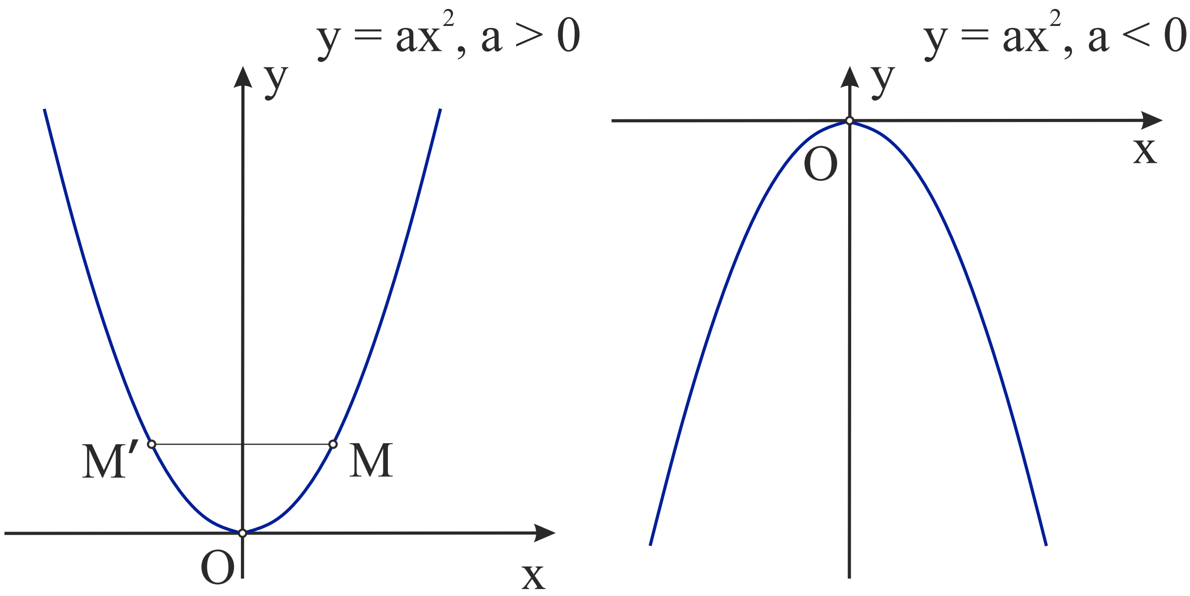
CONIC SECTIONS

Throughout this theme, it is assumed that **a certain scale has been chosen in the plane** (in which all the figures considered below lie); only **rectangular coordinate systems with this scale are considered**.

1 Parabola

A parabola is known from a school mathematics course as a curve that is a graph (Fig. 1) of a function

$$y = ax^2, \qquad a \neq 0. \tag{1.1}$$



(a) (b)

Figure 1

The graph of any square trinomial

$$y = ax^2 + bx + c \tag{1.2}$$

is also a parabola; it is possible by just shifting the coordinate system Oxy (to some vector $\overrightarrow{OO'}$), i.e. by transformation

$$\begin{aligned} x &= x' + x_0, \\ y &= y' + y_0, \end{aligned} \tag{1.3}$$

make the graph of the function $y' = ax'^2$ (in the second coordinate system) coincide with the graph (1.2) (in the first coordinate system).

Indeed, let's substitute (1.3) into the equality (1.2). We get

$$y' + y_0 = a(x' + x_0)^2 + b(x' + x_0) + c,$$

i.e.

$$y' = ax'^2 + (2ax_0 + b)x' + (ax_0^2 + bx_0 + c - y_0).$$

We want to choose x_0 and y_0 so that the coefficient of x' and the free term of the polynomial (with respect to x') on the right side of this equality are equal to zero. To do this, we determine x_0 from the equation

$$2ax_0 + b = 0,$$

which gives

$$x_0 = -\frac{b}{2a}.$$

Now we determine y_0 from condition

$$ax_0^2 + bx_0 + c - y_0 = 0,$$

into which we substitute the value $x_0 = -\dfrac{b}{2a}$. We have

$$\frac{b^2}{4a} - \frac{b^2}{2a} + c = y_0,$$

i.e.

$$y_0 = \frac{4ac - b^2}{4a}.$$

So, by shifting (1.3), in which

$$x_0 = -\frac{b}{2a}, \qquad y_0 = \frac{4ac - b^2}{4a},$$

we moved to a new coordinate system (Fig. 2), in which the equation of the parabola (1.2) becomes[1]

$$y' = ax'^2.$$

Let's go back to the equation (1.1). It can serve as a definition of a parabola. We recall its simplest properties. The curve has an axis of symmetry: if the point

[1] The same result can also be obtained by the "complement to the full square" method known from school.

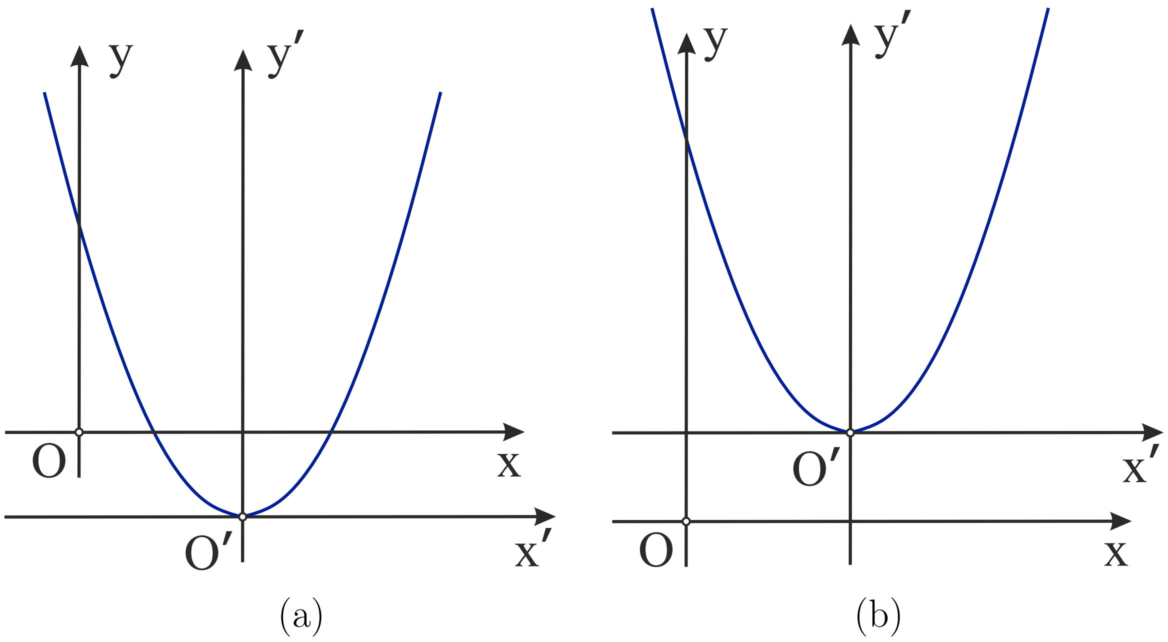


Figure 2

$M = (x, y)$ satisfies the equation (1.1), then the point $M' = (-x, y)$, which is symmetric to the point M with respect to the y-axis , also satisfies the equation (1.1) — the curve is symmetrical about the y-axis (Fig. 1).

If $a > 0$, then the parabola (1.1) lies in the upper semiplane $y \geqslant 0$, having the only common point O with the abscissa. As the modulus of the abscissa x increases indefinitely, the ordinate $y = ax^2$ also increases indefinitely. The general view of the curve is given in Fig. 1a.

If $a < 0$ (Fig. 1b), then the curve is located in the lower semiplane symmetrically with respect to the x-axis to the curve $y = |a|x^2$.

If we switch to a new coordinate system obtained from the old one by replacing the positive direction of the y-axis with the opposite one[2], then the parabola, which has the equation $y = ax^2$ in the old system, will receive the equation $y = -ax^2$ in the new coordinate system. Therefore, when studying parabolas, one can confine oneself to equations (1.1) in which $a > 0$.

Finally, let's change the names of the axes, i.e. let's move on to a new coordinate system, in which the y-axis will be the old abscissa axis, and the x-axis will be the old axis of ordinates. In this new system, the equation (1.1) will be written as

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

[2]Denoting the coordinates of some point M by x, y, we obtain in the new system for the same point the coordinates x, $-y$.

or, if the number $1/a$ is denoted by $2p$, in the form (Fig. 3)

$$\boxed{y^2 = 2px, \qquad p > 0.} \tag{1.4}$$

The equation (1.4) is called in analytic geometry the **canonical equation of a parabola**; the rectangular coordinate system in which the given parabola has the equation (1.4) is called the **canonical coordinate system** (for that parabola).

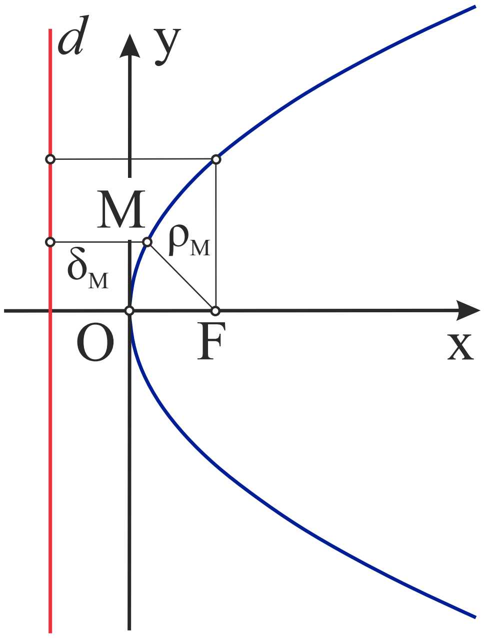


Figure 3

Now let's establish the geometric meaning of the coefficient p. To do this, we take the point

$$F = \left(\frac{p}{2}, 0\right), \tag{1.5}$$

called the **focus** of the parabola (1.4), and the line d, defined by equation

$$x = -\frac{p}{2} \tag{1.6}$$

This line is called the **directrix** of the parabola (1.4) (see Fig. 3).

Let $M = (x, y)$ be an arbitrary point of the parabola (1.4). It follows from the equation (1.4) that $x \geqslant 0$. Therefore, the distance of the point M from the directrix d is the number

$$\delta_M = \frac{p}{2} + x. \tag{1.7}$$

The distance of M from the focus F is

$$r = \rho(F, M) = \rho_M = +\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = +\sqrt{x^2 - px + \frac{p^2}{4} + y^2}. \tag{1.8}$$

But $y^2 = 2px$, therefore

$$r = +\sqrt{x^2 - px + \frac{p^2}{4} + 2px} = x + \frac{p}{2} = \delta_M.$$

So, all points M of the parabola are equidistant from its focus and directrix:

$$r = \delta_M. \tag{1.9}$$

Conversely, each point M that satisfies the condition (1.9) lies on the parabola (1.4).

Indeed,

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

Therefore,

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

or

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

and, after opening the brackets and collecting of similar terms,

$$y^2 = 2px.$$

> We have proved that *each parabola (1.4) is the locus of points that are equidistant from the focus F and from the directrix d of this parabola.*
>
> At the same time, we also established the *geometric meaning of the coefficient p in the equation (1.4): the number p is equal to the distance between the focus and the directrix of the parabola.*

Now let a point F and a line d not passing through this point be given arbitrarily on the plane. Let us prove that there exists a parabola with focus F and directrix d. To do this, draw a line g through the point F (Fig. 4), perpendicular to the line d; the point of intersection of both lines will be denoted by D; the distance $|DF|$ (that is, the distance between the point F and the line d) will be denoted by p. We turn the line g into an axis, taking the direction $\overrightarrow{DF}$ on it as positive. Let's make this axis the x-axis of a rectangular coordinate system whose origin is the midpoint O of the segment DF and the straight line d gets the equation $x = -\dfrac{p}{2}$.

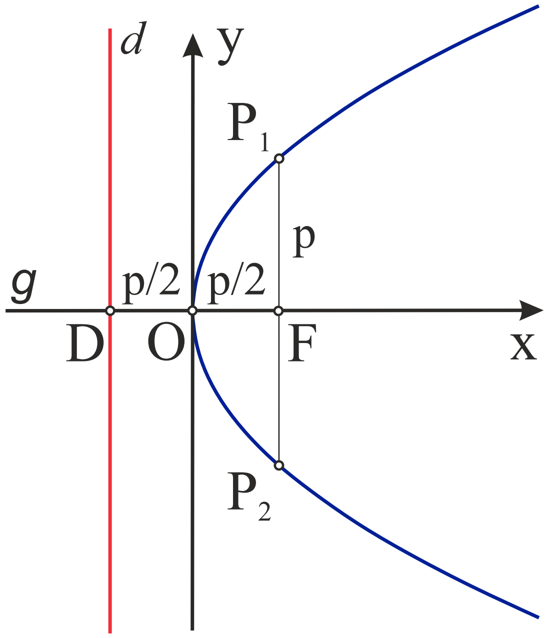


Figure 4

Now we can write the canonical equation of parabola in the chosen coordinate system:

$$y^2 = 2px,$$

where the point F is the focus, and the line d is the directrix of the parabola (1.4).

We established above that a parabola is the locus of points M that are equidistant from the point F and the line d. So, we can give such a geometric (i.e., independent of any coordinate system) definition of a parabola.

Definition. *A parabola is the locus of points equidistant from some fixed point (the "focus" of the parabola) and some fixed line (the "directrix" of the parabola).*

Denoting the distance between the focus and the directrix of the parabola as p, we can always find a rectangular coordinate system that is canonical for the given parabola, i.e. one in which the equation of the parabola has a canonical form (1.4).

Conversely, any curve that has such an equation in some rectangular coordinate system is a parabola (in the geometric meaning just established).

The distance p between the focus and the directrix of the parabola is called **focal parameter**.

The line passing through the focus perpendicular to the parabola's directrix is called its **focal axis** (or simply **axis**); it is the axis of symmetry of the parabola — this follows from the fact that the axis of the parabola is the x-axis in the coordinate system with respect to which the parabola's equation is (1.4). If the point $M = (x, y)$ satisfies the equation (1.4), then the point $M' = (x, -y)$, which is symmetric to the point M with respect to the x-axis, also satisfies this equation.

The point of intersection of a parabola with its axis is called the **vertex** of the parabola; it is the origin of the coordinate system that is canonical for the given parabola.

Let us give one more geometric interpretation of the focal parameter of parabola.

Let us draw a straight line through the focus of the parabola perpendicular to the axis of the parabola; it intersects the parabola at two points P_1 and P_2 (see Fig. 4) and defines the so-called **focal chord** P_1P_2 of the parabola (that is, the chord passing through focus parallel to the directrix of the parabola[3]). *Half the length of the focal chord is the focal parameter of the parabola.*

Indeed, half the length of the focal chord is the absolute value of the ordinate of any of the points P_1, P_2, the abscissa of each of which is equal to the abscissa of the focus, i.e. $p/2$. Therefore, for the ordinate y of each of the points P_1, P_2 we have

$$y^2 = 2p \cdot \frac{p}{2} = p^2, \qquad |y| = p,$$

which was to be proved.

2 Ellipse

2.1 Definition and canonical equation of an ellipse

Definition. *An* ***ellipse*** *is a locus of points in a plane whose distances from two fixed points F_1 and F_2 (Fig. 5) have a constant sum; we denote this sum by $2a$. The points F_1 and F_2 are called the* ***foci*** *of the ellipse; the distance between them is denoted by $2c$ and is called the* ***focal distance****. The line on which the foci of the ellipse lie is called the* ***first (or focal) axis****. The point O on the focal axis midway between the foci is the* ***center*** *of the ellipse. A straight line passing through the center of the ellipse perpendicular to the focal axis is called the* ***second axis*** *of the ellipse. As with circles and parabolas, a line segment with endpoints on an ellipse is a* ***chord*** *of the ellipse. The chord lying on the focal axis is the* ***major axis*** *of the ellipse. The chord through the center perpendicular to the focal axis is the* ***minor axis*** *of the ellipse.*

Let M be some point of the ellipse. Since

$$2a = |F_1M| + |F_2M| \geqslant |F_1F_2| = 2c,$$

[3]Or perpendicular to the axis.

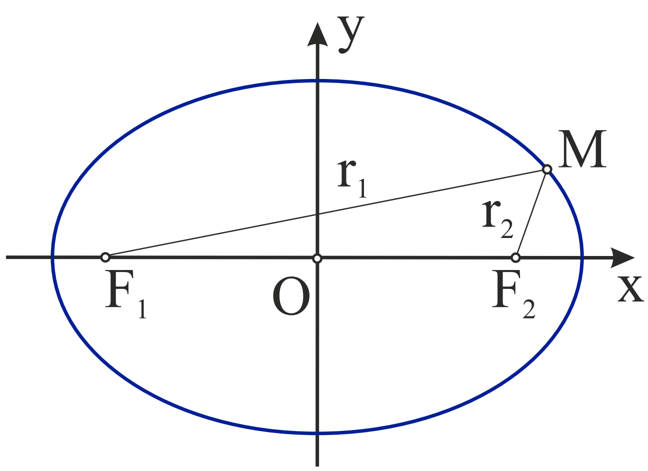


Figure 5

then $a \geqslant c$. However, if $a = c$, then we get the set of all points M for which

$$|F_1M| + |F_2M| = |F_1F_2|,$$

i.e., the segment $|F_1F_2|$. We will not consider this case further and therefore we will assume that $a > c$.

Definition. *Number* $e = \dfrac{c}{a}$ *called the eccentricity of the ellipse; it is always* < 1.

The eccentricity of the ellipse is zero if and only if the foci of the ellipse coincide: $F_1 = F_2$. In this case, the ellipse becomes the locus of points M, whose distance from the point $F_1 = F_2$ is equal to a, i.e., into a circle of radius a with center $O = F_1 = F_2$; by an axis of a circle we mean any straight line passing through its center O.

Let an ellipse be given; hence, its foci F_1 and F_2 are given, and its **semi-major axis** a is given. We also know the number $c < a$, which is equal to half the distance between the foci.

Let's construct a rectangular coordinate system on the plane, which we will call the **canonical system** for the given ellipse. Its origin O is the center of the ellipse, and the x-axis coincides with the focal axis. The positive direction of it is the direction of the vector $\overrightarrow{F_1F_2}$. The positive direction on the y-axis is chosen arbitrarily.

In this coordinate system we have $F_1 = (-c, 0)$, $F_2 = (c, 0)$; focus F_1 is conditionally called **left** focus, focus F_2 — **right** focus.

Suppose now that $M = (x, y)$ is an arbitrary point of the ellipse. Let $r_1 = \rho(F_1, M)$ and $r_2 = \rho(F_2, M)$ be the distances of the point M to the foci F_1 and F_2, respectively. The numbers r_1 and r_2 are called the **focal radii** of the point

M. We have:

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

The point $M = (x, y)$ is a point of the ellipse if and only if

$$r_1 + r_2 = 2a, \tag{2.1}$$

i.e.,

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a. \tag{2.2}$$

This equation (in which both roots are positive) is the equation of the ellipse in the chosen coordinate system.

Let's transform the equation (2.1) into a form called the **canonical equation of an ellipse**. To do this, we move the second radical to the right side. After squaring both sides of the equation, we get

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

or (after opening the brackets and collecting similar terms)

$$a\sqrt{(x-c)^2 + y^2} = a^2 - cx. \tag{2.3}$$

We square both sides of the equation (2.3) again and get

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

or (after obvious simplifications)

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2). \tag{2.4}$$

Since $a > c$, the number $a^2 - c^2$ is positive; denote it by b^2, calling **the number $b = +\sqrt{a^2 - c^2}$ the semi-minor axis of the ellipse**. Now the equality (2.4) can be rewritten as

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

or

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.} \tag{2.5}$$

So far, we have only proved that every point $M = (x, y)$ that satisfies the equation (2.2) also satisfies the equation (2.5). Let us now show that the equation (2.5) is indeed the equation of our ellipse. It remains to prove the converse, namely, that each point $M = (x, y)$ that satisfies the equation (2.5) is a point of the ellipse, i.e., that it satisfies the condition

$$r_1 + r_2 = 2a;$$

this is not obvious, since when going from the equation (2.2) to the equation (2.5), we squared both sides of the equation twice. So, let $M = (x, y)$ be an arbitrary point satisfying the equation (2.5). Let's find the distances r_1, r_2 of the point M from the foci F_1 and F_2. We have

$$r_1 = +\sqrt{(x+c)^2 + y^2}, \tag{2.6}$$

and from (2.5) we have

$$y^2 = b^2 \left(1 - \frac{x^2}{a^2}\right). \tag{2.7}$$

But $b^2 = a^2 - c^2$; substituting this into (2.7), we get

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

We substitute this y value into (2.6) and get

$$r_1 = +\sqrt{2cx + a^2 + \frac{c^2}{a^2} x^2} = +\sqrt{\left(a + \frac{c}{a} x\right)^2},$$

whence

$$r_1 = \pm \left(a + \frac{c}{a} x\right) = \pm (a + ex). \tag{2.8}$$

The left side of the equation is a positive number r_1; on the right, one must take such a sign that the right-hand side is also positive. But from (2.5) it follows that $|x| \leqslant a$; moreover, $0 \leqslant e < 1$; hence, $|ex| < a$, i.e. always $a + ex > 0$, so on the right in (2.8) we must take the $+$ sign and we get

$$\boxed{r_1 = a + ex.} \tag{2.9}$$

The same way we get

$$r_2 = +\sqrt{(x-c)^2 + y^2}. \tag{2.10}$$

This equality differs from (2.6) only in that it has $-c$ instead of c. Therefore, without repeating calculations similar to those just carried out, we can get the value for r_2 immediately by replacing c in (2.8) with $-c$. We get

$$r_2 = \pm (a - ex). \tag{2.11}$$

Again we have $|ex| < a$, so always $a - ex > 0$; on the right side of the equation, we must again take the sign $+$:

$$\boxed{r_2 = a - ex.} \tag{2.12}$$

From (2.9) and (2.12) we get $r_1 + r_2 = 2a$, the point $M = (x, y)$ belongs to our ellipse.

So, we have proved that the equation (2.5) is indeed the equation of an ellipse; it is called the **canonical equation of an ellipse**.

The formulas (2.9) and (2.12) very simply (and even linearly) express the focal radii of any point of the ellipse through the x-coordinate of this point.

In addition, note that $b^2 = a^2 - c^2$ implies $c^2 = a^2 - b^2$; it means that

$$e^2 = \frac{c^2}{a^2} = 1 - \frac{b^2}{a^2}, \qquad \frac{b^2}{a^2} = 1 - e^2; \tag{2.13}$$

$$e = \sqrt{1 - \frac{b^2}{a^2}} = \frac{\sqrt{a^2 - b^2}}{a}, \qquad \frac{b}{a} = \sqrt{1 - e^2}. \tag{2.14}$$

We will need these formulas later.

From the equation (2.5), some properties of the ellipse are easily deduced. First of all, if the point $M = (x, y)$ lies on our ellipse, i.e. it satisfies the equation (2.5), then the point $M' = (x, -y)$ (Fig. 6) is symmetrical to the point M relative to the x-axis (i.e. the focal axis of the ellipse), as well as the point $M'' = (-x, y)$ is symmetrical to the point M with respect to the y-axis. So, **both axes of an ellipse are its axes of symmetry**.

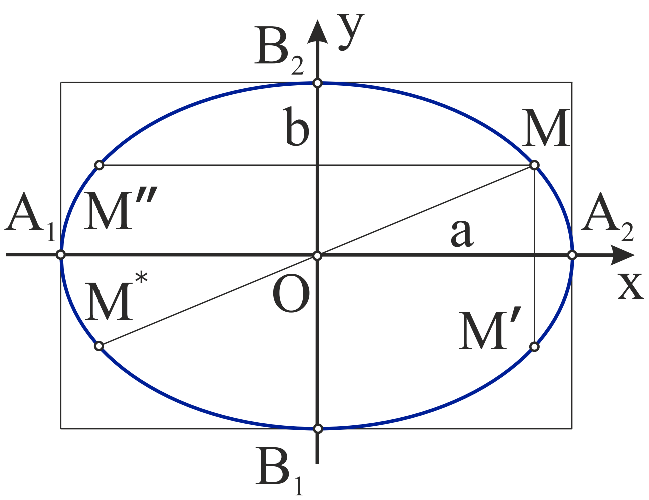


Figure 6

The center of an ellipse is its center of symmetry: indeed, with our choice of coordinate system, the center is the origin O; if the point $M = (x, y)$ satisfies the equation (2.5), then the point $M^* = (-x, -y)$, which is symmetric to the point M with respect to the center O , also satisfies the equation (2.5).

Finally, note that, due to the equation (2.5), which is satisfied by all points of the ellipse, for each point $M = (x, y)$ of the ellipse we have

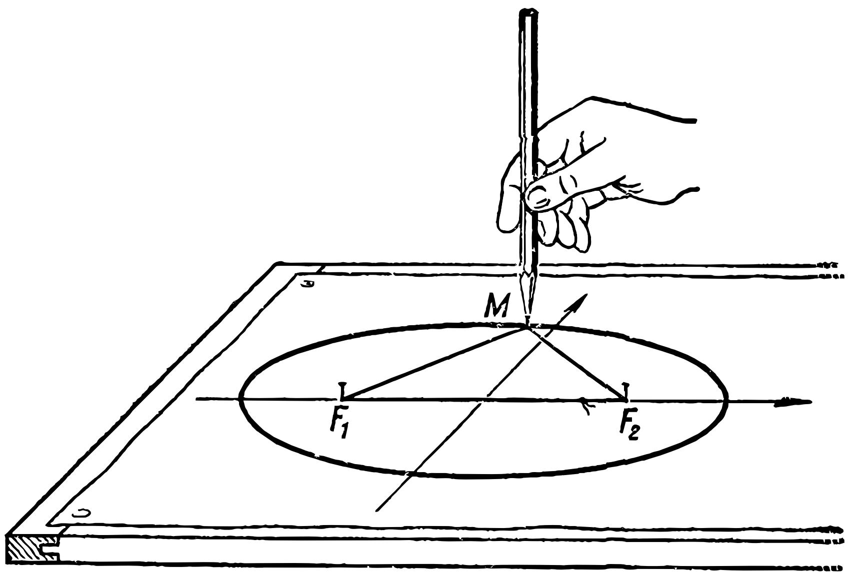
$$\frac{x^2}{a^2} \leqslant 1, \qquad \frac{y^2}{b^2} \leqslant 1,$$

i.e., $|x| \leqslant a$, $|y| \leqslant b$. This means that the entire ellipse lies in the rectangle bounded by the lines $x = \pm a$, $y = \pm b$, parallel to the axes of the ellipse and separated from them by distances a and b, respectively. This rectangle is called the **fundamental rectangle** for the given ellipse.

The points $A_1 = (-a, 0)$, $A_2 = (a, 0)$, as well as the points $B_1 = (0, -b)$, $B_2 = (0, b)$, i.e. points of intersection of an ellipse with its axes are called **ellipse vertices.**

Thus, an ellipse (which is not a circle) has four vertices.

Remark. Directly from the definition of an ellipse follows the well-known method of its construction. Two nails are driven into the drawing board at the points F_1 and F_2 (with a distance $2c$ between them). A closed thread of length $2a + 2c$ is thrown over them; pulling this thread by attaching the tip of a pencil to it, move the pencil so that the thread is taut all the time. In this case, the pencil will draw an ellipse as the locus of points, the sum of the distances of which from the two foci F_1 and F_2 is obviously equal to $2a$.



2.2 Parametric equation of an ellipse; building an ellipse from points

Let an ellipse be given in the plane by its canonical equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{2.15}$$

Construct two circles ("large" and "small") with a common center at the origin and radii a for the "large circle" and b for the "small circle" (Fig. 7). Draw an arbitrary ray from the origin. The angle of its inclination to the x-axis will be denoted by φ. The point of intersection of our ray with the "large circle" will be denoted by L, and with the "small circle" by K. Through the points K and L we draw lines k and l parallel to the y-axis (crossing the x-axis at the points K' and L'). Through the point K we draw a line parallel to the axis Ox. The point of intersection of this line with the line l is denoted by M. For $\varphi = 0$ and $\varphi = \pi$ we get $M = L$; for $\varphi = \pi/2$, $\varphi = 3\pi/2$ we get $M = K$; they are the vertices of the ellipse. In general, for $0 < \varphi < \pi$, our construction gives some point of the upper semiplane, and for $\pi < \varphi < 2\pi$, a point of the lower semiplane, so that on each vertical line $x = m$, $-a < m < a$, we get exactly two points of M.

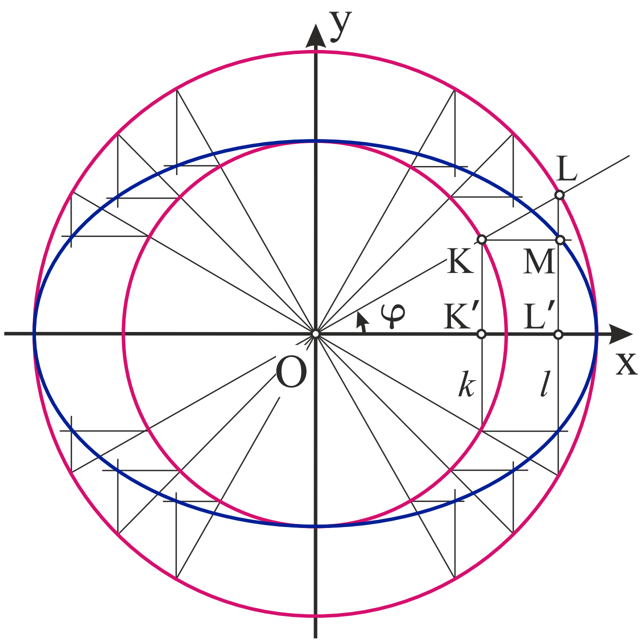


Figure 7

Let's prove that each of these points lies on the ellipse (2.15). Indeed, for the coordinates x, y of the point M we have

$$\begin{aligned} x &= |OL'| = |OL| \cos\varphi = a\cos\varphi, \\ y &= |L'M| = |K'K| = |OK| \sin\varphi = b\sin\varphi, \end{aligned} \tag{2.16}$$

so

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

which proves the statement.

Let us prove that our construction — and hence the set of equations (2.16) — for some φ gives us any point of the ellipse.

Indeed, let $N = (x_0, y_0)$ be some point of the ellipse, $-a < x_0 < a$. On the line $x = x_0$, our construction gives us two points M, and they, satisfying the equation (2.15), lie on our ellipse. If none of them coincided with the point N, then there would be three points of the ellipse on the line $x = x_0$, which cannot be, since, assuming in the equation (2.15) $x = x_0$, we obtain a quadratic equation to determine the ordinate y_0 of the intersection point of the line $x = x_0$ with the ellipse (2.15).

We have proved the following proposition:

All points of the ellipse (2.15) and only these points satisfy the system of equations

$$\boxed{\left.\begin{aligned} x &= a\cos\varphi, \\ y &= b\sin\varphi, \end{aligned}\right\} \quad 0 \leqslant \varphi < 2\pi.} \tag{2.17}$$

This system of equations is called **"parametric equation of an ellipse"** (although there are two equations, not one).

At the same time, we have given a way to construct an ellipse "by points", or rather, a way to construct an arbitrarily large number of points of an ellipse. But we have proved, in addition, one important and visual property of the ellipse. Indeed, let's go back to Fig. 7. The point $M = (x, y)$ is an arbitrary point of the ellipse, and $L = (x, Y)$ is a point of the "large circle" having the same x-coordinate, i.e., it lying on the same vertical line as the point M. Then $Y = |L'L|$ and $y = |L'M| = |K'K|$. From the similarity of triangles OKK' and OLL' we have

$$\frac{|K'K|}{|L'L|} = \frac{|OK|}{|OL|} = \frac{b}{a},$$

i.e., $y = \dfrac{b}{a}Y$. Thus, our ellipse was obtained from a "large" circle by a plane transformation, which means that all points of the x-axis remain in place, and each point of P (Fig. 8) that does not lie on this axis, goes to point P^* with the same x-coordinate, but with the y-coordinate obtained from y-coordinate of point P multiplied by the number b/a, — the ordinate of each point "compressed" with respect to b/a. Such a transformation is called **uniform compression of the plane to the x-axis in relation b/a.**

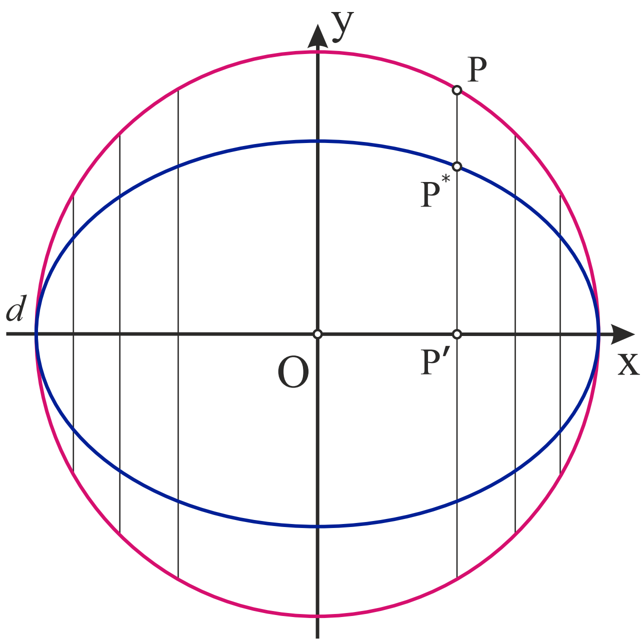


Figure 8

2.3 Ellipse as a result of compressing a circle to one of its diameters

In general, **uniform compression of a plane to a given line** d **with respect to** $k > 0$ is a plane transformation consisting of the following: each point of the line d remains in place; an arbitrary point P not lying on the line d passes to a point P^* lying on the perpendicular PP' to the line d (see Fig. 8) by the same side of this line as the point P and defined by the condition

$$|P'P^*| = k|P'P|,$$

where P' is the base of the perpendicular, i.e., its intersection with the line d.

The term "compression" is clearly justified only when $k < 1$ (see Fig. 8); for $k > 1$ it is not compressing but "stretching". However, we will use the term "compression" for any k. You can prove for yourself that our ellipse can be obtained from a small circle by "compressing" it it with respect to a/b (that is, more precisely, by stretching it) to the y-axis.

So the following proposition has been proven:

Any ellipse is obtained by compressing a circle to one of its diameters[4] (see Fig. 8).

[4]This diameter itself turns out to be the major axis of the ellipse for $k < 1$ (then $k = b/a$), and for $k > 1$ it becomes the minor axis ($k = a/b$).

2.4 Ellipse as a projection of a circle and as a section of a round cylinder

Let two planes α and β be given. They intersect along some line d (Fig. 9). The acute angle between the planes α and β will be denoted by φ.

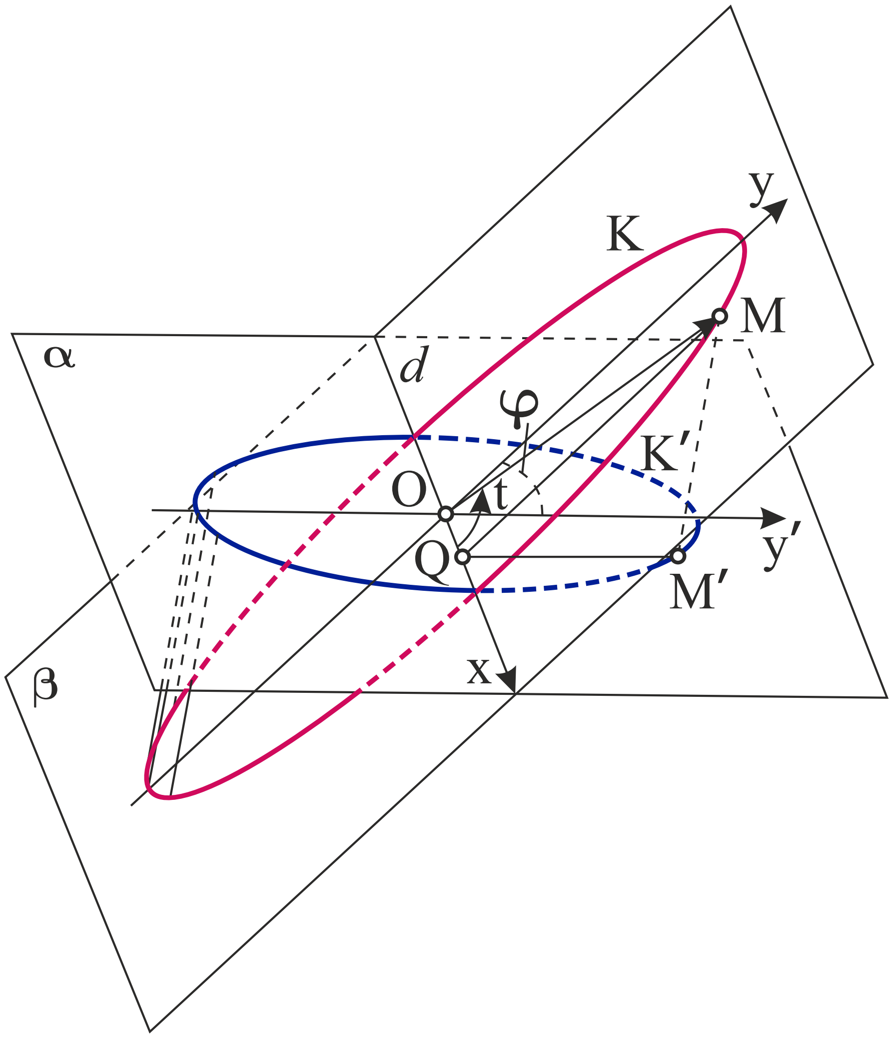


Figure 9

A circle K of radius a with center O, which (for convenience) we assume to lie on the line d, is given in the plane β. Let us prove that **the orthogonal projection K' of the circle K onto the plane α is an ellipse.**

To prove this, we choose in each plane α and β rectangular coordinate systems with a common origin O and a common x-axis that goes along the line d. We direct the y-axes in the α and β planes so that the angle between their unit vectors is equal to the angle φ.

Let M be an arbitrary point of the circle K; we denote its projection onto the plane α by M'. The angle of slope of the vector $\overrightarrow{OM}$ (in the β plane) to the x-axis will be denoted by t. Denote by Q the orthogonal projection of the point

M' (and the point M) onto the x-axis. Then for the coordinates x, y of the point M' on the plane α, we have

$$
\begin{aligned}
x &= |OQ| = |\overrightarrow{OM}|\cos t = a\cos t,\\
y &= |QM'| = |\overrightarrow{QM}|\cos\varphi = |\overrightarrow{OM}|\sin t\cos\varphi = a\cos\varphi\sin t.
\end{aligned}
$$

Denoting the constant $a\cos\varphi$ by b, we have

$$
\begin{aligned}
x &= a\cos t,\\
y &= b\sin t,
\end{aligned}
$$

i.e., we have a parametric form of the ellipse equation, which is satisfied by the coordinates of any point M' of our curve K' (defined as the projection of the circle K onto the plane α). Therefore, the curve K' is an ellipse, and our assertion is proved.

Let us give a rigorous proof of the fact that **the section of a circular cylinder by a plane not parallel to the axis of the cylinder is an ellipse.**

The cutting plane will be denoted by α, the point of its intersection with the axis of the cylinder will be denoted by O (Fig. 10). We will make this point the origin of the rectangular coordinate system $Ox'y$, whose y-axis goes along the line of intersection d passing through the point O perpendicular to the axis of the cylinder. Denoting by Ox the projection of the Ox' axis onto the β plane, we obtain a rectangular coordinate system Oxy in the β plane. The acute angle between the planes α and β is denoted by φ.

The plane β intersects the cylinder along the circle K (which is an orthogonal projection onto the plane β of the curve K', which is the curve of intersection of the cylinder with the plane α).

Let M' be an arbitrary point of the curve K', and M be its projection onto the plane β (the point M lies on the circle K of radius b). The points M and M' lie on a line perpendicular to the plane β. If we draw the plane through this line, perpendicular to the y-axis, we obtain, at the intersection with the latter, the point Q, which is the projection onto the Oy of both points M and M'. The angle of slope of the vector $\overrightarrow{OM}$ to the Ox will be denoted by t. Then we have for the coordinates x', y of the point M'

$$
\begin{aligned}
y &= |OQ| = |\overrightarrow{OM}|\sin t = b\sin t,\\
x' &= |QM'| = \frac{|\overrightarrow{QM}|}{\cos\varphi} = \frac{|\overrightarrow{OM}|\cos t}{\cos\varphi} = \frac{b}{\cos\varphi}\cos t.
\end{aligned}
\tag{2.18}
$$

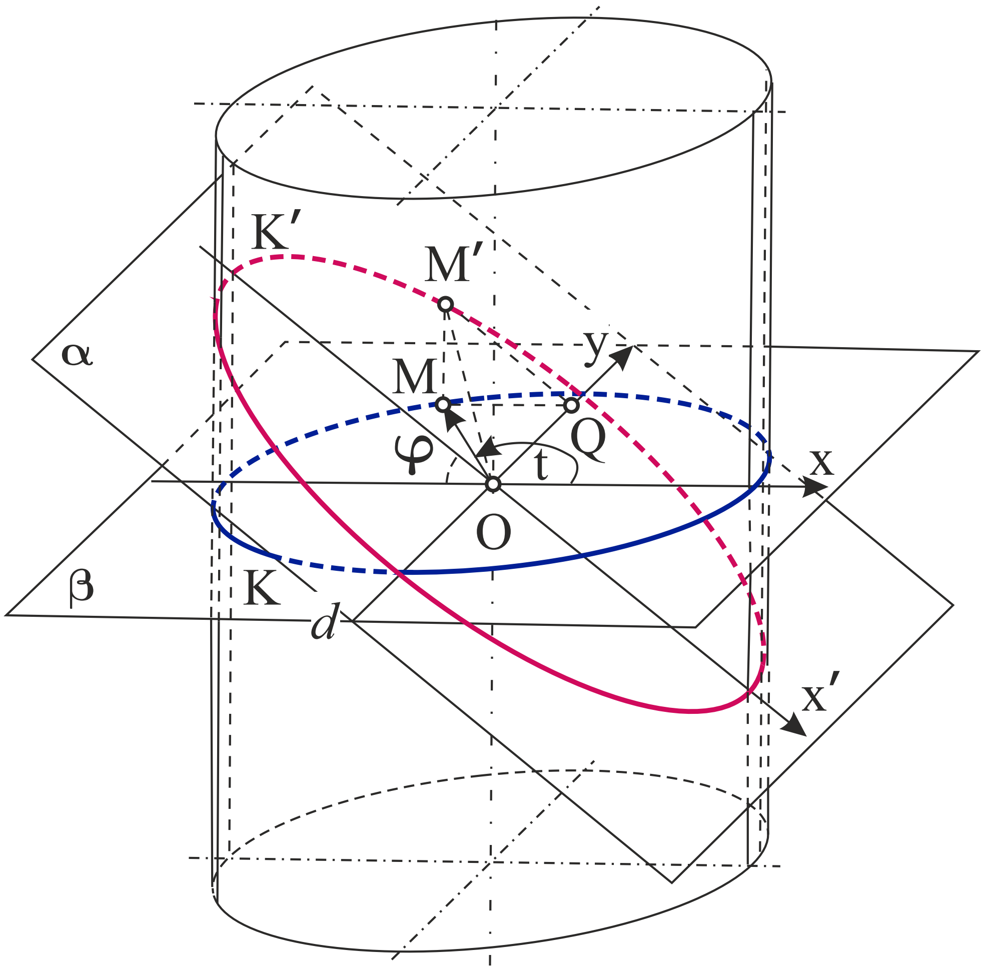


Figure 10

Let's denote the constant $\dfrac{b}{\cos\varphi}$ by a, which, when substituted into (2.18), gives

$$x' = a\cos t,$$
$$y = b\sin t,$$

we again obtain the parametric equation of an ellipse as the equation of the curve K'. The theorem has been proven.

3 Hyperbola

3.1 Definition and canonical equation of a hyperbola

Definition. *A **hyperbola** is a locus of points in a plane, the absolute difference between the distances of each of them to two fixed points F_1 and F_2 (Fig. 11) is a positive constant. We denote this constant by $2a$. The points F_1 and F_2 are called the **foci** of the hyperbola. The distance between them is denoted by $2c$ and is called the **focal distance**. As in the case of an ellipse, the midpoint of the segment F_1F_2 is called the **center** of the hyperbola. The line on which the foci of the hyperbola lie is called the **focal or first (or real) axis** of the hyperbola. The line passing through the center perpendicular to the focal axis of the hyperbola is called its **second (or imaginary) axis**.*

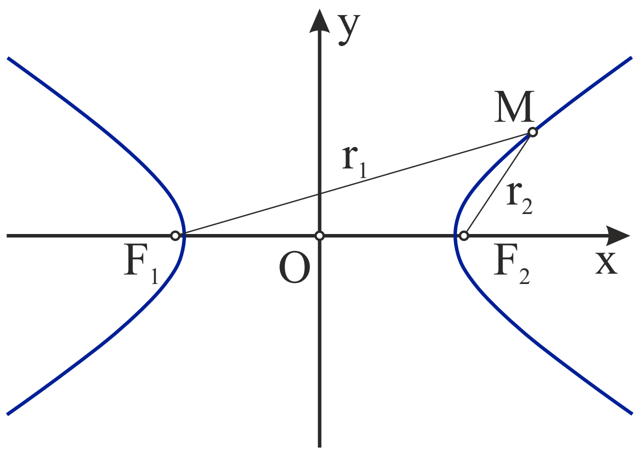


Figure 11

From Fig. 11 it is clear that

$$|F_1F_2| \geqslant ||MF_1| - |MF_2||,$$

i.e. $c \geqslant a$.

If $c = a$, then we get points M for which either

$$|MF_1| - |MF_2| = |F_1F_2|$$

or

$$|MF_2| - |MF_1| = |F_1F_2|.$$

These M points fill two half-lines that complement the segment F_1F_2 to the entire line. Therefore, we will not consider the case $c = a$. We assume that $c > a$. As in the case of an ellipse, we call the number c/a the **eccentricity of the hyperbola** and denote it by e. We have

$$e = \frac{c}{a} > 1.$$

Let a hyperbola be given, i.e., we have its foci F_1 and F_2, as well as the numbers a and c. As in the case of an ellipse, we construct a rectangular coordinate system on the plane, which we will call **canonical** for the given hyperbola. The origin of this coordinate system lies at the center O of the hyperbola, the x-axis coincides with the focal axis of the hyperbola. For the positive direction of the x-axis, we take the direction of the vector $\overrightarrow{F_1F_2}$. Then $F_1 = (-c, 0)$, $F_2 = (c, 0)$.

Let $M = (x, y)$ be an arbitrary point of the hyperbola. Denote by $r_1 = \rho(F_1, M)$ and $r_2 = \rho(F_2, M)$ the distances of the point $M = (x, y)$ to the foci F_1 and F_2, respectively. The numbers r_1 and r_2 are called the **focal radii** of the point M.

We have

$$r_1 = \sqrt{(x+c)^2 + y^2}, \qquad r_2 = \sqrt{(x-c)^2 + y^2}. \tag{3.1}$$

A point $M = (x, y)$ is a point of a hyperbola if and only if

$$|r_1 - r_2| = 2a \tag{3.2}$$

or

$$r_1 - r_2 = \pm 2a.$$

If we take into account the equalities (3.1), then we have (assuming square roots are positive)

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \pm 2a. \tag{3.3}$$

This is the equation of our hyperbola in the chosen coordinate system.

Let's transform the equation (3.3) into a form called **canonical**. To do this, we move the second radical to the right side. Then squaring both sides of the equation gives

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

or (after simple transformations)

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

We square again both parts of the last equality and get

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

or

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2). \tag{3.4}$$

Since $c > a$, the number $c^2 - a^2$ is positive; denote it by b^2, assuming $b = +\sqrt{c^2 - a^2}$. The equality (3.4) can be rewritten as

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

or

$$\boxed{\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.} \tag{3.5}$$

It remains to show that the equation (3.5) is indeed the equation of our hyperbola: so far we have only proved that each point $M = (x, y)$ satisfying the equation (3.2) also satisfies the equation (3.5); as in the case of an ellipse, we still need to prove that every point $M = (x, y)$ that satisfies the equation (3.5) also satisfies the equation (3.1).

Let $M = (x, y)$ be an arbitrary point satisfying the equation (3.5). Let us find the focal radii r_1 and r_2 of the point M. We have

$$r_1 = \sqrt{(x + c)^2 + y^2}. \tag{3.6}$$

From (3.5) we find (remembering that $b^2 = c^2 - a^2$)

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

Substituting this expression for y^2 into (3.6) and taking into account that $e = c/a$ we have

$$r_1 = \sqrt{x^2 + 2eax + a^2e^2 + e^2x^2 - x^2 - a^2e^2 + a^2} = \sqrt{(a + ex)^2},$$

i.e.

$$r_1 = \pm(a + ex). \tag{3.7}$$

In a similar manner we have

$$r_2 = \pm(a - ex). \tag{3.8}$$

Since r_1 and r_2 are positive numbers, we need to choose the sign before the brackets so that the right parts of the equalities (3.7) and (3.8) are positive. To do this, we examine various possible cases represented by the equalities (3.7) and (3.8). From the equation (3.5) we find first of all that $|x| \geqslant a > 0$. Therefore, we have two main cases: depending on whether the point $M = (x, y)$ lies in the right half-plane $x > 0$ or in the left half-plane $x < 0$. Since $e > 1$, in both cases we have

$$|ex| > a. \tag{3.9}$$

For $x > 0$, the number inside the bracket in (3.7) is positive, so the bracket must be taken with the $+$ sign, and we get

$$\boxed{r_1 = a + ex \text{ if } x > 0.} \tag{3.10}$$

It follows from (3.9) that if $x > 0$ there is a negative number inside the bracket (3.8), the bracket must be taken with the $-$ sign, so that

$$\boxed{r_2 = -a + ex \text{ if } x > 0.} \tag{3.11}$$

It follows from (3.10) and (3.11) that for $x > 0$ we have

$$r_1 - r_2 = 2a,$$

and the point $M = (x, y)$, which satisfies the equation (3.5), lies on a hyperbola.

Let $x < 0$; from (3.9) it follows that now inside the bracket (3.7) there is a negative number, which means that we need to take the $-$ sign before the bracket, so

$$\boxed{r_1 = -a - ex \text{ if } x < 0.} \tag{3.12}$$

But number inside the bracket in (3.8) now is positive, which means

$$\boxed{r_2 = a - ex \text{ if } x < 0.} \tag{3.13}$$

We have

$$r_2 - r_1 = 2a.$$

So, in all cases, every point that satisfies the equation (3.5) lies on the hyperbola. We have proved that the equation (3.5) is indeed the equation of our hyperbola. It's called the **canonical equation of a hyperbola**.

Formulas (3.10), (3.12), (3.11) and (3.13) linearly express the focal radii of any point of the hyperbola through its abscissa.

Noting that $c^2 = a^2 + b^2$, we get

$$e^2 = \frac{c^2}{a^2} = \frac{a^2 + b^2}{a^2} = 1 + \left(\frac{b}{a}\right)^2, \tag{3.14}$$

$$\frac{b}{a} = \sqrt{e^2 - 1}, \qquad e = \sqrt{1 + \left(\frac{b}{a}\right)^2}. \tag{3.15}$$

From the equation (3.5) it follows (as in the case of an ellipse) that both axes of the hyperbola are its axes of symmetry, and the center of the hyperbola is its center of symmetry.

Rewriting the equation of the hyperbola (3.5) as

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

and noticing that its left-hand side is always $\geqslant 0$, we see that for the points of the hyperbola there must be

$$\frac{x^2}{a^2} - 1 \geqslant 0,$$

i.e. $|x| \geqslant a$. In other words, in the region $-a < x < a$ limited by the lines $x = \pm a$ (this region is shaded in Fig. 12), in particular, the second axis $x = 0$ does not contain hyperbola points: they all lie either to the right of the line $x = a$ or to the left of the line $x = -a$, except for two points $A_1 = (-a, 0)$, $A_2 = (a, 0)$ lying on these lines themselves and being the intersection points of the hyperbola with its focal axis. These two points are called **vertices** of the hyperbola. A line segment with endpoints on a hyperbola is a **chord** of the hyperbola. The chord lying on the focal axis connecting the vertices is the **transverse axis** of the hyperbola. The length of the transverse axis is $2a$. The line segment of length $2b$ that is perpendicular to the focal axis and that has the center of the hyperbola as its midpoint is the **conjugate axis** of the hyperbola. The number a is the **semitransverse axis**, and b is the **semiconjugate axis**.

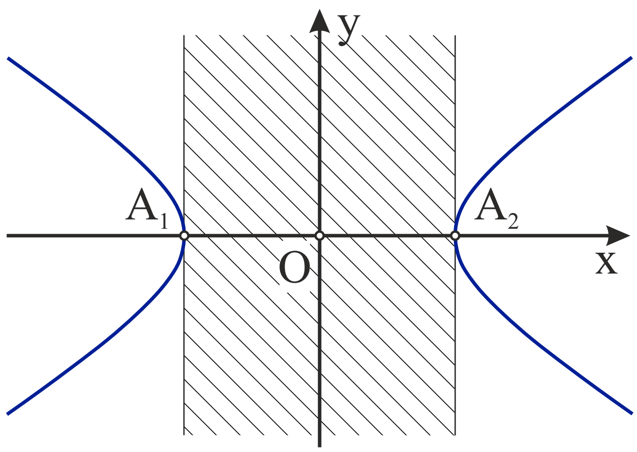
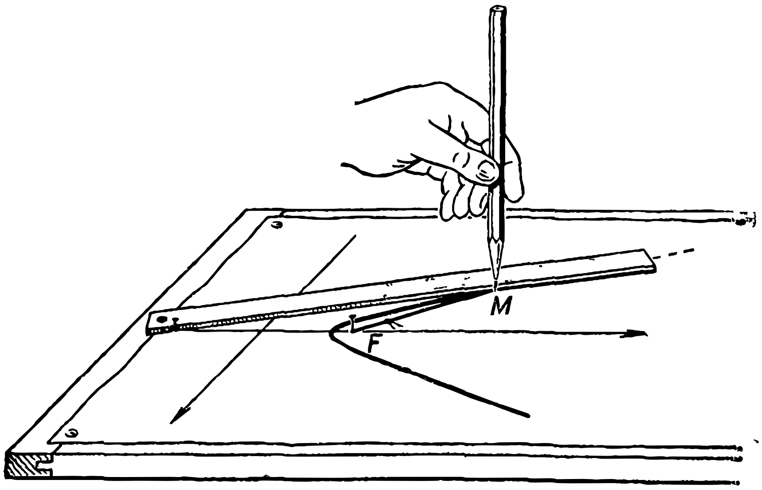


Figure 12

So, the hyperbola splits into two branches: “right”, for whose points the abscissa is $x \geqslant a$, and “left”, for whose points $x \leqslant -a$.

To get better acquainted with the general form of a hyperbola, it is necessary to define lines, called its asymptotes.

Remark. The definition of a hyperbola implies the following way of drawing it. We take a ruler whose length is greater than $2a$, and attach a thread of such length to one end of it so that the difference between the length of the ruler and the length of the thread is equal to $2a$. The second end of the ruler is fixed in one focus so that the ruler can freely rotate around it, and the second end of the thread is fixed in another focus. If you hold the thread with the tip of the pencil in the position stretched along the ruler, then when the ruler rotates, the pencil's tip also moves and describes the branch of the hyperbola, inside of which lies the focus in which the thread is fixed.



3.2 Fundamental rectangle and hyperbola asymptotes

As in the case of an ellipse, the **fundamental (or central) rectangle** of a hyperbola is a rectangle limited by lines parallel to the second and first axes of the hyperbola and spaced respectively by distances a and b (Fig. 13). In the canonical coordinate system, the equations of these lines are $x = \pm a$, $y = \pm b$, while the equation of the hyperbola itself has the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{3.16}$$

Diagonals of the fundamental rectangle[5] are lines whose equations have the form (in the same coordinate system, canonical for a given hyperbola)

$$y = \pm\frac{b}{a}x. \tag{3.17}$$

These lines are called **asymptotes** of the hyperbola. The line $y = \dfrac{b}{a}x$ will be called the **first** asymptote, and the line $y = -\dfrac{b}{a}x$ — the **second** asymptote.

[5]The diagonal of a rectangle (and a polygon in general) is understood as the entire infinite line passing through two given (not adjacent) vertices of the polygon.

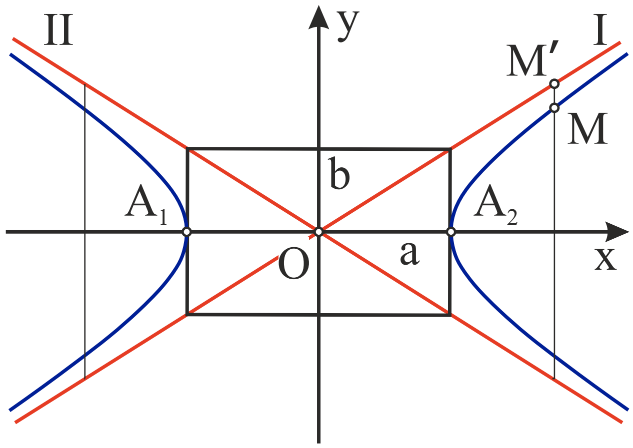


Figure 13

Let's take some value of the variable x, $x \geqslant a$. It corresponds in the upper half-plane to a point M of a hyperbola with abscissa x (see Fig. 13) and a point M' of the (first) asymptote with the same abscissa x:

$$M = (x, y), \qquad M' = (x, y').$$

In this case

$$y' = \frac{b}{a}x. \tag{3.18}$$

To find the ordinate y of the point M on the hyperbola, it is enough to solve the equation (3.16) with respect to y; we get

$$y = \frac{b}{a}\sqrt{x^2 - a^2}. \tag{3.19}$$

In this case (since we are in the upper half-plane), the radical must be taken with the + sign. Comparing the right-hand sides of the equalities (3.18) and (3.19), we first of all see that

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

In other words, the point M, having the same abscissa as the point M', has a smaller ordinate, i.e. lies under the point M'. Let's estimate the difference of ordinates:

$$0 < y' -' y = \frac{b}{a}(x - \sqrt{x^2 - a^2}) = \frac{b}{a}\frac{(x - \sqrt{x^2 - a^2})(x + \sqrt{x^2 - a^2})}{x + \sqrt{x^2 - a^2}} = \frac{ab}{x + \sqrt{x^2 - a^2}}.$$

As x increases without limit, the difference

$$y' - y = \frac{ab}{x + \sqrt{x^2 - a^2}}$$

while remaining positive, monotonically decreases and tends to zero, i.e., the points M and M', going to infinity (as their common coordinate x increases indefinitely), approach each other indefinitely. In this case, the point M of the hyperbola always remains under the point M' of the asymptote.

On the lower half-plane, the situation is similar[6]; as the (positive) coordinate x increases indefinitely, the point $M = (x, y)$ of the hyperbola (which has the abscissa x and lies in the lower half-plane) approaches the point $M' = (x, y')$ of the second asymptote indefinitely, at that

$$y = -\frac{b}{a}\sqrt{x^2 - a^2}, \qquad y' = -\frac{b}{a}x,$$

$$0 < y - y' = \frac{ab}{x + \sqrt{x^2 - a^2}} \to 0,$$

the point M lies above the point M'.

We have studied the mutual arrangement of the points of the hyperbola and the pair of its asymptotes for $x \geqslant a$. The picture for $x \leqslant -a$ is symmetric (since the figure consisting of a hyperbola and its two asymptotes is also symmetric with respect to the y-axis). In total, the general form of the hyperbola is clear from Fig. 13.

From the study of the relative position of the hyperbola and its asymptotes, it follows that the hyperbola does not have a single point in common with any of its asymptotes. This geometric fact can also be easily proved algebraically: if there were a common point $M = (x, y)$ of the hyperbola and its asymptotes, then the coordinates x, y of this point would have to simultaneously satisfy the equations (3.16) and (3.17), which are incompatible (substituting (3.17) into (3.16), we get $\frac{x^2}{a^2} - \frac{b^2}{a^2}\frac{x^2}{b^2} = 1$, i.e. $0 = 1$).

Remark. From the formulas (3.15) it follows that the eccentricity of the hyperbola

$$e = \sqrt{1 + \frac{b^2}{a^2}} = \frac{\sqrt{a^2 + b^2}}{a}$$

(always greater than one) is equal to the ratio of the length of the diagonal[7] of the fundamental rectangle to its base (i.e., to the side parallel to the focal axis of the hyperbola (Fig. 14)). The eccentricity is the smaller, the smaller the ratio of the height of the fundamental rectangle to its base, i.e. the sharper the angle between the asymptotes.

[6] That follows from the symmetry of the figure, composed of a hyperbola and a pair of its asymptotes, with respect to the x-axis.

[7] In an elementary geometric sense.

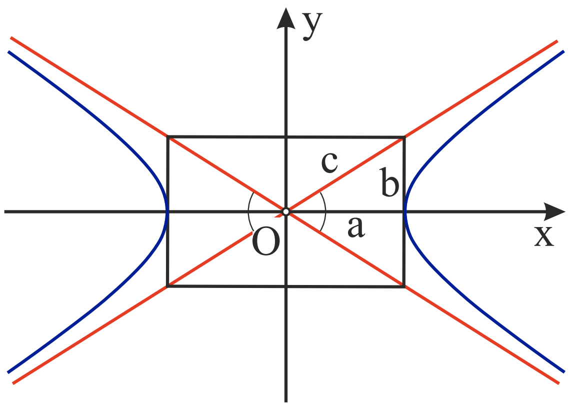


Figure 14

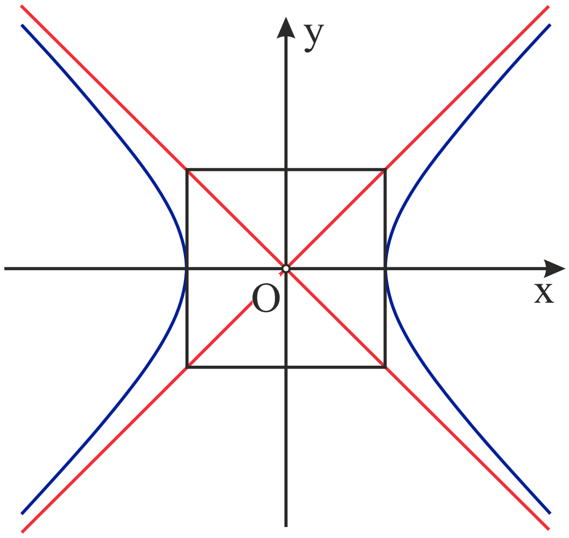


Figure 15

If $b = a$, i.e. if the main rectangle is a square (Fig. 15), then the eccentricity of the hyperbola is equal to $\sqrt{2}$ and the asymptotes are mutually perpendicular. In this case, the hyperbola is called **rectangular** or **equilateral**; its equation is

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

If we take the asymptotes of an equilateral hyperbola as the coordinate axes, then the equation of this hyperbola will be

$$xy = \frac{a^2}{2}$$

or, assuming $a^2/2 = k$,

$$y = \frac{k}{x},$$

— we get the equation of a hyperbola, like the "graph of inverse proportion" known from a high school course.

3.3 Parametric equation of a hyperbola

Let's rewrite the hyperbola equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{3.20}$$

as

$$\left(\frac{x}{a} - \frac{y}{b}\right)\left(\frac{x}{a} + \frac{y}{b}\right) = 1.$$

From here it is clear that

$$\frac{x}{a} - \frac{y}{b} \neq 0, \qquad \frac{x}{a} + \frac{y}{b} \neq 0.$$

Let's put

$$\frac{x}{a} + \frac{y}{b} = t;$$

then $t \neq 0$ and

$$\frac{x}{a} - \frac{y}{b} = \frac{1}{t};$$

hence,

$$\boxed{x = \frac{a}{2}\left(t + \frac{1}{t}\right), \qquad y = \frac{b}{2}\left(t - \frac{1}{t}\right).} \tag{3.21}$$

So, the coordinates of any point of the hyperbola can be represented as (3.21), where $t \neq 0$. Conversely, for any $t \neq 0$, the point with coordinates (3.21) lies on the hyperbola (3.20), which can be seen by substituting into the equation (3.20) instead of x and y their expressions from formulas (3.21). Therefore, the equations (3.21) are parametric equations of the hyperbola.

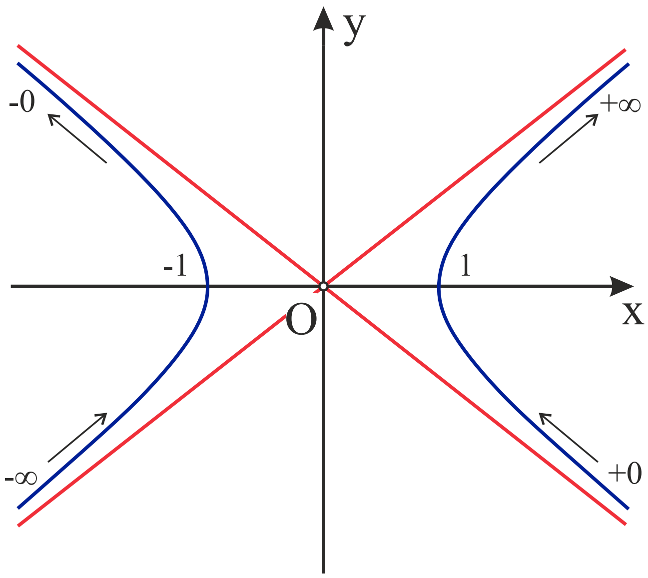


Figure 16

If the point $M(x, y)$ lies on the branch of the hyperbola whose points $x \geqslant a$, then

$$\frac{a}{2}\left(t + \frac{1}{t}\right) \geqslant 0, \quad \text{or} \quad \frac{a}{2}\frac{t^2 + 1}{t} \geqslant 0,$$

hence $t > 0$. Conversely, if $t > 0$, then $x \geqslant a$, i.e., the point $M(x, y)$ lies on the indicated branch of the hyperbola. When t changes in the half-interval $(0, 1]$, the value of x decreases from $+\infty$ to a, and the value of y increases from $-\infty$ to 0; when t in the interval $[1, +\infty)$ the value of x increases from a to $+\infty$, and the value of y increases from 0 to $+\infty$ (Fig. 16). For $t = 1$ we obtain the right vertex $(a, 0)$ of the hyperbola. For negative values of t, we get the left branch. Moreover, if t changes in the half-interval $(-\infty, -1]$, then the value of x increases from $-\infty$ to $-a$, and the value of y increases from $-\infty$ to 0, and if t changes in the half-interval $[-1, 0)$, then the value of x decreases from $-a$ to $-\infty$, and the value of y increases from 0 to $+\infty$.

Using the hyperbolic functions a parametric equation of the hyperbola (3.21) can be rewritten as

$$\boxed{\begin{cases} x = \pm a \cosh \tau, \\ y = b \sinh \tau, \end{cases} \quad -\infty < \tau < +\infty.} \tag{3.22}$$

3.4 Conjugated hyperbolas

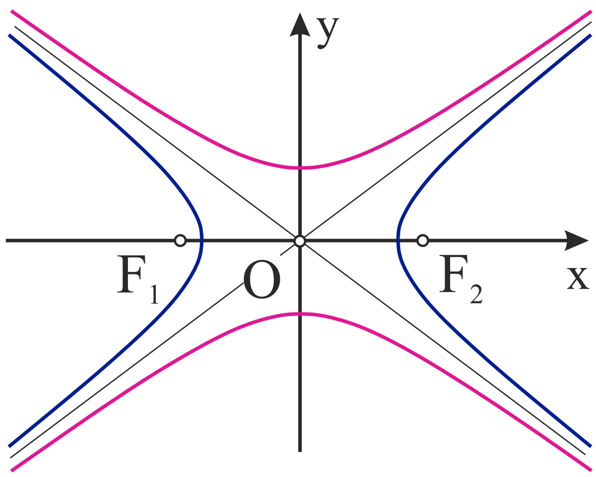


Figure 17

Two hyperbolas given by equations

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{3.23}$$

and

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1 \tag{3.24}$$

in the same Cartesian rectangular coordinate system with the same values of the semiaxes a and b are called **conjugate** (Fig. 17).

It was proved above that any hyperbola (3.23) can be expressed by parametric equations

$$x = \frac{a}{2}\left(t + \frac{1}{t}\right), \qquad y = \frac{b}{2}\left(t - \frac{1}{t}\right);$$

parametric equations of a hyperbola conjugate to a given one will be

$$x = \frac{a}{2}\left(t - \frac{1}{t}\right), \qquad y = \frac{b}{2}\left(t + \frac{1}{t}\right).$$

3.5 The equation of a hyperbola with respect to its asymptotes

Let a hyperbola be given in its canonical coordinate system $O\mathbf{e}_1\mathbf{e}_2$ (Fig. 18) by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{3.25}$$

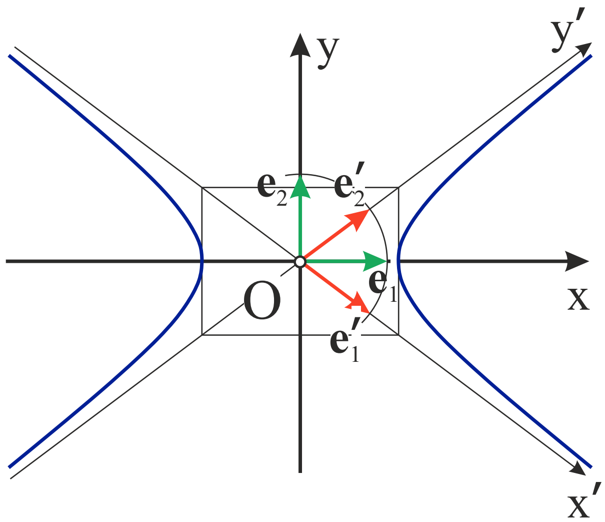


Figure 18

We pass to the new coordinate system $O\mathbf{e}'_1\mathbf{e}'_2$, where

$$\mathbf{e}'_1 = \left(\frac{a}{c}, -\frac{b}{c}\right), \quad \mathbf{e}'_2 = \left(\frac{a}{c}, \frac{b}{c}\right), \quad c = \sqrt{a^2 + b^2}$$

(i.e., we take as new unit vectors the orts directed along the asymptotes of the hyperbola). The direct transition matrix from the old basis to the new one will have a form

$$A = \begin{pmatrix} \dfrac{a}{c} & \dfrac{a}{c} \\ -\dfrac{b}{c} & \dfrac{b}{c} \end{pmatrix},$$

i.e.

$$x = \frac{a}{c}x' + \frac{a}{c}y',$$
$$y = -\frac{b}{c}x' + \frac{b}{c}y';$$

substituting these expressions in the equation (3.25), we obtain after obvious transformations

$$x'y' = \frac{c^2}{4}. \tag{3.26}$$

This equation is called the **natural equation of a hyperbola** with respect to its asymptotes (which have become the axes of the new affine coordinate system).

In particular, for an equilateral hyperbola

$$x^2 - y^2 = a^2 \tag{3.27}$$

we have

$$x'y' = \frac{a^2}{2}. \tag{3.28}$$

If in the general case of an arbitrary hyperbola we take its asymptotes as new axes, without imposing any restrictions on the lengths of unit vectors, then the equation of the hyperbola will have the form

$$x'y' = k. \tag{3.29}$$

The number k can be any real number other than zero.

4 Ellipse and hyperbola directrices

The **directrix** of an ellipse (hyperbola) corresponding to a given focus F is a line d perpendicular to the focal axis of the curve, spaced from the center by a distance $\dfrac{a}{e}$ and lying along the same side from the center as the focus F (Figs. 19 and 20).

Thus, both the ellipse (which is not a circle) and the hyperbola have two directrices.

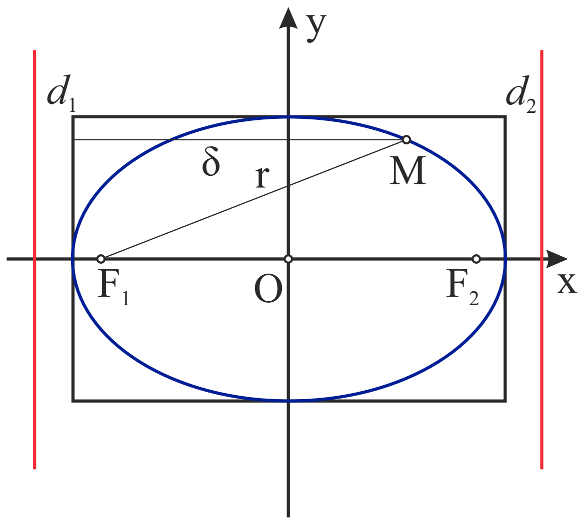


Figure 19

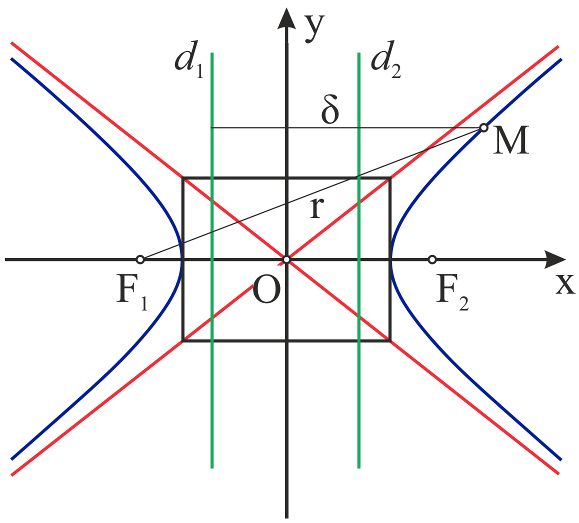


Figure 20

If the canonical rectangular coordinate system for the given curve is taken, then the equation of directrices d_1, d_2 (corresponding to foci F_1, F_2) will be respectively

$$x = -\frac{a}{e}, \tag{4.1}$$

$$x = \frac{a}{e}. \tag{4.2}$$

For an ellipse $e < 1$, so the directrices of the ellipse are at a distance greater than a from the center, they are located outside the fundamental rectangle (see Fig. 19).

For a hyperbola $e > 1$, therefore the directrices of the hyperbola are at a distance less than a from the center, they intersect the fundamental rectangle and pass between the center and the corresponding vertex of the hyperbola (see Fig. 20).

Finally, note that the distance Δ of the directrix from the corresponding focus is

1. in the case of an ellipse

$$\Delta = \frac{a}{e} - ae = a\frac{1-e^2}{e} = \frac{1}{e} \cdot \frac{b^2}{a};$$

2. in case of a hyperbole

$$\Delta = ae - \frac{a}{e} = a\frac{e^2-1}{e} = \frac{1}{e} \cdot \frac{b^2}{a}.$$

So, for an ellipse and for a hyperbola we have

$$\Delta = \frac{1}{e} \cdot \frac{b^2}{a}. \tag{4.3}$$

If in the case of a hyperbola (with given a) the focal length c, and hence the eccentricity $e = \dfrac{c}{a}$, increase, then the acute angle between the asymptotes decreases, and the directrices are all closer to the second axis (and approach each other).

If in the case of an ellipse (with given a) the focal length c, and hence the eccentricity $e = \dfrac{c}{a}$, decrease, then the ellipse becomes more and more like a circle, and its directrices move farther and farther from the second axis (and from each other). Finally, for the circle $e = 0$ and the directrices disappear ("going to infinity") — the circle has no directrices.

Let some ellipse or hyperbola C be given; we denote one of the foci of the curve C by F, and the directrix corresponding to it by d. For an arbitrary point M we denote by r the distance of this point M from the point F, by δ the distance of the point M from the line d. Let us prove that for all points M of the curve C we will have

$$\frac{r}{\delta} = e. \tag{4.4}$$

It suffices to prove this equality for the case when $F = F_1$ is the first (left) focus (the coordinate system is canonical).

Then we have

$$r = |a + ex|, \qquad \delta = \left| x + \frac{a}{e} \right|,$$

where

$$\frac{r}{\delta} = e.$$

Thus, the equality (4.4) is true for all points of the curve C.

Let us prove the converse: if the equality (4.4) is true for some point $M = (x, y)$ of the plane, then the point M lies on the curve C.

Indeed, let F again be the left focus of the curve C, i.e., $F(-c, 0)$, and let the line d have the equation

$$x = -\frac{a}{e}.$$

Then

$$r^2 = (x + c)^2 + y^2,$$

$$\delta^2 = \left(x + \frac{a}{e}\right)^2 = \left(x + \frac{a^2}{c}\right)^2.$$

By assumption, the condition (4.4) is satisfied for the point M, so that

$$\frac{r^2}{\delta^2} = e^2 = \frac{c^2}{a^2},$$

i.e.

$$\frac{(x+c)^2 + y^2}{\left(x + \dfrac{a^2}{c}\right)^2} = \frac{c^2}{a^2}$$

or

$$a^2(x+c)^2 + a^2y^2 = c^2\left(x + \frac{a^2}{c}\right)^2,$$

which, after some transformations, becomes

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2). \tag{4.5}$$

If the curve C is an ellipse, then $e = \dfrac{c}{a} < 1$, $a^2 - c^2 = b^2$ and equation (4.5) rewritten as

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

or

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

— the point M lies on the ellipse C.

If the curve C is a hyperbola, then $e = \dfrac{c}{a} > 1$, $c^2 - a^2 = b^2$ and the equation (4.5) can be written in the form

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

or in the form

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

— the point M lies on the hyperbola C. So, the following theorem is proved:

Theorem. *Both the ellipse and the hyperbola C with eccentricity e are the locus of points M of the plane satisfying the following condition: the ratio of the distance of the point M to an arbitrarily chosen focus of the curve to the distance of the point M to the directrix corresponding to this focus is equal to e.*

Let now (in the plane) be given a point F, a line d not passing through this point, and a positive number $e \neq 1$.

Let us prove that for $e < 1$ there is an ellipse and for $e > 1$ there is a hyperbola with eccentricity e, focus F, and corresponding directrix d.

Indeed, let us drop the perpendicular FD from the point F to the line d and denote by A the point dividing the segment $\overrightarrow{FD}$ in the ratio equal to e, and by A' we denote the point dividing the same segment $\overrightarrow{FD}$ in the ratio equal to $-e$, so that

$$\frac{\overrightarrow{FA}}{\overrightarrow{AD}} = e, \qquad \frac{\overrightarrow{FA'}}{\overrightarrow{A'D}} = -e. \tag{4.6}$$

It is easy to show that then the midpoint O of the segment $\overrightarrow{AA'}$ divides the segment $\overrightarrow{FD}$ in the ratio equal to $-e^2$:

$$\frac{\overrightarrow{FO}}{\overrightarrow{OD}} = -e^2,$$

i.e.

$$\overrightarrow{OF} = e^2\overrightarrow{OD}. \tag{4.7}$$

It follows from the equalities (4.6) and (4.7) that the points F, D and A lie on the same side of the point O.

Let's choose a rectangular coordinate system Oxy with origin at point O and positive direction $\overrightarrow{OF}$ of axis Ox. Suppose in this system

$$F = (c, 0), \qquad D = (d, 0), \qquad A = (a, 0).$$

Since the points A, F and D lie on the positive ray of the x-axis, then all three numbers a, c and d are positive, and

$$a = \frac{c + ed}{1 + e}, \tag{4.8}$$

$$c = e^2 \cdot d. \tag{4.9}$$

To establish that the point F and the line d are the focus and directrix of the curve with center O, major semiaxis a and eccentricity e, it suffices to show that

$$ae = c, \qquad \frac{a}{e} = d.$$

We get

$$a \cdot e = \frac{c + ed}{1 + e} \cdot e = \frac{ec + e^2 d}{1 + e} = \frac{ec + c}{1 + e} = c$$

and

$$\frac{a}{e} = \frac{c + ed}{(1 + e) \cdot e} = \frac{ed^2 + ed}{(1 + e) \cdot e} = d.$$

The statement is proven.

The eccentricity of an ellipse (not a circle) is a positive number $e < 1$; eccentricity of the hyperbola is $e > 1$.

We define the eccentricity for any parabola by setting it equal to $e = 1$. Now any positive number e is the eccentricity of either an ellipse, or a parabola, or a hyperbola, and as a result we get:

The class of curves that are ellipses (except circles), parabolas or hyperbolas can be defined as follows:
Every curve C of this class (and only a curve of this class) is the locus of points M for which the ratio $\frac{r_M}{\delta_M}$ of the distance r_M of point M from some fixed point F ("focus of curve C") to the distance δ_M of point M from some fixed line ("directrix of curve C") is a constant positive number e (for all points M of curve C),

$$e = \frac{r_M}{\delta_M},$$

which is called the eccentricity of the curve C.
Curve C is

- *an ellipse if $e < 1$,*
- *a parabola if $e = 1$,*
- *a hyperbole if $e > 1$.*

5 Focal parameter of ellipse and hyperbola. Curve equation at vertex

5.1 Focal parameter

Let C be an ellipse or a hyperbola. As before (in the case of a parabola), we draw a straight line through (some) focus F of the curve C perpendicular to its focal axis. This line will intersect the curve C at two points P and P' located "above" and "below" the focus F. The length of the focal chord PP' is denoted by $2p$; the quantity p is called the **focal parameter**[8] of the curve C.

The focal parameter of a circle is obviously equal to its radius.

[8]Due to the symmetry of the curve about its second axis, it does not matter which one of the two foci we choose.

Take the canonical coordinate system for the given curve C; then the focal parameter of the curve C is equal to the absolute value of the ordinate of each of the points P, P'. Let's calculate it.

If the curve C is an ellipse with the equation

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

then for any point $M = (x, y)$ of this ellipse

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

Substituting here the abscissa of the focus F, i.e.

$$x = \pm c,$$

we obtain for the ordinates of the points P, P' the values

$$y = \pm\frac{b}{a}\sqrt{a^2 - c^2} = \pm\frac{b^2}{a}.$$

So the focal parameter of the ellipse is

$$p = \frac{b^2}{a}. \tag{5.1}$$

For a point $M = (x, y)$ of a hyperbola

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

we have

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

For $x = \pm c$ we get

$$y = \pm\frac{b}{a}\sqrt{c^2 - a^2} = \pm\frac{b^2}{a}.$$

The focal parameter of the hyperbola is also

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

Recall that we found (see Sec. 4) for the distance Δ between the focus and the corresponding directrix of both the ellipse and the hyperbola the expression

$$\Delta = \frac{1}{e} \cdot \frac{b^2}{a}.$$

Now we see that this distance can also be expressed in terms of the focal parameter:

$$\Delta = \frac{p}{e}, \tag{5.2}$$

and this expression is suitable not only for the ellipse and hyperbola, but also for the parabola (for which, as we know, $e = 1$ and $\Delta = p$).

Thus, for all our curves (except the circle) *the focal parameter p can be defined as the number $p = e\Delta$, where Δ is the distance from the focus to the directrix, and e is the eccentricity.*

5.2 Curve equation at the vertex. Explanation of the names "parabola", "ellipse", "hyperbola"

Let's use the focal parameter to find a coordinate system in which the equations of all three curves (ellipse, hyperbola and parabola) have a similar form.

This coordinate system has the vertex of the curve as its origin, and the focal axis as its x-axis; for a parabola, this coordinate system has already been considered by us in Section 1 under the name "canonical"; with respect to this system, the parabola has the equation

$$y^2 = 2px.$$

In the case of an ellipse, we take the "left" vertex A_1 as the origin of the new coordinate system, retaining the unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ of the old system (the canonical system for given curve C (Fig. 21)). Denoting the coordinates of some point M in the new system as x', y', we have formulas for the transition from the old coordinates x, y to the new x', y':

$$\begin{aligned} x &= -a + x', \\ y &= y'. \end{aligned}$$

Substituting these values into the canonical equation of the ellipse

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

we get

$$\frac{(-a + x')^2}{a^2} + \frac{y'^2}{b^2} = 1$$

or, after transformations,

$$-2\frac{x'}{a} + \frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 0,$$

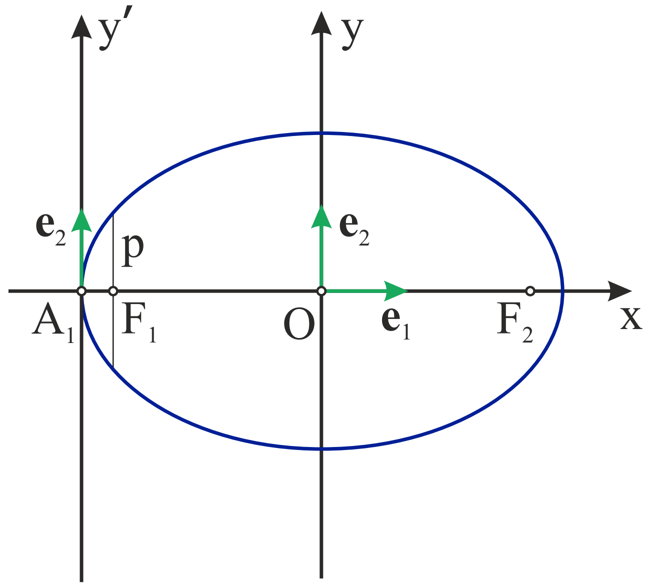


Figure 21

or
$$y'^2 = 2\frac{b^2}{a}x' - \frac{b^2}{a^2}x'^2. \tag{5.3}$$
But $\dfrac{b^2}{a} = p$, $\dfrac{b^2}{a^2} = 1 - e^2$, so the equation (5.3) can be rewritten as
$$y'^2 = 2px' + (e^2 - 1)x'^2$$
or, assuming
$$q = e^2 - 1, \tag{5.4}$$
as
$$y'^2 = 2px' + qx'^2. \tag{5.5}$$
Now let the curve C be a hyperbola
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{5.6}$$
We pass to a new coordinate system, transferring the origin of the canonical (for the curve C) system to the right vertex $A_2 = (a, 0)$ of the hyperbola and preserving the unit vectors of the canonical system (Fig. 22).

Now
$$x = x' + a,$$
$$y = y'.$$
Substituting this into the hyperbola equation (5.6), we get
$$\frac{x'^2 + 2ax' + a^2}{a^2} - \frac{y'^2}{b^2} = 1$$

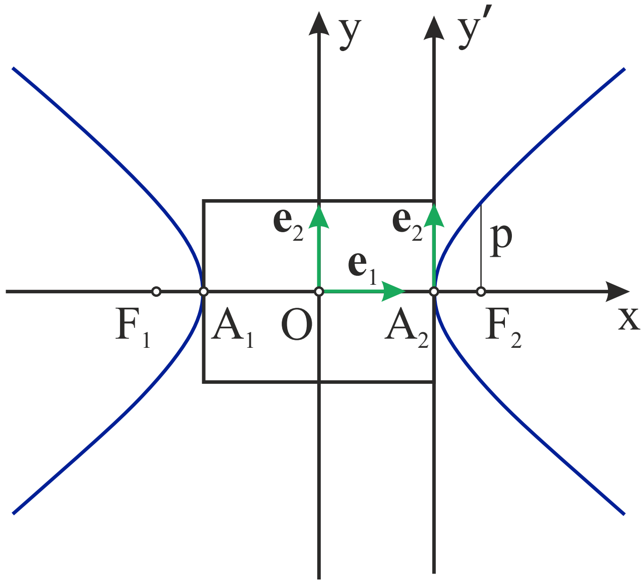


Figure 22

or

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

thus (remembering that in the case of a hyperbola $b^2/a^2 = e^2 - 1$)

$$y'^2 = 2px' + (e^2 - 1)x'^2,$$

or finally (with the notation (5.4)),

$$y'^2 = 2px' + qx'^2,$$

So, *in a properly chosen coordinate system, all three curves (parabola, ellipse and hyperbola) have equations of the same form, namely:*

$$\boxed{y^2 = 2px + qx^2,} \tag{5.7}$$

where p *is the focal parameter of the curve, and*

$$q = e^2 - 1$$

(as always, e *denotes the eccentricity of the curve).*

It follows from these notations that $q < 0$ for an ellipse, $q = 0$ for a parabola, and $q > 0$ for a hyperbola. The origin of this coordinate system is the vertex of the curve[9] (left for an ellipse, right for a hyperbola), and the x-axis is focal axis directed from the (selected) vertex to the (corresponding) focus.

[9]Therefore, the equation (5.7) is called the "curve equation at the vertex".

Figure 23 shows the curves of a family consisting of ellipses, parabolas and hyperbolas with a common vertex, a common focal axis, with the same focal parameter p and eccentricity e, taking on ever increasing values: starting from zero (circle) to 1 (parabola) and further (hyperbolas).

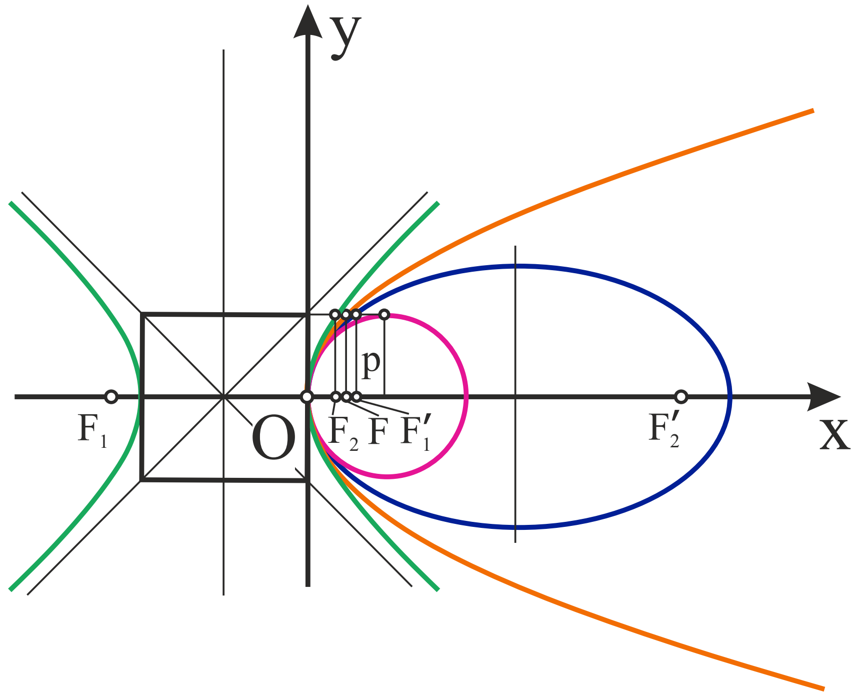


Figure 23

Remark. The geometers of ancient Greece did not know the method of coordinates and did not know analytical geometry; but the properties of the ellipse, hyperbola and parabola were known; they obtained this knowledge by the methods of so-called synthetic geometry, i.e. by direct geometric construction, defining these curves (ellipse, hyperbola, parabola) as conic sections, i.e. as plane sections of a round cone (about this definition. The properties of our curves, expressed by the equation (5.7), the Greeks also knew well; but the ordinate y was for them half a vertical (that is, perpendicular to the focal axis) chord, one of the ends of which was the given point P, and the abscissa x of this point was the distance from the tangent drawn to the curve at its vertex. Greek geometers knew well also the focal parameter p of a conic section.

They called "application" of a given square to a given segment the construction of a rectangle with a given base that is equal in area to the given square (i.e., in fact, the construction of the height of this rectangle). Formula

$$y^2 = 2px$$

solves (for the case of a parabola) the problem of applying the square y^2 (with the above geometric interpretation of the ordinate of an arbitrary point y) to the base $2p$ (i.e. to the focal chord of this curve) — the height of the desired rectangle is x, i.e. the distance from the point M to the tangent at the vertex. The problem is solved so well for the parabola, hence its name: παραβολή means "finding something nearby", "application".

For an ellipse and a hyperbola, the task no longer goes so smoothly: an additional term qx^2 is needed, positive (i.e., being an excess, ὑπερβολή)) in the case of a hyperbola, negative (i.e., being a deficit, ἔλλειψις) in the case of an ellipse. The words "hyperbole" and "ellipse" mean "excess" and "lack"[10].

6 A line tangent to conics

The equation of the tangent line at a non-singular point $(x_0,\, y_0)$ to the line given by the implicit equation

$$F(x, y) = 0,$$

is written as

$$F'_x(x_0, y_0)(x - x_0) + F'_y(x_0, y_0)(y - y_0) = 0, \tag{6.1}$$

where $F'_x(x_0, y_0)$ and $F'_y(x_0, y_0)$ are the the values of partial derivatives of the function $F(x, y)$ at the point $(x_0,\, y_0)$.

For an *ellipse* given by the canonical equation

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

the equation of the tangent at the point $(x_0,\, y_0)$ lying on this ellipse has the form

$$\frac{2x_0}{a^2}(x - x_0) + \frac{2y_0}{b^2}(y - y_0) = 0$$

or (taking into account that $x_0^2/a^2 + y_0^2/b^2 = 1$)

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1.$$

The equation of the tangent line to the *hyperbola* given by the canonical equation

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

[10] These words are also used in this sense in stylistics and rhetoric; the word "hyperbole" means exaggeration, and "ellipsis" (an exact reproduction of the Greek word) is a rhetorical figure, consisting in the omission of some member of the sentence.

at the point $(x_0,\, y_0)$ can be obtained by making similar calculations:

$$\frac{2x_0}{a^2}(x - x_0) - \frac{2y_0}{b^2}(y - y_0) = 0, \text{ or } \frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1.$$

If the *parabola* is given by the equation

$$y^2 = 2px,$$

then the equation of the tangent line to it at the point $(x_0,\, y_0)$ has the form

$$y_0 y = p(x - x_0).$$

7 Optical property of conics

7.1 Ellipse

Theorem. The tangent line to the ellipse at its arbitrary point M_0 is the bisector of the outer angle M_0 of the triangle $F_1F_2M_0$, which has the foci F_1 and F_2 of the ellipse and this point M_0 as its vertices.

Proof. Let's consider the equation of the tangent line to the ellipse

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

at the given point on it $M_0(x_0, y_0)$:

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} - 1 = 0.$$

The ratio of the distances h_1 and h_2 from the foci $F_1(-c, 0)$ and $F_2(c, 0)$ of the ellipse to the tangent line at the point $M_0(x_0, y_0)$ is equal to the ratio of the modules of the results of substituting the coordinates of the foci F_1 and F_2 in the left side of the tangent equation:

$$h_1 : h_2 = \left|-\frac{cx_0}{a^2} - 1\right| : \left|\frac{cx_0}{a^2} - 1\right| = \left|1 + \frac{ex_0}{a}\right| : \left|1 - \frac{ex_0}{a}\right| = \\ |a + ex_0| : |a - ex_0| = r_1 : r_2.$$

Note that the results of substitutions $|-cx_0/a^2 - 1|$ and $|cx_0/a^2 - 1|$ of the coordinates of the foci $F_1(-c, 0)$ and $F_2(c, 0)$ to the left side of the tangent equation are numbers of one sign:

$$-\frac{cx_0}{a^2} - 1 = -\frac{ex_0}{a} - 1 = -\frac{ex_0 + a}{a} = -\frac{r_1}{a} < 0,$$

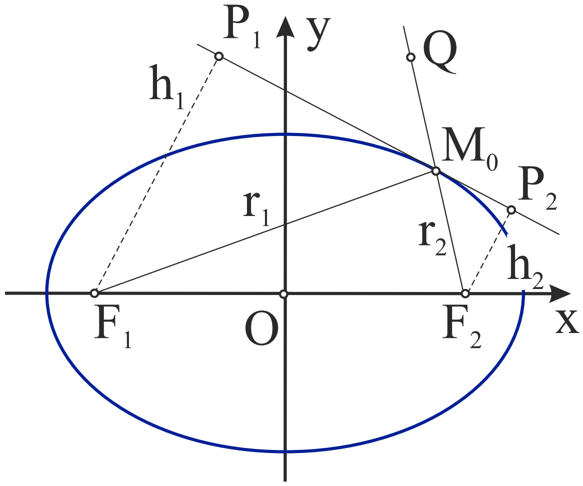


Figure 24

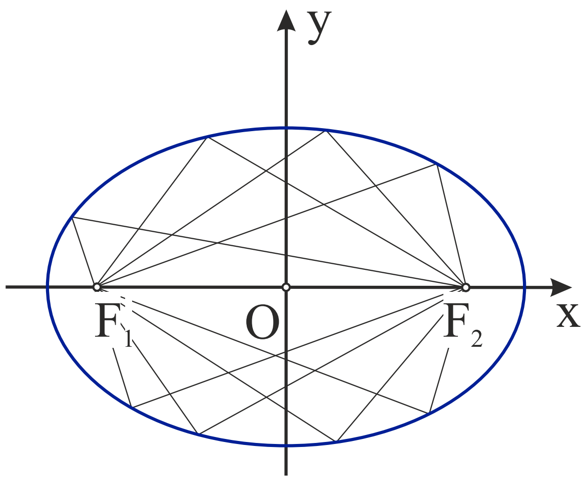


Figure 25

$$\frac{cx_0}{a^2} - 1 = \frac{ex_0}{a} - 1 = -\frac{a - ex_0}{a} = -\frac{r_2}{a} < 0,$$

therefore, both foci F_1 and F_2 are located on the same side of the tangent line to the ellipse at arbitrary point.

Denote by P_1 and P_2 the bases of the perpendiculars omitted from the foci of the ellipse on the tangent line (Fig. 24). Then $\triangle F_1P_1M_0 \sim \triangle F_2P_2M_0$, since both are rectangular, and according to the proven

$$\frac{|F_1P_1|}{|F_2P_2|} = \frac{|F_1M_0|}{|F_2M_0|},$$

therefore

$$\angle F_1M_0P_1 = \angle F_2M_0P_2,$$

so, $\angle F_1M_0P_1 = \angle P_1M_0Q$, where the point Q lies on the continuation of the segment F_2M_0 beyond the point M_0. $\square$

This theorem directly implies the method of constructing a tangent to an ellipse at an arbitrary point.

The proved theorem can be given the following physical interpretation: if a light source is placed in one of the foci of the ellipse, then the rays after reflection from the ellipse will gather in another focus, since the light beam is reflected from the ellipse as from a tangent line drawn to the ellipse at the point of incidence of the beam (Fig. 25).

7.2 Hyperbola

Theorem. The tangent line to the hyperbola at the arbitrary point M_0 is the bisector of the inner angle M_0 of the triangle $F_1M_0F_2$, which has the foci F_1 and F_2 of the hyperbola and point M_0 as its vertices.

Proof. We omit from the foci F_1 and F_2 the perpendiculars F_1P_1 and F_2P_2 on the tangent line (Fig. 26). Just like for an ellipse, it can be proved that

$$\frac{|F_1P_1|}{|F_2P_2|} = \frac{|F_1M_0|}{|F_2M_0|},$$

therefore, $\triangle F_1M_0P_1 \sim \triangle F_2M_0P_2$, and we get

$$\angle F_1M_0P_1 = \angle F_2M_0P_2. \quad \square$$

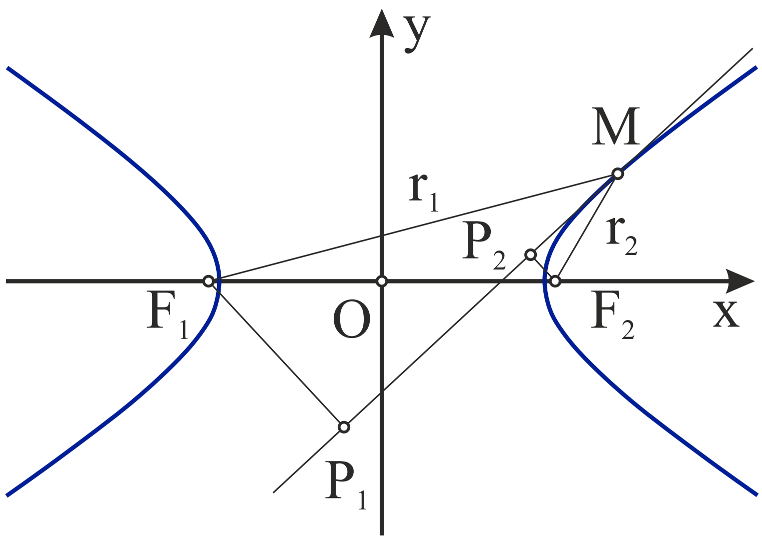


Figure 26

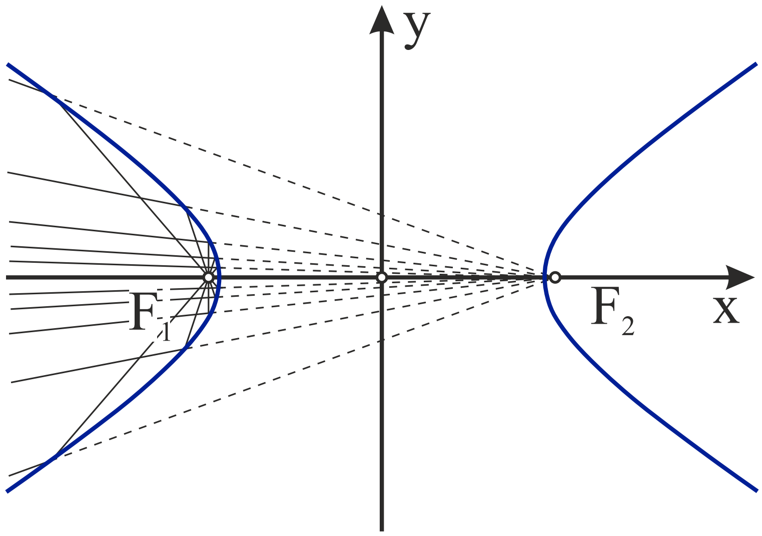
The results of substituting the coordinates of the foci $F_1(-c, 0)$ and $F_2(c, 0)$ into the expression

$$\frac{x_0x}{a^2} - \frac{y_0y}{b^2} - 1$$

are numbers of different signs, which implies that the foci of the hyperbola lie on different sides of any tangent line to it.

This geometric property allows us to construct a tangent line to the hyperbola at an arbitrary point M_0: we connect the point M_0 with the foci F_1 and F_2 of the hyperbola and divide the angle $\angle F_1M_0F_2$ in half; the bisector of this angle is tangent line to the hyperbola at the point M_0.

The proved theorem can be given an optical interpretation, similar to what was given for the ellipse (Fig. 27).



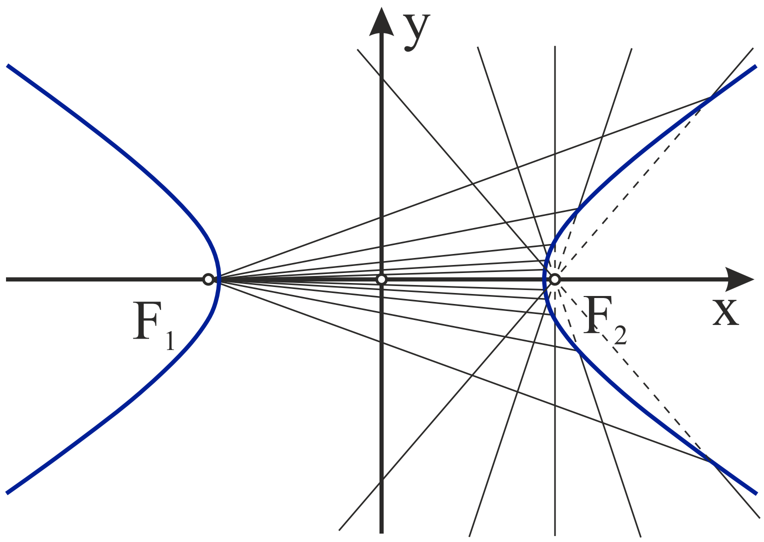


Figure 27

7.3 Parabola

Theorem. The tangent to the parabola is the bisector of the angle $\angle FMD$ between the focal radius MF of the tangent point and the perpendicular MD from it to the directrix.

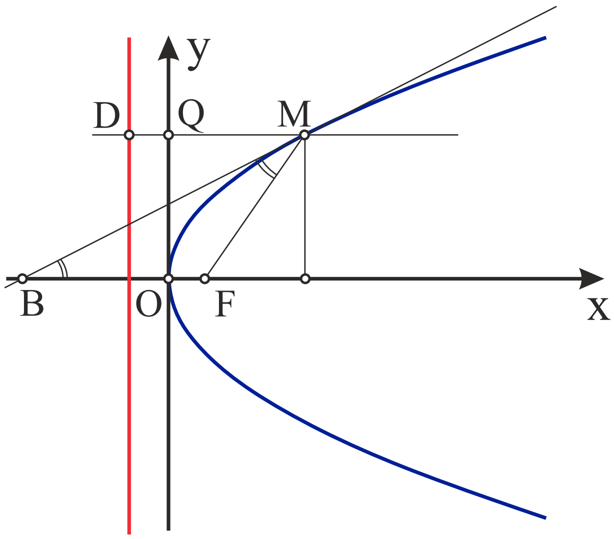


Figure 28

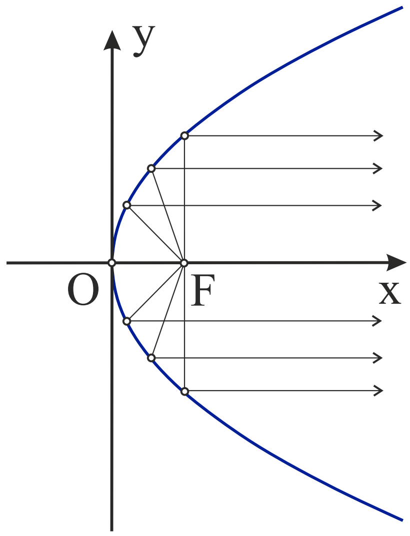


Figure 29

Proof. We have (Fig. 28):

$$MD = FM, \qquad MD = MQ + QD = x_0 + QD.$$

But

$$x_0 = OB, \qquad QD = OF,$$

therefore,

$$MD = OB + OF = FB.$$

So the triangle BFM is isosceles and, therefore,

$$\angle FMB = \angle FBM;$$

but $\angle FBM = \angle BMD \Rightarrow \angle FMB = \angle BMD$. $\square$

This theorem has the following optical interpretation, if a light source is placed at the focus F of a parabolic mirror, then rays reflecting from a mirror form a beam of parallel rays (parallel to the axis of the parabola). The indicated property of a parabolic mirror is used when constructing mirror spotlights (Fig. 29).

8 Ellipse, hyperbola and parabola equations in polar coordinates

The focal parameter also finds its application in the derivation of the equations of the ellipse, hyperbola and parabola in polar coordinates. These equations are constantly used in astronomy and in many problems of mechanics.

The origin of the polar coordinate system is placed at the focus F (left in the case of an ellipse (Fig. 30), right in the case of a hyperbola (Fig. 31)[11] and to a single focus in the case of a parabola (Fig. 32)); the polar axis is directed away from the pole and away from the corresponding directrix d (i.e., in the same way as the x-axis is directed in the canonical coordinate system for the given curve).

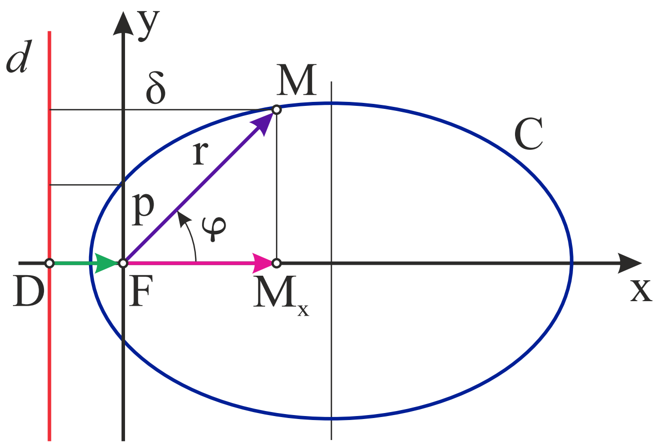


Figure 30

For any point M of our curve we denote by r the distance from M to the focus F, by δ we denote the distance from M to d. Our curve C is the locus of points M for which $\dfrac{r}{\delta} = e$, i.e.

$$r = \delta e. \tag{8.1}$$

But r is the polar radius of the point M. Let's calculate δ. Denoting by D the point of intersection of the directrix d with the focal axis, and by M_x the projection of the point M onto this axis, we see that δ is the length of the vector $\overrightarrow{DM_x}$ lying on x-axis (in canonical system). For the algebraic values of the vectors on this axis, we have

$$|DM_x| = |DF| + |FM_x|, \tag{8.2}$$

but

$$|DF| = \frac{p}{e},$$

[11]In the case of a hyperbola the equation in polar coordinates does not determine the entire curve, but only one right branch (with our choice of origin).

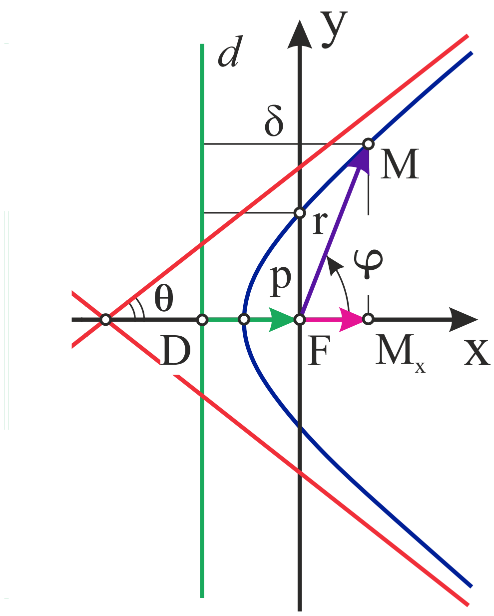


Figure 31

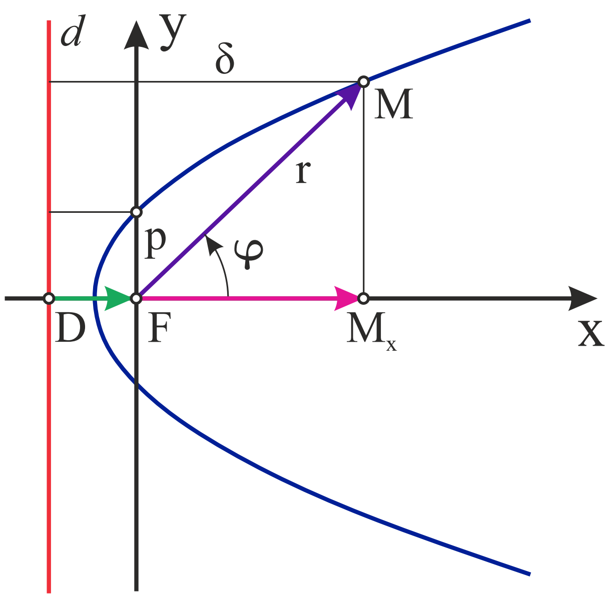


Figure 32

whereas

$$|FM_x| = r\cos\varphi,$$

where φ is the angle of slope of the vector $\overrightarrow{FM}$ to the polar axis, i.e., the polar angle of the point M. On the curve C (in the case of a hyperbola on its right branch) we have $|DM_x| = \delta$. Substituting into the equality (8.2) the found values, we obtain

$$\delta = \frac{p}{e} + r\cos\varphi = \frac{p + er\cos\varphi}{e}.$$

Finally, substituting this δ value into (8.1), we have

$$r = p + er\cos\varphi,$$

i.e. $r(1 - e\cos\varphi) = p$ or

$$\boxed{r = \frac{p}{1 - e\cos\varphi}.} \tag{8.3}$$

This is the equation of the parabola, ellipse and (right branch of) hyperbola in polar coordinates.

For a parabola, we simply get

$$r = \frac{p}{1 - \cos\varphi}; \tag{8.4}$$

here φ takes all the values of $0 < \varphi < 2\pi$; the value $\varphi = 0$ is not suitable, which is natural, since no point of the parabola corresponds to it.

In the case of an ellipse, all values of $0 \leqslant \varphi < 2\pi$ are suitable (since $e\cos\varphi \leqslant e < 1$ is always true).

For a hyperbola, one can take the values of φ for which

$$\cos\varphi < \frac{1}{e} = \frac{a}{\sqrt{a^2 + b^2}} = \cos\theta,$$

where θ is the (acute) angle between the asymptote and the focal axis of the hyperbola; for all points of the right branch of the hyperbola the polar angle φ lies within

$$\theta < \varphi < 2\pi - \theta,$$

so

$$\cos\theta > \cos\varphi.$$