solución de ecuaciones que contienen funciones trigonométricas inversas (véase p. 3.4).

Ejemplo. Expresar la suma de dos funciones trigonométricas inversas

$$\arcsen x + \arcsen y$$

a través de una función trigonométrica inversa (por ejemplo, arco seno).

Designemos la suma dada más arriba con la letra z:

$$\arcsen x + \arcsen y = z. \tag{1}$$

La igualdad dada puede considerarse como una ecuación con tres variables x, y, y z. La ecuación (1) es consecuencia de la ecuación siguiente:

$$\operatorname{sen}(\arcsen x + \arcsen y) = \operatorname{sen} z,$$

la cual se reduce a la forma

$$\operatorname{sen}(\arcsen x)\cdot\cos(\arcsen y) + \operatorname{sen}(\arcsen y)\cos(\arcsen x) =$$

$$\operatorname{sen} z \Leftrightarrow x\sqrt{1-y^2} + y\sqrt{1-x^2} = \operatorname{sen} z.$$

Ahora es necesario resolver la ecuación respecto a la variable z, teniendo en cuenta que esta variable pertenece al intervalo $[-\pi;\ \pi]$.

El conjunto de soluciones de esta ecuación tiene la forma

$$z = \arcsen\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$$

para $xy \leqslant 0$ ó $x^2 + y^2 \leqslant 1$;

$$z = \pi - \arcsen\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$$

para $x > 0$, $y > 0$ y $x^2 + y^2 > 1$;

$$z = -\pi - \arcsen\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$$

para $x < 0, y < 0$ y $x^2 + y^2 > 1$.

Las expresiones obtenidas para la variable z dan la expresión buscada de la suma de dos funciones trigonométricas inversas a través de una función trigonométrica inversa.

§ 3. Solución de ecuaciones trigonométricas y desigualdades

3.1. Ecuaciones trigonométricas simples.

Solución de la ecuación sen $x = a$. Examinemos una función $y =$ sen x en el intervalo $[-\pi/2; \pi/2]$. En este intervalo la función $y =$ sen x crece, cambiando desde su valor mínimo igual a -1, en el extremo izquierdo del intervalo,

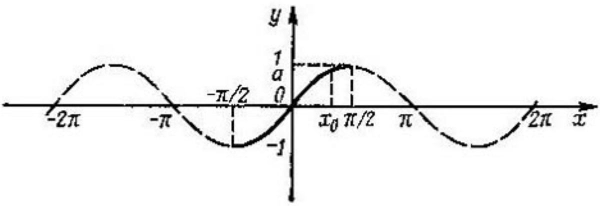


Fig. 7.18.

hasta su valor máximo igual a 1, en el extremo derecho del intervalo (fig. 7.18). Debido a la continuidad de la función $y =$ sen x, a cada valor $y = a$ que satisface la condición $|a| \leqslant 1$ le corresponde un solo valor $x_0 \in [-\pi/2; \pi/2]$ tal, que

$$\text{sen } x_0 = a.$$

El ángulo x_0 es el arco seno del número a (se designa arcsen a).

La ecuación de la forma sen $x = a$ para $|a| \leqslant 1$ tiene un conjunto de soluciones $x = (-1)^n$ arcsen $a + \pi n$, $n \in \mathbf{Z}$; para $|a| > 1$ la ecuación no tiene soluciones (es decir, el conjunto de soluciones es un conjunto vacío).

Solución de la ecuación cos $x = a$. Examinemos la función $y = \cos x$ en el intervalo $[0; \pi]$. En este intervalo la función $y = \cos x$ decrece, cambiando desde su valor máximo, igual a 1, en el extremo izquierdo del intervalo hasta su valor mínimo, igual a -1, en el extremo derecho del intervalo (fig. 7.19). Debido a la continuidad de la función $y = \cos x$, a cada valor $y = a$ que satisface la condición $|a| \leqslant 1$ le corresponde un solo valor $x_0 \in [0; \pi]$ tal, que

$$\cos x_0 = a.$$

El ángulo x_0 es el arco coseno del número a (se designa arccos a).

La ecuación de la forma $\cos x = a$ para $|a| \leqslant 1$ tiene un conjunto de soluciones $x = \pm \arccos a + 2\pi n$, $n \in \mathbf{Z}$;

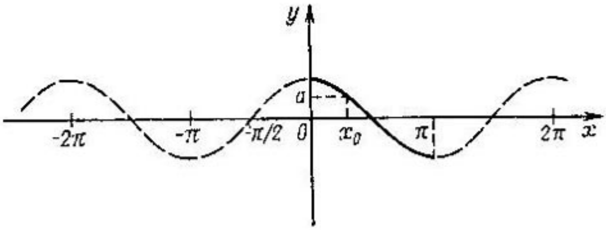


Fig. 7.19.

para $|a| > 1$ la ecuación no tiene soluciones (el conjunto de soluciones es un conjunto vacío).

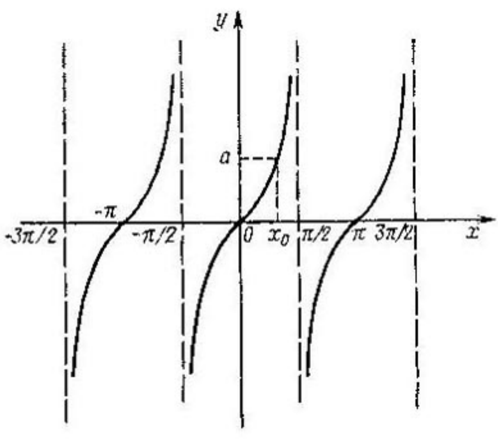


Fig. 7.20.

Solución de la ecuación $\operatorname{tg} x = a$. La función $y = \operatorname{tg} x$ en el intervalo $(-\pi/2; \pi/2)$ crece y obtiene todos sus valores (fig. 7.20). En vigor de la continuidad de la función $y = \operatorname{tg} x$ a cada valor $y = a$ le corresponde un solo valor

$x_0 \in (-\pi/2;\ \pi/2)$ tal, que

$$\operatorname{tg} x_0 = a.$$

El ángulo x_0 es el arco tangente del número a (se designa arctg a).

La ecuación de la forma tg $x = a$ para cualquier a tiene un conjunto de soluciones

$$x = \operatorname{arctg} a + \pi n,\ n \in \mathbf{Z}.$$

Solución de la ecuación ctg $x = a$. La función $y = \operatorname{ctg} x$ en el intervalo $(0;\ \pi)$ decrece y obtiene todos sus valores

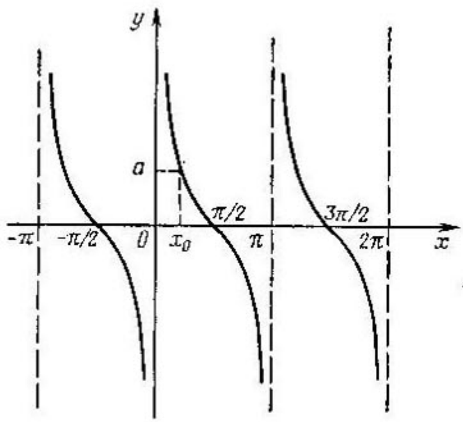


Fig. 7.21.

(fig. 7.21). Debido a la continuidad de la función $y = \operatorname{ctg} x$, a cada valor $y = a$ le corresponde un solo valor $x_0 \in (0;\pi)$ tal, que

$$\operatorname{ctg} x_0 = a.$$

El ángulo x_0 es el arco cotangente del número a (se designa arcctg a).

La ecuación de la forma ctg $x = a$ para cualquier a tiene un conjunto de soluciones $x = \operatorname{arcctg} a + \pi n,\ n \in \mathbf{Z}$.

3.2. Ejemplos de ecuaciones trigonométricas más comlicadas.

1) La ecuación trigonométrica de la forma

$$R\ (\operatorname{sen} kx, \cos nx, \operatorname{tg} mx, \operatorname{ctg} lx) = 0, \qquad (1)$$

donde R es el polinomio de los argumentos indicados (k, n, m y l son números naturales), con ayuda de las fórmulas para las funciones trigonométricas de la suma de los ángulos (en particular, de las fórmulas del ángulo doble y triple) se puede reducir a una ecuación racional respecto a los argumentos $\operatorname{sen} x$, $\cos x$, $\operatorname{tg} x$ y $\operatorname{ctg} x$, después de lo cual la ecuación (1) puede ser reducida a la ecuación racional respecto a la incógnita $t = \operatorname{tg}\frac{x}{2}$, mediante las fórmulas de *sustitución universal*

$$\operatorname{sen} x = \frac{2\operatorname{tg}\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}}; \quad \cos x = \frac{1-\operatorname{tg}^2\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}};$$

$$\operatorname{tg} x = \frac{2\operatorname{tg}\frac{x}{2}}{1-\operatorname{tg}^2\frac{x}{2}}; \quad \operatorname{ctg} x = \frac{1-\operatorname{tg}^2\frac{x}{2}}{2\operatorname{tg}\frac{x}{2}};$$

Ejemplo 1. Resolver la ecuación

$$\operatorname{sen} 2x + \operatorname{tg} x = 2.$$

Expresando $\operatorname{sen} 2x$ a través de $\operatorname{tg} x$, obtenemos la ecuación

$$\frac{2\operatorname{tg} x}{1+\operatorname{tg}^2 x} + \operatorname{tg} x = 2,$$

la cual se reduce a la ecuación cúbica de una variable relativa $t = tg\ x$:

$$t^3 - 2t^2 + 3t - 2 = 0 \Leftrightarrow (t-1)(t^2 - t + 2) = 0.$$

Es fácil convencerse que la ecuación dada tiene sólo una raíz real: $t = 1$.

Resolviendo la ecuación $\operatorname{tg} x = 1$, obtenemos un conjunto de soluciones de la ecuación inicial:

$$x = \frac{\pi}{4} + \pi k, \quad k \in Z.$$

Observemos, que el método general expuesto aquí de reducción de la ecuación trigonométrica a una ecuación entera no siempre es cómodo, puesto que en una serie de casos puede reducir la solución de la ecuación trigonométrica a la obtención de raíces de un polinomio de una potencia bastante elevada.

2) La ecuación que tiene la forma

$$R\,(\operatorname{sen} x + \cos x,\ \operatorname{sen} x \cos x) = 0, \tag{2}$$

donde R es el polinomio de los argumentos indicados, puede ser reducida a la ecuación respecto a la incógnita $t = \operatorname{sen} x + \cos x$, si aplicamos la identidad trigonométrica

$$(\operatorname{sen} x + \cos x)^2 = \operatorname{sen}^2 x + \cos^2 x + 2 \operatorname{sen} x \cos x = \\ = 1 + 2 \operatorname{sen} x \cos x,$$

de la cual se desprende que

$$\operatorname{sen} x \cos x = \frac{t^2-1}{2}. \tag{3}$$

Teniendo en cuenta la relación (3), la ecuación (2) se puede reducir a la forma

$$R\left(t,\ \frac{t^2-1}{2}\right) = 0.$$

Análogamente, la ecuación que tiene la forma

$$R\,(\operatorname{sen} x - \cos x,\ \operatorname{sen} x \cos x) = 0$$

por sustitución de $\operatorname{sen} x - \cos x = t$ se reduce a la ecuación

$$R\left(t,\ \frac{1-t^2}{2}\right) = 0.$$

Ejemplo 2. Resolver la ecuación

$$\operatorname{sen} x + \cos x - 2\sqrt{2} \operatorname{sen} x \cos x = 0.$$

Designando $\operatorname{sen} x + \cos x = t$ y utilizando la igualdad

$$\operatorname{sen} x \cos x = \frac{t^2-1}{2},$$

reducimos la ecuación a la ecuación cuadrática siguiente respecto a la incógnita t:

$$\sqrt{2}\, t^2 - t - \sqrt{2} = 0;$$

las raíces de esta ecuación cuadrática son los números

$$t_1=\sqrt{2} \text{ y } t_2=-1/\sqrt{2}.$$

De esta manera, la solución de la ecuación inicial se reduce a la solución de dos ecuaciones trigonométricas más simples:

$$\operatorname{sen} x+\cos x=\sqrt{2} \text{ y } \operatorname{sen} x+\cos x=-\frac{1}{\sqrt{2}}.$$

Multiplicando ambos miembros de las ecuaciones obtenidas por el número $1/\sqrt{2}$, reducimos las ecuaciones a dos ecuaciones trigonométricas más simples:

$$\frac{1}{\sqrt{2}}\operatorname{sen} x+\frac{1}{\sqrt{2}}\cos x=1 \Leftrightarrow \operatorname{sen} x\cos\frac{\pi}{4}+$$
$$+\operatorname{sen}\frac{\pi}{4}\cos x=1 \Leftrightarrow \operatorname{sen}\left(x+\frac{\pi}{4}\right)=1,$$
$$\frac{1}{\sqrt{2}}\operatorname{sen} x+\frac{1}{\sqrt{2}}\cos x=-\frac{1}{2} \Leftrightarrow \operatorname{sen}\left(x+\frac{\pi}{4}\right)=-\frac{1}{2}.$$

Los conjuntos de soluciones de las ecuaciones $\operatorname{sen}\left(x+\frac{\pi}{4}\right)= 1$ y $\operatorname{sen}\left(x+\frac{\pi}{4}\right)=-\frac{1}{2}$ respectivamente tendrán la forma

$$x=\frac{\pi}{4}+2\pi k, \quad k\in \mathbf{Z},$$
$$x=(-1)^{n+1}\frac{\pi}{6}-\frac{\pi}{4}+\pi n, \quad n\in \mathbf{Z}.$$

3) La ecuación que tiene la forma $a \operatorname{sen} x+b\cos x=c$ (donde a, b y c son ciertos números) puede ser resuelta mediante las fórmulas de sustitución universal (véase 1).

Mostremos un procedimiento más de solución de esta ecuación (el cual a veces se llama método del *ángulo complementario*).

Dividamos ambos miembros de la ecuación inicial entre $\sqrt{a^2+b^2}\neq 0$. (Si $a=b=0$, entonces la ecuación se convierte bien en identidad (para $c=0$) o bien no tiene soluciones (para $c\neq 0$)). De resultas, después de la división obtenemos una ecuación, equivalente a la inicial:

$$\frac{a}{\sqrt{a^2+b^2}} \operatorname{sen} x + \frac{b}{\sqrt{a^2+b^2}} \cos x = \frac{c}{\sqrt{a^2+b^2}}. \qquad (4)$$

Es fácil verificar que los coeficientes $a/\sqrt{a^2+b^2}$ y y $b/\sqrt{a^2+b^2}$ se relacionan mediante la igualdad

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

y por eso se pueden considerar como valores del seno y coseno de cierto ángulo φ:

$$\frac{a}{\sqrt{a^2+b^2}} = \operatorname{sen} \varphi, \quad \frac{b}{\sqrt{a^2+b^2}} = \cos \varphi.$$

Así pues, la ecuación (4) se reduce a la ecuación

$$\cos x \cos \varphi + \operatorname{sen} x \operatorname{sen} \varphi = \frac{c}{\sqrt{a^2+b^2}} \Leftrightarrow \cos (x-\varphi) = \frac{c}{\sqrt{a^2+b^2}},$$

cuyo conjunto de soluciones para $|c| \leqslant \sqrt{a^2+b^2}$ tiene la forma

$$x = \varphi \pm \arccos \frac{c}{\sqrt{a^2+b^2}} + 2\pi n, \quad n \in Z.$$

Queda por hallar cierto valor del ángulo φ, que es la solución del sistema de ecuaciones trigonométricas

$$\operatorname{sen} \varphi = \frac{a}{\sqrt{a^2+b^2}},$$

$$\cos \varphi = \frac{b}{\sqrt{a^2+b^2}}.$$

Aquí son posibles los casos siguientes (para φ se han elegido las más simples de varias de las expresiones posibles):

Si $a > 0$, $b > 0$, entonces

$$\varphi = \operatorname{arcsen} \frac{a}{\sqrt{a^2+b^2}} = \arccos \frac{b}{\sqrt{a^2+b^2}} = \operatorname{arctg} \frac{a}{b};$$

si $a > 0$, $b < 0$, entonces $\varphi = \arccos \dfrac{b}{\sqrt{a^2+b^2}}$;

si $a < 0$, $b < 0$, entonces $\varphi = \pi + \operatorname{arctg} \dfrac{a}{b}$;

si $a < 0$, $b > 0$, entonces $\varphi = \operatorname{arcsen} \dfrac{a}{\sqrt{a^2+b^2}} = \operatorname{arctg} \dfrac{a}{b}$.

4) La ecuación que tiene la forma

$$\operatorname{sen}^p x + \cos^k x = 1 \quad (p,\ k = 3,\ 4,\ 5,\ \ldots)$$

puede ser resuelta de la manera siguiente.

De la definición de las funciones sen x y cos x se deduce que para todos los valores del argumento $x \neq \frac{\pi n}{2}$, $n \in \mathbf{Z}$, existen las desigualdades

$$\operatorname{sen}^p x \leqslant |\operatorname{sen} x|^p < \operatorname{sen}^2 x,$$
$$\cos^k x \leqslant |\cos x|^k < \cos^2 x,$$
$$\operatorname{sen}^p x + \cos^k x \leqslant |\operatorname{sen} x|^p + |\cos x|^k < 1.$$

Por consiguiente, las soluciones de la ecuación inicial pueden ser sólo números del conjunto de valores del argumento $x = \pi n/2$, $n \in \mathbf{Z}$.

Luego, el proceso de obtención de raíces se reduce a la verificación, de cuáles de los valores indicados del argumento son raíces de la ecuación dada.

3.3. Solución de las desigualdades trigonométricas simples.

Tabla 5

Tipo de desigualdad	Conjunto de soluciones de la desigualdad ($n \in \mathbf{Z}$)
$\operatorname{sen} x > a$ $(\lvert a\rvert < 1)$	$x \in (\operatorname{arcsen} a + 2\pi n;\ \pi - \operatorname{arcsen} a + 2\pi n)$
$\operatorname{sen} x < a$ $(\lvert a\rvert < 1)$	$x \in (-\pi - \operatorname{arcsen} a + 2\pi n;\ \operatorname{arcsen} a + 2\pi n)$
$\cos x > a$ $(\lvert a\rvert < 1)$	$x \in (-\arccos a + 2\pi n;\ \arccos a + 2\pi n)$
$\cos x < a$ $(\lvert a\rvert < 1)$	$x \in (\arccos a + 2\pi n;\ 2\pi - \arccos a + 2\pi n)$
$\operatorname{tg} x > a$	$x \in \left(\operatorname{arctg} a + \pi n;\ \frac{\pi}{2} + \pi n\right)$
$\operatorname{tg} x < a$	$x \in \left(-\frac{\pi}{2} + \pi n;\ \operatorname{arctg} a + \pi n\right)$
$\operatorname{ctg} x > a$	$x \in (\pi n;\ \operatorname{arcctg} a + \pi n)$
$\operatorname{ctg} x < a$	$x \in (\operatorname{arcctg} a + \pi n;\ \pi + \pi n)$

En la tabla 5 se muestra un conjunto de soluciones de las desigualdades trigonométricas simples. Las desigualdades trigonométricas más complicadas se resuelven mediante métodos, que son semejantes a los métodos de resolución de las ecuaciones trigonométricas.

3.4. Ejemplos de solución de ecuaciones y desigualdades que contienen funciones trigonométricas inversas.

Ejemplo 1. Resolver la ecuación

$$\operatorname{arcsen} x + \operatorname{arcsen} 2x = \frac{\pi}{3}. \tag{5}$$

El conjunto de valores admisibles de la incógnita x es el intervalo $[-1/2;\ 1/2]$. Designando $\operatorname{arcsen} 2x = \alpha$, $\operatorname{arcsen} x = \beta$, obtenemos

$$\begin{aligned} &\operatorname{sen} \alpha = 2x, \quad \alpha \in [-\pi/2;\ \pi/2]; \\ &\operatorname{sen} \beta = x, \quad \beta \in [-\pi/6;\ \pi/6]. \end{aligned} \tag{6}$$

En las nuevas denotaciones la ecuación (6) se escribe en la forma

$$\alpha + \beta = \frac{\pi}{3} \Longleftrightarrow \alpha = \frac{\pi}{3} - \beta. \tag{7}$$

De las condiciones (6) se deduce que los ángulos α y $\frac{\pi}{3} - \beta$ pertenecen al intervalo $[-\pi/2;\ \pi/2)$. Por consiguiente, la ecuación (7) es equivalente a la ecuación

$$\operatorname{sen} \alpha = \operatorname{sen}\left(\frac{\pi}{3} - \beta\right) \Longleftrightarrow 2x = \operatorname{sen} \frac{\pi}{3} \cos \beta - \operatorname{sen} \beta \cos \frac{\pi}{3}. \tag{8}$$

El seno y coseno de un mismo ángulo se relacionan por la igualdad

$$\operatorname{sen}^2 \beta + \cos^2 \beta = 1,$$

de la cual, teniendo en cuenta las condiciones (6), se desprende que

$$\cos \beta = \sqrt{1 - x^2}.$$

Por consiguiente, la ecuación (8) puede escribirse en forma de

$$2x = \frac{\sqrt{3}}{2} \sqrt{1 - x^2} - \frac{x}{2} \Longleftrightarrow 4x = \sqrt{3}\sqrt{1 - x^2} - x \Longleftrightarrow 5x = \sqrt{3}\sqrt{1 - x^2}.$$

Elevando ambos miembros de la última ecuación al cuadrado, hallamos la única solución de esta ecuación irracional:

$$x = \frac{1}{2}\sqrt{\frac{3}{7}}.$$

Esto es, precisamente, la solución de la ecuación trigonométrica inicial (5):

Ejemplo 2. Hallar el campo de definición de la función

$$f(x) = \text{arcsen}\ (\text{arcsen}\ x) + \text{arccos}\left(\frac{2\ \text{arccos}\ x}{\pi - 2}\right).$$

El campo de definición de la función dada se halla como un conjunto de soluciones del siguiente sistema de desigualdades:

$$-1 \leqslant x \leqslant 1,$$
$$-1 \leqslant \text{arcsen}\ x \leqslant 1,$$
$$-1 \leqslant \frac{2\ \text{arccos}\ x}{\pi - 2} \leqslant 1.$$

Resolvamos la segunda desigualdad del sistema:

$$-1 \leqslant \text{arcsen}\ x \leqslant 1. \tag{9}$$

Puesto que la región de cambio de la función arco seno es el intervalo $[-\pi/2;\ \pi/2]$, entonces todos los ángulos, que figuran en la desigualdad (9) (-1, arcsen x y 1) pertenecen al intervalo de la función seno creciente monótona. Por consiguiente, la desigualdad (9) es equivalente a la desigualdad sen $(-1) \leqslant$ sen (arcsen x) $\leqslant$ sen $1 \Leftrightarrow -$ sen $1 \leqslant x \leqslant$ sen 1.

Resolvamos la tercera desigualdad del sistema:

$$-1 \leqslant \frac{2\ \text{arccos}\ x}{\pi - 2} \leqslant 1. \tag{10}$$

Multiplicando todos los tres miembros de la desigualdad doble por el valor $\frac{\pi - 2}{2}$ (mayor que cero), obtenemos una desigualdad que es equivalente a la desigualdad (10):

$$-\frac{\pi - 2}{2} \leqslant \text{arccos}\ x \leqslant \frac{\pi - 2}{2}.$$

Esta doble desigualdad es equivalente al sistema de dos desigualdades

$$\text{arccos}\ x \geqslant -\frac{\pi - 2}{2},$$
$$\text{arccos}\ x \leqslant \frac{\pi - 2}{2},$$

además, la primera desigualdad es válida para todos los valores $x \in [-1; 1]$, puesto que el conjunto de valores de la función arcocoseno es el intervalo $[0; \pi]$.

Resolvamos la segunda desigualdad de sistema. Puesto que los valores que se encuentran en ambos miembros de la desigualdad pertenecen al decrecimiento monótono de la función $\cos x$, la desigualdad es equivalente a la desigualdad

$$\cos(\arccos x) \geqslant \cos\left(\frac{\pi}{2} - 1\right) = \operatorname{sen} 1.$$

De esta manera, para hallar el campo de definición de la función dada es necesario hallar un conjunto de valores x, que satisfacen el sistema

$$\begin{aligned} -1 \leqslant x &\leqslant 1, \\ -\operatorname{sen} 1 \leqslant x &\leqslant \operatorname{sen} 1, \\ x &\geqslant \operatorname{sen} 1. \end{aligned}$$

Este sistema tiene una sola solución $x = \operatorname{sen} 1$. Así pues, el campo de definición de la función dada $f(x)$ consta de un punto $x = \operatorname{sen} 1$.

§ 4. Relación entre los elementos del triángulo

4.1. Fórmulas fundamentales. Designaciones: a, b, c son los lados; α, β, γ, los ángulos (fig. 7.22); $p = \frac{a+b+c}{2}$, el semiperímetro; S, el área; R es el radio de la circunferencia circunscrita; r es el radio de la circunferencia inscrita; h, la altura; m, la mediana; l, la bisectriz; los índices a, b, c para h, m y l concretizan una de las tres alturas, medianas y bisectrices, respectivamente (por ejemplo, m_a es la mediana, trazada al lado a); r_a es el radio de la circunferencia circunscrita por fuera, tangente al lado a y a la prolongación de los lados b y c.

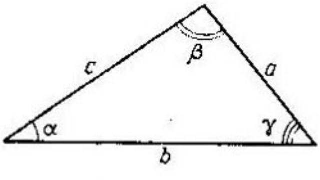


Fig. 7.22.

1. *Teorema de los cosenos.* El cuadrado de un lado del triángulo es igual a la suma de los cuadrados de los otros

dos lados menos el producto duplicado de estos lados por el coseno del ángulo entre ellos:

$$a^2 = b^2 + c^2 - 2bc \cos \alpha;$$
$$b^2 = a^2 + c^2 - 2ac \cos \beta;$$
$$c^2 = a^2 + b^2 - 2ab \cos \gamma.$$

2. *Teorema de los senos.* Los lados del triángulo son proporcionales a los senos de los ángulos opuestos:

$$\frac{a}{\operatorname{sen} \alpha} = \frac{b}{\operatorname{sen} \beta} = \frac{c}{\operatorname{sen} \gamma} = 2R.$$

3. *Teorema de las tangentes:*

$$\frac{a+b}{a-b} = \frac{\operatorname{tg} \frac{\alpha+\beta}{2}}{\operatorname{tg} \frac{\alpha-\beta}{2}} = \frac{\operatorname{ctg} \frac{\gamma}{2}}{\operatorname{tg} \frac{\alpha-\beta}{2}};$$

$$\frac{a+c}{a-c} = \frac{\operatorname{tg} \frac{\alpha+\gamma}{2}}{\operatorname{tg} \frac{\alpha-\gamma}{2}} = \frac{\operatorname{ctg} \frac{\beta}{2}}{\operatorname{tg} \frac{\alpha-\gamma}{2}};$$

$$\frac{b+c}{b-c} = \frac{\operatorname{tg} \frac{\beta+\gamma}{2}}{\operatorname{tg} \frac{\beta-\gamma}{2}} = \frac{\operatorname{ctg} \frac{\alpha}{2}}{\operatorname{tg} \frac{\beta-\gamma}{2}}.$$

4. *Fórmulas para calcular el área de un triángulo.*

$$S = \frac{1}{2} ab \operatorname{sen} \gamma;$$

$$S = \sqrt{p(p-a)(p-b)(p-c)};$$

$$S = p^2 \operatorname{tg} \frac{\alpha}{2} \operatorname{tg} \frac{\beta}{2} \operatorname{tg} \frac{\gamma}{2};$$

$$S = p(p-a) \operatorname{tg} \frac{\alpha}{2} = p(p-b) \operatorname{tg} \frac{\beta}{2} = p(p-c) \operatorname{tg} \frac{\gamma}{2};$$

$$S = \frac{abc}{4R};$$

$$S = pr;$$

$$S = \sqrt{r r_a r_b r_c}.$$

4.2. Cálculo de los elementos del triángulo. Un triángulo puede ser definido o por tres lados, o por un lado y dos ángulos, o por dos lados y el ángulo entre ellos. En calidad de terna de los elementos que definen el triángulo, se pueden elegir también algunas otras combinaciones de los elementos del triángulo. Por ejemplo, un triángulo puede ser defininido por la base a, la altura h_a y el ángulo de la base. Teniendo una terna de elementos que definen el triángulo, con ayuda del teorema de los senos y del teorema de los cosenos se pueden calcular todos los demás elementos del triángulo.

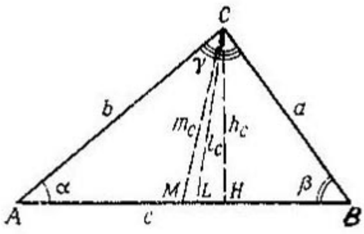


Fig. 7.23.

Ejemplo 1. Sea que el triángulo ABC está definido por tres lados a, b y c (fig. 7.23). Es necesario calcular todos los demás elementos del triángulo.

Los ángulos α y β del triángulo dado pueden ser hallados mediante el teorema de los cosenos:

$$a^2 = b^2 + c^2 - 2bc\cos\alpha \Rightarrow \cos\alpha =$$
$$= \frac{b^2+c^2-a^2}{2bc} \Rightarrow \alpha = \arccos\frac{b^2+c^2-a^2}{2bc},$$
$$b^2 = a^2 + c^2 - 2ac\cos\beta \Rightarrow \cos\beta =$$
$$= \frac{a^2+c^2-b^2}{2ac} \Rightarrow \beta = \arccos\frac{a^2+c^2-b^2}{2ac}.$$

El tercer ángulo del triángulo γ se puede calcular de la manera siguiente:

$$\alpha+\beta+\gamma = 180° \Rightarrow \gamma = 180° - \arccos\frac{b^2+c^2-a^2}{2bc} -$$
$$- \arccos\frac{a^2+c^2-b^2}{2ac}.$$

La altura del triángulo h_c, bajada al lado c (véase fig. 7.23) se halla de la manera siguiente:

$$\frac{h_c}{a} = \operatorname{sen}\beta \Rightarrow h_c = a\operatorname{sen}\beta = a\operatorname{sen}\left[\arccos\frac{a^2+c^2-b^2}{2ac}\right] \Rightarrow$$
$$\Rightarrow h_c = a\sqrt{1-\left(\frac{a^2+c^2-b^2}{2ac}\right)^2}.$$

Hallemos las medianas m_c valiéndonos de que el punto M divide el lado c por la mitad ($|AM| = |MB|$), y por eso en el triángulo MBC resultan conocidos los dos lados $|MB| = c/2$, $|BC| = a$ y el ángulo entre ellos β. Aplicando el teorema de los cosenos para el triángulo MBC, obtenemos

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

La bisectriz l_c puede hallarse del triángulo LBC. En el triángulo LBC son conocidos el lado BC ($|BC| = a$) y el ángulo β. El ángulo $LCB = \gamma/2$ es igual:

$$\frac{\gamma}{2} = 90^\circ - \frac{1}{2}\operatorname{arccos}\frac{b^2+c^2-a^2}{2bc} - \frac{1}{2}\operatorname{arccos}\frac{a^2+c^2-b^2}{2ac}.$$

De aquí

$$\angle CLB = 180^\circ - \angle LCB - \angle LBC = 180^\circ -$$
$$-\left[90^\circ - \frac{1}{2}\operatorname{arccos}\frac{b^2+c^2-a^2}{2bc} - \frac{1}{2}\operatorname{arccos}\frac{a^2+c^2-b^2}{2ac}\right] -$$
$$-\operatorname{arccos}\frac{a^2+c^2-b^2}{2ac} = 90^\circ + \frac{1}{2}\operatorname{arccos}\frac{b^2+c^2-a^2}{2bc} -$$
$$-\frac{1}{2}\operatorname{arccos}\frac{a^2+c^2-b^2}{2ac}.$$

Utilicemos ahora el teorema de los senos para hallar la bisectriz l_c;

$$\frac{l_c}{\operatorname{sen}\beta} = \frac{a}{\operatorname{sen}\angle CLB} \Rightarrow l_c = \frac{a\operatorname{sen}\beta}{\operatorname{sen}\angle CLB}.$$

Es fácil observar que el cálculo del seno del ángulo CLB que se encuentra en el denominador de la última fracción es un problema bastante difícil. La bisectriz del triángulo se puede calcular también mediante otro procedimiento.

Como se sabe, la bisectriz divide el lado del triángulo en una razón, proporcional a los lados adyacentes:

$$\frac{|AL|}{|AC|} = \frac{|LB|}{|CB|}.$$

Designando $|LB| = x$, esta relación se puede escribir en la forma

$$\frac{c-x}{b} = \frac{x}{a} \Rightarrow ac - ax = bx \Rightarrow ac = x(a+b) \Rightarrow x + \frac{ac}{a+b}.$$

Ahora en el triángulo LBC se conocen dos lados $|BC| = a$, $|LB| = \frac{ac}{a+b}$ y el ángulo entre ellos β. Usando el teorema de los cosenos obtenemos

$$l_c^2 = |BC|^2 + |LB|^2 - 2\,|BC|\cdot|LB|\cos\beta \Rightarrow l_c^2 =$$

$$= a^2 + \left(\frac{ac}{a+b}\right)^2 - \frac{2a^2c}{a+b}\cos\beta \Rightarrow$$

$$\Rightarrow l_c^2 = a^2 + \left(\frac{ac}{a+b}\right)^2 - \frac{2a^2c}{a+b}\cdot\frac{a^2+c^2-b^2}{2ac} \Rightarrow$$

$$\Rightarrow l_c = \sqrt{a^2 + \left(\frac{ac}{a+b}\right)^2 - \frac{a\,(a^2+c^2-b^2)}{a+b}}.$$

El área del triángulo dado ABC puede ser hallada, por ejemplo, mediante la fórmula de Gerón:

$$S_{\triangle ABC} = \sqrt{p\,(p-a)\,(p-b)\,(p-c)},$$

y los radios de las circunferencias inscrita y circunscrita pueden ser hallados de las fórmulas

$$S_{\triangle ABC} = pr \Rightarrow r = \frac{S_{\triangle ABC}}{p},$$

$$S_{\triangle ABC} = \frac{abc}{4R} \Rightarrow R = \frac{abc}{4S_{\triangle ABC}}.$$

Ejemplo 2. Examinemos ahora el triángulo, del cual se dan el lado a y dos ángulos α y β (fig. 7.24).

El tercer ángulo γ se encuentra de la manera siguiente: $\alpha + \beta + \gamma = 180° \Rightarrow \gamma = 180° - \alpha - \beta$. Los lados AB y AC del triángulo se hallan mediante el teorema de los senos:

$$\frac{|AB|}{\operatorname{sen}\gamma} = \frac{a}{\operatorname{sen}\alpha} = \frac{|AB|}{\operatorname{sen}(180° - \alpha - \beta)} = \frac{a}{\operatorname{sen}\alpha} \Rightarrow \frac{|AB|}{\operatorname{sen}(\alpha+\beta)} =$$

$$= \frac{a}{\operatorname{sen}\alpha} \Rightarrow |AB| = \frac{a\operatorname{sen}(\alpha+\beta)}{\operatorname{sen}\alpha},$$

$$\frac{|AC|}{\operatorname{sen}\beta} = \frac{a}{\operatorname{sen}\alpha} \Rightarrow |AC| = \frac{a\operatorname{sen}\beta}{\operatorname{sen}\alpha}.$$

De esta manera, en el triángulo hemos hallado tres lados; la obtención de todos los demás elementos del triángulo (mediana, altura, bisectriz, etc.) se puede hacer de la misma manera que en el problema anterior. Podemos observar, que el radio de la circunferencia circunscrita para el conjunto dado de elementos que definen el triángulo es cómodo hallarlo mediante la fórmula

$$\frac{a}{\operatorname{sen}\alpha} = 2R \Rightarrow R = \frac{a}{2\operatorname{sen}\alpha}.$$

Ejemplo 3. Si en el triángulo están dados dos lados b, c y el ángulo entre ellos α (fig. 7.25), entonces mediante el

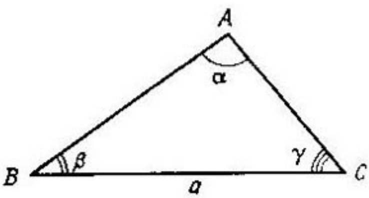


Fig. 7.24.

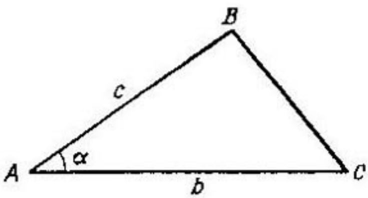


Fig. 7.25.

teorema de los cosenos se puede hallar el tercer lado del triángulo:

$$|BC|^2 = b^2 + c^2 - 2bc\cos\alpha,$$

y después, todos los demás elementos del triángulo (ángulos, alturas, medianas, bisectricas, etc.) se hallan de la misma manera que en el ejemplo 1.

CAPITULO 8

Teoría de los límites

Por primera vez la definición del concepto de un límite fue introducida en la obra de Wallis «Aritmética de los valores infinitos» (en el siglo XVII). Actualmente mediante el concepto de límite se efectúa la construcción estricta de todo el análisis matemático. Sin embargo, históricamente este concepto no fue la base del cálculo diferencial e integral. Los elementos del análisis matemático (del cálculo diferencial e integral) fueron elaborados por Newton, pero sólo en las obras de Cauchy (en el siglo XIX) la teoría de los límites sirvió de base para una argumentación estricta del análisis matemático.

§ 1. Sucesiones numéricas

1.1. Concepto de sucesión numérica. Examinemos un conjunto infinito de números

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

Consideremos que cada número de este conjunto tiene un número ordinal que corresponde al lugar que él mismo ocupa en la expresión (1). Así, consideremos que el número 1 que se encuentra en el primer lugar tiene un número ordinal 1, el número $-1/2$ que se encuentra en el segundo lugar tiene el número 2, el número $1/3$ tiene el número 3, etc. Y a la inversa, en vigor de este acuerdo sobre la numeración de un conjunto de números (1), cualquier número ordinal que sea tiene en el conjunto de los números (1) su único número ordinal indicado. Por ejemplo, en el conjunto (1) el número ordinal 4 tiene el número $-1/4$, el 10 tiene el número $-1/10$.

De esta manera, en el conjunto de los números (1) citado aquí, cada número tiene un número ordinal bien determinado y se define totalmente por este número.

Se llama *sucesión numérica* al conjunto de números, cada uno de los cuales tiene su número ordinal.

Los elementos de este conjunto numérico se llaman *términos* de la sucesión y se suelen denotar de la manera siguiente: el primer término a_1, el segundo a_2, el término n con a_n, etc. Toda la sucesión numérica se denota

$$a_1, a_2, a_3, \ldots, a_n, \ldots \text{o } (a_n).$$

La sucesión numérica

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

no es otra cosa que un conjunto de números numerados, ordenado como una serie natural, es decir, colocados en orden de crecimiento de los números. Una sucesión puede contener tanto un número de términos finito como infinito.

Una sucesión que se compone de un número finito de términos se llama *finita*, y la sucesión que se compone de un número (numerable) infinito de términos se llama sucesión *infinita*. Más adelante, en el capítulo presente, hablaremos sobre las sucesiones infinitas, las cuales llamaremos para mayor brevedad, *sucesiones*.

El concepto de una sucesión numérica puede ser introducido también mediante el concepto de aplicación de un conjunto numérico sobre el otro.

Sea **N** el conjunto de todos los números naturales, A cierto conjunto numérico numerable o finito. Demos la aplicación f del conjunto de todos los números naturales **N** sobre el conjunto A. Entonces a cada número natulal $n \in$ **N** le corresponde cierto elemento bien determinado del conjunto numérico A, el cual designamos con a_n, donde el índice n indica que el número a_n corresponde a n, y, a la inversa, a cada elemento a_n le corresponde por lo menos un número natural n del conjunto de todos los números naturales **N**.

La aplicación del conjunto **N** sobre cierto conjunto numérico numerable o finito A se llama *sucesión numérica infinita*; los números a_n se llaman *elementos* o *términos* de la sucesión numérica, y el índice n, *número* (ordinal) del término de la sucesión.

Representar una sucesión numérica infinita significa representar una aplicación, para la cual a cada número natural n le corresponde sólo un elemento del conjunto numérico A.

A veces la sucesión numérica infinita la introducen, utilizando el concepto de función: se llama sucesión numérica infinita a la función numérica, definida en el conjunto de todos los números naturales.

La sucesión (s_n), cuyos términos son sumas de los términos de las sucesiones (x_n) e (y_n) que tienen los mismos números ordinales:

$$s_n = x_n + y_n, \quad n \in \mathbf{N},$$

se llama *suma* de dos sucesiones (x_n) e (y_n).

Se llama *producto*, *diferencia* y *cociente* de dos sucesiones (x_n) e (y_n) respectivamente las sucesiones (m_n), (q_n) y (r_n), cuyos términos se calculan mediante las reglas:

$$m_n = x_n \cdot y_n;$$

$$q_n = x_n - y_n;$$

$$r_n = \frac{x_n}{y_n}, \quad n \in \mathbf{N},$$

y además, el cociente de dos sucesiones (es decir, la sucesión (r_n)) está definido sólo para tales sucesiones (y_n), en las cuales ningún término se anula.

1.2. Ciertos métodos de representación de las sucesiones.

1) *Método analítico de representación de una sucesión.* La sucesión se representa por una fórmula que permite por el número del término de la sucesión determinar este término. La fórmula que permite calcular cualquier término de una sucesión por su número, se llama *fórmula del término general* de una sucesión numérica.

Por ejemplo, las fórmulas del término general

$$a_n = \frac{(-1)^{n-1}}{n},$$

$$b_n = \operatorname{sen} n,$$

$$c_n = \frac{1+(-1)^n}{n}$$

representan sucesiones numéricas

$$1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ \ldots,$$
$$\operatorname{sen} 1,\ \operatorname{sen} 2,\ \operatorname{sen} 3,\ \operatorname{sen} 4,\ \ldots,$$
$$0,\ 1,\ 0,\ \frac{1}{2},\ 0,\ \frac{1}{3},\ \ldots$$

2) *Método recurrente de representación de una sucesión.* Se dan varios términos primeros de la sucesión y se indica la fórmula para calcular los términos sucesivos de la sucesión partiendo de los primeros términos dados.

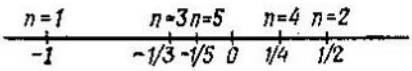


Fig. 8.1.

Ejemplos de representación recurrente de sucesiones:

$$a_1 = 1,\ a_{n+1} = n \cdot a_n,\ n \in \mathbf{N};$$
$$b_1 = 0,\ b_2 = 1, \quad b_{n+2} = b_n + b_{n+1}, \quad n \in \mathbf{N}.$$

La sucesión (b_n) se denomina sucesión de los *números de Fibonacci* (véase también § 4 del capítulo 1).

3) La sucesión puede ser representada también mediante la descripción de sus términos. Así, por ejemplo, se dice que la sucesión

$$3{,}1;\ 3{,}14;\ 3{,}141;\ 3{,}1415;\ \ldots$$

se forma de los valores aproximados del número π con insuficiencia de una precisión hasta 0,1; 0,01; 0,001; 0,0001, etc. En tales casos, como regla, es imposible indicar ni la fórmula del término general de la sucesión ni el método recurrente de calculación de sus términos.

1.3. Representación geométrica de los términos de una sucesión. Generalmente los términos de una sucesión se representan como puntos de una recta numérica. Por ejemplo, los cinco primeros términos de la sucesión $a_n = (-1)^n/n$ se muestran en la fig. 8.1. Más raramente los términos de una sucesión considerada como una función de un argumento natural, se representan mediante puntos en el sistema de coordenadas cartesiano rectangular Oxy. En la fig. 8.2, en

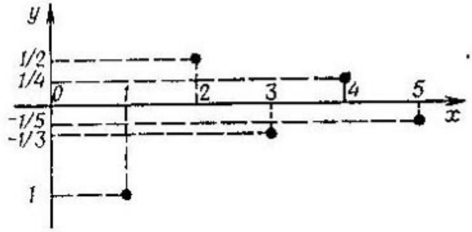


Fig. 8.2.

el sistema de coordenadas cartesiano rectangular, se muestran los cinco primeros términos de la sucesión $a_n = (-1)^n/n$.

1.4. Sucesiones acotadas. La sucesión (x_n) se llama *acotada*, si existen tales dos números m y M, que para todos $n \in \mathbf{N}$ se cumple la desigualdad

$$m \leqslant x_n \leqslant M. \tag{2}$$

La sucesión (x_n) se llama *acotada superiormente*, si existe tal número M, que para todos $n \in \mathbf{N}$ se cumple la desigualdad $x_n \leqslant M$.

La sucesión (x_n) se llama *acotada inferiormente*, si existe tal número m, que para todos $n \in \mathbf{N}$ se cumple la desigualdad

$$x_n \geqslant m.$$

Para que la sucesión sea acotada, es necesario y suficiente que ella sea acotada superiormente y acotada inferiormente. La demostración de la acotación de sucesión se reduce a hallar tales números m y M, para los cuales la desigualdad (2) es válida para todos $n \in \mathbf{N}$.

Ejemplos 1. La sucesión (x_n) que se representa mediante la fórmula del término general $x_n = n$, está acotada inferiormente (por ejemplo, por el número 0) y no está acotada superiormente.

2. La sucesión (x_n) que se representa mediante la fórmula del término general $x_n = -n^2/(n+1)$, está acotada superiormente (por) ejemplo, por el número $-1/2$) y no está acotada inferiormente.

3. Las sucesiones que se representan mediante las fórmulas del término general $x_n = \operatorname{sen} n$, $y_n = (-1)^n/n$, están acotadas por los números -1; 1 y -1; $1/2$ respectivamente.

1.5. Sucesiones monótonas. La sucesión (x_n) se llama *creciente*, si cada término suyo, comenzando por el segundo, es mayor que el anterior, es decir, si para cualquier n natural se cumple la desigualdad

$$x_{n+1} > x_n.$$

La sucesión (x_n) se llama *decreciente*, si cada término suyo, comenzando por el segundo, es menor que el anterior, es decir, si para cualquier n natural se cumple la desigualdad

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

La sucesión (x_n) se llama *no decreciente*, si cada término suyo, comenzando por el segundo, no es menor que el anterior, es decir, si para cualquier n natural se cumple la desigualdad

$$x_{n+1} \geqslant x_n.$$

La sucesión (x_n) se llama *no creciente*, si cada término suyo, comenzando por el segundo, no es mayor que el anterior, es decir, si para cualquier n natural se cumple la desigualdad

$$x_{n+1} \leqslant x_n.$$

Las sucesiones crecientes, decrecientes, no crecientes y no decrecientes forman la clase de las *sucesiones monótonas* *).

$$x_{n+1} > x_n \quad (x_{n+1} < x_n).$$

Tales sucesiones son monótones, comenzando por cierto número.

Ejemplos 1. Demostrar la monotonía de una sucesión que se representa mediante la fórmula del término general

$$x_n = \frac{3n+1}{2n-1}, \quad n \in \mathrm{N}.$$

Examinemos la diferencia entre el $(n+1)$-ésimo térmi-

*) A veces dan tal definición de *la sucesión monótona*: la sucesión (x_n) se llama monótona creciente (decreciente), si se halla un número N_0 tal, que para todos los términos de la sucesión con los números, mayores que N_0, se cumple la desigualdad $x_{n+1} > x_n$ ($x_{n+1} < x_n$). Tales sucesiones son monótonas, comenzando por cierto número.

no de la sucesión y su n-ésimo término:

$$x_{n+1}-x_n=\frac{3n+4}{2n+1}-\frac{3n+1}{2n-1}=\frac{-5}{(2n+1)(2n-1)}.$$

El numerador de la fracción que se encuentra en el segundo miembro de la igualdad es negativo, el denominador es positivo para cualquier n natural, y, por consiguiente, toda la fracción es negativa para cualquier $n \in \mathbf{N}$, y por lo tanto

$$x_{n+1}-x_n<0 \Leftrightarrow x_{n+1}<x_n,$$

es decir, la sucesión dada es decreciente.

2. Investigar la monotonía de una sucesión representada por la fórmula del término general

$$x_n = n^2 - 9n - 100.$$

Examinemos la diferencia entre los términos $(n + 1)$-ésimo y n-ésimo de la sucesión:

$$x_{n+1}-x_n = [(n+1)^2 - 9(n+1) - 100] - [n^2 - 9n - 100] = 2n - 8.$$

La expresión que se encuentra en el miembro segundo de la igualdad es negativa para $n = 1, 2, 3$, igual a cero para $n = 4$, y positiva para los demás valores de $n \in \mathbf{N}$.

Conforme a la definición de las sucesiones monótonas, es necesario que para todos n se cumpla bien la desigualdad $x_{n+1} < x_n$ (sucesión decreciente) o bien $x_{n+1} - x_n > 0$ (sucesión creciente), mientras que para la sucesión dada ni ésta ni la otra desigualdad se cumplen a la vez para todos los valores $n \in \mathbf{N}$, por lo que la sucesión que examinamos no es monótona. (La sucesión dada es monótona creciente comenzando por el quinto término.)

§ 2. Límite de una sucesión

2.1. Concepto de límite de sucesión. Examinemos una sucesión numérica infinita

$$1+\frac{1}{1},\ 1-\frac{1}{2},\ 1+\frac{1}{3},\ 1-\frac{1}{4},\ \ldots$$

que se representa por la fórmula del término general

$$x_n = 1+\frac{(-1)^{n-1}}{n},\quad n\in\mathbf{N}.$$

Representemos los términos de esta sucesión por puntos en una recta numérica (fig. 8.3). Con el crecimiento de los números los términos de la sucesión «se aproximan» cada vez más cerca al punto del eje numérico, que tiene la abscisa 1.

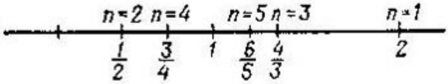


Fig. 8.3.

De esto nos podemos convencer, si examinamos el valor absoluto de la diferencia $x_n - 1$:

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

Los cálculos, efectuados más arriba, muestran que el valor absoluto de la diferencia $x_n - 1$ con el crecimiento de n se hace cada vez menor aproximándose indefinidamente a cero. Por ejemplo, para todos los términos de la sucesión, comenzando por el undécimo, el valor absoluto de la diferencia $x_n - 1$ es menor que 0,1; para todos los términos de la sucesión, comenzando por 101, el valor absoluto es menor que 0,01, etc. En general, por pequeño que sea el número positivo ε escogido, en la sucesión dada se hallará tal término (cuyo número lo denotaremos por N_0) que para todos los términos de la sucesión con números, mayores que N_0, se cumplirá la desigualdad

$$|x_n - 1| < \varepsilon.$$

De tal manera, para cualquier término positivo ε existe tal número N_0, comenzando por el cual todos los términos de la sucesión dada pertenecen al intervalo $(1 - \varepsilon; 1 + \varepsilon)$.

El número 1 se llama *límite de sucesión* con el término general

$$x_n = 1 + \frac{(-1)^{n-1}}{n}$$

y se escribe

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

Sea (x_n) cierta sucesión numérica infinita. El número a se llama *límite de una sucesión numérica* (x_n), si para cualquier número positivo ε, por pequeño que sea, existe tal número N_0, que todos los términos de la sucesión (x_n) con los números mayores que N_0, satisfacen la desigualdad

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

El hecho de que el número a es el límite de la sucesión (x_n) se escribe de la manera siguiente *):

$$\lim_{n\to\infty} n_n = a \text{ o } \lim x_n = a.$$

Si el número a es el límite de la sucesión (x_n), entonces sobre la sucesión (x_n) se dice que élla *converge* hacia el número a o tiene por *límite* el número a; las sucesiones que no tienen límite se llaman *divergentes*.

La desigualdad (1) también puede ser escrita en la forma

$$a - \varepsilon \leqslant x_n < a + \varepsilon.$$

El intervalo $(a - \varepsilon; a + \varepsilon)$ se llama *ε-entorno* del punto a. Usando el concepto de ε-entorno del punto a, se puede dar también la siguiente definición del límite de una sucesión:

El número a se llama *límite* de una sucesión numérica infinita (x_n), si para cualquier número positivo ε existe el número N_0 tal, que todos los términos de la sucesión, comenzando por x_{N_0+1}, pertenecen a ε-entorno del número a.

Formulando el concepto de límite de sucesión, es necesario tener en cuenta que el número N_0, en general, no puede ser indicado de una vez para siempre; éste depende de la elección del número ε. Para subrayar esto, frecuentemente en vez de N_0 se escribe $N_0(\varepsilon)$. Cuando disminuye el número positivo ε, el número respectivo $N_0(\varepsilon)$, en general, aumenta. Es excepción del caso, cuando todos los términos de la sucesión (x_n) son iguales:

$$x_1 = x_2 = x_3 = \ldots = x_n = x_{n+1} = \ldots$$

El límite de tal sucesión

$$a = \lim_{n\to\infty} x_n$$

*) lím es la abreviación de la palabra latina «limes» que significa «límite».

es igual a los términos de esta sucesión. En este caso la desigualdad (1) se cumple para cualquier $\varepsilon > 0$ a la vez para todos los valores $n \in \mathbf{N}$.

Ejemplo. Demostrar que el límite de la sucesión numérica que se representa por la fórmula del término general

$$x_n = \frac{n}{n+1}, \quad n \in \mathbf{N},$$

es igual a 1.

Examinemos el valor absoluto de la diferencia $x_n - 1$:

$$|x_n - 1| = \left|\frac{n}{n+1} - 1\right| = \left|-\frac{1}{n+1}\right| = \frac{1}{n+1}.$$

Mostremos que para cualquier $\varepsilon > 0$ existe tal número N^0 que $|x_n - 1| < \varepsilon$ para todos $n > N_0$. Puesto que

$$|x_n - 1| = \frac{1}{n+1}, \tag{2}$$

entonces la desigualdad $|x_n - 1| < \varepsilon$ se puede escribir en la forma

$$\frac{1}{n+1} < \varepsilon, \tag{3}$$

de donde

$$1 < \varepsilon(n+1) \Leftrightarrow 1 < \varepsilon n + \varepsilon \Leftrightarrow 1 - \varepsilon < \varepsilon n \Leftrightarrow n > \frac{1-\varepsilon}{\varepsilon}. \tag{4}$$

De la última desigualdad se deduce que para cualquier número positivo ε se halla tal número $N_0(\varepsilon)$, que todos los términos de la sucesión con los números $n > N_0(\varepsilon)$ satisfacen a la desigualdad *)

$$|x_n - 1| < \varepsilon.$$

---

*) En calidad de ejemplo $N_0(\varepsilon)$ $\left(\text{para } \frac{1-\varepsilon}{\varepsilon} > 0\right)$ se puede, por ejemplo, tomar el número $\left[\frac{1-\varepsilon}{\varepsilon}\right] + 1$. Si $\frac{1-\varepsilon}{\varepsilon} < 0$, entonces en calidad de N_0 se puede tomar cualquier número natural. (Con el símbolo $\left[\frac{1-\varepsilon}{\varepsilon}\right]$ se denota la parte entera del número $\frac{1-\varepsilon}{\varepsilon}$. Se llama *parte entera* del número real a al número entero p tal, que $p \leqslant a < p + 1$).

Efectivamente, de la desigualdad (4) se ve que cualquiera que sea $\varepsilon > 0$, la desigualdad (4) se satisface por todos los números naturales n, comenzando por cierto $N_0(\varepsilon)$. Pero la desigualdad (4) es equivalente a la desigualdad (3), la cual considerando la igualdad (2) puede ser secrita en la forma

$$|x_n - 1| < \varepsilon. \tag{5}$$

De esta manera, para cualquier $\varepsilon > 0$ todos los términos de la sucesión con los números que satisfacen la desigualdad (4) satisfarán la desigualdad (5), es decir, el número 1 satisface la definición formulada de límite y es el límite de la sucesión dada.

2.2. Condición necesaria para la convergencia de una sucesión. Sea dada cierta sucesión (x_n), convergente hacia el número a.

$$\lim_{n\to\infty} x_n = a.$$

Entonces, conforme a la definición de límite, para cualquier $\varepsilon > 0$ todos los términos de la sucesión dada, menos, puede ser, su número finito, entran en ε-entorno del punto a, es decir,

$$a - \varepsilon < x_n < a + \varepsilon.$$

Sea el número de términos de esta sucesión, que no entraron en el entorno del punto a, igual a N_0 y sea que algunos de estos términos se encuentran a la derecha del punto $a + \varepsilon$, y otros se encuentran a la izquierda del punto $a - \varepsilon$. Puesto que los términos que se encuentran a la derecha del punto $a + \varepsilon$ son un número finito, entonces, al examinarlos, podemos elegir un término máximo, el cual lo denotamos con M. Análogamente, designamos el mínimo de los términos de la sucesión, que se encuentran a la izquierda del punto $a - \varepsilon$, con m. Puede suceder que a la derecha del punto $a + \varepsilon$ (o a la izquierda del punto $a - \varepsilon$) no halla ningún término. En este caso en calidad de número M se puede tomar el número $a + \varepsilon$ y, respectivamente, como número m, el número $a - \varepsilon$.

Así pues, si x_n es una sucesión convergente, entonces existen dos números tales, m y M, que para todos n

$$m \leqslant x_n \leqslant M$$

y, por consiguiente, (x_n) es la sucesión acotada.

Examinemos una sucesión representada por la fórmula del término general

$$x_n = (-1)^n, \quad n \in \mathbf{N}.$$

La sucesión dada será acotada (inferiormente, por el número —1, y superiormente, por el número 1). Aplicando la definición de límite, demostremos que esta sucesión no tiene límite.

Cada término de la sucesión que tiene un número impar se representa en el eje numérico por el punto con una abscisa igual a —1, y los términos con números pares, por el punto con la abscisa 1. De aquí se deduce que si la sucesión dada tiene un límite, entonces este límite es igual o a —1, ó a 1. Supongamos que el límite es igual a —1. Entonces para cualquier $\varepsilon > 0$ todos los términos de la sucesión, comenzando por cierto término deben entrar en el ε-entorno del punto —1. Tomemos $\varepsilon = 0{,}1$. En el intervalo $(-1{,}1;\ -0{,}9)$ entran todos los términos de la sucesión, que se encuentran en los sitios impares, pero no entra ninguno de los términos que se encuentran en los sitios pares, es decir, no existe tal número $N_0(\varepsilon)$, comenzando por el cual todos los términos de la sucesión entran en el ε-entorno dado del punto —1, y, por consiguiente, el número —1 no es límite de la sucesión dada. Análogamente se demuestra que también el número 1 no es límite de la sucesión. De esta manera, resulta que la sucesión

$$-1, 1, \ -1, \ 1, \ -1, \ 1, \ \ldots$$

no tiene límite.

De la actoación de la sucesión *no se deduce* la convergencia de una sucesión.

Los ejemplos examinados permiten formular la *condición necesaria de convergencia* de una sucesión numérica infinita arbitraria:

Para que la sucesión converja es necesario que sea acotada.

2.3. Teoremas sobre los límites de las sucesiones.

1) La sucesión no puede converger hacia dos límites distintos (*teorema de la unicidad del límite*).

2) El límite de una sucesión, todos los términos de la cual son iguales al mismo valor, es igual a este valor.

3) Si la sucesión (x_n) tiende al número a, y $a > p$ ($a < q$), entonces también todos los términos de la sucesión, comenzando por cierto término, serán mayor que p (menor que q).

4) Si la sucesión (x_n) tiene límite, entonces es acotada.

5) Si las sucesiones (x_n) e (y_n) tienen límites (convergen), entonces su suma (diferencia) también tiene límite, además, el límite de la suma (diferencia) es igual a la suma (diferencia) de los límites:

$$\lim_{n\to\infty} (x_n \pm y_n) = \lim_{n\to\infty} x_n \pm \lim_{n\to\infty} y_n.$$

6) Si las sucesiones (x_n) e (y_n) tienen límites (convergen), entonces su producto también tiene límite y el límite del producto es igual al producto de los límites:

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

El multiplicador constante puede sacarse del signo del límite:

$$\lim_{n\to\infty} (cx_n) = c \lim_{n\to\infty} x_n.$$

7) Si las sucesiones (x_n) e (y_n) tienen límites y además $y_n \neq 0$, $n \in \mathrm{N}$, y el límite de la sucesión (y_n) es diferente de cero, entonces su razón también tiene un límite y el límite de la razón es igual a la razón de los límites:

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{\lim\limits_{n\to\infty} x_n}{\lim\limits_{n\to\infty} y_n}.$$

Mostremos ahora en un ejemplo, cómo mediante los teoremas formulados el problema de hallar la sucesión puede reducirse a hallar los límites de las sucesiones simples.

Ejemplo. Calcular el límite $\lim\limits_{n\to\infty} x_n$ de la sucesión (x_n) que se representa por la fórmula del término general

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

Para calcular el límite de una sucesión dada aplicamos los teoremas sobre los límites de las sucesiones. Observemos, que no se puede utilizar de inmediato el teorema sobre el límite de un cociente para calcular el límite

$$\lim_{n\to\infty} \frac{2n^2}{n^2+n+1},$$

puesto que los límites de las sucesiones que se encuentran en el numerador y en el denominador de la fracción no existen (las sucesiones $y_n = 2n^2$ y $z_n = n^2 + n + 1$, $n \in \mathbf{N}$, son no acotadas y, por consiguiente, son divergentes). Efectuemos previamente la siguiente transformación de la fracción que se encuentra bajo el signo del límite, es decir, dividamos el numerador y el denominador de la fracción por n^2:

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

Luego, aplicamos sucesivamente los teorema 7) y 5):

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

Utilizando la definición de límite, es fácil demostrar que

$$\lim_{n\to\infty} \frac{1}{n} = 0 \text{ y } \lim_{n\to\infty} \frac{1}{n^2} = 0;$$

de resultas obtenemos

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

Es necesario subrayar, que cuando hallamos el límite de una sucesión dada, no hemos aplicado con toda precisión los teoremas 7) y 5). Por ejemplo, aplicando el teorema sobre el límite de un cociente, hemos puesto el signo de igualdad entre las expresiones

$$\lim_{n\to\infty} \frac{2}{1+\frac{1}{n}+\frac{1}{n^2}} \text{ y } \frac{\lim\limits_{n\to\infty} 2}{\lim\limits_{n\to\infty} \left(1+\frac{1}{n}+\frac{1}{n^2}\right)},$$

aunque para este momento no sabíamos todavía que los límites del númerador y del denominador de la segunda fracción existen. La convergencia de la sucesión que se encuentra en el denominador de la segunda fracción se demuestra más tarde, después de aplicar «ilegalmente», de la misma manera, el teorema 5). El que nosotros de inmediato pongamos el signo de igualdad, significa realmente que primeramente suponemos que las condiciones de los teoremas que aplicamos están cumplidas, y después, regresando atrás por la cadena de igualdades, argumentamos la legalidad del signo de igualdad entre estas expresiones.

2.4. Condición suficiente para la convergencia de una sucesión (teorema de Weierstrass). La acotación de una sucesión es la condición necesaria de su convergencia, pues de la convergencia de una sucesión se deduce su acotación; sin embargo, la acotación de una sucesión no es condición suficiente de su convergencia. La monotonía de sucesión parece no estar relacionada con su convergencia, ya que la sucesión monótona puede ser tanto convergente, como divergente. Por ejemplo, la sucesión representada por la fórmula del término general $x_n = n$ es monótona y divergente, y la sucesión $x_n = 1/n$ es monótona y convergente.

Por otra parte, las sucesiones no monótonas pueden ser tanto convergentes, como divergentes. Así, la sucesión $x_n = (-1)^n/n$ será convergente, y la sucesión $x_n = (-1)^n n$, divergente.

Sin embargo, las sucesiones que al mismo tiempo son monótonas y acotadas, son, obligatoriamente, convergentes. Este resultado que a veces se llama *condición de convergencia* de una sucesión se formula en forma del siguiente teorema, el cual fue demostrado por primera vez por K. Weierstrass:

Si la sucesión es monótona y acotada, entonces converge.

El teorema de Weierstrass es un teorema sobre la existencia de un límite de sucesión y no da le método para hallar el límite de sucesión. Sin embrago, en una serie de casos, sabiendo que el límite de sucesión existe, se pueden indicar los procedimientos para calcularlo. Para esto muy a menudo se aplica, el que para cualquiera sucesión convergente (que no es obligatoriamente monótona) (x_n) es válida la igualdad

$$\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n. \tag{6}$$

Antes de pasar a los ejemplos, observemos que para el cálculo de los límites mediante la igualdad (6) es cómodo tener una forma recurrente y no analítica de representación de la sucesión, y por eso en todos los ejemplos que vamos a examinár más abajo las sucesiones que se representan analíticamente se escribirán, según la necesidad, en la forma recurrente*).

*) No hace falta pensar que cualquier sucesión, representada por el método analítico, puede ser escrita en la forma recurrente.

Ejemplos. Calcular el límite de sucesión

$$x_n = q^n, \quad 0 < q < 1.$$

La sucesión $x_n = q^n$ es monótona decreciente, puesto que para cualquier $n \in \mathrm{N}$ cuando $0 < q < 1$ se cumple la desigualdad $x_n > x_{n+1}$, y es acotada, puesto que $0 < q^n < 1$. Según el teorema de Weierstrass la sucesión dada tiene límite. Designemos este límite por a:

$$\lim_{n\to\infty} x_n = a. \tag{7}$$

Es fácil observar que la sucesión dada puede ser escrita en la forma recurrente:

$$x_{n+1} = qx_n. \tag{8}$$

En base a la igualdad (6) y a la fórmula recurrente (8) tenemos

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} \Rightarrow \lim_{n\to\infty} x_n = \lim_{n\to\infty} (qx_n).$$

Puesto que el multiplicador constante se puede sacar del signo del límite, entonces, aplicando la anotación (7), se puede escribir la última igualdad en la forma

$$\lim_{n\to\infty} x_n = q \lim_{n\to\infty} x_n \Rightarrow a = qa \Rightarrow a(1-q) = 0.$$

Pero, puesto que $q \neq 1$, entonces de la última igualdad obtendremos $a = 0$. Así pues, el límite de la sucesión dada es igual a cero.

2. Calcular el límite de la sucesión

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

La sucesión dada es monótona decreciente, puesto que

$$x_{n+1} - x_n = -\frac{5}{(2n+1)(2n-1)} < 0$$

para todos $n \in \mathrm{N}$. Es fácil observar que $x_n > 0$ para todos $n \in \mathrm{N}$. Por otra parte, en vigor del decrecimiento monótono cualquier término de la sucesión, comenzando por el segundo, es menor que el primer término x_1, el cual es igual a cuatro. Por eso, cualquier término de la sucesión dada es mayor que cero y menor que cuatro, y, por consiguiente, la sucesión es acotada.

La sucesión dada satisface las condiciones del teorema de Weierstrass y, por consiguiente, tiene límite, el cual lo designemos mediante a. Vamos a calcular este límite con ayuda de la fórmula (6), pasando del procedimiento analítico de representación de una sucesión, al recurrente. Para esto resolvemos la razón

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

respecto a n:

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

Sustituyendo la expresión obtenida para n en la expresión para el término $(n+1)$

$$x_{n+1}=\frac{3n+4}{2n+1},$$

obtendremos la fórmula recurrente de representación de una sucesión dada:

$$x_{n+1}=\frac{11x_n-9}{4x_n-1}. \tag{9}$$

Sustituyendo en el primer miembro de la igualdad (6) x_{n+1} por su expresión a través de x_n, conforme con la fórmula (9) obtendremos

$$\lim_{n\to\infty}\frac{11x_n-9}{4x_n-1}=\lim_{n\to\infty}x_n. \tag{10}$$

Aplicando el teorema sobre el límite del cociente y teniendo en cuenta que el límite de la sucesión (x_n) fue designado mediante a, de la igualdad (10) obtendremos la ecuación para hallar el número a:

$$\frac{11a-9}{4a-1}=a.$$

La única solución de la ecuación dada es: $a=3/2$. De esta manera, el límite de la sucesión dada es igual a $3/2$.

Observemos, que el límite de la sucesión $x_n=\dfrac{3n+1}{2n-1}$ se puede calcular también sin aplicar el teorema de Weierstrass, utilizando los teoremas sobre los límites de las sucesiones.

3. Calcular el límite de la sucesión, que se representa por la fórmula del término general

$$x_n=\underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{c}}}}_{n\ \text{raíces}}.$$

Demostremos que la sucesión dada tiene un límite y calculemos este límite mediante la igualdad (6). La sucesión dada puede ser representada por la fórmula recurrente siguiente:

$$x_{n+1}=\sqrt{c+x_n}. \tag{11}$$

Demostremos que la sucesión dada crece monótonamente, es decir, que $x_{n+1}>x_n$ para todos $n\in\mathbf{N}$. Aplicando la fórmula recurrente (11), la última desigualdad se puede escribir en la forma $\sqrt{c+x_n}>$ $>x_n$, lo que es equivalente (teniendo en cuenta la condición $x_n>0$)

a la desigualdad $x_n^2 - x_n - c < 0$. De esta manera,

$$c+\sqrt{\underbrace{c+\sqrt{c+\ldots+\sqrt{c}}}_{n-1\text{ raíz}}}-\sqrt{\underbrace{c+\sqrt{c+\ldots+\sqrt{c}}}_{n\text{ raíces}}}-c<0 \Leftrightarrow$$

$$\Leftrightarrow \sqrt{\underbrace{c+\sqrt{c+\ldots+\sqrt{c}}}_{n-1\text{ raíz}}}<\sqrt{\underbrace{c+\sqrt{c+\ldots+\sqrt{c}}}_{n\text{ raíces}}}.$$

Elevando ambos miembros de la última desigualdad $n - 1$ veces al cuadrado, obtendremos la desigualdad

$$0 < \sqrt{c}.$$

Por consiguiente, la desigualdad inicial es válida para todos los valores $c > 0$, y la sucesión dada es creciente.

La sucesión dada es acotada superiormente, por ejemplo, por el número $\sqrt{c} + 1$. Efectivamente, $x_1 = c$ es menor que este número. Si ahora admitimos que $x_n < \sqrt{c} + 1$, entonces también para el término siguiente de la sucesión x_{n+1} tendremos

$$x_{n+1} < \sqrt{c+\sqrt{c}+1} < \sqrt{c+2\sqrt{c}+1} = \sqrt{c}+1,$$

y la acotación de la sucesión superiormente puede ser demostrada mediante el método de inducción matemática. La acotación de la sucesión inferiormente se deduce de que $x_n > 0$ para todos $n \in \mathbf{N}$.

Así pues, la sucesión dada es creciente y es acotada; por consiguiente, según el teorema de Weierstrass ella tiene un límite, el cual se designa mediante a. Para calcular el número a aplicamos la fórmula recurrente (11), escribiéndola en forma

$$x_{n+1}^2 = c + x_n.$$

En vigor de la convergencia de nuestra sucesión

$$\lim_{n\to\infty} x_{n+1}^2 = \lim_{n\to\infty} (c + x_n),$$

de donde sigue que el valor a satisface la ecuación cuadrática

$$a^2 = c + a.$$

Un conjunto de soluciones de la ecuación cuadrática dada es:

$$a = \frac{1+\sqrt{4c+1}}{2}, \quad a = \frac{1-\sqrt{4c+1}}{2}.$$

El número $\frac{1-\sqrt{4c+1}}{2}$ para cualquier $c > 0$ es negativo. Pero puesto que un número negativo no puede ser límite de una sucesión con términos positivos, entonces el número

$$\frac{1+\sqrt{4c+1}}{2}$$

será el límite de la sucesión que examinamos.

4. *El número neperiano e como límite de una sucesión numérica infinita**). Examinemos una sucesión representada por la fórmula del término general

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

Demostremos la monotonía de esta sucesión. Desarrollando la expresión $\left(1+\frac{1}{n}\right)^n$ según la fórmula del binomio de Newton, obtendremos

$$x_n=\left(1+\frac{1}{n}\right)^n=1+n\cdot\frac{1}{n}+\frac{n(n-1)}{2!}\cdot\frac{1}{n^2}+$$

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

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

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

Si ahora escribimos el desarrollo análogo para x_{n+1}, entonces al desarrollo obtenido se añade, ante todo, el término positivo $(n+2)$; cada uno de los términos $n+1$ escritos aumenta, puesto que cualquier multiplicador entre paréntesis de la forma $1-\frac{s}{n}$ se sustituye por un multiplicador mayor $1-\frac{s}{n+1}$. De aquí se desprende que

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

es decir, la sucesión (x_n) dada resulta ser creciente.

Antes de pasar a demostrar la acotación de la sucesión dada, demostremos una desigualdad, es decir, mostremos que para cualquier $n\in\mathbf{N}$

$$2^{n-1}\leqslant n!. \tag{12}$$

De la validez de esta desigualdad es fácil convencerse, si demostramos que la sucesión

$$a_n=\frac{2^{n-1}}{n!}$$

*) Sobre el número e véase también el p. 1.5 del capítulo 9.

es no creciente, y su primer término (es muy fácil de verificar) es igual a 1. Efectivamente, la sucesión $\frac{2^{n-1}}{n!}$ es no creciente, puesto que para cualquier $n \in \mathbf{N}$

$$a_{n+1}-a_n=\frac{2^n}{(n+1)!}-\frac{2^{n-1}}{n!}=\frac{2^n-2^{n-1}(n+1)}{(n+1)!}=\frac{2^{n-1}(1-n)}{(n+1)!}\leqslant 0.$$

Demostremos ahora que la sucesión (x_n) es acotada. Omitiendo en la expresión

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

todos los multiplicadores que se encuentran entre paréntesis, aumentamos el segundo miembro de la igualdad (puesto que cada uno de los multiplicadores omitidos es menor que la unidad); tenemos

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

En vigor de la desigualdad (12)

$$2\leqslant 2!;\quad 2^2\leqslant 3!;\quad 2^3\leqslant 4!;\ \ldots;\quad 2^{n-1}\leqslant n!$$

y, por consiguiente,

$$\frac{1}{2}\geqslant\frac{1}{2!};\ \frac{1}{2^2}\geqslant\frac{1}{3!};\ \frac{1}{2^3}\geqslant\frac{1}{4!};\ \ldots;\ \frac{1}{2^{n-1}}\geqslant\frac{1}{n!}.$$

Considerando las últimas desigualdades, tenemos

$$x_n<2+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\ldots+\frac{1}{n!}\leqslant 2+$$
$$+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\ldots+\frac{1}{2^{n-1}}.$$

Es fácil observar, que en el segundo miembro de la última desigualdad la suma de todos los términos, comenzando por el segundo, es de $n-1$ terminos de una progresión geométrica infinitamente decreciente. La suma de todos los términos de la progresión infinitamente decreciente

$$\frac{1}{2^2};\ \frac{1}{2^3};\ \ldots;\ \frac{1}{2^{n-1}};\ \ldots$$

es igual a 1, y por eso para cualquier $n \in \mathbf{N}$

$$x_n<3$$

De esta manera, fue demostrado que nuestra sucesión es monótona creciente y acotada y, por consiguiente, en vigor del teorema de Weierstrass tiene un límite. Este límite tiene una *designación especial*, e:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e = 2{,}71\,828\ldots$$

2.5. Condición necesaria y suficiente para la convergencia de una sucesión numérica infinita. Hemos examinado más arriba dos procedimientos para demostrar la convergencia de las sucesiones. Uno de los procedimientos está basado en la aplicación de los teoremas sobre los límites de las sucesiones (véase p. 2.3), y el segundo, en la aplicación del teorema de Weierstrass (véase p. 2.4). Sin embargo, ambos procedimientos de demostración de la convergencia de la sucesión son aplicables para examinar un número pequeño de sucesiones bastante simples.

Surge una pregunta: ¿cómo demostrar la convergencia, por ejemplo, de una sucesión no monótona, la cual se da no por el procedimiento analítico, sino por cualquier otro procedimiento (por ejemplo, por el recurrente), o de una sucesión que se da por una fórmula analítica, cuyo límite es imposible calcular, limitándose a los teoremas sobre los límites de las sucesiones? Está claro que en este caso ya no podemos usar la definición del límite de una sucesión para demostrar su convergencia, puesto que en esta definición ya participa el mismo número a, cuya existencia tenemos que demostrar. Además, en una serie de casos nos puede interesar solamente la cuestión sobre la convergencia de la sucesión, pero no la cuestión sobre hacia qué número converge.

De suerte que, sería muy deseable tener cierto teorema que pudiera contestar a la pregunta sobre la existencia de un límite de sucesión y en la formulación del cual participara sólo lo que se da, es decir, los términos de nuestra sucesión. Tal teorema fue formulado y demostrado por Bolzano y Cauchy:

Teorema de Bolzano—Cauchy (criterio de convergencia):

Para que la sucesión (x_n) tenga un límite finito, es necesario y suficiente que para cada número positivo ε exista tal número $N_0(\varepsilon)$, que la desigualdad

$$|x_n - x_{n+p}| < \varepsilon$$

se cumpla para todos $n > N_0$ y $p \geqslant 1$ (es decir, para cualesquiera números naturales n y p que satisfacen estas desigualdades).

§ 3. Series numéricas

3:1. Concepto de serie numérica. Sea dada una sucesión numérica infinita

$$x_1;\ x_2;\ x_3;\ \ldots;\ x_n\ \ldots \qquad (1)$$

Planteemos un problema: calcular la suma de todos los términos de una sucesión infinita (1). Para solucionar tal problema surge, naturalmente, una pregunta: ¿cómo calcular la suma de un número infinito de sumandos? No se puede calcular la suma de un número infinito de sumandos de la misma manera que se calcula la suma de un número finito puesto que nunca podremos finalizar el proceso de adición. Sin embargo resulta que, aplicando el concepto de límite de una sucesión, se puede introducir el concepto de suma de un número infinito de sumandos. Para esto tomemos el primer término de la sucesión (x_n) y designémoslo con S_1:

$$S_1 = x_1.$$

Sumemos el primer término de la sucesión x_1 con el segundo término x_2 y designemos su suma con S_2:

$$S_2 = x_1 + x_2.$$

Sumemos los tres primeros términos de la sucesión (1) y designemos su suma con S_3:

$$S_3 = x_1 + x_2 + x_3.$$

Continuando el proceso iniciado, es decir, sumando sucesivamente cuatro, cinco, seis, etc. primeros términos de la sucesión (x_n), denotemos las sumas de cuatro, cinco, seis, etc. primeros términos mediante S_4, S_5, S_6, . . .:

$$S_4 = x_1 + x_2 + x_3 + x_4,$$
$$S_5 = x_1 + x_2 + x_3 + x_4 + x_5,$$
$$S_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

El conjunto de las sumas S_n, $n \in \mathbf{N}$, forma una sucesión infinita. Si el límite de esta sucesión existe, entonces se llama suma de un número infinito de sumandos

$$x_1, x_2, x_3, \ldots, x_n, \ldots$$

Así, para la sucesión numérica dada (x_n) la expresión

$$x_1 + x_2 + x_3 + \ldots + x_n \ldots \tag{2}$$

se llama *serie numérica*, y x_n — n se llama *término* x_n — n-ésimo de la serie. Para designar la serie numérica (2) se usa la expresión

$$\sum_{n=1}^{\infty} x_n.$$

La suma n de los primeros términos de la serie se llama n-ésima *suma parcial* de esta serie y se denota con S_n:

$$S_n = x_1 + x_2 + x_3 + \ldots + x_n.$$

Si existe un límite S de sucesión (S_n) de las sumas parciales de la serie (2), entonces se dice que esta serie converge, y el límite indicado S

se llama *suma* de la serie (2) y se escribe

$$S = x_1 + x_2 + x_3 + \ldots + x_n + \ldots = \sum_{n=1}^{\infty} x_n.$$

Si no existe un límite de sucesión de las sumas parciales (S_n), entonces la serie (2) se llama *divergente*. La serie infinita

$$x_{k+1} + x_{k+2} + \ldots + x_{k+l} + \ldots,$$

que se obtiene de la serie (3) mediante la eliminación de los primeros k términos, se llama *resto* de la serie

$$x_1 + x_2 + x_3 + \ldots + x_k + x_{k+1} + \ldots + x_{k+l} + \ldots \quad (3)$$

después del termino k.

Un ejemplo simple de una serie numérica infinita es la suma de todos los términos de una progresión geométrica con el primer término, igual a a y con el denominador igual a q

$$a + aq + aq^2 + \ldots + aq^{n-1} + \ldots; \quad (4)$$

la n-ésima suma parcial de esta serie es la *suma de n términos de la progresión geométrica* y se calcula mediante la fórmula

$$S_n = \frac{a(1-q^n)}{1-q} \quad (\text{para } q \neq 1).$$

Si el denominador de la progresión $|q| < 1$, entonces la sucesión de las sumas parciales (S_n) tiene límite

$$S = \frac{a}{1-q},$$

el cual se llama *suma* de todos los términos de la progresión geométrica infinitamente decreciente (es decir, la suma de la serie (4)).

Para $|q| \geqslant 1$ la misma expresión (4) da un ejemplo de una serie divergente.

Propiedades de las series numéricas infinitas.

1) Si la serie numérica converge, entonces converge también cualquiera de sus restos; a la inversa, de la convergencia del resto se desprende la convergencia de la serie inicial.

2) Si la serie converge, entonces el resto de la serie después del k-ésimo término tiende a cero cuando $k \to \infty$.

3) Si los términos de una serie convergente los multiplicamos por un mismo multiplicador c, entonces su convergencia no altera, y la suma se multiplica por c:

$$\sum_{n=1}^{\infty} cx_n = c \sum_{n=1}^{\infty} x_n.$$

4) Dos series convergentes

$$x_1+x_2+x_3+\ldots+x_n+\ldots=\sum_{n=1}^{\infty}x_n=X,$$

$$y_1+y_2+y_3+\ldots+y_n+\ldots=\sum_{n=1}^{\infty}y_n=Y$$

se pueden sumar término a término (o sustraer) y la serie

$$(x_1\pm y_1)+(x_2\pm y_2)+(x_3\pm y_3)+\ldots$$
$$\ldots+(x_n\pm y_n)+\ldots=X\pm Y$$

converge también, y su suma (diferencia) es igual a $X+Y$ (respectivamente $X-Y$).

5) El término general x_n de la serie convergente tiende a cero (*condición necesaria de convergencia* de una serie numérica infinita).

3.2. Series numéricas positivas. La serie numérica

$$x_1+x_2+x_3+\ldots+x_4+\ldots=\sum_{n=1}^{\infty}x_n$$

se llama *positiva*, si todos los términos de la serie son números positivos:

$$x_1>0,\ x_2>0,\ x_3>0,\ \ldots,\ x_n>0,\ \ldots$$

La sucesión de las sumas parciales de una serie positiva es una sucesión creciente. Una serie numérica positiva converge, si la sucesión de las sumas parciales de la serie está acotada superiormente, y diverge en caso contrario (*condición necesaria y suficiente de convergencia* de una serie numérica positiva).

La convergencia (o divergencia) de una serie numérica positiva muy a menudo se establece mediante la comparación de una serie numérica dada con otra, que es de antemano una serie numérica convergente o divergente. En la base de esta comparación se encuentra la siguiente *propiedad de las series numéricas*: sean dadas dos series numéricas positivas

$$\text{(A)}\quad x_1+x_2+x_3+\ldots+x_n+\ldots=\sum_{n=1}^{\infty}x_n,$$

$$\text{(B)}\quad y_1+y_2+y_3+\ldots+y_n+\ldots=\sum_{n=1}^{\infty}y_n.$$

Si, comenzando por cierto número, para todos $n>N_0$ se cumple la desigualdad $x_n\leqslant y_n$, entonces de la convergencia de la serie (B) se deduce la convergencia de la serie (A), y de la divergencia de la serie (A) se deduce la divergencia de la serie (B).

Sea dada la serie

$$x_1 + x_2 + x_3 + \ldots + x_n + \ldots = \sum_{n=1}^{\infty} x_n,$$

cuya convergencia es necesario examinar. Para demostrar la divergencia de la serie dada es necesario construir una serie tan notoriamente divergente

$$y_1 + y_2 + y_3 + \ldots + y_n + \ldots = \sum_{n=1}^{\infty} y_n,$$

que $y_n \leqslant x_n$ para todos $n > N_0$.

Para demostrar la convergencia de la serie dada es necesario construir una serie tan notoriamente convergente

$$z_1 + z_2 + z_3 + \ldots + z_n + \ldots = \sum_{n=1}^{\infty} z_n,$$

que $z \geqslant x_n$ para todos $n > N_0$.

Para un caso general es muy difícil dar cualquier recomendación concreta, acerca de cómo construir las series indicadas $\sum\limits_{n=1}^{\infty} y_n$ y $\sum\limits_{n=1}^{\infty} z_n$. Frecuentemente es cómodo tomar como tales series, a las series formadas de los términos de una progresión geométrica, las cuales en dependencia del valor del denominador pueden ser tanto convergentes, como divergentes.

3.3. Serie armónica. La serie numérica

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n}$$

se llama *serie armónica*. Cada término de tal serie, comenzando por el segundo, representa la media armónica de dos términos vecinos (el número c que satisface la igualdad $\frac{1}{c} = \frac{1}{2}\left(\frac{1}{a} + \frac{1}{b}\right)$ se llama *media armónica* de los números a y b).

Una serie numérica de la forma más general

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \ldots + \frac{1}{n^s} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n^s}, \tag{5}$$

donde s es cualquier número real, se llama también con frecuencia serie armónica.

La serie armónica de la forma (5) diverge cuando $s \leqslant 1$ y converge si $s > 1$.

§ 4. Productos infinitos

4.1. Concepto de producto infinito. Examinemos una sucesión numérica infinita $x_1, x_2, x_3, \ldots, x_n, \ldots$. Formemos de la sucesión numérica dada una sucesión, cuyos términos son los números

$$X_1 = x_1,\ X_2 = x_1 \cdot x_2,\ X_3 = x_1 \cdot x_2 \cdot x_3,\ \ldots,\ X_n = \\ = x_1 \cdot x_2 \cdot x_{3+} \cdot \ldots \cdot x_n \ldots .$$

Si el límite de la sucesión $X_1, X_2, \ldots, X_n, \ldots$ existe y es distinto de cero, entonces se llama valor del producto infinito

$$x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n \cdot \ldots .$$

Así:

1) para la sucesión numérica dada (x_n) la expresión

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

se llama *producto infinito*. El producto numérico infinito (1) se designa

$$\prod_{n=1}^{\infty} x_n$$

2) El producto de n primeros términos de la sucesión (x_n) se llama *producto parcial* del producto infinito dado:

$$X_n = x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n.$$

3) Si existe un límite X, distinto de cero de la sucesión (X_n) de los productos parciales, entonces se dice que el producto infinito *converge*; el límite X se llama *valor* del producto infinito (1) y es escribe

$$X = x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n \cdot \ldots .$$

En caso contrario (cuando el límite de la sucesión de los productos parciales no existe o es igual a cero) el producto infinito se llama *divergente*.

4.2. Relación entre los productos infinitos y las series. Examinemos una sucesión numérica infinita (x_n), todos los términos de la cual son positivos. Formemos para esta sucesión un producto infinito

$$x_1 \cdot x_2 \cdot x_n \cdot \ldots \cdot x_n \cdot \ldots = \prod_{n=1}^{\infty} x_n. \tag{2}$$

Para la convergencia del producto infinito (2) es necesario y suficiente, que converja la serie infinita

$$\sum_{n=1}^{\infty} \ln x_n. \tag{3}$$

Si la serie (3) converge hacia el número S, entonces también el producto infinito (2) converge hacia el número $X = e^S$.

§ 5. Progresiones

5.1. Progresión aritmética. Se llama *progresión aritmética* a la sucesión, de la cual se da el primer término a_1, y cada término siguiente, comenzando por el segundo, es igual al término anterior, sumado con un mismo número d. El número d se llama *diferencia* de la progresión aritmética. Si el número $d > 0$, entonces la progresión se llama *creciente*; si $d < 0$, decreciente.

Para que la sucesión (a_n) sea una progresión aritmética, es necesario y suficiente que para cualquier $n > 1$ se cumpla la igualdad

$$a_n = \frac{a_{n-1} + a_{n+1}}{2}.$$

Si se dan el primer término a_1 y la diferencia de la progresión aritmética d, entonces el n-ésimo término de la progresión aritmética se calcula mediante la fórmula

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

Los términos de la progresión aritmética a_{n-1}, a_n y a_{n+l} $(n > l)$ se relacionan entre sí por la igualdad

$$a_n = \frac{a_{n+l} + a_{n-l}}{2}.$$

La *suma de los primeros n términos* de la progresión aritmética (a_n) se calcula por la fórmula

$$S_n = \frac{a_1 + a_n}{2} \cdot n$$

o mediante la fórmula

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

Una serie, que está formada de los términos de la progresión aritmética

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n,$$

para cualesquiera valores numéricos del primer término a_1 y de la diferencia d será divergente (a excepción del caso $a_1 = d = 0$).

5.2. Progresión geométrica. Se llama *progresión geométrica* a la sucesión, de la cual se da el primer término u_1, y cada término siguiente, comenzando por el segundo, se obtiene multiplicando el término anterior por un mismo número constante q de la sucesión dada. El número q se llama *denominador* de la progresión geométrica.

Para que la sucesión (u_n) sea una progresión geométrica, es necesario y suficiente que para cualquier $n > 1$ se cumpla la igualdad

$$u_n^2 = u_{n-1} \cdot u_{n+1}.$$

Si se dan el primer término u_1 y el denominador de la progresión geométrica q, entonces el término n de la progresión geométrica se calcula mediante la fórmula

$$u_n = u_1 \cdot q^{n-1}.$$

Los términos de la progresión geométrica u_{n-l}, u_n y u_{n+l} $(n > l)$ se relacionan entre sí por la igualdad

$$u_n^2 = u_{n-l} \cdot u_{n+l}.$$

La *suma de los primeros n términos* de la progresión geométrica se calcula por la fórmula

$$S_n = \frac{u_1(1-q^n)}{1-q}.$$

La serie numérica infinita, que está formada de los términos de la progresión geométrica

$$u_1 + u_2 + u_3 + \ldots + u_n + \ldots = \sum_{n=1}^{\infty} u_n, \tag{1}$$

para $|q| < 1$ converge, y su suma S es igual

$$S = \frac{u_1}{1-q}; \tag{2}$$

para $|q| \geqslant 1$ la serie (1) diverge.

La fórmula (2) se llama también fórmula de la *suma de los términos* de una progresión geométrica decreciente infinita. Aquí por «progresión geométrica decreciente infinita» se comprende una progresión, cuyos términos decrecen de valor absoluto, es decir,

$$|u_1| > |u_2| > |u_3| > \ldots > |u_n| > \ldots.$$

§ 6. Funciones numéricas

6.1. Concepto de función numérica. Sea X e Y dos conjuntos de números reales. Los elementos de estos conjuntos se designan con las letras x e y respectivamente. Representar una correspondencia entre el conjunto X y el conjunto Y significa indicar una regla, según la cual para cada número x del conjunto X se elige uno, varios o infinitamente muchos números y del conjunto Y. Puede resultar que a ciertos números $x \in X$ no corresponde ningún número $y \in Y$.

Por ejemplo, si representamos una correspondencia entre los conjuntos de los números reales $X = \mathbf{R}$ e $Y = \mathbf{R}$ en forma de la fórmula

$$x^2 + y^2 = 1,$$

entonces a cada número $x \in (-1; 1)$ le corresponderán dos números

$$y = \sqrt{1-x^2} \quad \text{e} \quad y = -\sqrt{1-x^2}.$$

A los números $x = 1$ y $x = -1$ les corresponde un solo número $y = 0$, y a los demás $x \in \mathbf{R}$ no corresponde ningún y.

Las correspondencias entre el conjunto numérico X y el conjunto numérico Y suelen designarse con las letras del alfabeto latino f, F, g, ... y se escribe

$$X \xrightarrow{f} Y, \quad X \xrightarrow{F} Y, \quad X \xrightarrow{g} Y, \ \ldots$$

La correspondencia f, que asigna a cada número dado x del conjunto X un solo número y del conjunto Y, se llama *función numérica* del argumento x y se escribe

$$y = f(x).$$

Al mismo tiempo x se llama variable independiente, e y, variable dependiente; el conjunto X se llama *campo de definición* de la función $f(x)$, y el conjunto Y se llama *región de cambio* (o *conjunto de valores*) de la función $f(x)$. A veces el campo de definición y la región de cambio de la función $y = f(x)$ se designan $D(f)$ y $E(f)$ respectivamente

La función numérica $y = f(x)$ puede ser definida también como un conjunto de pares ordenados de los números reales $(x; f(x))$ tales, que para cada x en el conjunto de estos pares existe no más que un par con el primer elemento x. Por ejemplo, la función representada por la fórmula $y = x^2$, puede ser representada también como un conjunto de todos los pares ordenados de la forma $(x; x^2)$.

Si consideramos las variables x e y como coordenadas cartesianas de puntos en el plano, entonces el conjunto de los puntos del plano de coordenadas Oxy con las coordenadas $(x; f(x))$ se llama *gráfica de la función* $y = f(x)$.

6.2. Procedimientos para la representación de funciones. Una función puede ser representada por distintos procedimientos. En el análisis matemático con frecuencia la función se representa analíticamente.

Formas principales de representación analítica de una función:

1) *Forma explícita de representación de una función.* La función se representa con una fórmula que indica las operaciones (y la consecutividad de su realización) que es necesario hacer con el valor de una variable independiente y como resultado de las cuales se obtiene el valor de la variable dependiente.

Por ejemplo, sea que la función $y = f(x)$ tiene la forma

$$y = (\sqrt{x} - 1)^2.$$

El campo de definición de la función dada $D(f) = [0; +\infty)$.

Tomemos cierto valor $x_0 \in D(f)$. Para obtener el valor $y_0 = f(x_0)$ con el valor x_0 es necesario efectuar las operaciones siguientes, de resultas de las cuales se obtiene el valor y_0: a) extraer la raíz cuadrada del número x_0; b) sustraer del valor obtenido de la raíz cuadrada el número unidad; c) elevar la diferencia obtenida al cuadrado.

Para el método analítico de representación de una función, la función puede ser representada para unos valores del argumento $x \in D_1$ por una fórmula, y para otros valores del argumento $x \in D_2$, por otra fórmula ($D_1 \cup D_2 = D(f)$ y $D_1 \cap D_2 = \varnothing$).

La función que se define por las condiciones siguientes:

$$f(x) = 1, \quad \text{si } x > 0,$$
$$f(x) = -1, \quad \text{si } x < 0,$$
$$f(0) = 0$$

puede servir como ejemplo de tal función en el intervalo $(-\infty; +\infty)$. Esta función tiene una designación especial:

$$\operatorname{sign} x$$

(se lee; «signum x») y se llama «el signo del número x».

2) *Forma implícita de representación de una función.* Por representación implícita de una función se entiende la representación de la función en forma de una ecuación con dos variables:

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

La ecuación (1) representa una función sólo en el caso, cuando el conjunto de los pares ordenados $(x; y)$ que son la solución de la ecuación dada es tal, que para cualquier número x existe en este conjunto no más que un par $(x_0; y_0)$ con el primer elemento x_0. Por ejemplo, la ecuación

$$xy - 1 = 0$$

representa una función, mientras que la ecuación que tiene la forma

$$x^2 + y^2 = 1$$

representa no una función, sino una correspondencia, puesto que entre el conjunto de los pares ordenados $(x; y)$, que son la solución de la ecuación dada, se hallan, por ejemplo, dos tales pares con el primer elemento coincidente:

$$\left(\frac{\sqrt{3}}{2}; \frac{1}{2}\right) \text{ y } \left(\frac{\sqrt{3}}{2}; -\frac{1}{2}\right).$$

3) *Forma paramétrica de representación de una función.* Para la representación paramétrica de una función los valores correspondientes entre sí de las variables x e y se expresan mediante el tercer valor llamado *parámetro*:

$$x = \varphi(t), \quad y = \psi(t).$$

En ciertos casos la función, representada en forma paramétrica, puede ser escrita también en forma explícita. Por ejemplo, la función, representada en la forma paramétrica:

$$x = \operatorname{sen} t, \quad y = \cos t, \quad t \in [-\pi/2, \pi/2],$$

admite la expresión en la forma explícita:

$$y=\sqrt{1-x^2}.$$

Además del procedimiento analítico de representación, la función puede ser representada también:

1) por la tabla de sus valores o por la regla de cálculo de la tabla;

2) por una gráfica.

6.3. Suma, producto, diferencia y cociente de dos funciones.

Las funciones $f(x)$ y $g(x)$ se llaman *iguales* en la parte común de su campo de definición $F(f) \cap D(g)$, si $f(x) = g(x)$ para todos $x \in D(f) \cap D(g)$.

La función $S(x)$ que se representa por las condiciones:

1) El campo de definición de la función $S(x)$ (si no hay restricción especial) es parte común de los conjuntos $D(f)$ y $D(g)$, es decir,

$$D(S) = D(f) \cap D(g).$$

2) El valor de la función $S(x)$ en cada punto $x_0 \in D(S)$ se calcula como la suma de los valores de las funciones $f(x)$ y $g(x)$ en el punto x_0;

$$S(x_0) = f(x_0) + g(x_0),$$

se llama suma de dos funciones $f(x)$ y $g(x)$, definidas en el conjunto de los valores de la variable independiente $D(f)$ y $D(g)$ respectivamente.

Análogamente se define la diferencia, el producto y el cociente de dos funciones, además el cociente de dos funciones

$$\frac{f(x)}{g(x)}$$

está definido sólo para aquellos valores $x \in D(f) \cap D(g)$, para los cuales la función $g(x)$ no se anula.

6.4. Función compuesta (superposición de funciones). Sea $y = f(x)$ una función numérica con un campo de definición $D(f)$ y una región de cambio (conjunto de valores) $E(f)$, y $z = g(y)$, una función numérica, representada en el conjunto $E(f)$ o en cierto subconjunto suyo, con una región de cambio $E(g)$. La correspondencia que asigna a cada número dado x del conjunto $D(f)$ un solo número y del conjunto $E(f)$, y al número y un número solo z del conjunto $E(g)$, se llama *función compuesta* (o superposición de las funciones $y = f(x)$ y $z = g(y)$) y se escribe:

$$z = g(f(x)).$$

La mayoría de las funciones que se estudian en las matemáticas elementales y en el análisis matemático pueden ser consideradas como funciones compuestas.
Por ejemplo, la función

$$z = \sqrt{x} - 1$$

puede ser representada como una superposición de dos funciones:
$y = f(x)$ se llama *par*, si

$$y = \sqrt{x}, \quad z = y - 1.$$

6.5. Funciones pares e impares. La función numérica $y = f(x)$ se llama par, si

1) el campo de definición de la función es simétrico respecto al punto 0 del eje numérico (es decir, si el punto x_0 pertenece al campo de definición de la función);

2) para cualquier valor de una variable independiente, perteneciente al campo de definición de la función, se cumple la igualdad

$$f(x) = f(-x).$$

La función numérica $y = f(x)$ se llama *impar*, si

1) el campo de definición de una función es simétrico respecto al punto O del eje numérico (es decir, si el punto x_0 pertenece al campo de definición de la función, entonces el punto $-x_0$ pertenece también al campo de definición de la función);

2) para cualquier valor de una variable independiente, perteneciente al campo de definición de una función, se cumple la igualdad

$$f(x) = -f(-x).$$

La gráfica de una función par es simétrica respecto al eje de coordenadas, y la gráfica de una función impar es simétrica respecto al origen de coordenadas, es decir, es simétrica central.

Propiedades de funciones pares e impares:

1) la suma, diferencia, producto y cociente de dos funciones pares es una función par,

2) la suma y diferencia de las funciones impares es una función impar, y el producto y cociente es una función par.

La demostración de la paridad (o imparidad) de una función $y = f(x)$ se efectúa de la manera siguiente. Se aclara la simetría del campo de definición de la función $y = f(x)$ respecto al punto O del eje numérico. Si el campo de definición de la función no es simétrico respecto al punto O, entonces la función no es ni par, ni impar. Si el campo de definición de la función es simétrico respecto al punto O, entonces pasan a la verificación de la validez de las igualdades

$$\begin{aligned} f(x) &= f(-x), \\ f(x) &= -f(-x). \end{aligned} \tag{2}$$

Si se cumple la primera de las igualdades (2), entonces la función $f(x)$ es par; si se cumple la segunda, la función es impar. Si no se cumple ninguna de las igualdades expuestas aquí, entonces la función no es ni par ni impar.

Ejemplos. 1. La función $f(x) = \frac{1}{x+1}$ no es par y no es impar, puesto que su campo de definición no es simétrico respecto al punto O (en el punto $x = 1$ la función está determinada, y en el punto $x = -1$ no está determinada).

2. La función $f(x) = \frac{x^2+x}{x}$ tiene un campo de definición simétrica respecto al punto O, pero no es ni par ni impar, de lo que es fácil convencerse basándose en los cálculos siguientes:

$$f(-x) = \frac{(-x)^2+(-x)}{-x} = \frac{x^2-x}{-x} = -\frac{x^2-x}{x},$$

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

No es difícil observar que

$$f(-x) \neq f(x) \text{ y } -f(-x) \neq f(x).$$

3. La función $f(x) = \frac{1}{x^2-1}$ tiene un campo simétrico de definición respecto al punto O del eje numérico y satisface la primera de las igualdades (2):

$$f(-x) = \frac{1}{(-x)^2-1} = \frac{1}{x^2-1} = f(x).$$

Por consiguiente, la función $f(x) = \frac{1}{x^2-1}$ es una función par.

6.6. Funciones periódicas. La función $f(x)$ se llama *periódica* con el período T (T es cierto número real, distinto de cero), si

1) para cualquier valor del argumento x del campo de definición de la función los valores $x + T$ y $x - T$ pertenecen también al campo de definición de la función;

2) para cualquier x del campo de definición de la función

$$f(x) = f(x+T) \text{ o } f(x) = f(x-T).$$

Por *período* de una función se suele entender el mínimo de todos los períodos positivos (si tal período existe). En este caso todos los períodos de las funciones son múltiples de su mínimo período:

$$T = kT_0,$$

donde T_0 es el mínimo período positivo de una función, y k es un número entero, distinto de cero.

Existen funciones periódicas que no tienen un período mínimo positivo. La función de Dirichlet, definida por las condiciones

$$D(x) = \begin{cases} 0, \text{ si } x \text{ es racional,} \\ 1, \text{ si } x \text{ es irracional,} \end{cases}$$

es un ejemplo de tal función.

La demostración de la periodicidad (o no periodicidad) de una función consiste en hallar su período (o en demostrar que no existe el número T, que satisface las condiciones 1), 2)).

Ejemplos. 1. Hallar el período mínimo de la función $f(x) = \operatorname{sen}(ax + b)$, donde $a \neq 0$.

Resolviendo la ecuación sen $(ax + b)$ = sen $[a\ (x + T) + b]$ respecto al valor T, obtendremos

$$T = \frac{\pi + 2\pi k}{a} - 2x - 2 \cdot \frac{b}{a}, \quad k \in \mathbf{Z}; \tag{3}$$

$$T = \frac{2\pi n}{a}, \quad n \in \mathbf{Z}. \tag{4}$$

El valor T que se define por la fórmula (3) no satisface la definición de período, puesto que T es una función de x. De la fórmula (4) hallamos que el mínimo período positivo de la función sen $(ax + b)$ es el número $2\pi / |a|$.

2. Demostrar la no periodicidad de la función $f(x) =$ sen (x^2).

Supongamos que existe un número $T \neq 0$ tal, que sen (x^2) = sen $(x + T)^2$ para cualquier x. Sin embargo, resolviendo esta ecuación respecto al valor T, obtenemos que T depende de x, y, por consiguiente, la función es no periódica.

6.7. Funciones acotadas. La función $f(x)$ se llama *acotada superiormente*, si existe tal número M, que para todos los valores del argumento del campo de definición de una función se cumple la desigualdad $f(x) \leqslant M$, y acotada inferiormente, si existe tal número m, que para todos los valores del argumento del campo de definición de una función se cumple la desigualdad $f(x) \geqslant m$.

Una función acotada superior e inferiormente se llama simplemente *acotada*.

Ejemplos. 1. La función x^2 es, por ejemplo, acotada inferiormente por el número 0 y no acotada superiormente.

2. La función $-\sqrt{x}$ es, por ejemplo, acotada superiormente por el número 1 y no acotada inferiormente.

3. La función sen x es, por ejemplo, acotada superiormente por el número 1, e inferiormente por el número -1.

Ejemplos de funciones no acotadas:

$$f(x) = x^3, \quad f(x) = \operatorname{tg} x, \quad f(x) = \frac{1}{x}.$$

6.8. Funciones monótonas. Sea la función $f(x)$ definida en cierto intervalo $(a;\ b)$.

La función $f(x)$ se llama *creciente* en el intervalo $(a;\ b)$, si para cualquier par de valores $x_1 \in (a;\ b)$, $x_2 \in (a;\ b)$ de la desigualdad $x_1 > x_2$ se deduce la desigualdad $f(x_1) > f(x_2)$, es decir, a un valor mayor del argumento le corresponde un valor mayor de la función. Si de la desigualdad $x_1 > x_2$ se deduce la desigualdad $f(x_1) \geqslant f(x_2)$, entonces la función se llama *no decreciente*.

La función $f(x)$ se llama *decreciente* en el intervalo $(a; b)$, si para cualquier par de valores $x_1 \in (a; b)$, $x_2 \in (a; b)$ de la desigualdad $x_1 > x_2$ se deduce la desigualdad $f(x_1) < f(x_2)$, es decir, a un valor mayor del argumento le corresponde un valor menor de la función. Si de la desigualdad $x_1 > x_2$ se deduce la desigualdad $f(x_1) \leqslant f(x_2)$, entonces la función se llama *no creciente*. Las funciones crecientes y decrecientes se llaman funciones *monótonas*.

Puede suceder que la función $f(x)$ no es monótona en todo el campo de su definición, sin embargo para una partición conveniente del campo de definición de la función en cierto (posiblemente, infinito) número de intervalos, que no se intersecan, resulta que la función es monótona en cada uno de estos intervalos (en unos es decreciente, en otros es creciente). Tales intervalos se llaman intervalos de monotonía de una función.

Ejemplos. 1. La función $y = x^2$, definida en todo el eje numérico, en el intervalo $(-\infty; 0)$ decrece, y en el intervalo $[0; +\infty)$ crece.

2. Demostrar la monotonía de la función $f(x) = \sqrt{x}$. El campo de definición de la función es el intervalo $[0; +\infty)$.

Examinemos la diferencia

$$f(x_1) - f(x_2) = \sqrt{x_1} - \sqrt{x_2},$$

donde x_1 y x_2 son los valores del argumento del campo de definición de la función y $x_1 > x_2$. La diferencia $\sqrt{x_1} - \sqrt{x_2}$ se puede representar en la forma

$$\sqrt{x_1} - \sqrt{x_2} = \frac{(\sqrt{x_1} - \sqrt{x_2})(\sqrt{x_1} + \sqrt{x_2})}{\sqrt{x_1} + \sqrt{x_2}} = \frac{x_1 - x_2}{\sqrt{x_1} + \sqrt{x_2}}.$$

Puesto que $x_1 > x_2$ y $\sqrt{x_1} + \sqrt{x_2} > 0$, entonces $\sqrt{x_1} - \sqrt{x_2} > 0$ y, por consiguiente, la función $f(x) = \sqrt{x}$ es creciente.

Condición suficiente de monotonía de una función. Sea la función $f(x)$ definita y diferenciable en el intervalo abierto

$(a; b)$. Para el crecimiento de la función $f(x)$ en el intervalo $(a; b)$ es suficiente que la derivada de la función dada sea positiva para todos $x \in (a; b)$, y para el decrecimiento de la función es suficiente que la derivada sea negativa para todos $x \in (a; b)$.

La condición de monotonía formulada no es necesaria, pues se puede citar aquí ejemplos de funciones monótonas, cuya derivada se anula en un número finito (y hasta infinito) de los puntos del intervalo $(a; b)$. Un ejemplo simple de tal función es la función

$$f(x) = x^3.$$

Ella crece para todos $x \in \mathbf{R}$ y tiene una derivada, igual a cero en el punto O.

6.9. Funciones recíprocamente inversas. Sea $y = f(x)$ una función numérica con un campo de definición $D(f)$ y una región de cambio (conjunto de valores) $E(f)$. Más adelante, para hacerlo más simple, supongamos que los conjuntos $D(f)$ y $E(f)$ representan ciertos intervalos.

Elijamos un valor y_0 del conjunto de valores de la función $E(f)$; entonces en el conjunto $D(f)$ obligatoriamente se halla por lo menos un valor $x = x_0$ tal, que

$$f(x_0) = y_0.$$

Examinemos las funciones, las cuales a cada valor y_0 de la región de cambio de la función le corresponde un solo valor x_0 del campo de la definición de la función. La correspondencia que asigna a cada número dado y_0 del conjunto $E(f)$ un solo número x_0 del conjunto $D(f)$ se llama función *inversa* a la función f, y se designan con el símbolo f^{-1}:

$$x = f^{-1}(y).$$

No toda función tiene función inversa. Las funciones, para las cuales existen funciones inversas se llaman inversibles; las funciones f y f^{-1} se llaman *recíprocamenre inversas*.

Para que la función $y = f(x)$, definida en el intervalo $(a; b)$, tenga una función inversa, es necesario que sea continua *) y monótona (creciente o decreciente) en este intervalo. Entonces la función f^{-1} también será monótona y continua en el conjunto de valores de la función $f(x)$.

*) Sobre la continuidad de una función véase p. 9.1.

Puede ser que en cierto intervalo $(a; b)$ la función es continua, pero no es monótona en este intervalo. Partamos el intervalo $(a; b)$ en intervalos

$$X_1, X_2, X_3, \ldots$$

de tal manera que en cada uno de éstos la función dada sea monótona. Entonces en cada uno de los intervalos $X_1, X_2, X_3, \ldots$ la función dada tendrá una función inversa. Por ejemplo, la función $f(x) = x^2$ con la región de definición $(-\infty; +\infty)$ es ininvertible. Partiendo el campo de definiciones en dos intervalos $(-\infty; 0)$ y $[0; +\infty)$, es fácil convencerse de que en cada uno de estos intervalos la función es continua y monótona, y por eso en cada uno de los intervalos indicados para la función dada existe una función inversa. En el intervalo $(-\infty; 0)$ tal función será la función $x = -\sqrt{y}$, y en el intervalo $[0; +\infty)$, la función $x = \sqrt{y}$. (Aquí y juega el papel de la variable independiente, y x, el papel de la variable dependiente). Las gráficas de las funciones directa e inversa en el intervalo $[0; +\infty)$ se muestran en la fig. 8.4.

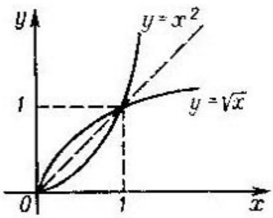


Fig. 8.4.

Propiedades de las funciones recíprocamente inversas:

1) Si la función f^{-1} es inversa para la función f, entonces también la función f es inversa para la función f^{-1}.

2) El campo de definición de la función f es la región de cambio de la función f^{-1}, y la región de cambio de la función f es el campo de definición de la función f^{-1}.

3) Las gráficas de las funciones recíprocamente inversas son simétricas con respecto a la bisectriz del primero y tercer ángulo de coordenadas del plano de coordenadas Oxy.

§ 7. Límite de una función

7.1. Concepto de límite de una función. Sea $f(x)$ una función numérica con campo de definición $D(f)$. En adelante, para mayor sencillez, supongamos que $D(f)$ representa cierto intervalo finito o infinito del eje numérico.

De los valores del argumento de la función $f(x)$ construimos una sucesión

$$x_1, x_2, x_3, \ldots, x_n, \ldots \ (x_n \in D(f),\ n \in \mathbf{N}), \qquad (1)$$

que converge hacia cierto número a, el cual, en general, puede no pertenecer al campo de definición de la función:

$$\lim_{n\to\infty} x_n = a.$$

Los valores de la función $f(x)$ en los puntos

$$x_1,\ x_2,\ x_3,\ \ldots,\ x_n,\ \ldots$$

forman una sucesión de valores de la función $f(x)$:

$$f(x_1),\ f(x_2),\ f(x_3),\ \ldots,\ f(x_n),\ \ldots,\ n \in \mathbf{N}. \qquad (2)$$

El número A se llama *límite de la función* $f(x)$ para x que tiende a a, si para cualquier sucesión que cojamos (1) con límite a, la sucesión que le corresponde de valores de la función (2) converge hacia el número A. El hecho de que el número A es límite de la función $f(x)$ para x, que tiende hacia a, se escribe como sigue:

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

Existe también otra definición equivalente de límite de una función:

El número A se llama *límite de la función* $f(x)$ para x, que tiende hacia a, si para cualquier número positivo ε se halla tal número positivo δ, que depende de ε, que para todos $x \neq a$, pertenecientes al campo de definición de la función y que satisfacen la desigualdad

$$|x - a| < \delta,$$

se cumple la desigualdad

$$|f(x) - A| < \varepsilon$$

Ejemplos. 1. Demostrar que el límite de la función $f(x) = \sqrt{x}$ para x, que tiende hacia a $(a > 0)$, es igual a $\sqrt{a}$.

La demostración la realizaremos aplicando la segunda definición de límite, es decir, mostraremos que para cualquier número $\varepsilon > 0$ se halla tal número $\delta > 0$ (para cada ε, es posible el suyo), que para todos x, pertenecientes al campo de definición de la función, de la desigualdad

$$|x - a| < \delta$$

se deduce la desigualdad

$$|\sqrt{x}-\sqrt{a}|<\varepsilon.$$

La expresión $|\sqrt{x}-\sqrt{a}|$ se puede reducir a la forma

$$|\sqrt{x}-\sqrt{a}|=\left|\frac{x-a}{\sqrt{x}+\sqrt{a}}\right|=\frac{|x-a|}{|\sqrt{x}+\sqrt{a}|}=\frac{|x-a|}{\sqrt{x}+\sqrt{a}}.$$

Suprimiendo en el denominador de la última fracción el sumando no negativo $\sqrt{x}$ y aumentando de esta manera la fracción, obtendremos

$$|\sqrt{x}-\sqrt{a}|\leqslant\frac{|x-a|}{\sqrt{a}}. \tag{3}$$

Luego, al tomar $\delta=\varepsilon\sqrt{a}$, obtenemos, que de la desigualdad

$$|x-a|<\delta$$

se deduce la desigualdad

$$\frac{|x-a|}{\sqrt{a}}<\frac{\delta}{\sqrt{a}}\Rightarrow\frac{|x-a|}{\sqrt{a}}<\varepsilon$$

y en vigor de la desigualdad (3)

$$|\sqrt{x}-\sqrt{a}|<\varepsilon.$$

Así pues, para cualquier $\varepsilon>0$ existe tal $\delta>0$ (es decir, $\delta=\varepsilon\sqrt{a}$), que para los x, que satisfacen la desigualdad $|x-a|<\delta$, se cumple la desigualdad $|\sqrt{x}-\sqrt{a}|<\varepsilon$. Sin embargo, con tal elección δ puede resultar que no para todos los valores x, que satisfacen la desigualdad $|x-a|<$ $<\delta$, se cumple la desigualdad $|\sqrt{x}-\sqrt{a}|<\varepsilon$, puesto que entre los valores x que satisfacen la desigualdad $|x-a|<\delta$, pueden existir valores, menores que cero, es decir, los que no entran en el campo de definición de la función $f(x)=\sqrt{x}$.

Eligiendo ahora en calidad de δ el mínimo de los números $\varepsilon\sqrt{a}$ y a, obtenemos, que para todos los valores x que satisfacen la desigualdad

$$|x-a|<\delta,$$

se cumplirá la desigualdad

$$|\sqrt{x}-\sqrt{a}|<\varepsilon,$$

es decir, el número $\sqrt{a}$ es realmente límite de la función $f(x)=\sqrt{x}$ para x que tiende hacia a.

2. Demostrar que el límite de la función

$$f(x)=\frac{3x^2-3}{x+1}$$

para x, que tiende hacia -1, es igual a -6. (El campo de definición de la función dada: $(-\infty;\ -1)\cup(-1;\ +\infty)$.)

Demostremos, que para cualquier $\varepsilon>0$ existe tal valor $\delta=\delta(\varepsilon)>0$, que para todos $x\neq-1$, que satisfacen la desigualdad $|x+1|<\delta$, se cumple la desigualdad

$$\left|\frac{3x^2-3}{x+1}+6\right|<\varepsilon.$$

Efectivamente, para $x\neq-1$ tenemos

$$\left|\frac{3x^2-3}{x+1}+6\right|=\left|\frac{3(x+1)(x-1)}{x+1}+6\right|=$$
$$=\left|\frac{3(x-1)}{1}+6\right|=3|x+1|.$$

Al tomar, por ejemplo, $\delta=\varepsilon/3$, obtendremos que de la desigualdad $|x+1|<\delta=\varepsilon/3$ se desprende la desigualdad $3|x+1|<\varepsilon$, o

$$\left|\frac{3x^2-3}{x+1}+6\right|<\varepsilon,$$

lo que queda completamente demostrado.

3. Demostrar que la función

$$f(x)=\operatorname{sen}\frac{1}{x}$$

no tiene límite para x que tiende a cero.

Realicemos la demostración, utilizando la primera definición de límite. Construyamos dos sucesiones de los valores del argumento:

$$\begin{gathered}\frac{2}{5\pi},\ \frac{2}{9\pi};\ \ldots;\ \frac{2}{\pi(1+4n)};\ \ldots,\\ \frac{2}{7\pi};\ \frac{1}{11\pi};\ \ldots;\ \frac{2}{\pi(3+4n)};\ \ldots\end{gathered} \tag{4}$$

Ambas sucesiones convergen hacia el número 0. Las sucesiones de los valores de la función

$$f(x)=\operatorname{sen}\frac{1}{x},$$

que corresponden a las sucesiones construidas (4) y tienen la forma

$$\begin{gathered}1;\ 1;\ 1;\ \ldots;\ 1;\ \ldots,\\ -1;\ -1;\ -1;\ \ldots;\ -1;\ \ldots\end{gathered}$$

Los límites de estas sucesiones serán los números 1 y —1 respectivamente.

De esta manera, están construidas dos sucesiones de los valores del argumento de la función

$$f(x)=\operatorname{sen}\frac{1}{x}$$

tales, que las sucesiones respectivas de los valores de la función dada tienen distintos límites; por consiguiente, la función dada no tiene límite para x, que tiende a cero.

7.2. Teorema sobre límites de funciones.

1) Si la función $f(x)$ tiene límite para x, que tiende hacia a, entonces este límite es único.

2) Si las funciones $f(x)$ y $g(x)$ tienen límites para x, que tiende hacia a, entonces las funciones

$$f(x)\pm g(x),\quad f(x)\cdot g(x),\quad \frac{f(x)}{g(x)}$$

también tienen límites para x, que tiende hacia a, además

$$\lim_{x\to a}[f(x)\pm g(x)]=\lim_{x\to a}f(x)\pm\lim_{x\to a}g(x),$$

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

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

(En el último caso se supone que la función $g(x)$ no se anula en un entorno suficientemente pequeño del punto a y $\lim\limits_{x \to a} g(x) \neq 0$.)

3) Si para x que tiende hacia a, la función $f(x)$ tiene límite, igual a A, y este límite es mayor que cierto número c, entonces para los valores bastante aproximados a a, x y la misma función $f(x)$ satisface la desigualdad

$$f(x) > c.$$

Utilizando los teoremas formulados más arriba, hallemos el límite de la función

$$f(x) = \frac{\sqrt{1+2x}-3}{\sqrt{x}-2}$$

para x, que tiende hacia 4. La aplicación directa del teorema sobre el límite del cociente es aquí imposible, puesto que el límite de la función que se encuentra en el denominador de la fracción es, para $x \to 4$, igual a cero. Por eso, antes de aplicar este teorema, realicemos las transformaciones siguientes:

$$\frac{\sqrt{1+2x}-3}{\sqrt{x}-2} = \frac{(\sqrt{1+2x}-3)(\sqrt{1+2x}-3)}{(\sqrt{x}-2)(\sqrt{1+2x}+3)} =$$

$$= \frac{2x-8}{(\sqrt{x}-2)(\sqrt{1+2x}+3)} = \frac{2(x-4)}{(\sqrt{x}-2)(\sqrt{1+2x}+3)} =$$

$$= \frac{2(\sqrt{x}+2)(\sqrt{x}-2)}{(\sqrt{x}-2)(\sqrt{1+2x}+3)} = \frac{2(\sqrt{x}+2)}{\sqrt{1+2x}+3}.$$

El límite de la función que se encuentra en el denominador de la última fracción para $x \to 4$ existe y es distinto de cero, y el límite de la función que se encuentra en el numerador de la fracción existe también. Por eso para la última expresión ya es aplicable el teorema sobre el límite del cociente:

$$\lim_{x \to 4} \frac{\sqrt{1+2x}-3}{\sqrt{x}-2} = \lim_{x \to 4} \frac{2(\sqrt{x}+2)}{\sqrt{1+2x}+3} = \frac{\lim\limits_{x \to 4} 2(\sqrt{x}+2)}{\lim\limits_{x \to 4} (\sqrt{1+2x}+3)}.$$

Utilizando en adelante los teoremas sobre los límites de la suma y del producto de las funciones, obtendremos

$$\frac{\lim\limits_{x\to 4}(2\sqrt{x}+4)}{\lim\limits_{x\to 4}(\sqrt{1+2x}+3)}=\frac{2\lim\limits_{x\to 4}\sqrt{x}+\lim\limits_{x\to 4}4}{\lim\limits_{x\to 4}\sqrt{1+2x}+\lim\limits_{x\to 4}3}.$$

Teniendo en cuenta ahora que $\lim\limits_{x\to a}\sqrt{x}=\sqrt{a}\ (a>0)$, tenemos

$$\frac{2\lim\limits_{x\to 4}\sqrt{x}+\lim\limits_{x\to 4}4}{\lim\limits_{x\to 4}\sqrt{1+2x}+\lim\limits_{x\to 4}3}=\frac{4+4}{3+3}=\frac{4}{3}.$$

Así pues,

$$\lim_{x\to 4}\frac{\sqrt{1+2x}-3}{\sqrt{x}-2}=\frac{4}{3}.$$

Sin embargo observemos, que aquí, al igual que para calcular los límites de las sucesiones, ponemos «ilegalmente» el signo de igualdad entre las dos expresiones, las cuales se obtienen como resultado de la aplicación del teorema sobre el límite del cociente, puesto que la afirmación de que existen límites del numerador y del denominador no fue demostrada. Sólo más tarde, al aclarar que estos límites existen, argumentamos la validez de la unión de estas dos expresiones por el signo de igualdad.

7.3. Condición necesaria y suficiente para que exista el límite de una función (criterio de Cauchy). Para que la función $f(x)$, para x que tiende hacia a, tenga un límite, es necesario y suficiente que para cualquier $\varepsilon>0$ exista un número positivo δ (en general, dependiente de ε) tal, que

$$|f(x')-f(x'')|<\varepsilon,$$

cuando $0<|x'-a|<\delta$ y $0<|x''-a|<\delta$, donde x' y x'' son puntos del campo de definición de la función $f(x)$.

7.4. Algunos límites básicos.

1) $\lim\limits_{x\to 0}\dfrac{\operatorname{sen} x}{x}=1,$

2) $\lim\limits_{x\to 0}(1+x)^{1/x}=e,$

3) $\lim\limits_{x\to 0}\dfrac{a^x-1}{x}=\ln a$, $\lim\limits_{x\to 0}\dfrac{e^x-1}{x}=1$,

4) $\lim\limits_{x\to 0}\dfrac{\log_a(1+x)}{x}=\log_a e$, $\lim\limits_{x\to 0}\dfrac{\ln(1+x)}{x}=1$.

§ 8. Infinitésimos

8.1. Concepto de variable infinitamente pequeña. Una función numérica $f(x)$ se llama *función infinitamente pequeña* (o *infinitésimo*), si cuando x tiende hacia a, el límite de la función $f(x)$ existe y es igual a cero.

Con otras palabras, la función $f(x)$ se llama *infinitamente pequeña* cuando x tiende hacia a, si para cualquier número $\varepsilon < 0$ se halla tal número positivo δ, dependiente de ε, que para todos $x \neq a$, pertenecientes al campo de definición de la función y que satisfacen a la desigualdad

$$|x-a|<\delta,$$

se cumple la desigualdad

$$|f(x)|<\varepsilon.$$

Subrayemos, que el infinitésimo se debe comprender como una variable, la cual en el proceso de cambio (para $x \to a$) se hace menor que el número arbitrario ε. Por eso, por ejemplo, la afirmación de que «una millonésima es una variable infinitamente pequeña» es errónea: no tiene sentido decir que un número es infinitamente pequeño.

Ejemplos de infinitésimos:

$f(x)=x^2$ para x, tendiente a cero;

$f(x)=\cos x$ para x, tendiente a $\pi/2$;

$f(x)=\ln x$ para x, tendiente a la unidad.

8.2. Comparacion de los infinitésimos. Los ejemplos citados de distintos infinitésimos nos llevan a la idea de, si unos infinitésimos no tenderán a cero «mas rápidamente» que otros. Ante todo, tenemos que determinar que debe entenderse por la palabra «más rápidamente», es decir, cómo comparar dos infinitésimos.

Dos infinitésimos $\alpha(x)$ y $\beta(x)$ para x, que tiende hacia a, se comparan de la manera siguiente. Examinemos la razón de estas variables

$$\frac{\beta(x)}{\alpha(x)}$$

(en este caso se supone que la variable que se encuentra en el denominador de la fracción es distinta de cero, por lo menos, para los valores x, bastante aproximados al número a) y cualculemos el límite de la razón

$$\frac{\beta(x)}{\alpha(x)}$$

para x, que tiende hacia a.

1) Si la razón $\frac{\beta(x)}{\alpha(x)}$ tiene un límite distinto de cero para x, que tiende hacia a, entonces los infinitésimos $\alpha(x)$ y $\beta(x)$ se consideran del mismo orden.

2) Si la razón $\frac{\beta(x)}{\alpha(x)}$ resulta ser ella misma un infinitésimo para x, que tiende hacia a, entonces el infinitésimo $\beta(x)$ se considera de orden superior en comparación con el infinitésimo $\alpha(x)$, y el infinitésimo $\alpha(x)$ se considera de orden inferior en comparación con la variable $\beta(x)$.

Por ejemplo, para x que tiende a cero serán infinitésimos de un mismo orden los siguientes pares de funciones:

$$\alpha(x)=x, \quad \beta(x)=\sqrt{1+x}-1;$$
$$\alpha(x)=x, \quad \beta(x)=\operatorname{sen} x.$$

Para x, que tiende a cero, el infinitésimo $\beta(x)=1-\cos x$ será un infinitésimo de orden superior en comparación con el infinitésimo $\alpha(x)=x$, puesto que

$$\lim_{x\to 0}\frac{\beta(x)}{\alpha(x)}=\lim_{x\to 0}\frac{1-\cos x}{x}=\lim_{x\to 0}\frac{2\operatorname{sen}^2\frac{x}{2}}{x}=$$

$$=\lim_{x\to 0}\frac{\operatorname{sen}^2\frac{x}{2}}{\frac{x}{2}}=\lim_{x\to 0}\frac{\operatorname{sen}\frac{x}{2}}{\frac{x}{2}}\cdot\lim_{x\to 0}\operatorname{sen}\frac{x}{2}=1\cdot 0=0.$$

3) Si para x, que tiende hacia a, el límite de la razón $\frac{\beta(x)}{\alpha(x)}$ no existe, y el límite de la razón $\frac{\alpha(x)}{\beta(x)}$ es igual a cero, entonces el infinitésimo $\alpha(x)$ se considera de orden superior en comparación con la variable $\beta(x)$.

4) Si para x, que tiende hacia a, ninguna de las dos razones

$$\frac{\alpha(x)}{\beta(x)}, \quad \frac{\beta(x)}{\alpha(x)}$$

no tiene límite, entonces los infinitésimos $\alpha(x)$ y $\beta(x)$ se consideran incomparables entre sí.

Por ejemplo, los infinitésimos $\alpha(x)=x$ y $\beta(x)=x\operatorname{sen}\frac{1}{x}$ para x, que tiende a cero, son incomparables entre sí.

El infinitésimo de orden superior con respecto al infinitésimo $\alpha(x)$ suele designarse con el símbolo $o(\alpha)$.

El infinitésimo $\beta(x)$ se llama *infinitésimo de k-ésimo orden* con respecto a la variable $\alpha(x)$, si las variables $\beta(x)$ y $\alpha(x)$ son del mismo orden.

Por ejemplo, para x, que tiende a cero, el infinitésimo $\beta(x) = 1 - \cos x$ con respecto al infinitésimo $\alpha(x) = x$ será un infinitésimo de segundo orden, puesto que

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

§ 9. Continuidad de las funciones

9.1. Concepto de continuidad de una función. Con el concepto de límite está estrechamente relacionado otro concepto importante del análisis matemático, el concepto de continuidad de una función.

Examinemos una función $f(x)$, definida en cierto intervalo. Sea x_0 cierto punto de este intervalo, en el cual la función tiene un valor bastante definido $f(x_0)$.

Cuando formulamos el concepto de límite de una función

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

se subrayaba, que el número x_0 puede no pertenecer al campo de definición de la función $f(x)$ (véase p. 7.1), y si es que el número x_0 pertenece al campo de definición, el valor de la función en este punto no nos interesaba para el estudio del límite.

Sin embargo, representa un gran interés precisamente el caso, cuando

$$\lim_{x\to x_0} f(x) = f(x_0). \tag{1}$$

Se dice que la función $f(x)$ es *continua* en el punto x_0, si se cumple la igualdad (1); si la igualdad (1) no se cumple, entonces se dice que en el punto x_0 la función tiene *discontinuidad*.

De esta manera, la aclaración del problema sobre la continuidad de la función $f(x)$ en cualquier punto x_0, perteneciente al campo de definición de la función, se reduce a calcular el límite de la función en el punto x_0

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

y a verificar la validez de la igualdad (1).

Si la función $f(x)$ está definida en el intervalo $(a; b)$ y es continua en cada punto del intervalo, entonces se dice que la función es continua en este intervalo.

9.2. Teoremas fundamentales sobre las funciones continuas.

1) Si las funciones $f(x)$ y $g(x)$ están definidas en el intervalo $(a; b)$ y son continuas en el punto $x_0 \in (a; b)$, entonces en este punto las funciones

$$f(x) \pm g(x), \quad f(x) \cdot g(x), \quad \frac{f(x)}{g(x)}$$

también son continuas (la última función es continua a condición de que $g(x_0) \neq 0$).

2) Si la función $y = f(x)$ es continua en el punto x_0, y la función $g(y)$ es continua en el punto $y_0 = f(x_0)$, entonces también la función compuesta $g(f(x))$ es continua en el punto x_0.

3) *Primer teorema de Bolzano — Cauchy.* Si la función $f(x)$ está dada en el intervalo $[a; b]$ y es continua en el intervalo $(a; b)$, y si en los extremos del intervalo los valores de la función $f(a)$ y $f(b)$ tienen los signos distintos, entonces entre a y b existe (por lo menos uno) tal valor x_0, para el cual $f(x)$ se anula:

$$f(x) = 0 \quad (x_0 \in (a; b)).$$

El teorema de Bolzano — Cauchy halla una aplicación bastante inesperada para la solución de las ecuaciones algebraicas, que permite, precisamente, demostrar que cualquier ecuación algebraica de potencia impar con coeficientes reales

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0$$

tiene por lo menos una raíz real.

Para los valores x, suficientemente grandes por su valor absoluto, el polinomio que se encuentra en el primer miembro de la ecuación tiene el signo del término mayor, es decir, para $x > 0$, el signo a_0, y para $x < 0$, el signo contrario. Puesto que un polinomio es una función continua en todo el conjunto de números reales, entonces, según el teorema de Bolzano — Cauchy, obligatoriamente se halla tal punto x_0, en el cual este polinomio se anula, y, por consiguiente, la ecuación algebraica se reduce a una identidad.

4) *Primer teorema de Weierstrass.* Si la función $f(x)$ está definida y es continua en el intervalo $[a; b]$, entonces es acotada en este intervalo, es decir, existen tales dos números m y M, que para cualquier $x \in [a; b]$ tiene lugar la desigualdad

$$m \leqslant f(x) \leqslant M.$$

9.3. Continuidad de las funciones elementales. Utilizando la definición de continuidad de la función en un punto, es fácil demostrar que las funciones $f(x) = c$ y $f(x) = x$ son continuas en cualquier punto del eje numérico. De la continuidad de estas funciones según el teorema 1) se deduce la continuidad de una función racional entera

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

y la continuidad de una función racional para todos los valores x que no anulan el denominador

$$f(x) = \frac{a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n}{b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m}.$$

De la definición de continuidad (1) se puede deducir la continuidad de las funciones sen x y cos x. De la continuidad de estas funciones en vigor del teorema 1) se desprende que las funciones tg x, ctg x, sec x y cosec x también son continuas para todos los valores del argumento, pertenecientes al campo de definición de las funciones.

La continuidad de las funciones trigonométricas inversas se deduce de las propiedades de las funciones inversas (véase p. 6.9).

La función potencial x^α, la función exponencial a^x y la función logarítmica $\log_a x$ son continuas en el campo de su definición.

CAPÍTULO 9

Elementos del cálculo integral y diferencial

§ 1. Derivada

1.1. Concepto de la derivada. Sea $f(x)$ cierta función, definida en el intervalo $(a; b)$, y x_0 cierto punto fijo de este intervalo. Tomemos un punto arbitrario x del intervalo $(a; b)$ y hagamos una diferencia

$$x - x_0.$$

La diferencia $x - x_0$ se llama *incremento del argumento* de la función $f(x)$ y se designa Δx:

$$x - x_0 = \Delta x. \tag{1}$$

La diferencia entre el valor de la función en el punto $x = x_0 + \Delta x$ y el valor de la función en el punto x_0:

$$f(x_0 + \Delta x) - f(x_0), \tag{2}$$

se llama *incremento de la función* $f(x)$ en el punto x_0. El incremento de la función $f(x)$ en el punto x_0 se designa $\Delta f(x_0)$:

$$f(x_0 + \Delta x) - f(x_0) = \Delta f(x_0).$$

Puesto que el punto x_0 se considera fijo, el incremento de la función $\Delta f(x_0)$ será la función del incremento del argumento Δx.

Se llama *derivada* de una función $f(x)$ en el punto x_0 al límite de la razón entre el incremento de la función $\Delta f(x_0)$ y el incremento de la variable independiente Δx para Δx que tiende a cero:

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Si este límite existe, entonces se dice que la función $f(x)$ *tiene una derivada* en el punto x_0 o que la función $f(x)$ *es diferenciable en el punto* x_0.

La derivada de la función en el punto x_0 se designa

$$f'(x_0) \text{ ó } \frac{df(x_0)}{dx} \text{ ó } f'(x)|_{x=x_0}.$$

De esta manera,

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x}.$$

Se dice que la función $f(x)$ *es diferenciable en el intervalo* $(a; b)$, si es diferenciable en cada punto de este intervalo. La derivada para el valor dado $x = x_0$ (si existe) es un número definido; si la derivada existe en todo el intervalo $(a; b)$, es decir, en cada punto del intervalo (a, b), entonces ella misma es función de x.

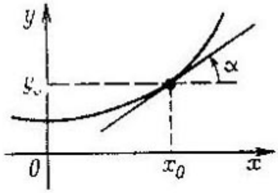


Fig. 9.1.

1.2. Ecuación de la tangente a la gráfica de una función, representada en forma explícita. Sea $f(x)$ cierta función, diferenciable en el punto x_0; $(x_0; y_0 = f(x_0))$ son las coordenadas del punto que pertenece a la gráfica de la función $y = f(x)$. En el sistema de coordenadas cartesianas rectangulares Oxy la ecuación de la tangente a la gráfica de la función $y = f(x)$ que pasa por el punto $(x_0; y_0)$ tiene la forma

$$y - y_0 = f'(x_0)(x = x_0).$$

El valor de la derivada $f'(x)$ en el punto x_0 es igual a la tangente del ángulo de inclinación de una tangente (fir. 9.1):

1.3. Sentido físico de la derivada. Supongamos que un punto material se mueve por una recta bajo la acción de ciertas fuerzas. Elegimos un momento cualquiera de tiempo t_0 y examinemos el intervalo de tiempo Δt desde el momento t_0 hasta el momento

$$t = t_0 + \Delta t.$$

Durante este intervalo de tiempo el punto hará un recorrido que designaremos $\Delta S(t_0)$. Según una conocida definición de física, la razón

$$\frac{\Delta S(t_0)}{\Delta t}$$

es la velocidad media del movimiento del punto en un tiempo igual a Δt. Examinemos los intervalos Δt cada vez más pequeños, tendiendo Δt hacia cero. El límite de la razón

$$\lim_{\Delta t \to 0} \frac{\Delta S(t_0)}{\Delta t}$$

es la velocidad momentánea del punto en el momento de tiempo t_0.

1.4. Teoremas de las derivadas.

1) Si la función $f(x)$ es diferenciable en el punto x_0, entonces y la función $c \cdot f(x)$ (c es constante) también es diferenciable en el punto x_0, además

$$(c \cdot f(x))'_{x=x_0} = c \cdot f'(x_0),$$

es decir, el factor constante se puede sacar del signo de la derivada.

2) Si las funciones $f(x)$ y $g(x)$ son diferenciables en el punto x_0, entonces también las funciones $f(x) + g(x)$ y $f(x) - g(x)$ son diferenciables en este punto, además

$$(f(x) \pm g(x))'_{x=x_0} = f'(x_0) \pm g'(x_0),$$

es decir, la derivada de la suma (diferencia) de dos funciones es igual a la suma (diferencia) de sus derivadas.

3) Si las funciones $f(x)$ y $g(x)$ son diferenciables en el punto x_0, entonces las funciones $f(x) \cdot g(x)$ y $\frac{f(x)}{g(x)}$ también son diferenciables en este punto, además

$$(f(x) \cdot g(x))'_{x=x_0} = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0),$$

$$\left(\frac{f(x)}{g(x)}\right)'_{x=x_0} = \frac{f'(x_0) \cdot g(x_0) - f(x_0) \cdot g'(x_0)}{g^2(x_0)} \; (g(x_0) \neq 0).$$

4) Si la función $y = f(x)$ es diferenciable en el punto x_0, y la función $z = g(y)$ es diferenciable en el punto $y_0 =$

$= f(x_0)$, entonces la función compuesta

$$z = g(f(x))$$

también es diferenciable en el punto x_0, además

$$(g(f(x)))'_{x=x_0} = g'_y(y_0) \cdot f'_x(x_0),$$

donde los índices y y x de las derivadas indican por qué argumento se calculan las derivadas.

5) Sea que la función $y = f(x)$ satisface las condiciones de existencia de la función inversa y en el punto x_0 tiene una derivada finita, distinta de cero. Entonces la función inversa $x = f^{-1}(y)$ en el punto correspondiente $y_0 = f(x_0)$ también tiene la derivada. Las derivadas de dos funciones recíprocamente inversas están relacionadas por la igualdad

$$(f^{-1}(y))'_{y=y_0} = \frac{1}{f'(x_0)}.$$

1.5. Cálculo de las derivadas de las funciones elementales. Aplicando la definición de una derivada, se puede demostrar que la derivada de la función $f(x) = c$ (c es constante) es igual a 0, y la derivada de la función $f(x) = x$ es igual a la unidad.

Calculemos la derivada de la función $f(x) = \operatorname{sen} x$. Sea x_0 un punto arbitrario de una recta numérica. Hagamos el incremento de la función

$$\Delta \operatorname{sen} x_0 = \operatorname{sen}(x_0 + \Delta x) - \operatorname{sen} x_0.$$

La diferencia, que se encuentra en el segundo miembro de la igualdad la escribimos en forma del producto

$$\operatorname{sen}(x_0 + \Delta x) - \operatorname{sen} x_0 = 2\cos\left(x_0 + \frac{\Delta x}{2}\right) \operatorname{sen}\frac{\Delta x}{2}.$$

Según la definición de la derivada

$$(\operatorname{sen} x)'_{x=x_0} = \lim_{\Delta x \to 0} \frac{\Delta \operatorname{sen} x_0}{\Delta x} = \lim_{\Delta x \to 0} \frac{2\cos\left(x_0 + \frac{\Delta x}{2}\right) \operatorname{sen}\frac{\Delta x}{2}}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \cos\left(x_0 + \frac{\Delta x}{2}\right) \lim_{\Delta x \to 0} \frac{\operatorname{sen}\frac{\Delta x}{2}}{\frac{\Delta x}{2}}.$$

En vigor de la continuidad de la función $\cos x$

$$\lim_{\Delta x \to 0} \cos\left(x_0 + \frac{\Delta x}{2}\right) = \cos x_0.$$

Empleando el límite 1) p. 7.4 del capítulo 8, obtenemos

$$\lim_{\Delta x \to 0} \frac{\operatorname{sen} \frac{\Delta x}{2}}{\frac{\Delta x}{2}} = 1.$$

De esta manera,

$$(\operatorname{sen} x)'_{x=x_0} = \cos x_0.$$

Puesto que el punto x_0 es un punto arbitrario del eje numérico, se escribe

$$(\operatorname{sen} x)' = \cos x.$$

Calculemos la derivada de la función exponencial $f(x) = a^x$ en el punto $x = 0$. Escribamos el incremento de la función en el punto $x = 0$:

$$a^{0+\Delta x} - a^0 = a^{\Delta x} - 1.$$

Según la definición de la derivada

$$(a^x)'_{x=0} = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

Basándose en la fórmula 3) p. 7.4 del capítulo 8, tenemos

$$\lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x} = \ln a.$$

Así, obtenemos

$$(a^x)'_{x=0} = \ln a.$$

La derivada de la función $f(x) = a^x$ en todos los demás puntos del eje numérico se calcula de la manera siguiente:

$$(a^x)' = \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x} = \lim_{\Delta x \to 0} \frac{a^x (a^{\Delta x} - 1)}{x}.$$

Puesto que el factor a^x que se encuentra bajo el signo de límite no depende de Δx, el último límite se puede escribir en forma

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

Así, definitivamente obtenemos

$$(a^x)'=a^x \ln a.$$

La regla para calcular la derivada de la función exponencial permite dar una definición del número e (véase p. 2.4 del capítulo 8). La base de la función exponencial, cuya derivada es igual a la misma función:

$$(e^x)'=e^x,$$

se llama *número e*.

Conociendo, que la derivada de la función $f(x)=\operatorname{sen} x$ es igual a $\cos x$, mediante los teoremas de las derivadas se pueden calcular las derivadas de las funciones

$$f(x)=\cos x,\ f(x)=\operatorname{tg} x,\ f(x)=\operatorname{ctg} x.$$

Por ejemplo, la derivada de la función $f(x)=\cos x$ se calcula de la manera siguiente: representemos la función $\cos x$ en la forma

$$\cos x=\operatorname{sen}\left(\frac{\pi}{2}-x\right);$$

entonces según el teorema 4) p.1.4 tenemos

$$(\cos x)'=\left(\operatorname{sen}\left(\frac{\pi}{2}-x\right)\right)'=$$
$$=\cos\left(\frac{\pi}{2}-x\right)\cdot\left(\frac{\pi}{2}-x\right)'=-\cos\left(\frac{\pi}{2}-x\right)=-\operatorname{sen} x.$$

Las derivadas de las funciones trigonométricas inversas $f(x)=\arcsen x$, $f(x)=\arccos x$, $f(x)=\operatorname{arctg} x$ y $f(x)=\operatorname{arcctg} x$ se calculan mediante el teorema 5) p. 1.4. Por ejemplo, la derivada de la función $f(x)=\operatorname{arcsen} x$ se calcula de la manera siguiente: para la función $y=\operatorname{arcsen} x$ la función inversa será $x=\operatorname{sen} y$; según el teorema de la derivada de una función inversa tenemos

$$(\operatorname{arcsen} x)'=\frac{1}{(\operatorname{sen} y)'}=\frac{1}{\cos y}=\frac{1}{\cos(\operatorname{arcsen} x)}.$$

Puesto que $\cos(\operatorname{arcsen} x) = \sqrt{1-x^2}$, entonces

$$(\operatorname{arcsen} x)' = \frac{1}{\sqrt{1-x^2}}.$$

Conociendo la derivada de la función exponencial $(a^x)' = a^x \ln a$, mediante el teorema 5) p. 1.4 se puede calcular la derivada de la función logarítmica $f(x) = \log_a x$:

$$(\log_a x)' = \frac{1}{x \ln a}.$$

La derivada de la función potencial $f(x) = x^\alpha$ se puede calcular como una derivada de la función compuesta, si representamos la función x^α en la forma

$$x^\alpha = e^{\ln(x^\alpha)} = e^{\alpha \ln x}.$$

Según la regla del cálculo de la derivada de una función compuesta

$$(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x}(\alpha \ln x)' = e^{\alpha \ln x} \cdot \alpha \cdot \frac{1}{x} =$$

$$= x^\alpha \cdot \alpha \cdot \frac{1}{x} = \alpha x^{\alpha-1}.$$

Las *derivadas de ciertas funciones simples* se dan en la tabla siguiente:

1) $f(x) = c, \quad f'(x) = 0;$
2) $f(x) = x^\alpha, \quad f'(x) = \alpha x^{\alpha-1};$
3) $f(x) = a^x, \quad f'(x) = a^x \ln a;$
   $f(x) = e^x, \quad f'(x) = e^x;$
4) $f(x) = \log_a x, \quad f'(x) = \dfrac{1}{x \cdot \ln a};$
   $f(x) = \ln x, \quad f'(x) = \dfrac{1}{x};$
5) $f(x) = \operatorname{sen} x, \quad f'(x) = \cos x;$
6) $f(x) = \cos x, \quad f'(x) = -\operatorname{sen} x;$
7) $f(x) = \operatorname{tg} x, \quad f'(x) = \dfrac{1}{\cos^2 x};$

8) $f(x)=\operatorname{ctg} x, \qquad f'(x)=-\dfrac{1}{\operatorname{sen}^2 x}$;

9) $f(x)=\operatorname{arcsen} x, \qquad f'(x)=\dfrac{1}{\sqrt{1-x^2}}$;

10) $f(x)=\arccos x, \qquad f'(x)=-\dfrac{1}{\sqrt{1-x^2}}$;

11) $f(x)=\operatorname{arctg} x, \qquad f'(x)=\dfrac{1}{1+x^2}$;

12) $f(x)=\operatorname{arcctg} x, \qquad f'(x)=-\dfrac{1}{1+x^2}$.

1.6. Derivadas de orden superior. Sea que la función $f(x)$ es diferenciable en el intervalo (a; b). La derivada de la función dada $f'(x)$ representa una función del argumento x Elijamos cierto punto $x_0 \in (a; b)$ y hagamos el incremento del argumento $\Delta x = x - x_0$ y el incremento de la función $f'(x)$ en el punto x_0:

$$\Delta f'(x_0) = f'(x_0+\Delta x) - f'(x_0).$$

Examinemos el límite

$$\lim_{\Delta x \to 0} \frac{f'(x_0+\Delta x)-f'(x_0)}{\Delta x}.$$

Si este límite existe, entonces se dice que la función $f(x)$ tiene una derivada de segundo orden (o una derivada segunda) en el punto x_0. La derivada del segundo orden de la función $f(x]$ en el punto x_0 se designa

$$f''(x_0) \text{ ó } \frac{d^2 f(x_0)}{dx^2} \text{ ó } f'(x)|_{x=x_0}.$$

De esta manera,

$$f''(x_0) = \lim_{\Delta x \to 0} \frac{f'(x_0+\Delta x)-f'(x_0)}{\Delta x}.$$

Luego, si para la función $f''(x)$ en cierto punto x_0 existe un límite

$$\lim_{\Delta x \to 0} \frac{f''(x_0+\Delta x)-f''(x_0)}{\Delta x},$$

entonces se dice que la función $f(x)$ tiene una derivada de tercer orden (o la derivada tercera) en el punto x_0:

$$f'''(x_0) = \lim_{\Delta x \to 0} \frac{f''(x_0 + \Delta x) - f''(x_0)}{\Delta x}.$$

Análogamente se definen las derivadas de cuarto, quinto. etc. orden. La derivada de n-ésimo orden de la función $f(x)$ suele designarse.

$$f^{(n)}(x) \text{ o } \frac{d^n f}{dx^n}.$$

Las fórmulas del cálculo de la derivada de n-ésimo orden de la suma y del producto de dos funciones $f(x)$ y $g(x)$ son:

$$(f \pm g)^{(n)} = f^{(n)} \pm g^{(n)},$$

$$(f \cdot g)^{(n)} = C_n^0 f^{(n)} \cdot g^{(0)} + C_n^1 \cdot f^{(n-1)} g' +$$
$$+ C_n^2 f^{(n-2)} g'' + \ldots + C_n^{n-1} f' g^{(n-1)} + C_n^n f^{(0)} g^{(n)} =$$
$$= \sum_{i=0}^{n} C_n^i f^{(n-i)} g^{(i)}.$$

En la última fórmula mediante $f^{(0)}$ y $g^{(0)}$ se designan las funciones $f(x)$ y $g(x)$ respectivamente, C_n^i es el número de combinaciones de n elementos por i elementos (coeficientes binomiales).

§ 2. Función primitiva. Integral indefinida

2.1. Concepto de función primitiva y de integral indefinida. Si un punto material se mueve por una recta, sin cambiar la dirección de su movimiento, y la distancia, recorrida por el punto, es una función conocida del tiempo $S(t)$, entonces la velocidad momentánea del punto se calcula como una derivada del camino recorrido:

$$v = S'(t).$$

Sea que ahora conocemos la ley del cambio de la velocidad del punto, es decir, está dada la función $v(t)$ y es necesario hallar la distancia recorrida por el punto material en cierto período de tiempo T. Para resolver este problema, es necesario por la derivada conocida (en nuestro caso la velocidad $v(t)$) reconstruir la misma función (en nuestro caso $S(t)$).

La reconstrucción de una función a partir de su derivada conocida es uno de los principales problemas del *cálculo integral*. Este problema es inverso al problema fundamental del *cálculo diferencial*, el cual consiste en hallar la derivada de la función dada.

La función $F(x)$ se llama función *primitiva* para la función $f(x)$ en el intervalo dado del eje numérico, si para todos los valores x de este intervalo la función $f(x)$ es la derivada de la función $F(x)$:

$$F'(x) = f(x) \tag{1}$$

Propiedad principal de las primitivas: si la función $F(x)$ es una función primitiva de la función $f(x)$, entonces también la función $F(x) + C$, donde C es una constante cualquiera, es una primitiva de la función $f(x)$. A la inversa, si $F(x)$ es cierta primitiva de la función $f(x)$, entonces cualquier primitiva de la función $f(x)$ puede ser escrita en la forma

$$F(x) + C,$$

donde C es cierta constante.

De esta manera, para hallar todas las primitivas de la función dada $f(x)$, es suficiente hallar sólo una primitiva $F(x)$; todas las demás primitivas se obtienen de la primitiva hallada mediante la adición de sumandos constantes.

La expresión $F(x) + C$ donde C es una constante arbitraria, representa la forma general de la función, la cual tiene la derivada $f(x)$. Esta expresión se llama *integral indefinida* de la función $f(x)$ y se escribe

$$F(x) + C = \int f(x)\,dx.$$

El producto $f(x)\,dx$ se llama *expresión subintegral*, y $f(x)$ se llama *función subintegral*.

La obtención de un conjunto de todas las primitivas de la función $f(x)$ se llama *integración* de esta función.

Aplicando la fórmula (1), de la tabla p. 1.5 se puede obtener la siguiente *tabla de integrales*:

1) $\int 0 \cdot dx = C$;

2) $\int 1 \cdot dx = x + C$;

3) $\int x^{\alpha}\, dx = \frac{x^{\alpha+1}}{\alpha+1} + C\ (\alpha \neq -1);$

4) $\int \frac{1}{x}\, dx = \ln|x| + C;$

5) $\int \frac{1}{1+x^2}\, dx = \operatorname{arctg} x + C;$

6) $\int \frac{1}{\sqrt{1-x^2}}\, dx = \operatorname{arcsen} x + C;$

7) $\int a^x\, dx = \frac{a^x}{\ln a} + C;$

8) $\int \operatorname{sen} x\, dx = -\cos x + C;$

9) $\int \cos x\, dx = \operatorname{sen} x + C;$

10) $\int \frac{1}{\cos^2 x}\, dx = \operatorname{tg} x + C;$

11) $\int \frac{1}{\operatorname{sen}^2 x}\, dx = -\operatorname{ctg} x + C.$

2.2. Reglas y métodos más simples de integración.

1) Sea $F'(x)$ una derivada de la función $F(x)$. Según la definición de la integral indefinida

$$\int F'(x)\, dx = F(x) + C.$$

2) El factor constante se puede sacar del signo de la integral:

$$\int c \cdot f(x)\, dx = c \cdot \int f(x)\, dx.$$

3) La integral de la suma de dos funciones es igual a la suma de las integrales de los sumandos:

$$\int [f(x) \pm g(x)]\, dx = \int f(x)\, dx \pm \int g(x)\, dx.$$

4) Si la función $F(x)$ es la primitiva de la función $f(x)$, entonces la primitiva de la función $f(ax+b)$ será la función $\frac{1}{a} F(ax+b)$.

Ejemplos. 1. Calcular la integral indefinida

$$\int (6x^3 + x - 1)\, dx.$$

Según las reglas 2), 3) la integral indefinida

$$\int (6x^3 + x - 1)\, dx$$

puede representarse en forma de la suma de las integrales:

$$\int (6x^3 + x - 1)\, dx = 6\int x^3\, dx + \int x\, dx - \int 1 \cdot dx.$$

Aplicando la fórmula 3) de la tabla de las integrales indefinidas, obtendremos

$$\int (6x^3 + x - 1)\, dx = \frac{3}{2} x^2 + \frac{1}{2} x^2 - x + C.$$

2. Integrar la función

$$f(x) = \frac{x - \sqrt{x}}{\sqrt[3]{x}}.$$

Representemos la función dada $f(x)$ en la forma

$$\frac{x - \sqrt{x}}{\sqrt[3]{x}} = \frac{x}{\sqrt[3]{x}} - \frac{x}{\sqrt[3]{x}} = \frac{x}{x^{1/3}} - \frac{x^{1/2}}{x^{1/3}} = x^{2/3} - x^{1/6}.$$

Aplicando las reglas 2), 3) y la fórmula 3) de la tabla de las integrales indefinidas, obtendremos

$$\int \frac{x - \sqrt{x}}{\sqrt[3]{x}}\, dx = \int [x^{2/3} - x^{1/6}]\, dx = \int x^{2/3}\, dx -$$

$$- \int x^{1/6}\, dx = \frac{3}{5} x^{5/3} - \frac{6}{7} x^{7/6} + C.$$

Se debe subrayar que la aplicación de las reglas 2), 3) da la posibilidad de integrar sólo las funciones muy simples. Existen métodos mucho más efectivos (pero también más complicados) de integración de distintas funciones. Uno de tales métodos es el de integración por partes. Sea $u(x)$ y $v(x)$ dos funciones continuas que tienen derivadas continuas. El método de *integración por partes* se basa en las dos igualdades siguientes:

$$(u \cdot v)' = u' \cdot v + u \cdot v',$$

$$\int (u \cdot v)'\, dx = u \cdot v + C,$$

De las igualdades citadas más arriba se deduce que

$$u \cdot v = \int u' \cdot v \, dx + \int u \cdot v' \, dx,$$

ó

$$\int u \cdot v' \, dx = u \cdot v - \int u' \cdot v \, dx.$$

Ejemplo. Calcular la integral

$$\int x \cos x \, dx.$$

Vamos a integrar la integral dada mediante el método de integración por partes. Supongamos que la función subintegral representa un producto de dos factores de la función $u(x) = x$ y de la derivada de cierta función $v(x)$:

$$v'(x) = \cos x.$$

La derivada de la función $u(x)$ es igual a la unidad:

$$u'(x) = 1,$$

y la primitiva de la función $v'(x)$ es igual a sen x:

$$v(x) = \int \cos dx = \operatorname{sen} x.$$

Sustituyendo $u(x)$, $v(x)$ y $u'(x)$ en el segundo miembro de la igualdad (2), obtendremos

$$\int x \cos x \, dx = x \operatorname{sen} x - \int \operatorname{sen} x \, dx.$$

De esta manera, el cálculo de la integral dada se reduce al cálculo de la integral de la tabla

$$\int \operatorname{sen} x \, dx = -\cos x + C;$$

definitivamente obtenemos

$$\int x \cos x \, dx = x \operatorname{sen} x + \cos x + C.$$

Observemos que en casos más complicados el desarrollo de la función subintegral en factores puede ser no el único y es necesario elegirlo tal, que la integración de la función $v'(x)$ no presente dificultades y que el cálculo de la integral que se encuentra en el segundo miembro de la fórmula (2) sea más simple, que el cálculo de la integral que se encuentra en el primer miembro.

Ciertas integrales se calculan mediante el método de integración por partes.

Las integrales de la forma

$$\int x^k \ln^m x\,dx, \quad \int x^k \operatorname{sen} bx\,dx, \quad \int x^k e^{ax}\,dx,$$

donde k y m son números naturales, y a y b son números reales cualesquiera, se calculan mediante la integración por partes.

Ejemplos 1. Calcular la integral

$$\int x^3 \ln x\,dx.$$

Supongamos $u = \ln x$, $v' = x^3$, tal, que $u' = \frac{1}{x}$, $v = \frac{1}{4}x^4$. Entonces según la fórmula de integración por partes obtenemos

$$\int x^3 \ln x\,dx = \frac{1}{4}x^4 \ln x - \int \frac{1}{4}x^4 \cdot \frac{1}{x}\,dx =$$
$$= \frac{1}{4}x^4 \ln x - \frac{1}{4}\int x^3\,dx = \frac{1}{4}x^4 \ln x - \frac{1}{16}x^4 + C.$$

2. Calcular la integral

$$\int x^2 \operatorname{sen} x\,dx.$$

Supongamos $u = x^2$, $v' = \operatorname{sen} x$, tal que $u' = 2x$, $v = -\cos x$. Basándose en la fórmula de integración por partes

$$\int x^2 \operatorname{sen} x\,dx = -x^2 \cos x + 2\int x \cos x\,dx.$$

Aplicando a la integral que se encuentra a la derecha otra vez la fórmula de integración por partes, definitivamente obtendremos

$$\int x^2 \operatorname{sen} x\,dx = -x^2 \cos x + 2x \operatorname{sen} x + 2\cos x + C.$$

3. Calcular la integral

$$\int xe^{2x}\,dx.$$

Supongamos $u = x$, $v' = e^{2x}$. Entonces $u' = 1$, $v = \frac{1}{2}e^{2x}$. En vigor de la fórmula (2) la integral inicial se escribe en la forma

$$\int xe^{2x}\,dx = \frac{1}{2}xe^{2x} - \frac{1}{2}\int e^{2x}\,dx = \frac{1}{2}xe^{2x} - \frac{1}{4}e^{2x} + C,$$

De la misma manera, mediante la aplicación sucesiva de la integración por partes, se calculan las integrales

$$\int P(x)\, e^{ax}\, dx, \quad \int P(x) \operatorname{sen} bx\, dx, \quad \int P(x) \cos bx\, dx,$$

donde $P(x)$ es un polinomio respecto a x.

Mediante el método de integración por partes se puede calcular también las integrales

$$\int e^{ax} \operatorname{sen} bx\, dx, \quad \int e^{ax} \cos bx\, dx.$$

Si a las integrales dadas aplicamos el método de integración por partes (tomando en ambos casos $v' = e^{ax} \Rightarrow v = \frac{1}{a} e^{ax}$), entonces obtendremos las dos igualdades siguientes:

$$\int e^{ax} \cos bx\, dx = \frac{1}{a} e^{ax} \cos bx + \frac{b}{a} \int e^{ax} \operatorname{sen} bx\, dx, \qquad (3)$$

$$\int e^{ax} \operatorname{sen} bx\, dx = \frac{1}{a} e^{ax} \operatorname{sen} bx - \frac{b}{a} \int e^{ax} \cos bx\, dx. \qquad (4)$$

Multiplicando ambos miembros de la igualdad (4) por el valor b/a y sumando la igualdad obtenida con la igualdad (3), obtenemos una ecuación lineal respecto a la integral $\int e^{ax} \cos bx\, dx$, de resultas de la solución de la cual hallamos la integral que necesitamos:

$$\int e^{ax} \cos bx\, dx = \frac{b \operatorname{sen} bx + a \cos bx}{a^2 + b^2} \cdot e^{ax} + C.$$

Análogamente hallamos también la segunda integral:

$$\int e^{ax} \operatorname{sen} bx\, dx = \frac{a \operatorname{sen} bx - b \cos bx}{a^2 + b^2} \cdot e^{ax} + C'.$$

§ 3. Integral definida

3.1. Problema para calcular el área de una figura plana. Sea que en cierto segmento $[a; b]$ se da una función continua $f(x)$, la cual obtiene sólo valores positivos (negativos). Construyamos la gráfica de la función $y = f(x)$ en el sistema de coordenadas cartesiano rectangular Oxy (fig. 9.2).

Examinemos la figura $ABCD$, limitada por la curva $y = f(x)$, dos rectas, que se representan por las ecuaciones $x = a$ y $x = b$, y el eje Ox. A esta figura se la suele llamar *trapecio curvilíneo*.

Dividamos la base AB del trapecio en n partes (es posible, desiguales) y por los puntos de división tracemos unas rectas perpendiculares las cuales parten el trapecio curvilíneo en una serie de franjas. Para cada franja construimos un rectángulo, cuya base es la misma que la de la franja, y la altura coincide con uno de los lados laterales de la franja,

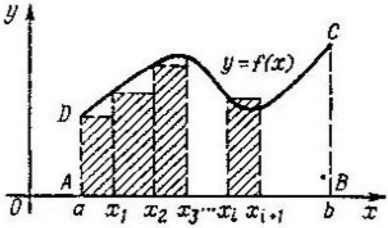


Fig. 9.2.

por ejemplo, con el izquierdo (en la fig. 9.2 estos rectángulos están rayados). Calculemos la suma de las áreas de todos los rectángulos construidos.

Sean las abscisas de los puntos de división

$$x_0,\ x_1,\ x_2,\ \ldots,\ x_n,$$

además,

$$x_0 = a < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b.$$

La base del i-ésimo rectángulo ($i = 0, 1, 2, \ldots, n-1$) es igual a la diferencia $x_{i+1} - x_i$, la cual designamos Δx_i, y la altura, conforme a lo expuesto más arriba, es igual a $y_i = f(x_i)$. Por eso el área del i-ésimo rectángulo es igual a $y_i \cdot \Delta x_i = f(x_i) \cdot \Delta x_i$. La suma de las áreas S de todos los rectángulos es igual a

$$S_n = f(x_0)\,\Delta x_0 + f(x_1)\,\Delta x_1 + \ldots + f(x_{n-1})\,\Delta x_{n-1} = \sum_{i=0}^{n-1} f(x_i)\,\Delta x_i.$$

El área S_n se llama valor aproximado del área del trapecio curvilíneo $ABCD$.

Por área del trapecio curvilíneo $ABCD$ se toma el número que es igual al límite de las áreas S_n en la suposición de que todas las longitudes Δx_i tienden al mismo tiempo a cero, cuando el número de puntos de partición del segmento AB aumenta ilimitadamente.

Todo lo expuesto anteriormente es válido también para el caso, cuando la función $f(x)$ puede obtener valores negativos. En este caso se considera que las áreas de unas partes de la figura, situadas más abajo del eje Ox, son negativas.

La definición dada más arriba del área de un trapecio curvilíneo sirve sólo para una clase bastante estrecha de funciones, precisamente para las funciones continuas y acotadas. La extensión del concepto de área de un trapecio curvilíneo para una clase más amplia de funciones conduce al concepto de integral definida.

3.2. Integral definida. Sea la función $f(x)$ dada en cierto segmento $[a;\ b]$. Partimos este segmento de manera arbitraria en n partes, poniendo entre los puntos a y b los puntos de división

$$a = x_0 < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b.$$

Tomemos en cada uno de los segmentos $[x_i;\ x_{i+1}]$ un punto arbitrario ξ_i:

$$x_i \leqslant \xi_i \leqslant x_{i+1},$$

y hagamos una suma llamada *suma integral*:

$$S = \sum_{i=0}^{n-1} f(\xi_i)\cdot\Delta x_i, \qquad (1)$$

que corresponde a la partición dada del segmento $[a;\ b]$. La suma (1) depende de la elección de los puntos de partición x_i del segmento $[a;\ b]$ y de la elección de los puntos ξ_i.

Construyamos las distintas particiones del segmento $[a;\ b]$, aumentando ilimitadamente el número de puntos de partición de tal manera, que las longitudes de todos los segmentos $[x_i;\ x_{i+1}]$ tiendan a cero. Si en este caso la sucesión de las sumas integrales respectivas tiende a un mismo límite, independientemente de la elección de los puntos $\xi_i \in [x_i;\ x_{i+1}]$, entonces se dice que la función $f(x)$ es *integrable* en el segmento $[a;\ b]$; el límite indicado de las sumas de integrales se llama *integral definida* de la función $f(x)$ y se designa

$$\int_a^b f(x)\,dx.$$

Los números a y b se llaman respectivamente *límites inferior y superior de integración*.

Para la integrabilidad de la función es necesario que ella sea acotada en el segmento $[a; b]$.

Serán funciones integrables, por ejemplo, las funciones continuas y las funciones acotadas que tienen en el segmento $[a; b]$ un número finito de puntos de discontinuidad.

3.3. Propiedades de las integrales definidas.

1) Si la función $f(x)$ es integrable en el segmento $[a; b]$, entonces también la función $k \cdot f(x)$ (k es constante) es integrable en este segmento, al mismo tiempo

$$\int_a^b k \cdot f(x)\,dx = k \cdot \int_a^b f(x)\,dx,$$

es decir, el factor constante se puede sacar del signo de la integral.

2) Si las funciones $f(x)$ y $g(x)$ son integrables en el segmento $[a; b]$, entonces también la función $f(x) \pm g(x)$ es integrable en el segmento, además

$$\int_a^b [f(x) \pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$$

3) En la reordenación de los límites de integración la integral cambia el signo por el contrario:

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

4) Si la función $f(x)$ es integrable en el segmento mayor $[a; b]$, $[a; c]$ y $[c; b]$, entonces ella es integrable y en los otros dos segmentos. Para esto

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

30—01477

5) Si la función no negativa $f(x)$ es integrable en el segmento $[a; b]$, entonces

$$\int_a^b f(x)\,dx \geqslant 0.$$

6) Si las funciones $f(x)$ y $g(x)$ son integrables en el segmento $[a; b]$ y para todos $x \in [a; b]$

$$f(x) \leqslant g(x),$$

entonces

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx.$$

7) Si la función es integrable en el segmento $[a; b]$, donde $a < b$, y en todo el segmento es válida la desigualdad

$$m \leqslant f(x) \leqslant M$$

entonces

$$m\cdot(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M\cdot(b-a).$$

3.4. La integral definida como función del límite superior. Sea la función $f(x)$ continua (y, por consiguiente, integrable) en el segmento $[a; b]$. Entonces según la propiedad de la integral definida ella es integrable también en el segmento $[a; x]$; donde x es un número cualquiera del segmento $[a; b]$. La integral

$$\int_a^x f(t)\,dt$$

puede considerarse como una función del límite superior variable:

$$\Phi(x) = \int_a^x f(t)\,dt.$$

La función $\Phi(x)$ es una función continua, y su derivada en cualquier punto del intervalo $(a; b)$ es la función dada

$f(x)$:

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

Con otras palabras, la derivada de la integral

$$\int_a^x f(t)\,dt$$

del límite superior variable x, en todas partes del intervalo abierto $(a; b)$ es igual al valor $f(x)$ de la función subintegral en este intervalo.

De esta manera, la integral con un límite superior variable representa una primitiva de la función subintegral.

3.5. Fórmula principal del cálculo integral. Sea la función $f(x)$ integrable en el segmento $[a; b]$ y $F(x)$, la primitiva de la función $f(x)$. La integral definida de la función $f(x)$ en el segmento $[a; b]$ y su primitiva $F(x)$ se relacionan mediante la igualdad

$$\int_a^b f(x)\,dx = F(b) - F(a). \tag{2}$$

La diferencia que se encuentra en el segundo miembro de la fórmula suele designarse

$$F(x)\Big|_a^b,$$

y la fórmula (2) se escribe en la forma

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b.$$

La fórmula (2) se llama *fórmula principal* del cálculo integral. Esta fórmula en muchos casos permite reducir el cálculo de la integral definida al cálculo de los valores de la primitiva.

Ejemplo. Calcular la integral definida

$$\int_0^{\pi/2} \operatorname{sen} x\,dx.$$

La primitiva de la función sen x será la función $-\cos x$. Según la fórmula principal del cálculo integral

$$\int_0^{\pi/2} \operatorname{sen} x\, dx = -\cos x \Big|_0^{\pi/2} = -\cos\frac{\pi}{2} -$$
$$-(-\cos 0) = -0 + 1 = 1.$$

Análogamente por la fórmula principal del cálculo integral se calculan las integrales siguientes:

$$\int_0^1 x^2\, dx = \frac{1}{3} x^3 \Big|_0^1 = \frac{1}{3} - 0 = \frac{1}{3},$$

$$\int_{1/2}^5 \frac{1}{x}\, dx = \ln x \Big|_{1/2}^5 = \ln 5 - \ln\frac{1}{2} = \ln 10,$$

$$\int_{-1}^1 e^x\, dx = e^x \Big|_{-1}^1 = e^1 - e^{-1} = e - \frac{1}{e}.$$

§ 4. Ecuaciones diferenciales

4.1. Concepto de dependencia funcional entre varias variables. En el § 1 fue introducido el concepto de función numérica de una variable. Sin embargo, en distintas secciones de las matemáticas y de la física muy a menudo se encuentran relaciones, en las cuales participan varios valores variables. Así, por ejemplo, del curso de geometría es conocido que el volumen de un cilindro V se calcula mediante la fórmula

$$V = \pi R^2 H,$$

donde R es el radio de la base del cilindro y H es su altura. Esta fórmula da un ejemplo más simple de una función numérica de dos variables: el volumen depende tanto del radio de la base del cilindro, como de su altura.

Del curso de física es conocida la fórmula de Klaiperon

$$pV = RT$$

que relaciona tres características físicas del estado del gas: la presión p, el volumen específico V y la temperatura absoluta T (R es la constante universal de los gases). Representando dos de los tres valores indicados, se puede obtener un valor bastante determinado del tercer valor. De esta manera, en la aplicación de esta fórmula de nuevo nos encontramos con el concepto de función de varias variables.

La variable z se llama *función* numérica *de dos variables* x e y, la cual se da en el conjunto de pares ordenados de los números M, si a cada par ordenado de números $(x; y) \in M$, según cierta regla (ley), se le asigna respectivamente el valor definido de la variable z.

El conjunto M de los pares ordenados de los números $(x; y)$ se llama *campo de definición* de la función z, y el conjunto de todos los valores posibles de la variable z se llama *conjunto de valores* de una función.

El hecho de que la variable z es función de dos variables x e y suele escribirse como sigue:

$$z = f(x, y).$$

Mientras que para la función numérica de una variable el campo de definición es el conjunto de los puntos del eje numérico, el campo de definición de una función de dos variables representa cierto conjunto de puntos del plano numérico.

Al igual que la función $y = f(x)$ fue ilustrada geométricamente por su gráfica en el sistema de coordenadas cartesianas rectangulares Oxy, la función de dos variables $z = f(x, y)$ en el sistema espacial de coordenadas cartesianas rectangulares $Oxyz$ representa un conjunto de puntos, el cual también es una gráfica espacial de la función $z = f(x, y)$. La gráfica espacial de la función $z = f(x, y)$ suele llamarse *superficie*, y la igualdad $z = f(x, y)$, *ecuación de la superficie*.

El ejemplo más simple de tal superficie es un plano en el espacio, el cual en el sistema de coordenadas cartesianas rectangulares se define por la ecuación

$$Ax + By + Cz + D = 0.$$

Si $C \neq 0$, entonces la ecuación puede ser escrita en la forma

$$z = -\frac{A}{C}x - \frac{B}{C}y - \frac{D}{C}.$$

Todos los conceptos principales de la teoría de las funciones de una variable, tales, como el concepto de límite de una función, concepto de la continuidad de una función, concepto de diferenciabilidad, etc., se generalizan también para el caso de las funciones de dos variables independientes.

4.2. Concepto de ecuación diferencial ordinaria. Las ecuaciones, en las cuales las funciones que entran en la ecuación junto con sus derivadas son incógnitas, se llaman *ecuaciones diferenciales ordinarias*. Las ecuaciones siguientes

$$f'(x) + f^2(x) = 0, \quad f''(x) + 2 \cdot f'(x) = x - 1, \quad f''(x) \cdot f(x) = \frac{1}{x}$$

son ejemplos de ecuaciones diferenciales.

El orden de la derivada superior que entra en la ecuación diferencial se llama *orden* de la ecuación diferencial. Así, la ecuación diferencial

$$f'(x) - f^3(x) - x = 0$$

será una ecuación diferencial de primer orden; la ecuación diferencial

$$f'''(x) - x^2 = 0$$

es una ecuación diferencial del tercer orden.

Se llama *solución* de una ecuación diferencial tal función de la variable independiente x, definida en el intervalo $(a; b)$, para la sustitución de la cual en la ecuación diferencial, esta última se invierte a una identidad para todas $x \in (a; b)$. Seguidamente serán expuestos ciertos conocimientos sobre las ecuaciones diferenciales de primero y segundo orden, y también se resuelven algunas ecuaciones diferenciales simples de primero y segundo orden. Vamos a resolver estas ecuaciones mediante «la elección», es decir, vamos a suponer que cierta función concreta es la solución de la ecuación diferencial dada, y después, mediante la sustitución de esta función en la ecuación, convencernos de que esto es realmente así. Para tal procedimiento de «hallar» la solución de la ecuación diferencial surgen, naturalmente, dos preguntas. En primer lugar ¿cómo hemos adivinado que es precisamente esta función concreta la solución de la ecuación dada? En segundo lugar ¿de dónde sabemos que la ecuación diferencial dada no tiene otras soluciones? A la segunda pregunta se puede contestar del modo siguiente: las ecuaciones que vamos a examinar satisfacen el, así llamado, teorema de existencia y unicidad de la solución de la ecuación diferencial y por eso no hay otras soluciones, distintas de las halladas. Contestar a la primera pregunta es mucho más difícil: para ciertos tipos particulares de ecuaciones diferenciales existe un sistema bien elaborado de obtención de las soluciones; las soluciones de algunas otras ecuaciones diferenciales se hallan mediante ciertos procedimientos especiales; las soluciones de muchas ecuaciones diferenciales no son conocidas todavía, y para ellas aún no está demostrado el teorema de existencia y unicidad de la solución. Nuestro procedimiento de «elección» de las soluciones se aplica para las ecuaciones diferenciales con un sistema bien elaborado de obtención de sus soluciones.

4.3. Ecuación diferencial de primer orden. Ejemplos de problemas que se describen por las ecuaciones diferenciales de primer orden. Se llama ecuación diferencial de *primer orden* a la que contiene una variable independiente x, una función incógnita $y(x)$, y también la primera derivada de la función incógnita $y'(x)$:

$$F(x, y, y') = 0, \tag{1}$$

donde F es la función dada de tres variables indicadas. La función F puede ser representada no para todos los valores de su argumento, por eso se dice del campo de representación de la función F como de un conjunto de puntos del espacio de coordenadas de tres variables x, y, y'.

Se llama *solución* de la ecuación diferencial (1) tal función $\varphi(x)$ de la variable independiente x, definida en cierto intervalo $(a; b)$, para cuya sustitución por y en la ecuación (1), la ecuación indicada se invierte en identidad en todo el intervalo $(a; b)$.

Es evidente, que la sustitución $y = \varphi(x)$ en la ecuación (1) es posible sólo cuando la función $\varphi(x)$ en todo el intervalo $(a; b)$ es diferenciable. Para que la sustitución $y = \varphi(x)$ en la ecuación (1) sea posible, es necesario también que para un valor arbitrario de la variable $x \in (a; b)$ el punto con las coordenadas $(x; \varphi(x)\ \varphi'(x))$ pertenezca al campo de definición de la función F.

La ecuación (1) relaciona tres valores variables x, y, y'. En aquellos casos, cuando ella define la variable y' como una función de las variables independientes x, y:

$$y' = f(x, y), \tag{2}$$

a ecuación diferencial se llama *despejada* respecto a la derivada.

Para la ecuación diferencial (2) en ciertas, bastante lógicas, suposiciones sobre el tipo de campo de definición de la función $f(x; y)$ resulta ser válido el teorema de existencia y unicidad de la solución:

1) Para cualquier punto $(x_0; y_0)$ del campo de definición de la función $f(x; y)$ se halla la solución $y = \varphi(x)$ de la ecuación (2) que satisface la condición

$$\varphi(x_0) = y_0.$$

2) Si dos soluciones $y = \varphi(x)$ e $y = \psi(x)$ de las ecuaciones (2) coinciden por lo menos para un valor $x = x_0$, es decir, si

$$\varphi(x_0) = \psi(x_0),$$

entonces estas soluciones son idénticamente iguales para todos los valores de la variable independiente x, para los cuales ambas están definidas.

Los números x_0, y_0 se llaman *valores iniciales* para la solución $y = \varphi(x)$, y la relación $\varphi(x_0) = y_0$ se llama *condición inicial* para esta solución. Se dice también que la solución $y = \varphi(x)$ satisface la condición inicial $y_0 = \varphi(x_0)$ o que ella tiene los datos iniciales x_0, y_0.

Resolvamos la ecuación diferencial de primer orden

$$y' = \alpha \cdot y, \tag{3}$$

donde α es un número real, $y = y(x)$; el campo de definición de la función que se encuentra en el segundo miembro es todo el plano numérico Oxy, y la ecuación (3) satisface las condiciones del teorema de existencia y unicidad de la solución. Mediante la sustitución directa se puede verificar, que cada función

$$y = c \cdot e^{\alpha x}, \tag{4}$$

donde c, un número arbitrario, es la solución de la ecuación (3), y del teorema de existencia y unicidad de la solución se deduce que cualquier solución de la ecuación (3) es representable en la forma (4).

1. *Proceso de desintegración radiactiva de la sustancia*. Designemos por $m(t)$ la cantidad (masa) de sustancia radiactiva, donde t es el tiempo. De las consideraciones físicas se desprende que, si no hay condiciones para el surgimiento de la reacción en cadena, la velocidad de desintegración, es decir, la derivada (el cambio instantáneo de la masa de la sustancia) $m'(t)$ es proporcional a la cantidad que se queda de la sustancia radiactiva no desintegrada

$$m'(t) = -\beta \cdot m(t). \tag{5}$$

Aquí β es un coeficiente constante positivo de la proporcionalidad, que depende de las propiedades de la sustancia radiactiva, y el signo menos indica que con el tiempo la masa de la sustancia radiactiva disminuye, convirtiéndose en radiación ($m' < 0$ para cualquier t). La ecuación (5) con una precisión hasta las designaciones coincide con la ecuación (3), y, por consiguiente, su solución general tiene la forma

$$m(t) = ce^{-\beta t}.$$

Para la definición de la constante c es suficiente indicar algunas condiciones iniciales. Por ejemplo, si se sabe, que en el momento de tiempo $t = t_0$ la masa de la sustancia fue igual a m_0, entonces la constante c es igual a

$$c = m_0 e^{\beta t_0},$$

y la solución de la ecuación (5) con las condiciones iniciales (t_0, m_0) tiene la forma

$$m(t) = m_0 e^{-\beta(t-t_0)}.$$

En la ecuación (5) la velocidad de desintegración se caracteriza por el valor β. Muy a menudo la velocidad de desintegración se caracteriza no por el valor β, sino por el así llamado «período de semidesintegración», es decir, por el tiempo, durante el cual la masa de la sustancia radiactiva disminuye en dos veces. Designando el período de semidesintegración por T_0 obtendremos, que el coeficiente β y el período de semidesintegración T_0 se relacionan por la igualdad

$$T_0 = \frac{1}{\beta} \ln 2.$$

2. *Problema del calentamiento de un cuerpo.* Examinemos un problema sobre el calentamiento (enfriamiento) de un cuerpo que tiene en un momento de tiempo $t = 0$ cierta temperatura inicial T_1 y que está sumergido en un medio con una temperatura constante θ. En ciertos casos se puede suponer que la velocidad del cambio de la temperatura de un cuerpo $T(t)$ es proporcional a la diferencia de temperaturas, es decir, satisface la ecuación diferencial

$$T'(k) = -k(T - \theta), \tag{6}$$

donde k es cierta constante positiva. El signo menos indica que para $T_0 > \theta$ el cuerpo se enfría, y para $T_1 < \theta$ se calienta.

Observemos que una simulación del proceso físico indicado mediante la ecuación (6) es bastante aproximado y se basa en las suposiciones siguientes: se considera, que en cualquier momento de tiempo el cuerpo está calentado uniformemente en todo el volumen; se considera, que la temperatura del medio es constante y es igual a la temperatura inicial θ, a pesar de que el cuerpo calienta el medio. Si tomamos que el proceso dado se describe por la ecuación (6), entonces la solución de la ecuación se obtiene análogamente que en el ejemplo anterior y tiene la forma

$$T(t) = (T_1 - \theta) \cdot e^{-kt} + \theta.$$

4.4 Ecuación diferencial de segundo orden. Ecuación de las oscilaciones armónicas. Se llama ecuación diferencial ordinaria de *segundo orden* la ecuación que contiene una variable independiente x, una función incógnita $y(x)$, así como la primera y segunda derivadas y' e y'':

$$F(x, y, y', y'') = 0 \tag{7}$$

donde F es una función dada de las variables x, y, y', y''. En aquellos casos, cuando la ecuación (7) representa la variable y'' como una función de las variables x, y e y':

$$y'' = f(x, y, y'), \tag{8}$$

la ecuación diferencial se llama *resuelta* respecto a la derivada mayor.

Se llama *solución* de la ecuación diferencial (8) la función continua $\varphi(x)$, definida en cierto intervalo $(a; b)$, en la cual al sustituir y en la ecuación (8), esta última se convierte en identidad para todas $x \in (a; b)$. La sustitución $y = \varphi(x)$ en la ecuación (8) es posible sólo cuando la función $\varphi(x)$ (para todos $x \in (a; b)$) tiene las derivadas hasta el segundo orden inclusive. Para que la sustitución $y = \varphi(x)$ en la ecuación (8) sea posible, es necesario también que el punto que tiene las coordenadas

$$x, \ \varphi(x), \ \varphi'(x)$$

pertenezca al campo de definición de la función f para todos $x \in (a; b)$.

Para la ecuación diferencial de segundo orden (8), resuelta respecto a la derivada, para ciertas suposiciones adicionales*) sobre el tipo de función $f(x, y, y')$ es válido el teorema de existencia y unicidad de la solución, es decir, para cada punto $(x_0; y_0; y')$ existe una solución $y = \varphi(x)$ de la ecuación (8) que satisface las *condiciones iniciales*

$$\varphi(x_0) = y_0, \quad \varphi'(x_0) = y',$$

o, como también se dice, una solución con los *datos iniciales*

$$x_0, \ y_0, \ y'_0,$$

además, esta solución es única.

Ecuación de oscilaciones armónicas. Examinemos una ecuación diferencial de segundo orden que describe, en particular, el proceso de oscilaciones mecánicas, o sea, la ecuación de las oscilaciones armónicas

$$y'' + \omega^2 y = 0, \tag{9}$$

donde ω es una constante dada, e y se considera la función del tiempo t. Es fácil convencerse de que la función

$$y = A\cos(\omega t + \theta), \tag{10}$$

---

*) No vamos a citar aquí estas condiciones y la formulación completa del teorema de existencia y unicidad de la solución, puesto que las ecuaciones diferenciales de segundo orden, las cuales se examinan en lo posterior, satisfacen las condiciones de existencia y unicidad de la solución.

donde A y θ, constantes arbitrarias, son la solución de la ecuación diferencial (9). La constante A se llama a menudo *amplitud de oscilación*, ω se llama *frecuencia* de oscilación, y θ, *fase inicial* de oscilación.

La fórmula (10) da el conjunto de todas las soluciones de la ecuación diferencial (9). Para formar de un conjunto infinito de soluciones (10) una única solución, es necesario dar las condiciones iniciales siguientes:

$$y(t_0) = y_0, \quad y'(t_0) = y_0',$$

es decir, representar el valor de la función $y(t)$ de su derivada $y'(t)$ en cierto momento de tiempo t_0, el cual se llama momento inicial de tiempo. La representación de los valores y_0 e y' permite hallar unívocamente las constantes A y θ, formando de un conjunto infinito de soluciones (10) una única solución.

Ejemplos. 1. Examinemos el movimiento de un punto material de la masa m que se desliza por la recta l bajo la acción de la fuerza F, que atrae un punto material al punto O situado en la misma recta.

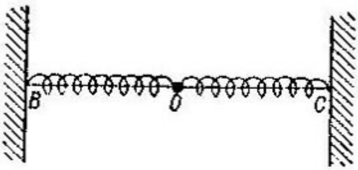


Fig. 9.3.

Sea el valor de la fuerza F proporcional al valor de la desviación del punto material del punto O con cierto coeficiente de proporcionalidad k. (En la práctica tal fuerza F puede ser realizada mediante la construcción que consta de un globo y de dos espirales (fig. 9.3).) Aquí por el punto O (origen de coordenadas) se toma el punto medio del segmento BC. La desviación de un punto material del punto medio la vamos a designar con $x(t)$, donde t es el tiempo. Entonces, en vigor de la segunda ley de Newton la ecuación del movimiento de un punto material tendrá la forma

$$mx'' = -kx, \tag{11}$$

donde $-kx$ es una fuerza que actúa del lado de la espiral sobre el punto material, y k es el coeficiente de elasticidad de la espiral.

Según la fórmula (10) la solución de la ecuación (11) tiene la forma

$$x(t) = A\cos\left(\sqrt{\frac{k}{m}}\,t + \varphi_0\right),$$

de donde se deduce que la frecuencia de oscilaciones del punto material

$$\omega = \sqrt{\frac{k}{m}}$$

se define sólo por su masa y la elasticidad de la espiral k. Dando las condiciones iniciales del problema, es decir, la posición de un punto material en cierto momento de tiempo t_0 y su velocidad en el momento inicial de tiempo t_0:

$$x(t_0) \quad \text{y} \quad x'(t_0),$$

se puede hallar los valores de las constantes A y θ, por lo tanto se define la ley del movimiento de un punto material.

2. Hagamos y resolvamos aproximadamente la ecuación del péndulo matemático. El péndulo matemático representa un punto de la masa m, el cual bajo la acción de la fuerza de gravedad se mueve

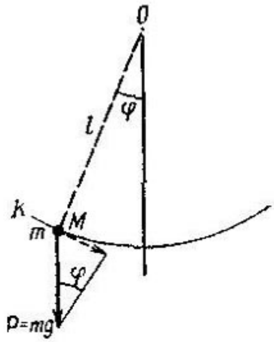


Fig. 9.4.

por el arco de la circunferencia K de radio l que se encuentra en el plano vertical (fig. 9.4). El valor l se llama longitud del péndulo. En el plano de la circunferencia K introducimos el sistema polar de coordenadas con centro en el centro de la circunferencia y el eje polar, dirigido perpendicularmente hacia abajo. El ángulo variable que caracteriza la desviación del péndulo de la perpendicular se designa con $\varphi = \varphi(t)$. El punto M se encuentra en el campo de la fuerza de gravedad $P = mg$, dirigida perpendicularmente hacia abajo. La componente de la fuerza de gravedad, dirigida por el radio de la circunferencia, se equilibra por la reacción de conexión (de la circunferencia o del hilo que hace al punto moverse por la circunferencia); la componente, dirigida por la tangente a la circunferencia en el punto M, es igual a mg sen φ. Es conocido de física que la aceleración lineal del punto M se relaciona con la aceleración angular φ'' por la fórmula

$$a = l\varphi''.$$

El movimiento de un punto material se describe por la segunda ley de Newton:

$$ml\varphi'' = -mg \operatorname{sen} \varphi. \tag{12}$$

Suponiendo, que las oscilaciones que realiza un punto material son pequeñas (es decir, que el ángulo de desviación $\varphi(t)$ es pequeño) se puede escribir aproximadamente

$$\operatorname{sen}\varphi \approx \varphi.$$

En este caso la ecuación diferencial (12) obtiene la forma

$$ml\varphi'' = -mg\varphi. \tag{13}$$

La función

$$\varphi(t) = A\cos\left(\sqrt{\frac{g}{l}}\cdot t + \varphi_0\right)$$

es la solución de la ecuación diferencial (13). La frecuencia de «las oscilaciones pequeñas» del péndulo matemático se define por la aceleración de la caída libre g y por la longitud del péndulo l:

$$\omega = \sqrt{\frac{g}{l}}.$$

Para el cálculo de las constantes A y φ_0 es necesario dar la posición y la velocidad del punto M en cierto momento de tiempo t_0, después de lo cual la ley de movimiento del péndulo matemático se hace completamente definida.

CAPÍTULO 10

Funciones elementales

Son funciones elementales las funciones racionales*), potenciales, exponenciales y logarítmicas, así como las funciones trigonométricas y las trigonométricas inversas. Además, también pertenecen a la clase de funciones elementales, las funciones compuestas, formadas de las funciones elementales enumeradas más arriba.

Las funciones exponencial y logarítmica, las funciones trigonométricas de argumento numérico y las funciones trigonométricas inversas, como también las funciones compuestas, formadas mediante estas funciones, se llaman funciones trascendentes elementales.

§ 1. Investigación de funciones

1.1. Función constante. Supongamos que la función está definida y es diferenciable en el intervalo $(a; b)$. Para que la función en el intervalo $(a; b)$ sea constante, es necesario y suficiente que su derivada sea igual a cero en todo el intervalo $(a; b)$:

$$f'(x) = 0.$$

1.2. Condición de monotonía de una función. Sea la función $f(x)$ definida y diferenciable en el intervalo $(a; b)$. Para que la función sea creciente en el intervalo $(a; b)$, es suficiente que

$$f'(x) > 0 \text{ para todos } x \in (a; b).$$

---

*) La función que tiene la forma $P(x)/Q(x)$, donde $P(x)$ y $Q(x)$ son polinomios, se llama *función racional.*

Para el decrecimiento de la función es suficiente que

$$f'(x) < 0 \text{ para todos } x \in (a;\ b).$$

La condición formulada no es una condición necesaria. Por ejemplo, la función $f(x) = x^3$ es creciente, sin embargo, en el punto $x = 0$ su derivada es igual a cero.

1.3. Máximo y mínimo de una función. Se dice que la función $f(x)$ tiene en el punto x_0 el *máximo* (o el *mínimo*), si se halla tal δ-entorno del punto x_0, perteneciente al campo de definición de la función, que para todos $x \neq x_0$, pertenecientes al intervalo $(x_0 - \delta;\ x_0 + \delta)$, se cumple la desigualdad

$$f(x) < f(x_0) \text{ (respectivamente } f(x) > f(x_0)).$$

En la fig. 10.1 se muestra una función, definida en el intervalo $[a;\ b]$, que tiene en los puntos x_1 y x_3 los máximos, y en los puntos x_2 y x_4, los mínimos.

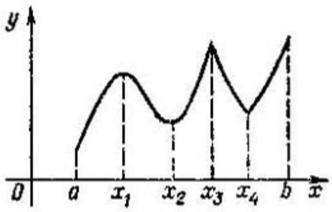


Fig. 10.1.

Los puntos del máximo y del mínimo se llaman puntos del *extremo* de una función, y los valores de la función en estos puntos se llaman *valores extremales* de una función.

Observemos que los puntos a y b no son puntos del extremo de una función, puesto que en cualquiera tanto como se quiera de pequeño entorno de estos puntos se hallan puntos, no pertenecientes al campo de definición $f(x)$.

Condición necesaria de existencia del extremo de una función. Sea la función $f(x)$ diferenciable en el intervalo $(a;\ b)$. Si en cierto punto $x_0 \in (a;\ b)$ la función $f(x)$ tiene extremo, entonces en este punto la derivada es igual a cero:

$$f'(x_0) = 0.$$

Condición suficiente de existencia del extremo de una función. Sea la función $f(x)$ definida y continua en el intervalo $(a;\ b)$, y diferenciable en todo el intervalo $(a;\ b)$ (a excepción, puede ser, de un número finito de puntos).

Los puntos, pertenecientes al intervalo $(a;\ b)$, en los cuales la derivada de la función es igual a cero o no existe, se llaman puntos *críticos* de la función $f(x)$.

Sea el punto $x_0 \in (a; b)$ un punto crítico de la función $f(x)$ y sea que en cierto δ-entorno del punto x_0 la derivada $f'(x)$ a la izquierda y a la derecha de x_0 conserva el signo (posiblemente en los distintos lados, distinto).

Entonces pueden tener lugar los tres casos siguientes:

1) $f'(x) > 0$ para $x < x_0$ y $f'(x) < 0$ para $x > x_0$, es decir, la derivada $f'(x)$ al pasar por el punto x_0 cambia el signo más por el menos. En este caso en el intervalo $(x_0 - \delta; x_0)$ la función $f(x)$ crece, y en el intervalo $(x_0; x_0 + \delta)$ decrece, así que el valor $f(x_0)$ es el mayor en el intervalo $(x_0 - \delta; x_0 + \delta)$, es decir, en el punto x_0 la función tiene un máximo.

2) $f'(x) < 0$ para $x < x_0$ y $f'(x) > 0$ para $x > x_0$, es decir, la derivada $f'(x)$ al pasar por el punto x_0 cambia el signo menos por el más. En este caso en el punto x_0 la función tiene un mínimo.

3) $f'(x) > 0$ tanto para $x > x_0$, como también para $x < < x_0$, o bien $f'(x) < 0$ a la izquierda y a la derecha de x_0, es decir, al pasar por el punto x_0 la derivada $f'(x)$ no cambia el signo. Entonces la función o todo el tiempo crece, o todo el tiempo decrece, en cualquiera tanto como se quiera de pequeño entorno del punto x_0, de un lado, se hallan los puntos x, en los cuales $f(x) > f(x_0)$ y de otro lado se hallan los puntos x, en los cuales $f(x) < f(x_0)$, y en el punto x_0 no hay extremo.

Así pues, obtuvimos la primera *regla* para la verificación del punto crítico x_0 en el extremo: sustituyendo en la derivada $f'(x)$ primeramente $x < x_0$, y después $x > x_0$, determinamos el signo de la derivada en el entorno del punto x_0 a la izquierda y a la derecha del mismo, si para esto la derivada $f'(x)$ cambia el signo de más a menos, entonces en el puntos x_0 hay máximo y si cambia del menos al más, entonces hay mínimo; si el signo no cambia, entonces no hay extremo.

En la fig. 10.2 la derivada de la función $f_1(x)$ al pasar por el punto x_0 cambia el signo de menos a más y la función $f_1(x)$ en el punto x_0 tiene el mínimo, mientras que la función $f_2(x)$, cuya derivada al pasar por el punto x_0 cambia el signo de más a menos, tiene el máximo.

En la fig. 10.3. tanto la función $f_1(x)$, como la función $f_2(x)$ tienen en el punto x_0 una derivada que es igual a cero; al paso por el punto x_0 las derivadas de estas funciones no

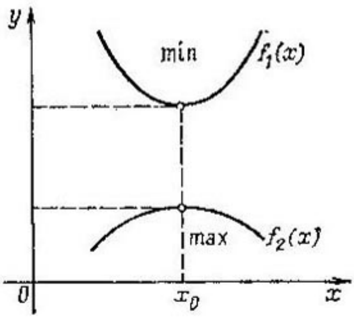


Fig. 10.2.

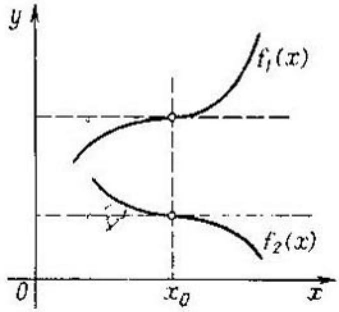


Fig. 10.3.

cambian el signo y en el punto x_0 las funciones $f_1(x)$ y $f_2(x)$ no tienen extremos.

En la fig. 10.4 y 10.5 se muestran las funciones que no tienen derivadas en el punto x_0. Conforme a las reglas de

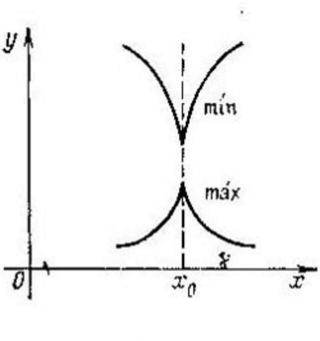


Fig. 10.4.

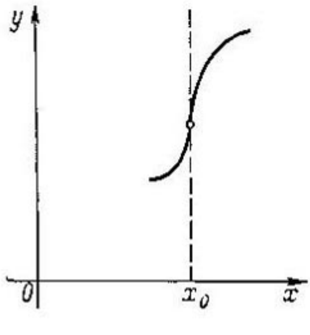


Fig. 10.5.

cambio de signo de las derivadas para estas funciones, ellas tienen mínimo y máximo respectivamente y no tienen extremo en el punto x_0.

Otra vez subrayamos que la condición suficiente de existencia del extremo de la función en el punto x_0 formulada más arriba es usable tanto en los casos, cuando la derivada en el punto x_0 se anula, como en los casos, cuando en este punto la derivada no existe.

Formulemos una condición suficiente más de existencia del extremo en un punto. El punto x_0 se llama *estacionario* de la función $f(x)$, si en este punto la función $f(x)$ es diferenciable y su derivada es igual a cero.

Sea x_0 un punto estacionario de la función $f(x)$ y en cierto entorno del punto x_0, incluyendo este mismo punto, la función $f(x)$ tiene la derivada, y en el punto x_0 existe también la segunda derivada $f''(x_0)$. Si la segunda derivada f'' en el punto x_0 es positiva, entonces la función $f(x)$ en el punto x_0 tiene mínimo; si la segunda derivada f'' en el punto x_0 es negativa, entonces la función tiene máximo.

La regla formulada tiene, en general, una aplicación bastante estrecha en comparación con la primera regla; esta regla no es aplicable para los puntos, en los cuales no existe la primera derivada. En aquellos casos, cuando la segunda derivada en el punto x_0 se anula, la regla tampoco da nada y la solución del problema de existencia en este punto del extremo depende de la conducta de las derivadas superiores.

1.4. Valores máximos y mínimos de una función. Sea que la función $f(x)$ está definida y es continua en un intervalo cerrado finito $[a; b]$.

Para hallar el valor máximo (mínimo) de la función es necesario hallar todos los valores máximos (mínimos) de la función en el intervalo $(a; b)$, elegir de ellos el valor máximo (mínimo) y compararlo con los valores de la función en los puntos a y b. El máximo (mínimo) de estos números es, precisamente, el valor máximo (mínimo) de la función $f(x)$ en el segmento $[a; b]$. Cuando se halla el valor máximo o mínimo de una función puede resultar que dentro del segmento $[a; b]$ la derivada existe en todos los puntos del segmento y en ningún punto del segmento se anula (es decir, faltan los puntos críticos de la función). Esto quiere decir que en el intervalo que examinamos la función crece o decrece y, por consiguiente, adquiere su valor máximo y mínimo en los extremos del segmento.

Examinemos dos problemas para hallar el valor máximo y mínimo de una función.

Ejemplos. 1. Hallar los valores máximo y mínimo de la función

$$f(x) = \frac{4}{3}x^3 - 4x$$

en el segmento $[0; 2]$.

Calculemos la derivada de la función dada

$$f'(x) = 4x^2 - 4.$$

Para todos $x \in (0; 2)$ la derivada existe; la derivada es igual al cero en los puntos $x = 1$, $x = -1$. En el intervalo $(0; 2)$ se encuentra sólo uno de estos puntos: $x = 1$. El valor de la función en este punto es igual a $-8/3$. En el punto $x = 1$ la derivada cambia el signo de menos a más, y por eso en el punto $x = 1$ tenemos el mínimo de la función $f(x)$. Los valores de la función en los extremos del intervalo son: $f(0) = 0$, $f(2) = 8/3$. Por consiguiente, el valor máximo de la función dada $f(x)$ se adquiere en el extremo derecho del intervalo en el punto 2, y el mínimo se adquiere dentro del intervalo en el punto $x = 1$.

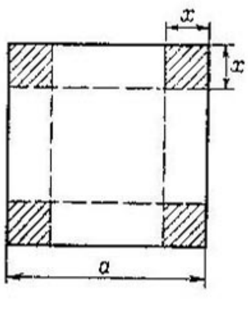


Fig. 10.6.

2. De una hoja cuadrada con un lado a, cortando por los ángulos cuadrados iguales y doblando los bordes (fig. 10.6), se compone una caja abierta rectangular. ¿Cómo obtener una caja de máximo volumen?

Designemos un lado del cuadrado cortado con x. Entonces el volumen de la caja V es igual a

$$x(a - 2x)^2,$$

además, x cambia en el intervalo $[0; a/2)$. La derivada de la función dada $V'(x) = (a - 2x)(a - 6x)$ en todo el intervalo $[0; a/2)$ se anula en el punto $x = a/6$. Calculemos la segunda derivada $V''(x) = 24x - 8a$. En el punto $x = a/6$ la segunda derivada es negativa, y, por consiguiente, en el punto $x = a/6$ tiene el máximo. En los extremos del segmento $[0; a/2]$ la función $V(x)$ se anula. De esta manera, el volumen de la caja es máximo para $x = a/6$.

1.5. Dirección de la concavidad de una curva. Sea la función $f(x)$ definida y continua en el intérvalo $(a; b)$, y que en el punto $x_0 \in (a; b)$ tiene la derivada. Entonces a la gráfica de la función dada en el punto $M_0(x_0; f(x_0))$ se puede trazar una tangente, cuya ecuación tiene la forma

$$y - y_0 = f'(x_0)(x - x_0).$$

Se dice que en el punto M_0 la curva (la gráfica de la función $y = f(x)$) está dirigida *con la concavidad* a un lado definido (hacia las y positivas o hacia las y negativas) de la tangente, si en el entorno suficientemente pequeño del punto M_0 todos los puntos de la curva están situados a un lado de la tangente (fig. 10.7 y 10.8). El punto M_0 se llama *punto de*

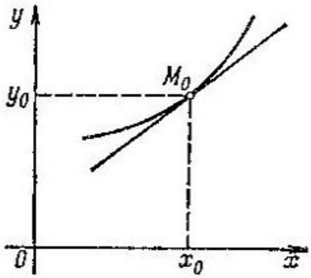


Fig. 10.7.

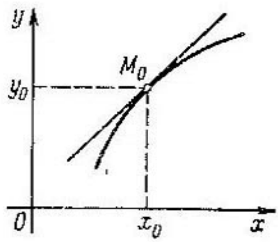


Fig. 10.8.

inflexión, si en su entorno suficientemente pequeño los puntos de la curva con las abscisas $x < x_0$ se encuentran a un lado de la tangente, y los puntos con las abscisas $x > x_0$ están situados al otro lado, es decir, en el punto M_0 la curva pasa de un lado de la tangente al otro (fig. 10.9).

La dirección de la concavidad de la curva $y = f(x)$ en cierto punto x_0 se aclara mediante la regla siguiente:

1) Si en cierto δ-entorno del punto x_0 la segunda derivada $f''(x)$ conserva el signo más (tanto a la izquierda, como a la derecha del punto x_0), entonces, la curva en el punto x_0 está dirigida con la concavidad hacia las y positivas (fig. 10.7).

2) Si en cierto δ-entorno del punto x_0 la segunda derivada conserva el signo menos (tanto a la izquierda, como a la derecha del punto x_0), entonces la curva en el punto x_0 está dirigida con la concavidad hacia las y negativas (fig. 10.8).

3) Si la segunda derivada cambia el signo al pasar por el punto x_0, entonces en el punto x_0 la gráfica de la función $f(x_0)$ tiene inflexión (fig. 10.9).

De esta manera, si nos limitamos a un δ-entorno suficientemente pequeño del punto x_0, el punto de inflexión M_0 es como si separara los puntos, donde la concavidad de la curva está dirigida hacia las y

positivas, de los puntos con la dirección de la concavidad hacia las y negativas.

La regla formulada es la condición suficiente (pero no necesaria) de inflexión de la curva. Así pues, la función $f(x) = x^4$ en el punto

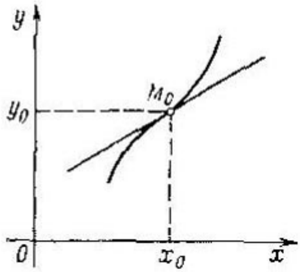


Fig. 10.9.

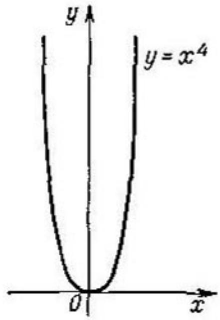


Fig. 10.10.

$x = 0$ tiene la segunda derivada que es igual a cero, pero no obstante en este punto (y en general para todos $x \in R$) la curva está dirigida con la concavidad hacia las y positivas (fig. 10.10).

Ejemplo. Examinemos la función $f(x) = \operatorname{sen} x$. La segunda derivada de esta función:

$$f''(x) = -\operatorname{sen} x.$$

En los intervalos $(2\pi k;\ \pi + 2\pi k)$ sen x conserva el signo más, $f''(x) = -\operatorname{sen} x < 0$ y la sinusoide está dirigida con la concavidad hacia

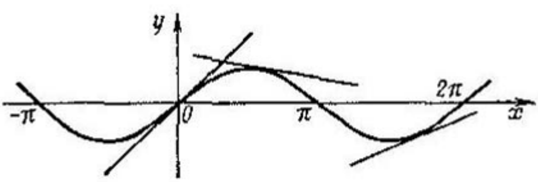


Fig. 10.11.

las y negativas. En los intervalos $(\pi + 2\pi k;\ 2\pi + 2\pi k)$ sen $x < 0$ también la sinusoide está dirigida con la concavidad hacia las y positivas. Para $x = \pi k$ la segunda derivada se anula, cambiando en este caso el signo, en estos puntos la gráfica de la función $y = \operatorname{sen} x$ tiene inflexión (fig. 10.11).

§ 2. Construcción de la gráfica de una función

La investigación de una función y la construcción de su gráfica se compone de los puntos siguientes:

1) Se halla el campo de definición de la función.

2) Se aclara, si la función dada es periódica o no.

3) Se aclara, si la función dada es par (o impar). Si la función resulta ser par, entonces su gráfica es simétrica respecto al eje Oy del sistema de coordenadas cartesianas rectangulares. Si la función es impar, entonces su gráfica es simétricamente central respecto al origen de coordenadas.

4) En el caso, cuando el campo de definición de la función representa bien toda la recta numérica o bien contiene los rayos $[a; +\infty)$ o $(-\infty; b]$, se aclara la conducta de la función en el infinito, calculando los límites respectivos

$$\lim_{x\to\infty} f(x), \quad \lim_{x\to-\infty} f(x).$$

5) Se halla la intersección de la curva con el eje Oy, calculando $f(0)$, y con el eje Ox, resolviendo la ecuación $f(x) = 0$.

6) Se hallan los puntos del máximo y del mínimo, determinando las regiones de crecimiento y decrecimiento de la función.

7) Se halla la región de cambio de la función $E(f)$.

8) Se hallan los puntos de inflexión de la función, determinando las regiones, donde la curva está dirigida con la concavidad hacia las y positivas o negativas, al mismo tiempo, en los puntos de inflexión se calcula el ángulo de inclinación de la tangente.

Para la investigación de las funciones concretas no es obligatorio seguir el orden indicado de investigación de las funciones. Se puede omitir la aclaración de unas u otras propiedades, si ellas son bastante evidentes. Así, para la investigación de las funciones racionales se suele omitir el estudio de su periodicidad debido a que de todas las funciones elementales sólo las funciones trigonométricas de argumento numérico tienen propiedad de periodicidad.

Ejemplo. Examinar las propiedades de la función

$$f(x) = \frac{2x^2+3x-4}{x^2}$$

y construir su gráfica,

1) El campo de definición de la función es toda la recta numérica, a excepción del punto $x = 0$.

2) La función no es ni par ni impar, puesto que

$$f(x) \neq f(-x) \text{ y } f(x) \neq -f(-x).$$

3) Para aclarar la conducta de la función cuando el argumento crece indefinidamente (y decrece), calculemos los límites

$$\lim_{x \to +\infty} \frac{2x^2+3x-4}{x^2} = 2,$$

$$\lim_{x \to -\infty} \frac{2x^2+3x-4}{x^2} = 2.$$

4) La gráfica de la función dada no interseca el eje Oy. Los puntos de intersección de la gráfica de la función con el eje Ox se definen como raíces de la ecuación

$$\frac{2x^2+3x-4}{x^2} = 0,$$

las cuales son las siguientes:

$$x_1 = \frac{-3+\sqrt{41}}{4}, \quad x_2 = \frac{-3-\sqrt{41}}{4}.$$

5) Puntos del máximo y del mínimo de una función. Calculemos la derivada de la función dada:

$$f'(x) = \frac{-3x^2+8x}{x^4}.$$

6) La región de cambio de la función (conjunto de valores de la función) es el intervalo

$$(-\infty;\ 41/16].$$

7) Calculemos la segunda derivada de la función:

$$f''(x) = \frac{6x-24}{x^4}.$$

La segunda derivada en el intervalo $(4;\ +\infty)$ es positiva, y, por consiguiente, en este intervalo la gráfica de la función está dirigida con la concavidad hacia las y positivas. En los y la gráfica de la función está dirigida con la concavidad hacia las y negativas. En el punto $x = 4$ la gráfica de la fun-

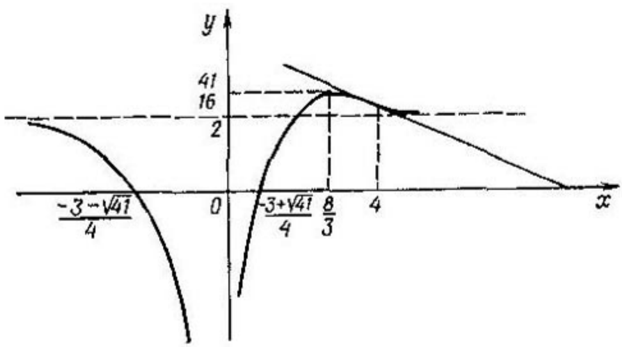


Fig. 10.12.

ción tiene inflexión. Sobre la base de la investigación realizada se puede dibujar la gráfica de la función dada (fig. 10.12).

§ 3. Transformaciones simples de la gráfica de una función

1) *Traslación de la gráfica paralelamente al eje de ordenadas.* La gráfica de la función $y = f(x) + a$ se obtiene de la gráfica de la función $y = f(x)$ mediante la traslación paralela $r = (0;\ a)$. Si el número a es positivo, entonces la gráfica de la función $y = f(x) + a$ se obtiene mediante el desplazamiento de la gráfica de la función $y = f(x)$ en a unidades hacia arriba; si a es negativo, en a unidades hacia abajo (fig. 10.13).

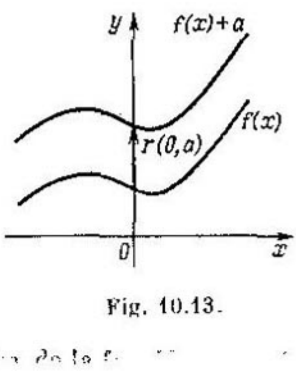


Fig. 10.13.

2) *Traslación de la gráfica paralelamente al eje de abscisas.* La gráfica de la función $y = f(x - a)$ se obtiene de la gráfica de la f

entonces la gráfica se traslada a la izquierda, si $a > 0$, a la derecha (fig. 10.14).

3) *Alargamientos y contracciones de las gráficas de las funciones.* Sea dada la gráfica de la función $y = f(x)$ y sea que el punto con las coordenadas $(x_0; y_0)$ pertenece a la gráfica de la función $y = f(x)$.

Examinemos la función $y = kf(x)$ (k es un número positivo). El punto de la gráfica de la función $y = kf(x)$ con la abscisa x_0 tiene la ordenada ky_0. Para $k > 1$ las ordenadas

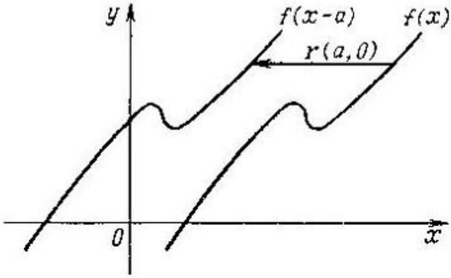


Fig. 10.14.

de los puntos de la gráfica de la función $y = kf(x)$ crecen en k veces en comparación con las ordenadas de los puntos de la gráfica de la función $y = f(x)$, que tienen la misma abscisa. En este caso se dice que la gráfica de la función $y = kf(x)$ se obtiene de la gráfica de la función $y = f(x)$ mediante el alargamiento del eje de abscisas en k veces.

Para $0 < k < 1$ las ordenadas de los puntos de la gráfica de la función $y = kf(x)$ decrecen en comparación con las ordenadas de los puntos de la gráfica de la función $y = f(x)$ que tienen la misma abscisa, y se dice que la gráfica de la función $y = kf(x)$ se obtiene de la gráfica de la función $y = f(x)$ mediante la contracción hacia el eje de abscisas en k veces.

Por ejemplo, la gráfica de la función $y = 2x^2$ se obtiene de la gráfica de la función $y = f(x)$ mediante el alargamiento del eje de abscisas en 2 veces, y la gráfica de la función $y = {}^1/_2x^2$ mediante la contracción hacia el eje de abscisas en 2 veces (fig. 10.15).

Examinemos ahora la función $y = f(kx)$ (k es un número positivo). El punto de la gráfica de la función $y = f(kx)$ con la ordenada y_0 tiene la abscisa x_0/k. Para $k > 1$ las abscisas de todos los puntos de la gráfica de la función $y = f(kx)$ decrecen en k veces en comparación con las abscisas de los puntos de la gráfica de la función $y = f(x)$, que tienen la misma ordenada. En este caso se dice que la gráfica

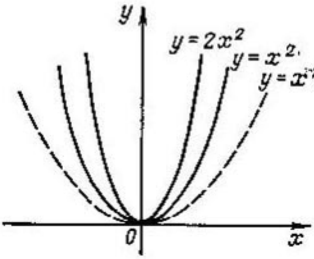


Fig. 10.15.

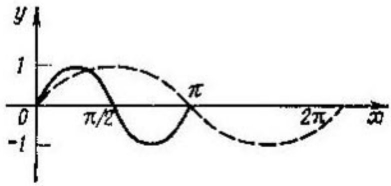


Fig. 10.16.

de la función $y = f(kx)$ se obtiene de la gráfica de la función $y = f(x)$ mediante la contracción hacia el eje de ordenadas en k veces.

Para $0 < k < 1$ las abscisas de los puntos de la gráfica de la función $y = f(kx)$ aumentan en $1/k$ veces en comparación con las abscisas de los puntos de la gráfica de la función $y = f(x)$, que tienen la misma ordenada. En este caso se dice que la gráfica de la función $y = f(kx)$ se obtiene de la gráfica de la función $y = f(x)$ mediante el alargamiento del eje de ordenadas en $1/k$ veces. Por ejemplo, la gráfica de la función $y = \operatorname{sen} 2x$ se obtiene de la gráfica de la función $y = \operatorname{sen} x$ mediante la contracción hacia el eje de ordenadas en 2 veces (fig. 10.16).

4) La gráfica de la función $y = -f(x)$ se obtiene de la gráfica de la función $y = f(x)$ mediante la reflexión simétrica respecto al eje Ox (fig. 10.17).

5) La gráfica de la función $y = f(-x)$ se obtiene de la gráfica de la función $y = f(x)$ mediante la reflexión simétrica respecto al eje Oy (fig. 10.18).

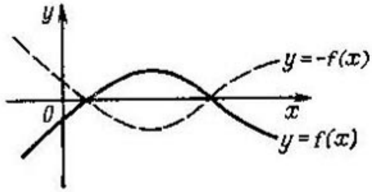


Fig. 10.17.

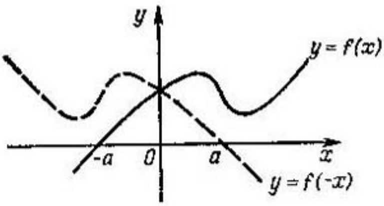


Fig. 10.18.

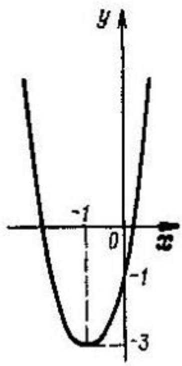
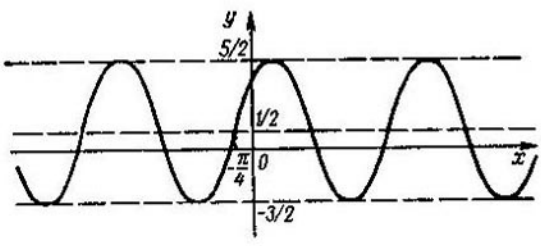


Fig. 10.19.



Ejemplos. 1. Mediante las transformaciones de las gráficas de las funciones obténgase de la gráfica de la función $y = x^2$ la gráfica de la función $y = 2x^2 + 4x - 1$.

Escribimos el trinomio de segundo grado $2x^2 + 4x - 1$ en la forma $2(x + 1)^2 - 3$. La gráfica del trinomio de segundo grado $y = 2(x + 1)^2 - 3$ se obtiene de la gráfica del trinomio de segundo grado $y = x^2$ como resultado de las transformaciones siguientes:

a) de la traslación paralela $\boldsymbol{r} = (-1; 0)$;

b) del alargamiento del eje de las abscisas en 2 veces;

c) de la traslación paralela $\boldsymbol{r} = (0; -3)$.

La gráfica de la función $y = 2x^2 + 4x - 1$ se muestra en la fig. 10.19.

2. Mediante las transformaciones de las gráficas de las funciones obténgase de la gráfica de la función $y = \operatorname{sen} x$ la gráfica de la función $y = 2 \operatorname{sen}\left(x + \frac{\pi}{4}\right) + {}^1/_2$.

La gráfica de la función $y = 2 \operatorname{sen}\left(x + \frac{\pi}{4}\right) + {}^1/_2$ se obtiene de la gráfica de la función $y = \operatorname{sen} x$ como resultado de las transformaciones siguientes:

a) de la traslación paralela $\boldsymbol{r} = (-\pi/4; 0)$;

b) del alargamiento del eje de las abscisas en 2 veces;

c) de la traslación paralela $\boldsymbol{r} = (0; {}^1/_2)$.

La gráfica de la función $y = 2 \operatorname{sen}\left(x + \frac{\pi}{4}\right) + {}^1/_2$ se muestra en la fig. 10.20.

§ 4. Función lineal

La función que tiene la forma

$$f(x) = kx + b,$$

donde k y b son ciertos números, se llama *función lineal*.

Propiedades de la función lineal.

1) El campo de definición de la función es el conjunto de todos los números reales $\mathbf{R}$.

2) La región de cambio (el conjunto de valores) de la función para $k \neq 0$ es el conjunto de todos los números reales. Para $k = 0$ el conjunto de los valores de la función cons-

3) Para $k \neq 0$ y $b \neq 0$ la función no es ni par ni impar. Para $k \neq 0$ (b es cualquiera) la función es par. Para $b = 0$ (k es cualquiera) la función es impar.

4) La función lineal $f(x) = kx + b$ es continua y diferenciable en todo el eje numérico; su derivada en cada punto es igual a k.

5) La función lineal no tiene extremo para ningún valor de k y b. Para $k \neq 0$ no hay puntos críticos. Para $k \neq 0$ cada punto es un punto crítico de la función.

6) Para $k > 0$ la función lineal crece para todos $x \in \mathbf{R}$. Para $k < 0$ la función lineal decrece para todos $x \in \mathbf{R}$. Para $k = 0$ la función lineal es constante.

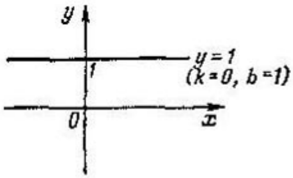


Fig. 10.21.

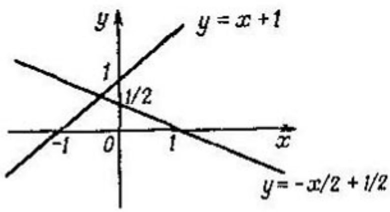


Fig. 10.22.

7) La gráfica de la función lineal $f(x) = kx + b$ interseca el eje Oy en el punto $y = b$. La gráfica de la función lineal $f(x) = kx + b$ para $k \neq 0$ interseca el eje Ox en el punto $x = -b/k$; para $k = 0$ la gráfica de la función lineal es paralela al eje Ox.

En las figs. 10.21, 10.22 se muestran las gráficas de las funciones lineales.

La función lineal $f(x) = kx + b$ puede tener como gráfica suya cualquier recta del plano de coordenadas Oxy, a excepción de las rectas perpendiculares.

Proporcionalidad directa. La variable y se llama variable proporcional directa x con el coeficiente de proporcionalidad k, si estas variables están relacionadas por la proporción $y = kx$, donde k es cierto número real, distinto de cero.

Si consideramos x una variable independiente, e y, dependiente, entonces la fórmula $y = kx$ define y como la función x, La recta que pasa por el origen de coordenadas, con un coeficiente angular k (la tangente del ángulo de inclinación de la recta al eje Ox es igual a k) es la gráfica de esta función. La dependencia proporcional directa es un caso particular de la función lineal.

§ 5. Dependencia inversamente proporcional

La variable y se llama *inversamente proporcional* a la variable x, si los valores de estas variables están relacionados mediante la igualdad $y = k/x$, donde k es cierto número real, distinto de cero. El número k se llama coeficiente de proporcionalidad inversa.

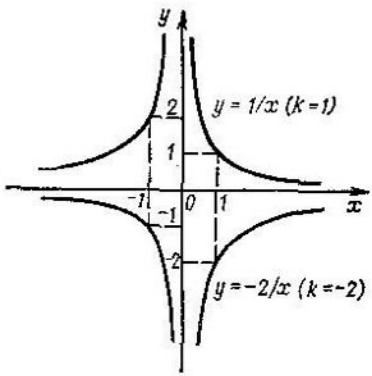


Fig. 10.23.

Si consideramos x una variable independiente, e y, dependiente, entonces la fórmula $y = k/x$ define y como una función de x. La gráfica de la función $y = k/x$ se llama *hipérbola*.

Propiedades de la función $f(x) = k/x$.

1) El campo de definición de la función es un conjunto de todos los números reales, a excepción del número 0.

2) La región de cambio (conjunto de valores) de la función es un conjunto de todos los números reales, a excepción del número 0.

3) La función $f(x) = k/x$ es impar, y su gráfica es simétrica respecto al origen de coordenadas. La función $f(x) = k/x$ es continua y es diferenciable en toda la región de definición. $f'(x) = -k/x^2$. La función no tiene puntos críticos.

4) La función $f(x) = k/x$ para $k > 0$ decrece monótonamente en $(-\infty, 0)$ y $(0, +\infty)$, y para $k < 0$ es monótona en los mismos intervalos.

5) La gráfica de la función $f(x) = k/x$ para $k > 0$ en el intervalo $(0; +\infty)$ está dirigida con la concavidad hacia

las y positivas, y en el intervalo $(-\infty; 0)$, con la concavidad hacia las y negativas. Para $k < 0$ el intervalo de la concavidad hacia las y positivas es $(-\infty; 0)$, el intervalo de la concavidad hacia las y negativas es $(0; +\infty)$.

Las gráficas de la función $f(x) = k/x$ para dos valores k se muestran en la fig. 10.23.

§ 6. Función lineal fraccional

La función que tiene la forma $f(x) = \frac{ax+b}{cx+d}$ se llama función *lineal fraccional* (a, b, c, d son ciertos números, $c \neq 0$)*). Transformemos la expresión lineal fraccional

$$\frac{ax+b}{cx+d} = \frac{a}{c}\cdot\frac{x+\frac{b}{a}}{x+\frac{d}{c}} = \frac{a}{c}\cdot\left[1+\frac{\frac{b}{a}-\frac{d}{c}}{x+\frac{d}{c}}\right] = \frac{\frac{d}{c}-\frac{da}{c^2}}{x+\frac{d}{c}}+\frac{a}{c}.$$

La gráfica de la función lineal fraccional $f(x) = \frac{ax+b}{cx+d}$ puede ser obtenida de la gráfica de la dependencia inversamente proporcional mediante las transformaciones siguientes:

a) por traslación paralela de $\boldsymbol{r} = (-d/c;\ 0)$;

b) por contracción (o alargamiento) hacia el eje de abscisas en $\left(\frac{b}{c}-\frac{da}{c^2}\right)$ veces;

c) por traslación paralela de $\boldsymbol{r} = (0;\ a/c)$.

La transformación sucesiva de la gráfica de la función $f(x) = \frac{1}{x}$ en gráfica de la función $f(x) = \frac{2x+5}{x+1}$ se muestra en las figs. 10.24-10.27. En la fig. 10.24 se muestra la gráfica de la función $f(x) = \frac{1}{x}$. En la fig. 10.25 se muestra la gráfica de la función $f(x) = \frac{1}{x+1}$ que se obtiene de la gráfica de la función $f(x) = \frac{1}{x}$ mediante la traslación paralela de $\boldsymbol{r} = (-1;\ 0)$. En la fig. 10.26 se muestra la gráfica de la función $f(x) = \frac{3}{x+1}$, obtenida de la gráfica $f(x) =$

*) Si $c = 0$, entonces la función lineal fraccional se convierte en función lineal.

$=\frac{1}{x+1}$ mediante el alargamiento del eje de abscisas en 3 veces. Y, por fin, en la fig. 10.27 se muestra la gráfica de la

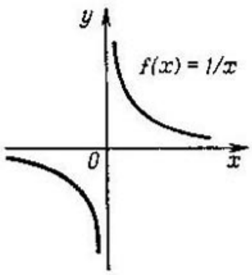


Fig. 10.24.

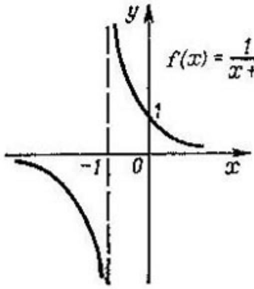


Fig. 10.25.

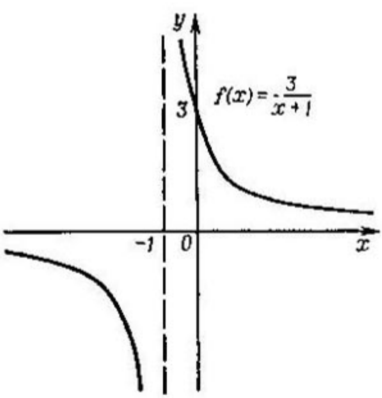


Fig. 10.26.

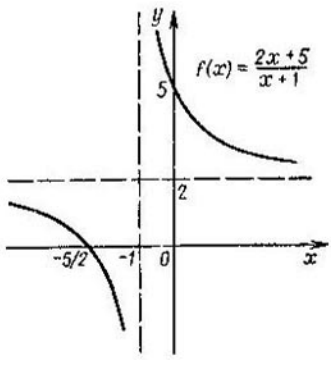


Fig. 10.27.

función lineal fraccional $f(x)=\frac{2x+5}{x+1}$, la cual se obtiene de la gráfica $f(x)=\frac{3}{x+1}$ mediante la traslación paralela de $\boldsymbol{r}=(0;\ 2)$.

§ 7. Función cuadrática

La función $f(x) = ax^2 + bx + c$, donde a, b, c son ciertos números reales ($a \neq 0$), se llama *función cuadrática*. La gráfica de la función cuadrática se llama *parábola*.

La función cuadrática puede ser reducida a la forma

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

La expresión $b^2 - 4ac$ se llama *discriminante* del trinomio de segundo grado. La representación de la función cuadrada en la forma (1) se llama la formación de un *cuadrado perfecto*.

Propiedades de la función cuadrática y su gráfica.

1) El campo de definición de la función cuadrática es toda la recta numérica.

2) Para $b \neq 0$ la función no es par y no es impar. Para $b = 0$ la función cuadrática es par.

3) La función cuadrática es continua y diferenciable en toda la región de definición.

4) La función tiene un solo punto crítico $x = -b/(2a)$. Si $a > 0$, entonces en el punto $x = -b/(2a)$ la función tiene el mínimo. Para $x < -b/(2a)$ la función decrece monótonamente, para $x > -b/(2a)$ crece monótonamente.

Si $a < 0$, entonces en el punto $x = -b/(2a)$ la función tiene el máximo. Para $x < -b/(2a)$ la función crece monótonamente, para $x > -b/(2a)$ decrece monótonamente.

El punto de la gráfica de la función cuadrática con la abscisa $x = -b/(2a)$ y la ordenada $y = -\frac{b^2 - 4ac}{4a}$ se llama *vértice de una parábola*.

5) La región de cambio de la función es: para $a > 0$ un conjunto de valores de la función $\left[-\frac{b^2 - 4ac}{4a}; +\infty\right)$; para $a < 0$, un conjunto de valores de la función $\left(-\infty; -\frac{b^2 - 4ac}{4a}\right]$.

6) La gráfica de la función cuadrática se interseca con el eje Oy en el punto $y = c$. En el caso cuando $b^2 - 4ac > 0$, la gráfica de la función cuadrática interseca el eje Ox en dos puntos (raíces reales distintas de la ecuación cuadrática); si $b^2 - 4ac = 0$ (la ecuación cuadrática tiene una raíz múltiple de 2), la gráfica de la función cuadrática es

tangente al eje Ox en el punto $x = -b/(2a)$; si $b^2 - 4ac < 0$, no existe intersección con el eje Ox.

De la representación de la función cuadrática en la forma (1) también se deduce que la gráfica de la función es simétrica respecto a la recta $x = -b/(2a)$, es decir, a la ima-

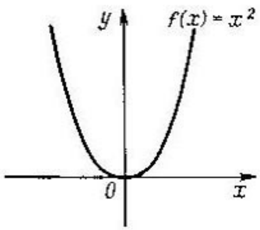


Fig. 10.28.

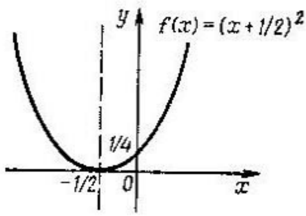


Fig. 10.29.

gen del eje de ordenadas para la traslación paralela $\boldsymbol{r} = (-b/(2a);\ 0)$.

La gráfica de la función

$$f(x) = ax^2 + bx + c$$

$\left(\text{o } f(x) = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}\right)$ puede ser obtenida de la gráfica de la función $f(x) = x^2$ mediante las transformaciones siguientes:

a) por la traslación paralela de $\boldsymbol{r} = (-b/(2a);\ 0)$;

b) por la contracción (o el alargamiento) hacia el eje de abscisas en a veces;

c) por la traslación paralela de $\boldsymbol{r} = \left(0;\ -\frac{b^2 - 4ac}{4a}\right)$.

Ejemplo. Construir la gráfica de la función

$$f(x) = 2x^2 + 2x + 2$$

Reducimos el trinomio de segundo grado a la forma (1):

$$f(x) = 2x^2 + 2x + 2 = 2\left(x + \frac{1}{2}\right)^2 + \frac{3}{2}.$$

Ahora la gráfica de la función $f(x) = 2x^2 + 2x + 2$ puede ser obtenida de la gráfica de la función

$$f(x) = x^2$$

mediante las transformaciones siguientes:

a) por la traslación paralela de $\boldsymbol{r} = (-1/2;\ 0)$;
b) por el alargamiento del eje de abscisas en 2 veces;
c) por la traslación paralela de $\boldsymbol{r} = (0;\ 3/2)$.

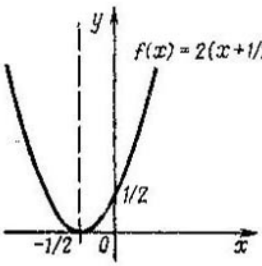


Fig. 10.30.

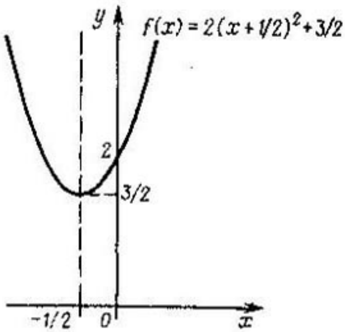


Fig. 10.31.

Los resultados de la aplicación sucesiva de las transformaciones indicadas se muestran en las figs. 10.28—10.31.

§ 8. Función potencial

La función que tiene la forma $f(x) = x^\alpha$, donde α es un número real cualquiera, llamado *exponente de una potencia*, se llama *función potencial*.

Propiedades de una función potencial.

1) El campo de definición de la función potencial es el conjunto de todos los números positivos. La región de cambio de la función potencial es el conjunto de todos los números positivos.

2) La función potencial es aperiódica, no es par y no es impar.

3) La función potencial es continua en todo el campo de definición.

4) La función potencial es diferenciable en todo el campo de definición, y su derivada se calcula mediante la fórmula

$$(x^\alpha)' = \alpha \cdot x^{\alpha-1}.$$

5) La función potencial x^α crece monótonamente en todo el campo de definición para $\alpha > 0$ y decrece monótonamente para $\alpha < 0$.

6) Para $\alpha < 0$ y $\alpha > 1$ la gráfica de la función potencial está dirigida con la concavidad hacia las y positivas, y para $0 < \alpha < 1$ con la concavidad hacia las y negativas.

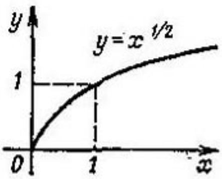


Fig. 10.32.

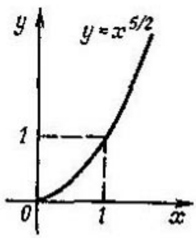


Fig. 10.33.

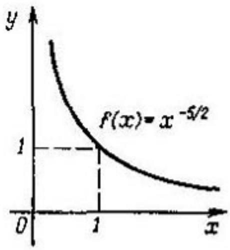


Fig. 10.34.

Las gráficas de la función potencial para ciertos valores α se muestran en las figs. 10.32—10.34*).

§ 9. Función exponencial

La función que tiene la forma de $f(x) = a^x$, donde a es cierto número real positivo que se llama *base de una potencia*, se llama *función exponencial*. Para $a = 1$ el valor de una función exponencial para cualquier valor del argumento es igual a la unidad, y el caso $a = 1$ no se examinará a continuación.

Propiedades de la función exponencial.

1) El campo de definición de la función es toda la recta numérica.

---

*) Aquí la función potencial se define sólo en el semieje positivo. La función potencial se define a veces en toda la recta numérica (para ciertas α concretas).

2) La región de cambio (el conjunto de valores) de la función es el conjunto de todos los números positivos.

3) La función es continua y diferenciable en todo el campo de definición. La derivada de la función exponencial se calcula mediante la fórmula

$$(a^x)' = a^x \ln a.$$

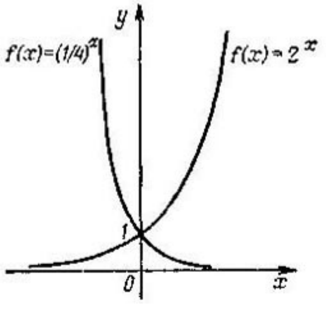


Fig. 10.35.

4) Para $a > 1$ la función crece monótonamente, para $a < 1$ decrece monótonamente.

5) La función exponencial tiene una función inversa que se llama función logarítmica.

6) La gráfica de una función exponencial cualquiera interseca el eje Oy en el punto $y = 1$.

7) La gráfica de la función exponencial es una curva, dirigida con la concavidad hacia las y positivas.

Las gráficas de la función exponencial para los valores $a = 2$ y $a = 1/4$ se muestran en la fig. 10.35.

§ 10. Función logarítmica

La función que es inversa a la función exponencial $y = a^x$ se llama función *logarítmica* y se designa

$$y = \log_a x.$$

El número a se llama *base* de la función logarítmica. La función logarítmica con la base 10 se designa

$$\lg x,$$

y la función logarítmica con la base e se designa

$$\ln x.$$

Propiedades de una función logarítmica:

1) El campo de definición de una función logarítmica es el intervalo $(0; +\infty)$.

2) La región de cambio (el conjunto de valores) de la función logarítmica es toda la recta numérica.

3) La función logarítmica es continua y diferenciable en todo el campo de definición. La derivada de la función logarítmica se calcula según la fórmula

$$(\log_a x)' = \frac{1}{x \ln a}.$$

4) La función logarítmica crece monótonamente, si $a > 1$. Para $0 < a < 1$ la función logarítmica con la base a decrece monótonamente.

5) Para cualquiera base $a > 0$, $a \neq 1$, existen las igualdades

$$\log_a 1 = 0, \quad \log_a a = 1.$$

6) Para $a > 1$ la gráfica de la función logarítmica es una curva, dirigida con la concavidad hacia las y negativas;

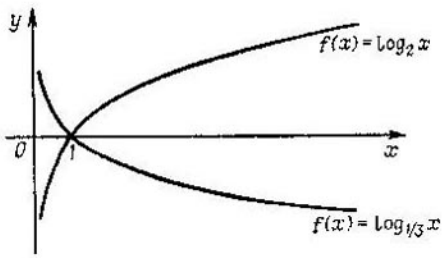


Fig. 10.36.

para $0 < a < 1$ es una curva, dirigida con la concavidad hacia las y positivas.

Las gráficas de la función logarítmica para $a = 2$ y $a = 1/3$ se muestran en la fig. 10.36.

Identidad logarítmica principal. La función logarítmica $x = \log_a y$ es una función inversa para la función exponencial $y = a^x$. Según las propiedades de las funciones recíprocamente inversas f y f^{-1}

$$f(f^{-1}(y)) = y \tag{1}$$

para todas x del campo de definición de la función $f^{-1}(x)$. En particular, para las funciones logarítmica y exponencial la igualdad (1) obtiene la forma

$$a^{\log_a y} = y. \qquad (2)$$

La igualdad (2) se llama con frecuencia *identidad logarítmica básica*.

Para una función logarítmica, cualesquiera que sean x e y positivas, son válidas las igualdades siguientes, que pueden ser obtenidas como consecuencia de la identidad logarítmica básica (2) y de la propiedad de la función exponencial:

$$\log_a (x \cdot y) = \log_a x + \log_a y;$$

$$\log_a \left(\frac{x}{y}\right) = \log_a x - \log_a y;$$

$$\log_a (x^\alpha) = \alpha \cdot \log_a x \text{ (α es un número real cualquiera)}$$

$$\log_a a = 1;$$

$$\log_a x = \frac{\log_b x}{\log_b a} \text{ (b es un número real, $b > 0$, $b \neq 1$).}$$

En particular, de la última fórmula para $a = e$, $b = 10$ se obtiene la igualdad

$$\ln x = \frac{1}{\lg e} \lg x. \qquad (3)$$

El número $\lg e$ se llama *módulo* de paso de los logaritmos naturales a los decimales, y se designa con la letra M, mientras que la fórmula (3) suele escribirse en la forma

$$\lg x = M \cdot \ln x.$$

El número M con una precisión hasta el sexto signo es igual a

$$M = 0{,}434294\ \ldots$$

Tabla de las fórmulas más frecuentes

Combinatoria

El número de permutaciones de n elementos es:

$$P_n = n!$$

El número de variaciones de n elementos por m elementos es:

$$A_n^m = \frac{n!}{(n-m)!}.$$

Número de combinaciones de n elementos por m elementos:

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

Las fórmulas para el número de combinaciones son:

$$C_n^m = C_n^{n-m};$$

$$C_{n+1}^{m+1} = C_n^m + C_n^{m+1};$$

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^{n-1} + C_n^n = 2^n.$$

Binomio de Newton:

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

Potencias y logaritmos

Potencias:

$$a^0 = 1\ (a \neq 0);$$

$$a^x \cdot a^y = a^{x+y};$$

$$\frac{a_x}{a_y} = a^{x-y};$$

$$(a^x)^y = a^{xy};$$
$$(a \cdot b)^x = a^x \cdot b^x;$$
$$\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}.$$

Logaritmos:

$$\log_a a = 1 \quad (a > 0), \quad (a \neq 1);$$
$$\log_a (a^k) = k \; (a > 0, \; a \neq 1),$$
$$\log_a (M_1 M_2) = \log_a M_1 + \log_a M_2 \; (M_1 > 0, \; M_2 > 0);$$
$$\log_a \left(\frac{M_1}{M_2}\right) = \log_a M_1 - \log_a M_2 \; (M_1 > 0, \; M_2 > 0);$$
$$\log_a (b^c) = c \log_a b;$$
$$\log_a c = \frac{\log_b c}{\log_b a}\; ; \; \log_a c = \frac{1}{\log_c a}.$$

Álgebra

Fórmulas de la multiplicación abreviada:

$$(x + c)\,(x - c) = x^2 - c^2;$$
$$(x + c)\,(x^2 - xc + c^2) = x^3 + c^3;$$
$$(x - c)\,(x^2 + xc + c^2) = x^3 - c^3.$$

Fórmulas de Viète para las raíces:

$$x_1 + x_2 = -p, \; x_1 \cdot x_2 = q.$$

Raíces de la ecuación cuadrática $ax^2 + bx + c = 0$;

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Raíces de la ecuación cuadrática con el segundo coeficiente par $ax^2 + 2kx + c = 0$:

$$x_{1,2} = \frac{-k \pm \sqrt{k^2 - ac}}{a}.$$

Raíces de la ecuación cuadrática reducida $x^2 + px + q = 0$;

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

Desigualdades simples:

$$|a+b| \leqslant |a| + |b|;$$

$$|a-b| \geqslant ||a| - |b||;$$

$$a^2 + b^2 \geqslant 2ab;$$

$$\frac{a}{b} + \frac{b}{a} \geqslant 2 \text{ (} a \text{ y } b \text{ son números del mismo signo);}$$

$$\frac{a+b}{2} \geqslant \sqrt{ab} \quad (a \geqslant 0,\ b \geqslant 0).$$

Vectores

El módulo del vector $\boldsymbol{a} = (x_1;\ y_1;\ z_1)$ es:

$$|\boldsymbol{a}| = \sqrt{x_1^2 + y_1^2 + z_1^2}.$$

El producto escalar de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es:

$$(\boldsymbol{a},\ \boldsymbol{b}) = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cos \widehat{(\boldsymbol{a},\ \boldsymbol{b})}.$$

El producto escalar de los vectores, representados por las coordenadas es:

$$(\boldsymbol{a},\ \boldsymbol{b}) = x_1x_2 + y_1y_2 + z_1z_2.$$

Geometría

El *triángulo.*

Area de un triángulo:

$$S = \frac{1}{2} ah_a = \frac{1}{2} bh_b = \frac{1}{2} ch_c;$$

$$S = \sqrt{p(p-a)(p-b)(p-c)} \quad \textit{(fórmula de Geron)};$$

$$S = \frac{1}{2} ab \operatorname{sen} \gamma;$$

$$S = \frac{abc}{4R};$$

$$S = pr.$$

Aquí a, b, c son los lados del triángulo, h_a, h_b, h_c, las alturas, $p = \frac{1}{2}(a+b+c)$, el semiperímetro, R es el radio de la circunferencia, descrita alrededor del triángulo y r es el radio de la circunferencia, inscrita en el triángulo, γ es el ángulo situado entre los lados a y b.

El *rectángulo*.

El área del rectángulo es:

$$S = ab,$$

donde a y b son los lados adyacentes del rectángulo.

El área del cuadrado es:

$$S = a^2,$$

donde a es un lado del cuadrado.

El *trapecio*.

El área del trapecio es:

$$S = \frac{a+b}{2} h,$$

donde a y b son las bases del trapecio y h es la altura del trapecio.

El *círculo* y la *circunferencia*.

El área del círculo de radio r es:

$$S = \pi r^2.$$

La longitud de la circunferencia de radio r es:

$$l = 2\pi r.$$

El área del sector con un valor angular del arco α^0 es:

$$S_{\text{sect.}} = \frac{\pi r^2 \alpha}{360}.$$

El *polígono*.

El área de un n-polígono regular es igual a la mitad del producto de su perímetro por el radio de la circunferencia inscrita:

$$S_n = \frac{1}{2} P_n r.$$

El área de un n-polígono regular es:

$$S = \frac{1}{2} R^2 n \operatorname{sen} \frac{360^\circ}{n},$$

donde R es el radio de la circunferencia descrita.

El *prisma*.

El área de la superficie lateral es:

$$S_{\text{lat.}} = P_p \cdot | A_1A_2 |,$$

donde P_n es el perímetro de la sección perpendicular del prisma, $| A_1A_2 |$, la longitud de la arista lateral.

Volumen del prisma:

$$V = S_p \cdot | A_1A_2 |,$$

donde S_n es el área de la sección perpendicular del prisma, $| A_1A_2 |$ es la longitud de la arista lateral,
o

$$V = S_{\text{base}} \cdot H,$$

donde S_{base} es el área de la base del prisma y H es la altura del prisma.

Volumen de un paralelepípedo rectangular:

$$V = a \cdot b \cdot c$$

(a, b, c son las dimensiones del paralelepípedo).

Volumen de un cubo:

$$V = a^3$$

(a es la dimensión del cubo).

La *pirámide*.

El área de la superficie lateral de una pirámide regular es:

$$S = \frac{1}{2} Ph,$$

donde P es el perímetro de la base de la pirámide y h es la apotema.

Volumen de una pirámide:

$$V = \frac{1}{3} SH,$$

donde S es el área de la base de la pirámide y H es la altura de la pirámide.

La *pirámide truncada*.

El área de la superficie lateral de la pirámide truncada regular es:

$$S = \frac{1}{2} (P + p) h,$$

donde P, p son los perímetros de las bases de la pirámide, h es la apotema.

Volumen de una pirámide truncada:

$$V = \frac{1}{3} H \left(S_1 + \sqrt{S_1 S_2} + S_2\right),$$

donde H es la altura de la pirámide truncada, S_1 y S_2 son los áreas de las bases de la pirámide truncada.

El *cilindro*.

Volumen del cilindro:

$$V = \pi R^2 H,$$

donde R es el radio de la base del cilindro y H es la altura del cilindro.

Área de la superficie lateral y total del cilindro:

$$S_{\text{lat}} = 2\pi RH;$$

$$S_{\text{cil}} = 2\pi RH + 2\pi R^2,$$

donde R es el radio de la base del cilindro y H es la altura del cilindro.

El *cono*.

Volumen del cono:

$$V_{\text{con}} = \frac{1}{3} \pi R^2 H,$$

don... R es el radio de la base del cono y H es la altura del cono. El área de la superficie lateral del cono es:

$$S_{\text{lat}} = \pi RL,$$

donde L es la generatriz del cono.

El *cono truncado*.

Volumen del cono truncado:

$$V = \frac{1}{3} \pi H \left(R_1^2 + R_1 R_2 + R_2^2\right),$$

donde H es la altura del cono truncado, R_1 y R_2 son los radios de la base superior e inferior del cono truncado.

El área de la superficie lateral del cono truncado es:

$$S_{\text{lat}} = \pi \left(R_1 + R_2\right) L,$$

donde L es la generatriz del cono truncado.

La *esfera*.
Area de una esfera de radio R:

$$S = 4\pi R^2.$$

Volumen de una esfera de radio R:

$$V = \frac{4}{3}\pi R^3.$$

Trigonometría

$$\operatorname{sen}^2\alpha + \cos^2\alpha = 1;$$

$$\operatorname{tg}\alpha\cdot\operatorname{ctg}\alpha = 1;$$

$$1 + \operatorname{tg}^2\alpha = \frac{1}{\cos^2\alpha};$$

$$1 + \operatorname{ctg}^2\alpha = \frac{1}{\operatorname{sen}^2\alpha};$$

$$\operatorname{sen}(\alpha \pm \beta) = \operatorname{sen}\alpha\cos\beta \pm \cos\alpha\operatorname{sen}\beta;$$

$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \operatorname{sen}\alpha\operatorname{sen}\beta;$$

$$\operatorname{tg}(\alpha \pm \beta) = \frac{\operatorname{tg}\alpha \pm \operatorname{tg}\beta}{1 \mp \operatorname{tg}\alpha\operatorname{tg}\beta};$$

$$\operatorname{sen}2\alpha = 2\operatorname{sen}\alpha\cos\alpha;$$

$$\cos 2\alpha = \cos^2\alpha - \operatorname{sen}^2\alpha = 1 - 2\operatorname{sen}^2\alpha = 2\cos^2\alpha - 1;$$

$$\operatorname{tg}2\alpha = \frac{2\operatorname{tg}\alpha}{1 - \operatorname{tg}^2\alpha};$$

$$1 + \cos\alpha = 2\cos^2\frac{\alpha}{2}; \qquad 1 - \cos\alpha = 2\operatorname{sen}^2\frac{\alpha}{2};$$

$$\cos\alpha = \frac{1 - \operatorname{tg}^2\dfrac{\alpha}{2}}{1 + \operatorname{tg}^2\dfrac{\alpha}{2}}; \qquad \operatorname{sen}\alpha = \frac{2\operatorname{tg}\dfrac{\alpha}{2}}{1 + \operatorname{tg}^2\dfrac{\alpha}{2}};$$

$$\operatorname{sen}\alpha + \operatorname{sen}\beta = 2\operatorname{sen}\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2};$$

$$\operatorname{sen}\alpha - \operatorname{sen}\beta = 2\cos\frac{\alpha+\beta}{2}\operatorname{sen}\frac{\alpha-\beta}{2};$$

$$\cos\alpha + \cos\beta = 2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2};$$

$$\cos\alpha - \cos\beta = -2\,\mathrm{sen}\,\frac{\alpha+\beta}{2}\,\mathrm{sen}\,\frac{\alpha-\beta}{2} = 2\,\mathrm{sen}\,\frac{\alpha+\beta}{2}\,\mathrm{sen}\,\frac{\beta-\alpha}{2};$$

$$\mathrm{tg}\,\alpha \pm \mathrm{tg}\,\beta = \frac{\mathrm{sen}\,(\alpha\pm\beta)}{\cos\alpha\cos\beta}; \quad \mathrm{ctg}\,\alpha \pm \mathrm{ctg}\,\beta = \frac{\mathrm{sen}\,(\beta\pm\alpha)}{\mathrm{sen}\,\alpha\,\mathrm{sen}\,\beta};$$

$$\mathrm{sen}\,\alpha\,\mathrm{sen}\,\beta = \frac{1}{2}\,[\cos(\alpha-\beta) - \cos(\alpha+\beta)];$$

$$\cos\alpha\cos\beta = \frac{1}{2}\,[\cos(\alpha-\beta) + \cos(\alpha+\beta)];$$

$$\mathrm{sen}\,\alpha\cos\beta = \frac{1}{2}\,[\mathrm{sen}\,(\alpha-\beta) + \mathrm{sen}\,(\alpha+\beta)].$$

Teorema de los senos:

$$\frac{a}{\mathrm{sen}\,\alpha} = \frac{b}{\mathrm{sen}\,\beta} = \frac{c}{\mathrm{sen}\,\gamma} = 2R.$$

Teorema de los cosenos:

$$a^2 = b^2 + c^2 - 2bc\cos\alpha;$$
$$b^2 = a^2 + c^2 - 2ac\cos\beta;$$
$$c^2 = b^2 + a^2 - 2ab\cos\gamma;$$

(a, b, c son los lados del triángulo que se encuentran frente a los ángulos α, β, γ respectivamente, R es el radio de la circunferencia circunscrita).

Progresión aritmética.

La fórmula del n-ésimo término y la suma de los primeros n términos son:

$$a_n = a_1 + d\,(n-1), \qquad S_n = \frac{2a_1 + d\,(n-1)}{2}\cdot n.$$

Progresión geométrica.

La fórmula del n-ésimo término y la suma de los primeros n términos son:

$$u_n = u_1\cdot q^{n-1}, \qquad S_n = \frac{u_1\,(1-q^n)}{1-q}.$$

La suma de los términos de una progresión geométrica decreciente infinitamente es:

$$S = \frac{u_1}{1-q}.$$

Las *derivadas.*

$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
c	0	$\operatorname{sen} x$	$\cos x$
x^{α}	$\alpha x^{\alpha-1}$	$\cos x$	$-\operatorname{sen} x$
a^{x}	$a^{x}\ln a$	$\operatorname{tg} x$	$\frac{1}{\cos^2 x}$
e^{x}	e^{x}		
$\log_a x$	$\frac{1}{x\cdot\ln a}$	$\operatorname{ctg} x$	$-\frac{1}{\operatorname{sen}^2 x}$
$\ln x$	$\frac{1}{x}$	$\operatorname{arcsen} x$	$\frac{1}{\sqrt{1-x^2}}$

Integrales indefinidas.

$$\int 0\cdot dx = C;$$

$$\int 1\cdot dx = x + C;$$

$$\int x^{\alpha}\,dx = \frac{x^{\alpha+1}}{\alpha+1} + C\ (\alpha \neq -1);$$

$$\int \frac{1}{x}\,dx = \ln|x| + C;$$

$$\int a^{x}\,dx = \frac{a^{x}}{\ln a} + C;$$

$$\int \operatorname{sen} x\,dx = -\cos x + C;$$

$$\int \cos x\,dx = \operatorname{sen} x + C;$$

$$\int \frac{1}{\cos^2 x}\,dx = \operatorname{tg} x + C;$$

$$\int \frac{1}{\operatorname{sen}^2 x}\,dx = -\operatorname{ctg} x + C.$$

Anexo

Sistemas de numeración

Conjuntamente con el sistema de numeración posicional decimal que se utiliza en la práctica, para el trabajo en las máquinas computadoras electrónicas se suelen utilizar los sistemas de numeración binarios, de base 8 (octonario) y de base 16 (así como sistemas mixtos de numeración). Al exponer los conocimientos básicos sobre los sistemas octonarios y binarios de numeración y los procedimientos de conversión de los números de un sistema de numeración a otro, en vez de las palabras «el número escrito en el sistema decimal (octonario, binario) de numeración» vamos a decir simplemente «el número decimal (respectivamente octonario, binario)».

Sistema octonario de numeración. En el sistema posicional octonario de numeración para la escritura de los números se usan ocho cifras distintas: 0, 1, 2, 3, 4, 5, 6, 7. El número «ocho» se escribe 10. En el sistema posicional octonario de numeración la anotación

$$a_n a_{n-1} \dots a_1 a_0, \quad a_{-1} \dots a_{-k} \dots, \tag{1}$$

donde $a_n, a_{n-1}, \dots, a_1, a_0, a_{-1}, \dots, a_{-k}, \dots$ son cifras del sistema octonario de numeración y por lo menos una cifra es distinta de cero, representa una anotación abreviada del número positivo a:

$$a = a_n \cdot 8^n + a_{n-1} \cdot 8^{n-1} + \dots + a_1 \cdot 8^1 + a_0 \cdot 8^0 + \\ + a_{-1} \cdot 8^{-1} + \dots + a_{-k} \cdot 8^{-k} + \dots \tag{2}$$

Para la anotación del número negativo $-a$ delante de la sucesión de las cifras (1) se pone el signo menos. Todos los conceptos introducidos para las fracciones decimales, que son una forma de anotación de los números en el sistema de numeración posicional decimal, se trasladan también literalmente a las expresiones del tipo (1), las cuales se pueden llamar fracciones octonarias.

Reglas de conversión de los números decimales en octonarios:

Conversión de la parte entera de un número. Dividamos la parte entera de un número decimal por el número 8. La cifra obtenida en el resto es la última cifra de la anotación octonaria del número dado. El cociente obtenido lo dividimos por el número 8; la cifra, que es el resto, es la penúltima cifra de la anotación octonaria del número. Continuando el proceso de división de los cocientes que obtenemos, escribimos sucesivamente las cifras de la anotación octonaria del número. El proceso de división termina, cuando en el cociente se obtiene una cifra, menor que 8, la cual será la primera cifra de la anotación octonaria del número dado.

Ejemplo 1. Escribir el número decimal 201 en el sistema octonario de numeración:

$$
\begin{array}{r|l}
-201 & \;\;8 \\
\cline{2-2}
\underline{16} & \begin{array}{r|l} -25 & 8 \\ \cline{2-2} \underline{24} & 3 \\ 1 & \end{array} \\
-41 & \\
\underline{40} & \\
1 &
\end{array}
$$

El proceso de división de los cocientes se termina en el número 3 puesto que $3 < 8$. El número decimal 201 en el sistema octonario de numeración se escribe como 311.

Conversión de la parte fraccionaria de un número. Multiplicamos la parte fraccionaria del número por 8. La parte entera del producto es la primera cifra de la anotación octonaria de la parte fraccional del número. Multiplicando la parte fraccionaria del producto obtenido por 8, obtenemos un número, cuya parte entera es la segunda cifra de la anotación octonaria de la fracción. El proceso de multiplicación continúa hasta que la parte fraccionaria del producto no esté compuesta sólo de ceros o hasta que obtengamos la cantidad necesaria de cifras de la anotación octonaria de la parte fraccionaria del número. Observemos, que el número, representado por una fracción decimal finita en el sistema octonario, puede resultar ser representado en forma de una fracción octonaria periódica infinita (véase el ejemplo 3).

Ejemplo 2. Escribir el número decimal 0,1875 en el sistema octonario de numeración.

Representemos el proceso de multiplicación en forma de dos columnas de números, donde la columna izquierda es la parte entera del número, la cual al realizar la siguiente multiplicación se sustituye por cero.

$$
\begin{array}{r|r}
\times 0 & 1875 \\
 & 8 \\
\hline
\times 1 & 5000 \\
 & 8 \\
\hline
4 & 0000
\end{array}
$$

El número decimal 0,1875 se escribe en forma de fracción octonaria 0,14.

Ejemplo 3. Escribir el número decimal 0,3 en el sistema octonario de numeración.

La conversión del número en el sistema octonario de numeración se realiza de la misma manera que en el ejemplo anterior:

$\times 0$	3 8
$\times 2$	4 8
$\times 3$	2 8
$\times 1$	6 8
$\times 4$	8 8
$\times 6$	4 8
$\times 3$	2 8

Es fácil ver que comenzando con el paso de multiplicación, marcado con la segunda estrellita, en la columna izquierda aparece la combinación de las cifras 3146, la cual representa el período de la fracción octonaria. De esta manera, el número decimal $0,\bar{3}$ en el sistema octonario de numeración se escribe en forma de fracción 0,2 (3146).

La conversión de los números escritos en el sistema octonario de numeración al sistema decimal se suele realizar mediante la fórmula (2).

Ejemplo 4. Escribir el número octonario 163,2 en el sistema decimal de numeración.

Según la fórmula (2) la anotación 163,2 es la anotación abreviada del número

$$1\cdot 8^2 + 6\cdot 8^1 + 3\cdot 8^0 + 2\cdot 8^{-1} = 64 + 48 + 3 + \frac{1}{4} =$$

$$= 64 + 48 + 3 + 0,25 = 115,25.$$

Así pues, el número octonario 163,2 en el sistema decimal de numeración se escribe como 115,25.

Sistema binario de numeración. En el sistema posicional binario de numeración para la anotación de los números se usan dos símbolos (cifras): 0 (el cero) y 1 (la unidad). El número «dos» en el sistema binario de numeración se escribe 10. En el sistema binario de numeración la anotación

$$a_n a_{n-1} \ldots a_1 a_0, \quad a_{-1} \ldots a_{-k} \ldots, \tag{3}$$

donde $a_n, a_{n-1}, \ldots, a_1, a_0, a_{-1}, \ldots, a_{-k} \ldots$ son cifras del sistema binario de numeración (0 ó 1) y por lo menos una de las cifras es distinta de cero, representa una anotación abreviada del número positivo a:

$$a = a_n \cdot 2^n + a_{n-1} \cdot 2^{n-1} + \dots + a_1 \cdot 2^1 + a_0 \cdot 2^0 + \\ + a_{-1} \cdot 2^{-1} + \dots + a_{-k} \cdot 2^{-k} + \dots \qquad (4)$$

Para la anotación del número negativo $-a$, delante de la sucesión de las cifras (3) se pone el signo menos. Todos los conceptos que se introducen para las fracciones decimales se trasladan literalmente a las expresiones que tienen la forma (3).

Las tablas binarias de adición y multiplicación tienen la forma:

La tabla de adición	La tabla de multiplicación
$0 + 0 = 0$	$0 \times 0 = 0$
$0 + 1 = 1$	$0 \times 1 = 0$
$1 + 0 = 1$	$1 \times 0 = 0$
$1 + 1 = 10$	$1 \times 1 = 1$

La conversión de los números enteros y fraccionarios, escritos en el sistema decimal de numeración, al sistema binario se puede efectuar de la misma manera que se hace para la conversión de los números decimales en números octonarios.

Ejemplo 5. Escribir el número decimal 19 en el sistema binario de numeración:

$$\begin{array}{r|l}
19 & 2 \\
-18 & 9 \;|\; 2 \\
1 & -8 \quad 4 \;|\; 2 \\
 & \;\;1 \quad -4 \quad 2 \;|\; 2 \\
 & \qquad 0 \quad -2 \quad 1 \\
 & \qquad\qquad 0
\end{array}$$

De esta manera, el número 19 en el sistema binario de numeración se escribe como 10011.

La conversión de los números del sistema decimal de numeración al sistema binario se puede realizar también mediante otro procedimiento. Este procedimiento se basa en que entre los números 0, 1, 2, 3, 4, 5, 6, 7, y los números binarios que se componen de tres cifras existe una correspondencia biunívoca, es decir, los números enumerados en el sistema binario de numeración se escriben respectivamente como

000, 001, 010, 011, 100, 101, 110, 111

Teniendo en cuenta lo expuesto más arriba, el número decimal se pasa al sistema binario en dos etapas. Primeramente se escribe el número decimal en el sistema octonario de numeración. Después cada cifra octonaria se sustituye por el grupo respectivo de tres cifras binarias.

Ejemplo 6. Escribir el número decimal 201, 1875 en el sistema binario de numeración.

El número decimal dado en el sistema octonario de numeración se escribe como 311,14 (véase los ejemplos 1, 2). Cada cifra del número octonario 311,14 se sustituye por un grupo de tres cifras binarias:

$$3 \to 011; \quad 1 \to 001; \quad 4 \to 100.$$

De resultas el número octonario 311,14 en el sistema binario de numeración se escribe en la forma

$$011\ 001\ 001{,}001\ 100$$

Despreciando los ceros que se encuentran en el principio de la parte entera y en el final de la parte fraccionaria del número binario, definitivamente obtenemos que el número decimal 201,1875 se escribe en el sistema binario de numeración en la forma 11001001,0011.

La conversión de los números del sistema binario de numeración al sistema decimal es cómodo realizarla mediante la fórmula (4).

Ejemplo 7. Escribir el número binario 10110,11 en el sistema decimal de numeración.

Según la fórmula (4) la anotación 10110,11 es la anotación abreviada del número

$$1\cdot 2^4 + 0\cdot 2^3 + 1\cdot 2^2 + 1\cdot 2^1 + 0\cdot 2^0 + 1\cdot 2^{-1} + 1\cdot 2^{-2} =$$
$$= 16 + 4 + 2 + \frac{1}{2} + \frac{1}{4} = 22 + \frac{3}{4} = 22{,}75$$

Así pues, el número binario 10110,11 en el sistema decimal de numeración se escribe como 22,75.

Lista de las principales notaciones

$[a; b]$	es un intervalo cerrado (segmento): los puntos (números) a y b pertenecen al intervalo
$(a; b]$	y $[a; b)$ son intervalos semiabiertos: el punto a no pertenece al intervalo; el punto b pertenece al intervalo; respectivamente, el punto a pertenece, y el b no pertenece al intervalo
$(a; b)$	es un intervalo abierto: los puntos a y b no pertenecen al intervalo
$\Rightarrow$	es el símbolo que significa «sigue en una dirección»
$\Leftrightarrow$	es el símbolo que significa «sigue en ambas direcciones»
$\in$, $\ni$	son los símbolos de pertenencia
$\notin$	es el símbolo «no pertenece»
$\varnothing$	es el símbolo del conjunto vacío
$\cup$	es el símbolo de la unión de conjuntos
$\cap$	es el símbolo de la intersección de conjuntos
$\setminus$	es el símbolo del complemento de conjuntos
$\sim$	es el signo de equivalencia (en geometría es el símbolo de semejanza)
$\{a_1; a_2;$	$\ldots; a_n\}$, es un conjunto, formado de los elementos $a_1, a_2, \ldots, a_n$
$(a_1; a_2;$	$\ldots; a_n)$, un conjunto ordenado, formado de los elementos $a_1, a_2, \ldots, a_n$
f y f^{-1},	g y g^{-1}, etc. son las designaciones de las aplicaciones recíprocamente inversas
$\leqslant$	en la teoría de conjuntos es el símbolo «no supera»
$n!$	es la notación del producto $1\cdot2\cdot3\ldots\cdot n$ (se lee «en factorial»)

P_n,	el número de reordenaciones de n elementos
C_n^m o $\binom{n}{m}$	es el número de combinaciones de n elementos por m elementos
A_n^m,	el número de variaciones de n elementos por m elementos
$\Vert a_{ij} \Vert$	o (a_{ij}) o $[a_{ij}]$ es la matriz, compuesta de los números a_{ij}
$\vert a_{ij} \vert$ o det $\Vert a_{ij} \Vert$,	el determinante de la matriz cuadrada $\Vert a_{ij} \Vert$
N	es el conjunto de todos los números naturales
Z	el conjunto de todos los números enteros
Q	es el conjunto de todos los números racionales
R,	el conjunto de todos los números reales
$\{m, n\}$	es el mínimo común múltiplo de los números m y n
(m, n),	el máximo común divisor de los números m y n
$>, <, \geqslant, \leqslant$	son los símbolos «mayor», «menor», «mayor o igual que», «menor o igual que»
$\vert a \vert$	es el valor absoluto (módulo) de un número real
$\sqrt[n]{}$	es la raíz n-ésima potencia
$\sqrt{}$	es la raíz cuadrada
$\log_a M$	son los logaritmos del número M de base a
$\lg M$	es el logaritmo del número M de base 10
$\ln M$,	el logaritmo del número M de base e
N, n_1n_2	$\ldots n_k \ldots$ es una fracción decimal con la parte entera N
N, n_1n_2	$\ldots n_k (n_{k+1} \ldots n_{k+p})$ es una fracción decimal con un período de $n_{k+1} \ldots n_{k+p}$
z y $\bar{z}$	son números conjugados complejamente
i	es una unidad imaginaria
Re z,	la parte real del número complejo z
Im z,	la parte imaginaria del número complejo z
$\vert z \vert$,	el módulo del número complejo z
Arg z	es el argumento del número complejo z
$\boldsymbol{i}$, $\boldsymbol{j}$	son los vectores básicos del sistema de coordenadas cartesiano rectangular en el plano
$\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$	son los vectores básicos del sistema de coordenadas cartesiano rectangular en el espacio

$\widehat{(\boldsymbol{a}, \boldsymbol{b})}$	es el ángulo entre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$
$(\boldsymbol{a}, \boldsymbol{b})$,	el producto escalar de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$
$[\boldsymbol{a}, \boldsymbol{b}]$,	el producto vectorial de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$
$(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$,	el producto mixto (escalar-vectorial) de los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$
$\lvert AB \rvert$,	la longitud del segmento AB
$\angle$	es el símbolo del ángulo
$^\circ$	es un grado
$'$	es un minuto
$''$	es un segundo
d	es la designación del ángulo recto
$\widehat{ABC}$	es el valor del ángulo $\angle ABC$
$\uparrow\uparrow$	es el símbolo de la codirección (de los rayos)
$\downarrow\uparrow$,	el símbolo de la dirección opuesta
$\parallel$,	el símbolo de paralelismo
$\perp$,	el símbolo de perpendicularidad
$\angle(a, \alpha)$	es un ángulo entre la recta a y el plano α
$\angle(\alpha, \beta)$,	el ángulo entre los planos α y β
$\cup ACB$	es un arco con los extremos A y B
$\overset{\frown}{ACB}$,	el valor angular del arco $\cup ACB$
Z_0	es la simetría central del plano con centro O
S_l,	la simetría axial del espacio (plano) con el eje de simetría l
S_α,	la simetría del espacio respecto al plano α
sen α	es el seno α
cos α,	el coseno α
tg α,	la tangente α
ctg α,	la cotangente α
sec α,	la secante α
cosec α,	la cosecante α
arcsen a	es el arco seno del número a
arccos a,	el arco coseno del número a
arctg a,	el arco tangente del número a
arcctg a,	el arco cotangente del número a
(a_n)	es una sucesión con los términos $a_1, a_2, \ldots, a_n, \ldots$
lím	es el símbolo de un límite
Δx	es el incremento del argumento
Δf,	el incremento de una función

$f'(x)$	o $\frac{df}{dx}$ es la primera derivada de la función $f(x)$
$f^{(n)}(x)$	o $\frac{d^n f}{dx^n}$, la n-ésima derivada de la función $f(x)$
$\int$	es el símbolo de la integral indefinida
$\int_a^b$,	el símbolo de la integral definida
máx	es el máximo
mín,	el mínimo

Número e = 2,718281828...
Número π = 3,141592653... (π es la razón entre la longitud de la circunferencia y el diámetro)
Número M (módulo de conversión) = 0,434294481...

Indice alfabético de materias