Independent University of Moscow
Translation Series

A. N. Zemlyakov

Mathematical Analysis of Reality

Differential Equations for School Students

Part I

Continuous Processes and Differential Equations

MCCME, IUM
Moscow, 2025

Zemlyakov A. N.

Mathematical Analysis of Reality: Differential Equations for School Students. Part I: Continuous Processes and Differential Equations. — Moscow: MCCME: IUM, 2025. — 219 p.

ISBN 978-5-4439-4702-0

The book provides numerous examples of mathematical modelling of reality that can be understood and comprehended at the school level of mathematics.

The book is intended for high school students choosing the direction of their professional education and inclined to understand the actual role of mathematics in science and practice. This book will also be useful for undergraduate students who are studying differential equations and mathematical models.

Translated into English on the initiative and with financial support of Alexander Gerko

Translated from the Russian language edition
Земляков А. Н. Математический анализ реальности. Дифференциальные уравнения для школьников. — М.: МЦНМО, 2013.

ISBN 978-5-4439-4702-0

Contents

Preface

This book is aimed at demonstrating — using specific examples of school mathematics (the course of algebra and elementary calculus) — how mathematics can be applied to the analysis of the real world around us, which is revealed in a variety of natural processes, from "simple" (mechanical) motion to biological evolution. Since its very origin, mathematics has been largely concerned with solving practical problems. However, it was not until the XVII century that the remarkable effectiveness of mathematics in the study of natural sciences became fully apparent. To a large extent, this was due to the birth and development of calculus as a powerful tool not only for explaining reality, but also for scientific forecasting.

For Newton, an indispensable part of a method of mathematical research invented by him and Leibniz were differential equations, i.e., equations in which unknowns are not numbers — as in algebraic equations, which have been solved as early as five millennia BC in Babylon and Egypt — but functions describing various processes: motion (from a thrown stone to rotation of planets), evolutionary changes (from reproduction of bacteria to the biocoenosis), etc. In modern language, differential equations are the most important mathematical models of real-world processes, helping to compute space trajectories and nuclear reactions and predict the evolution of processes.

This book provides numerous examples of mathematical modelling of reality that can be understood and comprehended at the school level of mathematics. The book is intended for senior school students who are choosing the direction of their professional education and who are eager to understand the actual role of mathematics in science and practice.

Chapter I

Evolutionary Models

§ 1.1. Differential Equations and Evolutionary Models

A great many of processes in the animate and inanimate nature, as well as in social systems, can be described as changes in time of some *parameters* of a considered system. Thus, a varying, "dynamic" system *evolves* in time, and one of the tasks of mathematics, which has played a dominant role since the XVII century (since the discovery of *calculus* methods), is the development of mathematical models of evolution (changes in time), the so-called *evolutionary models.*

Evolution, from Latin *evolutio* 'unrolling', is in the broad sense a synonym of the term 'development'; in the narrow sense, which is assumed in our case, just any change; and most specifically, a continuous, gradual (in time!) quantitative change of a particular system, some of its numerical characteristics and/or parameters.[1] In some cases it is possible to find out (e.g., experimentally) an explicit *functional* dependence of the considered parameters on time — then we know a *variation law*, i.e., a set of some *functions* $x_i = x_i(t)$ of the time variable $t \in \mathbb{R}$. This set is just what constitutes a *mathematical model of evolution.* Much more often we cannot specify in advance, before a mathematical investigation, the dependence of the parameters on time, but we can find somehow (e.g., experimentally) the *rate of change* and construct a mathematical model based on this information. Such models are what we shall call *evolutionary* models.

In this section we describe a way to construct evolutionary models; also, in this section and in the next one, we consider general and particular examples of evolutionary models.

1.1.1. Meaning of the derivative: rate of change. Recall (see textbooks on elementary calculus) that the *derivative of a numerical function* $f : x \mapsto y = f(x)$ at a point x such that f is defined in some its *neigh-*

[1] In contrast to *revolution*, from Latin *revolutio*, meaning 'turn', 'flip-over'.

bourhood $\mathrm{Nghb}_x = (x - r, x + r)$ is the *limit*

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = f'(x), \tag{1}$$

provided that *the limit exists.* The limit equality (1) means that the difference

$$\frac{f(x+h) - f(x)}{h} - f'(x) = \alpha_x(h) \tag{2}$$

tends to zero as h tends to zero; or, in other words, the difference $\alpha_x(h)$ is an *infinitesimal* at zero. This, in turn, means that $\alpha_x(h)$ becomes *arbitrarily close to zero for a sufficiently small* h: whatever a positive number ε, there exists a number $\delta > 0$ such that for any $h \in (-\delta, \delta)$ the quantity $\alpha_x(h)$ differs from zero by less than ε, which can be written using *quantifiers*:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall h \in (-\delta, \delta) \quad |\alpha_x(h)| < \varepsilon$$

('$\forall$' is the *universal quantifier*: 'for any...', 'for all...'; '$\exists$' is the *existential quantifier*: 'there exists...').

The concept of the derivative, which is one of the two basic concepts of *calculus*, has been introduced independently of each other in the late XVII century by Sir Isaac Newton (1643–1727) in England and Gottfried Wilhelm Leibniz (1646–1716) in Germany. A hundred years later, the French mathematicians Jean le Rond d'Alembert (1717–1783) and Joseph-Louis Lagrange (1736–1813) defined the derivative using an intuitive idea of a limit (both the term '*derivative*' itself and the notation for it were proposed by Lagrange; Newton called it 'fluxion', and Leibniz related it to the notion of a 'differential', for which see Ch. IV). Finally, another century later, the German mathematician Karl Theodor Wilhelm Weierstrass (1815–1897) established a rigorous foundation for the notion of the derivative and gave precise definitions for both the derivative and the limit, inventing the "(ε, δ)-language" used above.

We are not as much concerned with strict definitions of the derivative, limit, and infinitesimal as we are with the *essential meaning* of the concept of the derivative. Its primary meaning was most clearly identified by Newton himself: *the derivative of a function* $y = f(x)$ *at a point* x *is the* rate of change *of the variable* y *depending on the variable* x *at a particular point* x. What does this mean?

The simplest way to explain the above *essential meaning* of the derivative is as follows. If the *argument* x of a function f changes by some h (the *increment* $\Delta x = h$ *of the argument* can as well be negative), then the value of the function changes by $\Delta y = f(x+h) - f(x)$, and it is natural to regard the ratio of the increment of the function to the increment of the argument as the *average rate of change* of y over the interval from x to $x + h$. The limit of the average rate as $h \to 0$ is precisely the *rate of change of the function at a point*, which is the derivative $y'(x) = f'(x)$.

This explanation becomes quite natural if, like Newton, we consider that x is the *universal variable of Time.* However, velocity in mechanics (kinematics) is a *vector*, whereas the above-introduced rate of change of a function is a *number*, a *scalar*. A scalar quantity, unlike a vector quantity, has no *direction*. Instead, it has a *sign*, which shows "in which direction" — increasing or decreasing — the value of y changes. Thus, from the essential meaning of the derivative it is clear that if the derivative $f'(x)$ is positive on some interval I, then the function f is increasing on it, and if $f'(x) < 0$ on I, then the function is decreasing.[1]

Although physicists and mathematicians have been dealing with vector and scalar quantities for a long time and managed them well (we can mention Kepler with his second law, as well as Newton), the terms themselves and precisely defined notions of vector and scalar were introduced only in 1845 by the Irish mathematician and astronomer Sir William Rowan Hamilton (1805–1865), who is famous for the invention of the so-called *quaternions*, an analogue of numbers in four-dimensional space. The term *vector* is from Latin *vector*, 'carrier'; in mathematics it means a *directed segment* (or an element of the so-called *vector space*). The term *scalar* is from Latin *scalaris*, 'staircase'; this is the name given to numerical quantities, with or without a sign (lengths, areas, volumes, temperature, etc.).

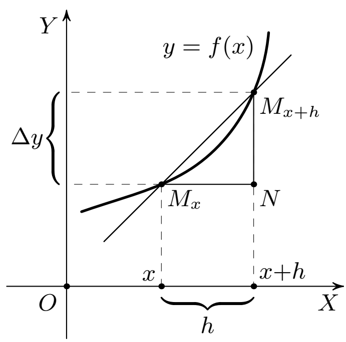


Fig. 1.

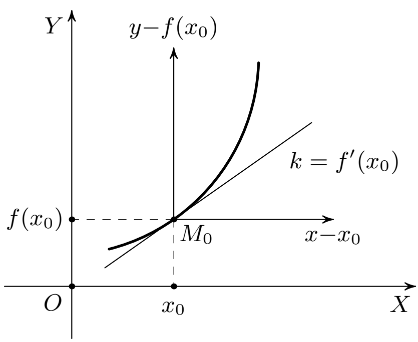


Fig. 2.

1.1.2. Tangent line and geometric interpretation of the derivative. We shall return to the kinematic notion of velocity vector and its relation to the derivative later, but for now let us recall the *geometric interpretation of the derivative.* On the graph of the function $y = f(x)$ we consider two points, $M_x(x; f(x))$ and $M_{x+h}(x+h; f(x+h))$, and the associated "characteristic triangle" $M_xM_{x+h}N$ (Fig. 1). This shows that the ratio of Δy to $\Delta x = h$ ("former" average rate) is the *slope* (*angular coefficient*) $k = k(x,h)$ *of the secant line* M_hM_{x+h}. It is intuitively clear that

[1]Recall that a *rigorous* proof of these *monotonicity criteria* requires using the *Lagrange theorem*, one of the *fundamental* theorems of differential calculus, and in fact of all mathematical analysis.

for "good" functions, this secant tends to the *tangent line* to the graph as $h \to 0$. Noting that the slope $k = k(x, h)$ tends to the derivative $f'(x)$ (of course, if the derivative exists), it is usually *assumed by definition that the tangent to the graph $y = f(x)$ at the point with x-value $x = x_0$ is the line passing through this point with the slope equal to the derivative $k_0 = f'(x_0)$.* By drawing *auxiliary coordinate axes* through the point $M_0(x_0; f(x_0))$ so that the coordinates of $(x; y)$ with respect to them are $(x - x_0; y - f(x_0))$ (Fig. 2), we can immediately write the equation of the tangent:[1]

$$y - f(x_0) = k_0(x - x_0) \quad \Leftrightarrow \quad y = f(x_0) + f'(x_0)(x - x_0). \tag{3}$$

Just *by the definition of a tangent line*, the derivative of a function at x_0 is the *slope of the tangent* to the graph of the function at the point on the graph with the given x-value. This is the (rather tautological with our definition of the tangent line!) geometrical meaning of the derivative.

1.1.3. Notion of a linear approximation. Now we come to something new and important. The linear function (3) that is tangent to the graph of $y = f(x)$ at the point with abscissa x_0 is called the *linear approximation to $f(x)$ at x_0*. This is among the most basic concepts in the analysis of numerical functions and its various generalisations (such as multivariable calculus or functional analysis). To clarify the meaning of a *linear approximation*, let us consider the difference between the function $f(x)$ itself and its linear approximation (3), i.e., the quantity

$$R = f(x) - f(x_0)f'(x_0)(x - x_0)) = (x - x_0)\Big(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)\Big). \tag{4}$$

Do not you notice something familiar in the expression in the parentheses, something you have already encountered?

Certainly, this is the infinitesimal difference (as $x - x_0 \to 0$, i.e., as $x \to x_0$) between the average velocity over the interval from x_0 to x and the derivative $f'(x_0)$, i.e., the rate of change of f at x_0, which appeared in equation (2). After changing the notation — using x instead of x_0, $x + h$ instead of x, so that $x - x_0$ is now simply equal to h — we can rewrite (4) in the form

$$R = f(x + h) - f(x) - f'(x)h = h\Big(\frac{f(x + h) - f(x)}{h} - f'(x)\Big) = h\alpha_x(h),$$

whence follows the representation

$$f(x + h) = f(x) + f'(x)h + h\alpha_x(h). \tag{5}$$

In the notation of equation (4), the above formula looks like

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + (x - x_0)\alpha(x). \tag{6}$$

[1] In the auxiliary coordinates $(\overline{x}; \overline{y})$, this will be the *direct proportionality* equation $\overline{y} = k_0\overline{x}$.

These two equations are called *formulae for the linear approximation of a function* f *at a point* x_0 (equation (5)) or *at a point* x (equation (6)). By discarding the last terms in these formulae, we obtain *approximate formulae* (linear approximations)

$$f(x+h) \approx f(x) + f'(x)h \tag{7}$$

and

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0), \tag{8}$$

where the absolute accuracy of computation is $|R|$, and the error with respect to the increment of the argument (h or $x - x_0$) tends to zero (as $h \to 0$ or $x \to x_0$).

Alternatively, equations (5) and (6) are called first-order *Taylor formulae* with *remainder term* $R = h\alpha_x(h)$ or $(x - x_0)\alpha(x)$. If a function f in the neighbourhood of the point where the linear approximation is considered is "good enough", then the remainder term is *of order* h^2. For example, if f has a second derivative $f''(x) \overset{\text{def}}{=} (f'(x))'$,[1] it can be written as $R = \frac{1}{2}f''(c)h^2$, where c is some point between x and $x+h$ (said to be the *Lagrange-form remainder term*). Certainly, for small values of h, the term of order h^2 can be neglected as compared to the linear term $f'(x)h$ (of course, in the case where it is *non-zero*, i.e., $f'(x) \neq 0$), and approximate formulae (7) and (8) are rather accurate.

Brook Taylor (1685–1731) was an English mathematician and philosopher, a proponent of Newton in his "dispute" with Leibniz on the priority in the discovery of calculus, member and secretary of the London Royal Society, who made his name immortal by the discovery (in 1712) of the above formula (of arbitrary order) and the corresponding infinite *power series* (see § 3.1 and § 3.3 below).

Example 1. For the function $f(x) = \sqrt{x}$ with $x > 0$, equation (5) is of the form

$$\sqrt{x+h} \approx \sqrt{x} + \frac{1}{2\sqrt{x}} \cdot h.$$

Let us apply this formula to find an approximate value of $\sqrt{17}$. To this end, we represent 17 as $16 + 1 = x + h$. Then $\sqrt{x} = \sqrt{16} = 4$, and linear approximation yields

$$\sqrt{17} = \sqrt{16+1} \approx \sqrt{x} + \frac{1}{2\sqrt{x}} \cdot h = 4 + \frac{1}{2 \cdot 4} \cdot 1 = 4 + \frac{1}{8} = 4.125.$$

Precise computations for $\sqrt{17}$ give $4.123105\ldots$, so the absolute error of the linear approximation formula in this case is less than two thousandths! □

Note that if for some number k we can write

$$f(x+h) = f(x) + kh + h\alpha(h) \tag{9}$$

[1] The symbol 'def' means 'equality by definition'.

with $\alpha(h)$ an infinitesimal at zero, then equation (9) implies

$$\frac{f(x+h)-f(x)}{h}-k=\alpha(h)\to 0 \quad \text{as} \quad h\to 0,$$

so the above ratio has a limit as $h\to 0$ and this limit equals k.

In other words, the derivative $f'(x)=k$ exists (see Sec. 1.1.1).

Hence, the existence of a representation (*linear approximation*) of the form (9) is equivalent to the differentiability condition (existence of the derivative) for the function, and the coefficient k in the linear approximation is precisely the derivative $f'(x)$.

It was precisely the existence of a linear approximation of the form (9) that Karl Weierstrass chose as the *definition of differentiability* in his lectures of 1861. He claimed that it was this definition that reflected the *"true meaning of the derivative"*. Note that our reasoning implies that *of all the lines passing through the point* $M_0(x_0; f(x_0))$, *it is the* tangent — *whose angular coefficient is equal to the derivative* (by the definition of the tangent line!) — *that is the closest to the graph of the function* $y=f(x)$ *in the neighbourhood of this point*, in the sense that for the difference $R=f(x)-f(x_0)-k(x-x_0)$ we have the relation

$$\lim_{x\to x_0}\frac{R}{x-x_0}=0 \quad \Leftrightarrow \quad k=f'(x_0)$$

(think it over). This statement, called the *linear approximation theorem*, is, according to Weierstrass, the essence of the definition of the derivative. Actually, we do not need Weierstrass's approach to the "true meaning" of the derivative further on; the essential, geometric, and *kinematic* (presented below) interpretations, or "meanings", will suffice. However, the linear approximation theorem and the series of formulae (5)–(8) will be useful in analysing the simplest evolutionary and dynamic models (the latter will be discussed in the next section).

The already mentioned German mathematician Karl Theodor Wilhelm Weierstrass (1815–1897; in particular, the teacher of Sofya Kovalevskaya) paid special attention to a rigorous justification of calculus. Thus, in 1861 he for the first time strictly formulated the definition of a limit in modern language, introduced the notion of the neighbourhood of a point, and in general "brought complete order" to the presentation of calculus. It was Weierstrass who introduced the notion of the absolute value (modulus) of a real number and the very notation $|x|$. His rigorous introduction of rational and irrational real numbers was based on the notion of the limit of a sequence of sums of fractions, in particular decimal fractions (so that, as is still customary in school textbooks, a real number is treated as an infinite decimal fraction). The *Weierstrass extreme value theorem* (one of many that bear his name), which is a part of the school course of elementary calculus, also belongs to this circle of questions. David Hilbert (1862–1943), one of the founders of the XX century mathematics, greatly appreciated Weierstrass's contribution to mathematics education and to mathematical science itself, noting in 1900: *"Besides it is an error to believe that rigor in the proof is the enemy of simplicity. On the contrary we find it confirmed by numerous examples that the rigorous method is at the same time the simpler and the more easily comprehended. The very effort for rigor forces us to find out simpler methods of proof"*. We shall not dispute this opinion.

1.1.4. Motion along a straight line. To describe the motion of a point particle (or simply a point) *along a straight line*, it suffices to introduce coordinates on this line, or, in other words, to regard the line as a *coordinate axis*, say Ox, and to consider the numerical function $x=x(t)$ specifying the coordinate of the point at time t. Then the *velocity of motion* — which, according to the definition of velocity accepted in physics (mechanics) is equal to the time derivative of the coordinate — is a numerical function $v=v(t)=x'(t)$. So is the rate of change of the velocity, i.e., *acceleration*, a scalar function $a=a(t)=v'(t)=x''(t)$, the second derivative of the coordinate function $x(t)$, also called the *law of motion.* Thus, we are ready to describe the simplest kinematic models of motion along a straight line.

Motion along a straight line (or in a straight line) is, in fact, the simplest of motions; however, its description in *dynamic* models related to motions under a known force is far from simple (see § 1.2 and Ch. V). For the moment, we shall confine ourselves to the simplest kinematic model.

Kinematics, from Greek κινεω [kineo], 'to move'; κινημα(τος) [kinema(tos)],'motion' ('moving'), is a branch of *mechanics* in which motion is considered from the geometric point of view only, irrespective of its "causes", i.e., *forces.*

Dynamics, from Greek δυναμις [dunamis], 'power'; δυναμικος [dunamikos], 'powerful', is a branch of *mechanics* that studies motion depending on the acting forces. Dynamics also refers to a state or progress of a change of some kind (e.g., *population dynamics*; see below).

Lastly, the word *mechanics* comes from Greek μηχανη [mekhane], 'machine', 'instrument' (and also 'invention', 'trick', 'tool', 'device'); μηχανικος [mekhanikos], 'mechanical' in the sense of 'pertaining to machines' (and also 'inventive', 'tricky'!). *Classical mechanics* (in contrast to *quantum* mechanics) studies movements *in space* from different points of view;[1] *theoretical mechanics* deals with general laws of motion and in a sense can be considered as a special branch of mathematics (one of the objects of study in mechanics is precisely *differential equations*).

Example 2. Motion without acceleration, i.e., *with zero acceleration*, is given by

$$a(t)=v'(t)\equiv 0, \quad \text{or} \quad x''(t)\equiv 0.$$

From the point of view of mathematics, the above relations are *differential equations*, i.e., equations in which unknowns are not numbers, as in algebraic equations, but *functions*, and we are given information (written as equations) about derivatives of the desired functions.

The *first-order* (involving the first derivative) differential equation $v'=0$ means that the rate of change of a function $v(t)$ is zero, i.e., the function "does not change", or is *constant*,[2] $v(t)\equiv 0=\text{const}$. A rigorous proof of this fact (*criterion of constancy of a function*: if $f'(x)\equiv 0$ holds on an

[1] We so far consider motion *along a straight line*, which is a *particular case* of motion in space; see Sec. 1.1.6 below.

[2] From Latin *constans*, *constantis*, 'standing firm'; hence also the name Constantine!

interval I, then $f(x) = \text{const}$ on I) uses the *Lagrange mean value theorem* (or *finite-increment formula*); the proof is given in § 2.1.

The *second-order* (involving the *second* derivative) differential equation $x'' = 0$ can be written as a "chain" (or *system*) of first-order differential equations:

$$x' = v, \qquad v' = 0.$$

We have already solved the second of these equations, $v(t) = v_0 = \text{const}$, so it remains to solve the first one, which takes the form $x' \equiv v_0$. A *constant rate* change can be described, for instance, by the *direct proportionality relation* $x = v_0 t$ (in this case $x' = (v_0 t)' = v_0$), but, as above, we can add an *arbitrary constant* to this function $v_0 t$. Thus, the coordinate varies according to the linear law $x(t) = v_0 t + \text{const}$, with $x(0) = 0 + \text{const} = \text{const}$, so the kinematical meaning of the constant is that it is the coordinate at zero time, $x(0) = x_0$.

Thus, the law of motion without acceleration is given by a *linear* function $x(t) = At + B$, where the constants A and B are uniquely determined by the *initial conditions*: $A = x'(0) = v(0) = v_0$, $B = x(0) = x_0$. This motion is said to be *uniform*. □

1.1.5. One-dimensional evolutionary models. In many cases, the state of a system is characterised by a single parameter, $x = x(t)$. From the point of view of mathematics, study of such a system is equivalent to the analysis of the *motion of an imaginary point* $X_t = X_t(x(t))$ *along a straight line*, the Ox axis. In this case the Ox axis plays the role of a set of all possible *configurations* (or *states*) of the system, generally called its *configuration space*. Here, the configuration space is *one-dimensional*, since the state is described by only one number, and we speak about *one-dimensional systems* (or systems *with one degree of freedom*) and about the corresponding *one-dimensional models*.

As we have already mentioned at the beginning of this section, one often knows (either from experiment or from theoretical considerations) some information about the rate of change of $x(t)$, i.e., about the derivative $x'(t)$. As a rule, the rate of change depends only on the value of the quantity itself, and we can write the corresponding *differential equation*

$$x'(t) = F(x(t)), \quad \text{or} \quad x' = F(x), \tag{10}$$

which is actually a *mathematical model* of the evolving one-dimensional system.

This situation is typical, for example, for models of evolution of biological systems belonging to a *single population*.[1] They are formulated as *laws of change of the population size* $x = x(t)$ written through the *rate* $x'(t)$ of change of this population size (measured in certain units; the population is assumed to be *sufficiently large* for the change of the population size $x(t)$ to be considered *continuous* and to speak about the *rate of change*). Let us give some simplest examples.

Example 3. *Linear model*, which yields, as will be seen in Ch. III, a fast (*exponential*) growth of the population size under *linear dependence of the growth rate on the population size* (under sufficient "living" condition, i.e., given enough living space and nutrient resources), is described by a *linear* differential equation

$$x'(t) = \alpha x(t).$$

A particular example: *"not too many fish in a very big pond"*.

This law of change is sometimes referred to as the *Malthus equation*, because it was presumably first derived in 1798 by the English economist and clergyman T.R. Malthus (1766–1734). Malthus applied this law to the description of the demographic situation on the global scale,[1] predicting overpopulation with a relative scarcity of people's means of subsistence (this concept was called *Malthusianism*). □

Example 4. *Logistic model* (or *Verhulst equation*) takes into account that at large population sizes, along with population *growth* proportional to the population size $n = n(t)$, there is also population *loss* proportional to the squared population size:

$$n'(t) = \alpha n(t) - \gamma n^2(t),$$

or, making the change $x = \gamma n$,

$$x'(t) = x(t)(\alpha - x(t)).$$

We will analyse this remarkable model in detail in Ch. IV. A particular example: *"fish in a not too large pond"*. □

Example 5. *Quadratic model*, where the growth rate of a population size is proportional to the squared population size:

$$n'(t) = \alpha n^2(t),$$

[1] Population (from Latin *populus*, 'people'), a group of individuals of the same species with a common habitat, capable of long-term persistence in time and space, as well as self-reproduction as an elementary unit of the evolutionary process. (This definition really lacks only a differential equation, does not it?!).

[1] *Demography* (from Greek δημος [demos], 'people,' and γραφω [grapho], 'write', 'record') is the science of human population.

or, after the change $x = \alpha n$,

$$x'(t) = x^2(t).$$

This model results, as will be shown in § 4.1, in a population "explosion". A similar equation describes some chemical reactions, predicting a real explosion. □

In all these examples, populations evolved on their own — without external influences, or interferences from the outside. Therefore, all the three differential equations were of the form (10), $x' = F(x)$, where the function on the right-hand side depends only on the value of the varying parameter x and does not depend on time. Such equations and models are said to be *autonomous.*[1] For real-world natural systems, this is typical; quoting the words of the famous Russian mathematician Vladimir Arnold,[2] "the laws of nature do not depend on time".

However, in some cases we have to take into account time-dependent external factors, for example, catching fish in the pond or, on the contrary, putting them into the pond (from time to time!). Models of such systems will now be *non-autonomous* differential equations of the form

$$x' = F(x, t) \quad (x'(t) = F(x(t), t). \tag{11}$$

Particular examples of systems of this kind will be given in Chs. III and V.

What is the goal of a researcher once the model of an evolutionary process has already been constructed (the corresponding differential equation has been written)? First, it is to *find all solutions* of a differential equation. Finding all solutions of this or that differential equation is also traditionally called *integration* of this differential equation, and before 1770s (before the works of Lagrange) the *solutions* themselves were called *integrals.* The reason for this will be clarified in the next chapter.

As will be shown in Ch. IV, any autonomous equation of the form (10) can be "integrated" in a certain sense. After that, one has to analyse properties of solutions of the equation depending on particular *initial conditions*, interpret them in terms of the original applied problem, and, possibly, predict potential "behaviour" of the system in question (which is one-dimensional in this case). After that, of course, it would be reasonable to verify the conclusions drawn; however, as we will see in § 3.4, experimental verification of theoretical conclusions might be hazardous.

[1] *Autonomous* (Greek αυτονομος [autonomos], from αυτο [auto], 'self' and νομος [nomos], 'custom', 'law') means 'self-governing', 'independent'.

[2] One of the world's foremost experts on differential equations and dynamical systems.

As will be clear from Ch. IV, the possibility to integrate a differential equation of the form (10) does not mean that its solutions can be expressed in *elementary functions* (which include rational and irrational algebraic functions; trigonometric, logarithmic, and exponential functions; as well as all functions obtained from these with the use of arithmetic operations and taking the *composition*, i.e., "function of a function"). Moreover, non-autonomous equations of the form (11) may turn out to be non-integrable at all (in the sense explained in Ch. IV). In both cases, *qualitative analysis* of the behaviour of solutions of the differential equation becomes necessary. We shall consider the simplest methods of such analysis in the next chapter (in § 2.2), but in general there exists a special qualitative *theory of differential equations*, which we shall mention in Ch. V.

1.1.6. Derivatives and velocities in mechanics (kinematics). Let us return to the relationship between the concepts of the derivative and the velocity. Real motions of real bodies take place in space. Modelled motions of point particles should also be considered in space, or sometimes (for instance, when we consider the motion of a stone thrown at an angle to the horizon) in a plane. For a kinematic description of such motions, a single numerical function is no longer sufficient. There are two ways out of the situation.

The first possibility is to consider instead of a numerical function a *vector-valued* function of time $t \mapsto \overline{r}(t) = \overline{OM_t}$, where O is a chosen *reference point*, and M_t is the position of the moving point at time t. The vector $\overline{r} = \overline{OM}$ is often called the *radius vector* of the point M, and the vector-valued function $\overline{r} = \overline{r}(t)$ is simply referred to as a *vector function.*

The term *radius vector*, which in fact appeared already in *Kepler's 2nd law* (1618), was proposed in 1853 by one of the founders of rigorous calculus, the famous French mathematician Augustin-Louis Cauchy (1789–1857). Latin *radius* means 'spoke' (in a wheel), 'ray'; it is noteworthy that the term *radius* was introduced relatively recently, only in 1569, by the French mathematician and philosopher Ramus (real name Pierre de La Ramée, 1515–24.8.1572; killed on St. Bartholomew's Day massacre).

The second approach is to choose a (Cartesian) coordinate system $Oxyz$ in space and consider *three* numerical functions

$$(x; y; z) = (x(t); y(t); z(t)),$$

which are the corresponding coordinates of the moving point M_t.

Of course, these approaches are interrelated: the coordinates of the vector $\overline{r}(t)$ are exactly $(x(t); y(t); z(t))$. However, while the second approach intrinsically suggests a coordinatewise definition of the velocity of motion

$$(v_x; v_y; v_z) = (x'(t); y'(t); z'(t)),$$

when describing the motion with the help of vector-valued functions it will be necessary to independently introduce the notion of their derivatives. This definition is quite natural and exactly follows the definition of *instantaneous velocity* in physics (mechanics).

Namely, to define the instantaneous velocity at time t, we consider the *increment of the position* over the time from t to $t+h$, i.e., the vector

$$\overline{\Delta r} = \overline{r}(t+h) - \overline{r}(t) = \overline{OM_{t+h}} - \overline{OM_t} = \overline{M_t M_{t+h}},$$

then by dividing by h (in mathematics,[1] by multiplying by h^{-1}) we compute the average velocity on this time interval, and then we find the *limit of the average velocity vector* as $h \to 0$; this *vector* (if the limit exists) is what is called the instantaneous velocity of motion, or the derivative of the vector-valued function $\overline{r}(t)$ at point (time instant) t:

$$\overline{v}(t) = \overline{r}\,'(t) = \lim_{h\to 0} \frac{\overline{\Delta r}}{h} = \lim_{h\to 0} \frac{1}{h}(\overline{r}(t+h) - \overline{r}(t)).$$

Here, the definition of the *limit of a vector-valued function* needs to be explained. It repeats verbatim the definition of the limit of an ordinary number-valued function and can be most simply written in the "Weierstrass (ε, δ)-language" (cf. the definition of an infinitesimal in Sec. 1.1.1): for a vector-valued function $\overline{a} = \overline{a}(h)$ and a vector $\overline{c}$ we have

$$\lim_{h\to 0} \overline{a}(h) = \overline{c} \quad \Leftrightarrow \quad \forall \varepsilon > 0\ \exists \delta > 0\ \forall h \in (-\delta, \delta)\ |\overline{a}(h) - \overline{c}| < \varepsilon$$

(as usual, the modulus, or the magnitude, or the absolute value of a vector is simply its length).

The question arises: how is the vector definition of velocity related to the coordinate definition? It is better to formulate the answer in terms of limits.

Theorem 1 (on the limit of a vector-valued function in coordinates)**.** *If a vector-valued function $\overline{a}(h)$ is written in the $Oxyz$ coordinate system by coordinate functions $(x(h); y(h); z(h))$ and a vector $\overline{c}$ has coordinates $(x_c; y_c; z_c)$, then the limit of $\overline{a}(h)$ as $h \to 0$ exists and equals $\overline{c}$ if and only if the limits of all the three coordinate functions $x(h); y(h); z(h)$ exist and are equal to $x_c; y_c; z_c$, respectively.*

Proof. Let us write the absolute value of the vector difference $|\overline{a}(h) - \overline{c}|$ in the definition of the limit through the coordinate functions:

$$|\overline{a}(h) - \overline{c}| = \sqrt{(x(h) - x_c)^2 + (y(h) - y_c)^2 + (z(h) - z_c)^2}.$$

This equality implies that if all the three squared coordinate differences tend to zero, then the absolute value of the vector difference under consideration tends to zero: existence of the limits of the coordinate functions implies existence of the limit of the vector-valued function, and the coordinates of this limit (vector) are the limits of the coordinate functions.

[1] In mathematics, vectors are *not divided* but only *multiplied* by numbers.

To prove the converse, note that we always have

$$\max\{|x(h)-x_c|,|y(h)-y_c|,|z(h)-z_c|\} \\ \le \sqrt{(x(h)-x_c)^2+(y(h)-y_c)^2+(z(h)-z_c)^2} = |\overline{a}(h)-\overline{c}|.$$

Hence, if the absolute value of the vector difference tends to zero, so do the absolute values of the coordinate differences. Therefore, the existence of the limit of the vector-valued function implies the existence of the corresponding limits of the coordinate functions, as remained to be established. □

Corollary 1 (on the derivative of a vector-valued function in coordinates)**.** *A vector-valued function $\overline{r}(t)$ is written in the Oxyz coordinate system by its coordinate functions $(x(t);y(t);z(t))$. The derivative $\overline{r}'(t)$ exists if and only if there exist derivatives $x'(t);y'(t)';z'(t)$ of the coordinate functions, and the coordinates of the vector $\overline{r}'(t)$ are $(x'(t);y'(t)';z'(t))$.*

Thus, it makes no difference whether we compute the velocity as a vector or compute it coordinatewise.

1.1.7. Geometric (kinematic) meaning of the velocity vector. The set of all positions M_t, $t\in\mathbb{R}$ (or t belonging to some time interval), is called the (mechanical) *trajectory* of a point's motion. If this motion is described by a vector-valued function $\overline{r}=\overline{r}(t)$, the trajectory is also referred to as the *hodograph* of this vector-valued function.

Trajectory is from Latin *trajectorius*, 'pertaining to throwing across'. The word *hodograph* comes from Greek words οδος [hodos], 'way', and γραφω [grapho], 'write', 'draw', so that its meaning is a 'road scribed by something'. Both the concept and the term were introduced by Sir William Hamilton, mentioned above, when studying the curvilinear motion of a point particle.

Note that in the case of non-zero displacement, the displacement vector $\overline{\Delta r}=\overline{M_tM_{t+h}}$ is directed along the secant M_tM_{t+h} to the trajectory (Fig. 3). The average velocity vector $t^{-1}\overline{\Delta r}$ and the instantaneous velocity vector $\overline{v}(t)=\overline{r}'(t)$ are directed similarly, provided that they are non-zero. In this case, *there exists a limit position of the secant M_tM_{t+h}* as $h\to 0$, i.e., as $M_{t+h}\to M_t$. Just as for function graphs (see Sec. 1.1.2), the limit position of the secant is called the tangent to the graph, so it follows from the above reasoning that *the velocity vector of a moving point particle is directed along the tangent line* (Fig. 4).

On the other hand, the magnitude of the vector of the derivative, i.e., $|\overline{v}(t)|=|\overline{r}'(t)|$, the absolute value of the velocity, turns out to be equal to the *rate of change of the travelled path length*, i.e., to the velocity of motion of a point along the hodograph. We will not prove this intuitively clear statement.

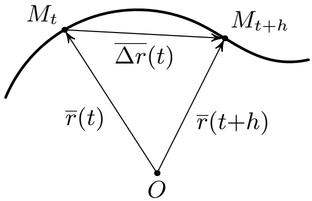


Fig. 3.

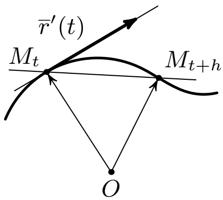


Fig. 4.

Example 6. One of the basic types of motion is the *rotational motion in a plane*, i.e., rotation about the centre O (which is convenient to choose as the initial point of a radius vector) with a constant angular velocity $\omega \, \frac{\text{rad}}{\text{s}}$ or simply $\omega \, \text{s}^{-1}$. It is convenient to represent it using rotations of vectors:

$$M_t = R^{\omega t} M_0, \quad \text{or} \quad \overline{r}(t) = R^{\omega t}(\overline{OM_0}) = r_0 \cdot R^{\omega t}\overline{e},$$

where $r_0 = OM_0$ is the distance from the rotation centre to the point and $\overline{e}$ is a unit vector directed along the OM_0 line. If we choose a (Cartesian) coordinate system Oxy in the rotation plane pointing the Ox axis along $\overline{e}$, then, *by the definition of cosine and sine*, the coordinates of the vector $R^{\omega t}\overline{e}$ are $(\cos \omega t; \sin \omega t)$, so the coordinate representation of the above vector-valued function is

$$\overline{r}(t) = (r_0 \cos \omega t; r_0 \sin \omega t).$$

Hence by Corollary 1 it follows that

$$\overline{r}'(t) = (r_0 (\cos \omega t)'; r_0 (\sin \omega t)').$$

On the other hand, the velocity vector $\overline{r}'(t)$ when rotating with angular velocity ω is easy to find using its geometrical (kinematic) meaning: it is tangent to the trajectory, i.e., to the circle, in the rotation direction, and its absolute value is the linear velocity of motion along the circle, i.e., $v_0 = \omega r_0$ (Fig. 5).

Since the tangent to a circle at a point M_t is perpendicular to the radius OM_t, the vector $\overline{r}'(t)$ is obtained from $\overline{r}(t)$ by rotating by $90°$, or $\frac{\pi}{2}$ radians followed by multiplication by ω, so we can write

$$\overline{r}'(t) = \omega r_0 R^{\omega t + \frac{\pi}{2}} \overline{e} = \Big(\omega r_0 \cos\big(\omega t + \tfrac{\pi}{2}\big); \Big(\omega r_0 \sin\big(\omega t + \tfrac{\pi}{2}\big)\Big).$$

Comparing this formula with the preceding one, we obtain *beautiful formulae for differentiating the cosine and sine*:

$$(\cos \omega t)' = \omega \cos\big(\omega t + \tfrac{\pi}{2}\big), \qquad (\sin \omega t)' = \omega \sin\big(\omega t + \tfrac{\pi}{2}\big).$$

Of course, using the *reduction formulae*, one can also derive the standard expressions from them: $(\cos \omega t)' = -\omega \sin \omega t$ and $(\sin \omega t)' = \omega \cos \omega t$. These

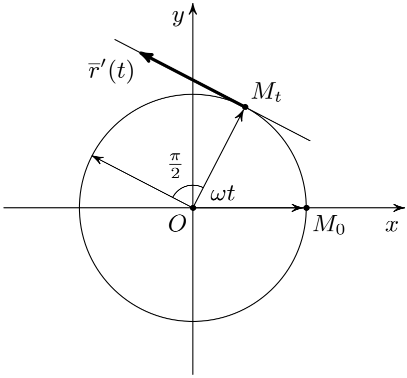


Fig. 5.

are some remarkable results that can be obtained from simple geometric-kinematic considerations. □

1.1.8. Two-dimensional Lotka–Volterra evolutionary model. When studying evolving systems, one often has to consider several parameters, say two, characterising their states. Then one should consider *two* functions, $x = x(t)$ and $y = y(t)$, describing the variation of both parameters in time. With the pair of functions we can associate an *imaginary* point M_t with coordinates $(x(t); y(t))$ and consider its *imaginary* motion in a *configuration space*, which in this case is *two-dimensional*; i.e., it is the *configuration plane* Oxy or some part of it.

Example 7. Consider the simplest two-dimensional evolutionary model of a *biocoenosis*[1] consisting of two interacting populations, where *carp and pike coexist in a pond.* Of course, this is not peaceful coexistence: *predators* (pikes) live only by eating *prey* (carp). The prey themselves are *"herbivores"*, i.e., they feed only on plant food, which is plentiful.

Denote the size of the prey population by $x = x(t)$ and the size of the predator population by $y = y(t)$. In the absence of predators, the *x-population* would grow according to the Malthus law: $x' = ax$. In the presence of the *y-population* of predators, along with the growth of the x-population at a rate of ax, the "loss" of prey must be taken into account. The simplest hypothesis regarding the loss is based on the "encounter theory", i.e., on the assumption that the rate of decrease in the number of prey is proportional to the number of possible encounters between a predator and a prey, i.e., the product of the sizes of the x-population and the y-population (it is

[1] *Biocoenosis* (from Greek βιος [bios], 'life', and κοινος [koinos], 'common') is an area of common habitat for animals, plants, and microorganisms.

assumed that *each* predator can encounter *each* prey and with some probability eat it!). This hypothesis leads to the following differential equation for the x-population size:

$$x'(t) = ax(t) - bx(t)y(t)$$

(the constants a and b, as well as the constants c and d appearing below, are positive).

Now let us look at predators. Their growth occurs only by eating prey, and, according to the same "encounter theory", the rate of growth is proportional to the product of sizes of the x-population and the y-population. This gives a term of the form $c \cdot x \cdot y$ in the rate $y'(t)$ of change in the size of the y-population. However, we must also take into account the natural loss of predators, which is proportional to their number and plays a significant (as will be seen later, *stabilising*) role when the y-population is large (numerous) or the x-population is small (few), i.e., food (prey) becomes scarce. Thus, the rate $y'(t)$ of change in the y-population is written as

$$y'(t) = cx(t)y(t) - dy(t).$$

Thus, we come to a system of two equations

$$\begin{cases} x'(t) = ax - bxy, \\ y'(t) = cxy - dy \end{cases} \Leftrightarrow \begin{cases} x'(t) = x(a - by), \\ y'(t) = y(cx - d), \end{cases}$$

considered only in quadrant I ($x > 0,\ y > 0$) of the Oxy plane (the "prey–predator" plane). This system of differential equations is the simplest model of the *"struggle for existence"*, or, otherwise called, model of the *"predator–prey" system.* This model (the corresponding system of equations) is called the *Lotka–Volterra model.*

Vito Volterra (1860–1940) was an Italian mathematician who worked in both "pure" and applied mathematics: mathematical physics, applications of mathematics in biology and in social research. It was he who developed the mathematical theory of the struggle for existence and population dynamics. In the 1920s and later, Volterra proposed not only the above model of biocoenosis development but also a number of others (he published about 30 papers on these topics).

Alfred James Lotka (1880–1949) was an American mathematician who proposed the same model of biocoenosis independently of Volterra and later continued his developments.

We will analyse the Lotka–Volterra model in a special section at the end of Part I of this book (in § 4.4). □

The derived system of equations is a special case of an *autonomous system of two differential equations* written as

$$\begin{cases} x'(t) = A(x, y), \\ y'(t) = B(x, y) \end{cases}$$

and treated as a *planar differential equation.* Solutions of such a system on some time interval, $t \in I$, are *pairs* of differentiable functions $x(t)$, $y(t)$ such that for any $t \in I$ we have

$$\begin{cases} x'(t) = A(x(t), y(t)), \\ y'(t) = B(x(t), y(t)). \end{cases}$$

If along with the imaginary point M_t with coordinates $(x(t); y(t))$ in the configuration plane Oxy (or in its part, as in the Lotka–Volterra model) we consider the corresponding vector-valued function $\overline{r}(t) = \overline{OM_t}$ and attach to each point $M(x; y)$ of the plane a *vector* $\overline{V}(x, y)$ with coordinates $(A(x, y); B(x, y))$, then, according to Corollary 1, the above system of differential equations can be written as a *single vector differential equation*

$$\overline{r}'(t) = \overline{V}(x, y) = \overline{V}(\overline{r}) \tag{12}$$

(we have replaced the point $M(x; y)$ with its radius vector $\overline{r} = \overline{OM}$). It follows from the geometric meaning of the derivative of a vector function that the (imaginary) trajectory $t \mapsto M_t$ corresponding to a solution $\overline{r}(t)$ of the above equation *is at each of its points tangent to the vector* $\overline{V}(x, y) = \overline{V}(\overline{r})$ *attached to this point.* This is the *geometric meaning of the vector differential equation* (12) or of the original system of differential equations.

Note that if the state of the system is described by three parameters, it is analogously identified with a point in space. If the parameters are four, then the configuration space is *four-dimensional*, and so on.

The general notion of a *configuration space in arbitrarily many dimensions*, which turned out to be very convenient and fruitful for mathematical research, was introduced at the end of the XIX century by the prominent American physicist, mechanic, and mathematician Josiah Willard Gibbs (1839–1903), a Yale University graduate and professor, one of the founders of thermodynamics and rigorous statistical mechanics.

§ 1.2. Concept of Dynamical Systems

In modern mathematics, dynamical systems are *any* systems (of any dimension) whose evolution is governed by an appropriate system of differential equations. In this section we will consider the simplest dynamical systems of classical mechanics: motion of a point particle along a straight line identified with the Ox coordinate axis subject to a force directed along this line. In this case there is no need to consider vector-valued functions; as was explained in Sec. 1.1.4, we may limit ourselves to scalar functions, namely, to the coordinate function $x(t)$ and its derivatives: the scalar velocity $v(t) = x'(t)$ and acceleration $a(t) = v'(t) = x''(t)$.

In this case, the force can also be considered not as a vector but as some scalar function of, generally speaking, coordinate, velocity, and time,

$F = F(x, v, t)$. The *fundamental law of dynamics*, i.e., *Newton's second law of motion*, can be expressed as a *second-order differential equation* $ma = mx'' = F(x, v, t)$. Assuming the mass m entering the equation to be 1,[1] we write *Newton's differential equation*, which is the main one in this section and in Ch. V, as

$$x'' = F(x, v, t). \tag{1}$$

This differential equation is actually the mathematical model of the corresponding dynamical system. Such a system is called a *dynamical system with one degree of freedom*, but from the point of view outlined in the preceding section, it is a *two-dimensional evolutionary model.* We will start with explanations on this point.

1.2.1. Newton's system of equations. Phase plane. A numerical function $x(t)$ is said to be a *solution* of differential equation (1) on an interval $I \in \mathbb{R}$ if it is twice differentiable (i.e., has a second derivative) and

$$\forall t \in I \quad x''(t) = F(x(t), x'(t), t).$$

As shows the simplest Example 2 from the preceding section (the case of $F \equiv 0$, equation $x'' = 0$), to *uniquely* determine a solution of Newton's equation, two *initial conditions* must be specified: initial coordinate and velocity. It is not obligatory to set them exactly at $t = 0$; any time $t = t_0$ can be chosen as the initial one. As physicists believe, initial conditions, i.e., the quantities

$$x(t_0) = x_0, \qquad x'(t_0) = v(t_0) = v_0, \tag{2}$$

together with differential equation (1) uniquely determine *the law of motion* $x = x(t)$ *at all times* $t \in \mathbb{R}$. In good cases (and these are the only cases we will consider) this is indeed true.

Alternatively, the need for two initial conditions can (and should) be interpreted as the need for *two* quantities, coordinate and velocity, to precisely, unambiguously describe the state of our one-dimensional dynamical system. For each of these quantities, x and v, as in Sec. 1.1.8, we can write a *first-order* differential equation. For this purpose, as in Example 2 in Sec. 1.1.4, we "split" the second-order differential equation (1) i.e., write it in a "chain", as a system:

$$\begin{cases} x' = v, \\ v' = F(x, v, t). \end{cases} \tag{3}$$

[1] Or, alternatively, dividing both parts of the equation by m and introducing the same notation F for $\frac{1}{m}F$.

This system is called *Newton's system of differential equations*, and it is clear that it makes no difference whether we solve equation (1) or the system of equations (3).

Since the state of a one-dimensional dynamical system is described by a pair of parameters $(x; v)$, it is natural to interpret this pair as an (imaginary) point $M(x; v)$ in the *phase plane Oxy*. A solution $x(t)$ of equation (1) satisfying the initial conditions (2) corresponds to a solution $(x(t); v(t)) = (x(t); x'(t))$ of Newton's system (3) and an imaginary (and depictable!) trajectory $t \mapsto M_t(x(t); v(t))$ in the phase plane, which in this case is called the *phase trajectory*. Studying a one-dimensional dynamical system assumes finding or, at least, qualitatively describing *all* phase trajectories that form the *phase portrait* of the system in the phase plane. In the general case this is an intractable problem. However, it can be solved if we impose some reasonable conditions on the force $F = F(x, v, t)$. Before formulating them, we will consider a particular example.

The notion of a *phase plane*,[1] as well as the more general notion of a *phase space*, was introduced by J. Willard Gibbs, mentioned above, when developing statistical mechanics and vector (or *multivariate*) calculus. Note that the state of the simplest spatial dynamical system *"one point particle in space"* is described by *six* parameters: three coordinates of the particle and three coordinates of the velocity vector, so in this case the phase space is *six-dimensional*. When considering *two* particles in space, the phase space will be twelve-dimensional, etc.

This "multivariance" of problems considered by Gibbs prompted him to develop vector calculus, a relatively new at that time branch of calculus in which various generalisations of differentiation and integration operations on "multivariate objects" (vector functions, functions of vectors, vector fields, etc.) are studied. This branch was developed in the middle XIX century, mainly by Irish mathematician and astronomer William Rowan Hamilton (1805–1865; about him, see above) and German mathematician, physicist, and philologist Hermann Günther Grassmann (1809–1877), who introduced *multidimensional spaces* (and multidimensional vectors), discovered the laws of colour addition (he interpreted colour perceptions as three-dimensional vectors), and translated from Sanskrit the *Rigveda*, a monument of ancient Indian literature. Hamilton and Grassmann were the first to develop a rigorous theory of the so-called *complex numbers* (see § 5.7). Vector calculus was finally formalised into a rigorous mathematical science (and given this name) by Gibbs in 1881–1901.

1.2.2. Uniformly accelerated motion and free fall. The simplest case after the above example $F \equiv 0$ is the situation where the acting force is *constant* and hence the acceleration is also constant: $x'' = F \equiv a = \text{const} \neq 0$. Writing this equation as a chain of first-order equations (system (3) again)

$$x' = v, \qquad v' = a,$$

[1] This term is due to the fact that formerly *states* of a system were often referred to as *phases*. *Phase* is from Greek φασις [phasis], 'appearance'. In § 5.3 this word will appear in yet another sense.

we solve them also in a chain, starting from the last equation. As we have already seen in Sec. 1.1.4,

$$v'(t) \equiv a \quad \Leftrightarrow \quad v(t) = at + B,$$

where $B \in \mathbb{R}$ is an arbitrary constant. To solve the resulting first equation $x' = at + B$, we separately solve the equations $x_1' = at$ and $x_2' = B$. From the latter equation, as above, we find $x_2(t) = Bt + B_1$. The first equation is satisfied by a quadratic function $x_2 = \frac{1}{2}at^2$, and along with it by any function of the form $x_2(t) = \frac{1}{2}at^2 + B_2$ (B_2, as well as B_1, are arbitrary constants). Summing up, we obtain the general solution of the "aggregate" equation:

$$x' = at + B \quad \Leftrightarrow \quad x(t) = \tfrac{1}{2}at^2 + Bt + C.$$

Thus, we have derived the *general law of uniformly accelerated motion*; it turns out to be quadratic and involves two arbitrary constants, B and C, which are determined from the initial conditions (2). Assuming $t_0 = 0$ in them, we obtain

$$x(0) = C = x_0, \quad v(0) = B = v_0 \quad \Rightarrow \quad x(t) = \tfrac{1}{2}at^2 + v_0 t + x_0.$$

This law (as a position-time dependence) was first discovered experimentally by Galileo Galilei (1564–1642) when performing experiments on throwing pebbles from the famous Leaning Tower of Pisa and then studying the rolling of a ball down an inclined plane. In both cases, the main force is the gravity force $F = -mg$ (what is the force when rolling along a plane inclined by angle α above the horizon?). After the discovery of calculus, solving the corresponding problem became a trivial theoretical computation.

Example 1. Let us construct the *phase portrait* of the system describing the motion of a point particle of mass m attracted by the gravity force $F = -mg$ *along the vertical axis* Oz directed *upwards*. This motion obeys the differential equation of the form just considered:

$$z'' = -g \quad \Leftrightarrow \quad z(t) = -\tfrac{1}{2}gt^2 + Bt + C, \quad v(t) = z'(t) = -gt + B.$$

To draw the phase trajectories in the Ozv plane, we express t through v from the last equation and substitute this expression into the law of motion $z = z(t)$:

$$t = -\tfrac{1}{g}(v - B) \quad \Rightarrow \quad z = -\tfrac{1}{2}g\tfrac{1}{g^2}(v - B)^2 - B\tfrac{1}{g}(v - B) + C$$
$$= -\tfrac{1}{2g}v^2 + \tfrac{1}{g}Bv + \text{const} - \tfrac{1}{g}Bv + \text{const} + C = -\tfrac{1}{2g}v^2 + \text{const}$$

(in these calculations, the same symbol const denotes different and, generally speaking, *unequal* constants, whose sum is replaced by a single constant in the last expression). Thus, the phase trajectories in the Ozv plane geometrically represent horizontally placed parabolas obtained from the parabola $z = -\frac{1}{2g}v^2$ by all possible parallel translations along the Oz axis (Fig. 6). A phase point moves along them according to the derived expressions for $(x(t); y(t))$: *to the right* in the upper half-plane ($v > 0$, so that the coordinate $x(t)$ increases) and *to the left* in the lower half-plane (according to $v < 0$).

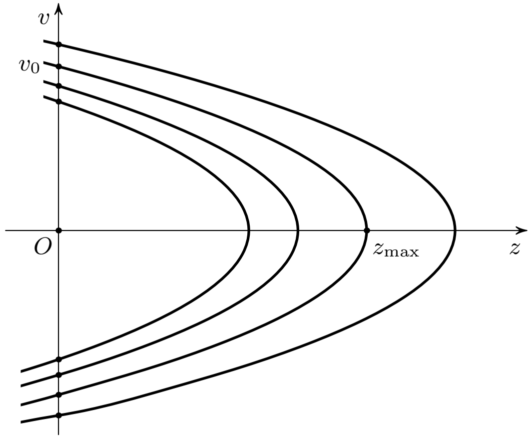


Fig. 6.

The vertex of each parabola corresponds to the maximum possible value of z, i.e., the maximum height of a stone thrown upwards from the ground level $z = 0$. For the trajectory $z = -\frac{1}{2g}v^2 + \text{const}$, we obviously have $z_{\max} = \text{const}$, and if we want to find the maximum height given the initial velocity v_0 at the ground level $z = 0$, we must substitute the pair $(z; v) = (0; v_0)$ into the formula for the constant:

$$z_{\max} = \text{const} = z + \frac{1}{2g}v^2 = 0 + \frac{1}{2g}v_0^2 = \frac{v_0^2}{2g}.$$

Note a remarkable equality that holds on *any phase trajectory*:

$$z + \frac{1}{2g}v^2 = \text{const} \Leftrightarrow gz + \frac{1}{2}v^2 = \text{const} \quad \Leftrightarrow \quad mgz + \frac{1}{2}mv^2 = \text{const},$$

or, in other words, $U(z) + T(v) = \text{const}$, where $U(z) = mgz$ is the *potential energy* of a point particle at height z above the ground level (which is taken as the reference level of potential energy), and $T(v) = \frac{1}{2}mv^2$ is the *kinetic energy* of a particle of mass m moving with velocity v. Thus, *the equations of phase trajectories in the phase plane Ozv coincide with the energy conservation law* for the considered system!

This very important fact allows one to construct phase portraits of much more general one-dimensional dynamical systems and to find (in a certain sense) laws of motion in them. We will return to the energy conservation law at the beginning of Ch. V.

In this case, the energy conservation law explains why the maximum height of the thrown stone is exactly as described above: from the equality $\frac{1}{2}mv^2 + mgz = \text{const}$ it follows that the height is maximum at $v = 0$, i.e., when *all the kinetic energy* $T = \frac{1}{2}mv_0^2$ *imparted to the stone when throwing it from level* $z = 0$ *"flows" into the potential energy* $U = mgz_{\max}$:

$$mgz_{\max} = \tfrac{1}{2}mv_0^2,$$

whence the above formula for $z_{\max}$. □

1.2.3. Two-dimensional dynamical system (example). Now it is a good time to consider the behaviour of a dynamical system with *two* degrees of freedom, well known from the school course of physics: the motion of a stone (or a projectile; but actually just a point particle) thrown at an angle to the horizon. This motion takes place in a vertical plane, to which we associate a coordinate system Oxz by choosing the origin at the "launch point" and pointing the Ox axis horizontally in the direction of motion and the Oz axis upwards. The motion obeys Newton's *vector* differential equation, which is conveniently written using coordinates:

$$m(x;z)'' = \overline{F} = (0; -mg) \quad \Leftrightarrow \quad x'' = 0 \quad \text{and} \quad z'' = -g \tag{4}$$

(here we have taken into account that the gravity vector is directed downwards).

The phase space of this system is *four-dimensional*: its state is determined by two coordinates, x and z, and two components (coordinates along the chosen axes) of the velocity, $v_x = x'$ and $v_z = z'$, so consideration of the *phase trajectory* is in this case difficult (see, however, the *Task* at the end of this subsection). We will consider not an imaginary phase trajectory but the true *configuration trajectory* described by the moving point $P_t(x(t); z(t))$ in the configuration plane Oxz. To this end, we have to solve a system of Newton's second-order differential equations (4). Previously, we split each of these equations into systems of two first-order equations:

$$x'' = 0 \quad \text{and} \quad z'' = -g \quad \Leftrightarrow \quad \begin{cases} x' = v_x, \\ v_x' = 0 \end{cases} \quad \text{and} \quad \begin{cases} z' = v_z, \\ v_z' = -g, \end{cases}$$

which we then solved "in a chain". In this case this is not necessary, since the equations in system (4) are *independent*, i.e., each of them can be considered and solved separately from the other.

Thus, using the results of Sec. 1.1.4 and Example 1 just discussed and taking into account the initial conditions $x(0) = z(0) = 0$, $v_x(0) = x'(0) = v_1$, and $v_z(0) = z'(0) = v_2$ (v_1 and v_2 being the horizontal and vertical components of the initial velocity, respectively), we write

$$x'' = 0 \quad \text{and} \quad z'' = -g$$
$$\Leftrightarrow \quad x(t) = At + B = v_1 t \quad \text{and} \quad z(t) = -\tfrac{1}{2}gt^2 + Ct + D = -\tfrac{1}{2}gt^2 + v_2 t.$$

Thus, it turns out that the x-coordinate depends on time linearly and the z-coordinate depends on time quadratically. Since (when $v_1 \neq 0$) the x-coordinate and time are simply *proportional*, the trajectory is described by a *quadratic dependence* of z on x, so *a thrown stone follows a parabola.*[1] Of course, we can write this dependence, investigate it, and construct the configuration trajectory $z = z(x)$ (Fig. 7), but all essential characteristics of this parabola can be found from the laws of variation $x = x(t)$ and $z = z(t)$ already described, and this way it will be even easier (you can compare it yourselves).

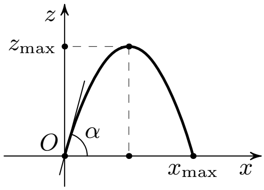


Fig. 7.

Let us first find the largest throwing distance $x_{\max}$, assuming that the absolute value v of the initial velocity is fixed, so that $v_1 = v\cos\alpha$, $v_2 = v\sin\alpha$, where α is the angle between the direction of the "throw" and the horizon. The quantity $x_{\max}$ corresponds to the case where the height z of the stone becomes zero again. From this the corresponding time $t_{\max}$ (the total duration of flying) and then the value of $x_{\max}$ are found:

$$z(t) = -\frac{1}{2}gt^2 + v_2 t = 0 \quad \Leftrightarrow \quad t_{\max} = \frac{2v_2}{g},$$
$$x_{\max} = x(t_{\max}) = v_1 t_{\max} = \frac{2v_1 v_2}{g} = \frac{v^2}{g} \cdot 2\cos\alpha\sin\alpha = \frac{v^2}{g}\sin 2\alpha.$$

Note that if the angle α varies from $0°$ to $90°$, then 2α varies from $0°$ to $180°$, so the obtained value of $x_{\max}$ is the greatest when the argument of the sine in the last expression is $90°$, i.e., $2\alpha = 90°$, $\alpha = 45°$. Thus we arrive at a well-known result: *the greatest flight distance is when the stone is thrown at an angle of* $45°$ *to the horizon.*

Let us also find the maximum height of the thrown stone for a given value of the angle α. This is the coordinate $z_{\max} = z_0$ of the vertex of the parabola $z = z(x)$, and to find it, it suffices to find the time t_0 when it is reached. In principle, for this purpose we can equate the vertical component of velocity $v_z = z' = -gt + v_2$ to zero, but even without this, for reasons of symmetry of the parabola and uniformity of motion along it in the horizontal direction ($v_x \equiv v_1$), it is easy to realise that the maximum height is reached

[1]This fact was also first discovered by Galileo Galilei (and also *by experiments*).

halfway to the touchdown, i.e., t_0 is *half as large* as $t_{\max}$:

$$t_{\max} = \frac{2v_2}{g} \quad \Rightarrow \quad t_0 = \frac{v_2}{g}$$
$$\Rightarrow \quad z_{\max} = z_0 = z(t_0) = -\frac{1}{2}gt_0^2 + v_2 t_0 = -\frac{1}{2g}v_2^2 + \frac{1}{g}v_2^2 = \frac{1}{2g}v_2^2 = \frac{v^2}{2g}\sin^2\alpha.$$

The formula shows that the maximum height is the greater the closer the angle of throwing to 90°; this is quite natural: the more "vertically" we throw the stone, the greater the height it will reach.

The same can be obtained from the energy conservation law in this system: the squared velocity in this case is $|\overline{v}|^2 = v_x^2 + v_z^2$, the potential energy is $U = U(z) = mgz$, and the energy conservation law can be written as

$$U + T = mgz + \frac{1}{2}m(v_x^2 + v_z^2) = \text{const.}$$

The horizontal velocity $v_x \equiv v_1$ is also constant, so

$$U + T_z = mgz + \frac{1}{2}mv_z^2 = \text{const;}$$

i.e., the conservation law for vertical motion "splits off" from the general conservation law. Hence, this motion is exactly the same as was described in Example 1, with the "throwing upwards" initial velocity $v_z(0) = v_2$. The maximum height will be reached when all the initial kinetic energy of the vertical motion turns into potential energy, i.e., when

$$T_z(0) = \frac{1}{2}mv_2^2 = U(z_{\max}) = mgz_{\max},$$

whence we obtain the same expression for $z_{\max} = z_0$.

Let us nevertheless try to realise what the *phase* trajectory of the considered motion in a four-dimensional phase space with coordinates $(x; y; v_x; v_y)$ looks like. We can write the laws of variation for all coordinates:

$$x(t) = v_1(t), \qquad z(t) = -\frac{1}{2}gt^2 + v_2 t, \qquad v_x(t) = v_1, \qquad v_z(t) = -gt + v_2.$$

From the third equation we obtain that the coordinate v_x remains unchanged, so we get a correct idea of the phase trajectory by considering the motion of a point $\widetilde{M}_t(x(t); z(t); v_z(t))$ in a standard three-dimensional space with coordinates $(x; z; v_z)$. (We could say it another way: just as the equation, for instance, $y = \text{const}$ identifies a *plane* [parallel to the coordinate plane Oxz] in three-dimensional space, the equation $v_x \equiv v_1 = \text{const}$ also identifies a *three-dimensional subspace* [parallel to the aforementioned *three-dimensional coordinate subspace* $Oxzv_z$!] in the four-dimensional phase space.)

Task. Describe the phase trajectory

$$x(t) = v_1(t), \qquad z(t) = -\frac{1}{2}gt^2 + v_2 t, \qquad v_z(t) = -gt + v_2$$

in the three-dimensional space $Oxzv_z$.

Hint. This is a *plane curve*. Which one?

Now let us return to considering one-dimensional dynamical systems.

1.2.4. Conservative one-dimensional systems. Ball-in-a-trough model. In the case where a mechanical system is not affected by external influences, the force F entering Newton's equation (1) is not explicitly time-dependent,[1] i.e., $F = F(x, v)$ is a function of only the state of the system, i.e., of the coordinate and velocity (or phase point $(x; v)$).

As an example of a velocity-dependent force, we can mention the *viscous friction force*, $F_{\text{fr}} = -\lambda v$, directed opposite to the velocity of motion and proportional to it (see Sec. 3.1.3). The presence of such a force leads to the so-called *dissipation of energy*,[2] roughly speaking, transformation of the energy of motion into heat. In the ideal case, friction is assumed to be absent and the force F to depend *only on the coordinate*, $F = F(x)$. Such forces and the corresponding dynamical systems are called *conservative*. This name is due to the fact that such systems *obey the energy conservation law* (the word *conservative* is from Latin *conservare*, 'to protect', 'to preserve'; *conservativus*, 'protective', 'preserving'), which we shall explain and prove in detail at the beginning of Ch. V (in § 5.1).

Newton's differential equation for conservative systems

$$x'' = F(x) \tag{5}$$

can be replaced by a system of Newton's equations

$$\begin{cases} x' = v, \\ v' = F(x), \end{cases} \tag{6}$$

which, though simpler than the general system (3), does not admit an easy solution: the system "tangles" the variables x and v with each other.

We devote a special chapter to the study of such and somewhat more complicated equations and systems (Ch. V). But here we are going to consider a very general illustrative model, which will later allow us to find and analyse phase portraits of arbitrary conservative systems.

Let a (point) ball be rolling without friction in a flat trough having the form of a graph of some *differentiable* function $h = u(x)$ (Fig. 8). The position of the ball is determined by its coordinate on the Ox axis, and the force acting on the ball also depends only on x: it depends on the derivative of the function $u(x)$ and is pointed in the decreasing direction of this function.

[1] Recall: *the laws of nature do not depend on time*!

[2] From Latin *dissipatio*, 'scattering'.

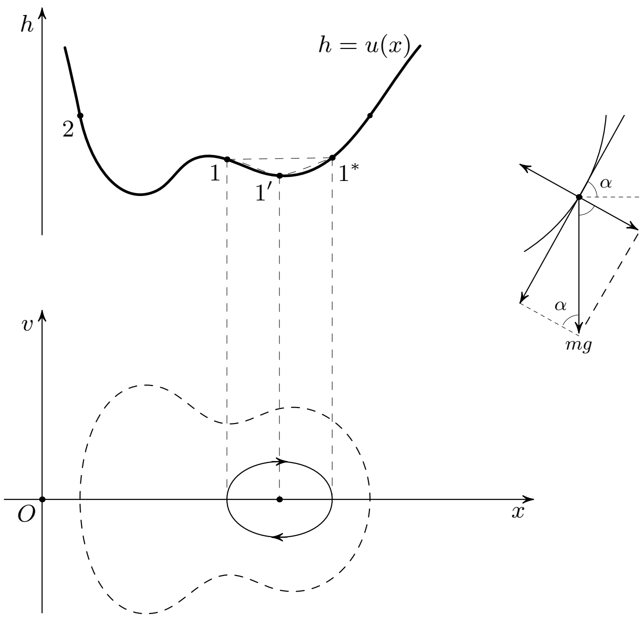


Fig. 8.

Indeed, there are two forces acting on the ball: the downward force of gravity of magnitude mg and the *support reaction force*, directed along the normal (perpendicular)[1] to the graph $h = u(x)$. This force is balanced by the normal component of the force of gravity, equal to $mg \cos\alpha$, where α is the angle between the tangent and the Ox axis (see Fig. 8, right), so that the *resultant force is the tangential*[2] component of the gravity force, i.e., $-mg\sin\alpha$, $\alpha = \alpha(x) = \arctan u'(x)$. The minus sign in front of the tangential component shows the *direction* of the force: if $u' > 0$, the force is directed *to the left*, and if $u' < 0$, it is directed *to the right.* Note that when $u' = 0$, i.e., when the tangent to the "trough profile" $h = u(x)$ is horizontal, the resultant is zero, so the *critical points*[3] of $u(x)$ correspond to *equilibrium positions* (*equilibrium points*, or simply *equilibria*) of the ball-in-a-trough system.

We will not write Newton's equation for this system, since this requires to consider the so-called *generalised coordinate*, but will confine ourselves

[1]The *normal* (from Latin *normalis*, 'made according to a carpenter's square') to a curve is the perpendicular to the tangent to the curve drawn at the tangency point.

[2]From Latin *tangens*, *tangentis*, 'touching'.

[3]Recall that for a differentiable function, critical points are zeros of its derivative.

to a qualitative analysis of the behaviour of the ball under various initial conditions.

It is clear that if the ball is simply released (with zero velocity) at some point 1 "on the right slope of the hill on the trough" (see Fig. 8), it will roll to the right, its velocity will increase until it reaches the "pit in the trough" (equilibrium point $1'$ in Fig. 8), and then decrease during the time that the ball climbs up the "right slope of the pit". This will continue until the ball reaches on the right slope the height 1^* from which it was released — at point 1^* the instantaneous velocity will be zero, and the ball will immediately start to roll back, return to point 1, then roll to the right again, and so on. We can sketch a rough phase curve $t \mapsto M_t(x(t); v(t))$: it is a closed line extending from point 1 to point 1^* and back, and the phase point M_t periodically moves clockwise along it (Fig. 8, bottom).

This motion of the ball and the corresponding imaginary motion of the phase point along the phase curve can be explained using the *energy conservation law* for the ball-in-a-trough system. The kinetic energy of the motion is $T = \frac{1}{2}mv^2$ (the velocity v is tangent to the graph $h = u(x)$); the potential energy of the ball at height h counted from the zero energy level, for which we will take the Ox axis, is $U = mgh = mgu(x)$; *the total mechanical energy is*

$$E = U + T = mgu(x) + \frac{1}{2}mv^2 = \text{const}.$$

As the height $h = u(x)$ decreases, the potential energy of the ball decreases, turning into kinetic energy, and the velocity increases. The minimum points of $u(x)$ correspond to minima of the potential energy and, accordingly, to the largest values of the velocity v, and so on.

Task. Describe the character of the ball's motion after it is released at point 2 on the "leftmost slope of the trough". Will the ball roll over the hill in the middle?

Hint. Use the phase curve shown by the dashed line in the Oxv plane (Fig. 8, bottom).

1.2.5. Example of a conservative system: ball on a spring. Let us consider a very concrete and easily modelled (both experimentally and theoretically) one-dimensional dynamical system of oscillation of a ball attached to a spring with a given stiffness (or elasticity) coefficient.

Example 2. Let a ball (point particle) of mass m be attached to the center of a horizontally placed spring with fixed (stationary) ends. Let us take the center of the spring as the origin O of the coordinate axis Ox directed along the spring (Fig. 9). According to *Hooke's law*, when the ball is displaced from the equilibrium point O to a point with coordinate x, it is subject to the *elastic force* proportional to the displacement x and pointed

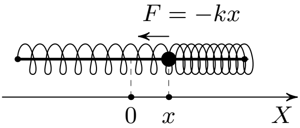


Fig. 9.

in the opposite direction: $F = F(x) = -kx$, where $k > 0$ is the elasticity coefficient of the spring (this law and formula are valid for relatively small displacements x, though comparable to the length of the spring, which, in fact, was established experimentally by *Robert Hooke* in 1660).

Thus, we have a conservative dynamical system modelled by Newton's equation of the form

$$mx'' = -kx, \quad \text{or} \quad x'' = -\frac{k}{m}x = -\omega^2 x, \quad \text{where} \quad \omega = \sqrt{\frac{k}{m}}. \qquad (7)$$

How to find its solutions?

In this case, unexpected help comes from the kinematics side: we know from Example 6 of Sec. 1.1.7 that when differentiating the functions $a\cos\omega t$ and $b\sin\omega t$, the argument (often called the *phase*) of the function is increased by $\frac{\pi}{2}$ and a factor of ω appears in front of the function. After the second differentiation we obtain that

$$(a\cos\omega t)'' = \omega^2 \cdot a\cos(\omega t + \pi), \qquad (b\sin\omega t)'' = \omega^2 \cdot a\sin(\omega t + \pi).$$

But $\cos\alpha$ and $\sin\alpha$ are by the definition the coordinates of the vector $R^\alpha \overline{e}$ obtained from the unit vector $\overline{e}(1;0)$ by rotating by angle α, and the additional rotation by π is *symmetry about the origin* and changes the signs of both coordinates, so the above formulae for the second derivatives can be rewritten as

$$(a\cos\omega t)'' = -\omega^2 \cdot a\cos\omega t, \qquad (b\sin\omega t)'' = -\omega^2 \cdot a\sin\omega t.$$

Hence, both functions $x_1(t) = a\cos\omega t$ and $x_2(t) = b\sin\omega t$, *and thus also their sum* $x(t) = x_1 + x_2 = a\cos\omega t + b\sin\omega t$,[1] are solutions of the differential equation (4). It can be proved, and we shall do it in § 5.3, that these are all solutions of equation (4).

Question. It is easy to check that an arbitrary function of the form $\tilde{x}(t) = A\cos(\omega t + \varphi)$ satisfies the differential equation (4). Does not this contradict the above phrase "It can be proved..."? □

Quantities varying according to the above laws (formulae) are said to be *subject to harmonic oscillations* or simply called *harmonic oscillations.* Harmonic oscillations arise whenever a one-dimensional system is described by a differential equation of the form $x'' = -\omega^2 x$, or $x'' = -\lambda x$, $\lambda > 0$. This situation is quite typical when considering motions of a system *near its*

[1] Note this fact: *the sum of solutions of the differential equation* (4) $x'' = -\omega^2 x$ *is also a solution of this equation.* This is by far not always the case.

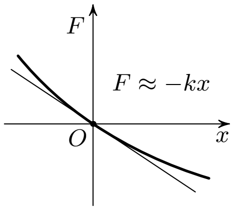


Fig. 10.

equilibrium. For instance, if the ball in the trough from the preceding subsection moves in the vicinity of a *stable equilibrium point*, i.e., in the "pit" corresponding to the minimum of the trough profile $h=u(x)$, then when x deviates from the minimum point, which can be taken as the origin of the Ox coordinate axis, the force $F=F(x)$ is directed towards the equilibrium point, i.e., $F(x)<0$ when $x>0$, $F(x)>0$ when $x<0$, and $F(0)=0$ (Fig. 10).

Assume that a function $F=F(x)$ possessing the above property is *differentiable* at $x=0$, i.e., there exists a *negative* derivative $F'(0)=-\lambda$, $\lambda>0$. Then in some neighbourhood of zero *the function can be replaced by its linear approximation*:

$$F(x)\approx F(0)+F'(0)(x-0)=-\lambda x$$

(see Sec. 1.1.3). Hence, Newton's equation can also be replaced by its approximation:

$$mx''=F(x)\approx -\lambda x.$$

There we have Hooke's law!

Robert Hooke (1635–1703) was an English encyclopaedic scientist, secretary and curator of experiments of the Royal Society of London, the discoverer of the law of universal gravitation. Regarding the latter, it should be noted that at about the same time this law was discovered and discussed not only by Hooke and Isaac Newton, but also by Christopher Wren (1632–1723), English mathematician and astronomer, a founding member of the Royal Society and its president in 1681–1683 (from 1703 until his death, the permanent president of the Royal Society was Isaac Newton; from 1705 onwards, *Sir* Isaac Newton), chief architect of London for 30 years (it was Wren who supervised the construction of the famous St. Paul's Cathedral in London in 1675–1710). Wren cooperated with Hooke in resolving questions about Kepler's laws and the law of gravitation, as well as in the development of London — Robert Hooke was also an architect!

Hooke is also famous for the improvement of the microscope, the discovery of the cellular structure of organisms (it was he who introduced the term *cell*), the invention of the barometer, and others. The meaning of his discovery of the *law of elasticity for solids* is, as has already been mentioned, that this law ($F=-kx$) is valid *in a rather wide range of displacements* x.

Exercises, Problems, and Tasks to Chapter I

Tangent lines and linear approximations

1. Find linear approximations:

(1) For the function x^2 at $x_0=1$;

(2) For the function $\sqrt{x}$ at $x_0=1$;

(3) For the function $3x-2$ at $x_0=1$.

2. Can it be that

(1) The graph of a function coincides with a tangent to it? For which functions?

(2) The tangent to a graph intersects it in more than one point?

Give examples.

3. Draw graphs of continuous functions such that

(1) $D(f) = [-4; 4]$, $f(-3) = f'(0) = 0$, $f'(-3) = f(0) = 1$, $f(3) = -3$, $f'(3) = -2$;

(2) $D(g) = [-1; 5]$, $g(0) = g'(1) = 0$, $g'(2) = g(2) = 1$, $g(3) = -1$, $g'(4) = -\frac{1}{2}$.

4. Write equations of tangent lines to the given functions at the given points and draw sketches of the graphs and the tangents:

(1) $y = x^3$, $x_0 = 1$;

(2) $y = \frac{3}{x}$, $x_0 = -1$;

(3) $y = \sqrt{x}$, $x_0 = 4$.

By definition, the angle between a function graph and a straight line at the point of their intersection is the angle between the tangent to the graph at that point and the given line.

Similarly, the angle between two intersecting graphs is the angle between the tangents to these graphs at the intersection point.

5. Find the angles at which the graphs of the following functions intersect the x-axis:

1) $y = x^3 - 3x$;

2) $y = x^3 - 3x + 2$.

6. Find the angles at which the graphs of the following functions intersect the y-axis:

(1) $y = \frac{1}{2}(x-1)^2$;

(2) $y = \dfrac{1}{x-1}$.

7. At which points the tangents to the graph of the function $y = x^3 - x$ are parallel to the straight lines:

(1) $y = x$;

(2) $y = 2x$;

(3) $y = -x$;

(4) $y = -2x$?

8. For which values of the parameter a a does the graph of the function $y = \frac{1}{4}(ax - x^3)$ intersect the x-axis at an angle of $45°$ (in at least one point)?

9. Find the distance between the nearest points on the graphs of the functions:

(1) $y = -2x + 1$ and $y = x^2 - 8x + 16$;

(2) $y = 2x - 1$ and $y = x^4 + 3x^2 + 2x$.

10. Write approximate formulae for evaluating the following quantities:

(1) $(x_0 + \Delta x)^n$, $(1 + h)^n$ $(n \in \mathbb{N})$;

(2) $\sqrt[n]{x_0 + \Delta x}$, $\sqrt[n]{1 + h}$, $\sqrt[n]{a^n + h}$ $(n \in \mathbb{N})$;

(3) $\dfrac{1}{x_0+\Delta x}$, $\dfrac{1}{1+h}$, $\dfrac{1}{1-h}$, $\dfrac{1+h}{1-h}$;

(4) $\dfrac{1}{(x_0+\Delta x)^n}$, $\dfrac{1}{(1+h)^n}$ $(n\in\mathbb{N})$.

11. Find approximated to two decimal places:

(1) $\sqrt[3]{9}$, (3) $\sqrt[7]{100}$, (5) $\sqrt[3]{3}$, (7) $\sqrt[4]{17}$,

(2) $\sqrt[4]{80}$, (4) $\sqrt[5]{33}$, (6) $\sqrt[3]{24}$, (8) $\sqrt[4]{9}$.

12. The side of the square is 4 ± 0.1 m. What is the maximum relative error in calculating the area of the square?

13. What relative error is allowed in measuring the radius R of a ball so that its volume can be determined with accuracy to 1%?

14. Assuming f and f' to be known, write linear approximations and derive from them formulae for the derivatives of the following functions:

(1) $f^2(x)$, $f(2x)$, $f(x^2)$; (3) $f(x^2+2x)$, $f(1/x)$, $1/f(x)$;

(2) $f^3(x)$, $f(x^3)$, $f(3x-1)$; (4) $\sqrt{f(x)}$, $f(\sqrt{x})$.

15. Prove that a non-vertical line $y=kx+m$ and a parabola $y=ax^2+bx+c$ are tangent if and only if the line and the parabola have a single common point.

16. Prove that a non-vertical and non-horizontal line $y=kx+m$, $k\neq 0$, and a hyperbola $y=\frac{a}{x}$ are tangent if and only if the line and the hyperbola have a single common point.

17. Write equations of the tangents to the graph of the function $y=x^2-4x+1$ passing through the points (1) $(0;0)$; (2) $(-1;-3)$.

18. Find the intersection point of the tangent to the parabola $y=ax^2$ at the point $x=z$ and the x-axis. How can we obtain from this a method of exact geometric construction of the tangent to the parabola?

19. $M_1(x_1;ax_1^2)$ and $M_1(x_2;ax_2^2)$ are two points on the parabola $y=ax^2$. What is the point at which the tangent is parallel to the secant M_1M_2?

20. Prove that the midpoints of all possible chords parallel to a given chord M_1M_2 of the parabola $y=ax^2$ lie on a single straight line.

21. Prove that the x-value of the intersection point of the tangents to the parabola $y=ax^2$ passing through two given points $M_1(x_1;ax_1^2)$ and $M_1(x_2;ax_2^2)$ divides the segment $[x_1,x_2]$ in half.

22. Prove that the tangent to the hyperbola $y=\dfrac{a^2}{x}$ forms with the coordinate axes a triangle of constant area (which one?) and that the tangency point is the centre of the hypotenuse of this triangle. How can we obtain from this a method of exact geometric construction of a tangent to the hyperbola?

23. (1) Find the angle between the lines given by their equations $y=k_1x+b_1$ and $y=k_2x+b_2$.

(2) Prove that a necessary and sufficient condition for these lines to be perpendicular is the equality $k_1k_2=-1$.

24. Prove that the tangent to the semicircle — the graph of the function $y=\sqrt{R^2-x^2}$ — is perpendicular (as is the rule in geometry!) to the radius drawn to the tangency point.

25. Write the equation of a *normal* (perpendicular to the tangent line) to the graph of a given differentiable function f drawn through the point on the graph with x-value $x=z$.

26. Find the shortest distance from the points $(0;0)$ and $(-1;-3)$ to the graph of the function $y=x^2-4x+1$.

27. What are the values of the parameter a for which the tangents to the graph of the function $y=x^3-a^2x$ at the points with x-values 0 and a are perpendicular to each other?

28. Find the locus of points from which the parabola $y=ax^2$ is seen at right angles (i.e., the tangents to the graph drawn from any of these points form a right angle).

29. Write the equation of a common tangent to the graphs of the functions

$$y=x^2+4x+8 \qquad \text{and} \qquad y=x^2+8x+4.$$

30. At what point on the graph of the function $y=ax^2+bx+c$ does the tangent to it pass through the origin?

* * *

31. Find the derivatives f', f'', f''', $\ldots$, $f^{(k)}$ (the k-th derivative) for the function $f(x)=(a+x)^n$ *for all values of* $k\in\mathbb{N}$.

32. Find the derivatives $f'(0)$, $f''(0)$, $f'''(0)$, $\ldots$, $f^{(k)}(0)$ for the polynomial function

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$$

for all values of $k\in\mathbb{N}$.

33. Let

$$(a+x)^n=a_0+a_1x+a_2x^2+\ldots+a_nx^n.$$

Find the coefficients a_k, $k=0,1,2,\ldots,n$.

The coefficients from the last problem are usually written in the form $a_k=\binom{n}{k}\cdot a^{n-k}$; the numbers $\binom{n}{k}$ are called *binomial coefficients*.

34. (1) Using factorials,[1] write a concise formula for the binomial coefficients (the formula itself for $(a+x)^n$ expressed through the binomial coeffi-

[1] Just in case: $n!$ (pronounced "n factorial") denotes the product of all natural numbers from 1 to n; besides, $0!=1!=1$.

cients is called *Newton's binomial expansion* formula, or *Newton's binomial theorem*).

(2) Find a relation between the numbers $\binom{n+1}{k}$ and $\binom{n}{k}$.

Hint: use the identity $(a+x)^{n+1}=(a+x)^n\cdot(a+x)$.

Isaac Newton (1643–1727) is an English physicist, mathematician, mechanic, and astronomer, one of the creators of differential and integral calculus (Leibniz published his discovery of the "infinitesimal calculus" 28 years earlier, but Newton established the existence of two mutually related calculi 10 years earlier; his corresponding work *Method of Fluxions* was published only in 1736). Newton's contribution to "Newton's binomial formula" consists in its generalisation to the case of arbitrary (fractional or real) powers of a binomial.

Vector-valued and trigonometric functions and their derivatives

35. Using rotations with variable centre, write a vector-valued function corresponding to the rotation of a point with angular velocity ω about a centre moving rectilinearly and uniformly. Classify hodographs of such functions.

36. Using rotations, write the vector-valued functions corresponding to the motion of a point on a circle of radius ρ rolling without slipping:

1) along a straight line;

2) along another circle of radius ρ_0 while being outside or inside it.

Which form are hodographs of these functions?

These hodographs are called *cycloids*, *epicycloids*, and *hypocycloids*, respectively. The name *cycloid* was given in 1598 by Galileo Galilei (1564–1642), the famous Italian physicist, mathematician, one of the founders of exact natural science, poet, philologist, and literary critic. The word is derived from Greek words κυκλος [kuklos], 'circle', and ειδος [eidos], meaning 'origination' in compound words, so the literal meaning of the term is 'generated by a circle'. Since the XVII century the cycloid has been studied by many mathematicians, who discovered a number of its remarkable properties.

Construction of some epicycloids and hypocycloids were first found in the book *Underweysung der Messung, mit dem Zirckel und Richtscheyt* (Instructions for Measuring with Compass and Ruler; 1525) by Albrecht Dürer (1471–1528), a German artist, the author of the famous engraving *Melencolia I*. The terms contain the Greek prepositions επι [epi], 'upon', 'over', used in names of figures constructed with a given curve as a base, and υπο [hupo], 'under'. These lines attracted the attention of many major mathematicians, including Newton, Johann Bernoulli, and Euler.

Consider separately epicycloids with $\rho=\frac{1}{n}\rho_0$ $(n\in\mathbb{N})$,[1] and analogous hypocycloids (for $n\geq 2$).[2]

[1] When $n=1$, i.e., $\rho=\rho_0$, we obtain the so-called *cardiod*, a 'heart-shaped' closed curve.

[2] The hypocycloid with $\rho=\frac{1}{4}\rho_0$ is called the *astroid*, 'star-shaped'.

37. In which case are the trajectories of points in the preceding problem (36.2) closed (accordingly, the motion along them being periodic)? Find a necessary and sufficient condition for the radii ρ and ρ_0.

38. Let, as above, a circle of radius ρ be rolling without sliding

1) along a straight line;

2) along another circle of radius ρ_0, while being outside or inside it. The first circle has a point rigidly attached to it at a distance h from its centre.

Using rotations, write the vector-valued function describing the motion of this point. What do hodographs of these functions look like?[1] Consider separately the case $\rho_0 = \rho$,[2] and the case $\rho_0 = 2\rho$.

39. In which case are the trajectories of points in the preceding problem (38.2) closed (the motion along them being periodic)?

* * *

40. Construct hodographs of a vector-valued function $\overline{r}(t)$ and its derivative $\overline{v}(t) = \overline{r}'(t)$ if the function $\overline{r}(t)$ is given by its coordinates $(x(t); y(t))$, $t \in \mathbb{R}$:

(1) $(t; t)$;	(5) $(-t; t^2)$;	(9) $(t; t^3)$;	(13) $(t^3; t^2)$;
(2) $(t^2; t^2)$;	(6) $(-2t; 4t^2)$;	(10) $(-t; -t^3)$;	(14) $(t^2; -t^4)$;
(3) $(t^3; t^3)$;	(7) $(t^2; t)$;	(11) $(t^3; t)$;	(15) $(-t^2; t^4)$;
(4) $(t; t^2)$;	(8) $(t^2; -t)$;	(12) $(t^2; t^3)$;	(16) $(t^4; t^2)$.

(On the hodographs you should mark the initial point corresponding to $t = 0$ and the direction of motion of the point P_t, and for the velocity, the direction of motion of a point V_t such that $\overline{OV_t} = \overline{v}(t)$.)

41. Construct hodographs of the vector-valued functions $\overline{r}(t)$ given by their coordinates:

(1) $\left(\frac{1}{t}; t^2 - 2\right)$, $t \in (0; +\infty)$;	(5) $\left(\sqrt{1+t}; \sqrt{t}\right)$, $t \in [0; +\infty)$;
(2) $\left(t^2 - \frac{1}{2}; t^4 - \frac{1}{4}\right)$, $t \in \mathbb{R}$;	(6) $\left(t^2; \sqrt{1+t^2}\right)$, $t \in \mathbb{R}$;
(3) $\left(\frac{1-t^2}{1+t^2}; \frac{2t}{1+t^2}\right)$, $t \in \mathbb{R}$;	(7) $\left(\sqrt{1-t^2}; \sqrt{1+t^2}\right)$, $t \in [-1, 1]$;
(4) $\left(\sqrt{t}; \sqrt{1-t}\right)$, $t \in [0; 1]$;	(8) $\left(\sqrt[3]{t^2}; (t+1)^3\right)$, $t \in \mathbb{R}$.

(On the hodographs you should mark the initial point corresponding to $t = 0$ and the direction of motion of P_t.)

* * *

42. Choose a convenient coordinate system and write the coordinate functions for the vector functions from Exercises 36–38. Find the derivatives of these functions (in coordinates). What are the hodographs of the derivatives?

[1] These hodographs are called *trochoids*, *epitrochoids*, and *hypotrochoids*, respectively.

[2] These curves (for any value of h) are called *Pascal's limaçons*, or *Pascal's snails*.

43. For each point Q on the upper semicircle of the unit circle (centred at the origin), we draw a straight line connecting this point to the origin and intersecting the line $y=2$ at a point R. Let P be the centre of the segment QR. Define by coordinate functions the trajectory of P (the point Q moves *somehow* along the upper semicircle).

44. Let $M=M(-1;0)$; through each point Q of the upper semicircle of the unit circle (centred at the origin) we draw a line MQ and choose a point P on it such that $PQ=1$. Define by coordinate functions the trajectory of P when Q traverses the upper semicircle.

45. Suppose that a thread of length $2\pi r$ is unwound from a disc (circle) of radius r (imagine, for example, a spool) with a stationary centre. Describe in coordinates the motion of the free end of the thread.[1]

* * *

46. Let $\lambda(t)$ and $\overline{r}(t)$ be a scalar (i.e., numerical) and a vector-valued functions differentiable at a point t_0. Prove that then the vector function $\overline{R}(t)=\lambda(t)\cdot\overline{r}(t)$ is also differentiable at t_0; write and prove a formula for the derivative of this product $\overline{R}(t)$.

47. Let $\overline{r}_1(t)$ and $\overline{r}_2(t)$ be vector-valued functions differentiable at a point t_0. Prove that the scalar function $s(t)=\overline{r}_1(t)\cdot\overline{r}_2(t)$ (the dot product of the vector functions $\overline{r}_1$ and $\overline{r}_2$) is then also differentiable at t_0; write and prove a formula for the derivative of $s(t)$.

48. Let the hodograph of a vector-function $\overline{r}(t)=(x(t);y(t))$ differentiable at a point t coincide in the neighbourhood of this point with the graph of a function $y=f(x)$ differentiable at the point $x=x(t)$. Prove that if the derivative $x'(t)$ is different from 0, the formula

$$f'(x)=\frac{y'(t)}{x'(t)}, \quad \text{or} \quad f'_x(x)=\frac{y'_t(t)}{x'_t(t)}$$

takes place (the second variant of the formula is given to emphasize the fact that the derivatives on the left- and right-hand sides of the equality are with respect to quite different variables). Give a geometric interpretation of this formula.

[1]The trajectory described by the free end of the thread is called the *involute* (also known as *evolvent*) of the circle, and the circle itself is the *evolute* for this curve. The terms are derived from the Latin *evolvo*, ‘roll out’, ‘unfold’. Tangents to the evolute are normals (perpendiculars) to the involute (prove!). These curves were first considered and studied in 1654 by Christiaan Huygens (1629–1695), a Dutch physicist and mathematician, inventor of the pendulum clock; he was the first president of Académie des sciences de Paris (the Paris Academy of Sciences).

* * *

49. Since a rotation by $2\pi k$ radians is an identity transformation ($R^{2\pi k} = \mathrm{id}$, i.e., all points are invariant) for any $k \in \mathbb{Z}$, the values of the functions $\cos t$ and $\sin t$ periodically repeat every 2π; what property of the graphs of these functions corresponds to their periodicity? Draw the graphs of the functions $y = \cos x$ and $y = \sin x$ without finding derivatives but rather based on general considerations and the periodicity. Prove that these graphs are congruent as geometric figures (i.e., one can be transformed into another by an isometry; what kind of isometry?). Based on the form of the graphs, sketch below them graphs of their derivatives $y = \cos' x$ and $y = \sin' x$.

(*Hint.* When drawing the graphs of the functions $\cos x$ and $\sin x$, follow the standard scheme: determine zeros of the functions, maximum and minimum points — for example, for $\sin x$ these are, respectively, the points

$$\sin x = 0 \Leftrightarrow x = \pi n, \quad n \in \mathbb{Z};$$
$$\sin x = 1 \Leftrightarrow x = \tfrac{\pi}{2} + 2\pi k, \quad k \in \mathbb{Z};$$
$$\sin x = -1 \Leftrightarrow x = \tfrac{3\pi}{2} + 2\pi k, \quad k \in \mathbb{Z}.$$

Then from the definition and elementary geometric considerations find monotonicity intervals).

50. Given the coordinates of a point $M(x; y)$, find the coordinates of the point $R^{\pi/2}M$. Using the obtained result, express the derivatives $\cos' t$ and $\sin' t$ through the functions $\cos t$ and $\sin t$ themselves (functions of the argument t!).

(*Hint.* Rotation by the angle $t + \frac{\pi}{2}$ can be represented as a composition of rotations by t and $\frac{\pi}{2}$:

$$R^{t+\pi/2} = R^{\pi/2} \circ R^{t}.$$

Now apply this equality to the point $(1; 0)$ and note the definitions of the functions $\cos t$ and $\sin t$ together with the "kinematic" formulae for their derivatives).

* * *

Definition 1. The trigonometric functions *tangent*, *cotangent*, *secant*, and *cosecant* of a numerical argument $x \in \mathbb{R}$ are the functions defined respectively by the formulae

$$\tan x = \frac{\sin x}{\cos x}, \qquad \cot x = \frac{\cos x}{\sin x}, \qquad \sec x = \frac{1}{\cos x}, \qquad \csc x = \frac{1}{\sin x}.$$

The trigonometric function *sine* is found in the Indian *Siddhantas*, anonymous works on astronomy of the IV or V century, as well as in *Aryabhatiyam* (499 AD), a work by the Indian astronomer and mathematician Aryabhata (c. 476–550). The origin of the name of this function is curious. The sine line (from the point P_α to its projection on the

x-axis) was called *ardhajiva* — *ardha* means 'half' and *jiva* means 'bow string', chord. Afterwards, this name became shortened to *jiva.* In Arabic literature the term was changed into *jiba*, and in the IX century this word, meaningless in everyday life, was replaced by the Arabic word *jayb*, which means 'bosom', 'pocket', or 'fold'. In the XII century this word was literally translated into Latin by the term *sinus*! The notation sin (as well as cos and tan) became generally accepted thanks to the authority of the famous Swiss mathematician Leonhard Euler (1707–1783; for a long time he worked in St. Petersburg, Russia, and was buried there).

The term *cosine* (as well as *cotangent*) was introduced in 1620 by Edmund Gunther (1581–1626), an English mathematician and astronomer, inventor of the logarithmic slide rule, as an abbreviation of the Latin *complementi sinus*, 'sine of the complementary angle' ($\cos\alpha = \sin(90^\circ - \alpha)$).

The trigonometric function *tangent*, as a shadow of a vertical pole, was introduced in the X century by the Baghdad mathematician and astronomer Muhammad ibn Muhammad Abu al-Wafa (940–988; born in Khorasan). He also introduced the *cotangent*, *secant*, and *cosecant* functions. Afterwards, tangents were repeatedly rediscovered: by the Englishman Thomas Bradwardine (c. 1300–1349; Archbishop of Canterbury), by the German mathematician Regiomontanus (1436–1476); but even 100 years after Regiomontanus, Copernicus still was not aware of this discovery. Originally, for the cotangent and tangent, the names *umbra recta* and *umbra versa* ('straight shadow' and 'turned shadow') were used. In 1583, in his book *Geometria rotundi* (Geometry of the Round), the Danish physician, astronomer, and mathematician Thomas Fincke (1561–1656) introduced the term *tangens*, which is Latin for 'touching', 'segment of a touching line'. (What exactly is this segment?) The term *secans* was also introduced there; it is Latin for 'cutting', 'segment of a cutting line', from *seco*, 'to cut'. (What cutting line are we talking about? Have a think!) After Abu al-Wafa, the secant was also rediscovered, in particular by the Swiss Georg Rheticus (1514–1574), the first to tabulate it, who was a pupil and friend of Copernicus.

51. Using the formulae for the derivatives of $\cos x$ and $\sin x$ (from the preceding problem), find the derivatives of the functions introduced above:

(1) $\tan x$, (2) $\cot x$, (3) $\sec x$, (4) $\csc x$.

Draw their graphs. Find out if there are any equal graphs among them.

52. Find the derivatives of the functions:

(1) $\cos 2x$, (5) $\tan 3$,
(2) $\sin 3x$, (6) $\tan \frac{x}{2}$,
(3) $\sin \frac{x}{3}$, (7) $\sec \frac{x}{2}$,
(4) $\sin\left(2x + \frac{3\pi}{4}\right)$, (8) $\cot\left(3x - \frac{\pi}{4}\right)$.

53. Draw graphs of the functions:

(1) $y = \sin 2x$, (3) $y = \sin \frac{x}{3}$,
(2) $y = \cos 3x$, (4) $y = \sin\left(2x + \frac{3\pi}{4}\right)$.

Hint. Use the same reasoning as when graphing $\cos x$ and $\sin x$ (zeros of functions, maximum points, minimum points, etc.).

Questions to the solution. Which geometric transformations can be used to obtain the graph of a function $y = f(ax)$ from the graph of $y = f(x)$?

(Consider separately the cases $a > 1$, $0 < a < 1$, and $a < 0$.) The same question for the graph of the function $y = f(x + b)$.

General case: by which transformations is the graph of a function $y = f(ax + b)$ obtained from the graph of $y = f(x)$? (Note that horizontal translations and stretching/shrinking operations do not commute in this case!)

* * *

54. Construct hodographs of the vector-valued functions $\overline{r}(t)$ defined by their coordinates:

(1) $(\cos t; \cos t)$;	(5) $(\sin t; \cos t)$;	(9) $(2\cos t; \sin t)$;
(2) $(\cos t; \sin t)$;	(6) $(\cos t; -\sin t)$;	(10) $(a\cos t; b\sin t)$;
(3) $(\cos t; \cos 2t)$;	(7) $(\sin t; \cos 2t)$;	(11) $(\cos t; \cos 3t)$;
(4) $(\cos t; \sin 2t)$;	(8) $(\cos t; 2\sin t)$;	(12) $(\cos t; \sin 3t)$.

Chapter II

Integral Calculus and Differential Equations

§ 2.1. Analysis of the Differential Equation $y' = f(x)$

Among continuous mathematical models written as differential equations of the form

$$x' = F(x,t), \tag{1}$$

especially distinguished are equations in which the right-hand side is a given (known) function of a continuous variable (time) t, $F(x,t) \equiv f(t)$, and *autonomous equations*, in which the derivative is given as a known function of a continuously varying quantity $x(t)$, $x' = F(x)$, i.e., $x'(t) \equiv F(x(t))$.

We shall deal with equations of the latter type (autonomous equations) in Chs. III and IV, and in this chapter we focus on equations of the type $x'(t) = f(t)$, or, using the notation y for x and x for t, on the analysis of the differential equation

$$y' = f(x): \quad \forall x \in D \quad y' = f(x). \tag{2}$$

Here f is a *given function* and D is the domain of the differential equation: a number set, most often a finite or infinite interval or union of intervals (if this set is not specified, it is assumed that $D = D_f$ is the domain of f).

2.1.1. Integration as solving a differential equation. The problem of solving (finding all solutions) of the differential equation (2) is exactly the problem of *recovering a function given its derivative*, i.e., the *main problem of integral calculus* (the *inverse* of the problem of finding the derivative). Let us reconcile the terminology related to differential equations with the terminology used in integral calculus: give definitions, examples, and formulations of statements known from the course of elementary calculus in terms of differential equations.

Definition 1. A function $F(x)$ is called an *antiderivative* (or a *primitive*) of a function $f(x)$ on a set D, an interval or a union of intervals, if it is a solution of the differential equation (2), $y' = f(x)$, on this set, i.e.,

$$\forall x \in D \quad \exists F'(x) \quad \text{and} \quad F'(x) = f(x).$$

Example 1. The differential equation $y' = x$, $x \in \mathbb{R}$, is satisfied by the function $y = \frac{x^2}{2}$, and along with it by any function of the form $y = \frac{x^2}{2} + C$, since *if we add a constant to a function, its derivative does not change.* □

Example 2. The differential equation $y' = x^2$, $x \in \mathbb{R}$, has solutions $y = \frac{x^3}{3} + C$, where C is an *arbitrary additive constant.*[1] □

Example 3. The differential equation $y' = \frac{1}{x^2}$, $x \neq 0$, is satisfied by any functions of the form $y = -\frac{1}{x} + C$. □

The question arises: have we identified (guessed, "found") *all* the solutions of the differential equation (2) for the considered functions f? The answer is known.

Proposition 1 (fundamental property of solutions of the differential equation $y' = f(x)$)**.** *If $y_1(x)$ and $y_2(x)$ are two solutions of a differential equation $y' = f(x)$, $x \in I$, where I is an arbitrary (possibly, infinite) interval, then $\forall x \in I$ we have $y_2(x) = y_1(x) + C$, where C is a constant.*

To prove this, we need an auxiliary fact.

Lemma 1 (constancy criterion)**.** *If a function $h(x)$ is differentiable on an interval $I \subset \mathbb{R}$ and its derivative is zero on the entire interval, then h is constant on I:*

$$h'(x) \equiv 0 \text{ on } I \quad \Rightarrow \quad h(x) \equiv \text{const on } I.$$

This immediately follows from *Lagrange's mean value theorem*, known from the course of elementary calculus:

$$\forall a, b \in I \quad \exists c \in I \quad f(b) - f(a) = f'(c)(b - a) = 0.$$

Proof of Proposition 1. Let us differentiate the difference of the solutions, i.e., the function $r(x) = y_2(x) - y_1(x)$. We obtain

$$\forall x \in I \quad r'(x) = y_2'(x) - y_1'(x) = f(x) - f(x) = 0.$$

Hence, by the *constancy criterion* (Lemma 1) this difference $r(x)$ is constant, i.e., equals some constant C *on the interval I*. Thus, $y_2(x) - y_1(x) = C$ *on this interval*, so $y_2(x) = y_1(x) + C$. □

We have specially emphasised that we are talking about solutions *on a single interval*, since the *constancy criterion* is valid only for functions considered *on a single interval*, because its proof is based on Lagrange's mean value theorem: if $r'(x) \equiv 0$ on I, then $\forall a, b \in I$ there exists a point c lying between a and b such that

$$r(b) - r(a) = r'(c)(b - a) = 0 \quad \Rightarrow \quad r(b) = r(b).$$

[1] Recall: this means that the constant enters the solution as a summand; from Latin *additivus*, 'added', 'annexed'.

On the union of (non-overlapping) intervals it makes no sense to talk about Lagrange's theorem: a function with zero derivative is a constant only on a separate interval, but it would not be such in the whole, because each of the intervals has its own constant!

Thus, in the first two examples we have indeed specified all the solutions of the corresponding differential equations, but in Example 3 we have not, because this equation is considered on the domain of the function $f(x)=\frac{1}{x^2}$, i.e., on the union of intervals $(-\infty,0)\cup(0,+\infty)$. Therefore, the general form of solutions of this differential equation is given by the "compound" formula

$$y(x)=\begin{cases} -\frac{1}{x}+C_1, & x<0; \\ -\frac{1}{x}+C_2, & x>0. \end{cases}$$

In what follows we shall write such compound formulae shorter, in the form

$$y(x)=-\tfrac{1}{x}+\widehat{C},$$

where the "hat" above an arbitrary constant C means that on each of the intervals on which we consider the differential equation $y'=f(x)$ (as was said above, unless otherwise specified, this is the *entire natural domain of the function* $f\colon D=D(f)$), we may take a different constant.

Now we give a statement following from the proved proposition, which also applies to the differential equation (2) considered on a singular interval.

Corollary 1 (general form of solutions of the differential equation $y'=f(x)$)**.** *Any solution of the differential equation $y'=f(x)$, $x\in I$, where I is a given (possibly, infinite) interval, can be written as $y(x)=y_1(x)+C$, where $y_1(x)$ is an arbitrary* particular *solution of the equation and C is an arbitrary constant.*

In accordance with Definition 1, both of the above statements can be translated into the language of antiderivatives. Proposition 1 states that *any two antiderivatives of the same function considered on a single interval differ on it by a constant.* Corollary 1 is usually referred to as the *fundamental property of antiderivatives*: *if a function f has some antiderivative F on an interval I, then it has* infinitely many *antiderivatives on this interval, and* each *antiderivative of f on this interval can be represented as $F(x)+C$,* $C=\mathrm{const}\in\mathbb{R}$.

Newton did not use the notion of the *antiderivative* (*primitive*); he understood the problem of finding it as *solving a differential equation*: in his terminology, *to solve a differential equation* meant to find a *"fluent"* (a "flowing" function; from Latin *fluentum*, 'flow [of water]', 'current', 'flood') given its *"fluxion"* (i.e., its derivative, "velocity"; from Latin *fluxio* or *fluctio*, 'flowing'). Moreover, Newton did not add an *arbitrary constant* (however, he mentioned that solutions "include" arbitrary constants). The first to write it was Leibniz (in 1694).

The term *primitive* was introduced by Lagrange in 1797 in his *Théorie des fonctions analytiques* (Theory of Analytic Functions), together with the term *derivative.* The terms originate from Latin *primitivus*, 'first of its kind', 'original', and *derivativus*, 'led or drawn off [from its source]'. The notation $y' = f'(x)$ (or $x'(t)$), $y'' = f''$, etc. was also introduced by Lagrange (1770). Newton instead of the strokes (primes) used dots over the fluents: $\dot{x}$, $\ddot{x}$; in mechanics, Newton's notation is still used today for time derivatives (although in England it has been abolished since 1915 for "typographical reasons"!). Leibniz's notation and terminology are discussed in Ch. IV.

2.1.2. Uniqueness theorem for solutions of the equation $y' = f(x)$ and properties of antiderivatives. We have seen that if a function $f(x)$ on an interval I has *at least one* antiderivative $F(x)$, then $f(x)$ has *infinitely many* antiderivatives, and all of them differ from $F(x)$ by an additive constant C. If we additionally impose an *initial condition* on the antiderivative, i.e., on the solution $y(x)$ of a differential equation $y' = f(x)$, $x \in I$, which means that we specify the value of $y(x)$ at some point $x_0 \in I$, then the antiderivative (the solution of the differential equation) is determined *uniquely* — the constant C is found from the initial condition:

$$y(x_0) = y_0 \quad \Rightarrow \quad y_0 = F(x_0) + C \quad \Rightarrow \quad C = y_0 - F(x_0).$$

In terms of differential equations, this is formulated as the *uniqueness theorem.*

Theorem 1 (uniqueness of solutions of a differential equation $y' = f(x)$). *If a solution of a differential equation $y' = f(x)$, $x \in I$, where I is a given* (*possibly, infinite*) *interval,* exists, *it is* uniquely *determined by an* initial condition $y(x_0) = y_0$.

In other words: *if there exists a function $y(x)$, $x \in I$, satisfying the "system of conditions"*

$$\begin{cases} y'(x) = f(x) & (\forall x \in I); \\ y(x_0) = y_0 & (x_0 \in I,\ y_0 \in \mathbb{R}), \end{cases}$$

it is unique.

Proof. Actually, the proof has already been given above. □

Note the formula relating the *desired* solution $y(x)$ (satisfying the given initial condition) to an arbitrary antiderivative $F(x)$ (assumed to exist):

$$y(x) = F(x) + C = F(x) + (y_0 - F(x_0)) = y_0 + (F(x) - F(x_0))$$
$$\Leftrightarrow \quad y(x) - y(x_0) = F(x) - F(x_0).$$

Let us formulate the corresponding statement (on the *equality of increments* of any two antiderivatives on the segment from x_0 to x) as a separate theorem (having slightly changed the notation).

Theorem 2 (on the equality of increments of any antiderivatives of a given function)**.** *If a function $f(x)$ has two antiderivatives $F(x)$ and $\overline{F}(x)$ on an interval I, then for any two points $a, b \in I$, increments of these antiderivatives on the segment from a to b coincide*:

$$F(b) - F(a) \overset{\text{des}}{=} F(x)\Big|_a^b = \overline{F}(x)\Big|_a^b \overset{\text{des}}{=} \overline{F}(b) - \overline{F}(a).$$[1]

Proof. By the fundamental property of antiderivatives, if a function $f(x)$ has an antiderivative $F(x)$ on an interval I, then *any* its antiderivative $\overline{F}(x)$ differs from $F(x)$ by an additive constant C: $\overline{F}(x) = F(x) + C$. But in this case we have

$$\overline{F}(x)\Big|_a^b = \overline{F}(b) - \overline{F}(a) = (F(b) + C) - (F(a) + C) = F(b) - F(a) = F(x)\Big|_a^b. \ \square$$

2.1.3. Existence of solutions of the equation $y' = f(x)$, or antiderivatives. All the statements above have been proved *assuming* that solutions or antiderivatives in question *exist.* However, by far not every function has an antiderivative. Let us recall a relevant example.

Example 4. Let us prove that *the* $\operatorname{sgn} x$ *function does not have an antiderivative on* $\mathbb{R}$. Recall that this function is defined by the following "compound formula":

$$\operatorname{sgn} x = \begin{cases} -1, & x < 0; \\ \ \ 1, & x > 0; \\ \ \ 0, & x = 0. \end{cases}$$

Assume that this function has some antiderivative $F(x)$ on the entire number axis $\mathbb{R}$. Then $\forall x \in \mathbb{R}\ \exists F'(x) = \operatorname{sgn} x$. In particular, on the interval $(-\infty, 0)$ we have $F'(x) = -1$, and since $(-x)' = -1$ on the same interval, the fundamental property of antiderivatives implies that on the interval $(-\infty, 0)$ we have $F(x) = x + C$ for some constant C. Similarly, on the interval $(0, \infty)$ we have $F(x) = x + C_1$ for some constant C_1.

Next, since the antiderivative $F(x)$ is differentiable everywhere, including the point $x = 0$, it is *continuous* at zero, and therefore

$$\lim_{x \to 0} F(x) = F(0).$$

[1] The symbol 'des' above the equality sign stands for 'designation'. In this case, it is a convenient notation for the *increment of a function F on a segment from a to b*: $F(x)|_a^b$. It was introduced in 1848 by the French mathematician P.F. Sarrus.

Hence, $C_1 = F(0) = C$, and the function $F(x)$ can be represented by the following compound formula:

$$F(x) = \begin{cases} -x + C, & x < 0; \\ x + C, & x > 0; \\ C, & x = 0 \end{cases} \quad \Leftrightarrow \quad F(x) = |x| + C.$$

However, this function cannot be *differentiable everywhere*: otherwise, so would also be the function $F(x) - C = |x|$, but we know that the function $|x|$ has no derivative at $x = 0$.

The obtained contradiction proves that the function $\operatorname{sgn} x$ has no antiderivative on $\mathbb{R}$. □

In Example 4, the function with no antiderivative was *discontinuous.* The following results are known to hold:

1. *A function which is continuous on an interval always has an antiderivative on this interval* — this will be proved in § 2.3;

2. On the other hand, discontinuity of a function *by no means imply* that it does not have an antiderivative; it well *may have* one, as the following example shows.

Example 5. The function

$$f(x) = \begin{cases} 2x \cos \frac{1}{x} + \sin \frac{1}{x}, & x \neq 0; \\ 0, & x = 0, \end{cases}$$

is discontinuous at $x = 0$ (it is easily seen that $\lim_{x \to 0} f(x)$ does not exist), but at the same time this function is the derivative for

$$F(x) = \begin{cases} x^2 \cos \frac{1}{x}, & x \neq 0; \\ 0, & x = 0, \end{cases}$$

which, in turn, is an antiderivative for the (discontinuous) function f on the entire number axis. □

§ 2.2. Geometric Interpretation of the Equation $y' = F(x, y)$

Before addressing the question of existence of antiderivatives or solutions of a differential equation of the form $y' = f(x)$, let us consider one of the most important approaches to the analysis and approximate solution of differential equations, the so-called *Euler method* (1768), in the general case applied to an arbitrary evolutionary differential equation

$$x' = F(x, t), \tag{1}$$

which we shall write in another notation as

$$y' = F(x, y), \quad \text{i.e.,} \quad y'(x) = F(x, y(x)), \tag{2}$$

where $F = F(x, y)$ is a given function of two variables, i.e., a function on the Oxy coordinate plane.

2.2.1. Directional fields and integral curves of differential equations. Equation (2) has a simple geometric interpretation: since the derivative $y'(x)$ is the slope of the tangent to the graph of a solution $y = y(x)$ at the point $(x; y) = (x; y(x))$, and it is *known* to us thanks to equation (2) (it is the value of the given function $F(x, y)$), to solve equation (2) means to recover the function $y(x)$ (or its graph) given directions of its tangent lines at all points. All possible tangents to the graphs of solutions of the differential equation (2) form the so-called *directional field* in the Oxy coordinate plane (Fig. 11). Strictly speaking, a directional field is a mapping

$$\boldsymbol{L}\colon\ M(x; y) \mapsto \boldsymbol{L}_{F(x,y)}(M),$$

where by $\boldsymbol{L}_k(M)$ we denote the straight line passing through point M and having slope k. Graphs of solutions of equation (2) at each of their points are tangent to this directional field (see Fig. 11).

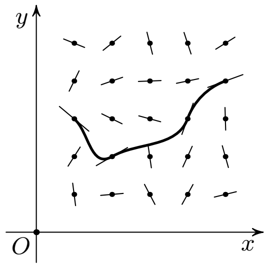


Fig. 11.

Curves that are tangent to a given directional field at each point are called *integral curves* of this directional field (or of the corresponding differential equation). Thus, the problem of solving the differential equation (2) is equivalent to the problem of finding all integral curves corresponding to this directional field equation.

Clearly, to *uniquely* determine a solution of the differential equation (2), it is required to specify an *initial condition*, i.e., a point $M_0(x_0; y_0)$ through which an integral curve passes. Usually, an initial condition is written in the form

$$y(x_0) = y_0. \tag{3}$$

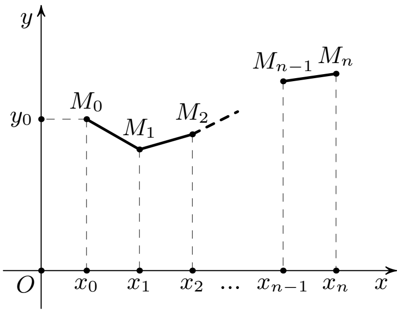


Fig. 12.

2.2.2. Euler's method for constructing integral curves. Euler's idea is to find integral curves *approximately*, replacing them by *polygonal lines* constructed in a certain way from a given directional field. Suppose we want to find a solution $y = y(x)$ of differential equation (2) with initial condition (3). Let us try to *approximately* draw through the point $M_0(x_0; y_0)$ an integral curve, i.e., a graph of a solution, up to the point with abscissa x. For this purpose, we divide the segment $[x_0, x]$ into n equal parts by points $x_1, x_2, \ldots, x_{n-1}, x_n = x$ (clearly, $x_m = x_0 + m\frac{x - x_0}{n}$) and draw *Euler's polygonal line* $M_0M_1M_2 \ldots M_n$, where $M_m = (x_m; y_m)$ and the points $M_1, M_2, \ldots$ are chosen so that each segment M_mM_{m+1} goes along the directional field $\boldsymbol{L}$ at M_m (Fig. 12). Then the slope of the line M_mM_{m+1} is $F(x_m, y_m)$, and its equation can be written as

$$y = y_m + F(x_m, y_m)(x - x_m),$$

so we obtain

$$y_{m+1} = y_m + F(x_m, y_m)(x_{m+1} - x_m). \tag{4}$$

By successively substituting the recurrence formulae (4) one into another for $m = 1, 2, \ldots, n$, we obtain a formula for the ordinate of the last point of the Euler polygonal line:

$$y_n = y_0 + \sum_{m=0}^{n-1} F(x_m, y_m)(x_{m+1} - x_m),$$

which in the limit as $n \to \infty$ *should hypothetically give* the value of a solution of equation (2) with initial condition (3) at the point x:

$$\lim_{n\to\infty} y_n = y(x).$$

It turns out that under certain conditions on the function $F = F(x, y)$, this is indeed the case, but the proof of the corresponding theorem (and even its exact formulation) is far beyond the scope of this course. In the next section

(§ 2.3) we will apply Euler's method to the antiderivative equation $y' = f(x)$, and in the next chapter (in § 3.1) we will see what Euler's approach gives in the case of a *homogeneous linear differential equation* $y' = ky$.

But first let us consider a (sometimes useful) geometric (graphical) method for finding solutions of differential equations of the form (2).

2.2.3. Isoclines of directional fields and graphical integration of differential equations. If for a given differential equation (2) we are able to represent its directional field in reasonable detail (in this case, instead of the line $\boldsymbol{L}_{F(x,y)}(M)$, a small segment is drawn through the point M), such a picture can sometimes give us an idea of the *qualitative behaviour* of integral curves, i.e., graphs of solutions of the equation. Long before Euler (in 1694), Johann Bernoulli, a Swiss mathematician and a close associate of Leibniz, proposed to use *isoclines* to construct directional fields.

Definition 1. *Isoclines* of a differential equation $y' = F(x, y)$ or of the corresponding directional field are *sets of points of the same slope* of the directional field:

$$\Gamma_k = \{(x; y) \mid F(x, y) = k\};$$

in other words, the *isoclines*[1] are *level curves* of the function F given by the equations $F(x, y) = k$, along which the directional field has a *constant slope*: the tangent of the angle of inclination is k (e.g., for $k = 0$, the isocline Γ_0 gives all points at which the tangents to the integral curves are *horizontal*, i.e., all critical points of solutions $y = y(x)$ of the differential equation $y' = F(x, y)$).

The *isocline method* for constructing directional fields is to draw a family of some isoclines Γ_k and along each of them construct a field of directions (straight line segments) with slopes k.

Example 1. For the differential equation $y' = \dfrac{ax + by}{cx + dy}$, $(c; d) \neq (0; 0)$, $ad - bc \neq 0$,[2] the isocline equation is written as

$$\frac{ax + by}{cx + dy} = k \quad \Leftrightarrow \quad ax + by = k(cx + dy) \quad \text{and} \quad cx + dy \neq 0 \quad \Leftrightarrow \quad Ax + By = 0$$

($cx + dy \neq 0$; $A = a - kb$, $B = b - kd$). The last equation defines a line passing through the origin, but with the origin *punctured*, i.e., actually a pair of resulting rays (which can as well be vertical, i.e., coincide with the positive and negative half-lines of the y-axis). It is easier to construct the directional field along isoclines in a somewhat reversed manner: for an oblique line $y = \lambda x$ find the slope by the formula

$$k = F(x, y) = F(x, \lambda x) = \frac{ax + b\lambda x}{cx + d\lambda x} = \frac{a + b\lambda}{c + d\lambda},$$

[1] From Greek ισος [isos], 'same', and κλινω [klino], 'make to slope'.

[2] Ponder what this condition means. What happens in the case $ad - bc = 0$?

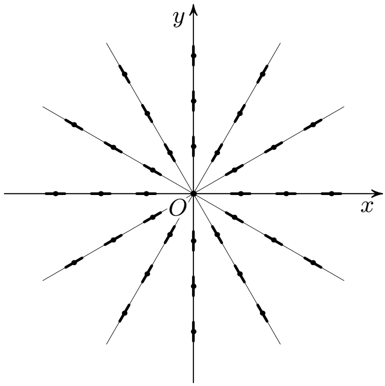


Fig. 13.

and then, along the two rays of this line, construct “directions” with this slope $k = k(\lambda)$.

At the origin the directional field is *undefined*—this point is called a *singular point* of the differential equation. But at the points of the line $cx + dy = 0$ different from the origin, it makes sense to consider the direction of the vector field to be *vertical*: as a point $(x; y)$ outside this line tends to a point on the line, the slope $k = F(x, y)$ tends to infinity. □

Now we will show by concrete examples what can be done after constructing a part of isoclines and the directional field along them.

Example 2. For the differential equation $(*)$ $y' = \frac{y}{x}$, the slope along the isocline $y = \lambda x$ is

$$k = k(\lambda) = \frac{y}{x} = \frac{\lambda x}{x} = \lambda.$$

Hence, the directional field of this differential equation is *codirectional* with the isoclines (Fig. 13) and is vertical along the Oy axis ($x = 0$ on it). Therefore, the rays passing through the origin are *themselves* integral curves.They are given by the same equation $y = \lambda x$, and the differential equation $(*)$ can be checked directly:

$$y' = (\lambda x)' = \lambda, \qquad \frac{y}{x} = \frac{\lambda x}{x} = \lambda = y',$$

as desired. □

Example 3. For the differential equation $(*)$ $y' = \frac{x}{y}$, the slope along the isocline $y = \lambda x$ is

$$k = k(\lambda) = \frac{x}{y} = \frac{x}{\lambda x} = \frac{1}{\lambda}.$$

In this case the directional field is codirectional with the isoclines only if

$$k(\lambda) = \lambda \quad \Leftrightarrow \quad \frac{1}{\lambda} = \lambda \quad \Leftrightarrow \quad \lambda^2 = 1, \quad \lambda = \pm 1.$$

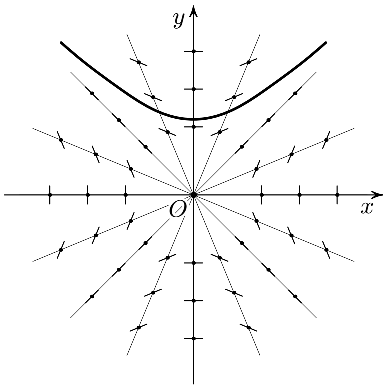


Fig. 14.

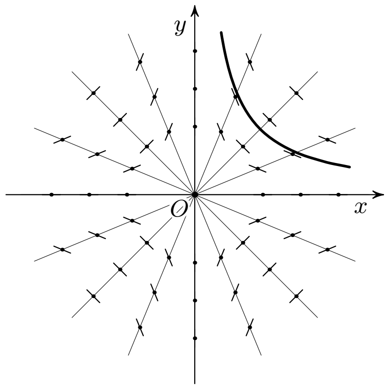


Fig. 15.

Hence, the half-lines of the two lines $y = \pm x$ (with the origin punctured) are integral curves of equation $(*)$ (Fig. 14), and their equations define solutions.

To get an idea of how other integral curves behave, note that along the Oy axis the directional field is horizontal, along the Ox axis it is vertical, and along other rays $y = \lambda x$ it is codirectional with the lines $y = \lambda^{-1} x$ *symmetric* to them about the line $y = x$ (recall: the symmetric graph defines the *inverse function*, and its slope is precisely λ^{-1} (see Fig. 14)). As the rays get closer to the lines $y = \pm x$, the directions tend to the directions of these rays, and it is natural to assume that the integral curves also tend to them as to *asymptotes*. Try to guess a formula for the solutions. □

Example 4. For the differential equation $(*)$ $y' = -\frac{y}{x}$, the slope along the isocline $y = \lambda x$ is

$$k = k(\lambda) = -\frac{y}{x} = -\frac{\lambda x}{x} = -\lambda.$$

Hence, the directional field of this differential equation along the ray $y = \lambda x$ is codirectional with the rays $y = -\lambda x$ *symmetric* to them about the x-axis (Fig. 15). Along the semi-axes of the Ox and Oy axes, the directional field is *codirectional* with them, so the semi-axes are integral curves of equation $(*)$. As the rays $y = \lambda x$ approach the semi-axes, the directions along them tend to the directions of the semi-axes, and again it is natural to assume that the integral curves also tend to them as to *asymptotes* (see Fig. 15). We know such curves: the *hyperbolas* $xy = a = \text{const}$ have the necessary properties. It is easy to check that the corresponding functions do satisfy the differential equation $(*)$:

$$y = \frac{a}{x} \quad \Rightarrow \quad y' = -\frac{a}{x^2}; \quad -\frac{y}{x} = -\frac{a}{x} \cdot \frac{1}{x} = -\frac{a}{x^2} = y'.$$

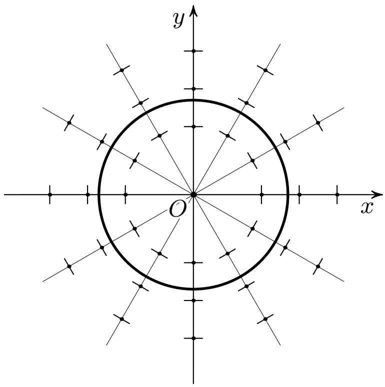


Fig. 16.

Equation (∗) can also be easily checked by differentiating the equality $xy = a$:

$$xy(x) \equiv a = \text{const} \quad \Rightarrow \quad (xy(x))' = y(x) + xy'(x) \equiv 0 \quad \Rightarrow \quad y'(x) = -\frac{y(x)}{x},$$

as required. □

Example 5. For the differential equation (∗) $y' = -\frac{x}{y}$, the slope along the isocline $y = \lambda x$ is

$$k = k(\lambda) = -\frac{x}{y} = -\frac{x}{\lambda x} = -\frac{1}{\lambda}.$$

In this case, the directional field is never codirectional with the isoclines:

$$k(\lambda) = \lambda \quad \Leftrightarrow \quad -\frac{1}{\lambda} = \lambda \quad \Leftrightarrow \quad \lambda^2 = -1.$$

On the rays $y = \pm x$, the directional field is perpendicular to them, as well as on the semi-axes (Fig. 16). The direction of the field on an arbitrary ray $y = \lambda x$ is obtained from the direction of this ray by two symmetries: about the line $y = x$ (then, as in Example 3, λ turns into λ^{-1}) and about the Ox axis (after which we finally obtain the slope $k(\lambda) = -\lambda^{-1}$). It is easily seen that *the composition of these symmetries is a rotation by* $-90°$, and so *the directional field along any ray is perpendicular to this ray* (see Fig. 16).

Hence it is clear that integral curves of the directional field of equation (∗) are all circles $x^2 + y^2 = R^2 = \text{const}$ centred at the origin, and solutions of the equation are functions of the form $y(x) = \pm\sqrt{R^2 - x^2}$. The latter can easily be verified by differentiating these functions:

$$y'(x) = (\pm\sqrt{R^2 - x^2})' = \frac{-2x}{\pm 2\sqrt{R^2 - x^2}} = -\frac{x}{y(x)},$$

or, as in the previous example, by differentiating the *"implicit function"* given by $x^2 + y^2(x) \equiv \text{const}$:

$$(x^2 + y^2(x))' = 2x + 2y(x) \cdot y'(x) \equiv 0 \quad \Rightarrow \quad y'(x) = -\frac{x}{y(x)},$$

as required. □

Below, in Ch. IV, we will show a method, dating back to Leibniz, for solving the differential equations of Examples 2–5, which "unveils the mystery" of our "guessing".

§ 2.3. Euler's Polygonal Lines, Solutions of the Equation $y' = f(x)$, and Integral

Now we will show how a geometric interpretation of differential equations and Euler's method outlined in the previous section can be applied to the analysis of a differential equation of the form

$$y' = f(x): \quad \forall x \in I \quad y'(x) = f(x), \tag{1}$$

where $I \subset D_f$ is a considered interval in the domain of $f(x)$. In particular, on this way we will also obtain the *existence theorem* for solutions of the differential equation (1), together with numerous applications.

2.3.1. Euler's polygonal lines and integral sums. Consider the directional field $M \mapsto \boldsymbol{L}_k(M)$ of a differential equation of the form (1). In this case the slope of the direction at a point $M(x; y)$ is $k(M) = f(x)$, and it depends only on the x-value of M. Geometrically this means that the directional field is *constant along vertical lines* $x = \text{const}$ (in other words, isoclines of such equations are composed of vertical lines). Hence we conclude that the directional field is *invariant* (i.e., maps to itself) under translations along these lines, and each integral curve maps to an integral curve — this corresponds to the fact that any two solutions of the differential equation $y' = f(x)$ (if the solutions exist!) differ from each other by a constant (Fig. 17). Hence a *uniqueness theorem* for solutions of the differential equation of the form (1) is also clear: if at least one solution of the equation $y' = f(x)$ $(x \in I)$ exists, then by a translation (shift) along the Oy axis we can make the shifted graph pass through the "initial point" $M_0(x_0; y_0)$ — obviously, such a translation is unique.

Now, using geometric arguments, let us try to *approximately* draw through the initial point $M_0(x_0; y_0)$ an integral curve, i.e., a graph of a solution of the differential equation $y' = f(x)$, $x \in I$, up to the point with x-value x. Recall (see Sec. 2.2.2) the method proposed by Leonhard Euler for constructing such *approximations* by n-segment polygonal lines $M_0M_1M_2 \ldots M_n$. These *Euler's polygonal lines* are obtained for any natural n in the following way.

We divide the segment from x_0 to x (the case $x < x_0$ is allowed as well, so we do not write "segment $[x_0, x]$") into n equal parts by points $x_1, x_2, \ldots, x_{n-1}, x_n = x$ $(x_m = x_0 + m\dfrac{x - x_0}{n})$, and starting from the point $M_0(x_0; y_0)$, we successively find points $M_m = (x_m; y_m)$, $m = 1, 2, \ldots, n$, so that each segment $M_m M_{m+1}$ goes along the directional field $\boldsymbol{L}$ at M_m (Fig. 18). The slope of the line $M_m M_{m+1}$ must be equal to $f(x_m)$, its equation is written as $y = y_m + f(x_m)(x - x_m)$, so we have $y_{m+1} = y_m + f(x_m)(x_{m+1} - x_m)$, and by substituting one of these recurrence formulae into another (for $m = 1, 2, \ldots, n$) we obtain the formula for the y-coordinate of the last point of the Euler polygonal line:

$$y_n = y_0 + \sum_{m=0}^{n-1} f(x_m)(x_{m+1} - x_m) = y_0 + S_n(f)_{x_0}^{x}. \tag{2}$$

The above sum, denoted by $S_n(f)_{x_0}^{x}$, is called the *n-th "left"*[1] *integral sum for the function f on the segment ∇ between x_0 and x.*

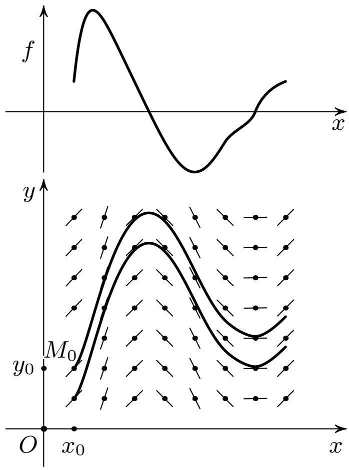


Fig. 17.

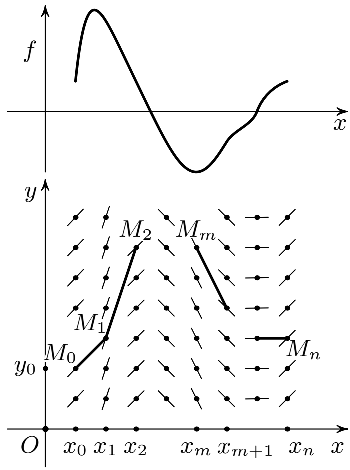


Fig. 18.

We *expect* that Euler's polygonal lines in the limit as $n \to \infty$ will give the desired integral curve $y = y(t)$, where t belongs to the segment ∇, and, in particular, that the value of the solution of the differential equation (1) with initial condition $y(x_0) = y_0$ at point x is

$$y(x) = \lim_{n\to\infty} y_n = \lim_{n\to\infty} (y_0 + S_n(x)),$$
$$S_n(x) = S_n(f)_{x_0}^{x} = \sum_{m=0}^{n-1} f(x_m)(x_{m+1} - x_m). \tag{3}$$

[1] To emphasise that, first, the segment is partitioned into n equal parts, and second, the m-th term in the sum (2) (the count starts from $m = 0$) is the product of the length of the m-th partition segment by the value of f at its *left-hand endpoint* x_m.

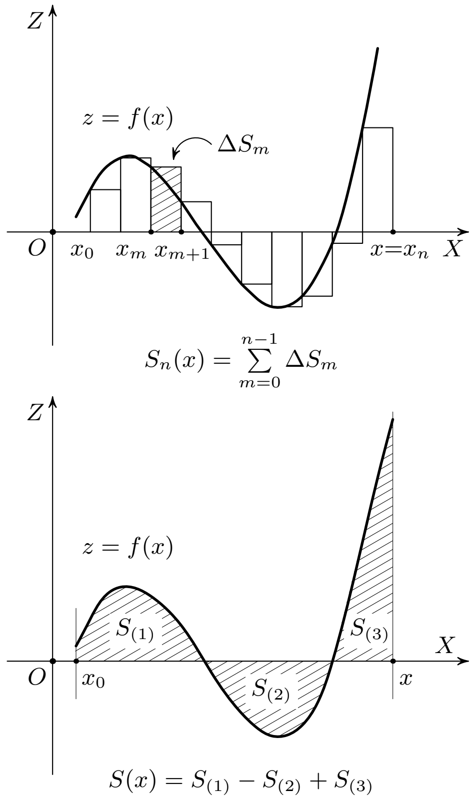


Fig. 19.

(Our belief is based on the fact that when constructing Euler's polygonal lines we "simply" replaced the graph of the solution with tangents to it on increasingly small segments.)

Another approach to finding solutions of the differential equation (1) on an interval I is *to consider for each* $x \in I$ the function $y(x)$ given by the limit formula (3) (if $x = x_0$, then we should assume that all integral sums are zero, so that $y(x_0) = y_0$) and *prove* that this function is a solution of equation (1). We will follow the second way; however, we *will not prove the existence of the limit* (3) but rather *find out what it should be equal to.*

2.3.2. Integral sums and quadratures (areas). Let us consider the integral sums $S_n(x) = S_n(f)_{x_0}^{x}$ more closely. Each summand $\Delta S_m = f(x_m)(x_{m+1} - x_m)$ of the sum $S_n(x)$ has a simple geometric interpretation on the graph of the function $z = f(x)$: if the function f is positive and

$x > x_0$, then the summand ΔS_m equals the *area* of a rectangle with the segment $[x_m, x_{m+1}]$ of the Ox axis as its base and with the upper left vertex $L_m(x_m; f(x_m))$ lying on the graph of $z = f(x)$ (Fig. 19, top). The entire integral sum equals the area of the *stepped figure* composed of these rectangles, shown in Fig. 19 (top). Note that if $f(x_m)$ is negative or if $x < x_0$, then the corresponding summand is still equal to the area of the same rectangle but *taken with the minus sign.*

It is natural to expect, at least for *continuous* functions f, that as $n \to \infty$, the area $S_n(x)$ of the stepped figure tends to the *area of the "curvilinear trapezium"* bounded by the Ox axis, the graph of the function $z = f(x)$, and the vertical lines drawn through the points $(x_0; 0)$ and $(x; 0)$ of the Ox axis (Fig. 19, bottom), with the area taken with the minus sign whenever the function is negative (or in the case of $x < x_0$): for $S_n(x) = S_n(f)_{x_0}^x$ we have

$$\lim_{n \to \infty} S_n(x) = S(x) = S(f)_{x_0}^x. \tag{4}$$

Comparing the hypothetic formulae (3) and (4), we obtain a *conjectural* formula for a solution $y(x)$ of the differential equation (1) on an interval I satisfying the initial condition $y(x_0) = y_0$:

$$y(x) = y_0 + S(x) = y_0 + S(f)_{x_0}^x.$$

Now let us *prove* this formula!

2.3.3. Fundamental theorem of calculus: derivative of a variable area.

Theorem 1 (fundamental theorem of calculus[1]). *If a function f is defined and continuous on an interval $I = (a, b)$, then the function $y(x) = y_0 + S(x) = y_0 + S(f)_{x_0}^x$ is a solution on this interval I of the differential equation* (1) *with the initial condition $y(x_0) = y_0$.*

Precisely in this form, as a *theorem on the derivative of a variable area* (but without using the terminology of differential equations), the *fundamental theorem of calculus* was published in *Lectiones* by Isaac Barrow (1667–1670). However, an exact *analytical* formulation of this theorem is due to Cauchy (1820s). The point is that before the XIX century mathematicians *"did not possess"* a precise (in the sense of rigour) *notion of continuity* (they as if "did not need" it). The definition of continuity of a function at a point was formulated by Cauchy (1821), and even earlier by Bolzano (1817), about whom a special mention should be made.

---

[1] Sometimes, the name "Fundamental Theorem of Calculus" is used to refer to the fundamental property of antiderivatives (§ 2.1, Corollary 1): *if a function $F(x)$ is an antiderivative of a function $f(x)$ on an interval I, then any its antiderivative $\widetilde{F}(x)$ differs from $F(x)$ by an additive constant C: $\widetilde{F}(x) = F(x) + C$.* As we shall see shortly, this statement is indeed very important, especially when combined with the *fundamental theorem* formulated here.

Bernard Bolzano (1781–1848) was a Bohemian monk and Catholic priest, theologian and religious philosopher, logician, and very profound mathematician who anticipated not only Cauchy's developments (he was the first to prove *"Cauchy's intermediate value theorem"*, where he *pioneered* the interval halving method, sometimes referred to as *Bolzano's method*) but also the research of Karl Weierstrass (1840–50s) and the set theory of Georg Cantor (1870–80s). Bolzano, 30 years before Weierstrass, gave an example of an *everywhere continuous nowhere differentiable function*, which shocked mathematicians to such an extent that even in 1905 the greatest French mathematician Henri Poincaré asked: "How can intuition deceive us on this point?", and his compatriot Charles Hermite wrote at the same time that he "turns away with fear and horror from that pitiful plague of functions without derivatives" (of course, his saying was quite ironic).

Bolzano contributed significantly to the *logical foundations of mathematics* — the theory of real numbers, mathematical logic, the justification of calculus (before Cauchy and Abel, he rather rigorously approached the notion of limit; Weierstrass followed his footsteps when introducing the concept of the continuum of numbers).

Proof. Since $S(x_0)=0$, the initial condition is obviously fulfilled. Let us prove that for any $x\in I$ there exists a derivative $S'(x)$ and that $S'(x)=f(x)$. We will assume that $x>x_0$ and $f(x)>0$. Then for any positive increment Δx we have

$$\Delta S(x)=S(x+\Delta x)-S(x)=S(f)_x^{x+\Delta x}\approx f(x)\Delta x$$

(Fig. 20), whence it follows that $\frac{\Delta S}{\Delta x}\approx f(x)$, and by the continuity of f this equality is the more accurate the smaller Δx. This is a rough idea of the proof of the equality $S'(x)=f(x)$. Let us conduct a strict proof.

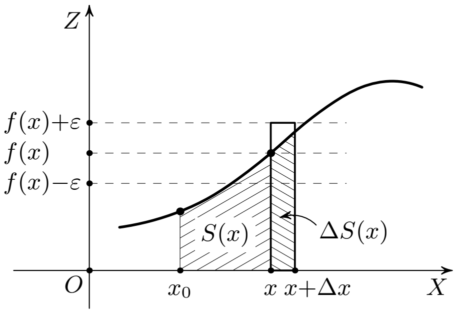


Fig. 20.

Assume that $x>x_0$ and $f(x)>0$. Since f is continuous at x, for any $\varepsilon>0$ there exists a neighbourhood Nghb_x such that if $x+h\in\mathrm{Nghb}_x$, then $|f(x+h)-f(x)|<\varepsilon$, or equivalently $f(x)-\varepsilon<f(x+h)<f(x)+\varepsilon$.

In the case where $\Delta x>0$ and $x+\Delta\in\mathrm{Nghb}_x$, this implies that the increment ΔS of the variable area $S(x)$ lies in the range $(f(x)-\varepsilon)\Delta x<\Delta S<(f(x)+\varepsilon)\Delta x$ (if we consider values of ε satisfying $f(x)-\varepsilon>0$, then the curvilinear trapezium between the vertical lines x and $x+\Delta x$ contains a rectangle with base $[x, x+\Delta x]$ and height $f(x)-\varepsilon$ and is contained in a rectangle with the same base and with height $f(x)+\varepsilon$; see Fig. 20). Hence, since $\Delta x>0$, by dividing all sides of the last inequality by Δx we obtain the estimate

$$f(x)-\varepsilon<\frac{\Delta S}{\Delta x}<f(x)+\varepsilon \quad \Rightarrow \quad \left|\frac{\Delta S}{\Delta x}-f(x)\right|<\varepsilon.$$

The same inequalities are valid in the case of $\Delta x<0$: then the increment ΔS is negative, and if $x+\Delta x \in \mathrm{Nghb}_x$, we have

$$(f(x)-\varepsilon)|\Delta x|<|\Delta S|<(f(x)+\varepsilon)|\Delta x|$$

(explain), whence it follows that

$$f(x)-\varepsilon<\frac{|\Delta S|}{|\Delta x|}=\frac{-\Delta S}{-\Delta x}=\frac{\Delta S}{\Delta x}<f(x)+\varepsilon.$$

Thus,

$$\forall \varepsilon>0 \quad \exists\, \mathrm{Nghb}_x \quad \forall x+\Delta x \in \mathrm{Nghb}_x \quad \left|\frac{\Delta S}{\Delta x}-f(x)\right|<\varepsilon,$$

and this precisely means that $\lim \dfrac{\Delta S}{\Delta x}=f(x)$ as $\Delta x \to 0$; i.e., the function (variable area) S is differentiable at x, and moreover, $S'(x)=f(x)$.

In a similar way but taking into account *"signed areas"*, the cases where $f(x)<0$ or $x<x_0$ can be considered; examine these cases on your own. □

2.3.4. Existence theorems for solutions of the equation $y'=f(x)$ and antiderivatives. The *fundamental theorem of calculus* just proved allows us to formulate *sufficient* (but by no means necessary! — see Example 5 in § 2.1) *existence conditions for antiderivatives*, or equivalently, *for solutions of the differential equation* $y'=f(x)$.

Corollary 1 (existence theorem for solutons of the differential equation $y'=f(x)$)**.** *If a function f is continuous on an interval or on a union of intervals D, then the differential equation $y'=f(x)$, $x \in D$, has solutions.*

Corollary 2 (existence theorem for antiderivatives)**.** *Any function f which is continuous on an interval or on a union of intervals D has infinitely many antiderivatives on D.*

2.3.5. Areas of curvilinear trapezia as increments of antiderivatives. We have shown above that *in order to find an antiderivative $y(x)$ of a function continuous on an interval I, it suffices to find the* area *$S(x)=S(f)_{x_0}^{x}$ of the curvilinear trapezium enclosed between the Ox axis, the graph of the function $z=f(x)$, and vertical lines at x_0 and x*: if we require in addition that $y(x_0)=y_0$, then the desired antiderivative is given by

$$y(x)=y_0+S(x)=y_0+S(f)_{x_0}^{x}. \tag{5}$$

However, as it often happens, in practice we do *the other way round*: not the antiderivatives are found by finding the areas (*variable* areas), but on the contrary, *areas are found with the help of antiderivatives.* To learn how to do this, let us rewrite equation (5) in the form

$$S(x) = S(f)_{x_0}^{x} = y(x) - y_0 = y(x) - y(x_0).$$

We obtain that the area of the curvilinear trapezium between the vertical lines at points x_0 and x *is equal to the increment of the antiderivative* $y(x)$ on the segment from x_0 to x:

$$S(f)_{x_0}^{x} = y(x) - y(x_0) = y(t)\Big|_{x_0}^{x}.$$

However, from the *theorem on the equality of increments of antiderivatives* (§ 2.1, Theorem 2) it follows that instead of the antiderivative $y(x)$, which by its very *definition* is related to areas, we may take *any* antiderivative F of f on the interval I: anyway, we have

$$S(f)_{x_0}^{x} = y(t)\Big|_{x_0}^{x} = F(t)\Big|_{x_0}^{x} = F(x) - F(x_0),$$

and an antiderivative F can be "guessed" (fitted, found) without regard to areas. Putting $x_0 = a$ and $x = b$, let us formulate the proved statement.

Theorem 2 (Barrow's theorem; 1669–1670)**.** *If a function f is defined and continuous on an interval I, then for any two points $a, b \in I$ the area $S(f)_a^b$ (taking into account the sign of the function f and the sing of the difference $b - a$) of the curvilinear trapezium bounded by*

$$y = 0, \qquad y = f(x), \qquad x = a, \qquad x = b,$$

is equal to the increment of (any) antiderivative F of f on the segment from a to b:

$$S(f)_a^b = F(t)\Big|_a^b = F(b) - F(a). \tag{6}$$

Isaac Barrow was the first to see the relationship between *areas and antiderivatives.* Newton interpreted this relationship using the idea of velocity: the path (fluent) is the area below the graph of velocity (fluxion). Leibniz came from quadratures (areas) to antiderivatives by 1695.

Example 1 (Archimedes' problem: *quadrature of the parabola*)**.** Let us find the area of a "curvilinear triangle" bounded by the parabola $y = \alpha x^2$, the Ox axis, and the vertical line $x = c > 0$ (Fig. 21). Clearly, this "triangle" can be considered a curvilinear trapezium: for it we have $a = 0$, $b = c$, and $f(x) = \alpha x^2$. Since (for example!) the function $F(x) = \frac{1}{3}\alpha x^3$ $(x \in \mathbb{R})$ is an antiderivative for f (since *any* antiderivative is suitable, we do not write an

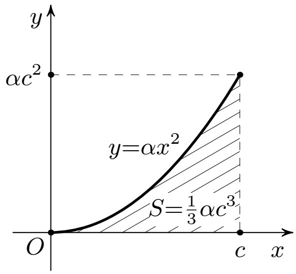


Fig. 21.

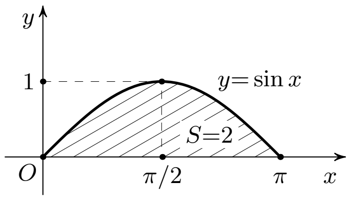


Fig. 22.

additive constant or assume it to be zero), from Barrow's formula (6) we obtain

$$S(\alpha x^2)_0^c = \tfrac{1}{3}\alpha x^3\Big|_0^c = \tfrac{1}{3}\alpha c^3 - 0 = \tfrac{1}{3}\alpha c^3.$$

The result can also be written as $S = \frac{1}{3} \cdot c \cdot \alpha c^2$: the area of a "parabolic triangle" is one third of the area of its circumscribed rectangle (bounded by the coordinate axes and the lines $x = c$ and $y = f(c) = \alpha c^2$; see Fig. 21). □

Archimedes in the III century BC indeed solved many problems that can be reduced to the considered quadrature; however, he did not see the generality of his methods. Among the problems solved by Archimedes was finding the area of an *arbitrary parabolic segment* bounded by a parabola and any of its chords. We will address this problem in the exercises to this chapter.

Example 2. (Pascal's problem: *quadrature of the sine curve.*) Let us find the area of the "curvilinear biangle" bounded by the Ox axis and the graph of the function $y = \sin x$ from 0 to π (Fig. 22). This "sine lobe" can be considered a curvilinear trapezium: for it we have $a = 0$, $b = \pi$, and $f(x) = \sin x$. As an antiderivative for $\sin x$ we take $F(x) = -\cos x$ $(x \in \mathbb{R})$, and from Barrow's formula (6) we obtain

$$S(\sin x)_0^\pi = [-\cos x]_0^\pi = (-\cos \pi) - (-\cos 0) = -(-1) - (-1) = 2.$$

An intriguing result. □

This problem was solved by Pascal in 1659 by summation in a more general case: for any part of a sine lobe cut off by a vertical line. This was, in the terminology to which we shall now turn, the first *integral of a trigonometric function* in mathematics. The solution itself, given by Pascal in *Traité des sinus du quart de cercle* (Treatise on the Sine of a Quarter of a Circle), played an outstanding role in the history of calculus: in 1672, thanks to a happy accident, Leibniz (in between his diplomatic affairs!), who wanted to get acquainted with contemporary mathematics, met Huygens, who immediately gave him Pascal's works. Leibniz later wrote to Johann Bernoulli: "I happened to come across a proof of Dettonville's [Pascal's pseudonym] that was of a supremely easy nature. But nothing astonished me so much as the fact that Pascal seemed to have had his eyes obscured by some evil fate; for I saw at a glance that the theorem [actually, a lemma leading to the quadrature in question] was a most general one for any kind of curve whatever" [Pascal was only concerned with the circle]. Thus, like Archimedes, Pascal did not see a general method in a particular problem. We will return to Pascal's activity below.

2.3.6. Integral sums and integral. Let us return to the consideration of integral sums for a function f on an interval from a to b:

$$S_n = S_n(f)_a^b = \sum_{m=0}^{n-1} f(x_m)\Delta x_m = \sum_{m=0}^{n-1} f(x_m)(x_{m+1} - x_m). \tag{7}$$

Here $\Delta x_m = x_{m+1} - x_m$, $x_0 = a$, $x_n = b$, and the intermediate points x_1, x_2, ..., x_{n-1} divide the segment between a and b (it is allowed that $a > b$) into n equal parts: $x_m = a + \frac{b-a}{n}$. We have given the integral sums (7) a geometric interpretation with the help of the graph of $z = f(x)$, considering the individual terms $\Delta S_m = f(x_m)\Delta x_m = f(x_m)(x_{m+1} - x_m)$ of the sum S_n as (signed) areas of the corresponding rectangles inscribed in the curvilinear trapezium. However, even ancient Greek mathematicians (Eudoxus, Archimedes), and then Cavalieri, Kepler, and other scientists of the "pre-Newtonian" era used integral sums with summands that were not areas but volumes, masses, etc. Although the sums S_n can still be interpreted as areas of step figures, it makes sense to introduce a limit of integral sums without regard to areas.

Definition 1. For a continuous function f on a segment from a to b, the limit of its integral sums (7) on this segment as $n \to \infty$ is called the *integral* of f from a to b (or on the segment from a to b) and is denoted by

$$\int_a^b f(x)\,dx \stackrel{\text{def}}{=} \lim_{n\to\infty} S_n(f)_a^b = \lim_{n\to\infty} \sum_{m=0}^{n-1} f(x_m)\Delta x_m.$$

The numbers a and b are called, respectively, the *lower* and *upper integration limits* (when $a = b$, the integral is defined to be zero), the letter x appearing in the notation for the *integrand* $f(x)\,dx$ is called the *integration variable*; instead of x we may choose any other letter — the integral (the number obtained in the limit) will not change:

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(u)\,du = \int_a^b f(\xi)\,d\xi$$

(however, you should avoid using the same letters that appear in the notation for the integration limits![1]).

[1] However, physicists quite often use the same letter in one and the same expression denoting different quantities by it; this does not embarrass them in the least.

The terms *integral*, *integration* were first used in 1690 by Jacob Bernoulli, who derived them, most likely, from Latin *integer*, 'whole', 'entire' (according to another assumption, Bernoulli derived them from the verb *integro*, 'restore', 'make whole'). Before that time, Cavalieri's terminology was used: "aggregate of all indivisibles" (see the next section) or simply "all lines", "sum of lines" (Latin *omnes lineae*, *summa lineae*).

Accordingly, for the integral itself, the abbreviation *omn* was used (e.g., by Wallis, and after him by Leibniz). From 1675 Leibniz, for the sake of even greater brevity, began to write the initial letter of the word *Summa*, whose stylised writing at that time was similar to the present-day integral sign. Initially Leibniz wrote $\int y$, but a month later he introduced the independent variable (variable of integration) into the notation: $\int y\,dx$. Why specifically d, it is often explained that this was a modified symbol Δ for the increment of the variable; however, the notation Δx, Δy for increments was introduced only in 1755 by Leonhard Euler. The meaning of Leibniz's notation will be explained in Ch. IV (§ 4.2).

The modern notation for integration limits appeared in Fourier's *Théorie Analytique de la Chaleur* (The Analytical Theory of Heat; 1822), replacing Euler's notation $\int y\,dx \begin{bmatrix} \text{ab } x=a \\ \text{ad } x=b \end{bmatrix}$ (it is written in Latin: 'from $x=a$ to $x=b$'). Before Euler, limits were written in words.

Independently of the concept of area, the following statement can be proved, which we will assume without proof.

Theorem 3 (Cauchy–Riemann). *For any continuous function f on a segment from a to b there exists a limit of the sequence of its integral sums*:

$$\lim_{n\to\infty} S_n(f)_a^b = \lim_{n\to\infty} \sum_{m=0}^{n-1} f(x_m)\Delta x_m.$$

Cauchy, who had a precise notion of continuity, was the first to replace the geometrical treatment of the integral with an *analytical* one, and he thought he had proved the above theorem (1823). However, he did not notice a gap, the necessity of a rather subtle property of functions continuous on a segment. This omission was noted only by Riemann, who considered more general integrability criteria than continuity (1853). The property of uniform continuity needed by Cauchy was proved only in 1870 by the German mathematician Heinrich Eduard Heine (1821–1881), who also gave a definition of the corresponding concept.

There is an *integrability condition* for differential equations named after Cauchy and Riemann, which in fact was derived much earlier by Euler and d'Alembert.

This theorem allows us to deal with integrals of continuous functions irrespective of the geometrical notion of area. However, if the notion of area is introduced in such a way that areas of curvilinear trapezia can also be handled, then the above limit is equal to the corresponding area, so *the area can be written through the integral*:

$$\int_a^b f(x)\,dx = S(f)_a^b. \tag{8}$$

Comparing equations (8) and (6), we obtain Barrow's formula, which is usually referred to as the *Newton–Leibniz theorem* (in spite of the fact that neither of them had anything to do with its discovery and publication — though in a more "geometric" notation — to which they never claimed!): *if F is any antiderivative for a continuous function f on an interval from a to b, then*

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b = F(b) - F(a). \tag{9}$$

Barrow's formula can be directly derived from the analytical definition of the integral if one formulates and proves the *fundamental theorem of calculus* independently of areas.

Theorem 4 (Cauchy–Moigno theorem [1823–1844] on the derivative of integral with variable upper limit; the "fundamental theorem of calculus")**.** *If a function f is continuous on an interval I, then for any $a \in I$ the derivative of the integral of f with lower limit a and variable upper limit $x \in I$ equals the integrand function $f(x)$:*

$$\left(\int_a^x f(u)\,du\right)'_x = f(x). \tag{10}$$

Conversely, this theorem immediately follows from Barrow's formula (9) if we put $b = x$ in it and then take the derivative (of the left- and right-hand sides). Barrow's formula can be interpreted as follows: *if a function F has a continuous derivative on a segment from a to b, then*

$$\int_a^x F'(x)\,dx = F(x)\Big|_a^b = F(b) - F(a). \tag{11}$$

Augustin Louis Cauchy (1789–1837) was the greatest French mathematician of the first half of the XIX century, the founder of many branches of modern mathematics. He graduated from Napoleon's École Polytechnique, and after the return of the Bourbons he was its professor until the Revolution of 1830. An ultra-royalist and clericalist (unlike the oppositionists d'Alembert and Jacobi, the revolutionaries Fourier, Monge, Galois, and the indiscriminate Laplace), Cauchy went into exile with the Bourbons to Turin and Prague after the 1830 revolution. But even there he did not stop doing mathematics. Extremely versatile and hardworking, he published more than 800 papers; there were times when Cauchy submitted a new memoir to the Paris Academy every week.

In addition to Cauchy's seminal work on "bringing rigour and order" to calculus, his most important contributions include the development of the *theory of functions of a complex variable*. Cauchy's work in this area was continued by Riemann, who also "put his hand" in calculus by proposing his approach to integrability of functions.

François-Napoléon-Marie Moigno was a French mathematician, Catholic priest, a student of Cauchy, who was the first to publish the above theorem in his *Leçons de calcul* *différentiel et de calcul intégral, rédigées d'après les méthodes et les ouvrages publiés ou inédits de A.L. Cauchy* (Lectures on Differential and Integral Calculus, Based on the Methods and Published or Unpublished Works of A.L. Cauchy; 1840–1844).

The operation of finding integrals is called integration, and the main meaning of one of the greatest discoveries in the history of mathematics — the invention of calculus by Newton and Leibniz, as well as their predecessors, contemporaries, and followers — is that they revealed the inverse nature of differentiation and integration; this very fact is reflected in equations (10) and (11).

§ 2.4. Geometric Applications of Integral

2.4.1. Main idea: applying Barrow's formula. In this section we write and apply various formulae using which *measures* (areas, volumes, and in the exercises also lengths) of various geometric figures can be found as *integrals* of some auxiliary functions. However, the introduction (definition) of measures (areas, volumes) themselves is assumed to be known from the corresponding topics of geometry, and all the figures under consideration are assumed to be "good enough", i.e., such that the question of their "measurability" (in the sense of existence of their areas, volumes, etc.) is not discussed at all (in concrete examples, their measurability will be obvious).

The *integral formulae* to be considered can be derived in different ways, according to the way we interpret the concept of integral itself. Four "interpretations" of integral have been considered above:

— Integral as a limit of integral sums (by definition);
— Integral as the (signed) area of a curvilinear trapezium;
— Integral as the increment of an antiderivative (Barrow's formula);
— Integral of a derivative as the increment of a function, or the "second" Barrow's formula: *if a function F has a continuous derivative on a segment from a to b, then*

$$F(b) - F(a) = F(x)\Big|_a^b = \int_a^b F'(x)\,dx. \qquad (1)$$

This integral formula allows us to reconstruct a function given its derivative; we will use it as a basis for integral formulae for evaluating areas and volumes of plane and spatial figures. For this purpose, we will first interpret these measures as *functions $y = y(x)$ of one variable*; then we will find (making standard intuitively clear assumptions) derivatives of these auxiliary functions, i.e., we will actually write down differential equations of the form $y' = f(x)$; and lastly we will obtain the desired *integral formulae* from Barrow's integral formula (1). These rather general considerations will become quite clear when we consider the very first example.

Note that in this case the differential equations will describe *not the process* of change of the measure (area, volume) in time but the dependence of a variable measure on a parameter, i.e., some auxiliary coordinate related to the geometric figure under consideration. Even though this coordinate can be tentatively considered as *time*, this plays no role for the efficiency of using information about the derivative of the desired function, i.e., for integrating the resulting differential equation.

2.4.2. Areas of plane figures. So, let us show how this approach can be used to derive an *integral formula for the area of a plane figure.* Let Φ be a "good"[1] bounded figure in a plane. Choose an arbitrary coordinate axis Ox in the plane and denote by $S(x)$ the area of the part of Φ that lies *to the left* (with respect to the positive direction of the Ox axis) of the line L_x drawn through the point x perpendicular to the axis (Fig. 23).

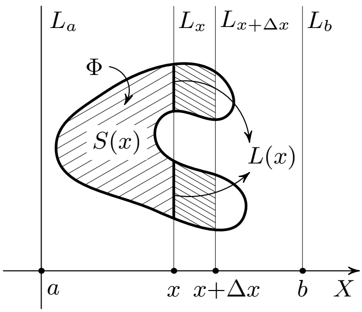


Fig. 23.

Assuming that the intersection of the line L_x with the given figure Φ can consist of at most finitely many segments (or separate points), let us denote their total length by $L(x)$. Then *it is natural to assume* that the *increment of the variable area* $S(x)$ over the interval from x to $x+\Delta x$, i.e., the area $\Delta S(x) = S(x+\Delta x) - S(x)$ of the part of the figure Φ enclosed between the lines L_x and $L_{x+\Delta x}$, is approximately

$$\Delta S(x) = S(x+\Delta x) - S(x) \approx L(x) \cdot \Delta x,$$

and *in the case where* $L(x)$ *is a continuous function*, this equality is the more exact the smaller Δx, so that after dividing by Δx we should obtain

[1]For example, bounded by finitely many arcs of graphs of differentiable functions.

an exact equality as $\Delta x \to 0$:

$$\lim_{\Delta x \to 0} \frac{\Delta S(x)}{\Delta x} \stackrel{\text{def}}{=} S'(x) = L(x). \tag{2}$$

Note this reasoning; in the next chapter (in § 3.1), almost the same reasoning will be used to derive linear differential equations for growth/decay. Note that in many applications of mathematics to practical problems, such plausible arguments are inevitable at some stage.

Since the figure Φ is bounded (which can be interpreted, for example, as lying entirely inside some circle), it is enclosed between some lines L_a and L_b perpendicular to the Ox axis, where $a < b$. But then for these a and b we have $S(a) = 0$ and $S(b) = S[\Phi]$, the area of the whole figure Φ, so $S[\Phi] = S(b) - S(a)$, which by Barrow's formula (1) and in accordance with equation (2), $S'(x) = L(x)$, *assuming the continuity of the "length of the variable section"*, i.e., of the function $L(x)$, precisely gives the *integral formula for the area of an arbitrary plane figure* Φ:

$$S[\Phi] = S(b) - S(a) = \int_a^b L(x)\, dx \tag{3}$$

(the area equals the integral of the lengths of "perpendicular" sections of the figure).

Of course, exactly in this form, this formula is almost never used "in practice". We need it for further, step by step, progress in geometric integral formulae. On the other hand, if with our choice of the Ox axis it turned out that the figure Φ is simply a curvilinear trapezium $\{a \le x \le b; 0 \le y \le f(x)\}$ for a non-negative function $f(x) = L(x)$, then equation (3) would not be any different from the usual Barrow formula. A real application of formula (3) is evaluation of areas of regions enclosed between two vertical lines $x = a$ and $x = b$ and two non-intersecting graphs of continuous functions $f_1(x) \le f_2(x)$, such as those shown in Fig. 24: then $L(x) = f_2(x) - f_1(x)$, and the area of such a region is

$$S[\Phi] = \int_a^b [f_2(x) - f_1(x)]\, dx, \tag{4}$$

independently of the signs of the functions$f_{1,2}$. (However, formula (4) can easily be derived directly from the geometrical meaning of integrals of non-negative functions: it suffices to move the region vertically so that it would lie above the Ox axis.)

Now we proceed to integral formulae for volumes.

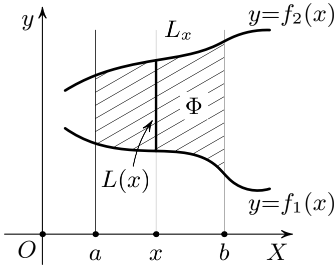


Fig. 24.

2.4.3. Volume of a general right cylinder.

Definition 1. Let Φ be a plane ("good", bounded) figure in space. A *general right cylinder* C_Φ with *base* Φ and *height* h is the union of segments MM' of length h drawn through all points $M \in \Phi$ perpendicular to the plane of the figure Φ on one side of it.[1]

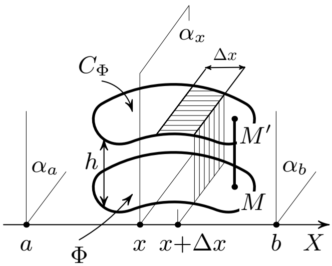


Fig. 25.

Let us derive a formula for the volume of such a cylinder. To this end, we choose an arbitrary coordinate axis Ox in the base plane and denote by $V(x)$ the volume of the part of the cylinder C_Φ that lies *to the left* (with respect to the positive direction of the Ox axis) of the plane α_x drawn through the point x of the Ox axis perpendicular to it (Fig. 25). Let us find the derivative of $V(x)$.

As above, we assume that the intersection of the plane α_x with the base Φ consists of at most finitely many segments of total length $L(x)$. For simplicity, we will assume that we have exactly one such segment (of length $L(x)$), as shown in Fig. 25. Then the increment of the variable volume $V(x)$ over the interval from x to $x + \Delta x$ is the volume $\Delta V(x) = V(x + \Delta x) - V(x)$ of the part of the cylinder C_Φ enclosed between the planes α_x and $\alpha_{x+\Delta x}$, which can be approximated (replaced) by a rectangular parallelepiped whose

[1] We obtain a "sole" with base Φ and height h, as shown in Fig. 25.

base is a rectangle with sides $L(x)$ and h, the intersection of the cylinder with the plane α_x, and whose height is Δx (if $\Delta x < 0$, the volume increment ΔV is negative, and this case leads to exactly the same conclusions as below). Therefore, we can approximately write

$$\Delta V(x) = V(x + \Delta x) - V(x) \approx L(x)h \cdot \Delta x,$$

and (assuming the continuity of $L(x)$) this equality is the more exact the smaller $|\Delta x|$, so that after dividing by Δx we should obtain an exact equality as $\Delta x \to 0$:

$$\lim_{\Delta x \to 0} \frac{\Delta V(x)}{\Delta x} \stackrel{\text{def}}{=} V'(x) = hL(x).$$

Choosing the numbers $a < b$ so that the general cylinder C_Φ is enclosed between the planes α_a and α_b, we apply to the last relation Barrow's formula (1) together with the already proved integral formula (3) for the area of a plane figure. This results in a *formula for the volume of the general right cylinder* C_Φ:

$$V[C_\Phi] = V(b) - V(a) = \int_a^b [hL(x)]\, dx = h \int_a^b L(x)\, dx = h \cdot S[\Phi] = S_\Phi \cdot h; \quad (5)$$

i.e., *the volume of the general right cylinder is the product of its base area and its height.*

Though this formula is also "intermediate" for us (preparatory to the next one), let us note two special cases of it:

1. The *volume of an* arbitrary right prism *is the product of its base area and its height*, $V = SH$;

2. The *volume of a* right circular cylinder *is the product of its base area and its height*, $V = \pi R^2 H$.

2.4.4. Integral formula for volumes (integral of cross-section areas). Let $\mathbb{B}$ be an arbitrary "good" bounded solid figure.[1] Let us derive an *integral formula* for its volume.

As above, we choose an arbitrary coordinate axis Ox in space and denote by $V(x)$ the volume of the part of $\mathbb{B}$ that lies *to the left* (with respect to the positive direction of the Ox axis) of the plane α_x drawn through the point x of the Ox axis perpendicular to it (Fig. 26). Next, let us find the derivative of $V(x)$.

Denote the plane figure obtained by intersection of the plane α_x with the given solid figure $\mathbb{B}$ by Φ_x. We assume that each of these plane figures Φ_x is good and has area $S[\Phi_x] = S(x)$. The increment of the variable volume $V(x)$ over the interval from x to $x + \Delta x$ is the volume $\Delta V(x) = V(x + \Delta x) - V(x)$

[1] Or a *solid body*; see textbooks in geometry.

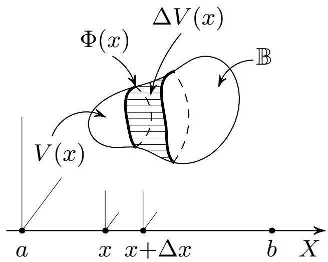


Fig. 26.

of the "segment" (part) $\Delta\mathbb{B}$ of the figure $\mathbb{B}$ enclosed between the planes α_x and $\alpha_{x+\Delta x}$, i.e., between the sections Φ_x and $\Phi_{x+\Delta x}$. *Assuming that the solid figure* $\mathbb{B}$ is "so good" that *at close values of x the sections Φ_x are also close* (in an intuitively clear sense), we replace this segment $\Delta\mathbb{B}$ by a close to it general right cylinder C_x with base Φ_x and height Δx. Then for the increment $\Delta V(x)$ of the variable volume we can write an approximate formula

$$\Delta V(x) = V(x+\Delta x) - V(x) \approx V[C_x] = S(x)\cdot\Delta x$$

(here we have applied formula (5)) for the volume of the general right cylinder already proved; obviously, the same approximate formula is valid in the case $\Delta x < 0$ too, since $\Delta V(x) < 0$ in this case). Again, under the assumption that $S(x)$ is continuous, this equality is *the more exact* the smaller $|\Delta x|$, so after dividing by Δx we should obtain an exact equality as $\Delta x \to 0$:

$$\lim_{\Delta x\to 0}\frac{\Delta V(x)}{\Delta x} \overset{\text{def}}{=} V'(x) = S(x).$$

Finally, choosing the numbers $a < b$ so that the solid figure $\mathbb{B}$ is entirely enclosed between the planes α_a and α_b, from Barrow's formula (1) we obtain an *integral formula for the volume of an arbitrary* (*"sufficiently good"*!) *solid figure*,

$$V[\mathbb{B}] = V(b) - V(a) = \int_a^b S(x)\,dx; \tag{6}$$

i.e., the *volume of a solid figure equals the integral of the area of "perpendicular" sections of the figure.*

The derived formula is sometimes briefly referred to as the *cross-section integral formula*, or simply the *cross-section integral.* Let us consider two important special cases of this formula.

2.4.5. Volume of a general cone.

Definition 2. Let Φ be a plane ("good", bounded) figure in space, and let P be a point not lying in the plane of Φ. A *general cone* K_Φ with *base* Φ and *vertex* P is the union of segments PM connecting the vertex P with all points M of the base Φ. The length h of the perpendicular PH drawn from the vertex of the cone to the base plane is called the *height* of the general cone (Fig. 27).[1]

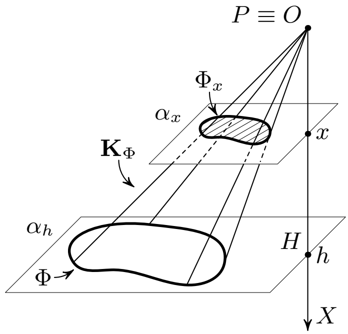


Fig. 27.

Note that arbitrary *pyramids* as well as *circular cones* (for instance, *right circular cones*) are *particular cases* of the general cone.

Using the cross-section integral, let us find the formula for the volume of the general cone. As a coordinate axis Ox we take the line PH containing the height of the cone K_Φ, choose the origin at the cone vertex P, and direct the axis towards the base plane (towards the point H). Then each of the required sections $\Phi_x = K_\Phi \cap \alpha_x$ of the cone is obtained from the cone base Φ by *homothety* with coefficient $k = \frac{x}{h}$ (explain!), so the area of the section is

$$S(x) = S[\Phi_x] = S[\Phi] \cdot \left(\frac{x}{h}\right)^2 = \frac{S_\Phi}{h^2} x^2.$$

Clearly, the integration limits in the cross-section integral (6) are in this case $a = 0$ and $b = h$, so we find the *volume of the cone* as the following integral:

$$V[K_\Phi] = \frac{S_\Phi}{h^2} \int_0^h x^2\,dx = \frac{S_\Phi}{h^2} \left[\frac{x^3}{3}\right]_0^h = \frac{S_\Phi}{h^2} \cdot \frac{h^3}{3} = \frac{1}{3} S_\Phi h; \qquad (7)$$

i.e., the volume of the general code is one third the product of the base area and the height.

[1]Note that the *segment PH itself* is also called in geometry the *height of the cone.*

For the above-mentioned particular cases of the general cone we obtain the following:

1 (Democritus–Eudoxus formula). *The volume of an* arbitrary pyramid *is one third the product of its base area and its height*: $V = \frac{1}{3}SH$;

2. *The volume of a* right circular cone *is one third the product of its base area and its height*: $V = \frac{1}{3}\pi R^2 H$.

Democritus of Abdera (c. 460–370 B.C.) was a prominent philosopher, materialist, and Thomist. His materialism was fiercely hated by the great Plato, who burned his writings. According to Marx and Engels, however, Democritus was "the first encyclopaedic mind among the Greeks". Based on his atomistic views, Democritus considered solid bodies to be composed of "parallel slices one atom thick", thus partly anticipating the very fruitful views of Cavalieri (XVII century; see below). In one way or another, without giving any proof, Democritus formulated both of the above statements.

Only a half-century later, Eudoxus of Cnidus (c. 408–355 B.C.), a mathematician and astronomer, an opponent of mysticism and astrology, gave a quite rigorous proof of them. Archimedes (about him, also see below) noted that it was important to Eudoxus that he knew in advance the answer given by Democritus!

2.4.6. Volume of a solid of revolution. Let $r = r(x)$ be a non-negative function continuous on a segment $[a, b]$, and let T be the corresponding curvilinear trapezium:

$$T = \{(x; y) \mid a \le x \le b,\ 0 \le y \le r(x)\}.$$

Consider a *solid of revolution* $\mathbb{B}_T$, i.e., a solid figure obtained by rotating the curvilinear trapezium T about the Ox axis. If someone does not like that we use in this "definition" the kinematic notion of *rotation*, this definition of a solid of revolution $\mathbb{B}_T$ can be replaced with a "purely static" definition by requiring that the intersection of $\mathbb{B}_T$ with each plane α_x drawn through a point x perpendicular to the Ox axis is a *circle of radius* $r = r(x)$ *with centre on the* Ox *axis* (Fig. 28). Let us write the general *integral formula* for the volume of this solid of revolution.

In this case each section Φ_x of $\mathbb{B}_T$ is a circle of radius $r = r(x)$, areas of the sections are accordingly $S(x) = \pi r^2 = \pi r^2(x)$, so by the cross-section integral (6) we immediately obtain the desired formula:

$$V[\mathbb{B}_T] = V(b) - V(a) = \int_a^b S(x)\,dx = \int_a^b \pi r^2(x)\,dx = \pi \int_a^b r^2(x)\,dx. \qquad (8)$$

That's the whole problem! Now we apply this formula to one single yet extremely important special case.

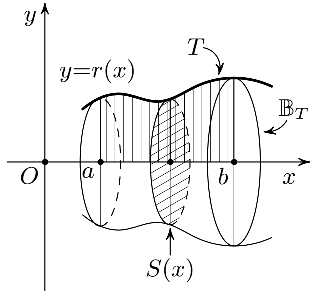


Fig. 28.

2.4.7. Volume of a sphere. A sphere B_R of radius R with center O can be considered as a solid of revolution of the semicircle

$$\{(x;y)\mid -R\le x\le R,\ 0\le y\le r(x)=\sqrt{R^2-x^2}\}$$

around the Ox axis (see Fig. 29). Applying formula (8) to this case, we find[1]

$$V(B_R)=\pi\int_{-R}^{R}(R^2-x^2)\,dx=\pi\Big[R^2x-\frac{x^3}{3}\Big]_{-R}^{R}=\pi\Big(2R^3-\frac{2}{3}R^3\Big)=\frac{4}{3}\pi R^3; \tag{9}$$

this is *Archimedes' first formula*!

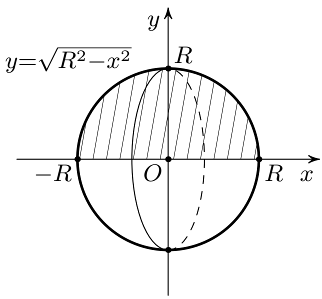


Fig. 29.

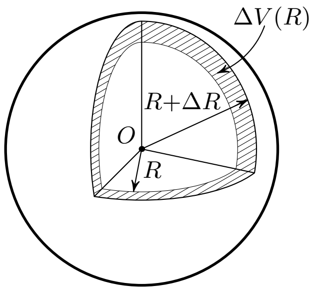


Fig. 30.

2.4.8. Remark on the area of a sphere. Now let us consider the volume of a sphere (more precisely, the volume of a ball, but classically

[1] Actually, this is one of those calculations that *everyone should do for themselves at least once in their life*!

referred to as the volume of a sphere) of radius R as a function of R:

$$V(R) = V(B_R) = \frac{4}{3}\pi R^3.$$

Obviously, the increment $\Delta V(R) = V(R + \Delta R) - V(R)$ of this function equals the volume (taken with the minus sign if $\Delta R < 0$) of the *spherical layer* of "width" $|\Delta R|$, i.e., the part of the space enclosed between two concentric spheres of radii R and $R + \Delta R$ (Fig. 30). Note that for a small increment ΔR of the radius, the volume of this layer *is approximately equal to the surface area $S(R)$ of the sphere B_R times the layer width ΔR*,

$$\Delta V(R) = V(R + \Delta R) - V(R) \approx S(R) \cdot \Delta R,$$

and (*by the continuity of the function $S(R)$!*) this equality is the more exact the less ΔR, so that after dividing by ΔR we should obtain an exact equality as $\Delta R \to 0$:

$$\lim_{\Delta R \to 0} \frac{\Delta V(R)}{\Delta R} \overset{\text{def}}{=} V'(R) = S(R).$$

But we know that $V(R) = \frac{4}{3}\pi R^3$, so

$$S(R) = V'(R) = \left(\frac{4}{3}\pi R^3\right)' = \frac{4}{3}\pi \cdot 3R^2 = 4\pi R^2; \tag{10}$$

that's *Archimedes' second formula*!

Archimedes of Syracuse (c. 287–212 B.C.) was the greatest mathematician, physicist, and mechanic of the Hellenistic era. Despite the most important discoveries and practical inventions in mechanics and physics, above all Archimedes placed his results from his work *On the Sphere and Cylinder* (letter to Dositheus, one of Euclid's pupils), which were formulated as follows: *if a sphere is inscribed in a cylinder, the volume of the sphere is $\frac{2}{3}$ of the volume of the cylinder, and its surface area is $\frac{2}{3}$ of the total surface of the cylinder*. Archimedes wished to have a drawing of this theorem depicted on his tomb, which was fulfilled by the Roman commander Marcellus.

Although Archimedes far surpassed Eudoxus in the methods he used to find areas and volumes, his methods were not universal (but not so long ago they were the methods used in school textbooks on stereometry!).

2.4.9. Geometric measures and integral sums. Cavalieri's principle. Note that all the formulae found above, both the general formulae (3) and (6)–(8) and "special" formulae (9) and (10), can as well be obtained by means of integral sums. For example, to find the volume of an arbitrary (though "good") solid (body) $\mathbb{B}$, we can divide it into n pieces (slices) by planes α_{x_i}, where the points x_i divide the segment $[a, b]$ into n equal parts (Fig. 31). The volume of the i-th slice is approximately $\Delta V_i = S(x_i)\Delta x_i$, where $\Delta x_i = x_{i+1} - x_i = \dfrac{b-a}{n}$ and where by $S(x)$, as above, we denote the area of the section $\mathbb{B} \cap \alpha_x$. Then for the volume V of $\mathbb{B}$ we can write an

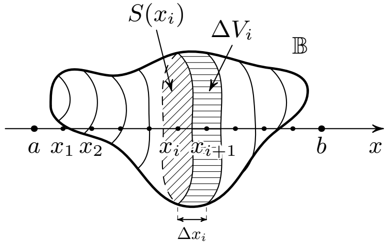


Fig. 31.

approximate formula

$$V \approx \sum_i S(x_i)\Delta x_i,$$

which *in the limit as* $n \to \infty$ *should give an exact formula*, and since the right-hand side of the formula is exactly the *integral sum* for the function $S(x)$ on the segment $[a, b]$, we obtain the *cross-section integral* (6).

Integral sums were composed and evaluated as early as Archimedes, but he considered individual sums for each particular problem he solved. The next advance in methods of finding areas and volumes was made only 18 centuries later by Kepler when formulating his second law (recall that it deals with areas of elliptic sectors swept out by radii vectors of planets) and deriving formulae for volumes of various solids of revolution. However, his reasoning was also very particular and rather far from introducing the general concept of integral.

Johannes Kepler (1571–1630) was a German astronomer and mathematician. From 1601 he was assistant to the famous Danish astronomer Tycho Brahe (1546–1601) and later succeeded him as “Imperial astronomer and mathematician” in Prague. When analysing the results of Brahe’s observations of planetary motions, he discovered the famous laws: the first law on the motion of planets along ellipses (1609, *Astronomia nova* [New Astronomy]), the second on the constancy of areal velocities in the motion of planets (1618, *Epitome Astronomiae Copernicanae* [Epitome of Copernican Astronomy]), the third on the proportionality of the squares of orbital periods of planets to the cubes of their mean distances from the Sun (1619, *Harmonice Mundi* [The Harmony of the World]); this grand treatise also contains a lot of Kepler’s fantastic ideas on the music of celestial motions, on the divine origin of geometry, etc.). However, the greatest contribution to mathematics was made by Kepler in the Austrian city of Linz, when in 1612 it was a fruitful season for grapes. On this occasion, Kepler devised a more general method for evaluating volumes than Archimedes’ and in 1615 published it, together with practical applications, in a wonderful work *Nova Stereometria doliorum vinariorum, in primis Austriaci, figurae omnium aptissimae* (New Stereometry of Wine Barrels, Predominantly Austrian, as Having the Most Advantageous Shape). In it he found volumes of 87 new (compared to those known to Greek mathematicians) solids, and at the same time proved that the “Austrian barrel” had the largest capacity for a given amount of material! Thus Kepler awakened geometry from its centuries-long hibernation, although he provoked the anger of “Archimedes’ defenders”.

Cavalieri came much closer to integral calculus. He considered sufficiently general plane and spatial figures, believing their areas (volumes) to be some *"aggregates of indivisible sections"* of parallel segments in the case of plane figures and parallel cross-sections for spatial figures. He himself was very vague about what "indivisibles" are, but he came to quite correct conclusions, which in modern notation can be written as integral formulae (3) and (6). Cavalieri formulated these results in the form of some "principles"; in simplified form, the most famous *Cavalieri's principle* reads: *if two solid figures are such that their cross-sections by planes parallel to a "fixed plane" have equal areas, then the two solids themselves have equal volumes.*

Of course, in integral calculus, Cavalieri's principle is a trivial consequence of the cross-section integral (7). However, the point is that Cavalieri was not only still far from integral calculus but even *did not actually consider* integral sums!

Bonaventura Cavalieri (1598–1647) was an Italian mathematician. His teacher Benedetto Castelli (1577–1644) recognised his student's outstanding ability and introduced him to his own teacher, Galileo Galilei. Cavalieri and Galileo began an active correspondence on the basis of their common interests in geometry, both intending to publish a work on the above-mentioned "indivisible quantities". In 1629 Galileo supported the candidacy of Cavalieri, who had already completed his work on the geometry of indivisibles, for a chair at the University of Bologna, describing him as "Archimedes' rival". In addition, Cavalieri received the patronage of the popes of the time, Urban VIII and Innocent X; to support the scientist financially, they also made him prior of a monastery. Like Kepler, Cavalieri was also an astrologer.

Despite the huge step forward towards integral calculus made by Cavalieri (the development of the "geometry of indivisibles"), the calculus itself was still quite far away. Among the mathematicians who contributed most to the progress in this field, we should mention, first of all, the famous French mathematicians Pierre Fermat (1601–1665) and Blaise Pascal (1623–1662), as well as Isaac Barrow (1630–1677), the direct predecessor of Newton and Leibniz (we have already mentioned all these scientists).

Example 1. Let us show how *Cavalieri's principle* can be used to derive the Archimedean formula for the volume of a sphere or, which is more convenient in this case, a *hemisphere* of radius R. Consider a cylinder with height R and base radius also R. First, let us inscribe in it a cone with its vertex at the centre O of the upper base of the cylinder and with its base coinciding with the lower base of the cylinder (Fig. 32). According to the "Democritus–Eudoxus" formula, the volume of the cone is one third the volume of the cylinder: $V_{\text{Con}} = \frac{1}{3} V_{\text{Cyl}}$.

Second, consider the "cup" obtained by cutting out of a cylinder a hemisphere with its base coinciding with the upper base of the cylinder, and consider its sections by planes α_x parallel to the base of the cylinder and distant from the upper base by x (see Fig. 32). Each such section

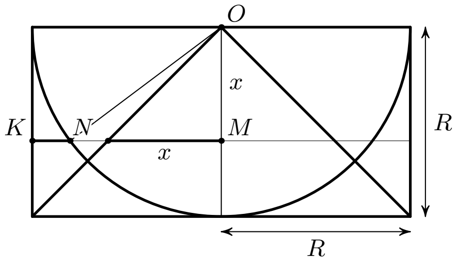


Fig. 32.

is a ring enclosed between two concentric circles of radii $MK = R$ and $MN = \sqrt{ON^2 - OM^2} = \sqrt{R^2 - x^2}$, so that its area is

$$S_{\text{Cup}}(x) = \pi R^2 - \pi(R^2 - x^2) = \pi x^2;$$

i.e., it is the same as the area $S_K(x)$ of the corresponding section of the cone (which is simply a circle of radius x).

According to Cavalieri's principle, the volumes of the cone and the cup are the same, and taking into account that the volume of the bowl is equal to the difference of volumes of the cylinder and the hemisphere, we obtain

$$V_{\text{Cup}} = V_{\text{Cyl}} - V_{\text{HSph}} = V_{\text{Con}} = \frac{1}{3}V_{\text{Cyl}} \quad \Rightarrow \quad V_{\text{HSph}} = \frac{2}{3}V_{\text{Cyl}};$$

this is exactly Archimedes' formula. □

Exercises, Problems, and Tasks to Chapter II

General remark. For finding antiderivatives, as well as for finding derivatives, one uses *formulae* and *rules* of integration, which are derived from formulae and rules of differentiation. We will *gradually* get acquainted with these formulae (for integrating functions of one or another "standard" type) and rules (for finding the antiderivatives of more complicated functions obtained from standard ones). At the starting level, for integration (finding antiderivatives, or solving differential equations of the form $y' = f(x)$), one should *"dare to guess"* what should be the form of the desired antiderivative, and then *check the guess by differentiation.* The word *find* in the conditions of the exercises so far means *give*, without assuming the use of any rules. Let us start with "visual integration".

1. Using the graphs of derivatives of the functions $y = F'(x) = f(x)$ given in Fig. 33(a–f), reconstruct an approximate form of the graphs of the functions $y = F(x)$ themselves (and show the correspondence between the "characteristic" points of the graphs of $y = f(x)$ and $y = F(x)$, placing the coordinate axes for the second graph exactly under the axes for the first).

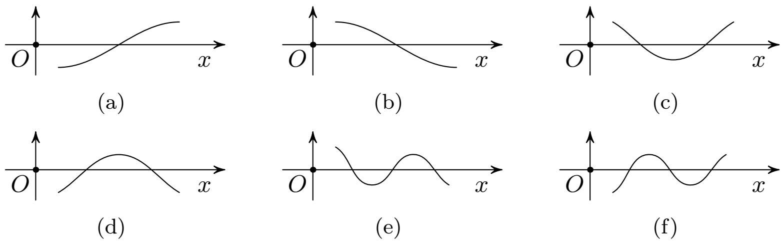


Fig. 33.

2. Find all solutions of the differential equations

(1) $y' = 2x$;
(2) $y' = 2x + 1$;
(3) $y' = x^2 + 1$;
(4) $y' = 6x^2 - 2x$;
(5) $y' = x^3$;
(6) $y' = 8x^3 + 6x^2$.

3. Find all antiderivatives of the following functions:[1]

(1) $f(x) = 10x^4$;
(2) $f(x) = 5x^4 - 6x^2$;
(3) $f(x) = 5x^4 + 6x^2 - 1$;
(4) $f(x) = 8x^3 + 6x$;
(5) $f(x) = x^5$;
(6) $f(x) = 6x^5 - 8x^3 + 6x$.

4. Let $F(x)$ be an antiderivative of a function $f(x)$ on the entire number axis $\mathbb{R}$. Which of the following statements are true?[2]

(1) If f has an even antiderivative F, then f itself is odd;
(2) If f is odd, then any its antiderivative F is even;
(3) If f has an odd antiderivative F, then f is even;
(4) If f is even, then any its antiderivative F is odd;
(5) If f is even, some of its antiderivatives F is odd.

Hint. Evennees/oddness conditions for a function h can be written as the identities $h(x) - h(-x) \equiv 0$ and $h(x) + h(-x) \equiv 0$.

5. Let $F(x)$ and $G(x)$ be antiderivatives of functions $f(x)$ and $g(x)$ (on a set $\mathbb{D}$).

(1) Give any antiderivative for the function $f_1(x) = \alpha f(x)$ ($\alpha \in \mathbb{R} \setminus \{0\}$ being a given number).

(2) Find all antiderivatives of the function $h_1(x) = 2f(x) + 1$.

(3) Find all antiderivatives of the function $h_2(x) = f(x) - 2$.

(4) Give any antiderivative for the function $s(x) = f(x) + g(x)$.

(5) Give any antiderivative for the function $d(x) = f(x) - g(x)$.

---

[1] In order to avoid diversity of formulations, starting from this problem we will more often use the *"language of antiderivatives"*. Of course, all such problems can also be formulated and solved in terms of the corresponding differential equations.

[2] If a statement is true, prove it; if not, give an appropriate *counterexample*.

(6) Is it true that *"an antiderivative of the sum of functions equals the sum of antiderivatives of the summands"*? How can we correct this statement so that to make it necessarily (i.e., always) true?

Comment. The statements *"the sum of antiderivatives of two functions is an antiderivative of the sum of these functions"* or *"an antiderivative of the sum of functions differs from the sum of antiderivatives of the summands by a constant"* are true.

6. Find all antiderivatives of the following functions:

(1) $f(x) = 2x(x+1)$;
(2) $f(x) = x(2x-1)$;
(3) $f(x) = 4x(x^2+1)$;
(4) $f(x) = x^2(x-3)$;
(5) $f(x) = (x+1)(x-1)$;
(6) $f(x) = (2x+1)(3x+2)$.

Comment. Surely, *the product of antiderivatives* of two functions in the general case *is by no means an antiderivative of the product* of these functions. As in the case of derivatives, when finding antiderivatives of a product it is *often* easiest to first remove the parentheses.

The next series of exercises shows what to do when integrating natural powers of binomials and similar expressions (in the cases where following the advice to remove parentheses is not really wanted!).

7. Find all antiderivatives of the functions

(1) $f(x) = (x+2)^5$;
(2) $f(x) = (1-x)^6$;
(3) $f(x) = (1-x)^7$;
(4) $f(x) = (2x+1)^9$;
(5) $f(x) = 2x(x^2+1)^5$;
(6) $f(x) = x^2(x^3-5)^5$.

Hint. Find[1] derivatives of functions of the form $F(x) = (ax+b)^n$ and $G(x) = (ax^m+b)^n$ $(m, n \in \mathbb{N})$.

8. Find all antiderivatives of the following functions:

(1) $y = (2x+1)(x^2+x+1)^5$;
(2) $y = 2x(x^2+x+1)^5$;
(3) $y = (3x^2-1)(x^3-x+5)^5$;
(4) $y = x^2(x^3-x+5)^5$;
(5) $y = x(ax^2+bx+c)^n$;
(6) $y = x^2(ax^3+bx^2+cx+d)^n$.

(Here $a, b, c, d \in \mathbb{R}$, $n \in \mathbb{N}$.)

Hint. Use the *method of reducing the problem to a previous one* (or *to previous ones*)!

To solve the next series of problems, you will need to recall the formulae for the derivatives of trigonometric functions and the basic trigonometric identities (i.e., formulae for trigonometric transformations).

[1] Following Leibniz (!); he surmised (though after Newton) that *integrable* functions are obtained by differentiation. This was the first step towards the creation of *Leibniz's calculus* (1675).

9. Find all antiderivatives of the functions

(1) $f(x) = -\cos x$;
(2) $f(x) = 5\cos x$;
(3) $f(x) = \cos 5x$;
(4) $f(x) = 5\cos\frac{x}{2}$;
(5) $f(x) = \sin x$;
(6) $f(x) = \sin\frac{x}{3}$;
(7) $f(x) = \sin 3x$;
(8) $f(x) = \frac{1}{2}\sin 2x$.

Remark. In this problem and the next ones, no transformations need to be made!

10. Find all antiderivatives of the functions

(1) $f(x) = \cos x + \sin 5x$;
(2) $f(x) = 5\cos x - \sin x + 2$;
(3) $f(x) = \cos 5 - 5x$;
(4) $f(x) = \cos 5x + x^5$;
(5) $f(x) = 5\cos x - \cos 5x$;
(6) $f(x) = 5\sin x - \sin 5x + 5$.

11. Let $F(x)$ be an antiderivative of $f(x)$ on the entire number axis $\mathbb{R}$. Which of the following statements are true?

(1) If a function f has a periodic antiderivative F, it is periodic itself.

(2) If a function f is periodic, its antiderivative F is also periodic.

(3) If a function F has period T, then its derivative f has period T too.

(4) If a function F has fundamental period T,[1] then its derivative f has fundamental period T too.

(5) If a function f has a periodic antiderivative F, then any antiderivative of f is periodic.

(6) If a periodic function f with fundamental period T has a periodic antiderivative F, then F has fundamental period T too.

(If a statement is true, prove it; if not, give a counterexample.)

Hint. The periodicity condition for a function h with period T can be written as the identity $h(x+T) - h(x) \equiv 0$.

12. Find all antiderivatives of the functions

(1) $f(x) = \cos^2 x$;
(2) $f(x) = \sin^2 x$;
(3) $f(x) = \cos^2 x - \sin^2 x$;
(4) $f(x) = \cos^2 x + \sin^2 x$;
(5) $f(x) = \sin x \cdot \cos x$;
(6) $f(x) = \sin^3 x$;
(7) $f(x) = \cos^3 x$;
(8) $f(x) = \cos^4 x$;
(9) $f(x) = \cos^4 x + \sin^4 x$;
(10) $f(x) = (\sin x + \cos x)^2$;
(11) $f(x) = (\sin x + \cos x)^3$;
(12) $f(x) = (\sin x + \cos x)^4$.

Hint. In this problem and in the next one, on the contrary, you will need to transform the integrand functions so that to reduce the problems to integration of functions of the form $a\cos\omega x$ and $a\sin\omega x$ (and, possibly, constants).

[1] Recall that the *fundamental* period is the *least positive period* (if exists).

13. Find all antiderivatives of the functions

(1) $f(x) = \sin x \cdot \cos 3x$;
(2) $f(x) = \cos x \cdot \cos 3x$;
(3) $f(x) = \sin x \cdot \sin 3x$;
(4) $f(x) = \cos 2x \cdot \cos 8x$;
(5) $f(x) = \cos x \cdot \sin^2 x$;
(6) $f(x) = \cos mx \cdot \cos nx$;
(7) $f(x) = \sin x \cdot \sin^2 x$;
(8) $f(x) = \sin mx \cdot \sin nx$.

(Here $m, n \in \mathbb{R} \setminus \{0\}$.)

14. Find all antiderivatives of the following functions:

(1) $f(x) = \dfrac{1}{\cos^2 x}$;
(2) $f(x) = \tan^2 x$;
(3) $f(x) = \dfrac{1}{\sin^2 x}$;
(4) $f(x) = \cot^2 x$;
(5) $f(x) = \dfrac{1}{\sin^2 2x}$;
(6) $f(x) = \tan^2 2x$;
(7) $f(x) = \dfrac{1}{\cos^2 \frac{x}{2}}$;
(8) $f(x) = \cot^2 \frac{x}{2}$.

Remark. Keep in mind that antiderivatives on a union of intervals differ not by a constant but by a *"pseudoconstant"*, a *multivalued* constant $\widehat{C}$ (which is "individual" for each of the intervals).

15. Find all antiderivatives of the following functions:

(1) $f(x) = x \cos(x^2)$;
(2) $f(x) = x \sin(x^2)$;
(3) $f(x) = \dfrac{x}{\cos^2(x^2)}$;
(4) $f(x) = \dfrac{2x}{\sin^2(x^2)}$;
(5) $f(x) = x \tan^2(x^2)$;
(6) $f(x) = x^2 \cos(x^3)$.

Hint. Guess an antiderivative. (*Leading question*: what is the derivative of a function of the form $\varphi(x^m)$, where φ is a known function and $m \in \mathbb{N}$?)

* * *

Consider the following sets of functions:

(1) $\mathbb{R}[x]$, the set of all polynomial functions, i.e., function given by *algebraic polynomials*

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

(of an arbitrary degree $n \geq 0$);

(2) $\mathbb{R}_T[x]$, the set of all functions given by *trigonometric polynomials*

$$T(x) = a_0 + \sum_{m=1}^{n} (a_m \cos mx + b_m \sin mx) \quad (n \geq 0 \text{ is arbitrary}).$$

(In both cases the coefficients a_k and b_k are arbitrary real numbers.)

16. Let $\mathbb{M}$ be one of the sets of functions defined above. Prove that if $f(x)$ and $g(x)$ belong to $\mathbb{M}$, then

(1) Their sum $f(x) + g(x)$ belongs to $\mathbb{M}$;
(2) Any *linear combination* $\alpha f(x) + \beta g(x)$ belongs to $\mathbb{M}$;
(3) Their product $f(x) g(x)$ belongs to $\mathbb{M}$.

17. Let $\mathbb{M}$ be one of the sets of functions defined above. Is it true that if a function $f(x)$ belongs to $\mathbb{M}$, then

(1) Its derivative $f'(x)$ belongs to $\mathbb{M}$?

(2) Any of its antiderivatives $F(x)$ belongs to $\mathbb{M}$?

(For each case, give an answer and justify it.)

18. For an arbitrary function $f(x) \in \mathbb{M}$, in each of the above cases, given a representation of $f(x)$ (according to the above formulae) find a representation (general form) for:

(1) the derivative $f'(x)$;

(2) an antiderivative $F(x)$ for $f(x)$.

What additional condition should be imposed on a function $f(x) \in \mathbb{M}$ so that its antiderivative also belongs to $\mathbb{M}$?

19. Explicitly write the following integral sums:

(1) $S_n(x) = S_n(t)_0^x$;

(2) $S_n(x) = S_n(t^2)_0^x$;

(3) $S_n(x) = S_n(t^3)_0^x$;

(4) $S_n(x) = S_n(t^m)_0^x \quad (m \in \mathbb{N})$;

(5) $S_n(x) = S_n(\sin t)_0^x$;

(6) $S_n = S_n(\sin t)_0^\pi$;

(7) $S_n(x) = S_n(\cos t)_0^x$;

(8) $S_n = S_n(\cos t)_0^\pi$;

(9) $S_n(x) = S_n\left(\frac{1}{t}\right)_1^x$.

(10) $S_n(x) = S_n\left(\frac{1}{t^2}\right)_1^x$.

Summation of sequences and difference calculus

In parallel with the development of calculus, its analogue for sequences was gradually developed in mathematics, which was called the *calculus of finite differences*. The role of the derivative for a sequence $f(n) = f_n$, which for the sake of convenience we will consider “two-sided”, i.e., defined not only for natural but also *for all integer* values of n, is played by the (*forward*) *difference* of the sequence f_n, i.e., a *new sequence* $g_n = g(n)$ denoted as

$$\Delta f(n) = \Delta f_n$$

and obtained as the *difference*:

$$g(n) = \Delta f(n) = f(n+1) - f(n) \quad (\text{or} \quad g_n = \Delta f_n = f_{n+1} - f_n).$$

For example,

$$\Delta(\text{const}) \equiv 0, \quad \Delta(kn + b) \equiv k, \quad \Delta(n^2) = 2n + 1, \quad \Delta(q^n) = (q-1)q^n$$

(explain, i.e., check!).

20. Find the forward differences of the following sequences:

(1) $f_n = an^2 + bn + c$;

(2) $f_n = n^3$;

(3) $f_n = n^4$;

(4) $f_n = 3^n - 17 \cdot 2^n$;

(5) $f_n = 2^{-n}$;

(6) $f_n = 2^{kn}$;

(7) $f_n = n2^n$;
(8) $f_n = \sin \alpha n$;
(9) $f_n = \cos \alpha n$;
(10) $f_n = \frac{1}{n}$.

Similarly to the antiderivative for a given function, one can define an *antidifference sequence* for a given sequence f_n: this is a sequence F_n such that

$$\forall n \in \mathbb{Z} \quad \Delta F_n = f_n.$$

21. Prove the following statements.

(1) If F_n and G_n are antidifference sequences for the same sequence f_n, then they differ by a constant: $G_n = F_n + C$;

(2) Any antidifference sequence of an arithmetic progression $f_n = f_0 + dn$ is a quadratic sequence $F_n = an^2 + bn + c$;

(3) Any antidifference sequence of a geometric progression $f_n = f_0 \cdot q^n$ is of the form $F_n = Aq^n + C$ (said to be a shifted geometric progression);

(4) If F_n is an antidifference sequence for a sequence f_n, then for any $n \in \mathbb{N}$ the sum $S_n = f_0 + f_1 + \ldots + f_{n-1}$ equals the difference $F_n - F_0$.

The last statement yields a basis for a popular method of summation (i.e., computation of sums) for sequences. For example, to find the sum of the first n terms of a geometric progression $f_n = aq^n$, i.e., $S_n = a + aq + \ldots + aq^{n-1}$, we find a sequence of the form $F_n = Aq^n$ such that $\Delta F_n = f_n$. Since $\Delta F_n = A(q-1)q^n$, we must take the coefficient A to be $\frac{a}{q-1}$. Thus,

$$a + aq + \ldots + aq^{n-1} = F_n - F_0 = Aq^n - Aq^0 = \frac{a}{q-1}(q^n - 1),$$

and in this case we have arrived at a well-known formula.

22. Using the described method, find the sums $S_n = \sum\limits_{k=0}^{n-1} f_k$ for the following sequences:

(1) $f_n = n$;
(2) $f_n = n^2$;
(3) $f_n = n^3$;
(4) $f_n = \cos n\alpha$;
(5) $f_n = \sin n\alpha$;
(6) $f_n = nq^n$.

23. Using formulae from the above problem, find limits of the integral sums from Exercise 19 (except for (4), (9), and (10)).

Variable areas

In the exercises below, by $s(x)$ we denote the area $S(f)_{x_0}^{x}$ of a curvilinear trapezium bounded by the Ox axis, the graph $z = f(x)$, and the vertical lines at points $(x_0; 0)$ and $(x; 0)$ of the Ox axis (recall that on the intervals where the function is negative or in the case $x < x_0$, the area is taken with the minus sign).

24. Draw the graph of the function $y = s(x)$ for given functions f and points x_0:

(1) $f(x) = 1$, $x_0 = 1$;
(2) $f(x) = -1$, $x_0 = 1$;
(3) $f(x) = x$, $x_0 = 0$;
(4) $f(x) = x$, $x_0 = 1$;
(5) $f(x) = x$, $x_0 = -1$;
(6) $f(x) = 1 - x$, $x_0 = 0$;
(7) $f(x) = 1 - x$, $x_0 = 1$;
(8) $f(x) = |x|$, $x_0 = 0$;
(9) $f(x) = |x|$, $x_0 = -1$;
(10) $f(x) = 1 - |x|$, $x_0 = 0$.

25. Find a function $f(x)$ and a point x_0 given that

(1) $s(x) = x + 1$;
(2) $s(x) = x^2$;
(3) $s(x) = x^2 - 1$;
(4) $s(x) = (x - 1)^2$;
(5) $s(x) = (x + 1)^2$;
(6) $s(x) = x^2 + 1$.

Is the formulated problem solvable for any (differentiable) function $s(x) = F(x)$? Are the function f and "reference point" x_0 uniquely determined?

26. Draw the graph of the function $y = s(x)$ for given functions f and points x_0:

(1) $f(x) = \theta(x)$,[1] $x_0 = -1$;
(2) $f(x) = \theta(x)$, $x_0 = 0$;
(3) $f(x) = \theta(x)$, $x_0 = 1$;
(4) $f(x) = \operatorname{sgn}(x)$, $x_0 = 0$;
(5) $f(x) = \operatorname{sgn}(x)$, $x_0 = -1$;
(6) $f(x) = \operatorname{sgn}(x)$, $x_0 = 1$.

(In this case, a curvilinear trapezium is understood as a figure composed of all vertical segments bounded by the x-axis and and the graph of the function that lie between the vertical lines x_0 and x.)

27. Find all antiderivatives for the functions

(1) $f(x) = |x|$;
(2) $f(x) = |x| - 1$;
(3) $f(x) = |x - 1|$;
(4) $f(x) = 1 - |x|$;
(5) $f(x) = \sin|x|$;
(6) $f(x) = |\sin x|$, $x \in [-2\pi, 2\pi]$.

28. For the function f, find antiderivatives $F(x)$ such that $F(0) = 0$:

(1) $f(x) = |1 - |x||$;
(2) $f(x) = x + |x|$;
(3) $f(x) = x - 3|x|$;
(4) $f(x) = \begin{cases} 1 - x^2, & |x| \le 1, \\ 1 - |x|, & |x| > 1; \end{cases}$
(5) $f(x) = \begin{cases} 1 - x^2, & |x| \le 1, \\ |x| - 1, & |x| > 1; \end{cases}$
(6) $f(x) = \begin{cases} 1, & x < 0, \\ 1 + x, & 0 \le x \le 1, \\ 2x, & x > 1. \end{cases}$

Draw graphs of $y = f(x)$ and $y = F(x)$.

[1] This is the notation for the *Heaviside theta function*: it is equal to 0 for $x < 0$ and to 1 for $x \ge 0$.

A *generalised antiderivative* on an interval $\mathbb{I}$ for a function f having a *finite* or *discrete*[1] set of discontinuity points in this interval (the function f may as well be undefined at these points) is a function $F(x)$ *defined and continuous* on $\mathbb{I}$ and such that $\exists F'(x) = f(x)$ *on all intervals of continuity of* f.

29. Find all generalised antiderivatives of the following functions (on the entire number axis):

(1) $f(x) = \theta(x)$;
(2) $f(x) = \operatorname{sgn} x$;
(3) $f(x) = x + \theta(x)$;
(4) $f(x) = x + \operatorname{sgn} x$;
(5) $f(x) = \operatorname{sgn}(x^2 - 1)$;
(6) $f(x) = \operatorname{sgn}(1 - x^2)$.

30. Let M be the set of discontinuity points of a function f defined on the entire number axis. Assume that M is a *finite* or *discrete* set and that f *has a generalised antiderivative* F (on the entire axis). Give (justified) answers to the following questions (take you time!).

(1) Can it be that f has a single discontinuity point, say $x = 0$? Or $M = \{-1\}$? $M = \{17\}$?

(2) Can it be that f has *two* discontinuity points, say $x = 0, 1$? Or $M = \{\pm 1\}$? $M = \{0, 200\}$?

(3) Can it be that f has *three* discontinuity points, say $0, 1, 2$? Or $M = \{0, 2, 1000\}$?

(4) Can it be that f has infinitely many discontinuity points? Give examples.

(5) Can it be that F has a single discontinuity point, say $x = 0$? Or $x = 2$? Or two discontinuity points?

31. Let $F(x)$ and $G(x)$ be, respectively, generalised antiderivatives of the functions $f(x) = [x]$ (*floor function*, or *integral part* of x) and $g(x) = \{x\}$ (*fractional part* x) such that $F(0) = G(0) = 0$.

(1) Are F and G uniquely determined by these conditions?
(2) Evaluate $F(1)$, $F(2)$, $F(10)$, $F(-2)$;
(3) Evaluate $G(1)$, $G(2)$, $G(4)$, $G(-8)$;
(4) Evaluate $F(100)$, $G(100)$, $F(100) \pm G(100)$, $F(200) \pm G(200)$.

Directional fields and solutions of non-autonomous differential equations

32. Let us denote the function $u(z) = \frac{|z|}{z}$ by $\operatorname{Sgn} z$ (this function differs from the *sign* function $\operatorname{sgn} z$ only by that it is *not defined* at $z = 0$). Construct the directional field and integral curves of the following differential equations:

[1] A number set is said to be *discrete* if each *bounded* interval of the number axis contains at most finitely many points of this set (from Latin *discretus*, 'separate', consisting of separate parts; an antonym for *continuous*).

(1) $y' = \operatorname{Sgn} x$;
(2) $y' = \operatorname{Sgn} x + 1$;
(3) $y' = \operatorname{Sgn}(x-1)$;
(4) $y' = -\operatorname{Sgn} x$;
(5) $y' = 2 - \operatorname{Sgn} x$;
(6) $y' = \operatorname{Sgn} y$;
(7) $y' = \operatorname{Sgn} y - 1$;
(8) $y' = \operatorname{Sgn}(y+1)$;
(9) $y' = -\operatorname{Sgn} y$;
(10) $y' = \operatorname{Sgn}(x-y)$;
(11) $y' = \operatorname{Sgn}(x+y)$;
(12) $y' = \operatorname{Sgn}(y-2x)$;
(13) $y' = \operatorname{Sgn}(xy)$;
(14) $y' = \operatorname{Sgn}(-xy)$;
(15) $y' = \operatorname{Sgn}(xy-1)$;
(16) $y' = \operatorname{Sgn}(1-xy)$;
(17) $y' = \operatorname{Sgn}(x^2-y)$;
(18) $y' = \operatorname{Sgn}(x^2-y^2)$;
(19) $y' = \operatorname{Sgn}(x^2+y^2-4)$;
(20) $y' = \operatorname{Sgn}(4-x^2-y^2)$;
(21) $y' = \operatorname{Sgn}(\sin x)$;
(22) $y' = \operatorname{Sgn}(\sin x \cdot \sin y)$.

Remark. The directional fields and the integral curves of the equation $y' = F(x,y)$ should be considered only in the regions of the Oxy plane where the right-hand side of the equation *is well defined*.

33. Sketch directional fields and integral curves of the differential equations

(1) $y' = x$;
(2) $y' = -x$;
(3) $y' = x-2$;
(4) $y' = -x+1$;
(5) $y' = y$;
(6) $y' = -y$;
(7) $y' = y-1$;
(8) $y' = -y+2$;
(9) $y' = x+y$;
(10) $y' = x-y$;
(11) $y' = y-x$;
(12) $y' = -y-x$.

34. Prove analytically (by explicitly finding the derivatives) that

(1) If a function $y = \varphi(x)$ satisfies a differential equation of the form $y' = f(x)$, then *any* function $y_C = \varphi(x) + C$ satisfies this equation too;

(2) If a function $y = \varphi(x)$ satisfies a differential equation of the form $y' = g(y)$, then *any* function $y_c = \varphi(y+c)$ satisfies this equation too;

(3) If a function $y = \varphi(x)$ satisfies a differential equation of the form $y' = k(x)y$, then *any* function $y_A = A\varphi(x)$ satisfies this equation too;

(4) If a function $y = \varphi(x)$ satisfies a differential equation of the form $y' = h\left(\frac{y}{x}\right)$, then *any* function $y_A = A\varphi\left(\frac{x}{A}\right)$ satisfies this equation too.

Isoclines and directional fields

35. Construct isoclines for the differential equations from Exercise 33. Use the constructed isoclines to draw directional fields of these equations.

36. Can an integral curve of a differential equation of the form $y' = F(x,y)$

(1) *coincide* with any of its isoclines?

(2) *be a part* of any of its isoclines?

37. Can all the family of integral curves of a differential equation of the form $y' = F(x,y)$ *coincide* with the family of its isoclines or their parts?

38. Find the form of the *family of isoclines* for the following differential equations ($u(z)$ is always a given function in one variable):

(1) $y' = u(x)$;
(2) $y' = u(y)$;
(3) $y' = u(x-y)$;
(4) $y' = u(x+y)$;
(5) $y' = u(y-2x)$;
(6) $y' = u(y-x^2)$;
(7) $y' = u\left(\frac{y}{x}\right)$;
(8) $y' = u\left(\frac{x}{y}\right)$;
(9) $y' = u(xy)$;
(10) $y' = u(x^2+y^2)$.

39. Prove that if the *family of isoclines* of a differential equation of the form $y' = F(x,y)$ is *invariant* under some transformation

$$L\colon (x;y) \mapsto (x_1;y_1)$$

of the coordinate plane Oxy (i.e., maps *to itself* under this *transformation*, e.g., under translation, rotation, symmetry, dilation,...), then the *family of integral curves* also maps to itself under this transformation, in the sense that each integral curve maps to an integral curve (generally speaking, into *another one*) or into several "pieces" of integral curves (this is what happens if the *graph* of the solution function $y = y(x)$ maps under L to a curve *that is not a graph*).

40. Sketch directional fields and integral curves of the differential equations

(1) $y' = x^2$;
(2) $y' = x^2+1$;
(3) $y' = x^2-1$;
(4) $y' = x^3$;
(5) $y' = -x^2$;
(6) $y' = -x^2+1$;
(7) $y' = -x^2-1$;
(8) $y' = -x^3$;
(9) $y' = |x|$;
(10) $y' = -|x|$;
(11) $y' = 1-|x|$;
(12) $y' = \sqrt{|x|}$;
(13) $y' = \sqrt[3]{x}$;
(14) $y' = \frac{1}{x}$;
(15) $y' = \frac{1}{x^2}$;
(16) $y' = -\frac{1}{x}$;
(17) $y' = -\frac{1}{x^2}$;
(18) $y' = \frac{1}{x-2}$;
(19) $y' = \frac{1}{2-x}$;
(20) $y' = \sin x$;
(21) $y' = -\sin x$;
(22) $y' = \tan x$.

41. Sketch directional fields and integral curves of the differential equations

(1) $y' = y^2$;
(2) $y' = -y^2$;
(3) $y' = y^2+1$;
(4) $y' = y^2-1$;
(5) $y' = y^2-y$;
(6) $y' = 1-y^2$;
(7) $y' = y^3$;
(8) $y' = -y^3$;
(9) $y' = \frac{1}{y}$;
(10) $y' = -\frac{1}{y}$;
(11) $y' = \frac{1}{y^2}$;
(12) $y' = \sqrt{|y|}$;
(13) $y' = \sqrt[3]{y}$;
(14) $y' = \cos y$;
(15) $y' = \cos^2 y$;
(16) $y' = \tan y$.

42. Sketch directional fields and integral curves of the differential equations

(1) $y' = \frac{y}{x}$;
(2) $y' = \frac{2y}{x}$;
(3) $y' = \frac{y}{2x}$;

(4) $y' = -\frac{y}{x}$;
(5) $y' = -\frac{2y}{x}$;
(6) $y' = -\frac{y}{2x}$;
(7) $y' = \frac{x}{y}$;
(8) $y' = \frac{2x}{y}$;
(9) $y' = \frac{x}{2y}$;
(10) $y' = -\frac{x}{y}$;
(11) $y' = -\frac{2x}{y}$;
(12) $y' = -\frac{x}{2y}$;
(13) $y' = \frac{y+1}{x-2}$;
(14) $y' = \frac{x+1}{y-2}$;
(15) $y' = \frac{y^2}{x}$;
(16) $y' = 2xy$.

First integrals of non-autonomous differential equations

Definition 3. A function $Y(x,y)$ which is *constant* on each integral curve $y=y(x)$, $x \in \mathbb{I}$, of a differential equation $y' = F(x,y)$, i.e., such that

$$\forall x \in \mathbb{I} \quad Y(x, y(x)) \equiv C \qquad (*)$$

(the value of C is, in general, different for each solution $y(x)$), is called a *first integral* of this differential equation.

43. Prove that the functions below are *first integrals* for the given differential equations:

(1) $Y(x,y) = x^2 - y, \quad y' = 2x$;
(2) $Y(x,y) = x + \frac{1}{y}, \quad y' = y^2$;
(3) $Y(x,y) = xy, \quad y' = -\frac{y}{x}$;
(4) $Y(x,y) = \frac{y}{x}, \quad y' = \frac{y}{x}$;
(5) $Y(x,y) = x^2 + y^2, \quad y' = -\frac{x}{y}$;
(6) $Y(x,y) = x^2 - y^2, \quad y' = \frac{x}{y}$.

Hint. Check that the derivative of the left-hand side of equation $(*)$ is *identically zero* for any solution of the corresponding differential equation.

44. Let a function $Y(x,y)$ be a first integral for the differential equation $y' = F(x,y)$. Prove that any integral curve of this equation entirely lies on some *level curve* of Y, i.e., is contained in the set

$$\mathbb{Y}_C = \{(x;y) \mid Y(x,y) = C\}.$$

Using this, find formulae for solutions $y = y(x)$ of the differential equations from Exercise 43 and draw *families* of their graphs.

Preliminary remark. One and the same function $y = \varphi(x)$ satisfies *many* (*infinitely many*!) differential equations. For example, the function $y = x^2$ on the positive semi-axis $x > 0$ satisfies the differential equations

$$y' = 2x, \qquad y' = x + \frac{y}{x}, \qquad y' = 17x - \frac{15y}{x}, \qquad y' = \frac{2y}{x}, \qquad y' = 2\sqrt{y}$$

etc. (check!).

45. For a given function $y=\varphi(x)$ considered (a) on the positive semi-axis only, (b) on its entire domain, give (find, think of, "guess") a differential equation of the form $y'=g(y)$ which this function would satisfy:

(1) $y=x^3$; (3) $y=\frac{1}{x}$; (5) $y=\frac{1}{x^3}$; (7) $y=\frac{1}{\sqrt[3]{x}}$;

(2) $y=\sqrt{x}$; (4) $y=-\frac{1}{x}$; (6) $y=\frac{1}{\sqrt{x}}$; (8) $y=\tan x$.

What other solutions to the obtained equations can be given?

46. For a given function $y=\varphi(x)$ considered (a) on the positive semi-axis only, (b) on $\mathbb{R}\setminus\{0\}$, give a differential equation of the form $y'=k(x)y$ which this function would satisfy:

(1) $y=x^3$; (3) $y=x^5$; (5) $y=\frac{1}{x}$; (7) $y=\frac{1}{x^2}$;

(2) $y=x^4$; (4) $y=\sqrt[3]{x}$; (6) $y=-\frac{1}{x}$; (8) $y=\frac{1}{x^3}$.

What other solutions to the obtained equations can be given?

47. Draw the *family of level curves* for the given functions $Z(x,y)$ (they are defined by equations $Z(x,y)=C$); write (find) a differential equation of the form $y'=F(x,y)$ (or $\mathcal{F}(x,y,y')=0$)[1] which is satisfied by the curves of this family (more precisely, those "parts" of them that are graphs $y=y(x)$ of differentiable functions):

(1) $x-y=C$;
(2) $x^2+y=C$;
(3) $x^n-y=C$, $n\in\mathbb{N}$;
(4) $x-y^2=C$;
(5) $x-y^n=C$, $n\in\mathbb{N}$;
(6) $x+y^n=C$, $n\in\mathbb{N}$;
(7) $xy=C$;
(8) $x^2y=C$;
(9) $x^ny=C$, $n\in\mathbb{N}$;
(10) $xy^2=C$;
(11) $xy^n=C$, $n\in\mathbb{N}$;
(12) $x^2y^3=C$;
(13) $x^2-y^3=C$;
(14) $x^2+y^2=C$;
(15) $4x^2+y^2=C$;
(16) $x^2+4y^2=C$;
(17) $x^2-y^2=C$;
(18) $2x^2-y^2=C$.

Hint. Notice that the left-hand sides $Z(x,y)$ of the equations of these families are *first integrals* of the desired differential equations.

Difference calculus: linear difference equations

Let us turn to differences once again. For a given sequence f_n, besides the "first" difference Δf_n we can define the *second difference* $\Delta^2 f_n$ as the *difference of the first difference*,

$$(\Delta^2 f)_n=\Delta(\Delta f)_n,$$

and then the third, fourth, etc. differences.

[1] Such equations are called *first-order differential equations* not resolved *with respect to the derivative* y'.

48. Write the second and third differences of a sequence f_n directly through the elements of the sequence.

49. Find the second and third differences for the following sequences:

(1) $f_n = n^2$; (3) $f_n = n^4$; (5) $f_n = 3^n$;
(2) $f_n = n^3$; (4) $f_n = 2^n$; (6) $f_n = \sin n\alpha$.

An analogue of differential equations for sequences are *difference equations*, which involve *differences of unknown sequences* rather than derivatives of unknown functions.

If an unknown sequence y_n and all of its differences Δy_n, $\Delta^2 y_n$, ... enter a difference equation *linearly*, then we speak of a *linear difference equation.*

50. Find all solutions of the difference equations

(1) $\Delta y_n = ky_n$; (3) $\Delta^2 y_n = 0$;
(2) $\Delta y_n = ky_n + b$; (4) $\Delta^3 y_n = 0$.

51. Prove that any *second-order homogeneous linear difference equation* with constant coefficients

$$\Delta^2 y_n + p\Delta y_n + qy_n = 0$$

can be represented as a recurrence relation of the form

$$y_{n+2} = ay_{n+1} + by_n, \tag{$*$}$$

where $a, b \in \mathbb{R}$ are constants. Is the converse true?

The theory of linear difference equations is in essence quite analogous to (and simpler than!) the theory of linear differential equations, discussed in the next chapter. We will not expose this beautiful analogy so far, confining ourselves to one example, which we expand into a small series of problems.

52. Prove that if sequences u_n and v_n are solutions of the difference equation $(*)$, then any *linear combination* $y_n = \alpha u_n + \beta v_n$ ($\alpha, \beta \in \mathbb{R}$ being arbitrary constants) is also a solution of equation $(*)$.

53. Prove that for any "initial conditions", i.e., for any given initial elements $y_0 = A$ and $y_1 = B$, a sequence y_n, $n \geq 0$, which is a solution of the difference equation $(*)$ exists and is unique. Is this true if we consider solutions y_n for all integer values of n?

54. Prove that if a linear difference equation $(*)$ has a solution that can be written as a geometric progression $q_n = \lambda^n$, then its common ratio λ is a root of the *characteristic equation*

$$\lambda^2 - a\lambda - b = 0 \tag{$**$}$$

(a and b being the coefficients of equation $(*)$).

55. Prove that if the discriminant of the characteristic equation $(**)$ is positive and λ_1, λ_2 are its roots, then any solution of the difference equation $(*)$

can be represented as

$$y_n = \alpha\lambda_1^n + \beta\lambda_2^n,$$

where the constants α and β are *uniquely* determined by the initial conditions $y_0 = A$, $y_1 = B$.

56. Give the general form of solutions of the *Fibonacci difference equation*:

$$y_{n+2} = y_{n+1} + y_n.$$

57. Find an analytical formula for the n-th *Fibonacci number* f_n: $f_0 = 0$, $f_1 = 1$, $f_2 = 1$, $f_3 = 2$, $f_4 = 3$, $f_5 = 5$, $\ldots$; $f_n = f_{n-1} + f_{n-2}$.

Comment. The characteristic equation for the Fibonacci equation is $\lambda^2 - \lambda - 1 = 0$, its roots are

$$\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2} = \left[\begin{matrix}\tau \\ \tau_1\end{matrix}\right.,$$

where the constant $\tau = \dfrac{1+\sqrt{5}}{2}$ is called the *gold ratio*, and the second root is $\tau_1 = 1 - \tau$ (e.g., by Vieta's formulae). The general solution of the Fibonacci equation is given by

$$y_n = \alpha\tau^n + \beta(1-\tau)^n.$$

For the Fibonacci numbers, we should require the initial conditions

$$y_0 = f_0 = \alpha + \beta = 0 \quad \Rightarrow \quad \beta = -\alpha;$$

hence,

$$y_1 = f_1 = \alpha\tau - \alpha(1-\tau) = \alpha(2\tau - 1) = \alpha \cdot \sqrt{5} = 1 \quad \Rightarrow \quad \alpha = \frac{1}{\sqrt{5}}.$$

Thus, it turns out that the number f_n of pairs of rabbits in the "Fibonacci model" after n months (we omit the description of this population growth model) can be found with the help of two geometric progressions, by *Binet's formula*:

$$f_n = \alpha(\tau^n - (1-\tau)^n) = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}.$$

(An amazing formula!)

58. Prove that the limit of the ratio of consecutive Fibonacci numbers equals the "golden ratio" τ.

This opens up a separate area of a kind of *mathematical aesthetics*, from proportions of the Parthenon and Botticelli to the structure of pineapple fruit or sunflower baskets. However, the discussion of the "divine proportion", as Leonardo da Vinci and Fra Luca Pacioli called the ratio $\tau : 1$ (in the book *De Divina Proportione*, 1509), is also beyond the scope of this course.

Leonardo of Pisa, alias Fibonacci (Italian *filius Bonacci*, 'son of good nature'), was a wonderful Italian mathematician, merchant, traveller, and writer (1180–1240). He was born in Pisa, was educated by the Arabs in Algeria, and wrote some of the most fundamental treatises on mathematics at that time, in one of which, *Liber Abaci* (The Book of Abacus; actually, not about the *abacus* itself but about all kinds of calculations), he considered the problem of rabbit reproduction (1202).

The relationship between the Fibonacci numbers and the golden ratio was pointed out by Johannes Kepler in his brilliant essay *Strena seu de Nive Sexangula* (On the Hexagonal Snowflake, 1611). Finally, the French mathematician, professor at École Polytechnique, algebraist, astronomer, and analyst Jacques Philippe Marie Binet (1786–1856) obtained the above formula in his study of general linear difference equations (1843). Fibonacci numbers were also dealt with by Lagrange and many others, and to the present day they find surprisingly diverse and very profound applications...

59. Draw curvilinear trapezia that correspond to the integrals and evaluate these integrals *without using* Barrow's formula (using geometric considerations only):

(1) $\int_{-1}^{2} x\,dx$; (3) $\int_{1}^{3} (x-2)\,dx$; (5) $\int_{1}^{3} (2-x)\,dx$;

(2) $\int_{1}^{2} x\,dx$; (4) $\int_{3}^{1} (x-2)\,dx$; (6) $\int_{-1}^{-3} (2-x)\,dx$.

Hint. Each trapezium consists of one or more "pieces", and the integral is the sum of their areas *but taken with appropriate signs.*

60. For each of the integrals in the preceding exercise, check the obtained "geometric" answer using Barrow's formula.

61. Evaluate the integrals and draw the corresponding curvilinear trapezia (indicating the signs of the areas of the components):

(1) $\int_{-2}^{2} x^2\,dx$; (3) $\int_{-2}^{3} x^2\,dx$; (5) $\int_{-2}^{3} x^3\,dx$;

(2) $\int_{-2}^{2} (x^2-1)\,dx$; (4) $\int_{0}^{3} (x^2-2x)\,dx$; (6) $\int_{-a}^{a} x^3\,dx$.

62. Draw the region cut off from quadrant I ($x \ge 0$, $y \ge 0$) of the Oxy coordinate plane by the graph of the given function $y = f(x)$ and evaluate its area (writing it as an integral):

(1) $y = 1 - x^\alpha \quad (\alpha > 0)$; (4) $y = (a-x)^m \quad (a, m > 0)$;
(2) $y = (1-x)^\alpha \quad (\alpha > 0)$; (5) $y = \cos \alpha x \quad (\alpha > 0)$;
(3) $y = a^m - x^m \quad (a, m > 0)$; (6) $y = \cos^2 \alpha x \quad (\alpha > 0)$.

63. Draw the region bounded by the given graphs (for $x \ge 0$) and evaluate its area (writing it with the use of integrals):

(1) $y = x^\alpha$, $y = x \quad (\alpha > 0)$;
(2) $y = x^\alpha$, $y = x^\beta \quad (\alpha > \beta > 0)$;
(3) $y = 1 - x^\alpha$, $y = 1 - x \quad (\alpha > 0)$;

(4) $y = (1-x)^{\alpha}$, $y = 1-x$ $(\alpha > 0)$;
(5) $y = \pi \sin \alpha x$, $y = 2\alpha x$ $(\alpha > 0)$;
(6) $y = \pi \cos \alpha x$, $y = \pi - 2\alpha x$ $(\alpha > 0)$.

General remark. When finding areas of regions bounded by some curves in the coordinate plane, these regions are represented as *unions* or *differences* of some *curvilinear trapezia*, and the desired areas are expressed as sums and differences of the corresponding integrals.

64. Draw the region bounded by the given curves and evaluate its area:
(1) $y = x^2$, $y = c^2$ $(c > 0)$;
(2) $y = ax^2$, $y = ac^2$ $(c > 0)$;
(3) $y = ax^2$, $y = kx$ $(a, k > 0)$;
(4) $y = h - x^2$, $y = 0$ $(h > 0)$;
(5) $y = x^2 - ax$, $y = 0$ $(a > 0)$;
(6) $y = (x-\alpha)(x-\beta)$, $y = 0$ $(\alpha < \beta)$.

Definition 4. A *parabolic segment* is a figure bounded by a parabola $y = p_2(x) = ax^2 + bx + c$ and its *chord*, a straight line segment AB connecting two points $A(x_A; y_A)$ and $B(x_B; y_B)$ of the parabola.

The *horizontal* distance between A and B, i.e., $d = |x_B - x_A|$, is called the *width* of the segment, and the length of the segment MP of the vertical line $x = \frac{x_A + x_B}{2}$ between the center M of the chord AB and the corresponding point P of the parabola is called the *height* of the segment.

The *parallelogram* bounded by the lines AB, $x = x_A$, $x = x_B$, and the tangent to the parabola parallel to the chord AB is said to be *circumscribed* about the segment ABP.

65. Prove that the area of the parabolic segment ABP equals
(1) $\frac{2}{3}$ of the product of the segment width by its height;
(2) $\frac{2}{3}$ of the area of its circumscribed parallelogram.

Finding antiderivatives of elementary functions (continued)

Previously, we gave almost no consideration to antiderivatives of *irrational algebraic functions*, i.e., of functions that can be obtained from the function x and various constants not only by means of four arithmetic operations (addition, subtraction, multiplication, and division), but also by means of the "algebraic operation" of *extracting the* (arithmetic) *root of arbitrary natural degree* $n \geq 2$ and by forming arbitrary compositions of already existing functions (and reapplying all these operations to the obtained functions any number of times). In the simplest cases, the above-mentioned methods are quite sufficient for finding antiderivatives of irrational algebraic functions.

66. Find all antiderivatives of the following functions:

(1) $f(x)=x(\sqrt{x}+1)$;
(2) $f(x)=\sqrt{x}(\sqrt{x}+2)$;
(3) $f(x)=\sqrt{x}(x+3)$;
(4) $f(x)=\dfrac{\sqrt{x}+4}{4}x$;
(5) $f(x)=\dfrac{\sqrt{x}+5}{10\sqrt{x}}$;
(6) $f(x)=\dfrac{x+6}{12\sqrt{x}}$.

67. Find all antiderivatives of the following functions:

(1) $f(x)=\sqrt{x+2}$;
(2) $f(x)=\sqrt{2x-1}$;
(3) $f(x)=(2x+1)\sqrt{4x+2}$;
(4) $f(x)=\dfrac{1}{\sqrt{x-2}}$;
(5) $f(x)=\dfrac{2}{\sqrt{2x+1}}$;
(6) $f(x)=\dfrac{4x-6}{\sqrt{2x-3}}$.

68. Find all antiderivatives of the following functions:

(1) $f(x)=x(\sqrt[3]{x}+1)$;
(2) $f(x)=\sqrt[3]{x}(\sqrt[3]{x}-1)$;
(3) $f(x)=\sqrt[3]{x}\cdot\sqrt{x}$;
(4) $f(x)=(\sqrt[3]{x}-3)(\sqrt{x}-2)$;
(5) $f(x)=\dfrac{\sqrt[3]{x}+1}{x^2}$;
(6) $f(x)=\dfrac{\sqrt{x}-1}{\sqrt[3]{x}}$.

69. Find all antiderivatives of the following functions:

(1) $f(x)=x\sqrt{x-2}$;
(2) $f(x)=(2x+3)\sqrt{1-x}$;
(3) $f(x)=3x\sqrt[3]{3x-2}$;
(4) $f(x)=\dfrac{2x}{\sqrt{x-1}}$;
(5) $f(x)=\dfrac{2x+1}{\sqrt{2-x}}$;
(6) $f(x)=\dfrac{3x-2}{\sqrt[3]{x-1}}$.

70. Find all antiderivatives of the following functions:

(1) $f(x)=4x\sqrt{x^2-1}$;
(2) $f(x)=\dfrac{x}{\sqrt{x^2+1}}$;
(3) $f(x)=x^2\sqrt{x^3+1}$;
(4) $f(x)=\dfrac{x^2}{\sqrt{1-x^3}}$;
(5) $f(x)=4x\sqrt[3]{1-x^2}$;
(6) $f(x)=\dfrac{x^2}{\sqrt[3]{x^3+17}}$.

Hint. Write the formula for the derivative of a function

$$f(x)=(p(x))^\alpha.$$

Besides the function $f(x)=x^{-1}$, whose antiderivative is the *transcendental* function $F(x)=\ln|x|+\widehat{C}$ (here, $\ln z$ is the *natural logarithm*, which will be discussed in the next chapter), there are many other rational and irrational algebraic functions whose antiderivatives are written through *transcendental* elementary functions. Recall: which of the elementary transcendental functions we have considered have *algebraic derivatives*?

71. Find all antiderivatives of the following functions:

(1) $f(x)=\dfrac{1}{1+x^2}$;
(2) $f(x)=\dfrac{x^2}{1+x^2}$;

(3) $f(x)=\dfrac{1}{4+x^2}$; (5) $f(x)=\dfrac{2x}{1+x^4}$;

(4) $f(x)=\dfrac{1}{2+2x+x^2}$; (6) $f(x)=\dfrac{ax^2+b}{c^2+d^2x^2}$ $(c,d>0)$.

72. Find all antiderivatives of the following functions:

(1) $f(x)=\dfrac{1}{\sqrt{1-x^2}}$; (4) $f(x)=\dfrac{1}{\sqrt{2x-x^2}}$;

(2) $f(x)=\dfrac{1}{\sqrt{9-x^2}}$; (5) $f(x)=\dfrac{x}{\sqrt{1-x^4}}$;

(3) $f(x)=\dfrac{1}{\sqrt{1-4x^2}}$; (6) $f(x)=\dfrac{ax+b}{c^2-d^2x^2}$ $(c,d>0)$.

As you may have guessed, the antiderivatives from the last two exercises are related to *inverse trigonometric functions*: $\arctan x$ and $\arcsin x$ (or $\operatorname{arccot} x=\frac{\pi}{2}-\arctan x$ and $\arccos x=\frac{\pi}{2}-\arcsin x$).

Integrals with variable integration limits

73. Find out at which $x\in\mathbb{R}$ the functions $J(x)$ defined in Exercise 34 are differentiable, and compute the derivatives $J'(x)$ at these points.

74. For the functions $F(x)$ defined as integrals, find the derivatives $F'(x_0)$ at the given points x_0:

(1) $F(x)=\int\limits_1^x \dfrac{du}{\ln u}$, $x_0=e$; (4) $F(x)=\int\limits_0^x e^{-u^2}\,du$, $x_0=\sqrt{\ln 2}$;

(2) $F(x)=\int\limits_1^x \dfrac{\sin u}{u}\,du$, $x_0=\frac{\pi}{2}$; (5) $F(x)=\int\limits_1^x u^u\,du$, $x_0=4$;

(3) $F(x)=\int\limits_1^x \dfrac{e^u}{u}\,du$, $x_0=\ln 2$; (6) $F(x)=\int\limits_{17}^x u^u\,du$, $x_0=4$.

75. Let a function $f(x)$ be continuous on the entire number axis. Find the derivatives of the functions $F(x)$ defined as integrals (on the entire axis):

(1) $F(x)=\int\limits_0^x f^2(u)\,du$; (4) $F(x)=\int\limits_x^{2x} f(u)\,du$;

(2) $F(x)=\int\limits_x^0 f(u)\,du$; (5) $F(x)=\int\limits_0^{x^2} f(u)\,du$;

(3) $F(x)=\int\limits_0^{2x} f(u)\,du$; (6) $F(x)=\int\limits_{-x}^{x} f(u)\,du$.

76. Let F be an antiderivative of some strictly monotonous function $f\colon (a,b)\to(\alpha,\beta)$ on an interval (a,b), and let $g\colon (\alpha,\beta)\to(a,b)$ be the inverse function of f. Prove that the function $G(x)=xg(x)-F(g(x))$ is an antiderivative for g on the interval (α,β). Apply this formula to the following functions:

(1) $f(x)=x^n$, $x\in(0,+\infty)$; (3) $f(x)=\arcsin x$, $x\in(-1,1)$;

(2) $f(x)=\sin x$, $x\in(-\frac{\pi}{2},\frac{\pi}{2})$; (4) $f(x)=\cos x$, $x\in(0,\pi)$.

Give a geometric interpretation for the formula for $G(x)$ (separately in the cases of increasing and decreasing functions f; you may first consider the case $a=\alpha=0$).

* * *

77. In the coordinate plane Oxy, draw regions bounded by the given curves and find their areas:

(1) $y=2x-x^2$, $y=0$;
(2) $y=x^2-3x$, $y=0$;
(3) $y=x^2-2$, $y=x$;
(4) $y=x^2$, $y=2x$;
(5) $y=x^2$, $y=x+2$;
(6) $y=x^2$, $y=x^3$;
(7) $y=-x^2$, $y=x^4-20$;
(8) $y=-x^2$, $y=x^3-12x$;
(9) $y=x^{1/3}$, $x=0$, $y=1$;
(10) $y=x^{1/3}$, $y=x^{1/5}$;
(11) $y=\frac{4}{x^2}$, $y=7-3x$;
(12) $y=x^2$, $y=2x-1$, $y=-2x-1$;
(13) $y=x^2$, $y=(x-1)^2$, $y=\frac{1}{9}$;
(14) $y=\sin^2 x$, $y=0$, $y=\frac{\pi}{2}$;
(15) $y=x^\alpha$, $y=x^\beta$, $(0<\alpha<\beta)$;
(16) $y=\sqrt{x}$, $y=x-2$, $x=0$;
(17) $y=\sqrt{a+x}$, $y=\sqrt{a-x}$, $y=0$;
(18) $y=\sqrt{x}$, $y=\sqrt{4-3x}$, $y=0$.

78. Find areas of the regions bounded by the curves
1) $y^2=x$, $y=2-x$;
2) $y=x$, $y=x^2+1$, $x=0$, $y=2$;
3) $y^2=x$, $y=\sqrt{x}(x^2-1)$;
4) $y=x^3$, $y=0$, tangent line to the graph of $y=x^3$ at $x=1$;
5) $y=\sqrt{x}$, $y=\frac{1}{4}x$, normal line to the graph of $y=\sqrt{x}$ at $x=1$;
6) $y=x^3-x$, tangent line to the graph of $y=x^3-x$ at $x=-1$.

79. Without using antiderivatives, evaluate the integrals

(1) $\int\limits_{-2}^{2} ||x|-1|\,dx$;
(2) $\int\limits_{0}^{5} \left|\{x\}-\frac{1}{2}\right| dx$;
(3) $\int\limits_{-2}^{0} \sqrt{4-x^2}\,dx$;
(4) $\int\limits_{-\pi}^{11\pi} \sin\frac{x}{3}\,dx$;
(5) $\int\limits_{-10}^{10} x^5\cos^{13} x\,dx$;
(6) $\int\limits_{0}^{\pi/2} \cos^2 x\,dx$.

* * *

80. The base of a solid is the triangle $\{0\le x\le 1,\ 0\le y\le x\}$ in the Oxy plane; each of its cross-sections perpendicular to the Ox axis is a square. Find the volume of the solid.

81. The base of a solid is the triangle $\{0\le x\le 1,\ 0\le y\le x\}$ in the Oxy plane; each of its cross-sections perpendicular to the Ox axis is a semicircle. Find the volume of the solid.

82. The base of a solid is the circle $x^2+y^2\le 1$; each of its cross-sections perpendicular to the Ox axis is a square. Find the volume of the solid.

83. Find the volumes of the solids of revolution of the region below the graph of $y=1-x^2$ from $x=0$ to $x=1$: (1) about the Ox axis; (2) about the Oy axis.

84. Find the volumes of the solids of revolution of the region bounded by the curves $y=\sqrt{x}$, $x=0$ and $y=1$: (1) about the Ox axis; (2) about the Oy axis.

85. Find the volumes of the solids of revolution of the region bounded by the curves $y=\sqrt{1-x^2}$, $y=x$ and $y=0$: (1) about the Ox axis; (2) about the Oy axis.

86. Find the volume of the solid of revolution of the region bounded by the curves $y=4-x^2$ and $y=3$ around the line $y=-1$.

87. Find the volume of the solid of revolution of the region bounded by the curves $y=x^3$ and $y=x$ around the line $x=-2$.

* * *

88. (1) Find the ratio of volumes of solids of revolution around the Ox and Oy axes of the curvilinear trapezium bounded by the arc of the graph of $y=x^2$ from $x=0$ to $x=a$, the corresponding axes, and the perpendiculars to them.

(2) Find the same ratio in the case of the graph of an arbitrary power function $y=x^\alpha$, $\alpha>0$.

(3) Solve the same problems for the graphs of the functions $y=Ax^2$ and $y=Ax^\alpha$ $(A,\alpha>0)$.

89. (1) Find the area of the region bounded by the ellipse with semi-axes a and b, i.e., the curve given by the equation

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

(2) Find the volume of the *ellipsoid of revolution* obtained by revolving this region about the Ox axis.

(3) Find the ratio of volumes of the ellipsoids of revolution of this region about the Ox and Oy axes.

90. Find the volume of an arbitrary ellipsoid given in the spatial Cartesian coordinate system $Oxyz$ by the inequality

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1.$$

91. Find volumes of solids of revolution about the Ox axis obtained by revolving

(1) the sine lobe $0\le x\le\pi$, $0\le y\le\sin x$;
(2) the circle $x^2+(y-2)^2\le 1$ (the resulting solid is called a *torus*);
(3) the infinite hyperbolic trapezium $x\ge 1$, $0\le y\le\frac{1}{x}$.

92. Apply the method used to find the area of a sphere to derive the formula for the lateral surface area of a right circular cylinder of radius R and height H.

93. Derive the formula for the volume of a solid of revolution of the curvilinear trapezium

$$0 \le a \le x \le b, \quad 0 \le y \le f(x)$$

about the Oy-axis ($f(x)$ being a continuous and non-negative function on a segment $[a, b]$).

(*Answer*: $V = 2\pi \int_a^b x f(x)\, dx$.)

94. Find volumes of solids of revolution about the Oy axis obtained by revolving

(1) the curvilinear triangle $0 \le x \le a$, $0 \le y \le x^\alpha$ $(\alpha > 0)$;
(2) the sine lobe $0 \le x \le \pi$, $0 \le |y| \le \sin x$;
(3) the circle $(x-2)^2 + y^2 \le 1$ (the solid of revolution is again a *torus*);
(4) the circle $(x-1)^2 + y^2 \le 1$ (and this is not a torus);
(5) the circle $(x-0.5)^2 + y^2 \le 1$ (not a torus at all!).

95. Derive the formula for the volume of the general truncated cone with base areas S_1 and S_2 and height h. (Recall that a *truncated cone* is a part of the [general] cone enclosed between the base plane and a cutting plane parallel to it. Solve the problem in two ways: using "similarity arguments" and using *Simpson's rule*).

(*Answer*: $V = \frac{1}{3}h(S_1 + \sqrt{S_1 S_2} + S_2)$.)

96. Find the volume of an "attic" whose base is a rectangle of size $a \times b$, top edge is c, and height is h.

97. Find the volume of an "obelisk" whose parallel bases are rectangles of sizes $A \times B$ and $a \times b$ (with respective sides parallel to each other) and whose height is h.

* * *

98. Derive the formula for the volume of a spherical segment of radius R and height H. (Recall that a *spherical segment* is a part of a solid sphere cut off from it by any plane intersecting it; the *radius of the segment* is the radius R of the original sphere, and the *height* is the distance H from the cutting plane to the plane parallel to it that touches the segment [of the sphere]. *Question*: within what range can the height of a segment of radius R vary?)

99. Derive the formula for the volume of a spherical sector corresponding to a spherical segment of radius R and height H. (This *spherical sector* is defined as the solid consisting of all radii of the sphere connecting the centre with points of the spherical surface that belong to the given spherical segment [these points form a *spherical cap*]. Note that a spherical sector can as well be *non-convex*!)

100. Using the technique used to find the area of a sphere, derive the formula for the area of a spherical segment of radius R and height H.

Recall that a *spherical zone* is a part of a spherical surface enclosed between two parallel planes intersecting it. The spherical zone is specified by the radius R of the sphere and two other parameters: the distances from the centre of the sphere to the planes *with signs taken into account.* If we assume that the sphere in a spatial Cartesian coordinate system $Oxyz$ is given by the equation $x^2+y^2+z^2=R^2$ and that parallel sections are perpendicular to the Ox axis, then the parameters are determined by the equations $x=a$ and $x=b$ with $a,b\in(-R,R)$.

101. Find the area of a spherical zone of radius R with parameters $a<b$.

Comment. The result obtained in this problem is quite remarkable: *the area of a spherical zone is determined only by its radius R and its* height H, *i.e., the distance between the cross-section planes*!

* * *

102. Derive the formula for the length of the arc of the graph of a differentiable function $y=f(x)$ from point $x=a$ to point $x=b>a$.

Hint. Consider the "variable length", i.e., the function $L(x)$ defined as the length of the arc of the graph from a to x; find the derivative $L'(x)$.

(*Answer*: $L=\int\limits_a^b \sqrt{1+(f'(x))^2}\,dx$.)

103. Find the arc lengths of the graphs of the given functions between the given points:

(1) $y=x\sqrt{x}$; $x=0$, $x=1$;

(2) $y=(x+1)^{3/2}$; $x=3$, $x=8$;

(3) $y=x^{2/3}-1$; $x=8$, $x=27$;

(4) $y=\frac{1}{3}(x^2+2)^{3/2}$; $x=\sqrt{2}$, $x=\sqrt{7}$;

(5) $y=\frac{2}{3}x\sqrt{x}-\frac{1}{2}\sqrt{x}$; $x=1$, $x=4$;

(6) $y=\frac{x^4}{4}+\frac{1}{8x^2}$; $x=1$, $x=2$;

(7) $y=\sqrt{x}$; $x=0$, $x=1$;

(8) $y=\ln x$; $x=\sqrt{3}$, $x=\sqrt{8}$.

* * *

104. Prove that the arc length of the hodograph of a differentiable vector-valued function $\overline{r}(t)=(x(t);y(t))$ from $t=\alpha$ to $t=\beta>\alpha$ is

$$L=\int\limits_{\alpha}^{\beta}\sqrt{(x'(t))^2+(y'(t))^2}\,dt.$$

105. Let a closed arc of the hodograph of a differentiable vector-valued function $\overline{r}(t)=\overline{OP_t}=(x(t);y(t))$ from $t=\alpha$ to $t=\beta>\alpha$ bound the *left* (as the point P_t moves with increasing t) side of some figure Φ. Prove that the area of this figure can be found by any of the following three formulae:

$$S=\int\limits_{\alpha}^{\beta}x(t)y'(t)\,dt,\qquad S=-\int\limits_{\alpha}^{\beta}y(t)x'(t)\,dt,$$

$$S=\frac{1}{2}\int\limits_{\alpha}^{\beta}(x(t)y'(t)-x'(t)y(t))\,dt.$$

Remark. Closeness of the arc means that $\overline{r}(\alpha)=\overline{r}(\beta)$.

106. Check the formulae from the last two problems for a parameterised circle of radius R: $\overline{r}(t)=(R\cos t;R\sin t)$, $t\in[0,2\pi]$.

Comment. Defining a plane curve as the hodograph of a vector-valued function $\overline{r}(t)=(x(t);y(t))$ is called *defining the curve in parametric form* (or *parametrically*). In other words, the representation of a curve by a hodograph is called its *parameterisation.* Note that the same curve can be parameterised in *infinitely many ways.* For example, by substituting into the parameterisation $\overline{r}(t)$, instead of the argument t, any function $\varphi(\tau)$ which is monotone on the entire numerical axis $\mathbb{R}$ and whose image coincides with $\mathbb{R}$ (for instance, a linear function $t=\varphi(\tau)=k\tau+b$), we obtain a different parameterisation of the same curve by means of the vector function $\overline{\rho}(\tau)=\overline{r}(\varphi(\tau))$.

107. A *cycloid* is the trajectory of a fixed point on a circle rolling along a straight line without slipping.

(1) Prove that the cycloid can be parameterised as

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

(2) Using this parameterisation, find the length of one arch of the cycloid $\overline{r}(t)=(x(t);y(t))$, $0\le t\le 2\pi$.

(3) Find the area of the region between this cycloid arch and the Ox axis.

108. An *astroid*[1] is a curve given in Cartesian coordinates by the equation

$$x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}} \quad (a > 0).$$

(1) Give any parameterisation of the astroid.

(2) Using an appropriate parameterisation of the astroid, find its length.

(3) Using an appropriate parameterisation of the astroid, find the area of the region that it bounds.

109. The curve given in Cartesian coordinates by the equation

$$x^3 + y^3 = 3axy \quad (a > 0),$$

is called the *folium of Descartes* (from Latin *folium*, 'leaf').[2]

(1) Draw the folium of Descartes (observe that this curve has a symmetry axis and an oblique asymptote; find them).

(2) Prove that the equations

$$x = \frac{3at}{1+t^3}, \qquad y = \frac{3at^2}{1+t^3}$$

give a parameterisation of the folium of Descartes.

(3) Prove that under this parametrisation, the three points on the folium of Descartes corresponding to the parameter values t_1, t_2, and t_3 lie on the same straight line if and only if $t_1 t_2 t_3 = -1$.

110. The *involute* (or *evolvent*) *of a circle* is the trajectory traced by the free end of a string of length $2\pi a$ unwound from a circle of radius a.

(1) Prove that the involute of the circle can be given parametrically by

$$x = a(\cos t + t\sin t), \quad y = a(\sin t - t\cos t) \quad (0 \le t \le 2\pi).$$

(2) Using this parameterisation, find the arc length of the involute.

(3) Find the area of the region bounded by the involute of the circle (in this parameterisation) and the line $x = a$ assuming that $y \le 0$.

---

[1] Literally '*star-shaped*', from Greek αστρον [astron], 'star' (hence the word *astronomy*), and ειδος [eidos], 'form', 'likeness' (respectively, νομος [nomos], 'law', 'custom').

[2] This curve was first mentioned by Descartes in a letter to Fermat (1638).

Chapter III

Exponential Function and Linear Differential Equations

§ 3.1. Linear Processes and Differential Equation $y' = ky$

Recall (see § 1.1) that in *continuous* mathematical models, where the state of a system is described by a function $x(t)$ of a continuously varying variable (time $t \in \mathbb{R}$), it is often possible to obtain information about the *rate* of change of $x(t)$ written as a *differential equation* of the form

$$x' = F(x, t). \tag{1}$$

The problem of studying the evolution of a system described by the differential equation (1) consists in finding all solutions of this equation and analysing them. The success in solving this general problem depends essentially on a particular kind of the function of two variables $F = F(x, t)$ appearing in the differential equation. If the function F depends only on t, i.e., $F(x, t) \equiv v(t)$, then we come to the basic problem of *integral calculus*; we have considered it in the preceding chapter. Now let us focus on the case where the function F depends only on x: $F(x, t) \equiv f(x)$, and moreover, depends *linearly* on x. The corresponding *first-order linear differential equations*,[1] the homogeneous equation

$$x' = kx \quad \text{(or, in other notation, } y'(x) = ky(x)\text{)}, \tag{2}$$

and the *non-homogeneous* equation

$$x' = kx + b \quad \text{(or } y'(x) = ky(x) + b\text{)}, \tag{3}$$

are encountered in many applied problems. To begin with, we give a few simple examples leading to the homogeneous linear differential equation (2). The corresponding processes are generally called *linear*.

3.1.1. Example: population growth. Let us denote by $N(t)$ the number of some population of bacteria or any unicellular individuals. If they are not killed and are provided with sufficient conditions for reproduction, i.e., living space and nutrient medium, then during a small time interval from t to $t + \Delta t$ each individual "produces offspring" (simply by division)

[1] This means that such equations involve only the first derivative.

with some probability proportional to Δt, so that the population increment is

$$\Delta N(t, \Delta t) = N(t + \Delta t) - N(t) \approx N(t) \cdot \alpha \Delta t = \alpha N(t) \cdot \Delta t.$$

It is natural to assume that this equality is the more exact the smaller Δt, so that after dividing by Δt, an exact equality should be obtained as $\Delta t \to 0$:

$$\lim_{\Delta t \to 0} \frac{\Delta N}{\Delta t} = N'(t) = \alpha N(t).$$

We arrive at a differential equation of the form (2) with a positive coefficient $k = \alpha$.

Note that in this example we have abstracted from the fact that the number of bacteria is *integer* (and hence the function $N(t)$ is stepwise and discontinuous, not to mention differentiable!). If N is a *very large* integer, then the appearance of new bacteria changes it quite little. Therefore, if we treat $N(t)$ as if it were a continuous or even differentiable function, we get a fairly accurate model of the real situation!

3.1.2. Example: radioactive decay. Let us take an example from physics. In 1896 Becquerel discovered the phenomenon of natural *radioactivity.*[1] The essence of this phenomenon is that atoms of some elements (called *radioactive*: uranium, radium, polonium, etc.) spontaneously decay (disintegrate), turning into other substances, and the radioactive decay is accompanied by radiation having much higher energy than light (a particular kind of electromagnetic radiation).

Antoine Henri Becquerel (1852–1908) was a French physicist. His father Alexandre-Edmond (1820–1891) and grandfather Antoine César (1788–1878) were also prominent physicists and studied various kinds of *radiation* such as *luminescence.* Antoine Becquerel, in a way continuing their research, conducted experiments to test the hypothesis of the famous French mathematician and physicist Henri Poincaré (1854–1912) that the *fluorescence* of uranium salts is of the same nature as the X-rays, discovered in 1895 by the German experimental physicist Wilhelm Röntgen (1845–1923; for this discovery he was awarded the Nobel Prize in 1901, the first among physicists). In 1896 Becquerel made an even bigger discovery when he found that uranium crystals emit without fluorescing. This phenomenon was investigated in Paris by the couple Pierre Curie (1859–1906; born and worked in Paris, tragically died in an accident) and Maria Skłodowska-Curie (1867–1934; born in Warsaw, studied in Paris, worked in Paris and Sèvres). By analysing the decay products of uranium ore (tons of ore!), they discovered two new radioactive elements in 1898: *polonium* (named after Poland, Skłodowska's homeland) and *radium*, whose radiation was stronger than that of uranium. At that time, they coined the term *radioactivity.*

In 1903, all the three pioneers of the study of radioactivity were awarded the Nobel Prize, and in 1911, Marie Skłodowska-Curie was awarded another Nobel Prize, this time

---

[1] From Latin *radiare*, 'to emit rays' (a word related to the mathematical term *radius*, in Latin 'spoke in a wheel', 'ray of light') and *activus*, from *actus*, 'doing'.

in chemistry, for her discovery of radium and polonium, synthesis of metallic radium, and her experiments with this element. (Skłodowska-Curie died of radiation sickness. Curie's daughter Irène also became an outstanding physicist; see about her in § 3.4).

The experimentally established *fundamental law of radioactive decay* is that the ratio of the number of atoms decayed per unit time to the total number of atoms (which is assumed to be large) is a constant, called the *decay probability*, depending only on the kind of atoms. Thus, if $M(t)$ denotes the number of atoms at time t and ω the decay probability, then the increment of the function $M(t)$ over the time interval Δt can be written as

$$\Delta M(t, \Delta t) = M(t + \Delta t) - M(t) \approx -\omega M(t) \cdot \Delta t$$

(the minus sign corresponds to the fact that the number of atoms decreases). As in the previous example, after dividing by Δt and passing to the limit as $\Delta t \to 0$, we obtain a differential equation of the form (2),

$$M'(t) = -\omega M(t),$$

but this time with a negative coefficient $k = -\omega$. Note that in introducing this model of radioactive decay, we also abstract from the discreteness of the change in the number of atoms; considering it very large, we assume that it is described by a continuous and differentiable function $M(t)$.

3.1.3. Example: viscous friction. The last example is from mechanics. When a body moves in a viscous medium (liquid or gas), it is subject to the *viscous friction force*, which is directed opposite to the velocity of motion and proportional to it:

$$F_{\mathrm{fr}} = -\lambda v = -\lambda x'$$

(as usual, we assume motion to be rectilinear and forces and velocities to be scalars with signs corresponding to their directions). The *friction coefficient* λ is proportional to the so-called *viscosity* η of the medium,[1] and depends also on the shape and size of the moving body.

For example, for a sphere of radius R, the value of λ is $6\pi R\eta$; this formula is called *Stokes' law.* Sir George Gabriel Stokes (1819–1903) was an English physicist and mathematician whose name, along with this law (1851), is given to the general equations of viscous fluid motion (the Navier–Stokes equations; Claude-Louis Marie Henri Navier [1785–1836] was a French mathematician and mechanic), as well as to one of the most important formulae of multivariate (vector) calculus.

If the motion is *free*, i.e., no external forces (except friction) act on the body, then Newton's second law,

$$mx'' = F = -\lambda x',$$

[1]For example, for air, $\eta = 0.00018 \frac{\mathrm{g}}{\mathrm{cm} \cdot \mathrm{s}}$, for water (at 20°), $\eta = 0.01 \frac{\mathrm{g}}{\mathrm{cm} \cdot \mathrm{s}}$.

gives for the velocity $v = x'$ a linear differential equation of the form (2):

$$v' = -\frac{\lambda}{m} v.$$

If the body is additionally subject to a constant external force, we arrive at a non-homogeneous linear differential equation of the form (3); for example, for a body falling in the Earth's gravity field with a parachute, we get the equation

$$mv' = -mg - \lambda v.$$

We will come back to these and other examples, but now we will turn to a *preliminary* general analysis of differential equations of the form (2).

3.1.4. Analysis of the differential equation $y'(x) = ky(x)$: Euler's approach. Recall (see § 2.2 and § 2.3), that a differential equation of the form (1), or, in other notation,

$$y' = F(x, y), \quad \text{i.e.,} \quad y'(x) = F(x, y(x)),$$

where $F = F(x, y)$ is a given function of two variables, i.e., a function on the Oxy coordinate plane, defines a *directional field* in this plane. Graphs $y = y(x)$ of solutions of the equation are *integral curves* of this directional field, and (in good cases) they can be approximated by *Euler's polygonal lines*. If $y(x_0) = y_0$, then vertices of an n-segment Euler polygonal line are the points

$$x_m = x_0 + m\frac{x - x_0}{n}, \qquad y_{m+1} = y_m + F(x_m, y_m)(x_{m+1} - x_m). \tag{4}$$

This gives for $y(x)$ the approximation

$$y(x) \approx y_n = y_0 + \sum_{m=0}^{n-1} F(x_m, y_m)(x_{m+1} - x_m).$$

For the equation $y'(x) = ky(x)$, the function $F(x, y) = ky$ depends only on y, so the directional field does not change along the horizontal lines $y = \text{const}$. It follows that any integral curve maps to an integral curve under a parallel translation along the Ox axis. Analytically, this means that if the function $y = y(x)$ is a solution of the differential equation (2), then the "shifted" function $y = y(x + c)$ is also a solution of equation (2). Therefore, we can replace the initial condition $y(x_0) = y_0$ by an initial condition at the point $x_0 = 0$, i.e., consider $y(0) = y_0$ to be given.

We can also notice that in the case of $k > 0$ the directional field goes the steeper the larger the value of y, and Euler's polygonal lines "rapidly and steeply" move away from the Ox axis (Fig. 34).

Now let us look at the formulae. In the recurrence relations (4), all brackets $(x_{m+1} - x_m)$ are the same — they are all equal to the n-th part

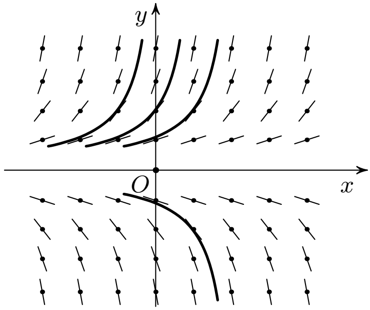


Fig. 34.

of the segment $[x_0, x] = [0, x]$, i.e., to $h = \frac{x}{n}$. Substituting this value of $x_{m+1} - x_m$, as well as $F(x_m, y_m) = ky_m$, into (4), instead of the additive recurrence relation we obtain a multiplicative one,

$$y_{m+1} = y_m + ky_m\frac{x}{n} = \Big(1 + \frac{kx}{n}\Big)y_m.$$

Clearly, it defines a geometric progression with first term y_0 and common ratio $q = 1 + \frac{kx}{n}$, and we are interested in the n-th element of the progression, which is

$$y_n = y_0 \cdot \Big(1 + \frac{kx}{n}\Big)^n.$$

Thus, according to our assumption, the solution of the differential equation $y' = ky$ satisfying the above initial condition should be given by the limit relation

$$y(x) = \lim_{n\to\infty} y_0 \cdot \Big(1 + \frac{kx}{n}\Big)^n. \qquad (5)$$

We can follow this way further and prove the existence of the limit (5), investigate the properties of the limit function, and so on.[1] We will proceed differently (somewhat simpler), returning to the geometric interpretation of the equation a little later. Now we will consider how the differential equation (2) was analysed by Newton.

3.1.5. Analysis of the equation $y'(x) = ky(x)$: Newton's approach. We will outline the Newton approach using a special case of equation (2) where $k = 1$ and under the unit initial condition:

$$y'(x) = y(x), \qquad y(0) = 1. \qquad (6)$$

[1]Note that y_n is the n-th term of a geometric progression whose common ratio q *depends* on n, so the question of the behaviour of the sequence y_n at $n \to \infty$ is by far not easy.

Assume (following Newton) that problem (6) has a solution $y = y(x)$ written as a *polynomial* in x:

$$y(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + a_4 \cdot x^4 + \dots \tag{7}$$

(the degree of the polynomial is unknown). From the initial condition $y(0) = 1$ we find the value of the constant term:

$$a_0 = 1. \tag{8}$$

Now let us differentiate expression (7):

$$y'(x) = a_1 + 2a_2 \cdot x + 3a_3 \cdot x^2 + 4a_4 \cdot x^3 + \dots,$$

and equate the result to the expression itself, according to equation (6). Equality of polynomials is equality of their respective coefficients; therefore, taking into account condition (8), we must have the equalities

$$\begin{aligned} a_1 = a_0 &= 1, \\ 2a_2 = a_1 = 1 &\quad\Rightarrow\quad a_2 = \frac{1}{2}, \\ 3a_3 = a_2 = \frac{1}{2} &\quad\Rightarrow\quad a_3 = \frac{1}{2 \cdot 3} = \frac{1}{3!}, \\ 4a_4 = a_3 = \frac{1}{3!} &\quad\Rightarrow\quad a_4 = \frac{1}{4!}, \end{aligned}$$

etc. The coefficient a_n of x^n is

$$a_n = \frac{1}{n!}.$$

Thus, the search for a solution to problem (10) in the form of a polynomial led to a "polynomial of infinite degree", but with very nice coefficients:

$$y(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots + \frac{1}{n!}x^n + \dots. \tag{9}$$

Expressions of this kind are called *power series*, and Newton succeeded in finding solutions to various differential equations in the form of (infinite) power series. One can prove that for any value of x the series (9) converges, in the sense of existence of the limit

$$y(x) = \lim_{N \to \infty} \sum_{n=0}^{N} \frac{1}{n!} x^n.$$

A little more difficult is to prove that the limit function $y(x)$ is differentiable and satisfies the differential equation (6). Even more difficult is to derive basic (and very interesting!) properties of the solution $y(x)$. In this case, as it often happens, it is more efficient to *proceed from the very differential equation* of the form (2). This is what we will do, but first consider another variant of approaching our differential equation.

3.1.6. Analysis of the equation $y'(x)=ky(x)$: difference analogue. Consider the *difference analogue* of the differential equation (2), i.e., a difference equation[1] corresponding to known information about the rate of change of a discretely varying quantity, i.e., of a sequence $y(n)=y_n$. The rate of change of the sequence is naturally defined as simply the *difference* $\Delta y(n)=\Delta y_n=y_{n+1}-y_n$.

The difference of a sequence is in many respects similar to the derivative of a function. For example, for the quadratic function x^2, the derivative is $(x^2)'=2x$, and for the corresponding quadratic sequence n^2, the difference is $\Delta n^2=(n+1)^2-n^2-n^2=2n+1$, almost the same! Accordingly, the differential equation $y'(x)=ky(x)$ should be replaced by the difference equation

$$\Delta y_n=ky_n. \tag{10}$$

Hence we obtain

$$y_{n+1}-y_n=ky_n \quad \Rightarrow \quad y_{n+1}=(1+k)y_n=qy_n, \quad q=1+k.$$

Therefore, solutions of the difference equation (10) are geometric progressions:

$$y_n=y_0\cdot(1+k)^n=Aq^n, \quad q=1+k.$$

A quite natural *conjecture* arises: since the solutions of the difference equation $\Delta y(n)=ky(n)$ are geometric progressions $y(n)=Aq^n=y_0\cdot q^n$, i.e., *exponential functions of a natural argument*, $y(n)=Aq^n=y_0\cdot q^n$, solutions of the differential equation $y'(x)=ky(x)$ should be *exponential functions* $y(x)=Aa^x=y(0)\cdot a^x$ *of a real argument* $x\in\mathbb{R}$ for some value of the base a that depends on the coefficient k. But if we *do not know* what is an exponential function of a real argument, i.e., a number a raised to an *arbitrary* real power, there is nothing to talk about. We know how to raise any number $a>0$ to a fractional (i.e., rational) power, but the corresponding definition is completely unsuitable for finding, say, $a^{\sqrt{2}}$.

One can proceed as follows: starting from *only the differential equation* (2), derive some (good!) properties of its solutions, in particular, try to justify the last conjecture somehow. And after that, turn to how to define an exponential function of an arbitrary real argument, what are its properties, its derivative, etc. This agenda is given in the practical exercise A in the Exercises and Tasks to this chapter; carry it out on your own. Now, we will again turn to the geometrical interpretation of the differential equation (2) in order to approach the problem of finding its solutions from a slightly different angle.

[1] For more details, see exercises to Ch. II.

§ 3.2. Natural Logarithm and Exponential Function

Let us first outline the *idea* of the approach to the analysis of the differential equation (2), $y' = ky$, $k \neq 0$, which is undertaken in this section. The directional field of equation (2) is *constant along the horizontal lines* $y = \text{const}$. Under symmetry about the bisector of coordinate angles I and III, i.e., the line $y = x$, this directional field, considered together with its integral curves $y = y(x)$ (Fig. 35), transforms into a directional field constant along *vertical lines* (Fig. 36). Such fields were considered in Ch. II; they correspond to differential equations defining the antiderivatives of the corresponding functions. Thus, the differential equation (2) must reduce to a differential equation of the form $x'(y) = f(y)$, which we already know how to solve.

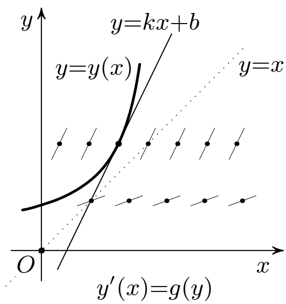


Fig. 35.

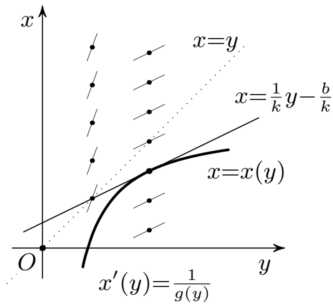


Fig. 36.

3.2.1. "Symmetric" differential equation. Let us find out more precisely what we get after symmetry. First, the Ox and Oy axes change their places, and each point $(x; y)$ moves to the point $(y; x)$.

Second, the line $y = kx + b$ defining the direction at the point (x, y) and having slope $K = g(y) = ky$, which we will consider to be non-zero, maps to the *symmetric line*. It defines the graph of the inverse function, the linear function $x = \frac{1}{K}y - \frac{b}{K}$ with slope $K_{\text{sym}} = \frac{1}{K}$. Thus, the symmetric directional field in the Oyx plane is given by

$$K_{\text{sym}} = F(y, x) = \frac{1}{K} = \frac{1}{g(y)} = f(y) = \frac{1}{ky} = \frac{1}{k} \cdot \frac{1}{y};$$

i.e., in the Oyx plane we obtain a *known* directional field constant along vertical lines.

And third, if a solution $y = y(x)$ of the differential equation (2) in the neighbourhood of the point x is *invertible*, then its graph maps exactly to the graph of the *inverse function* $x = x(y)$. It is intuitively clear that tangency is preserved under symmetry, so the inverse function $x(y)$ in the neighbourhood of y must satisfy the *"symmetric"* differential equation:

$$x'(y) = f(y) = \frac{1}{g(y)} = \frac{1}{k} \cdot \frac{1}{y} \quad \Leftrightarrow \quad x = \frac{1}{k} G(y) + C, \tag{1}$$

where $G(y)$ is some (any) antiderivative of the function $\lambda(y) = y^{-1}$ in the neighbourhood of y. It remains to express y through x from equation (3); owing to the assumption that the solution is invertible in the neighbourhood of x, this is essentially possible!

Thus, the analysis of solutions of the homogeneous linear equation (2) $y' = ky$ is reduced to the study of antiderivatives of $G(y)$ of the *inverse proportionality*, i.e., the function $\lambda(y) = y^{-1}$. Let us forget about the original differential equation (2) for the moment and consider the "symmetric" differential equation (3), or rather its special case, the equation

$$y'(x) = \lambda(x) = \frac{1}{x} \tag{2}$$

(for the sake of simplicity, we have swapped the variables y and x again).

3.2.2. Natural logarithm. According to what we have proved in Ch. IV, the continuity of the function $\lambda(x) = x^{-1}$ on the positive semi-axis $\mathbb{R}_+ = (0, +\infty)$ implies the existence of a solution of equation (4) on this semi-axis and its uniqueness under a given initial condition $y(x_0) = y_0$, $x_0 > 0$.

Definition 1. The solution $y(x)$, $x \in \mathbb{R}_+$, of the differential equation (4) satisfying the initial condition $y(1) = 0$ is called the *natural logarithm* and is denoted by $\ln(x)$ or $\ln x$.

In other words, the natural logarithm is a function that is (uniquely) defined on the positive semi-axis by two relations:

$$\forall x > 0 \;\; \ln' x = \frac{1}{x}, \qquad \ln 1 = 0. \tag{3}$$

It follows from the definition that $\ln' x > 0$ on the entire semi-axis, so *the natural logarithm increases on the positive semi-axis.* Therefore, the function $x \mapsto \ln x$ is *invertible* on the positive semi-axis and hence has an *inverse function*, called the *natural exponential function* (or simply the *natural exponential*) and denoted by exp:

$$x = \exp y \overset{\text{def}}{\Longleftrightarrow} y = \ln x.$$

Presumably, this function can be used to describe all solutions of the differential equation (2), so it should be considered in more detail, starting with the question of the domain of definition. Of course, the domain of the exp function coincides with the range of the natural logarithm, but so far we do not know much about it either.

Theorem 1 (fundamental property of the natural logarithm). *For any positive numbers a and b, we have*

$$\ln(ab) = \ln a + \ln b.$$

Proof. For an arbitrary number $a > 0$, consider the function $\varphi(x) = \ln(ax)$ defined on the entire positive semi-axis $x > 0$. By the chain rule formula for the derivative of a composite function and by condition (5), its derivative is

$$\varphi'(x) = (\ln(ax))' = \ln'(ax) \cdot (ax)' = \frac{1}{ax} \cdot a = \frac{1}{x} = \ln' x.$$

Therefore (according to Proposition 1 in § 2.1), the functions $\varphi(x) = \ln(ax)$ and $\ln x$ *differ by a constant*, i.e.,

$$\varphi(x) = \ln(ax) = \ln x + C.$$

Plugging $x = 1$ into this formula and taking into account the second condition in (5), we obtain

$$\varphi(a \cdot 1) = \ln 1 + C \quad \Leftrightarrow \quad \ln a = 0 + C \quad \Rightarrow \quad C = \ln a;$$

hence,

$$\forall x > 0 \;\; \ln(ax) = \ln x + C = \ln x + \ln a,$$

as required (take $x = b > 0$ and interchange the terms on the right-hand side). □

Corollary 1. *For any positive number a, we have* $\ln \frac{1}{a} = -\ln a$.

Proof. Since $a \cdot \frac{1}{a} = 1$, by Theorem 1 and condition (5) we obtain

$$0 = \ln 1 = \ln\Big(a \cdot \frac{1}{a}\Big) = \ln a + \ln \frac{1}{a} \quad \Rightarrow \quad \ln \frac{1}{a} = -\ln a. \qquad \square$$

Corollary 2. *For any $a > 0$ and any integer m we have*

$$\ln a^m = m \ln a.$$

Proof. If $m = n \in \mathbb{N}$ is a natural number, the claim can be proved by *induction on n*: if $\ln a^n = n \ln a$, then by Theorem 1 we have

$$\ln a^{n+1} = \ln(a^n \cdot a) = \ln a^n + \ln a = n \ln a + \ln a = (n+1) \ln a.$$

For $m = 0$ the claim coincides with condition (5). Lastly, if $m = -n$, where $n \in \mathbb{N}$ is a natural number, then by the corollary above and by what has been already proved in the case of $m = n \in \mathbb{N}$, we obtain

$$a^m = a^{-n} = \frac{1}{a^n} \quad \Rightarrow \quad \ln a^m = \ln \frac{1}{a^n} = -\ln a^n = -n \ln a = m \ln a. \qquad \square$$

Proposition 1. *The function* $\ln x$ *takes arbitrarily large* (*in magnitude*) *values, both positive and negative*; *moreover,*

$$\lim_{x \to +\infty} \ln x = +\infty, \qquad \lim_{x \to 0+} \ln x = -\infty.$$

Proof. Take an arbitrary $a>1$ and consider two sequences, $u_n=a^n$ and $v_n=a^{-n}$. By Corollary 2 we have

$$\ln u_n = n\ln a, \qquad \ln v_n = -n\ln a,$$

and since we have $\ln a>0$ when $a>1$, we obtain

$$\lim_{x\to\infty} \ln u_n = +\infty, \qquad \lim_{x\to\infty} \ln v_n = -\infty.$$

Since the function $\ln x$ is increasing on the positive semi-axis, and *by the theorem on the limit of a geometric progression* and according to the choice of a $(a>1)$ as $n\to\infty$ we have $u_n=a^n\to\infty$ and $v_n=\left(\frac{1}{a}\right)\to 0+$, the claim of Proposition 1 follows from what was proved above (conduct a formal ["rigorous"] proof on your own). □

Corollary 3. *The range of* $\ln x$ *is the entire number axis* $\mathbb{R}$: $E(\ln)=\mathbb{R}$.

Proof. By its very definition, the function $\ln x$ is *differentiable* and therefore *continuous* on the entire positive semi-axis $x>0$, and since by Proposition 1 it takes arbitrarily large (in magnitude) positive and negative values, *by the intermediate value theorem* we have

$$\forall z\in\mathbb{R}\ \exists x>0 \quad \ln x = z.$$ □

Note that all the properties of the natural logarithm proved above, including its increasing on its domain $D(\ln)=\mathbb{R}_+$, have been proved based only on the differential equation (4), whose solution is the natural logarithm (it was used in the proof of the *fundamental property*, Theorem 1, and in establishing the monotonicity), and on the initial condition satisfied by this solution ($y(1)=\ln 1=0$).

3.2.3. Natural exponential function. The facts proved above imply that the function $z=\ln x$ *is invertible on the positive semi-axis* $x>0$ *and takes all real values* $z\in\mathbb{R}$ *on it, so there exists an inverse function of* $\ln x$ *which is defined on the entire number axis* $\mathbb{R}$ *and takes any positive values.*

Definition 2. The function $\exp(z)$ (or $\exp z$, without parentheses) inverse to $z=\ln x$ on the positive semi-axis $x>0$ is called the *natural exponential function.*

Thus, the natural exponential $\exp(z)$ is defined on the entire number axis $\mathbb{R}$, strictly increases on it,[1] and takes any positive values; moreover, $\exp 0=1$ (this follows from condition (5): $\ln 1=0$). Furthermore, like any two *mutually inverse functions*, the functions ln and exp obey the relations

$$\forall x>0\ \exp(\ln x)=x; \qquad \forall z\in\mathbb{R}\ \ln(\exp z)=z. \tag{4}$$

[1] A function inverse to a strictly monotone function is strictly monotone itself, with the same "direction" of monotonicity.

Let us prove some more properties of the natural exponential function, first making two preliminary remarks.

Remark 1. To prove that two positive numbers coincide, it is sufficient to prove that their natural logarithms coincide.

Remark 2. To prove that arbitrary two numbers coincide, it is sufficient to prove that their natural exponentials coincide.

These facts follow from the strict monotonicity of the functions $\ln x$ and $\exp z$ (e.g., we can reason *by contradiction*). Note in addition that in both cases instead of the words "it is *sufficient* to prove" we could say "it is *necessary and sufficient* to prove"; however, the claim about the *"sufficiency"* suffices (actually, even Remark 1 would suffice).

Proposition 2 (derivative of the natural exponential)**.** *The function* $\exp x$ *is differentiable on the entire number axis*; *moreover,*

$$\forall x \in \mathbb{R} \quad \exp' x = \exp x.$$

Proof. By the *inverse function derivative theorem*, using the formula for the derivative of the natural logarithm (according to condition (5) $\forall x > 0$ $\ln' x = \frac{1}{x} > 0$), we find

$$\exp' x = \left(\frac{1}{\ln' y}\right)_{y=\exp x} = \frac{1}{\left(\frac{1}{y}\right)_{y=\exp x}} = y\big|_{y=\exp x} = \exp x. \qquad \square$$

Corollary 4. *The function* $y = y(x) = \exp x$ *is a solution of the differential equation*

$$y' = y$$

satisfying the "unit" initial condition

$$y(0) = 1.$$

Proof. The proof is obvious. □

Thus, we have found one of the solutions of one of the differential equations of the form (2): $y' = ky$ with $k = 1$. Can you find any *other* solutions to this equation? Give *all* its solutions? We will return to the discussion of these questions for arbitrary values of k later, at the end of the section.

Theorem 2 (fundamental property of the natural exponential)**.** *For any numbers* $a, b \in \mathbb{R}$, *we have*

$$\exp(a + b) = \exp a \cdot \exp b.$$

Proof. In fact, this theorem can be deduced from the corollary above, i.e., from the differential equation $y' = y$ taking into account the initial condition $y(0) = 1$, following the scheme from Exercise A to this chapter (to § 3.1). However, it is simpler to use directly the definition of the natural exponential function: according to Remark 1, it suffices to prove that $\ln(\exp(a+b)) = \ln(\exp a \cdot \exp b)$.

By the second identity in (6), the left-hand side of this equality is of the form $L = a + b$. Next, Theorem 1 (the *fundamental property* of the function $\ln z$) implies that the right-hand side is $R = \ln\exp a + \ln\exp b$, and the same identity (6) yields $R = a + b = L$. □

Note that the fundamental property of the natural exponential can be seen immediately, without the above formal proof. Indeed, the fundamental property of the logarithm (Theorem 1) means that the *mapping* (function) $\ln\colon \mathbb{R}_+ \to \mathbb{R}\colon x \mapsto \ln x$ *takes the multiplication operation* on $\mathbb{R}_+$ *to the addition operation* on $\mathbb{R}$. Hence, the *inverse mapping* $\exp\colon \mathbb{R} \to \mathbb{R}_+\colon x \mapsto \exp x$ *takes the addition operation* on $\mathbb{R}$ *to the multiplication operation* on $\mathbb{R}_+$, and this is just the statement of Theorem 2. Think over this reasoning.

3.2.4. Natural exponential and exponential functions.

Definition 3. The value of the natural exponential function at $x = 1$, i.e., the number $\exp(1)$, is called the *base of the natural logarithm*(*s*) (*Napier's constant*, or sometimes *Euler's number*) and is denoted by e.

Theorem 3 (on values of the natural exponential at rational points)**.** *For any* rational *number x, the value of* $\exp x$ *coincides with the value of the exponential function* e^x:

$$\forall x \in \mathbb{Q} \quad \exp x = \mathrm{e}^x.$$

Proof. Recall that for any positive number a, its rational power a^x, $x = \frac{m}{n} \in \mathbb{Q}$ ($m \in \mathbb{Z}$, $n \in \mathbb{N}$), is defined by $a^x = \sqrt[n]{a^m}$, and an integer power a^m with $m = -n < 0$ is defined by $a^m = a^{-n} = \frac{1}{a^n}$ (a natural power a^n is the product of n copies of a; $a^0 = 1$). Now we proceed with the proof. It consists of four "steps".

I. The fundamental property of the natural exponential implies that for an arbitrary $\alpha \in \mathbb{R}$ for any $n \in \mathbb{N}$ we have

$$\exp(n\alpha) = \exp\underbrace{(\alpha + \ldots + \alpha)}_{n} = \underbrace{\exp\alpha \cdot \ldots \cdot \exp\alpha}_{n} = (\exp\alpha)^n = \exp^n\alpha. \tag{5}$$

II. Since $\exp(0) = 1$, we obtain

$$\forall x \in \mathbb{R} \quad \exp(x)\cdot\exp(-x) = \exp(x + (-x)) = \exp(0) = 1,$$

whence it follows that

$$\forall x \in \mathbb{R} \quad \exp(-x) = \frac{1}{\exp(x)} = \exp^{-1} x.$$

This relation and property (5) yield

$$\forall x \in \mathbb{R}\ \forall m \in \mathbb{Z} \quad \exp(m\alpha) = \exp^m \alpha. \tag{6}$$

III. Letting $\alpha = \frac{c}{n}$ in (5), we conclude that

$$\exp(c) = \exp(n\alpha) = (\exp\alpha)^n = \left(\exp\left(\frac{c}{n}\right)\right)^n,$$

whence, taking into account that $\exp(x)$ is positive, we obtain

$$\forall c \in \mathbb{R}\ \forall n \in \mathbb{N}\ (n \geq 2) \quad \exp\left(\frac{c}{n}\right) = \sqrt[n]{\exp(c)} = \exp^{\frac{1}{n}}(c). \tag{7}$$

IV. Finally, for an arbitrary rational number $x = \frac{m}{n}$ from (6) and (7) we obtain

$$\exp(x) = \exp\left(\frac{m}{n}\right) = \sqrt[n]{\exp(m)} = \sqrt[n]{\exp(m \cdot 1)} = \sqrt[n]{\exp^m(1)} = \mathrm{e}^{\frac{m}{n}} = \mathrm{e}^x,$$

as required. □

Thus, the *everywhere defined* and strictly monotone function $\exp x$ coincides for all rational values of x with the *exponential function with base* $\mathrm{e} = \exp 1$. Exactly the same reasoning as in the proof of Theorem 3 shows that *for any* $\alpha \in \mathbb{R}$, *for any rational* $x \in \mathbb{Q}$, *we have* $\exp(\alpha x) = a^x$, *where* $a = \exp\alpha$, i.e., the function $\exp(\alpha x)$ at all rational values of x coincides with the *exponential function with base* $a = \exp\alpha$ (do the formal proof on your own). Now recall that for $a > 0$, $a \neq 1$, the *exponential function* $x \mapsto a^x$ *of an arbitrary real argument* x *is defined*[1] *as a strictly monotone function which for all rational values of the argument* $x = \dfrac{m}{n} \in \mathbb{Q}$ *coincides with the rational power* $a^x = \sqrt[n]{a^m}$, and note that *any positive number* a *can be written as* $a = \exp\alpha$, *where* $\alpha = \ln a$ (see formulae (4)). We obtain a fundamental relationship between an arbitrary exponential function and the natural exponential and logarithm: *for any* $a > 0$, *on the entire numerical axis we have the equality*

$$a^x = \exp(\alpha x), \quad \alpha = \ln a. \tag{8}$$

An exponential function with an arbitrary base $a > 0$, $a \neq 1$, is also called an *exponential* (exponential function) *to the base* a and is denoted by $a^x = \exp_a x$ (it reads *"exponential to the base* a *of* x*"*), just as the inverse function of a^x is the *logarithm to the base* a: $\log_a x$. Surely, an arbitrary logarithm can be expressed through the natural logarithm. By equation (8), we have

$$y = a^x = \exp((\ln a)x) \quad \Leftrightarrow \quad (\ln a)x = \ln y \quad \Leftrightarrow \quad x = \log_a y = \frac{\ln y}{\ln a}. \tag{9}$$

Exponential is from Latin *exponens*, *exponentis*, present participles of *exponere*, 'to put up', 'to display'. Originally, the German mathematician (first a monk and then a professor!) Michael Stifel (1487–1567) used the corresponding German word *Exponent*

[1]See a school textbook.

to refer to the exponent of a power, including a fractional power (1553). The term *exponential function* (or *exponential curve*) for the dependence $y = a^x$, $x \in \mathbb{R}$, was introduced by Gottfried Wilhelm Leibniz (1679).

Historically, logarithms appeared in mathematics earlier than exponents of an arbitrary argument. The term *logarithm* (Latin *logarithmus*) was introduced by one of the two "inventors" of logarithms, John Napier (1550–1617), a Scottish baron, a mathematician who worked on simplification and regularization of branches of mathematics of his time. Logarithms "appeared" in his work together with *tables of logarithms*, i.e., actually as a logarithmic dependence (function) (1614). The word *logarithmus* itself came from two Greek words: λογος [logos], in this case 'ratio', and αριθμος [arithmos], 'number', so the whole term can be translated as 'ratio numbers'. This is usually explained by comparing two progressions, the arithmetic 0, 1, 2, 3, 4, ..., and the geometric 1, a, a^2, a^3, a^4, ..., so that the ratio (division) in the second progression corresponds to the difference (subtraction) in the first, in which just the logarithms of powers are written. Napier and even earlier (in 1603; published only in 1620) the Swiss clockmaker and astronomer, Kepler's friend Joost Bürgi (1552–1632) came to logarithms starting from the purely practical challenge of *simplifying calculations*.

The first to realise the practical importance of decimal logarithms (to the base 10) and publish their tables (1617) was the English mathematician, Oxford professor Henry Briggs (1561–1630), a great admirer of Napier. The modern notation for decimal logarithms, $\log_{10} x = \lg x$, was introduced by Augustin-Louis Cauchy.

To simplify calculations, one uses the *fundamental property of logarithms*, i.e., the formula from Theorem 1, which is valid for logarithms to any base. Logarithms allow to reduce multiplication and division to the simpler operations of addition and subtraction (and also reduce the so-called third-stage operations of exponentiation and root extraction to the second-stage operations of multiplication and division). For example, multiplication with the help of logarithms is performed as follows: to find the product AB, we use the table of logarithms to find $\log_a A$ and $\log_a B$, add up these logarithms, and find in the table a number C the logarithm of which is equal to the calculated sum; this is precisely the product AB. The second of the above-mentioned tables should allow us, given the exponent (let it be α), to find the power a^α; i.e., this is the table of values of the exponential function $\exp_a \alpha$, which is usually (following Napier) called the table of *antilogarithms* (Bürgi in 1620 published precisely the table of *antilogarithms*, in contrast to the tables of logarithms of Napier and Briggs). Of course, here a is any number such that $a > 0$, $a \neq 1$ (in practice, the base $a = 10$ is often taken).

Note that before the invention of logarithms, the well-known trigonometric formula

$$\cos A \cos B = \frac{\cos(A+B) + \cos(A-B)}{2}$$

was used in practical calculations; for example, Kepler's "predecessor", the famous astronomer Tycho Brahe (1546–1601) virtuously used this formula. Think of how to calculate the product of two numbers using it (and the tables of cosines and "anticosines", i.e., arccosines).

3.2.5. Solutions of the differential equation $y' = ky$. Equations (8) and (9) and formulae for the derivatives of natural logarithmic and exponential functions imply formulae for differentiating arbitrary logarithms and exponentials:

$$\log_a x = \frac{\ln x}{\ln a} \quad \Rightarrow \quad (\log_a x)' = \frac{1}{\ln a} \cdot (\ln x)' = \frac{1}{x \ln a};$$

$$a^x = \exp(\alpha x) \quad \Rightarrow \quad (a^x)' = \exp'(\alpha x) \cdot (\alpha x)' = \alpha \exp(\alpha x) = \alpha a^x, \quad \alpha = \ln a.$$

When solving differential equations, it is more convenient to use the following variant of the last formula:

$$(\mathrm{e}^{kx})' = k\mathrm{e}^{kx}.$$

Hence, if $y(x) = \mathrm{e}^{kx}$, then $y'(x) = ky(x)$, i.e., this function is a solution of the differential equation (1), $y' = ky$.

Let us return to the question: aren't there any other solutions to this equation? — Of course there are; any function of the form $y = A\mathrm{e}^{kx}$, where A is an arbitrary *multiplicative constant*,[1] satisfies it. — Aren't there "completely different" solutions? — Turns out *there aren't*.

Theorem 4 (on existence and uniqueness of solutions of the differential equation $y' = ky$)**.** *A function $y(x)$ satisfies the differential equation*

$$y' = ky \tag{1}$$

if and only if it can be represented as

$$y(x) = A\mathrm{e}^{kx}, \tag{10}$$

where the multiplicative constant A is uniquely determined by the initial condition

$$y(x_0) = y_0 \quad (\textit{or } y(0) = y_0). \tag{11}$$

Proof. We have just seen that any function of the form (10) satisfies the differential equation (1). Let us prove the converse.

Assume that some function $y(x)$ is a solution of equation (1). Consider the function $h(x) = y(x)\mathrm{e}^{-kx}$. Since $y' = ky$, we have

$$\begin{aligned} h'(x) &= (y(x)\mathrm{e}^{-kx})' = y'(x) \cdot \mathrm{e}^{-kx} + y(x) \cdot (\mathrm{e}^{-kx})' \\ &= ky(x)\mathrm{e}^{-kx} + y(x) \cdot (-k)\mathrm{e}^{-kx} = ky(x)\mathrm{e}^{-kx} - ky(x)\mathrm{e}^{-kx} \equiv 0 \\ &\Rightarrow \quad h(x) = y(x)\mathrm{e}^{-kx} \equiv \mathrm{const} = A \quad \Rightarrow \quad y(x) = A\mathrm{e}^{kx}, \end{aligned}$$

as required.

Now take into account the initial condition (11):

$$y(x_0) = A\mathrm{e}^{kx_0} = y_0 \quad \Rightarrow \quad A = y_0\mathrm{e}^{-kx_0} \quad \Rightarrow \quad y(x) = y_0\mathrm{e}^{-kx_0}\mathrm{e}^{kx} = y_0\mathrm{e}^{k(x-x_0)}.$$

In particular, if $x_0 = 0$, we obtain a very simple formula

$$y(x) = y_0\mathrm{e}^{kx}.$$

The theorem is completely proved. □

[1] That is, a constant which is a *factor*; from Latin *multiplicare*, 'to increase'.

§ 3.3. Exponential Growth and Comparison Theorems

3.3.1. Newton's "exponential" polynomials. The analysis of the differential equation $y' = ky$ in the case $k = 1$ and under the initial condition $y(0) = 1$ carried out in § 3.1 led us, in particular, to a hypothetical representation of the solution in the form of the power series

$$y(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots.$$

The proof of convergence of this series to the desired solution is described step by step in Exercise B to this chapter. Now we will show that partial sums of the series, i.e., the polynomials

$$e_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} = \sum_{k=0}^{n} \frac{x^k}{k!},$$

on the positive semi-axis $x \geq 0$ are *not greater* than the natural exponential e^x. This fact is quite remarkable and important: it follows that *as* $x \to +\infty$, *the exponential function is greater than a polynomial of any degree.*

Lemma 1 (on Newton's approximations of the exponential function)**.** *For any* $n \in \mathbb{N} \cup \{0\}$ *and any* $x > 0$, *we have*

$$\mathrm{e}^x > e_n(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}. \tag{1}$$

Proof. We conduct the proof by induction on $n \geq 0$, noting first that

$$\forall n \in \mathbb{N} \quad e_n'(x) = e_{n-1}(x)$$

(check!). Besides, the proof will exploit the comparison theorem for comparing functions by their derivatives: *if functions f and g that are continuous at a point x_0 and differentiable at $x > x_0$ satisfy the conditions*

(a) $f(x_0) = g(x_0)$;

(b) $\forall x > x_0 \;\; f'(x) > g'(x)$,

then for any $x > x_0$ we have $f(x) > g(x)$ (to prove this, it suffices to apply the monotonicity criterion to the function $h(x) = f(x) - g(x)$, i.e., to show that for $x > x_0$ the value of $h(x)$ is greater than $h(x_0) = 0$).

Induction base. The inequality $\mathrm{e}^x > e_0(x) \equiv 1$ resulting in the case $n = 0$ is certainly valid for any $x > 0$ (this follows since the function e^x increases on the entire number axis, which in turn follows since the derivative $(\mathrm{e}^x)' = \mathrm{e}^x$ is positive).

Induction step. Assume that $\forall k < n \;\; \forall x > 0 \;\; \mathrm{e}^x > e_k(x)$; let us prove that $\forall x > 0 \;\; \mathrm{e}^x > e_n(x)$. Notice the following:

(1) For $x = 0$ we have $\mathrm{e}^x = \mathrm{e}^0 = e_n(x) = e_n(0) = 1$;

(2) For $x > 0$ by the assumption we have $(\mathrm{e}^x)' = \mathrm{e}^x > e_{n-1}(x) = (e_n(x))'$.

Hence, by the *comparison theorem*, $\forall x > 0\ \ e^x > e_n(x)$, as required.

By the *mathematical induction principle*, we conclude that $\forall n \in \mathbb{N} \cup \{0\}$ $\forall x > 0\ \ e^x > e_n(x)$. □

Now we will apply the simplest non-trivial variant of this estimate,

$$\forall x > 0 \quad e^x > e_2(x) = 1 + x + \frac{x^2}{2}, \tag{2}$$

to answer a question often encountered in practice.

3.3.2. How to distinguish between exponential and power-law growth? Suppose that for each object from a certain set (of objects), two of its numerical characteristics, x and y, are measured (for example, for stars, one can measure their temperatures and luminosities). Then the results of the measurements are plotted in $(x; y)$ coordinates as a set of points (for each object with parameters x and y, the point $M(x; y)$ in the Oxy plane is considered). It is quite possible that these points are located on some curve (or something like a curve; in the above-mentioned example of stellar temperatures and luminosities, the so-called *Hertzsprung–Russell diagram* is obtained, which plays a crucial role in astrophysics and stellar astronomy).

Ejnar Hertzsprung (1873–1967) was a Danish astronomer; in 1911, by his observations, he plotted the "colour" (spectrum, or temperature) vs. "luminosity" diagram for stars in the Hyades and Pleiades clusters. In 1913, the American astronomer Henry Norris Russell (1877–1957) constructed a similar diagram for a group of stars in the "neighbourhood" of the Sun, quite consistent with the Hertzsprung diagram. Based on this diagram, he put forward his concept of stellar evolution (1913–1914).

The question is what is the dependence of y on x, for instance, assuming that there are only *two possibilites*: *exponential dependence*

(I) $$y = Ae^{kx} = Aa^x, \quad k > 0 \quad (a > 1), \quad A > 0,$$

or *power-law dependence*

(II) $$y = Cx^\alpha, \quad \alpha > 0 \quad (\text{or } \alpha > 1, \text{ or even } \alpha \gg 1.^{1})$$

As can be seen, for example, from Fig. 37, at some distance from the origin the dependences (I) and (II) are difficult to distinguish from each other. What can we do?

The answer is quite simple: we need to pass to *logarithmic scales*, i.e., consider points not in the $(x; y)$ axes, but in the $(z; w) = (\ln x; \ln y)$ coordinates. Logarithmising the dependencies (I) and (II) (e.g., to the base e, although this is not essential), we arrive at the dependencies between the

[1] The symbol $\gg$ in physics, and sometimes in mathematics, is used to indicate that one quantity is "much greater" than another.

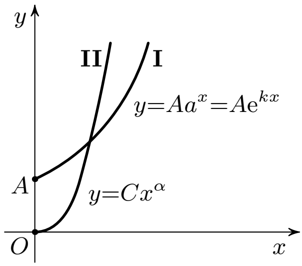


Fig. 37.

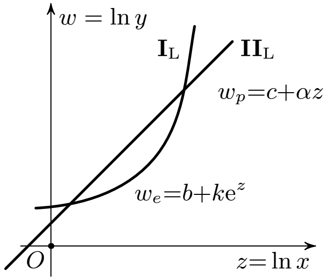


Fig. 38.

corresponding logarithms:

$$(\mathrm{I}_L) \qquad w = \ln y = \ln A + kx = b + ke^{\ln x} = b + ke^z \ (= w_e),$$
$$(\mathrm{II}_L) \qquad w = \ln y = \ln C + \alpha \ln x = c + \alpha z \ (= w_p).$$

These two dependencies are now easily distinguishable: for the power-law dependence w_p, the points are located on a straight line, whereas the exponential dependence in the logarithmic scales $(z; w)$ will still be exponential, $w_e = b + ke^z$; for sufficiently large values of the logarithmic variable z, its graph passes above the graph of the linear function $w_p = c + \alpha z$ (Fig. 38). To prove this, consider the difference $w_e - w_p$ and estimate it from below using inequality (2):

$$w_e - w_p = b - c + ke^z - \alpha z > b - c + k\left(1 + z + \frac{z^2}{2}\right)\alpha z = p_2(z).$$

Since $p_2(z)$ is a quadratic function with a positive coefficient $\frac{k}{2}$ of z^2, for sufficiently large values of z we have $p_2(z) > 0$ and, accordingly, $w_e - w_p > 0$, as claimed. In the general case, investigating whether points lie on the same straight line within a permissible error margin is not particularly difficult. However, we will refrain from delving into these questions, which belong to a special branch of mathematics, mathematical statistics.

3.3.3. What is exponential growth at infinity? When considering an exponential dependence

$$y = Ae^{kx} = Aa^x$$

of one variable on another in the case of $k > 0$ (or $a > 1$) and $A > 0$, we speak of the *exponential growth* of $y(x)$ as $x \to +\infty$ (or simply "at $+\infty$"). This is an essential characteristic of the dependence, meaning that the exponent $e^{kx} = a^x$ *grows as* $x \to +\infty$ *faster than any power function* x^α, in the sense specified in the following theorem.

Theorem 1 (first comparison theorem)**.** *For any $a>1$ and any $\alpha>0$, we have*

$$\lim_{x\to+\infty}\frac{x^\alpha}{a^x}=0.$$

Proof. Method I. Since we consider the limit at $+\infty$, we may treat only values $x>0$. Choose a natural number $n>\alpha$ (e.g., $n=[\alpha]+1$). Taking into account that if $x>0$, then $kx>0$, in the fraction

$$\frac{x^\alpha}{a^x}=\frac{x^\alpha}{\mathrm{e}^{kx}}$$

we replace the denominator e^{kx} by a smaller (according to the first lemma on Newton's approximations; see inequality (1)) expression

$$e_n(kx)=1+kx+\frac{k^2x^2}{2!}+\ldots+\frac{k^nx^n}{n!}.$$

The fraction will only increase, and we get

$$0<\frac{x^\alpha}{\mathrm{e}^{kx}}<\frac{x^\alpha}{1+kx+\frac{k^2x^2}{2!}+\ldots+\frac{k^nx^n}{n!}}<\frac{x^\alpha}{\frac{k^nx^n}{n!}}=\frac{n!}{k^n}x^{\alpha-n}.$$

Since the exponent $\alpha-n$ is negative by the choice of n, we obtain

$$\lim_{x\to+\infty}\frac{n!}{k^n}x^{\alpha-n}=0,$$

and by the squeeze theorem,[1]

$$\lim_{x\to+\infty}\frac{x^\alpha}{\mathrm{e}^{kx}}=0. \qquad \square$$

Method II. The second way of proving the theorem is to directly estimate the ratio of the power function and the exponential function.

Lemma 2. *If $a>1$, then for any $\alpha>0$ we have*

$$\exists M\in\mathbb{R}\quad \forall x>0\quad \frac{x^\alpha}{a^x}\le\frac{M}{x}. \tag{3}$$

Proof. Multiplying both parts of inequality (3) by x, we obtain that it suffices to prove that the function $\varphi(x)=x^{\alpha+1}\cdot a^{-x}$ is bounded from above on the interval $x>0$. To this end, let us examine the function $\varphi(x)$ for monotonicity. We have

$$(\varphi(x))'=(x^{\alpha+1}\cdot a^{-x})'=(\alpha+1)x^\alpha\cdot a^{-x}+x^{\alpha+1}\cdot\ln a\cdot a^{-x}\cdot(-1)$$
$$=x^\alpha a^{-x}((\alpha+1)-\ln a\cdot x).$$

Thus, the sign of the derivative $\varphi'(x)$ is determined by the sign of the expression in the parentheses, $L(x)=(\alpha+1)-\ln a\cdot x$, which is a linear

[1] *If three functions $\alpha(x)$, $\varphi(x)$, and $\beta(x)$ satisfy the inequalities $\alpha(x)<\varphi(x)<\beta(x)$, and the functions α and β* (minorant and majorant) *have the same limit* (*at a point or at infinity*), *then φ has the same limit too.*

function with a negative slope vanishing at $x=x_0=\dfrac{\alpha+1}{\ln a}$. Hence, when $x<x_0$, we have $\varphi'(x)>0$ and the function $\varphi(x)$ increases, and when $x>x_0$, we have $\varphi'(x)<0$ and the function $\varphi(x)$ decreases. Taking into account that $\varphi(x)$ is a continuous function, we obtain

$$\max_{(0,+\infty)} \varphi(x)=\varphi(x_0)=M,$$

so $\varphi(x)\le M$ for this M, for any $x>0$. □

Of course, the obtained estimate (3) immediately implies the claim of the theorem (in this case we can also refer to the squeeze theorem).

Control question: where in these proofs did we use that $k>0$ (or $a>1$)?

Corollary 1. *An exponential function q^x with base $q<1$ $(q>0)$ tends to zero* (*or decreases*) *at $+\infty$ faster than any power function with a negative exponent $x^{-\alpha}$*, $\alpha>0$, *in the sense that under these conditions*

$$\lim_{x\to+\infty}\frac{q^x}{x^{-\alpha}}=0. \tag{4}$$

Proof. Transforming the expression under the limit in (4), we obtain the fraction

$$\frac{q^x}{x^{-\alpha}}=\frac{x^\alpha}{\left(\frac{1}{q}\right)^x},$$

which by the first comparison theorem tends to zero as $x\to+\infty$ (note that $\frac{1}{q}>1$). □

Another consequence of the first theorem gives an essentially new statement, which in itself deserves to be called a theorem; we pass to it.

3.3.4. Comparison of power and logarithmic functions as $x\to+\infty$. Besides the problem of distinguishing between exponential and power-law dependences, in practice it is often needed to distinguish not “fast” but, on the contrary, slow growth. It is crucial to be able to distinguish between *logarithmic growth* by the law

(III) $\qquad y=A\log_a x=A\cdot\dfrac{\ln x}{\ln a}=B\ln x \quad (A>0,\ a>1;\ B>0)$

and *power-law growth*, like an arithmetic root, by the law

(IV) $\qquad y=Cx^\alpha \quad (\alpha>0;$ of interest is the case $\alpha<1)$.

In both of these cases (but only for $\alpha<1$), the functions $y(x)$ *increase* but *their growth rates decrease* (Fig. 39).

This situation can be clarified by switching to the logarithmic scale on the Ox axis, i.e., by considering these dependencies not in the $(x;y)$ axes but in the coordinates $(u;v)=(\ln x;y)$ (the scale of y does not change, but

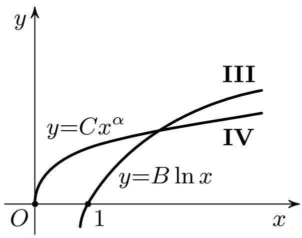


Fig. 39.

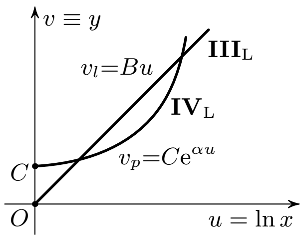


Fig. 40.

for convenience we have renamed the variable y by v). Dependences (III) and (IV) in the coordinates $(u; v)$ are written as

$$(\mathrm{III}_L) \qquad v = v_l = Bu$$

and

$$(\mathrm{IV}_L) \qquad v = v_e = C\mathrm{e}^{\alpha u} \quad (\text{since } u = \ln x \ \Rightarrow\ x = \mathrm{e}^u).$$

Of course, the linear dependence v_l can easily be distinguished from the exponential dependence v_e (Fig. 40), for example, based on the first comparison theorem: the second function grows at $+\infty$ "much faster" than the first (linear) function. Accordingly, we obtain that the logarithmic function $\log_a x$ (or $\ln x$) *grows as* $x \to +\infty$ *slower than any power function* x^α, $\alpha > 0$, in the sense specified in the following theorem.

Theorem 2 (second comparisom theorem)**.** *For any* $a > 1$ *and any* $\alpha > 0$, *we have*

$$\lim_{x\to+\infty} \frac{\log_a x}{x^\alpha} = 0.$$

Proof. Clearly, it suffices to show that for any $\alpha > 0$ we have

$$\lim_{x\to+\infty} \frac{\ln x}{x^\alpha} = 0. \tag{5}$$

Let us introduce a new variable $u = \ln x$; then $x = \mathrm{e}^u$, and since $x \to +\infty \Leftrightarrow u \to +\infty$, the limit (5) is equal to the limit

$$\lim_{u\to+\infty} \frac{u}{\mathrm{e}^{\alpha u}},$$

which is zero by the *first comparison theorem.* □

Remark. It is tempting to try to prove the "second theorem" using a direct estimate of the type of inequality (3) from the second proof method of the first comparison theorem: for some $M > 0$, we have

$$\forall x > 1 \quad \frac{\ln x}{x^\alpha} \le \frac{M}{x}, \tag{6}$$

or, in other words, *the function* $\varphi(x)=x^{1-\alpha}\cdot\ln x$ *is bounded on the interval* $(1,+\infty)$. However, it turns out that this is not true in the case of $\alpha<1$ (see the exercises). Nevertheless, the estimate (6) can be replaced by an estimate of the form

$$\forall x>1 \quad \frac{\ln x}{x^\alpha}\le\frac{M}{x^\beta}$$

for some $\beta>0$ (try to do it on your own!), and then the "second theorem" can as well be derived from the squeeze theorem.

3.3.5. Comparison of power and logarithmic functions as $x\to 0+$. In solving various kinds of problems, another question often arises, that of comparing the growth of a power function with a negative exponent, $x^{-\alpha}=\frac{1}{x^\alpha}$, $\alpha>0$, and an arbitrary logarithmic function $\log_a x$ *at zero*, i.e., as $x\to 0+$ (as x tends to zero from the right). Each of these functions tends to infinity (if $a>1$, then $\log_a x\to-\infty$) as $x\to 0+$, but which one is *faster* (say in absolute value)?

Theorem 3 (third comparison theorem)**.** *For any* $a>0$, $a\neq 1$, *and any* $\alpha>0$, *we have*

$$\lim_{x\to 0+}\frac{\log_a x}{x^{-\alpha}}=\lim_{x\to 0+}x^\alpha\log_a x=0;$$

i.e., any logarithmic function $\log_a x$ *grows* (*tends to infinity in magnitude*) *as* $x\to 0+$ *slower than any power function* $x^{-\alpha}=\frac{1}{x^\alpha}$, $\alpha>0$.

Proof. Introduce a new variable $z=\frac{1}{x}$; then $x\to 0+\ \Leftrightarrow\ z\to+\infty$, the limit in question can be rewritten as

$$\lim_{z\to+\infty}\left(\frac{1}{z}\right)^\alpha\log_a\frac{1}{z}=-\lim_{z\to+\infty}\frac{\log_a z}{z^\alpha},$$

and by the *second comparison theorem*,[1] it is zero. □

§ 3.4. Exponential Models

Now we are able to write down the evolution law $x=x(t)$ for any process modelled with a (homogeneous) linear differential equation of the form (1) $x'(t)=kx(t)$.[2] According to Theorem 4 in § 3.2, if $x(0)=X_0$, then x changes by the *exponential law*: $x(t)=X_0\mathrm{e}^{kt}$. According to what was proved in § 3.3, for $k>0$ this law gives a very fast *exponential* growth of $x(t)$, and for $k<0$, a very fast *exponential* decay. However, if the value of the coefficient k is close to zero, then the "decay" or "bursting" will proceed relatively slowly for

[1] Formally, the condition $a>1$ was required in that theorem, but this requirement is not essential due to the identity $\log_a N=-\log_{1/a}N$.

[2] We have again returned to the main variable t, i.e., *time*.

some time. Besides, it often happens that the linear evolution law $x' = kx$ adequately reflects (models) the process only in some range of variation of x. For example, the *Malthus model* of population growth $N'(t) = \alpha N(t)$ (see § 1.1) is applicable only under *favourable conditions*, when there is enough space and resources to consume, so that once the population size N exceeds some *critical level*, the linear Malthus law becomes inapplicable; it has to be replaced by some *non-linear* law. We will consider examples of non-linear processes in Ch. VI, and in this section we confine ourselves to linear processes, i.e., *exponential models.*

3.4.1. Example: radioactive decay. In § 3.1 we wrote the *fundamental law of radioactive decay* as a linear differential equation of the form (1): if $M(t)$ is the number of atoms of a radioactive element at time t (recall that it is assumed to be *large*) and ω is the *decay probability*, then we have

$$M'(t) = -\omega M(t). \tag{1}$$

From Theorem 4 of § 3.2 we obtain the *radioactive decay law*, now in the form of a time dependence of M: if $M(0) = M_0$, then

$$M(t) = M_0 \mathrm{e}^{-\omega t}. \tag{2}$$

In practice, instead of the decay probability, the notion of *half-life* is used.[1]

Proposition 1. *The number of atoms of a radioactive substance is* halved *every same period of time, called the* half-life, *regardless of the initial number of atoms.*

Proof. The corresponding time interval can be found by solving the equation

$$M(t) = M_0 \mathrm{e}^{-\omega t} = \frac{M_0}{2} \quad \Leftrightarrow \quad \mathrm{e}^{-\omega t} = \frac{1}{2}$$
$$\Leftrightarrow \quad -\omega t = -\ln 2 \quad \Leftrightarrow \quad t = T_{1/2} = \frac{\ln 2}{\omega}$$

Thus, the *half-life* $T_{1/2}$ does not depend on the initial amount M_0 of the substance and is related to the decay probability as $T_{1/2} = (\ln 2)/\omega$. □

Note also (in the spirit of the considerations in § 3.3) that to detect the exponential dependence (2), it is reasonable to use the *logarithmic scale* on the M axis; in this case we get a *linear dependence*

$$\ln M = \ln M_0 - \omega t,$$

the slope of which is precisely (but with the minus sign) the decay probability.

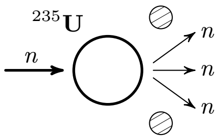
[1] The law of radioactive decay was formulated by Pierre Curie in 1903; he also introduced the notion of "half-life".

3.4.2. Example: nuclear fission (chain reaction). In the late 1930s, the Joliot-Curie couple in France, followed by Hahn and Strassmann in Germany, and also Enrico Fermi in Italy and then in the United States, experimentally discovered that when a "slow" neutron hits a uranium-235 nucleus, it splits into two large fragments and, in addition, emits 2 to 3 new neutrons, resulting in a large energy release, of the order of 6×10^{10} J/g (per 1 g of split uranium). The fragments first fly in opposite directions at velocities of about 10^9 cm/s and then decelerate with their kinetic energy being converted into heat. The aforementioned discoverers also "theoretically saw" the possibility of developing in uranium the so-called *chain reaction* of nuclear fission. Let us describe the simplest mathematical model of such a reaction, which, however, gives a clear picture of the processes occurring in nuclear fission.

Irène Curie (1897–1956) was a French physicist and radiochemist, daughter of Pierre Curie and Marie Skłodowska-Curie, winners of the 1903 Nobel Prize in Physics and founders of the science of radioactivity. Frédéric Joliot (1900–1958) was a French physicist. After his marriage in 1926 to Irène Curie, from 1934 the couple began to sign their joint works with the double surname Joliot-Curie. In 1935 they won the Nobel Prize in Chemistry for the synthesis of new radioactive elements. During the years of occupation they were active participants in the resistance movement, and after the war they were active in the World Peace Council.

Otto Hahn (1879–1968) was a German radiochemist and physicist who studied under Ernest Rutherford (1904–1905). Fritz Strassmann (1902–1980) was a German physicist and chemist. Together with Hahn, he discovered in 1938 the fission of uranium nuclei by "bombarding" them with neutrons; Hahn was awarded the 1944 Nobel Prize in Chemistry for this discovery.

Enrico Fermi (1901–1954) was an eminent theoretical and experimental physicist. Born in Rome, he worked for Max Born and Paul Ehrenfest in Germany (1923–1924) and then in Rome and Florence. In 1938 he emigrated to the USA, where he took part in the work on the atomic bomb. He built the *first* nuclear reactor (on 2 December 1942 he launched it, for the first time obtaining a self-sustained chain reaction; we will consider the mathematical model of the reactor below, at the end of this section). In 1938, he won the Nobel Prize in Physics for the discovery of nuclear reactions caused by the irradiation of nuclei with slow neutrons.



$N(t) \to N + \Delta^+ N$

Fig. 41.

Suppose we have a lump of uranium-235 (^{235}U) in which at time t there are $N = N(t)$ slow neutrons, and the nuclear fission described above *can take place* at collision of these neutrons with the uranium, resulting in an increase in the number of neutrons, including the slow neutrons we are interested in (Fig. 41).

Each of the neutrons with some probability enters the uranium-235 nucleus (the ^{235}U nucleus is a sphere of radius about 10^{-12} cm), which leads to the fission of the nucleus and to an increase in the number of neutrons, including slow ones, so that in the

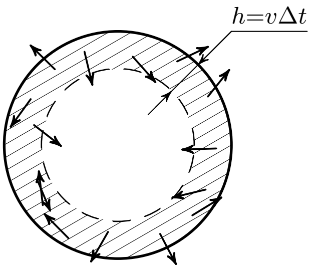


Fig. 42.

time interval from t to $t+\Delta$ the number $N(t)$ of neutrons increases by

$$\Delta^{+}N \approx \alpha N \cdot \Delta t. \tag{3}$$

However, some of the neutrons escape through the surface of the uranium lump and can no longer take part in nuclear fission. Namely, those neutrons escape which are close enough to the surface of the uranium lump and have the corresponding direction of motion. Let a lump of uranium-235 have the shape of a sphere of radius r. If v is the average velocity of neutrons (we mean the absolute value of velocity), then within the time interval from t to $t+\Delta$ only neutrons from the spherical layer of width $v \cdot \Delta t$ (Fig. 42) can escape from the sphere, and only a certain fraction $p:1$ of them, those that have the necessary directions of velocity to escape. The volume of the layer is $V_{\mathrm{L}} \approx S_{\mathrm{Sph}} \cdot v \cdot \Delta t$, where S_{Sph} is the surface area of the sphere. Since the density of neutrons in the sphere is

$$\rho = \frac{N}{V_{\mathrm{Sph}}}$$

(V_{Sph} being the volume of the uranium sphere), the number of neutrons in this layer is

$$\rho \cdot V_{\mathrm{L}} \approx \frac{N}{V_{\mathrm{Sph}}} \cdot S_{\mathrm{Sph}} \cdot v \cdot \Delta t = \frac{S_{\mathrm{Sph}}}{V_{\mathrm{Sph}}} \cdot v \cdot N\Delta t,$$

and the number of escaping neutrons is a fraction $p:1$ of this number, i.e.,

$$\Delta^{-}N \approx p\frac{S_{\mathrm{Sph}}}{V_{\mathrm{Sph}}} \cdot v \cdot N\Delta t = \frac{\beta}{r} \cdot N\Delta t, \tag{4}$$

where β, like α above, is a characteristic of the $^{235}\mathrm{U}$ substance itself; this constant is obtained by the following calculation:

$$p\frac{S_{\mathrm{Sph}}}{V_{\mathrm{Sph}}} \cdot v = pv\frac{4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{3pv}{r} = \frac{\beta}{r}.$$

Based on formulae (3) and (4) for the neutron growth and escape, we can write a final formula for the increment of the number N of neutrons over the time from t to $t+\Delta t$:

$$\Delta N(t,\Delta t) = N(t+\Delta t) - N(t) = \Delta^+ N - \Delta^- N$$
$$\approx \alpha N\Delta t - \frac{\beta}{r} N\Delta t = \Big(\alpha - \frac{\beta}{r}\Big) N\Delta t,$$

whence, after dividing by Δt and passing to the limit as $\Delta t \to 0$, we finally obtain the *differential equation for the number of slow neutrons in the spherical lump of* ^{235}U:

$$N'(t) = \Big(\alpha - \frac{\beta}{r}\Big) N(t).$$

Let us analyse this equation.

We write it in the form

$$N'(t) = k(r)N(t), \quad k(r) = \alpha - \frac{\beta}{r} \tag{5}$$

(recall that the coefficients α and β are characteristics of the *substance* itself, say ^{235}U, but not of a particular "piece" of it, a uranium sphere). Solutions of equation (5) are always of the form

$$N(t) = N_0 \cdot \mathrm{e}^{k(r)t} \quad (N_0 = N(0) \text{ being the initial number of neutrons}),$$

but in the case $k(r) < 0$ these solutions exponentially — hence, very fast — decrease, the "neutron fire" goes out; and in the case $k(r) > 0$, on the contrary, the number of neutrons, and along with it the number of nuclear fissions, and also the amount of energy released as heat — everything *increases exponentially* (*very fast*) (Fig. 43). This is exactly the "chain reaction" that leads to an atomic explosion. The condition for the explosion to occur means that

$$k(r) = \alpha - \frac{\beta}{r} > 0 \quad \Leftrightarrow \quad \alpha > \frac{\beta}{r} \quad \Leftrightarrow \quad r > \frac{\beta}{\alpha} = r_0,$$

i.e., the radius of the uranium sphere must be *greater* than some *critical radius* $r_0 = \frac{\beta}{\alpha}$; accordingly, the mass of the fissile material must be *greater than the critical mass*

$$m_{\text{cr}} = \rho \cdot V_{\text{Sph}}(r_0) = \rho \cdot \frac{4}{3}\pi r_0^3$$

(ρ is the density of the fissile material). For instance, if we are speaking of uranium-235, then $2r_0 \approx 17$ cm and $m_{\text{cr}} \approx 50$ kg.

We will not discuss this situation further (see the exercises), but rather turn to the case of $k(r) < 0$ (respectively, $r < r_0$ or $m < m_{\text{cr}}$), which apparently first interested Enrico Fermi, "fire goes out". But when the fire goes out, you throw fuel into it. So let us try to do the same thing with the "neutron fire": take a piece of uranium of mass less than the critical mass, i.e., for which $k(r) = -\gamma < 0$, and connect to it an external neutron

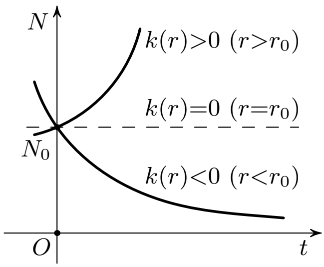


Fig. 43.

source (usually, radioactive plutonium). We will assume that the external source supplies neutrons at a constant rate of q neutrons per unit time. What would happen to the "neutron fire", the process of nuclear fission, in a *subcritical* mass of our material?

To answer this question, let us write for this case a differential equation that governs the number of neutrons in a subcritical mass in the presence of an external neutron source. Obviously, taking into account neutrons from an external source leads to the necessity to add to the right-hand side of equation (5) a constant term (independent of either r or t) equal to the rate q of neutron emission by the external source. Since we have denoted $k(r)$ by $-\gamma$, the equation is written as

$$N'(t) = -\gamma N(t) + q, \quad q > 0, \quad \gamma = \frac{\beta}{r} - \alpha > 0. \tag{6}$$

This is also a linear equation, though *non-homogeneous*, so before investigating solutions of the differential equation (6), let us consider *general approaches to solving non-homogeneous first-order linear equations.*

3.4.3. Non-homogeneous linear differential equation $y' = ky + f(x)$. It is convenient to consider directly the more general case where the homogeneity of the equation $y' = ky$ is broken by adding to its right-hand side not a constant but an *arbitrary function* $f(x)$, since the reasoning given below is also applicable in this case. Thus, consider a differential equation of the form

$$y' = ky + f(x), \tag{7}$$

where $k \in \mathbb{R}$ and $f(x)$ is the given function. The idea is simple: we consider *two* solutions of the non-homogeneous equation and, by subtracting one solution from the other, get rid of the inhomogeneity. This technique is useful not only in the case at hand, but also in similar situations in algebra, number theory, etc.

So, let $y(x)$ and $y_1(x)$ be two solutions of the non-homogeneous equation (7):

$$y'(x) = ky(x) + f(x),$$
$$y_1'(x) = ky_1(x) + f(x).$$

By subtracting the second equation from the first, we obtain that the inhomogeneity, the term $f(x)$, "disappears", and the difference $y(x) - y_1(x)$ of the solutions now satisfies a homogeneous equation of the form (1):

$$(y - y_1)' = k(y - y_1).$$

Hence (by Theorem 4 in § 3.2), this difference is $y - y_1 = Ae^{kx}$, and thus we have proved the following.

Theorem 1 (d'Alembert's theorem; fundamental property of solutions of a non-homogeneous linear differential equation)**.** *Let* $y_1(x)$ *be any* particular *solution of a non-homogeneous linear differential equation*

$$y' = ky + f(x).$$

Then the general (*i.e., any other*) *solution of this equation can be represented as*

$$y(x) = y_1(x) + Ae^{kx}$$

(in other words, the *general* solution of the *non-homogeneous* equation (7) is a sum of any *particular* solution of this equation and the *general* solution of the corresponding *homogeneous* equation (1)).

Hence, the problem of finding *all* solutions of the differential equation (7) reduces to finding some *particular* solution to it, which for each given function $f(x)$ can often be just appropriately chosen. However, a more detailed consideration of this equation in the case where $f(x) \neq \text{const}$ (and a very efficient *method* of finding particular solutions in some cases) is postponed to the next section; for now, let us concentrate on equations of the type (6), with inhomogeneity being a constant.

Jean le Rond d'Alembert (1717–1783) was a French mathematician, mechanic, and philosopher, by background a lawyer. Together with Denis Diderot, he worked until 1757 on the famous *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers* (Encyclopaedia, or a Systematic Dictionary of the Sciences, Arts, and Crafts), where in his article "Dimension" he first put forward the idea of time as a *fourth dimension.* He studied mechanics (dynamics), hydrodynamics, celestial mechanics, differential equations. In 1766 he established the above theorem in a somewhat more general case (see the exercises). He was also the first to write the string vibration equation $u''_{tt} = u''_{xx}$ (see Ch. VI) and its solutions in the form $u(x,t) = f(x+t) + g(x-t)$. He was also engaged in literary activity, in questions of musical theory and musical aesthetics... He was a member of the Paris Academy of Sciences and Académie Française (the latter was founded by Cardinal Richelieu in 1635 and was an association of prominent cultural, scientific, and political figures in France; D'Alembert was elected to the permanent membership of the "Forty Immortals" of this Academy, which also included Racine, Voltaire, Hugo, Pasteur, Frans...).

3.4.4. Non-homogeneous linear differential equation $y' = ky + b$ (const). Consider the differential equation

$$y' = ky + b, \tag{8}$$

where k and b are given constants. By d'Alembert's theorem, *the general solution of equation* (8) *is*

$$y(x) = y_1(x) + Ae^{kx},$$

where A is an arbitrary constant (it is uniquely determined by the initial condition, for instance, at a point $x = 0$: $y(0) = y_0$) and $y_1(x)$ is *any* particular solution of equation (8).

Since homogeneity in this case is broken by adding a constant, it is natural to look for a particular solution also in the form of a constant (in fact, this is a general principle — to "compensate" some function, i.e., "inhomogeneity", you should try to substitute into the equation a function of the same type or having a derivative "of the same type"!). Let us substitute the constant $y_1 \equiv C$ into equation (8). We obtain

$$y_1' = C' = 0, \quad ky_1 + b = kC + b; \quad \text{hence,} \quad 0 = kC + b \quad \Rightarrow \quad C = -\frac{b}{k}.$$

Therefore, we can take the constant function $y_1 = -\frac{b}{k}$ as a particular solution of equation (8), so *the general solution of equation* (8) *can be written as*

$$y(x) = -\frac{b}{k} + Ae^{kx}. \tag{9}$$

Now let us show how the same result can be obtained quite shortly, but using a "trick". Note that equation (9) can be written in a slightly different form:

$$y + \frac{b}{k} = Ae^{kx},$$

and this means that the function $z = y + \frac{b}{k}$ satisfies the equation $z' = kz$, or

$$\left(y + \frac{b}{k}\right)' = k\left(y + \frac{b}{k}\right). \tag{10}$$

But this equation can be *immediately* obtained from the original equation (8); it suffices to factor out the coefficient k on the right-hand side and add the constant $\frac{b}{k}$ to the left-hand side "under the derivative sign".

Hence the following reasoning. After reducing equation (8) to the form (10) and *changing the variable*, i.e., introducing a new function

$$y + \frac{b}{k} = z, \tag{11}$$

we obtain that this function $z=z(x)$ satisfies the equation $z'=kz$. Hence, $z(x)=Ae^{kx}$, and formula (11) immediately yields (9). (The "method of change" we have used is quite special, but it is not lacking in elegance and brevity.)

Anyway, we have found the general form (9) of the solution of the differential equation (8). Now let us take into account the *initial condition* $y(0)=y_0$. Substituting it into (9), we obtain

$$3y(0)=y_0=-\frac{b}{k}+Ae^{k\cdot 0}=\frac{b}{k}+A \quad\Rightarrow\quad A=y_0+\frac{b}{k}$$

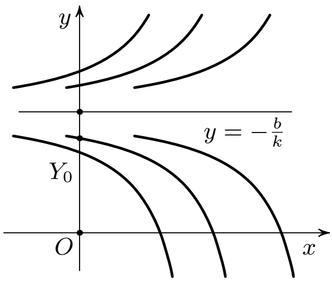
$$\Rightarrow\quad y(x)=-\frac{b}{k}+\Big(y_0+\frac{b}{k}\Big)e^{kx}, \quad (12)$$

or

$$y(x)=y_0e^{kx}+\frac{b}{k}(e^{kx}-1). \quad (13)$$

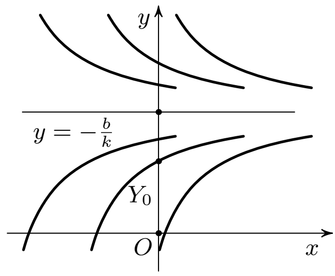
Both representations of the solution, (12) and (13), can be useful, depending on a particular problem.

It is also useful to have an idea of what the solution graphs look like. For each particular value of $k\neq 0$, *all* solution graphs have *the same* left (for $k>0$) or right (for $k<0$) *asymptote*, as is shown in Figs. 44 and 45. The easiest way to see this is to represent the solutions in the form (9); the asymptote is the line $y=-\frac{b}{k}$, *regardless of the initial conditions* (!).



$y(x)=-\frac{b}{k}+Ae^{kx},\ k>0$

Fig. 44.



$y(x)=-\frac{b}{k}+Ae^{kx},\ k<0$

Fig. 45.

Now we return to equation (6), which describes the change in the number of neutrons in a subcritical mass of a fissile material (e.g., ^{235}U) in the presence of an external neutron source.

3.4.5. Example: "Atomic reactor" equation. Recall what equation (6) looks like:

$$N'(t)=-\gamma N(t)+q, \quad q>0, \quad \gamma=\frac{\beta}{r}-\alpha>0.$$

Applying to it the results obtained above, we first find the *general law* of change in the number of neutrons,

$$N(t) = \frac{q}{\gamma} + A\mathrm{e}^{-\gamma t},$$

and then also a formula for $N(t)$ under initial condition $N(0) = N_0$: in this case,

$$\frac{q}{\gamma} + A = N_0 \quad \Rightarrow \quad N(t) = \frac{q}{\gamma} + \left(N_0 - \frac{q}{\gamma}\right)\mathrm{e}^{-\gamma t}.$$

Both formulae imply that, *regardless of the initial number of neutrons* N_0 (usually it is small), the number of neutrons *exponentially* (very fast!) tends to the limit value of $N_\infty = \frac{q}{\gamma}$. Hence very important conclusions follow.

First, we obtain that when an external source of neutrons is added, the total number of neutrons inside a uranium sphere of subcritical mass, and therefore the amount of *atomic energy* released as a result of nuclear fission, *stabilises*, i.e., quickly becomes almost constant. This energy, released as heat, can then be used, for example, to produce electricity. This is the process underlying the principle of operation of nuclear reactors and nuclear power plants. In this case, one speaks of a *controlled nuclear reaction*, as opposed to an *uncontrolled* nuclear reaction spontaneously occurring in a *supercritical* mass of fissile matter, resulting in an atomic explosion.

Second, if the external source supplies only q neutrons per unit time, then inside the uranium sphere, soon after the reaction starts, there will be about $N_\infty = q \cdot \frac{1}{\gamma}$ neutrons, which in the case of $\frac{1}{\gamma} > 1$, i.e., $\gamma < 1$, is greater than q. The ratio

$$\frac{N_\infty}{q} = \frac{1}{\gamma} = \varkappa$$

is called the *neutron multiplication factor* in the reactor; neutrons supplied by an external source "multiply" inside the subcritical mass. Note that the multiplication factor is the larger (i.e., the more "efficient" the atomic reactor) the smaller the factor $\gamma = \frac{\beta}{r} - \alpha$. But when γ is zero, this means that r is the critical radius, and when γ is close to zero, the radius is close to the critical radius, i.e., the subcritical mass becomes close to the critical mass! This is a very important conclusion: it turns out that the more efficient a nuclear reactor, the greater the risk of approaching the supercritical mass and, consequently, the greater the danger of explosion.

§ 3.5. Linear Differential Equations with Variable Coefficients

In § 3.1 and § 3.4 above, we have considered and solved linear differential equations $y' = ky\,[+\,b]$ with *constant* coefficients k and b and analysed the behaviour of their solutions. Solving such an equation with a "free term"

$b \neq 0$ was based on *d'Alembert's theorem* on an arbitrary solution of a non-homogeneous equation of a more general form $y' = ky + f(x)$, i.e., on representing an *arbitrary* solution in the form $y(x) = y_1(x) + Ae^{kx}$, where $y_1(x)$ is any *particular* solution of the equation and the second term Ae^{kx} is the general solution of the corresponding homogeneous equation $y' = kx$.

Our studies did not actually use *integration of functions*, i.e., the ability to write solutions of a differential equation of the form $y' = f(x)$ through the integral by *Barrow's second formula*: if $y(x_0) = y_0$, then

$$y(x) = y_0 + S(f)_{x_0}^{x} = y_0 + \int\limits_{x_0}^{x} f(u)\, du$$

(see Sec. 2.3.5). Note that, regardless of the possibility (or impossibility) of expressing the above integral through *elementary functions*, if solutions of a differential equation can be *written through integrals* of such functions, then the equation is considered to be *solvable*, as we say, *in quadratures.*[1]

Reduction of differential equations to finding integrals is one of the most important *general ideas* for solving differential equations. Now we will consider how it can be applied to *first-order linear differential equations with variable coefficients*, i.e., equations of the general form

$$y' = k(x)y + f(x) \quad (y'(x) \equiv k(x)y(x) + f(x)), \tag{1}$$

where $k(x)$ and $f(x)$ are given functions.

3.5.1. Lagrange's method: variation of constant. Let us first return to a non-homogeneous equation of the form

$$y' = ky + f(x) \tag{2}$$

($k = \text{const}$, and $f(x)$ is a given function). We know that the general solution of the corresponding homogeneous differential equation $y' = ky$ ($k = \text{const}$) can be written as $y = Ae^{kx}$, where A is an *arbitrary constant.* Let us try to find solutions of the differential equation (2) in the form

$$y = A(x)e^{kx}, \tag{3}$$

where $A = A(x)$ is now a *function* of an independent variable x, which we would like to find.

Let us differentiate the function $y = y(x)$ given by (3):

$$y' = (A(x)e^{kx})' = A'(x) \cdot e^{kx} + A(x) \cdot ke^{kx} = kA(x)e^{kx} + A'(x)e^{kx},$$

[1] Latin *quadratura* means 'division of land into squares', giving a squared shape to some figure in order to then *calculate the area of this figure.* For the Greeks, this meant *drawing a square equal in area to the given figure*; recall the famous "quadrature of circle" problem! Later on, any problem reduced to finding areas or integrals became to be called this way.

and substitute y' and y into equation (2):

$$kA(x)\mathrm{e}^{kx} + A'(x)\mathrm{e}^{kx} = k \cdot A\mathrm{e}^{kx} + f(x).$$

Since the first terms on both sides of the resulting equation (with the unknown function $A(x)$) coincide, this equation, and along with it the differential equation (2), are satisfied (for values of x in the interval I on which equation (2) is considered) *if and only if*

$$A'(x)\mathrm{e}^{kx} = f(x) \quad \Leftrightarrow \quad A'(x) = f(x)\mathrm{e}^{-kx}. \tag{4}$$

Thus, the non-homogeneous equation (2) has been reduced to quadratures: if we find at least one antiderivative $F(x)$ of the function $\varphi(x) = f(x)\mathrm{e}^{-kx}$ (again *on the considered interval*), then the general solution of the differential equation (4) can be written as $A(x) = F(x) + C$ (C being an arbitrary constant), and the general solution of equation (2), accordingly, can be written in the form (3):

$$y = A(x)\mathrm{e}^{kx} = (F(x) + C)\mathrm{e}^{kx}.$$

Since in the above reasoning we considered a *function* $A(x)$ instead of an arbitrary constant A in the formula for solutions of the homogeneous equation ($y' = ky \ \Leftrightarrow \ y = A\mathrm{e}^{kx}$), i.e., we "let the arbitrary constant vary", the above method of finding solutions of the non-homogeneous equation is called the *variation of constant* (or *variation of parameter*) *method*. This method in the general case was developed in 1775 by Lagrange, although in particular situations it was used as early as 1740 by Euler and Daniel Bernoulli. Let us consider some concrete examples.

Example 1. In § 3.4 we found solutions to the differential equation of the form (2) in the case where $f(x) = b = \text{const}$ by using two methods: change of variables and d'Alembert's method, i.e., by finding a particular solution. Now let us apply to the differential equation

$$y' = ky + b \quad (k \neq 0) \tag{5}$$

the "Lagrange method". As we have proved above, the general solution of the differential equation (5) can be written as $y(x) = A(x)\mathrm{e}^{kx}$, where $A(x)$ is the general solution of the equation of the form (4) with the function $f(x)$ replaced by a constant b:

$$A'(x) = b\mathrm{e}^{-kx} \quad \Leftrightarrow \quad A(x) = -\frac{b}{k}\mathrm{e}^{-kx} + C.$$

Hence, we may write

$$y(x) = A(x)\mathrm{e}^{kx} = \Big(-\frac{b}{k}\mathrm{e}^{-kx} + C\Big)\mathrm{e}^{kx} = -\frac{b}{k} + C\mathrm{e}^{kx},$$

and of course we obtain the same answer as above. □

Example 2. Let us apply the same method to the differential equation

$$y' = ky + b\mathrm{e}^{\lambda x}. \tag{6}$$

Consider the differential equation for the function $A(x)$:

$$A'(x) = b\mathrm{e}^{\lambda x} \cdot \mathrm{e}^{-kx} = b\mathrm{e}^{(\lambda-k)x} \quad \Leftrightarrow \quad A(x) = \frac{b}{\lambda - k}\mathrm{e}^{(\lambda-k)x} + C.$$

Hence, the general form of solutions of the differential equation (6) is

$$y(x) = A(x)\mathrm{e}^{kx} = \Big(\frac{b}{\lambda-k}\mathrm{e}^{(\lambda-k)x} + C\Big)\mathrm{e}^{kx} = \frac{b}{\lambda-k}\mathrm{e}^{\lambda x} + C\mathrm{e}^{kx}.$$

However, if $\lambda = k$, the above reasoning "fails": the equation for $A(x)$ *cannot* have solutions of this form. What can be done? — Analyse this case on your own. □

The *Lagrange method* is useful as a *general method* for solving differential equations of the form (2), but if a particular solution of the equation is "easy to guess", then *d'Alembert's method* is faster.

Now let us consider some applied examples from elementary mechanics leading to equations of the form $y' = f(x)$ or to non-homogeneous linear equations of the form $y' = ky + f(x)$. These will be *non-conservative* dynamical systems in the sense used in Sec. 1.2.4: they will involve the *viscous friction force*.

3.5.2. Example: free motion with friction. In the case of free motion under viscous friction, Newton's equation has the form

$$mx'' = -\lambda v, \quad \lambda > 0 \tag{7}$$

(see Sec. 3.1.3). It reduces to a "split" system of first-order differential equations:

$$\begin{cases} x' = v, \\ v' = -kv, \quad k = \dfrac{\lambda}{m} > 0. \end{cases} \tag{8}$$

The second of them is a homogeneous linear equation, and its solutions are decreasing (in magnitude) exponential functions: $v(t) = A\mathrm{e}^{-kt}$. Hence, the first equation in (8) is solvable in quadratures:

$$x' = v(t) = A\mathrm{e}^{-kt} \quad \Leftrightarrow \quad x(t) = -\frac{A}{k}\mathrm{e}^{-kt} + C = A_1\mathrm{e}^{-kt} + C$$

(A is an arbitrary constant, so the fraction $-\dfrac{A}{k}$ is also an arbitrary constant [i.e., it may take any values], and we have replaced this fraction by a new *arbitrary constant* A_1).

The peculiarity of the obtained solution is that the frictional force leads to "deceleration", so every solution (different from a constant; note that any constant $x(t) \equiv C$ satisfies the original equation (7)) *asymptotically* (and *exponentially*) tends to a constant; the point particle *does not stop* but "infinitely long tends to stop", and its velocity tends exponentially to zero:

$$\lim_{t\to+\infty} x(t) = \lim_{t\to+\infty} (A_1 e^{-kt} + C) = C,$$

and

$$x(t) - C = A_1 e^{-kt}; \qquad v(t) = x'(t) = -kA_1 e^{-kt}.$$

To find out what "limit position" the point particle tends to, we can determine the constants A_1 and C from the initial conditions $(x(0); x'(0)) = (x_0; v_0)$: since $x(0) = A_1 + C$, $x'(t) = -kA_1 e^{-kt}$ and $x'(0) = -kA_1$, the initial conditions give a system of equations

$$\begin{cases} A_1 + C = x_0, \\ -kA_1 = v_0 \end{cases} \Leftrightarrow \quad \begin{cases} A_1 = -\frac{v_0}{k}, \\ C = x_0 - A_1 = x_0 + \frac{v_0}{k} \end{cases}$$

for finding the constants. Hence, the law of motion under given initial conditions $(x_0; v_0)$ can be written as

$$x(t) = -\frac{v_0}{k} e^{-kt} + \left(x_0 + \frac{v_0}{k}\right) = x_0 + \frac{v_0}{k}(1 - e^{-kt}).$$

Thus, if a point particle of mass m is subject to a viscous friction force $F_{\text{fr}} = -\lambda v$, then due to its *initial momentum*, this point covers a *"stopping distance"* of $\Delta x = \frac{v_0}{k} = \frac{mv_0}{\lambda}$ *in infinite time* (Fig. 46).

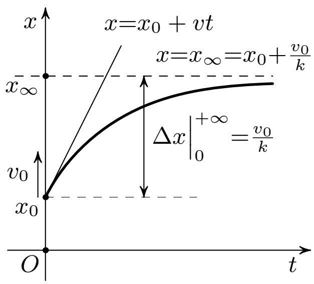


Fig. 46.

The same equation (7) can be solved in *another way*, which at first may seem too "artificial". However, both the next example and the considerations in the next chapter show that this way is a *method* rather than a mere "trick". In the case at hand, we rewrite the differential equation (7) using repeated differentiation and introduce an auxiliary variable z (we could do without it, but this way it is easier):

$$x'' = -kx' \quad \Leftrightarrow \quad x'' + kx' = 0 \quad \Leftrightarrow \quad (x' + kx)' = 0$$
$$\Leftrightarrow \quad \text{for } z = x' + kx \text{ we have } z' = 0 \quad \Leftrightarrow \quad z = C \text{ (const)}.$$

Thus, for the function $x=x(t)$ we obtain a simplest non-homogeneous linear equation like those considered in § 3.4:

$$x' = -kx + C.$$

Using the results of § 3.4 (see equations (11) and (12)), we, of course, obtain the same formula as above (though with a different notation for the constants) for the law of free motion with viscous friction.

3.5.3. Example: free fall with friction. Let us now find out what happens in *free fall* if the viscous friction force also acts. Recall that a mere free fall is *uniformly accelerated*, with velocity varying linearly and "height" varying quadratically:

$$z'' = -g \quad \Leftrightarrow \quad v = z' = -gt + v_0 \quad \text{and} \quad z = -\frac{1}{2}gt^2 + v_0 t + z_0$$

(see Ch. I, Sec. 1.2.2). Newton's differential equation in the presence of friction is written as follows:

$$mz'' = -mg - \lambda z' \quad \Leftrightarrow \quad z'' = -g - kz' \quad \left(k = \frac{\lambda}{m}\right). \tag{9}$$

This equation models a parachute jump.[1]

We will solve this equation in *three* ways. The *first way* is the "splitting method", where we first find the dependence $v=v(t)$, and then *integrate* the equation $z'=v(t)$ (i.e., find all its solutions).[2] In this case, we obtain a standard non-homogeneous linear equation for the velocity $v=z'$:

$$v' = -g - kv \quad \Leftrightarrow \quad v(t) = -\frac{g}{k} + A\mathrm{e}^{-kt}.$$

Then, the coordinate z is found by simple quadrature:

$$z' = v(t) = -\frac{g}{k} + A\mathrm{e}^{-kt}$$
$$\Leftrightarrow \quad z(t) = -\frac{g}{k}t - \frac{A}{k}\mathrm{e}^{-kt} + C = \left(-\frac{g}{k}t + C\right) + A_1\mathrm{e}^{-kt}.$$

The brackets enclose the "linear component" of the function $z=z(t)$ corresponding to the *oblique asymptote* $z_A = -\frac{g}{k}t + C$ of the graph of $z=z(t)$ in the Otz plane (Fig. 47): the difference $z(t) - z_A(t) = A_1\mathrm{e}^{-kt}$ tends to zero as $t \to +\infty$.

Thus, the presence of viscous friction causes the motion to become asymptotically "more and more" uniform, while the velocity tends to a constant $v_\infty = -\frac{g}{k}$.

[1] French *parachute*, from *parer*, 'to ward off', and *chute*, 'fall'.

[2] This terminology, *integrating* in the sense of finding solutions of a differential equation, is still used today. As we have already noted, before Lagrange the *solutions* of differential equations themselves were also called *integrals*.

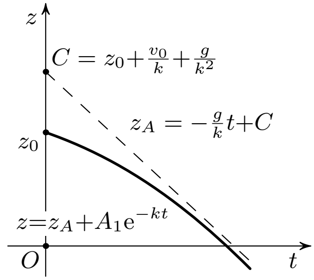


Fig. 47.

Initial conditions $(z(0); z'(0)) = (z_0; v_0)$ allow us to determine the constants A_1 and C: $z(0) = A_1 + C$; $z'(t) = -\frac{g}{k} - kA_1 \mathrm{e}^{-kt}$, whence it follows that $z'(0) = -\frac{g}{k} - kA_1$. We obtain a system of equations

$$\begin{cases} A_1 + C = z_0, \\ -\frac{g}{k} - kA_1 = v_0 \end{cases} \Leftrightarrow \quad \begin{cases} A_1 = -\frac{v_0}{k} - \frac{g}{k^2}, \\ C = z_0 - A_1 = z_0 + \frac{v_0}{k} + \frac{g}{k^2}. \end{cases}$$

Now, it is not difficult to write the law of motion in free fall; do it on your own.

The *second way* to find all solutions of the differential equation (9) is "reduction to repeated differentiation" (as in the preceding example). Let us transform equation (9):

$$(z' + kz)' = -g \quad \Leftrightarrow \quad z' + kz = -gt + C \quad \Leftrightarrow \quad z' = -kz - gt + C. \qquad (10)$$

We have obtained a non-homogeneous linear differential equation. We can apply the variation of constant method to it, but it is easier to find a particular solution $z = z_1(t)$ and use d'Alembert's theorem: $z = z_1(t) + A\mathrm{e}^{-kt}$. To "compensate for the inhomogeneity", which is the linear function $-gt + C$ entering the right-hand side of equation (10), we will look for a particular solution in the form of a linear function $z_1(t) = at + b$. To determine the coefficients a and b, we substitute this function into equation (10):

$$(at + b)' = -k(at + b) - gt + C \quad \Leftrightarrow \quad a = -kat - kb - gt + C$$
$$\Leftrightarrow \quad (ka + g)t + (a + kb + C) \equiv 0.$$

Since a linear function $\varphi(t) = \alpha t + \beta$ is identically zero if and only if both coefficients α and β are zero, from the last formula we obtain a system of equations to find a and b:

$$\begin{cases} ka + g = 0, \\ a + kb + C = 0 \end{cases} \Leftrightarrow \quad \begin{cases} a = -\frac{g}{k}, \\ b = \frac{1}{k}(C - a) = \frac{C}{k} + \frac{g}{k^2}. \end{cases}$$

Thus, the general solution of the equation $z' = -kz - gt + C$ is of the form

$$z(t) = -\frac{g}{k}t + \frac{C}{k} + \frac{g}{k^2} + A\mathrm{e}^{-kt} = \left(-\frac{g}{k}t + C_1\right) + A\mathrm{e}^{-kt},$$

which is, of course, equivalent to that obtained above.

The idea behind the *third way* is to apply *d'Alembert's method* directly to the *second-order* equation. Let us write the differential equation (9) in the form $z'' + kz' = -g$. If the right-hand side of this equation were zero, we would get a *second-order homogeneous linear differential equation* $z'' + kz' = 0$. We know how to solve it: this is the "free motion with friction" equation, and its general solution is written as $z = A\mathrm{e}^{-kt} + C$.

Next, the *linearity* of this equation means not just the fact that the unknown function $z = z(t)$ and its derivatives z' and z'' enter the expression $z'' + kz'$ *linearly*, i.e., with degree one (by the way, the function z itself does not enter the equation at all in this case). The most important property is, as physicists say, the *"superposition principle": if functions $z(t)$ and $z_1(t)$ are solutions of a* linear differential equation, *then any of their linear combinations $Z(t) = \alpha z(t) + \beta z_1(t)$ is also a solution of this* (*homogeneous*; the statement is not true for non-homogeneous equations) *linear equation.* In particular, along with $z(t)$ and $z_1(t)$, solutions are also $\alpha z(t)$ and $z(t) \pm z_1(t)$.

This property also holds for the *general second-order linear differential equation* (with constant coefficients), which can be written in other "main variables" $(x; y)$ as

$$y'' + py' + qy = 0. \tag{11}$$

This can be checked by direct substitution, using *linearity of differentiation*: for twice differentiable (on some interval I) functions $y(x)$ and $y_1(x)$, for any $\alpha, \beta \in \mathbb{R}$ we have

$$(\alpha y(x) + \beta y_1(x))' = \alpha y'(x) + \beta y_1'(x),$$
$$(\alpha y(x) + \beta y_1(x))'' = \alpha y''(x) + \beta y_1''(x).$$

Of course, the linearity-of-differentiation property is equivalent to the well-known *differentiation rules*:

$$(\alpha y(x))' = \alpha y'(x), \qquad (y(x) \pm y_1(x))' = y'(x) \pm y_1'(x).$$

Proposition 1 (d'Alembert's second theorem)**.** *The general solution $y(x)$ of a non-homogeneous differential equation*

$$y'' + py' + qy = f(x) \tag{12}$$

can be represented as a sum of any particular *solution of this equation and the general solution of the corresponding homogeneous equation* (11).

Proof. The proof of this statement is essentially the same as the proof of d'Alembert's "first" theorem (Theorem 1 in Sec. 3.4.3). Namely, if functions $y(x)$ and $y_1(x)$ both satisfy the differential equation (12), i.e.,

$$y''(x) + py'(x) + qy(x) = f(x),$$
$$y_1''(x) + py_1'(x) + qy_1(x) = f(x),$$

then after subtracting the second equation from the first, we obtain that the inhomogeneity $f(x)$ "disappears" and the difference $y(x) - y_1(x)$ of the solutions satisfies the homogeneous equation (11):

$$(y - y_1)'' + p(y - y_1)' + q(y - y_1) = 0.$$

Hence, if $z = z(x)$ is the "general solution" of the homogeneous equation (11), then $y - y_1 = z$, and $y = y_1 + z$. □

Thus, to find all solutions of a non-homogeneous linear differential equation of the form (12), we should, *first*, know the general form of solutions of the corresponding homogeneous equation (11), and *second*, find *at least one* (particular) solution of the non-homogeneous equation (12).

In our case, for the equation $z'' + kz' = -g$, the *general solution* of the corresponding homogeneous equation $z'' + kz' = 0$ is known from previous considerations: $z_0(t) = A\mathrm{e}^{-kt} + C$, where A and C are arbitrary constants. It remains to find at least one particular solution of the non-homogeneous equation, i.e., to choose a function z_1 so that to fulfil the equation $z'' + kz' = -g$. Of course, a constant cannot satisfy this equation: the left-hand side would be zero, while $-g \neq 0$. Let us try to find a solution in the form of a *linear function* $z_1(t) = at + b$. Then $z_1' = a$, $z_1'' \equiv 0$, so by substituting such a function into the differential equation we obtain

$$0 + k \cdot a = -g \quad \Leftrightarrow \quad a = -\frac{g}{k};$$

and the value of b can be any. Thus, we can take $z_1(t) = -\frac{g}{k}t$ as a particular solution, and it follows from the above proposition that the general solution of the non-homogeneous differential equation (9) can be written as

$$z(t) = z_1(t) + z_0(t) = -\frac{g}{k}t + A\mathrm{e}^{-kt} + C.$$

Of course, we have got the same answer as above.

3.5.4. Homogeneous linear equations with variable coefficients: $\boldsymbol{y' = k(x)y}$. We have considered, both in the general case and in a number of examples, a modification of the homogeneous first-order linear differential equation $y' = ky$ in the case where "inhomogeneity", i.e., a given function $f(x)$, is added to the right-hand side. Of interest is yet another generalisation of the equation $y' = ky$, where the coefficient k is not a constant but some function $k = k(x)$. The resulting differential equations are called *homogeneous linear equations with variable coefficients*:

$$y' = k(x)y. \tag{13}$$

Such equations have to be studied in evolutionary models when, for instance, the rate of change in the size of some population is time-dependent (e.g., the growth rate is subject to seasonal or other changes).

Note that solutions of the differential equation (13) satisfy the *superposition principle: if functions $y(x)$ and $y_1(x)$ are solutions of the differential equation* (13), *then any of their linear combinations $Y(x) = \alpha y(x) + \beta y_1(x)$ is also a solution of this equation.* In particular, along with $y(x)$ and $y_1(x)$, both $\alpha y(x)$ and $y(x) \pm y_1(x)$ are also solutions. This can be proved, as above, by direct substitution, using *linearity of differentiation.*

Hence, if we have found *at least one* nonzero (i.e., different from the identical zero) solution $y_1(x)$ of the differential equation (13), then from it we obtain *infinitely many* solutions of this equation of the form $y(x) = Ay_1(x)$ (A being an arbitrary multiplicative constant).

Now let us try to actually find at least one solution to the differential equation $y' = k(x)y$. In the case of $k(x) = k = \mathrm{const}$ discussed in § 3.4, the fundamental solution was $y = \mathrm{e}^{kx}$: after differentiating such a function, y turns into ky, where the number k appears *as the derivative of the exponent $z = kx$.* This suggests the idea to look for a solution among functions of the form $y_1(x) = \mathrm{e}^{z(x)}$ (along with it, all functions $y(x) = Ay_1(x)$ will be solutions, that is why we have not added a multiplicative constant to the "form" of $y_1(x)$).

Substituting the function $y_1(x) = \mathrm{e}^{z(x)}$ into the differential equation (13), we obtain

$$y_1' = z'(x)\mathrm{e}^{z(x)} = k(x)y_1 = k(x)\mathrm{e}^{z(x)} \quad \Leftrightarrow \quad z'(x) = k(x). \tag{14}$$

Thus, finding solutions to a differential equation in the above form leads to quadratures. If, for a function $k(x)$, equation (14), $z'(x) = k(x)$, has a solution $z = K(x)$ on some interval I, then, first, the function $y_1(x) = \mathrm{e}^{K(x)}$ is a solution of the differential equation (13), and second, according to the preliminary remark, *any* function of the form $y(x) = Ay_1(x) = A\mathrm{e}^{K(x)}$ is also a solution of (13).

The question arises: does not the differential equation (13) have any other solutions? In other words, we need to answer the *question of the uniqueness of solutions* of such an equation.

Theorem 1 (on existence and uniqueness of solutions of the differential equation $y' = k(x)y$)**.** *Assume that a function $k(x)$ is defined on some* (*finite or infinite*) *interval I and has an antiderivative on it, i.e., the differential equation* (14), *$z'(x) = k(x)$, has some solution $z = K(x)$ on this interval.*

Under this assumption, *the function* $y(x)$ *satisfies the differential equation* (13), $y' = k(x)y$, if and only if *it can be represented as*

$$y(x) = Ae^{K(x)}, \tag{15}$$

where the multiplicative constant A *is uniquely determined by the initial condition* $y(x_0) = y_0$ *(or* $y(0) = y_0$*).*

Proof. We have already checked that any function of the form (15) satisfies the differential equation (13). Let us prove the converse.

Assume that some function $y(x)$ is a solution of equation (13). Let us consider the function $h(x) = y(x)e^{-K(x)}$ and find its derivative (on the interval I). Taking into account that $y'(x) = k(x)y(x)$ and $K'(x) = k(x)$, we obtain

$$\begin{aligned} h'(x) = y'(x) \cdot e^{-K(x)} + y(x) \cdot e^{-K(x)} \cdot (-K'(x)) \\ = k(x)y(x)e^{-K(x)} - y(x)k(x)e^{-K(x)} \equiv 0 \\ \Rightarrow \quad h(x) = y(x)e^{-K(x)} \equiv \text{const} = A \quad \Rightarrow \quad y(x) = Ae^{K(x)}, \end{aligned}$$

as required.

Taking into account the initial condition $y(x_0) = y_0$, we obtain

$$y(x_0) = Ae^{K(x_0)} = y_0 \quad \Rightarrow \quad A = y_0 e^{-K(x_0)} \quad \Rightarrow \quad y(x) = y_0 e^{K(x) - K(x_0)}.$$

The theorem is completely proved. □

Example 3. To find all solutions of the differential equation $y' = xy$, we write and solve the equation for z:

$$z' = x \quad \Leftrightarrow \quad z(x) = \frac{1}{2}x^2 + C.$$

We are interested in *at least one* solution; therefore, we put $C = 0$. From the above theorem, we obtain

$$y' = xy \quad \Leftrightarrow \quad y(x) = Ae^{\frac{1}{2}x^2}$$

(the answer can be checked by differentiating). □

3.5.5. General linear equations with variable coefficients: $y' = k(x)y + f(x)$. Lastly, let us consider *non-homogeneous first-order linear differential equations with variable coefficients*, i.e., equations of the form

$$y' = k(x)y + f(x), \tag{16}$$

where $k(x)$ and $f(x)$ are given functions defined on some interval I.

Let us show that such equations can also be reduced to quadratures. First, *d'Alembert's theorem* remains valid for such equations.

Proposition 2 (d'Alembert's third theorem)**.** *The* general solution $y(x)$ *of the non-homogeneous differential equation* (16), $y' = k(x)y + f(x)$, *can be represented as a sum of any* particular *solution of this equation and the* general solution *of the corresponding homogeneous equation* (13), $y' = k(x)y$.

Proof. The proof of this statement essentially coincides with the proofs of d'Alembert's first and second theorems given above, so we omit it. □

Thus, as above, to find all solutions of a non-homogeneous linear differential equation of the form (16), we should, *first*, know the general form of solutions of the corresponding homogeneous equation (13), and *second*, find *at least one* (particular) solution of the non-homogeneous equation (16). The general form of solutions of the homogeneous equation (13) has been found above: if a function $K(x)$ is a solution of the differential equation $z' = k(x)$ (i.e., $\forall x \in I \ \exists K'(x) = k(x)$), then (on the given interval) we have

$$y' = k(x)y \quad \Leftrightarrow \quad y(x) = A\mathrm{e}^{K(x)} \quad (A \in \mathbb{R}).$$

As for the non-homogeneous linear equation (16), we apply to it the *variation of constant method*: we substitute into this equation the function $y = A(x)\mathrm{e}^{K(x)}$, where $K(x)$, as above, is the *antiderivative* of $k(x)$ and $A(x)$ is an unknown (sought for) function (the "variation" of the arbitrary constant A).

The derivative

$$y' = A'(x) \cdot \mathrm{e}^{K(x)} + A(x) \cdot k(x)\mathrm{e}^{K(x)} = k(x)A(x)\mathrm{e}^{K(x)} + A'(x)\mathrm{e}^{K(x)}$$

must be equal to $k(x)y + f(x)$; i.e.,

$$k(x)A(x)\mathrm{e}^{K(x)} + A'(x)\mathrm{e}^{K(x)} = k(x)A(x)\mathrm{e}^{K(x)} + f(x).$$

Since the first terms on both sides of the resulting equation (with the unknown function $A(x)$) coincide, this equation, and also the differential equation (16), are fulfilled (for values of x from the interval on which equation (16) is considered) *if and only if*

$$A'(x)\mathrm{e}^{K(x)} = f(x) \quad \Leftrightarrow \quad A'(x) = f(x)\mathrm{e}^{-K(x)}. \tag{17}$$

Thus, the non-homogeneous equation (16) is reduced to the quadrature (17): if at least one antiderivative $F(x)$ of the function $\varphi(x) = f(x)\mathrm{e}^{-K(x)}$ is found, then the general solution of the differential equation (17) is $A(x) = F(x) + C$ (C being an arbitrary constant); as a particular solution of equation (16) we may take $y_1 = F(x)\mathrm{e}^{K(x)}$ (check!); and finally, the general solution of equation (16) is of the form

$$y = F(x)\mathrm{e}^{K(x)} + A\mathrm{e}^{K(x)},$$

where A is an arbitrary constant.

Examples of application of the above methods are given in the exercises. At this point we will finish our introduction to *linear* first-order differential equations and proceed to the study of non-linear (autonomous and non-autonomous) first-order differential equations and related evolutionary models.

Exercises, Problems, and Tasks to Chapter III

Project task A (to § 3.1–3.2). Based *solely* on the differential equation

$$y' = ky(x), \tag{18}$$

it is required to prove that its solution for any nonzero initial condition $y(0) = A$ is a function that, for all rational values of x, coincides with Ac^x, where $c > 0$ is some number depending only on the coefficient k of equation (18). The proof can be conducted by following the plan below (or your own plan).

Assume that equation (18) has a solution *defined on the entire number axis* that satisfies a *non-zero* initial condition at zero: $y(0) = y_0 \neq 0$.

A-1. *Prove that an everywhere defined solution of equation* (18) *with a non-zero initial condition does not vanish at any point on the number axis.* In other words, if a function $f(x)$, $x \in \mathbb{R}$, satisfies the differential equation (18), i.e.,

$$\forall x \in \mathbb{R} \quad f'(x) = kf(x),$$

and if

$$f(0) = A_0 \neq 0,$$

then

$$\forall x \in \mathbb{R} \quad f(x) \neq 0.$$

A-2. *Prove that any solution $f(x)$ of equation* (18) *defined on the entire number axis that satisfies a non-zero initial condition $f(0) = A_0 \neq 0$ takes values of the same sign on the entire number axis.*

A-3. *Prove that any solution $f(x)$ of equation* (18) *defined on the entire number axis that satisfies a non-zero initial condition is either strictly increasing or strictly decreasing on the entire number axis.*

A-4 *Prove that* the general (an arbitrary) *solution of the differential equation* (18) *is proportional to any of its* particular *solutions.* In other words, if $f(x)$, $x \in \mathbb{R}$, is a solution of the differential equation (18) with a non-zero initial condition $f(0) = A_0 \neq 0$, then the function $y(x)$ is a solution of equation (18) *if and only if* it can be represented as

$$y(x) = Af(x)$$

for some constant A.

A-5. *Prove that if there exists a solution f of equation* (18) *defined on the entire number axis and satisfying a non-zero initial condition* $f(0)=A_0\neq 0$, *then* there exists a unique *solution satisfying the* unit *initial condition* $y(0)=1$.

Hereafter, we will call the solution of the differential equation (18) *satisfying the unit initial condition the* basis *solution and denote it by* $\boldsymbol{e}(x)$.

A-6. *Prove the* a priori[1] uniqueness theorem for solutions of the differential equation $y'=ky$: *any solution* $y=y(x)$, $x\in\mathbb{R}$, *of the differential equation* (18) *can be represented as*

$$y(x)=A\boldsymbol{e}(x),$$

where the multiplicative constant A is uniquely determined by the initial condition: $A=y(0)$.

A-7. *Prove the* main property of the basis solution of the differential equation $y'=ky$: *for any* $a,b\in\mathbb{R}$ *we have*

$$\mathrm{e}(a+b)=\mathrm{e}(a)+\mathrm{e}(b).$$

A-8. *Prove that*

$$\forall a\in\mathbb{R}\ \forall n\in\mathbb{N}\quad \mathrm{e}(na)=(\mathrm{e}(a))^n.$$

A-9. *Prove that the same equality holds valid for any* integer *multiples of* a:

$$\forall a\in\mathbb{R}\ \forall m\in\mathbb{Z}\quad \mathrm{e}(ma)=(\mathrm{e}(a))^m.$$

A-10. *Prove a similar property for fractional factors of the form* $\alpha=\frac{1}{n}$:

$$\forall c\in\mathbb{R}\ \forall n\in\mathbb{N}\ (n\geq 2)\quad \boldsymbol{e}\Big(\frac{1}{n}c\Big)=(\boldsymbol{e}(c))^{\frac{1}{n}}.$$

A-11. *Finally, prove a similar property for arbitrary fractional* (*i.e., rational*) *factors of the form* $\alpha=\frac{m}{n}$:

$$\forall a\in\mathbb{R}\ \forall m\in\mathbb{Z}\ \forall n\in\mathbb{N}\ (n\geq 2)\quad \boldsymbol{e}\Big(\frac{m}{n}a\Big)=(\boldsymbol{e}(a))^{\frac{m}{n}}.$$

A-12. *Prove the* theorem on values of a basis solution at rational points: *for a basis solution* $\boldsymbol{e}(x)$ *of the differential equation* $y'=ky$, *we have*

$$\forall x\in\mathbb{Q}\quad \boldsymbol{e}(x)=c^x,\quad \text{where}\quad c=\boldsymbol{e}(1).$$

[1]Latin *a priori*, 'from the earlier', in the sense of 'independently of any experience', 'before experience', as opposed to *a posteriori*, 'from the later', meaning 'from experience', 'based on experience'. In this case, we are talking about a property which is true *under the assumption of the existence* of a basis solution, although we have not proved it yet (we have only deduced it from the conjecture, accepted without proof, that some everywhere defined solution with a non-zero initial condition exists).

A-13. *Prove the original* (main) *statement*: a solution of equation (18) for any non-zero initial condition $y(0)=A$ is a function coinciding with Ac^x for all rational values of x, where $c=\boldsymbol{e}(1)$.

Problems and exercises

1. Solve the equations:

(1) $2^{2x}=5-x$; (3) $2^x+7^x=3^2$; (5) $3^x+4^x=5^x$;
(2) $5^x=27-x$; (4) $2^{2x}-3^x=1$; (6) $2^{1-x}=1+\lg x$.

Hint. Find one root by "guessing and trying", and then *prove that there are no other roots* using the monotonicity property (maybe, after some *transformation* of the equation).

2. Consider successive natural powers of 2:

$$2,\ 4,\ 8,\ 16,\ 32,\ 64,\ 128,\ 256,\ \ldots$$

(1) (*Easy problem.*) Prove that some (positive integer) power of 2 starts with the digit 7 (hereinafter we mean the *power written in decimal notation*).

(2) (*Not too difficult problem.*) Prove that some power of 2 starts with the digits 1999.

(3) (*Another not too difficult problem.*) Prove that some power of 2 starts with *any* given sequence of digits.

(4) (*Difficult problem.*) Find out with which digit, 7 or 8, the powers of 2 start *more often.*[1]

3. Find derivatives of the functions:

(1) e^{3-2x}; (2) e^{x^2}; (3) $e^{\cos x}$; (4) 2^{1+3x}; (5) $x3^x$; (6) $10^{\tan x}$;
(7) e^{e^x}; (8) $\sin e^x$; (9) x^{1-2x}; (10) $(x^2)^x$; (11) x^{x^2}; (12) $x^{\sin x}$;
(13) $\cosh x=\dfrac{e^x+e^{-x}}{2}$; (14) $\sinh x=\dfrac{e^x-e^{-x}}{2}$; (15) $\tanh x=\dfrac{\sinh x}{\cosh x}$; (16) $\coth x=\dfrac{\cosh x}{\sinh x}$.

Remark. The functions in the last four items are the so-called *hyperbolic functions,* respectively, the *hyperbolic cosine, sine, tangent, and cotangent.* Later (in § 5.6) we will explain the reason for such names.

4. Find derivatives of the functions:

(1) $\ln(1-2x)$; (2) $\ln(1-x^2)$; (3) $\log_2(1-3x)$; (4) $\ln\dfrac{1+x}{1-x}$;
(5) $\ln\dfrac{x+1}{x-1}$; (6) $\ln(\sqrt{x^2+1}+x)$; (7) $\ln(\sqrt{x^2+1}-x)$; (8) $\ln(x+\sqrt{x^2-1})$;
(9) $\ln(x-\sqrt{x^2-1})$; (10) $(\ln x)^2$; (11) $\frac{1}{\ln x}$; (12) $\ln(\ln x)$;

[1] In some natural sense that you give to the words *more often* on your own!

(13) $\ln(\sin x)$;
(14) $\ln(\cos x)$;
(15) $\ln(\tan x)$;
(16) $\cos(\ln x)$;
(17) $\log_x(1+x^2)$;
(18) $\ln\ln x$;
(19) $\ln\lg x$;
(20) $\lg\ln x$;
(21) $\ln(\ln(\ln x))$;
(22) $\ln 3^x$.

5. Write equations of the tangent and normal (i.e., perpendicular) to a given graph $y=f(x)$ at a given point x_0 (i.e., at the point on the graph with the given x-value):

(1) $y=\mathrm{e}^{-x}$, $x_0=-1$;
(2) $y=3^x$, $x_0=0$;
(3) $y=\dfrac{1}{\mathrm{e}^x+1}$, $x_0=0$;
(4) $y=\dfrac{x}{2^x}$, $x_0=0$;
(5) $y=\ln x$, $x_0=2$;
(6) $y=\dfrac{1}{\ln x}$, $x_0=\mathrm{e}$.

6. Draw the graph of the function (if possible, with complete analysis):

(1) $y=\mathrm{e}^{|x|}$;
(2) $y=\mathrm{e}^{-|x|}$;
(3) $y=\cosh x$;
(4) $y=\sinh x$;
(5) $y=\tanh x$;
(6) $y=\coth x$;
(7) $y=x+\mathrm{e}^x$;
(8) $y=x+\mathrm{e}^{-x}$;
(9) $y=\dfrac{1}{\mathrm{e}^x+1}$;
(10) $y=\dfrac{\mathrm{e}^x}{1+\mathrm{e}^x}$;
(11) $y=\dfrac{2^x-1}{3^x-1}$;
(12) $y=\dfrac{2^x+1}{3^x-1}$;
(13) $y=\dfrac{2^x-1}{3^x+1}$;
(14) $y=\dfrac{2^x+1}{3^x+1}$;
(15) $y=\dfrac{3^x-1}{2^x-1}$;
(16) $y=\dfrac{3^x+1}{2^x-1}$;
(17) $y=\dfrac{3^x-1}{2^x+1}$;
(18) $y=\dfrac{3^x+1}{2^x+1}$;
(19) $y=\mathrm{e}^{-x}\cos x$;
(20) $y=\mathrm{e}^{\cos x}$;
(21) $y=\mathrm{e}^{\sin x}$;
(22) $y=\mathrm{e}^{\tan x}$;
(23) $y=\ln x^2$;
(24) $y=(\ln x)^2$;
(25) $y=\ln\sin x$;
(26) $y=\ln|\sin x|$;
(27) $y=\ln(1+\cos x)$;
(28) $y=\ln\dfrac{x+1}{x-1}$;
(29) $y=\ln\dfrac{1-x}{1+x}$;
(30) $y=\ln(\sqrt{x^2+1}+x)$;
(31) $y=\ln(\sqrt{x^2+1}-x)$;
(32) $y=\ln(x+\sqrt{x^2-1})$;
(33) $y=\ln(x-\sqrt{x^2-1})$;
(34) $y=\ln 10^x$;
(35) $y=\lg 10^x$;
(36) $y=10^{\ln x}$.

7. Examine the function for monotonicity, find extremum points, make a "scheme of the graph" (signs, increasing, decreasing; do not find limits and asymptotes):

(1) $x\mathrm{e}^x$;
(2) $x^2\mathrm{e}^{-x}$;
(3) $x^2\mathrm{e}^{-x^2}$;
(4) $\dfrac{\mathrm{e}^x}{x}$;
(5) $x\ln x$;
(6) $x^2\ln x$;
(7) $x-\ln x$;
(8) $\dfrac{\ln x}{x}$;

(9) $\dfrac{x}{\ln x - 1}$;

(10) $\dfrac{x^2}{2^{-x}}$;

(11) $x^{\alpha}\mathrm{e}^{-x}$ $(\alpha > 0,\ x \geq 0)$;

(12) $x^2 - \ln x^2$.

8. Find the limit:

(1) $\lim\limits_{x\to 0} \dfrac{\mathrm{e}^{kx}-1}{x}$;

(2) $\lim\limits_{x\to 2} \dfrac{3^x-9}{x-2}$;

(3) $\lim\limits_{x\to 2} \dfrac{a^x-a^2}{x^2-3x+2}$;

(4) $\lim\limits_{x\to 0} \dfrac{a^x-b^x}{x}$;

(5) $\lim\limits_{x\to 0} \dfrac{a^x-1}{\sin bx}$;

(6) $\lim\limits_{x\to 0} \dfrac{\tan ax}{b^x-a^x}$;

(7) $\lim\limits_{x\to 0} \dfrac{\ln(1+ax)}{x}$;

(8) $\lim\limits_{x\to 0} \dfrac{\log_a(1+x)}{x}$;

(9) $\lim\limits_{x\to 0} \cot x \cdot \ln(1+x)$;

(10) $\lim\limits_{x\to 0} (1+\sin x)^{\cot x}$;

(11) $\lim\limits_{x\to +\infty} x\mathrm{e}^{-x}$;

(12) $\lim\limits_{x\to +\infty} x^2\mathrm{e}^{-x}$;

(13) $\lim\limits_{x\to +\infty} \dfrac{\ln x}{x}$;

(14) $\lim\limits_{x\to +\infty} \dfrac{(\ln x)^2}{x}$.

9. Find the limit of the sequence x_n (as $n\to\infty$) if it exists:

(1) $x_n = \left(1+\frac{\alpha}{n^2}\right)^n$;

(2) $x_n = \left(1+\frac{x}{n}+\frac{\alpha}{n^2}\right)^n$;

(3) $x_n = \left(\cos\frac{x}{n}\right)^n$;

(4) $x_n = \left(\cos\frac{x}{n}+\lambda\sin\frac{x}{n}\right)^n$.

10. Write linear approximation formulae in the neighbourhood of the point $x_0 = 0$ for the functions:

(1) e^x;

(2) 2^x;

(3) a^x $(a>0)$;

(4) $\ln(1+x)$;

(5) $\lg(1+x)$;

(6) $\log_a(1+x)$ $(a>0,\ a\neq 1)$.

Hint. See § 1.1 and exercises to Ch. I.

11. Using derivatives, prove the following inequalities (for any values of $x>0$):

(1) $\mathrm{e}^x > 1+x$;

(2) $\mathrm{e}^x > 1+x+\frac{x^2}{2}$;

(3) $\mathrm{e}^x > 1+x+\frac{x^2}{2!}+\frac{x^3}{3!}$;

(4) $\forall n\in\mathbb{N}\cup\{0\}\quad \mathrm{e}^x > 1+x+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}$.

12. Using derivatives, prove the following inequalities (for any values of $x>0$):

(1) $\mathrm{e}^x < 1+x\mathrm{e}^x$;

(2) $\mathrm{e}^x < 1+x+\dfrac{x^2\mathrm{e}^x}{2}$;

(3) $\mathrm{e}^x < 1+x+\dfrac{x^2}{2!}+\dfrac{x^3\mathrm{e}^x}{3!}$;

(4) $\forall n\in\mathbb{N}\cup\{0\}\quad \mathrm{e}^x < 1+x+\dfrac{x^2}{2!}+\ldots+\dfrac{x^n\mathrm{e}^x}{n!}$.

13. Using derivatives, prove the following inequalities (for any values of $x>0$):

(1) $x-\frac{x^2}{2}<\ln(1+x)<x$; (2) $1+2\ln x\le x^2$.

14. Assuming *Bernoulli's formula*

$$\mathrm{e}=1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\ldots+\frac{1}{n!}+\ldots$$

is known, prove the inequalities:

(1) $\mathrm{e}>1$; (2) $\mathrm{e}>2$; (3) (∗) $\mathrm{e}<4$; (4) (∗∗) $\mathrm{e}<3$.

15. Prove that the function $\varphi(x)=x^{1-\alpha}\cdot\ln x$ for any $\alpha<1$ is unbounded from above on the interval $(1,+\infty)$ (i.e., $\forall M\in\mathbb{R}\ \exists x>1\ \varphi(x)>M$).

16. Prove that for any $\alpha>0\ \exists M\in\mathbb{R}\ \forall x>1\ \frac{\ln x}{x^\alpha}\le\frac{M}{x^{\alpha/2}}$. Deduce the second comparison theorem from this.

17. Prove that for any $a>0$ we have $\lim\limits_{x\to+\infty}\frac{a^x}{x^x}=0$.

18. Draw graphs of the functions, examining them for monotonicity (if possible, for convexity/concavity) and analysing the behaviour of the functions at the boundaries of their domains (including the behaviour at infinity):

(1) $y=x\mathrm{e}^x$;
(2) $y=x\mathrm{e}^{-x}$;
(3) $y=x\mathrm{e}^{-x^2}$;
(4) $y=x^2\mathrm{e}^{-x}$;
(5) $y=x^2\mathrm{e}^{-x^2}$;
(6) $y=\frac{\mathrm{e}^x}{x}$;
(7) $y=\frac{x^2}{2^{-x}}$;
(8) $y=x^\alpha\mathrm{e}^{-x}$ $(\alpha>0,\ x\ge 0)$;
(9) $y=\ln(2^x+1)$;
(10) $y=x\ln x$;
(11) $y=x^2\ln x$;
(12) $y=x\ln^2 x$;
(13) $y=x^{x^2}$;
(14) $y=(x^x)^2$;
(15) $y=x-\ln x$;
(16) $y=\frac{\ln x}{x}$;
(17) $y=\frac{\ln^2 x}{x}$;
(18) $y=\frac{x}{\ln x-1}$;
(19) $y=\frac{\ln x^2}{1+\ln^2 x}$;
(20) $y=3\log_2^2 x-\log_2^3 x$;
(21) $y=x^2-2\ln x$;
(22) $y=x^2-\ln x^2$.

19. Find the number of roots of the equation depending on the value of the parameter $a\in\mathbb{R}$:

(1) $a^x=x$ $(a>0)$;
(2) $\ln x=ax$;
(3) $x^2\cdot\mathrm{e}^{2-|x|}=4a$;
(4) $3x\lg x=1+a\lg x$;
(5) (A.R. Zilberman's problem) $a^x=\log_a x$.

Hint to problem (5): First consider the case $a=\frac{1}{16}$; cannot you "guess" the roots?!

20. Write the *Taylor/Maclaurin series*, i.e., the *formal* power series

$$f(x) \mapsto f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots$$

for the following functions:

(1) $f(x) = e^{x^2}$;
(2) $f(x) = e^{-x}$;
(3) $f(x) = \dfrac{1}{1+x}$;
(4) $f(x) = \ln(1+x)$;
(5) $f(x) = \dfrac{1}{1+x^2}$;
(6) $f(x) = \arctan x$;
(7) $f(x) = \sin x$;
(8) $f(x) = \cos x$;
(9) $f(x) = \sinh x$;
(10) $f(x) = \cosh x$.

Hint. Use power series for the exponential function and for the sum of infinite geometric progression.

21. Prove that if for some series $f_0 + f_1 + \ldots + f_n + \ldots$ the series of absolute values $|f_0| + |f_1| + \ldots + |f_n| + \ldots$ converges, then the original series also converges (in this case the series $f_0 + f_1 + \ldots + f_n + \ldots$ is said to be *absolutely convergent*).

22. Using *majorisation of a series by an infinite geometric progression*, prove convergence of the following power series when $|x| < 1$, and specify the functions corresponding to these series:

(1) $1 - x + x^2 - x^3 + \ldots + (-1)^n x^n + \ldots$;
(2) $1 + 2x + 3x^2 + 4x^3 + \ldots + (n+1)x^n + \ldots$;
(3) $1 + x + 2x^2 + 3x^3 + \ldots + nx^n + \ldots$;
(4) $1 + x + \dfrac{x^2}{2} + \dfrac{x^3}{3} + \ldots + \dfrac{x^n}{n} + \ldots$;
(5) $1 - x + \dfrac{x^2}{2} - \dfrac{x^3}{3} + \ldots + (-1)^n \dfrac{x^n}{n} + \ldots$;
(6) $x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \ldots + (-1)^{n-1} \dfrac{x^n}{n} + \ldots$;
(7) $1 + x^2 + x^4 + x^6 + \ldots + x^{2n} + \ldots$;
(8) $1 - x^2 + x^4 - x^6 + \ldots + (-1)^n x^{2n} + \ldots$

Advice. To find the sum of an infinite series, try to differentiate it *formally* or to find a series for which the given series is a *formal derivative.*

23. Consider the function

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0; \\ 0, & x = 0. \end{cases}$$

(1) Prove that f is differentiable on the entire number axis and find its derivative $f'(x)$.

(2) Prove that f is *twice differentiable* on the entire number axis, i.e., has also the second derivative $f''(x)$.

(3) Prove that f has derivatives of all orders (on the entire number axis) and that for any $n \in \mathbb{N}$ the n-th derivative at $x = 0$ is zero: $f^{(n)}(0) = 0$.

Linear differential equations and linear models

24. Find the function $y = y(x)$ given that

(1) $y' = \frac{1}{2}y$, $y(0) = 4$;

(2) $y' = 4y$, $y(1) = \mathrm{e}$;

(3) $3y' = -2y$, $y(3) = \mathrm{e}$;

(4) $\dfrac{y'}{y} = \dfrac{2}{5}$, $y'(1) = \mathrm{e}$.

25. After 1 minute, a single bacterium in a flask produced 10 bacteria. How many bacteria will be in this flask after 10 minutes?

26. Twelve hours after the start of the experiment, the population had tripled in number. How many times will the number of bacteria increase after three days?

27. A substance has a half-life of 1 year. After what time will 1 g remain of 10 g of this substance?

28. In 5 days, 10% of a substance decays. What is its half-life?

29. The half-life of radium ^{226}Ra is 1620 years. If initially there was 200 g of radium, how much will remain after 300 years?

30. In 500 years, 10 g of radium decayed. How much was there initially?

31. Find how long it takes for 1, 10, 90, 99% of the initial amount of radium to decay?

32. The radium content of various rocks on Earth (the ratio of the number of radium atoms to the number of atoms in the rock) is on average close to 10^{-12}. What was the radium content 10 000 years ago; 10^6 years ago; 5×10^9 years ago (5×10^9 years is the age of the Earth)? Try to explain the answer.

33. Find the air pressure in the mine at depths of 1; 3; 10 km.

34. For some fissile substance, let the coefficients α and β be $2 \times 10^8\,\mathrm{s}^{-1}$ and 4×10^{10} cm/s, respectively.

(1) For what values of the radius r of a sphere of this substance will the number $N(t)$ of neutrons in the sphere increase?

(2) Find the smallest value of r at which a sphere of this radius will lose no more than half of its neutrons after one second, and find the ratio $\frac{r}{r_0}$ of this radius to the critical radius.

(3) For a given fissile substance, let $\alpha = 2 \times 10^8\,\mathrm{s}^{-1}$; assume that the critical radius is $r_0 = 50$ cm. A sphere of this substance of radius $r = 40$ cm contains N_0 free neutrons at time $t = 0$; how many neutrons does it contain at $t = 1$?

35. Let a sphere of radius $r = \frac{r_0}{2}$ (r_0 being the critical radius) be made of a fissile substance with fission rate $\alpha = 2 \times 10^8\,\mathrm{s}^{-1}$ and escape rate $\beta = 4 \times 10^{10}$ cm/s, and assume that a neutron source is supplying it with $q = 100$ neutrons per second. Find the time dependence $N(t)$ of the number of neutrons in the sphere and the limiting number N_1 of neutrons if the initial number of neutrons is $N(0)=N_0$.

36. Suppose there is a sphere of fissile matter with parameters $\beta=2\alpha$ and a source supplying it with qt^2 neutrons during time interval $[0,t]$. Let $N(t)$ be the number of free neutrons in the sphere at time t. Write the differential equation for $N(t)$ and find all its solutions.

37. Find (guess) particular solutions to the given differential equations and write their general solutions:

(1) $y'=-y+17$;
(2) $y'=y+\mathrm{e}^{-x}$;
(3) $y'=y+\mathrm{e}^{x}$;
(4) $y'=ky+\mathrm{e}^{\alpha x}$;
(5) $y'=y+\cos x$;
(6) $y'=ky+\sin\omega x$;
(7) $y'=y+2x$;
(8) $y'=2y-x$;
(9) $y'=ky+ax+b$;
(10) $y'=y+x^2$;
(11) $y'=y+x^2+x$;
(12) $y'=y+\frac{1}{x}$, $x>0$;
(13) $y'=2y+\mathrm{e}^{x}-x$;
(14) $y'=2y+\mathrm{e}^{x}+\mathrm{e}^{-x}$;
(15) $y'=2y+\mathrm{e}^{x}+\cos x$;
(16) $y'=2y+\mathrm{e}^{2x}+\sin x$;
(17) $y'=2y+\cos x+\sin x$;
(18) $y'=2y+\cos x+\cos 2x$.

Remark. If $y_1(x)$ and $y_2(x)$ are particular solutions of the equations $y'=ky+f_1(x)$ and $y'=ky+f_2(x)$, then $y_1(x)+y_2(x)$ is a particular solution of $y'=ky+f_1(x)+f_2(x)$ (prove).

38. Find the function $f(x)$ given that

(1) $f'(x)-2f(x)=x$ and $f(0)=\frac{1}{2}$;
(2) $f'(x)+f(x)=\mathrm{e}^{2x}$ and $f(0)=-1$;
(3) $f'(x)-\frac{1}{2}f(x)=\mathrm{e}^{x/2}$ and $f(0)=2$;
(4) $f'(x)+f(x)=\cos(\pi\mathrm{e}^{x})$ and $f(0)=\mathrm{e}$;
(5) $f'(x)+3f(x)=x\mathrm{e}^{x}$ and $f(0)=\frac{1}{\mathrm{e}}$.

39. As we know from physics, if an electric circuit with resistance R and inductance (self-induction coefficient) L and a voltage source that produces a *constant* EMF (electromotive force) $\mathcal{E}$ are connected in series in a loop, then the electric current $I=I(t)$ arising in this circuit satisfies the equation $LI'(t)+RI(t)=\mathcal{E}$. Find the time dependence of the current if there was no current at the initial time: $I(0)=0$.

40. Solve the same problem in the case where the circuit is connected with an EMF source varying according to the harmonic law: $\mathcal{E}=\mathcal{E}_0\sin\omega t$.

41. (*Radioactive decay chain.*) Let atoms of substance X turn into atoms of another radioactive substance Y in radioactive decay (the first substance is called the *mother*, the second is called the *daughter*), and atoms of Y turn into atoms of another substance Z, which itself is not radioactive.

(1) Write differential equations for the numbers $X(t)$ and $Y(t)$ of atoms of these substances if their decay probabilities are ω and ν.

(2) Find solutions $X(t)$ and $Y(t)$ of the resulting system of differential equations satisfying the initial conditions $X(0)=X_0$ and $Y(0)=0$. Examine them for monotonicity and extrema; plot the corresponding graphs (on the same plot).

(3) Under the same initial conditions, find out at what ratio between the decay probabilities (or between the half-lives T_X and T_Y) the number of atoms of substance Y will at some time be greater than the number of atoms of substance X.

(4) Show that in the case of a long-lived mother substance X and a short-lived daughter substance Y, after some sufficiently long time the "radioactive family" $X \to Y \to \dots$ passes to the so-called *stationary state* where

$$Y(t) \approx \frac{\omega}{\nu} X(t).$$

A real-world example of such a "radioactive family" is the chain $^{238}\mathrm{U} \to \dots \to {}^{226}\mathrm{Ra} \to \dots$ [uranium-238 decays into radium through a series of "intermediates"]. Note that the half-life of $^{238}\mathrm{U}$ is 4.5×10^9 years [and of radium $^{226}\mathrm{Ra}$, recall, 1620 years].

42. Through a hole in the bottom of a rectangular water tank with base area of $1\,\mathrm{m}^2$, water is poured out at a rate proportional to the height h of the water in the tank, and from above, water is poured into the tank at a rate of $10\,\mathrm{l/s}$. Compose and solve the differential equation for the function $h=h(t)$. Find out at what proportionality factor k for the pouring out rate will the water level in the tank rise? Can it grow to infinity?

43. Find the function $f(x)$ given that

(1) $f'(x)=-2x$, $f(1)=2$;

(2) $f'(x)=-x^2$, $f(0)=1$;

(3) $f'(x)=\mathrm{e}^{-2x}$, $f(0)=2$;

(4) $f'(x)=3^x$, $f(1)=3$;

(5) $f'(x)=-2\cos x$, $f(\pi)=17$;

(6) $f'(x)=\dfrac{1}{\sqrt{x}}$, $f(1)=2$ $(x>0)$;

(7) $f'(x)=-\dfrac{1}{x^2}$, $f(1)=2$ $(x>0)$;

(8) $f'(x)=-\dfrac{1}{x}$, $f(1)=1$ $(x>0)$;

(9) $f'(x)=3\sqrt{x}$, $f(1)=-2$ $(x>0)$;

(10) $f'(x)=\cos 2x$, $f(\pi)=-17$.

44. Prove that the differential equation $y' = \theta(x)$, where θ is the *Heaviside theta function*,[1] has no solutions defined on the entire number axis.

45. Find the general form of solutions of the differential equation

(1) $y' = x\cos x$; (3) $y' = x\mathrm{e}^{x}$; (5) $y' = x\mathrm{e}^{2x}$;
(2) $y' = x\sin x$; (4) $y' = x\mathrm{e}^{-x}$; (6) $y' = x\cos 2x$.

Hint for this and the next problem: try the "sequential search-and-try method", so to say. It is also useful to review the derivatives that were calculated previously.

46. Find the general form of solutions of the differential equation

(1) $y' = \cosh x$; (3) $y' = \ln x$; (5) $y' = \tan x$;
(2) $y' = \sinh x$; (4) $y' = \cot x$; (6) $y' = \sin^2 x$.

47. Find the general form of solutions of the *second-order* differential equation (it involves *two* arbitrary constants):

(1) $y'' = kx$; (3) $y'' = \cos x$; (5) $y'' = \sin 2x$;
(2) $y'' = \sin x$; (4) $y'' = 2 - 6x$; (6) $y'' = \sin^2 x$.

48. Find the general form of solutions of the *third-order* differential equation

(1) $y''' = 0$; (3) $y''' = 24x$; (5) $y''' = \sin x$;
(2) $y''' = 6$; (4) $y''' = \mathrm{e}^{-x}$; (6) $y''' = \sin 3x$.

49. Find the general form of solutions of the differential equation (y^{IV} denotes the *fourth* derivative of a function $y = y(x)$):

(1) $y''' = x^3$; (3) $y^{IV} = 0$; (5) $y^{IV} = 720x$.
(2) $y''' = y''$; (4) $y^{IV} = 24$. (6) $y^{IV} = y'''$.

50. Using the *variation of constant method*, find the general form of solutions of the differential equation

(1) $y' = -2y + \mathrm{e}^{-2x}$; (3) $y' = 2y - 2x\mathrm{e}^{2x}$; (5) $y' = -y - x\mathrm{e}^{x}$;
(2) $y' = -2y + \mathrm{e}^{-x}$; (4) $y' = 2y - x\mathrm{e}^{x}$; (6) $y' = -y + x\mathrm{e}^{-x}$.

51. Using the *change of variables* $y = A\mathrm{e}^{z(x)}$, find the general form of solutions of the differential equation

(1) $y' = 4xy$; (5) $y' = y\cos x$; (9) $y' = -\dfrac{y}{2\sqrt{2}}$ $(x>0)$;
(2) $y' = -2xy$; (6) $y' = y\cos 2x$;
(3) $y' = -3x^2 y$; (7) $y' = y\sin 2x$; (10) $y' = \dfrac{2y}{\sqrt{x}}$ $(x > 0)$.
(4) $y' = \mathrm{e}^{x}y$; (8) $y' = \dfrac{y}{2\sqrt{2}}$ $(x > 0)$;

[1] Recall that this function is defined by the compound formula: $\theta(x) = 0$ for $x < 0$ and $\theta(x) = 1$ for $x \geq 0$.

52. Using the same method, find the general form of solutions of the given differential equation and draw the family of its integral curves in the Oxy plane:

(1) $y' = \frac{y}{x}$; (3) $y' = \frac{y}{2x}$; (5) $y' = -\frac{2y}{x}$;

(2) $y' = \frac{2y}{x}$; (4) $y' = -\frac{y}{x}$; (6) $y' = -\frac{y}{2x}$.

53. Find the function $f(x)$ given that

(1) $f'(x) = 2xf(x)$, $f(0) = 1$;

(2) $f'(x) = 3x^2 f(x)$, $f(0) = 2$;

(3) $2\sqrt{x}f'(x) = f(x)$, $f(4) = 1$ $(x > 0)$;

(4) $x^2 f'(x) = -f(x)$, $\lim\limits_{x\to+\infty} f(x) = \mathrm{e}$ $(x > 0)$.

54. Using the *change of variables* $y = A\mathrm{e}^{z(x)}$ followed by the *variation of constant,* find the general form of solutions of the differential equation

(1) $y' = 2xy + b$; (4) $y' = -2xy + \mathrm{e}^{-x^2}$;

(2) $y' = 2xy + 2x\mathrm{e}^{x^2}$; (5) $y' = y\mathrm{e}^x + 2x\mathrm{e}^{\mathrm{e}^x}$;

(3) $y' = -2xy + 2x\mathrm{e}^{-x^2}$; (6) $y' = 2y\mathrm{e}^x + x\mathrm{e}^{\mathrm{e}^x}$.

Chapter IV

Variable Separable Models

§ 4.1. Analysis of the Evolutionary Equation $y' = g(y)$. Examples

In all the cases discussed in Chs. I–III, we succeeded in integrating *only because* the equations in question either were immediately reduced to quadratures (equations of the form $y' = f(x)$ from Ch. II) or were *linear* with respect to the unknown function $y = y(x)$ (as in Ch. III). This is very significant: it turns out that solutions of differential equations of the form $y' = F(x, y)$ that are *non-linear* with respect to y (i.e., the right-hand side of which is not representable as $k(x)y + f(x)$), *in the general case* cannot be written using elementary functions and quadratures (i.e., integrals). The simplest non-linear dependence is *quadratic*, and in 1841 the French mathematician Joseph Liouville rigorously proved the insolvability in elementary functions and quadratures of such a seemingly simple equation as $y' = y^2 - x$. However, also non-linear *first-order differential equations are solvable in elementary functions or in quadratures if their right-hand sides do not contain the independent variable x, i.e., if $F(x, y) = g(y)$* (and, of course, under certain conditions on the function $g(y)$). Recall (see § 3.1) that such equations are said to be *autonomous*; we will focus on them.

4.1.1. Directional fields of the equation $y' = g(y)$ and symmetric to them. As above, when analysing a differential equation of the form

$$y' = g(y) \tag{1}$$

(g is a given function, for the moment let it only be *continuous*), it will be helpful if we use the *geometric interpretation* of first-order equations by means of directional fields in the Oxy plane. Note that, just as for the equation $y' = ky$ in § 3.1, the directional field of equation (1) is *constant along the horizontal lines* $y = \text{const}$. Recall that we have already considered directional fields *constant along the vertical lines* $x = \text{const}$; such fields satisfy differential equations of the form

$$y' = f(x), \tag{2}$$

which reduce to quadratures. A comparison of these observations leads to the following idea: *by means of symmetry about the bisector of quadrants* I *and* III*, i.e., about the line* $y = x$*, the* directional field *of the differential equation* (1) *which is constant along the horizontals, together with the integral curve* $y = y(x)$ *of this field* (Fig. 48) *can be translated into a directional field* constant along the verticals. Thus, it would seem that the *differential equation* (1) *should be reduced to a differential equation of the form* (2), *which we already know how to solve.* Let us find out what we actually obtain after symmetry (Fig. 49).

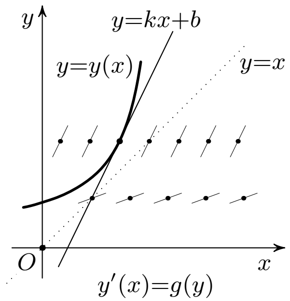


Fig. 48.

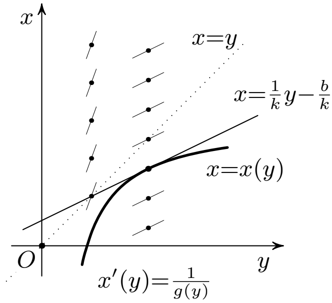


Fig. 49.

First, the Ox and Oy axes swap places, and the point $(x; y)$ maps to the point $(y; x)$.

Second, the line $y = kx + b$ that defines the direction at the point (x, y) and has slope $k = g(y)$, which we will assume to be non-zero, maps to its *symmetric line*; since the latter represents the graph of the inverse function, i.e., the linear function $x = \frac{1}{k}y - \frac{b}{k}$ with slope $k_{\text{sym}} = \frac{1}{k}$, the symmetric directional field in the Oyx plane is given by $k_{\text{sym}} = F(x, y) = \dfrac{1}{g(y)} = f(y)$; i.e., we obtain a directional field constant along the verticals in the Oyx plane. Note that if the direction of the original field at the point $(x; y)$ is horizontal, i.e., $k = g(y) = 0$, then the symmetric field at the corresponding point is vertical. We will call such a point a *singular* point, conventionally assuming that $k_{\text{sym}}(y, x) = \frac{1}{0} = \infty$.

And lastly third, if the solution $y = y(x)$ of the differential equation (1) in the neighbourhood of a point x is *invertible*, then its graph maps exactly to the graph of the *inverse function* $x = x(y)$. Intuitively, it is clear that the tangency is preserved under symmetry, so the *inverse function* $x(y)$ *in the neighbourhood of the point* y *should satisfy the* "symmetric" differential equation *of the form* (2): $x'_y = x'(y) = f(y) = \dfrac{1}{g(y)}$, *which is immediately* *integrable*:

$$x'(y) = \frac{1}{g(y)} \quad \Leftrightarrow \quad x = \int \frac{dy}{g(y)} = G(y) + C, \tag{3}$$

where the so-called *"indefinite integral"*[1] $\int \frac{dy}{g(y)}$ is merely a notation for an arbitrary (any) antiderivative $G(y)$ of $f(y) = \frac{1}{g(y)}$ in the neighbourhood of the point y (note that by assumption the function g is continuous and, since $g(y) \neq 0$, g is non-zero in the neighbourhood of y; therefore, the function $f = \frac{1}{g}$ is well defined, continuous, and hence *has an antiderivative* in this neighbourhood). It remains to express y through x from equation (3), and this is possible due to the assumption that the solution is invertible in the neighbourhood of x!

If we realise this *solution scheme*, we obtain, generally speaking, a dependence of the form $y = \varphi(x; C)$, where the value of the constant C is determined from the initial condition $y(x_0) = y_0$, i.e., from the equation $\varphi(x_0; C) = y_0$. It is often easier to take the initial condition into account directly during integration, by writing the dependence of x on y making use of another *Barrow's formula*:

$$x'(y) = \frac{1}{g(y)}, \quad y(x_0) = y_0 \quad \Leftrightarrow \quad x - x_0 = \int_{y_0}^{y} \frac{du}{g(u)}. \tag{4}$$

However, there is a gap in the above reasoning: it has not been rigorously proved that *tangency is preserved under symmetry* about the line $y = x$. Geometrically, this is clear, but the tangency and *tangent* to the graph were defined *analytically*, through the *derivative*, so finding solutions by the "scheme" (3) must also be justified analytically.

There is also one more obscurity: what about solutions $y = y(x)$ in the neighbourhood of points where the tangent is horizontal, i.e., $g(y) = 0$, and the function $y(x)$ itself can be both non-invertible and invertible (like the functions x^2 or x^3 in the neighbourhood of zero)?

Before addressing these questions, let us consider two already known examples.

4.1.2. Examples: linear equations $y' = ky\ [+ b]$. Let us apply the integration scheme (3) to first-order linear differential equations with constant coefficients, which have been studied in Ch. III above.

[1] We discuss the "indefinite integral" in more detail in the next section.

Example 1. The equation symmetric to $y' = ky$ $(k \neq 0)$ in the case of $k \neq 0$ is

$$x'(y) = \frac{1}{ky} \quad \Leftrightarrow \quad x = \int \frac{dy}{ky} = \frac{1}{k} \ln |y| + \widehat{C},$$

where $\widehat{C}$ is a double-valued constant in this case: $\widehat{C} = C_1, C_2$, depending on the sign of y. Hence, after some transformations and exponentiation, we express y through x:

$$\ln |y| = kx - k\widehat{C} \quad \Leftrightarrow \quad |y| = \mathrm{e}^{kx - k\widehat{C}} \quad \Leftrightarrow \quad y = \pm \mathrm{e}^{-k\widehat{C}} \mathrm{e}^{kx} = A\mathrm{e}^{kx},$$

where A is either a positive or negative constant. The special case $y = 0$ corresponds to a separate integral curve (stationary solution) $y(x) \equiv 0$; the latter corresponds to the vertical line $y = 0$ tangent to the symmetric vector field $k_{\mathrm{sym}} = \frac{1}{ky}$, which we will call a "singular solution" of the differential equation $x' = \frac{1}{ky}$ ("along it" we have as if "$x'_y = \infty$"). If we allow $A = 0$ in the above formula, we arrive at the already known family of solutions:

$$y' = ky \quad (x \in \mathbb{R}) \quad \Leftrightarrow \quad y = A\mathrm{e}^{kx}, \quad A \in \mathbb{R}. \qquad \square$$

Example 2. The non-homogeneous linear differential equation $y' = ky + b$ $(k \neq 0)$ can also be integrated similarly without much trouble (like finding a particular solution). The equation symmetric to it in the case $ky + b \neq 0$, i.e., $y \neq -\frac{b}{k}$, is

$$x'(y) = \frac{1}{ky+b} \quad \Leftrightarrow \quad x = \int \frac{dy}{ky+b} = \frac{1}{k} \ln |ky + b| + \widehat{C}$$

$$\Leftrightarrow \quad \ln |ky + b| = k(x - \widehat{C}) \quad \Leftrightarrow \quad |ky + b| = \mathrm{e}^{kx - k\widehat{C}}$$

$$\Leftrightarrow \quad y = -\frac{b}{k} \pm \frac{\mathrm{e}^{-k\widehat{C}}}{k} \mathrm{e}^{kx} = -\frac{b}{k} + A\mathrm{e}^{kx},$$

where $A = \pm \frac{\mathrm{e}^{-k\widehat{C}}}{k}$ is a non-zero constant. By adding $A = 0$, we also include the stationary solution $y \equiv -\frac{b}{k}$ in the formula for $y(x)$. $\square$

4.1.3. Theorems on solutions of the equation $y' = g(y)$. Stationary solutions.

Theorem 1 (on solutions of an autonomous first-order differential equation)**.** *Let the function $g(y)$ on the right-hand side of the differential equation* (1),

$$y' = g(y),$$

be continuous and non-zero on some interval J. Then, first, *if $y = y(x)$ is a solution of equation* (1) *on an interval I, then it is invertible on I, and its* *inverse function* $x = x(y)$ *satisfies the "symmetric" differential equation* (3),

$$x'_y = \frac{1}{g(y)};$$

and second, *if a function* $x = x(y)$ *satisfies equation* (3) *on* J*, then it is invertible on* J*, and the inverse function* $y = y(x)$ *is a solution of equation* (1).

Proof. First of all, note that the condition that g is non-continuous and non-zero on J implies that g *does not change its sign* on this interval, i.e., is everywhere positive or everywhere negative on J (this follows, for example, from the *intermediate value theorem*). Therefore, if $y = y(x)$ satisfies equation (1) on the interval I, then $\forall x \in I$ the derivative $y'(x) = g(y(x))$ has the same sign. Hence, the function $y = y(x)$ is strictly increasing or strictly decreasing on I and therefore *invertible* on I. According to the *inverse function derivative theorem* (see textbooks on algebra and elementary analysis), the inverse function $x = x(y)$ is differentiable at every point of its domain (it coincides with the range $J_1 \subset J$ of $y(x)$), and

$$x'(y) = \frac{1}{y'(x)|_{x=x(y)}} = \frac{1}{g(y(x))|_{x=x(y)}} = \frac{1}{g(y)},$$

since $\forall y \in J_1 \ y(x(y)) = y$ (the composition of two mutually inverse functions translates an initial point into itself). Thus, the first part of the statement of the theorem is proved.

The second part follows from the first; we only need to "swap" equations (1) and (3) and the variables x and y. □

Corollary 1 (existence and uniquemess theorem for solutions of an autonomous differential equation)**.** *If the function* $g(y)$ *on the right-hand of the differential equation* (1),

$$y' = g(y),$$

is continuous and non-zero on some interval J*, then for any initial condition* $y(x_0) = y_0 \in J$ *(*$x_0 \in \mathbb{R}$ *being an arbitrary number) there* exists *a solution* $y = y(x)$ *of the differential equation* (1) *which is defined on some interval* I *containing the point* x_0 *and satisfies the given initial condition, and this solution is* unique (*in the sense that any two solutions with this initial condition coincide in some neighbourhood of* x_0); *this solution satisfies* Barrow's formula (4):

$$x - x_0 = \int_{y_0}^{y(x)} \frac{du}{g(u)}.$$

Proof. Since the function $f(y) = \frac{1}{g(y)}$ is continuous, the *uniqueness* follows from the *uniqueness theorem for solutions of a differential equation of* *the form* (2) (§ 2.1, Theorem 1), the *existence* follows from the corresponding *existence theorem for solutions* of such an equation (§ 2.3, Corollary 1), and the formula follows from Barrow's usual formula (or directly from the *Cauchy–Moigno theorem*; § 2.3, Theorem 4). □

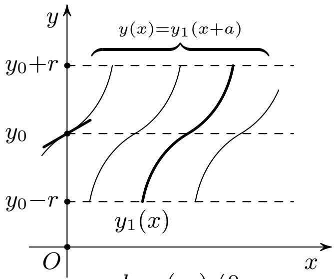
Corollary 2 (existence and uniqueness theorem in the neighbourhood of a non-singular point)**.** *If the function* $g(y)$ *on the right-hand side of the differential equation* (1), $y' = g(y)$, *is continuous and non-zero at some point* $y_0 \in \mathbb{R}$, *then for any* $x_0 \in \mathbb{R}$ *in some neighbourhood of* x_0 *there* exists *a solution* $y = y(x)$ *of equation* (1) *satisfying the initial condition* $y(x_0) = y_0$, *and this solution is* unique: *it satisfies* Barrow's formula (4).

Proof. By the continuity of $g(y)$ and by the condition $g(y_0) \neq 0$, the function g is non-zero (and even *has a constant sign*) in some neighbourhood J of y_0. Hence, the preceding corollary is applicable to equation (1) on J. □

If at a point y_0 the value of the function g is non-zero, $g(y_0) \neq 0$, then such a point is called a *non-singular* point for the differential equation (1). Corollary 2, in view of its *proof*, can be restated as follows: *a non-singular point* y_0 *has a neighbourhood* $J = (y_0 - r, y_0 + r)$ *such that through any point* $(x; y)$ *belonging to the* strip *of the* Oxy *plane enclosed between the horizontal lines* $y = y_0 - r$ *and* $y = y_0 + r$ *there passes a single integral curve of the directional field of equation* (1). Moreover, since the directional field $k = g(y)$ is constant along the horizontals, i.e., does not change under parallel translations along the Ox axis, all integral curves in this strip can be obtained from a single one by translations along the Ox axis (Fig. 50), and the corresponding solutions can be written via a single solution $y_1(x)$ by the formula $y(x) = y_1(x + a)$ (the value of $a \in \mathbb{R}$ being arbitrary). However, the same can aslo be seen from the integral formula (4).

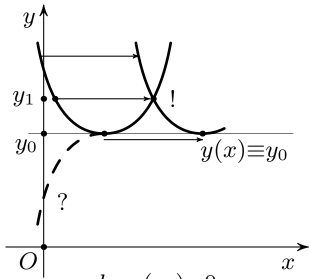
The points y for which the value of the function $g(y)$ *is zero* are called *singular points* of the differential equation (1). Each singular point y_0 corresponds to the so-called *stationary solution*, a constant: $y(x) \equiv y_0$. The question arises: are there any *non-stationary* solutions such that for some x_0 the value of $y'(x_0)$ is 0, i.e., $g(y(x_0)) = 0$?

Essentially, this is a question of *uniqueness of the solution* with initial condition $y_0 = y(x_0)$ at a singular point. In this case, formula (4) for solutions is no longer applicable, because the integral in it may not exist. Next, it follows from Corollary 2 that the graph of a nonstationary solution with this property cannot be tangent to the horizontal line $y = y_0$ (the graph of a stationary solution) and at the same time lie *on the same side* of this horizontal line in the neighbourhood of the tangency point x_0, similar to the position of the parabola $y = x^2$ with respect to the line $y = 0$ (the Ox axis):



$k = g(y_0) \neq 0$

Fig. 50.



$k = g(y_0) = 0$

Fig. 51.

as pointed out in the discussion following Corollary 2, for instance, the left "branch" of the corresponding integral curve would map, under a parallel translation to the right, to the integral curve intersecting the right branch (Fig. 51), and this contradicts the uniqueness of solutions in a band around the non-singular point y_1.

However, it is not too clear whether the graph of the solution can be located *on different sides* of the horizontal line $y = y_0$ tangent to it, like the graph of the function $y = x^3$ (with respect to the Ox axis). Try to figure this out on your own. (We will return to the question of the uniqueness of solutions at the end of the section, after considering a few particular examples.)

Another question related to singular points and their corresponding stationary solutions is that of their *stability*, i.e., of how the solutions of equation (1) behave if a given initial value $y(x_0) = y_1$ is close to a singular point y_0. For now, we only note that for the homogeneous linear equation $y = ky$ that we have investigated completely, we know the following. First, a solution with zero initial condition is unique: it is the stationary solution $y \equiv 0$. Second, this solution is *stable* for $k < 0$ (in this case, not only "close" but even *all* solutions $y = y(x)$ approach the stationary solution as $x \to +\infty$) and *unstable* for $k > 0$ (Figs. 52 and 53).

4.1.4. Autonomous equations as evolutionary models (reminder). In Ch. I, first-order differential equations, including autonomous equations of the form (1), were considered as *mathematical models of evolutionary processes*: if *time* t is taken as the independent variable x and the quantity describing the *state of the system* is considered as a *function of time* $x = x(t)$, then known information about *rate* v of change of $x(t)$ can be written as a *first-order ordinary differential equation* $x' = v(x, t)$. If the system under consideration is autonomous, i.e., does not undergo external influences, then the rate of evolution *should not* depend on time (again: *"the*

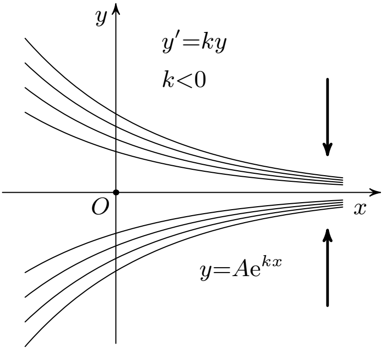


Fig. 52.

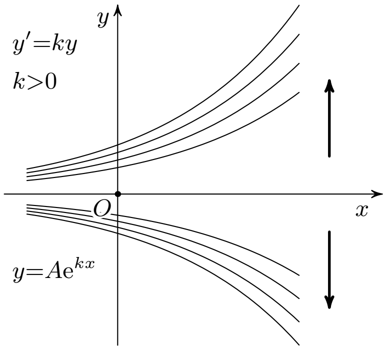


Fig. 53.

laws of nature do not depend on time"), so such a system is modelled by an *autonomous* differential equation of the form

$$x' = v(x).$$

In this notation, the above formulae, which can be used to find the *evolution law*, i.e., the solution $x = x(t)$, are rewritten as

$$t'_x = \frac{1}{v(x)} \quad \Leftrightarrow \quad t = \int \frac{dx}{v(x)},$$

or

$$t - t_0 = \int\limits_{x_0}^{x} \frac{du}{v(u)}.$$

By analogy with Newton's differential equation describing the motion of a point particle along the Ox line in a given *force field* $F(x)$, the equation $x' = v(x)$ can be interpreted in terms of mechanics as the motion of a particle in a *velocity field* on the Ox line given by the function $v(x)$ (with *velocity direction* determined by the *sign* of $v(x)$). Then singular points of this differential equation, i.e., zeros of the function v, which correspond to *stationary solutions*, can be called *equilibrium points* of the one-dimensional vector field $v(x)$, and the equilibrium at them can be either stable or unstable. If the field $v(x)$ is *smooth*, i.e., the function $v(x)$ has a continuous derivative, then its equilibrium point x_0 is called *non-degenerate* if the derivative $v'(x)$ at this point is non-zero. Then *in a small enough neighbourhood* of x_0 the field $v(x)$ can be replaced by its *linear approximation*: $v(x) \approx v(x_0) + v'(x_0)(x - x_0) = k(x - x_0)$ $(v(x_0) = 0,\ k = v'(x_0)$; see § 3.1). In new coordinates $\tilde{x} = x - x_0$, we replace equation (1) by the linear equation $\tilde{x}' = k\tilde{x}$ (it is called the *linear approximation for the differential equation* (1)). Then from the remark at the end of the preceding subsection we obtain a simple *stability criterion* for the equilibrium: a non-degenerate equilibrium point x_0 (where $v(x_0)=0$) is stable whenever $k=v'(x_0)<0$ and unstable whenever $k>0$.

We will accept this fact without a rigorous proof; it will help to better understand the motion of an "evolving particle" in a given velocity field. In general, an "evolutionary interpretation" of a differential equation of the form (1) using the velocity field as a vector field on a straight line helps in the qualitative analysis of differential equations.

4.1.5. Example: explosion equation $y' = \alpha y^2$. Now let us describe and analyse the simplest *nonlinear model* of evolution. In Ch. III (§ 3.1) we considered a model of the growth of a population of some individuals, in which the *growth rate is proportional to the current number of individuals*, i.e., the population size $N = N(t)$; the corresponding differential equation is linear: $N'(t) = \alpha N(t)$, and it has been fully analysed. Suppose now that the rate of increase is proportional not to the number of individuals but to the *number of pairs of individuals.* From N individuals, $\binom{N}{2} = \frac{N(N-1)}{2}$ pairs can be formed; neglecting here the "linear component" $-\frac{N}{2}$ as compared to the quadratic component, we obtain that the *growth rate is proportional to N^2*. If we consider individuals (or objects) of two types, half of type A and half of type B, each half of $\frac{N}{2}$ individuals, then the number of pairs of type (a, b) is $\frac{N}{2} \cdot \frac{N}{2} = \frac{N^2}{4}$, and again the growth rate N' turns out to be proportional to N^2. Thus we arrive at the *quadratic model* (Sec. 3.1.5, Example 5), the equation

$$N'(t) = \alpha N^2(t) \quad (\alpha > 0), \tag{5}$$

which also describes physicochemical reactions where the rate of reaction is proportional to the amounts of molecules of two "reactants", $\frac{N}{2}$ molecules each. For this and one other reason, discussed below, equation (5) is called the *explosion equation.*

Before solving this equation, we make a change of variables: if we multiply both sides of the equality by α, we get the equation $\alpha N' = \alpha^2 \cdot N^2$, i.e., $(\alpha N)' = (\alpha N)^2$, or

$$x' = x^2, \tag{6}$$

where $x = \alpha N$ (we may think that individuals are counted in "clumps": since $\alpha < 1$ [this coefficient is something like the *decay probability* in the model from § 3.1], the value $x = 1$ corresponds to the number $N = \frac{1}{\alpha} > 1$ of individuals).

The velocity field $v(x) = x^2$ defined by equation (6) has one singular point $x = 0$, and at all other points it has positive direction (see Fig. 54), so if the initial condition is $x_0 = x(t_0) < 0$, the integral curves $x = x(t)$ must approach the *stationary solution* $x(t) \equiv 0$, and if $x_0 > 0$ (this case is what matches "reality"), move away from it. Thus, in this case *the equilibrium position* $x = 0$ *is neither stable nor unstable.*

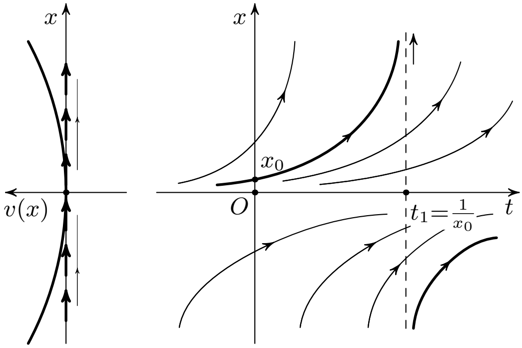


Fig. 54.

Now let us integrate the differential equation (6) for $x \neq 0$ following the "scheme" (2) and passing to the symmetric equation:

$$t'_x = \frac{1}{x^2} \quad \Leftrightarrow \quad t = -\frac{1}{x} + C \quad \Leftrightarrow \quad \frac{1}{x} = C - t \quad \Leftrightarrow \quad x = \frac{1}{C - t}. \tag{7}$$

This is the *general non-singular solution* of the differential equation (6). Given a (non-zero) initial condition at zero, $x(0) = x_0 \neq 0$, from relations (7) we find the value of the constant C and the "evolution law" $x = x(t)$ under this initial condition:

$$\frac{1}{x_0} = C \quad \Rightarrow \quad x(t) = \frac{1}{\frac{1}{x_0} - t} = \frac{x_0}{1 - x_0 t}.$$

Thus, we obtain a *fractional linear dependence*, whose graph consists of two branches of a hyperbola with a horizontal asymptote (at $\pm\infty$) $x = 0$ (it corresponds to a stationary solution and, unlike the case of the linear equation $x'(t) = kx$, cannot be included in the general solution, the family of functions (7)), and, *most important*, with the vertical asymptote $t = t_1 = \frac{1}{x_0}$: as $t \to t_1$, the value of $x(t)$ tends to $+\infty$ (see Fig. 54). It turns out that *over a finite time* equal to $t_1 = \frac{1}{x_0}$, the "conditional population size" *becomes infinite*, no matter how small the initial population size (i.e., the initial condition $x_0 > 0$)! This is the "explosion".

Note that the second branch of the hyperbola, with $t > t_1$, also satisfies equation (6): for it we have $x(t) < 0$, and starting "from $-\infty$" this solution tends to a stationary solution. Since the branches of the hyperbolas given by equation (7) cover both the upper and lower half-planes without having common points with the Ot axis, a *"global" uniqueness theorem for solutions* satisfying any initial condition $x(t_0) = x_0$ is valid for this equation.

Finally, we note that in *real-world situations* modelled by equations (5) and (6), *infinite growth in finite time* does not occur, because for large values of N or $x(t)$ such a model "does not work", i.e., does not match reality.

4.1.6. Example: logistic equation $y' = y(a - y)$. Moreover, the linear (or *exponential*) *model of evolutionary processes* analysed in Ch. III, for example, the equation $N' = \alpha N$ for population growth, also *does not work* for large values of $N(t)$, because it does not take into account the factor of *limited resources* (food, living space, etc.). The first approximation to reality is the so-called *logistic model*,[1] in which the coefficient α *depends* on N and is negative for large values of N; in the simplest case, α is assumed to be *linear* in N: $\alpha(N) = a - bN$, where $a, b > 0$. The *continuous logistic model* is described by the differential equation

$$N' = \alpha(N) \cdot N = N(a - bN),$$

or, after multiplying both parts of the equation by b and introducing a new variable $bN = x$,

$$x' = x(a - x). \tag{8}$$

Thus, the vector field of the logistic differential equation (8) is given by the *quadratic function* $v(x) = x(a - x)$, which is positive at $0 < x < a$ and negative outside the segment $[0, a]$ (Fig. 55). Hence, the singular point $x = 0$, in the neighbourhood of which the linear approximation of the function $v(x)$ is written in the form

$$v(0 + h) = v(h) = h(a - h) = ah - h^2 \approx ah,$$

i.e., as a linear function with a *positive* slope a (it characterises the *growth* of the population at low population size), is *"repelling"*, i.e., corresponding to the *unstable* stationary solution $x(t) \equiv 0$. The second singular point, $x = a$, in the neighbourhood of which the linear approximation $v(x)$ is written as a linear function with a *negative* slope, because for $x = a + h$ we have

$$v(x) = v(a + h) = (a + h)(a - (a + h)) = -(a + h)h = -ah - h^2 \approx -ah,$$

is *"attractive"*, i.e., it corresponds to the *stable* stationary solution $x(t) \equiv a$. From these remarks, the *qualitative* behaviour of the integral curves is clear; namely, besides the two stationary solutions, there "must be" solutions of the *three types* shown in Fig. 55: $x(t) > a$, and the solution tends to the

[1] This model of biological population growth was first proposed by P.F. Verhulst (1845), so the corresponding *logistic equation* (8) is also called the *Verhulst* equation.

singular point $x=a$ (seemingly *exponentially*); $x(t)<0$, and the solution "moves away" from the singular point $x=0$, remaining negative; $0<x(t)<a$, and the solution "rushes" from the singular point $x=0$ to the singular point $x=a$ (again *exponentially*).

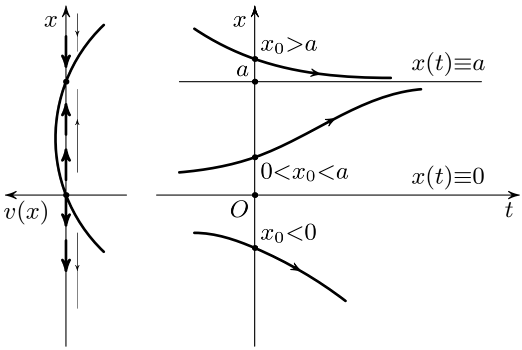


Fig. 55.

From the *qualitative description* of the behaviour of the system, let us move to *precise "quantitative"* analysis: let us integrate the symmetric differential equation

$$t'_x = \frac{1}{x(a-x)}. \tag{9}$$

To find the antiderivative for the function $f(x)=\dfrac{1}{x(a-x)}$, we represent it as a *linear combination of partial fractions* $\dfrac{1}{x}$ and $\dfrac{1}{x-a}$ with *undetermined* coefficients A and B. From the equality

$$f(x)=\frac{-1}{x(x-a)}=\frac{A}{x}+\frac{B}{x-a}=\frac{A(x-a)+Bx}{x(x-a)}=\frac{(A+B)x-aA}{x(x-a)}$$

we obtain a system of equations from which these coefficients can be found:

$$\begin{cases} A+B=0, \\ -aA=-1 \end{cases} \Leftrightarrow \begin{cases} A=\frac{1}{a}, \\ B=-A=-\frac{1}{a}; \end{cases}$$

hence,

$$f(x)=\frac{-1}{x(x-a)}=-\frac{1}{a}\left(\frac{1}{x-a}-\frac{1}{x}\right).$$

Now we find solutions of the differential equation (9):

$$\begin{aligned} t &= \int \frac{dx}{x(a-x)} = -\frac{1}{a}\int\left(\frac{1}{x-a}-\frac{1}{x}\right)dx \\ &= -\frac{1}{a}(\ln|x-a|-\ln|x|)+\widehat{C} \quad \Leftrightarrow \quad \ln|x-a|-\ln|x| = -at+a\widehat{C}. \end{aligned} \tag{10}$$

It is easier to analyse the solutions $x = x(t)$ *separately* in each of the above-mentioned cases, with the initial condition at zero being $x_0 = x(0) > a$, $0 < x_0 < a$, and $x_0 < 0$.

In the region (half-palne) $x > a$, the equation for the solutions of (9) is

$$\ln(x-a) - \ln x = -at + aC \quad \Leftrightarrow \quad \ln\frac{x-a}{x} = -at + aC$$
$$\Leftrightarrow \quad \frac{x-a}{x} = 1 - \frac{a}{x} = \mathrm{e}^{-at}\cdot \mathrm{e}^{aC} = A\mathrm{e}^{-at} \quad \Leftrightarrow \quad \frac{a}{x} = 1 - A\mathrm{e}^{-at},$$

whence we obtain the final answer:

$$x(t) = \frac{a}{1 - A\mathrm{e}^{-at}}, \tag{11}$$

where $A = \mathrm{e}^{aC} > 0$ is a constant which is found from the initial condition $x(0) = x_0 > a$. Substituting $t = 0$ into the equality $1 - \frac{a}{x} = A\mathrm{e}^{-at}$, we indeed obtain a positive constant $A = 1 - \dfrac{a}{x_0}$, which is, moreover, less than 1. Hence, for $t > 0$ the denominator of the fraction (10) does not vanish and, being strictly increasing, tends to 1 (since $A\mathrm{e}^{-at} \to 0$ as $t \to +\infty$), and the solution (11) itself, being decreasing, tends to a stable stationary solution, the constant $x \equiv a$.

However, the denominator vanishes at some *negative* $t = t_1$:

$$1 - A\mathrm{e}^{-at} = 0 \quad \Leftrightarrow \quad \mathrm{e}^{-at} = \frac{1}{A} \Leftrightarrow -at = -\ln A \quad \Leftrightarrow \quad t = t_1 = \frac{\ln A}{a}$$

(this t_1 is indeed negative, since $0 < A < 1$). Hence, for this value of t the solution (11) is undefined (!), and for $t < t_1$ it is negative! Since the exponent e^{-at} tends to $+\infty$ as $t \to -\infty$, we obtain that $x(t) \to 0-$: the horizontal asymptote at $-\infty$ does not coincide with the asymptote at $+\infty$. The line $t = t_1$ is the vertical asymptote for the *pair* of integral curves of our differential equation given by equation (11) (see Fig. 56).

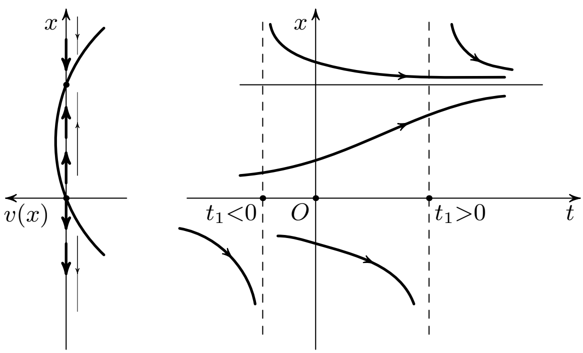


Fig. 56.

Note that expressing the "critical time" through the initial value $x_0 > a$ yields

$$t_1 = \frac{\ln A}{a} = \frac{1}{a}\ln\Big(1 - \frac{a}{x_0}\Big),$$

so t_1 tends to $\frac{1}{a}\ln 1 = 0$ as $x_0 \to +\infty$: the pair of integral curves shifts to the right so that the vertical asymptote coincides "in the limit" with the Ox axis ($t = 0$). But we can shift the integral curves and asymptotes even further: as usual (i.e., for any *autonomous* differential equation of the form (1)), we obtain integral curves which, however, correspond now to *negative* initial conditions $x_0 = x(0)$. Their branches lying in the half-plane $x > a$ *do not match any initial conditions at the point* $t = 0$; the "negative" branches are solutions that in finite time $t_1 > 0$ go to $-\infty$.

It remains to consider solutions whose graphs lie in the strip $0 < x < a$. In this case, equation (10) of the solutions has the form

$$\ln(x-a) - \ln x = -at + aC \quad \Leftrightarrow \quad \ln\frac{a-x}{x} = -at + aC$$
$$\Leftrightarrow \quad \frac{a-x}{x} = \frac{a}{x} - 1 = \mathrm{e}^{-at}\cdot \mathrm{e}^{aC} = A\mathrm{e}^{-at} \quad \Leftrightarrow \quad \frac{a}{x} = 1 + A\mathrm{e}^{-at},$$

whence it follows that

$$x(t) = \frac{a}{1 + A\mathrm{e}^{-at}}, \tag{12}$$

where the positive constant under the initial condition $x(0) = x_0 \in (0, a)$ is this time $A = \frac{a}{x_0} - 1$ and can take any value $A > 0$. The denominator of the fraction (12) cannot vanish, the corresponding solution $x(t)$ is well defined for all $t \in \mathbb{R}$, it is strictly increasing on the entire axis (since the denominator is positive and decreasing), and we have $x(t) \to a$ as $t \to +\infty$ and $x(t) \to 0$ as $t \to -\infty$ in accordance with our qualitative "predictions".

From the analysis of the logistic model we conclude that, whatever the initial population size $N_0 = N(0)$ (or the initial condition $x_0 = x(0) > 0$), the evolution *stabilises* over time, i.e., the system asymptotically approaches a stable equilibrium state $N \equiv \frac{a}{b}$ (or $x \equiv a$) depending only on the population parameters (but not on the initial conditions!).

The above consideration, as in the previous example (for the explosion equation), implies that *for the logistic equation, the global singularity theorem is valid.* But is this always the case?

4.1.7. Example of non-uniqueness of solutions of the equation $y' = g(y)$. Differentiating the function $x = t^3$ ($x' = 3t^2$), expressing t through x and substituting this expression into the formula for x',

$$t = \sqrt[3]{x} \quad \Rightarrow \quad x' = 3t^2 = 3(\sqrt[3]{x})^2 = 3\sqrt[3]{x^2},$$

we obtain that *the cubic function* $x(t) = t^3$ *is a solution of the differential equation*

$$x' = 3\sqrt[3]{x^2} \quad \Leftrightarrow \quad x' = 3|x|^{\frac{2}{3}}, \tag{13}$$

and moreover, this solution satisfies at $t = 0$ *the* zero initial condition $x(0) = 0$. So, finally we have found a solution of the autonomous equation with *continuous* right-hand side which is different from the stationary one ($x \equiv 0$)!

Let us apply the standard integration scheme (13) to equation (3): integrate the symmetric differential equation

$$t'_x = \frac{1}{3|x|^{\frac{2}{3}}} \quad \Leftrightarrow \quad t = \int \frac{dx}{3|x|^{\frac{2}{3}}} = \sqrt[3]{x} + C \quad \Leftrightarrow \quad x = (t - C)^3 \tag{14}$$

(the antiderivative of $f(x) = \dfrac{1}{3|x|^{\frac{2}{3}}}$ was found as follows: for $x > 0$, the function $f(x)$ is $\dfrac{1}{3|x|^{\frac{2}{3}}} = \frac{1}{3}x^{-\frac{2}{3}}$, and the antiderivative of any power function x^α for $x > 0$ and $\alpha \neq -1$ can be calculated by the formula

$$\int x^\alpha \, dx = \frac{x^{\alpha+1}}{\alpha+1} + C;$$

in this case $\int \dfrac{x^{-\frac{2}{3}}}{3}\, dx = \dfrac{x^{\frac{1}{3}}}{3 \cdot \frac{1}{3}} + C = x^{\frac{1}{3}} + C = \sqrt[3]{x} + C$, and the function $F(x) = \sqrt[3]{x}$ is everywhere defined and odd, so its derivative is even; therefore, the equality $F'(x) = f(x) = \dfrac{1}{3|x|^{\frac{2}{3}}}$ for $x > 0$ implies, by the evenness of f, that for negative values of x the derivative of $F'(x)$ is equal to $f(x)$).

Although we should have put $x \neq 0$ in the symmetric equation, all the obtained solutions intersect the Ot axis, i.e., the "forbidden" line $x = 0$; on the other hand, the family of solutions (14) does not include the stationary solution $x(t) \equiv 0$. Certainly, it can be added; it turns out that, besides this solution, the zero initial condition $x(t_0) = 0$ is also satisfied by an *non-stationary solution* of the form (14); indeed, by substituting $t = t_0$ and $x = 0$ into the "final" formula $x = (t - C)^3$, we find such a solution: $x(t) = (t - t_0)^3$. However, a doubt arises: did we find all solutions by integrating like when deriving formulae (14)?

It turns out that not all, and in general, functions of the form $F(x) = \sqrt[3]{x} + C$ *do not have a derivative* at the point $x = 0$, so they *are not solutions* of the symmetric differential equation; moreover, this equation itself does not make sense at $x = 0$. The above integration is valid only in the half-planes $x > 0$ and $x < 0$, so that the constant C must also be two-valued: $C = \widehat{C} = C_1, C_2$, depending on the sign of x, and the set of solutions in these half-planes should be written by the "compound" formula:

$$x(t)=\begin{cases}(t-C_1)^3 & \text{if } x>0, \text{ i.e., } t>C_1;\\(t-C_2)^3 & \text{if } x<0, \text{ i.e., } t<C_2.\end{cases} \tag{15}$$

If $C_1=C_2=C$ in this formula, then, assuming $x(C)$ to be zero, we obtain a solution differentiable at all $t\in\mathbb{R}$. A function defined by formulae (15) can be extended to a non-stationary solution on the entire t-axis for any values of the constants only if $C_1>C_2$ (how can we do this?). We can also augment the lower or upper "branches" of the solution with a zero constant; anyway, *for* any *initial condition* $x(t_0)=x_0\in\mathbb{R}$ *there are* infinitely many *solutions of the differential equation* (13) *satisfying this initial condition*!

The non-uniqueness of the solution in this case is explained by the fact that the velocity vector field $v(x)=3|x|^{\frac{2}{3}}$ decreases (in absolute value) "not fast enough" (say, as compared to the fields $v(x)=kx$ or $v(x)=x^2$) as $x\to 0$ (Fig. 57), so an imaginary moving particle can *in a finite time* "reach" a singular point starting from a non-singular one (or vice versa).[1] This can be seen from Barrow's formula for the equation $x'=v(x)$, which is used to calculate the time taken for a particle to move from position x_0 to position x in a given velocity field $v(x)$:

$$t-t_0=\int_{x_0}^{x}\frac{du}{v(u)}.$$

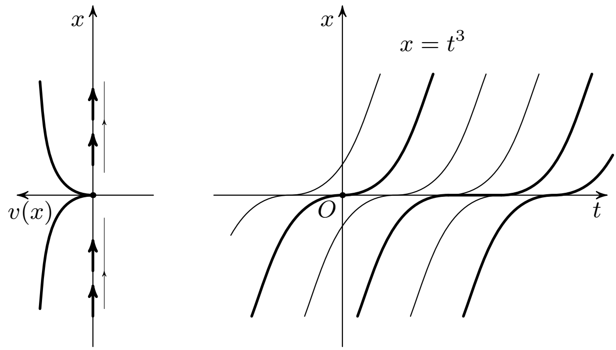


Fig. 57.

[1]At first glance, the difference between these fields is that the field of the differential equation (14) is not differentiable at $x=0$, but this is not the key point; for example, the velocity field $v(x)=|x|$ is not differentiable at zero either, but the "global uniqueness theorem" for the corresponding differential equation $x'=|x|$ is valid (prove this!).

Substituting here $t_0 = 0$, $x_0 = 0$, and $v(u) = 3u^{\frac{2}{3}}$, we arrive at the *improper integral*

$$t = \frac{1}{3}\int_0^x \frac{du}{u^{\frac{2}{3}}},$$

which in this case *converges* in the sense of the existence of the limit

$$\int_0^x \frac{du}{u^{\frac{2}{3}}} = \lim_{\varepsilon\to 0+} \int_\varepsilon^x \frac{du}{u^{\frac{2}{3}}},$$

and in the case $v(u) = u^2$ or $v(u) = |u|$ it *diverges* (the corresponding limits are infinite). Check this.

§ 4.2. Leibniz's Formalism and Variable Separable Equations

In the preceding section, when finding solutions to first-order autonomous differential equations, we began to use the *undefined integral* symbol $\int$ (like an integral, but without specifying the limits of integration — hence the name "indefinite"):

$$y'(x) = g(y) \quad\Leftrightarrow\quad x'(y) = \frac{1}{g(y)} \quad\Leftrightarrow\quad x = \int \frac{1}{g(y)}\,dy \overset{\text{des}}{=} \int \frac{dy}{g(y)} \qquad (1)$$

(the last expression is simply a *shorthand* for the previous one). In a similar way we can write a solution of the *basic equation of the integral calculus*

$$y'(x) = f(x) \quad\Leftrightarrow\quad y = \int f(x)\,dx,$$

so the scheme (1) is a particular case of the last expression (in the scheme (1) the variables x and y have been swapped). Using it, we will explain in more detail what is the "indefinite integral", the notion of (and notation for) which was introduced by Leibniz for *convenience of formal operation* in finding antiderivatives and solving differential equations. By comparing this formal expression with a substantive description of solutions of the equation $y' = f(x)$ with the use of the notion of antiderivative,

$$y'(x) = f(x) \quad\Leftrightarrow\quad y = F(x) + C,$$

where $F(x)$ is an arbitrary (any) antiderivative of $f(x)$, we obtain that the *meaning* of the indefinite integral is that we can write through it the *set of all antiderivatives* of the integrand $f(x)$. This understanding will suffice for now, but *there is more to it.*

The notion and notation, which cannot be considered methodologically satisfactory, go back to Leibniz. This notation, as we shall see, is very convenient from the point of view of *formalisation* (or *algorithmisation*) of integration. However, when we *start* with introducing the term *indefinite integral* (instead of *antiderivative*, as is commonly done in many calculus courses), the most important and profound fact of mathematical analysis that *differentiation and integration are mutually inverse operations*, or *the fundamental theorem of calculus*, i.e., the *relationship between the integral and the antiderivative*, turns into a tautology: we get a relationship between the integral and... integral!

To avoid confusion between the *integral* and *indefinite integral*, the famous French mathematician, physicist, and astronomer Pierre Simon Laplace proposed that the ordinary integral be called *definite* (1779). The exact analytical definitions of definite and indefinite integrals ((!) — read the beginning of this sentence again!) are attributed to the Parisian mathematician, Napoleon's associate in educational reforms in France, Academician Sylvestre François Lacroix (1765–1843).

4.2.1. Formal integration of the equation $y' = g(y)$. In the preceding section we presented a method for solving a differential equation of the form $y' = g(y)$ by reducing it to a *"symmetric"* differential equation $x'(y) = \frac{1}{g(y)}$. This method is very easily interpreted in *Leibniz's symbolism.* So far we have used both the terminology and notation of Lagrange: the *derivative* of a function $y = f(x)$ is written as y' or $f'(x)$. Newton called the derivative the *fluxion* and denoted it by a dot atop the function (*fluent*): $\dot{y}$ (see the commentary in Sec. 1.1.1). In contrast, Leibniz primarily considered the so-called *differentials*,[1] meaning by the *differential of a function* $y = f(x)$ the *formal expression*

$$dy = f'(x)\,dx, \tag{2}$$

which is usually interpreted as *the principal linear part of the function increment* (recall: *according to the linear approximation theorem* of Sec. 3.1.3, the increment of a function can be written as

$$\Delta y = f'(x)\Delta x + \alpha\Delta x,$$

where $\alpha = \alpha(x, \Delta x) \to 0$ as $\Delta x \to 0$; the first term in the formula for Δy is what is called the principal linear part of the increment, written as $dy = f'(x)\Delta x \overset{\text{des}}{=} f'(x)dx$). Operating with *symbols dx* and *dy* as with *variables*, Leibniz introduced the *notation* for the derivative $f'(x)$:

$$f'(x) \overset{\text{des}}{=} \frac{dy}{dx}$$

[1]This word was invented by Leibniz himself, being derived from Latin *differentia*, 'difference'.

(the right-hand side is read as "dee y dee x", not "dy divided by dx"). Note that under this approach, equality (2) becomes a *tautology*:[1]

$$dy = \frac{dy}{dx}\, dx;$$

that is, we "cancel out" dx, so to speak, as if this *symbol* were a variable or some *quantity*.

This purely *formal approach* to the derivative, *when used correctly*, leads, however, to *correct conclusions*. Leibniz became particularly convinced of the "rightness" of his approach when he noticed that formula (2) would hold even when x is not an independent variable but a function: if $x = x(t)$, then $dx = x'(t)dt$, and substituting these formulae into (2) gives the rule of differentiating (finding the differential of) a composite function: for $y = y(x(t))$ we have the equality

$$dy = y'(x)\, dx = y'(x(t))\, dx(t) = y'(x(t))x'(t)\, dt! \tag{3}$$

(We could also mention a quite "insane" application of Leibniz's symbolism to "derive" the chain rule, i.e., the formula for the derivative of a composite function: if $y = y(x(t))$, then

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt},$$

since dx is again "cancelled out"!). This property of the differential is called its *invariance*.[2]

Let us apply Leibniz's formal approach to a differential equation of the form $y' = g(y)$, for which we rewrite it in Leibniz's notation and transform it "without remorse". Namely,

$$\frac{dy}{dx} = g(y) \quad \Leftrightarrow \quad \frac{dy}{g(y)} = dx \quad \Leftrightarrow \quad \int \frac{dy}{g(y)} = \int dx.$$

The last equality, i.e., the coincidence of indefinite integrals, gives a relationship between the antiderivative of the integrand $\frac{1}{g(y)}$ on the left-hand side — denote it (any of the antiderivatives) by $G(y)$ — and the antiderivative of the integrand on the right-hand side, i.e., the antiderivative of $f(x) \equiv 1$, which we take to be $F(x) = x$. These antiderivatives *must differ by a constant*:

$$G(y) + C = x.$$

[1] Just in case we recall: Greek ταυτολογια, from Greek ταυτο [tauto], 'same', and λογος [logos], 'word', 'explanation'; a *tautology* is a repetition of just the same thing "in other words", or in this case, a statement which actually *repeats the definition*.

[2] That is, *independence* or *permanence*: Latin *invarians*, 'unchanging'. This term was coined in 1851 by the English mathematician James Joseph Sylvester (1814–1897; by the way, it was he who also introduced the term *discriminant*).

This yields the dependence of x on y, i.e., the function $x(y) = G(y) + C$, which obviously satisfies the "symmetric equation":

$$x'(y) = (G(y) + C)' = G'(y) = \frac{1}{g(y)}.$$

Thus, our formal transformations have led to a correct result. Could not we apply them in more general cases, for instance, for solving *nonautonomous* differential equations of the form $y' = F(x, y)$?

The notation invented by Leibniz for the integral $\int$ and differential df can be dated precisely (a rare case!): the integral sign was first written on 29 October 1675 (before that, Leibniz used for the integral $\int y$ the abbreviation *omn.y* of the expression "*omnia y*", i.e., "*all y*" [Latin]), and the differential was written on 11 November of the same year. Soon afterwards, Leibniz "brought" the differential under the integral sign in the notation "$\int y$": he began to systematically use the notation $\int y\,dx$; in this regard he wrote: *"I warn that care should be taken not to omit dx ... a mistake commonly committed, and which prevents going forward..."*, referring to possible applications of the invariance property of the differential.

Regarding the "symbolism", Bourbaki (Nicolas Bourbaki is the collective pseudonym of a group of mathematicians, authors of the fundamental *concept of the "architecture" of modern mathematics*) wrote in *Éléments d'histoire des mathématiques* (Elements of the History of Mathematics): "*...the properly Leibnizian tendency to a formal handling of symbols was to go on being emphasised throughout the whole of the* XVIII *century, well beyond what could be allowed by the resources of the analysis of the time. In particular, it must be recognised that the Leibnizian notion of a differential has, truth to tell no meaning; at the beginning of the* XIX *century, it fell into a disrepute from which it has only recovered little by little...*".

Indeed, Leibniz's appeal to integrate not just a function φ but the *expression* $\varphi(x)\,dx$ was far ahead of its time: the exact meaning of the expression $\varphi(x)\,dx$, as well as the *concept* of differential, was not given until the XX century. The integrand is now called in "higher mathematics" the *differential form* ($\varphi(x)\,dx$ is a 1-*form* on the real line $\mathbb{R}$), and the differential of a function is the so-called *"exact form"*.

In general, Leibniz paid very much attention to symbols and symbolism. Thus, in 1678 he wrote to the German mathematician, physicist, and philosopher Ehrenfried Walther von Tschirnhaus (1651–1708): *"In signs we must look to find convenience, which is most important when they express the inner nature of a thing in a few words and paint it as it were, for thus the labor of thinking is greatly reduced."* These aspirations of Leibniz are closely related to his essentially philosophical idea of creating a *"universal algorithm"*, which we shall discuss later on.

4.2.2. Separation of variables and Leibniz's formalism. The key point in the above formal solution of the differential equation $y' = g(y)$ is its reduction to the *equality between two differential forms*, $\psi(y)\,dy = \varphi(x)\,dx$ (on the left-hand side there is a form in y, and on the right-hand side there is a form in x), where after formally applying the same integral to the left- and right-hand sides, we get an *equality between indefinite integrals* written as a relation between functions (antiderivatives) of x and y: $\Psi(y)=\Phi(x)+C$, where C is an arbitrary constant. Obviously, for an equation of the form $y'=F(x,y)$, the above representation is possible only if the right-hand side of the differential equation can be written as a product of a function of x and a function of y: $F(x,y)=f(x)g(y)$. Then we say that $F(x,y)$ *admits separation of variables*, and for a *variable separable differential equation* we can make the following *formal transformations*:

$$\frac{dy}{dx}=f(x)g(y) \quad\Leftrightarrow\quad \frac{dy}{g(y)}=f(x)\,dx \quad\Leftrightarrow\quad \int\frac{dy}{g(y)}=\int f(x)\,dx.$$

From this we have now an *algebraic equation* relating x and y:

$$G(y)=F(x)+C, \quad \text{where} \quad G'(y)=\frac{1}{g(y)}, \quad F'(x)=f(x), \quad C=\text{const}.$$

It remains to express y through x from this equation (if possible), and we get a solution $y=y(x;C)$ of the original differential equation.

Example 1. Let us solve the differential equation $y'=6xy$. For nonzero values of y, we have

$$\frac{dy}{dx}=6xy \quad\Leftrightarrow\quad \frac{dy}{y}=6x\,dx \quad\Leftrightarrow\quad \int\frac{dy}{y}=\int 6x\,dx,$$

whence it follows that

$$\ln|y|=3x^2+C \quad\Leftrightarrow\quad |y|=\mathrm{e}^{3x^2}\mathrm{e}^C \quad\Leftrightarrow\quad y=\pm\mathrm{e}^C\mathrm{e}^{3x^2}.$$

Alternatively, this family of solutions can be written through an arbitrary nonzero multiplicative constant:

$$y=A\mathrm{e}^{3x^2}, \quad A\neq 0.$$

Furthermore, the function $y\equiv 0$ also satisfies the differential equation, and this solution can be included in the above family of solutions $y=A\mathrm{e}^{3x^2}$; it corresponds to $A=0$.

Recall that we *have not yet proved* that applying the *Leibniz formalism* to general variable separable equations always leads to a correct answer. But this is not a problem, since we can *directly check* that any function of the form $y=A\mathrm{e}^{3x^2}$ satisfies the differential equation $y'=6xy$. Indeed, for this function we have

$$y'=(A\mathrm{e}^{3x^2})'=A\mathrm{e}^{3x^2}\cdot(3x^2)'=A\mathrm{e}^{3x^2}\cdot 6x=6xy,$$

as required.

We can draw the family of all integral curves $y=y(x;A)$ of this differential equation in the Oxy plane; in this case we will call it the *phase portrait.*[1] Such a portrait for the equation $y'=6xy$ is shown in Fig. 58. Although this picture clearly shows that *through any point* $(x_0;y_0)$ of the Oxy plane *there passes a single integral curve*, it would also be nice to establish this fact analytically.

Task. Prove that

(1) *Any* solution of the differential equation $y'=6xy$ can be represented as $y=A\mathrm{e}^{3x^2}$;

(2) For *any* initial condition $(x_0;y_0)$ there exists a *unique* solution $y(x)$ of the differential equation $y'=6xy$ satisfying this initial condition (i.e., such that $y(x_0)=y_0$). □

Let us consider two more examples.

Example 2. The differential equation $y'=-\frac{y}{x}$ should have been considered only for $x\neq 0$, i.e, in the left and right (with respect to the Oy-axis) half-planes of the Oxy coordinate plane. At first, this is what we will suppose. Assuming also that $y\neq 0$, we apply to this equation the same formal method:

$$\frac{dy}{dx}=-\frac{y}{x} \quad\Leftrightarrow\quad \frac{dy}{y}=-\frac{dx}{x} \quad\Leftrightarrow\quad \int\frac{dy}{y}=-\int\frac{dx}{x},$$

whence we find that

$$\ln|y|=-\ln|x|+C \quad\Leftrightarrow\quad \ln|xy|=C\mathrm{e}^{C}.$$

Alternatively, the family of solutions can be written through an arbitrary non-zero constant:

$$y=\frac{A}{x}, \quad A\neq 0$$

($A=\pm\mathrm{e}^{C}\neq 0$; recall that $x\neq 0$). Besides, the *two* functions $y\equiv 0$, $x\neq 0$, (why *two*?) also satisfy this equation, and these solutions are included in the above family of solutions; they correspond to $A=0$.

Depicting the phase portrait of the equation $y'=-\frac{y}{x}$ (Fig. 59), we obtain a family of hyperbolas, or rather four families, one in each of the quadrants of the coordinate plane. These families are separated by two special integral curves: the half-lines $x<0$ and $x>0$ of the Ox axis, which play the role of *separatrices.*[2] To these separatrices we should add the rays $y>0$ and $y<0$ of the Oy axis and relax the restriction $x\neq 0$, i.e., reckon these rays among

[1] Previously, in § 1.2, we called the phase portrait the family of phase trajectories of a dynamical system in the $(x;v)$ (coordinate–velocity) phase plane. If the variable x is treated as *time*, then the Oxy plane is usually called the *extended* phase plane, and the family of solution graphs is referred to as the *generalised* or *extended* phase portrait.

[2] Latin *separatrix*, the feminine form of 'separator'.

integral curves, considering that the directional field of the differential equation $y' = -\frac{y}{x}$ for $x = 0$ and $y \neq 0$ is *vertical.* As for the point $(0; 0)$, at this point the directional field is *undefined*; this is a *singular point* of the equation. A singular point of this type is called a *saddle.*[1]

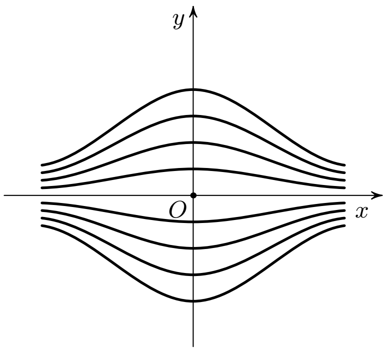


Fig. 58.

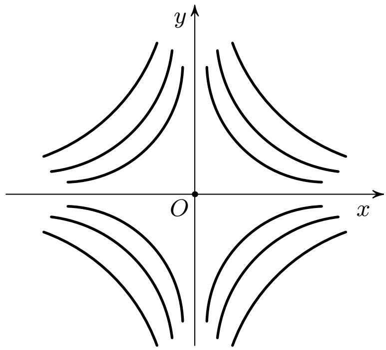


Fig. 59.

Note that the formula for integral curves can be rewritten as a *"conservation law"*:

$$H(x, y) = xy = A = \text{const.}$$

Similarly to the phase trajectories for Newton's conservative differential equations, which lie, as we shall see in Ch. V, on the *energy level curves* $E = E(x; v)$, the *integral curves under consideration lie on the level curves of the function* $H = H(x, y)$. The functions $H(x, y)$ that are *constant* "along the solutions" of the differential equation, i.e., such that for any solution $y = y(x)$ of this equation we have

$$h(x) = H(x, y(x)) = \text{const},$$

are called *first integrals* of the differential equation.[2] Thus, the first integral of the differential equation $y' = 6xy$ from Example 1 is the function $H(x, y) = y\mathrm{e}^{-3x^2}$ (explain!). Knowing the first integral of an equation of the form $y' = F(x, y)$ helps considerably in constructing and analysing the phase portrait of this equation.

Task. Prove that

(1) Every function of the form $y = \frac{A}{x}$ does *indeed* satisfy the original differential equation $y' = -\frac{y}{x}$;

(2) *Any* solution of this differential equation can be represented as $y = \frac{A}{x}$;

[1] We will explain the origin of this name later, in section 4.3.3.

[2] First integrals of Newton's differential equation are called *conservation laws* for the corresponding dynamical system.

(3) For *any* initial condition $(x_0; y_0)$ with $x_0 \neq 0$ there exists a *unique* solution $y(x)$ of the equation $y' = -\frac{y}{x}$ satisfying this initial condition. □

It should be noted that the differential equations in the examples above belong to the class of *homogeneous linear equations with variable coefficients*, i.e., equations of the form $y' = k(x)y$, which are discussed in detail in § 3.5 (see Theorem 1 and the exercises in this section). Thus, we could well solve these equations without using the Leibniz formalism. However, the next example does not belong to this class.

Example 3. Applying "Leibniz's method" (see comments at the end of the next subsection) to the differential equation $y' = -\frac{x}{4y}$, we obtain

$$\frac{dy}{dx} = -\frac{x}{4y} \quad \Leftrightarrow \quad 4y\,dy = -x\,dx \quad \Leftrightarrow \quad \int 4y\,dy = -\int x\,dx;$$

hence,

$$2y^2 = -\frac{x^2}{2} + C \quad \Leftrightarrow \quad y^2 + \frac{x^2}{4} = \frac{C}{2} = A = \text{const.}$$

Thus, we have found the *first integral* of this differential equation: $H(x, y) = \frac{x^2}{2^2}2^2 + y^2$. Its level lines are concentric (homothetic to each other) *ellipses* with the centre at the origin and with the ratio of semi-axes $2:1$, on which all the integral curves of the equation must lie. Taking into account that the right-hand side of the original equation is undefined for $y = 0$, we should think that each of the ellipses contains two integral curves:

$$y = \pm\sqrt{A - \frac{x^2}{4}}, \quad A = \text{const} > 0 \quad (|x| < A).$$

If, as in the previous example, we conventionally assume that for $y = 0$ and $x \neq 0$ the directional field of the differential equation $y' = -\frac{x}{4y}$ is vertical, then we may think that each of these ellipses is an *integral curve of the directional field* (or a *"generalised"* integral curve of the differential equation). Of course, the origin is a *singular point*, called a point of the *centre* type.

It is easy to prove that every function of the form $y = y(x) = \pm\sqrt{A - \frac{x^2}{4}}$ satisfies the original equation; however, this is somewhat "boring". So we will proceed in a different way: we will use the constancy of the first integral. Namely, we know that for any solution $y = y(x)$, the function

$$h(x) = H(x, y(x)) = \frac{x^2}{4} + y^2(x) \equiv A$$

is a constant; therefore, taking into account that $y(x) \neq 0$, we obtain

$$h'(x) = \frac{2x}{4} + 2y(x)y'(x) \equiv 0 \quad \Rightarrow \quad 2yy' = -\frac{x}{2} \quad \Rightarrow \quad y' = -\frac{x}{4y},$$

as required.

Task. Prove that

(1) *Any* solution $y = y(x)$ of the differential equation in question can be written as

$$y = y(x) = \pm\sqrt{A - \frac{x^2}{4}};$$

(2) For *any* initial condition $(x_0; y_0)$ with $y_0 \neq 0$ there exists a *unique* solution $y(x)$ of the equation $y' = -\frac{x}{4y}$ satisfying this initial condition. □

4.2.3. Theorems on variable separable equations. The reasoning from the last example shows how to justify the *Leibniz formalism* in the general case. First, let us formulate the corresponding statement in strict (*not formal*) mathematical language.

Theorem 1 (on solutions of variable separable equations)**.** *Let a function $f(x)$ be continuous on an interval I of the axis Ox, a function $g(y)$ be continuous and non-zero on an interval J of the Oy axis, $F(x)$ be an antiderivative of $f(x)$ on I, $G(y)$ be an antiderivative of $\frac{1}{g(y)}$ on J, and finally, let $y = \varphi(x)$, $x \in I$, be a function such that $\forall x \in I$ $y = \varphi(x) \in J$ and*

$$\forall x \in I \quad h(x) = G(\varphi(x)) - F(x) \equiv A = \text{const}.$$

Then this function $y = \varphi(x)$ is a solution of the differential equation

$$y' = f(x)g(y) \tag{4}$$

on the interval I.

Proof. Since $h(x) \equiv \text{const}$ on I, it follows that $h'(x) \equiv 0$ on I. But

$$h'(x) = G'(\varphi(x)) \cdot \varphi'(x) - F'(x) = \frac{\varphi'(x)}{g(\varphi(x))} - f(x) = \frac{\varphi'(x) - f(x) \cdot g(\varphi(x))}{g(\varphi(x))},$$

so on the interval I we have

$$\varphi'(x) - f(x) \cdot g(\varphi(x)) \equiv 0 \quad \Rightarrow \quad \varphi'(x) = f(x) \cdot g(\varphi(x));$$

i.e., the function $y = \varphi(x)$ satisfies the equation $y' = f(x)g(y)$, as required. □

Thus, any solution $y = y(x) = \varphi(x)$ (or $y = y(x; A)$) of the *algebraic*[1] *equation*

$$G(y) - F(x) = A \tag{5}$$

under the conditions of the theorem satisfies the *differential equation* (4). Let us prove also the converse.

Theorem 2 (on first integrals of variable separable equations)**.** *Let a function $f(x)$ be continuous on an interval I of the Ox axis, a function $g(y)$*

[1] In a wide sense (x and A in this equation should be regarded as *parameters*).

be continuous and non-zero on an interval J of the Oy axis, $F(x)$ be an antiderivative of $f(x)$ on I, and $G(y)$ be an antiderivative of $\frac{1}{g(y)}$ on J.

Then the function

$$H(x,y)=G(y)-F(x)$$

is a first integral of the differential equation (4) (*in a rectangular region* $R=\{(x;y)\mid x\in I, y\in J\}$); *i.e., for any solution $y=\varphi(x)$ of the differential equation* (4), $y'=f(x)g(y)$, *on I such that $\forall x\in I$ $y=\varphi(x)\in J$, the function $h(x)=G(\varphi(x))-F(x)$ is a constant.*

Proof. As above, we compute

$$h'(x)=G'(\varphi(x))\cdot\varphi'(x)-F'(x)=\frac{\varphi'(x)}{g(\varphi(x))}-f(x)=\frac{\varphi'(x)-f(x)\cdot g(\varphi(x))}{g(\varphi(x))}\equiv 0$$

on the interval I, since according to the differential equation (4) we have

$$\forall x\in I \quad \varphi'(x)=f(x)\cdot g(\varphi(x)) \quad \Rightarrow \quad \varphi'(x)-f(x)\cdot g(\varphi(x))\equiv 0,$$

which completes the proof. □

Thus, under the conditions of the theorem, *the differential equation* (4) *is equivalent to the algebraic equation* (5), and therefore, *the Leibniz formalism as applied to variable separable differential equations is justified.*[1]

Now let us explain why we call this solution method a *formalism.* The point is that the expression $\frac{dy}{dx}$ for the derivative $y'(x)$ *cannot be treated as a fraction*; this is just a *symbolic notation* for the derivative. It would be more correct to write

$$y'(x)=\frac{d}{dx}y,$$

where $\frac{d}{dx}$ is a "symbol" to denote the operation of "taking the derivative" (or a *differentiation operator*); actually, we could replace the "stroke" in the notation y' with this symbol,

$$y'(x)=y^{\frac{d}{dx}},$$

but, of course, this notation is too exotic (although it exactly conveys the *meaning* of the Leibniz notation!). The "operator" approach is very convenient for writing higher derivatives; thus, for the second derivative $y''(x)$, the "Leibniz" notation is

$$y''(x)=\frac{d}{dx}\left(\frac{d}{dx}y\right)=\left(\frac{d}{dx}\right)^2 y=\frac{d^2y}{dx^2}$$

[1] In the next section, the Leibniz formalism will be considered as applied to differential equations of a different type (more precisely, to *systems of differential equations* with two unknown functions).

(typically read as "dee two y dee x two", though some people read it as "dee square y dee x square").

By "splitting" the numerator and denominator, we formally operate with the expression $\frac{dy}{dx}$ as with a fraction — that is why we are talking about formalism. That kind of "operating" can well lead to nonsense, so one has to *"think about the sense"*,[1] *verify* the validity of particular results and *justify* the possibility of formal operation in each particular case (as is done above for variable separable differential equations).

The method of solving variable separable differential equations has been applied, moreover in a substantive (geometrical) form, as early as Isaac Barrow in his *Lectiones opticae* and *Lectiones geometricae* (Lectures on Optics and Geometry; 1669–1670). By the way, the same lectures contain the important inequality $(1+x)^n > 1+nx$ for $x>0$ and a natural $n>1$; this inequality nowadays bears the name of Jacob Bernoulli, who also published it, but in 1689. In the Leibniz notation, the method of Burrow consists in the transition

$$\frac{dx}{dy}=\frac{\psi(y)}{\varphi(x)} \quad \Leftrightarrow \quad \int \varphi(x)\,dx=\int \psi(y)\,dy.$$

Apparently independently of Barrow, in *Lectiones mathematicae, de methodo integralium...* (Mathematical Lectures on the Method of Integrals; 1691–1692) and *Modus generalis construendi omnes aequationes differentiales primi gradus* (General Method of Constructing all Differential Equations of the First Degree; 1694), Johann Bernoulli gave a method for solving several types of equations by separating the variables (see exercises). He was the first to use the term *separation* (Latin *separatio*), but only Leonhard Euler legitimised it (*separatio variabilium*). In 1749, d'Alembert applied separation of variables to find partial solutions ("standing waves"; see Ch. VI, § 6.2) of the *wave equation*; Fourier, after whom the method (for *partial* differential equations) is named, used such partial solutions to represent the general solution.

Leibniz's merit concerning variable separable equations (apart from methods for solving specific types of equations) was in inventing a *convenient way of formalising the separation of variables method.* In the late XIX century, the English engineer, physicist, and mathematician Oliver Heaviside (1850-1925) went much further, developing *symbolic* calculus and applying it to electrical and other physical calculations. However, he cared nothing for the "sense", i.e., for justifications, for which he was criticised many times by mathematicians; nevertheless, Heaviside once commented on it: *"Even Cambridge mathematicians deserve justice!"*

4.2.4. Singular points of differential equations. In Examples 2 and 3 we considered *singular points* of differential equations of the types *centre* and *saddle.* Without going into the question of *what*, in general, a "singular point" is, let us give some examples of differential equations with singular points of one more type, which is typical for the simplest variable separable equations.

[1] *"And the moral of* that *is — 'Take care of the sense and sounds will take care of themselves'"*, the Duchess told Alice. *Lewis Carroll*, Alice in Wonderland, Ch. IX.

Recall that Lewis Carroll was the pen name of the English mathematician, Oxford professor Charles Lutwidge Dodgson (1832–1898).

Example 4. Let us apply the Leibniz formalism to the differential equation $y' = \dfrac{3y}{x}$. For $y \neq 0$, we obtain

$$\frac{dy}{dx} = \frac{3y}{x} \quad \Leftrightarrow \quad \frac{dy}{y} = 3\frac{dx}{x} \quad \Leftrightarrow \quad \int \frac{dy}{y} = 3\int \frac{dx}{x},$$

whence it follows that

$$\ln|y| = 3\ln|x| + C \quad \Leftrightarrow \quad |y| = |x|^3 \cdot \mathrm{e}^C$$
$$\Leftrightarrow \quad y = \pm \mathrm{e}^C x^3 = Ax^3, \quad A \in \mathbb{R} \setminus \{0\}.$$

The case $A = 0$ corresponds to two more integral curves, the rays $y \equiv 0$, $x \neq 0$. Besides, assuming that for $x = 0$, $y \neq 0$, the directional field of the differential equation $y' = \dfrac{3y}{x}$ is vertical, we add two more "special" integral curves to the phase portrait (Fig. 60): the rays $y > 0$ and $y < 0$ of the Oy axis. In this case, the singular point $(0;0)$ of the differential equation is called a *node*: all its integral curves are "joined", "knotted" at the singular point. □

Example 5. Similarly, for the differential equation $y' = \dfrac{y}{4x}$ we find

$$\frac{dy}{dx} = \frac{y}{4x} \quad \Leftrightarrow \quad 4\frac{dy}{y} = \frac{dx}{x} \quad \Leftrightarrow \quad 4\int \frac{dy}{y} = \int \frac{dx}{x} \quad \Leftrightarrow \quad 4\ln|y| = \ln|x| + C$$
$$\Leftrightarrow \quad y^4 = |x| \cdot \mathrm{e}^C \quad \Leftrightarrow \quad y = \pm \mathrm{e}^C \sqrt[4]{|x|} = A|x|^{1/4}, \quad A \neq 0.$$

Adding to the phase portrait four more integral curves, the positive and negative rays of the coordinate axes, we again obtain a singular point of the *node* type (Fig. 61). □

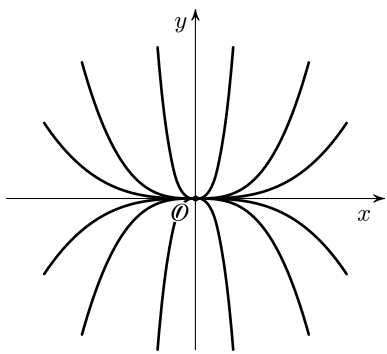


Fig. 60.

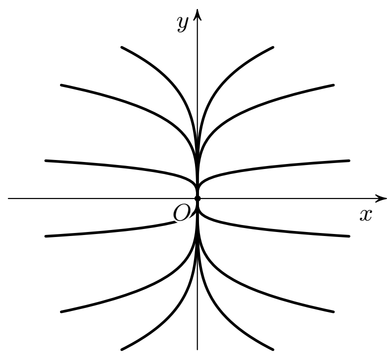


Fig. 61.

Thus, the simplest variable separable differential equations can have singular points of the types *centre*, *saddle*, and *node*. In the next chapter, we will consider one more type of singular points, review their general definition, and discuss their classification.

4.2.5. Separation of variables in linear equations of the form $y' = k(x)y\,[+f(x)]$. In § 3.5 we have already considered linear differential equations with variable coefficients

$$y' = k(x)y + f(x), \tag{6}$$

and discussed how to solve them using *d'Alembert's theorem* and *Lagrange's method* (variation of constants). Now we show how equations of the form (6) can be solved by *separation of variables.*

In the homogeneous equation

$$y' = k(x)y, \tag{7}$$

the variables are already separated, and the Leibniz formalism yields

$$\frac{dy}{dx} = k(x)y \quad \Leftrightarrow \quad \frac{dy}{y} = k(x)\,dx \quad \Leftrightarrow \quad |y| = \int k(x)\,dx = K(x) + C,$$

where $K(x)$ is any of antiderivatives of the variable coefficient $k(x)$ (which, of course, we assume to be *continuous*). Hence,

$$|y| = \mathrm{e}^{K(x)} \cdot \mathrm{e}^{C} \quad \Leftrightarrow \quad y = \pm\mathrm{e}^{C}\mathrm{e}^{K(x)} = A\mathrm{e}^{K(x)}, \quad A \neq 0.$$

Adding the *stationary solution* $y \equiv 0$ to this family, we obtain the general form of solutions of the homogeneous linear differential equation (7):

$$y = A\mathrm{e}^{K(x)}, \quad \text{where } A \in \mathbb{R} \text{ is any number.}$$

Quite naturally, we have obtained the same answer (and in fact, the same solution, in the sense of the "method of solution") as in § 3.5 (where we used *"variation of exponent"*, which reduced the equation of the form (7) to finding an antiderivative, i.e., to solving the differential equation $K'(x) = k(x)$).

The separation of variables in the non-homogeneous linear equation (6) is obtained by representing the sought-for function $y(x)$ as a product $u(x)v(x)$; then

$$y'(x) = u'(x)v(x) + u(x)v'(x),$$

so equation (6) can be rewritten as

$$u'(x)v(x) + u(x)v'(x) = k(x)u(x)v(x) + f(x)$$
$$\Leftrightarrow \quad u(x)v'(x) = \{(k(x)u(x) - u'(x)\}v(x) + f(x). \tag{8}$$

Let us find a function $u = u(x)$ such that the expression in curly brackets is zero, i.e., the differential equation $u'(x) = k(x)u(x)$ is satisfied (for the values of x under consideration). However, this is a homogeneous linear equation of the form (7), which can immediately be solved, as above, by separation of variables:

$$\frac{du}{dx} = k(x)u \quad \Leftrightarrow \quad u(x) = A\mathrm{e}^{K(x)}, \quad \text{where } K'(x) = k(x),\ A \in \mathbb{R}.$$

Taking the constant $A = 1$, for an arbitrary antiderivative $K(x)$ of $k(x)$ we obtain a *non-zero* solution $u(x) = \mathrm{e}^{K(x)}$ of the auxiliary equation $u' = k(x)u$. By substituting this solution into equation (8), we obtain a differential equation for the function $v = v(x)$, which can immediately be solved by integration:

$$\mathrm{e}^{K(x)}v'(x) = f(x) \quad \Leftrightarrow \quad v'(x) = f(x)\mathrm{e}^{-K(x)} \quad \Leftrightarrow \quad v(x) = F(x) + C,$$

where $F(x)$ is any antiderivative of the function $f(x)\mathrm{e}^{-K(x)}$. It remains to write the expression for the desired function $y = y(x)$:

$$y(x) = u(x)v(x) = \mathrm{e}^{K(x)}(F(x) + C) = F(x)\mathrm{e}^{K(x)} + C\mathrm{e}^{K(x)}.$$

Again, the answer is the same as in § 3.5 (where we used the variation of constants method).

It should be noted that Newton did not endeavour to find solutions of differential equations in elementary functions or even in quadratures; he wrote solutions in the form of infinite power series. Leibniz and his closest followers Jacob and Johann Bernoulli laid the foundation for the *classification* of differential equations and *methods* of their solution by reducing them to quadratures, i.e., to (indefinite) integrals. Their efforts were aimed primarily at reducing first-order differential equations to the form of variable separable equations.

One of the first to be found was a method for solving the so-called homogeneous differential equations (Johann Bernoulli, 1691–1692; Leibniz, 1693; see exercises). The above-discussed method for solving linear equations (by the substitution $y = u(x)v(x)$) is due to Leibniz (this method was described in a letter to L'Hôpital dated 6 January 1695, 80 years before Lagrange proposed the variation of constants method).

§ 4.3. Planar Differential Equations

In § 1.1 we have already considered differential equations in the Oxy plane, which can be written either in vector form,

$$\overline{r}' = \overline{V}(x, y, t) = \overline{V}(\overline{r}, t), \quad \overline{r} = \overline{r}(x; y),$$

or coordinate-wise, as systems of two scalar first-order differential equations

$$\begin{cases} x'(t) = A(x, y, t), \\ y'(t) = B(x, y, t). \end{cases}$$

Now we will consider them in more detail and learn how to solve ("integrate") some of them. Let us start from afar.

4.3.1. Planar equations, vector fields, phase portraits. All the first-order differential equations considered above were (or could be) mathematical models of evolution of systems whose *state* is described by a single parameter, $x=x(t)$, varying in time with a known *rate* $v=v(x,t)$, i.e., according to the ordinary differential equation $x'=v(x,t)$ (in other notation, $y'(x)=F(x,y)$).

Of course, this is an important but *simplest* case: states of real "systems" by far not always can be described by only one parameter; most often there are *many* parameters. We will confine ourselves to the case where the state is characterised by *two* quantities, say x and y, somehow varying with time. As above, suppose that we know some information about the *rates* of change of $x(t)$ and $y(t)$. This information can be written as a *system of first-order ordinary differential equations*:

$$\begin{cases} x'(t)=A(x,y,t), \\ y'(t)=B(x,y,t), \end{cases}$$

where the functions $A(x,y,t)$ and $B(x,y,t)$ of *three variables* are assumed to be known.

Let us make one more assumption: assume that the system under consideration is *autonomous*, i.e., it is not subject to external influences, and the evolution rate given by the functions A and B is *time-independent.* This means that $A=A(x,y)$, $B=B(x,y)$, and the evolution of such a system is described by an *autonomous system of differential equations* of the form

$$\begin{cases} x'(t)=A(x,y), \\ y'(t)=B(x,y). \end{cases} \tag{1}$$

Since a *pair* $(x;y)$ can be interpreted as the coordinates of a point X in the Oxy plane or as the coordinates of the corresponding vector $\mathbf{X}$ (or of the radius vector $\overline{OX}$) in this plane, we will call the system of differential equations (1) a *planar differential equation.* If we introduce one more vector, the "velocity vector" $\mathbf{V}=\overline{XV}$ attached to the point $X(x;y)$ and given by its coordinates $(A;B)=(A(x,y);B(x,y))$, or $\mathbf{V}(\mathbf{X})=(A(\mathbf{X});B(\mathbf{X}))$), then system (1) can indeed be written as a single *vector differential equation*

$$\mathbf{X}'=\mathbf{V}(\mathbf{X}) \quad (\mathbf{X}'=\mathbf{X}'(t)), \tag{2}$$

which looks (thanks to the notation!) quite analogous to the scalar autonomous equation $x'=v(x)$ (or $x'=x'(t)$; see § 3.1). However, this analogy does not go any further: the variables $\mathbf{X}$ (vector) and t (number) are by no means equivalent, so we cannot talk about a symmetric differential equation. Here we will again be helped by the *geometric interpretation* of the differential equation (2) and its solutions in the Oxy plane, which, as above, we will call the *phase plane* of the equation (or of the real-world evolving system described by it).

In this section, vectors will most often be written in (denoted by) *uppercase boldface letters*; this notation is commonly used in vector algebra, as well as in classical mechanics. This way of writing goes back to Oliver Heaviside and is quite convenient.

In general, the notation for vectors has a long history. The very first notation, by means of a bar ($\overline{a}$), was introduced in 1806 by Jean-Robert Argand (1768–1822), a self-taught Swiss mathematician and inventor of the geometrical interpretation of complex numbers. The notation $\overline{AB}$ was introduced by the German mathematician and astronomer August Ferdinand Möbius (1790–1868), the same one who proposed the simplest *one-sided* (or *unilateral*) *surface*, known as the *Möbius strip* (also the name of one of the stories, *el Anillo de Moebius*, by the famous XX century writer Julio Cortázar; 1914–1984). The Scottish physicist, mathematician, and astronomer James Clerk Maxwell (1831–1879), renowned for "his" equations for the electromagnetic field, as well as the distribution in molecular kinetic theory named after him, denoted vectors by Gothic letters (and Hamilton, by Greek letters). Outraged by Maxwell's choice, Heaviside remarked that *"This was an unfortunate choice, being by itself sufficient to prejudice readers against vectorial analysis"*, and suggested that vectors be denoted by boldface letters: **a**, **r**, **v** (1891; he used the "usual" letters a, r, v to denote the lengths of these vectors).

The derivative $\mathbf{X}' = \mathbf{X}'(t)$ of the left-hand side of equation (2) could be treated as the *derivative of the vector-valued function* $t \mapsto \mathbf{X}'(t)$ (see Sec. 1.1.6). The correspondence $\mathbf{X} \mapsto \mathbf{V}(\mathbf{X})$, which as if *attaches to each point* $X(x;y)$ *of the Oxy plane the corresponding vector* $\mathbf{V} = \mathbf{V}(\mathbf{X}) = \overline{XV}$, also has a simple geometric interpretation: this correspondence is a *vector field in the plane* (from the point of view of "pure" analysis, the above correspondence $\mathbf{X} \mapsto \mathbf{V}(\mathbf{X})$ is a *vector-valued function of two variables*: $(x;y) \mapsto \mathbf{V} = \mathbf{V}(x,y)$).[1]

The vector-valued function $\mathbf{X} = \mathbf{X}(t)$ corresponds to the *hodograph*, i.e., the trajectory described by the head $X_t = X(t) = (x(t);y(t))$ of the radius vector $\overline{OX}_t = \mathbf{X}(t)$. The derivative $\mathbf{X}'(t)$ is the velocity vector of an imaginary moving point X_t; in the case of $\mathbf{X}' \neq \mathbf{0}$,[2] it is tangent to the trajectory (hodograph), and its length is the absolute value (magnitude) of the velocity (or the scalar velocity of a point moving along the hodograph; see Sec. 1.1.7). Let a vector-valued function $\mathbf{X} = \mathbf{X}(t)$ be *a solution of the differential equation* (2) on some interval I, i.e.,

$$\forall t \in I \quad \mathbf{X}'(t) = \mathbf{V}(\mathbf{X}(t)) \tag{3}$$

[1] There are two other useful interpretations of a vector field, as *mappings of the plane to itself*: first, as a mapping in which every point $X(x;y)$ maps to the head $V(x+A;y+B)$ of the vector $\overline{XV} = \mathbf{V}(\mathbf{X})$ attached to X; second, as a mapping in which the point $X(x;y)$ maps to the point $Y(A;B)$, the head of the vector $\overline{OY} = \mathbf{V}(\mathbf{X})$ attached this time to the origin O.

[2] This is the notation for the *zero vector*, the vector $\mathbf{0}$ with coordinates $(0;0)$.

(in the notation of the system of scalar differential equations (1), this means that for $t \in I$ the equations

$$\begin{cases} x'(t) = A(x(t), y(t)), \\ y'(t) = B(x(t), y(t)) \end{cases} \tag{4}$$

are fulfilled, where $x(t)$ and $y(t)$ are the coordinate functions of the vector function $\mathbf{X}(t)$). According to the aforesaid, fulfilment of equality (3) means that *in the motion of the point X_t representing the state of the system, its velocity $\mathbf{X}'(t)$ at each time $t \in I$ is equal to the vector $\mathbf{V}(X_t) = \mathbf{V}(\mathbf{X}(t)) = \mathbf{V}(\overline{OX_t})$ specified at the point X_t.*

Thus, the vector differential equation (2) corresponds to a *vector field in the plane*, the *velocity field* of motion along the trajectories corresponding to the solutions of equation (2). We can think of a vector field as a velocity field of a steady flow of some (generally speaking, compressible) fluid in a plane, and trajectories as *"flow lines"*, i.e., curves along which "particles" of the fluid move. We will also call the trajectories *integral curves* or *phase trajectories*; the set of all phase trajectories (*with indicating the direction of motion* of the representing point X_t in time) is the *phase portrait* of the differential equation (2) or of the system of differential equations (1) (or of the corresponding vector field; or of a related real-world system whose model is equation (2) or the system of equations (1)).

The concept of a vector field (and the corresponding term) was introduced around 1830 by the famous self-taught English physicist Michael Faraday (1791–1867), a member of the Royal Society (and of the St. Petersburg Academy of Sciences), the discoverer of the law of electromagnetic induction and many other laws and phenomena; it was he who introduced the notions and terms *cathode, anode, ion, electrolysis, electrode....*

Faraday's major achievement was his development of the *concept of a field* (1852), which is the basis for the most important section of modern physics, *field theory.* He introduced the concepts of magnetic and electromagnetic fields, the idea of electric and magnetic *lines of force* (these are the same "flow lines"). Albert Einstein believed that the idea of the field was Faraday's most original idea, the most important discovery since Newton; he wrote: *"It needed great scientific imagination to realize that it is not the charges nor the particles but the field in the space between the charges and particles which is essential for the description of physical phenomena"*— for space embodies the electromagnetic field!

The notion of vector field is one of the basic concepts in *vector analysis*, the branch of mathematical analysis that studies various generalisations of differentiation and integration operations to "multidimensional objects" (vector functions, functions on vectors, vector fields, etc.; see comments in § 1.2).

4.3.2. Phase portraits and singular points: "nodes" and "saddles". Note that, as well as for ordinary autonomous differential equations of the form $x' = v(x)$, the points X at which the vector field $\mathbf{V} = \mathbf{V}(X)$ vanishes correspond to *stationary solutions*: if for some point $X_0(x_0; y_0)$ the equality $\mathbf{V}(X_0) = \mathbf{V}(x_0, y_0) = \mathbf{0}$ holds, then the *constant* vectors function $\mathbf{X}(t) \equiv \mathbf{X}_0 = \overline{OX}_0$ (for which the coordinate functions are $x(t) \equiv x_0 = \text{const}$ and $y(t) \equiv y_0 = \text{const}$) obviously satisfies the differential equation (2). Such points are called *singular points* of the differential equation (2) or of the corresponding vector field; they are found from the equations $A(x, y) = 0$, $B(x, y) = 0$. Singular points can be *different*, they are classified according to the behaviour of phase trajectories in their *neighbourhoods*; the neighbourhood of a point X_0 is defined to be any (open)[1] circle centred at X_0 (the ε-*neighbourhood* of a point X_0 is defined as the set of points $U_\varepsilon = \{X \mid X_0X < \varepsilon\}$). Before discussing the classification of singular points (recall that we discussed such points in the preceding section), let us consider some particular examples: construct phase portraits of several (simplest) differential equations of the form (2), i.e., systems of planar equations of the form (1).

Example 1. To construct the phase portrait of the system of differential equations

$$\begin{cases} x' = x, \\ y' = 3y, \end{cases}$$

we can simply *solve* this system, because its equations are "split": since the first equation involves only x and the second equation only y, the variables x and y vary *independently* of each other. The equations for $x(t)$ and $y(t)$ are linear and easy to solve:

$$x' = x \quad \Leftrightarrow \quad x = a\mathrm{e}^t \quad \text{with} \quad a = x(0);$$
$$y' = 3y \quad \Leftrightarrow \quad y = b\mathrm{e}^{3t}, \quad b = y(0).$$

Thus, we have found a vector function in coordinates which is the *general solution* of the original system

$$\mathbf{X}(t) = (a\mathrm{e}^t; b\mathrm{e}^{3t}),$$

where $(a; b) = (x_0; y_0)$ are the *initial conditions* at time $t = 0$. The form of the trajectory and its position in the phase plane are determined by the signs of the constants a and b. If $a \neq 0$, the variable y can easily be expressed through x:

$$x = a\mathrm{e}^t \quad \Rightarrow \quad \mathrm{e}^t = \frac{x}{a} \quad \Rightarrow \quad y = b\mathrm{e}^{3t} = b(\mathrm{e}^t)^3 = b \cdot \left(\frac{x}{a}\right)^3 = Cx^3, \quad C = \frac{b}{a^3};$$

in doing so, we have to take into account the *range of variation* of the variable $x = x(t) = a\mathrm{e}^t$, because it depends on the sign of a. Namely, if $a > 0$, then x varies over the positive semi-axis $x > 0$ so that when $t \to +\infty$, then

[1] That is, without boundary.

$x(t) \to +\infty$, and when $t \to -\infty$, then $x(t) \to 0+$; thus, we obtain families of right branches of the "cubic parabola" $y = Cx^3$ lying in quadrants I and IV ($C \neq 0$, $x > 0$; see Fig. 62), and also (for $C = 0$) the right half-line of the Ox axis. In this case, the representing point moves to the right (in the phase portrait, the direction of motion is indicated by "arrows").

If $a < 0$, then x varies over the negative semi-axis $x < 0$, when $t \to +\infty$, then $x(t) \to -\infty$, and when $t \to -\infty$, then $x(t) \to 0-$; again we obtain two families of parts of the graphs $y = Cx^3$, $x < 0$, lying in the left half-plane, together with the left half-line of the Ox axis; the representing point moves to the left.

Lastly, it remains to consider the case of $a = 0$. Then

$$x(t) \equiv 0, \qquad y(t) = b\mathrm{e}^{3t};$$

if $b = 0$, then also $y(t) \equiv 0$, and we obtain a stationary solution corresponding to the singular point $O(0;0)$. For $b \neq 0$ we obtain two more straight phase trajectories: the positive and negative rays of the Oy axis; the representing point, as for the other solutions, moves along them ("goes away") from the singular point to infinity. □

A *singular point* of a differential equation (or vector field) near which the behaviour of the trajectories is similar to that obtained in the above example is called an *unstable node* (in terms of a "fluid flow" such a singular point is called a *source*, because "flow lines" seem to "emerge" from this point). The *unstability* of a singular point of this type consists in the fact that if the initial conditions $(x(0); y(0))$ differ even "a little bit" from the *zero* initial conditions, the phase trajectory *"goes away"* from the singular point.

We have already encountered nodes when considering singular points of variable separable differential equations (§ 4.2, Examples 4 and 5). More-

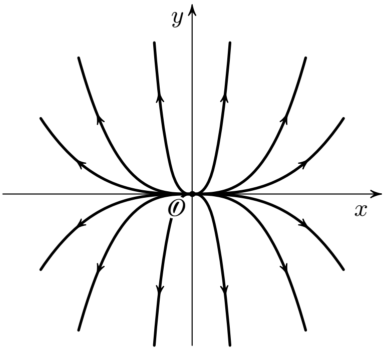


Fig. 62.

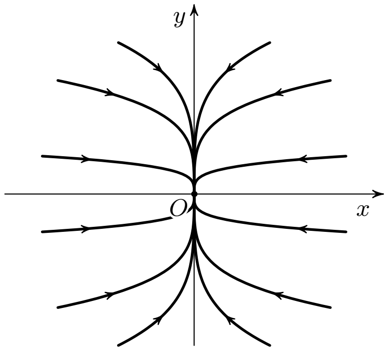


Fig. 63.

over, as one can easily see, the phase portrait of the system of equations just considered *geometrically* (ignoring the “arrows”) *coincides* with the phase portrait of the differential equation $y' = \frac{3y}{x}$ from Example 4 (§ 4.2; Fig. 60). Question: *is this accidental*? Let us consider some more examples.

Example 2. In the system of differential equations

$$\begin{cases} x' = -4x, \\ y' = -y \end{cases}$$

the variables x and y also vary independently of each other. Therefore, as in Example 1, we find

$$x' = -4x \quad \Leftrightarrow \quad x = a\mathrm{e}^{-4t}, \quad a = x(0);$$
$$y' = -y \quad \Leftrightarrow \quad y = b\mathrm{e}^{-t}, \quad b = y(0).$$

Hence, the general solution of the system in coordinates is written via exponential functions, though this time *decreasing*, so *for any initial conditions* $(x(0); y(0)) = (a; b)$ *the solution tends to the singular point* (“to the stationary solution”) $O(0;0)$:

$$\mathbf{X}(t) = (a\mathrm{e}^{-4t}; b\mathrm{e}^{-t}).$$

In this case it is simpler to express x through y: for $b \neq 0$ we have

$$\mathrm{e}^{-t} = \tfrac{y}{b} \quad \Rightarrow \quad x = a\mathrm{e}^{-4t} = a(\mathrm{e}^{-t})^4 = a \cdot \left(\tfrac{y}{b}\right)^4 = Cy^4, \quad C = \tfrac{a}{b^4}.$$

Though, we could do the opposite, i.e., express y through x:

$$x = a\mathrm{e}^{-4t} \quad \Rightarrow \quad \mathrm{e}^{-4t} = \frac{x}{a} \quad \Rightarrow \quad y = b\mathrm{e}^{-t} = b(\mathrm{e}^{-4t})^{1/4} = b \cdot \left|\frac{x}{a}\right|^{1/4} = D|x|^{1/4}.$$

Anyway, the result is four families of “biquadratic semi-parabolas” $x = Cy^4$, one in each of the four quadrants (see Fig. 63), and five more trajectories: positive and negative semi-axes of the coordinate axes and a stationary solution corresponding to the singular point $O(0;0)$. As follows from the formulae, for any non-zero initial conditions the *representing point* $X_t = X(t) = (x(t); y(t))$ *tends to the singular point*, so the singular point in this case is a *stable node*, and in terms of flows it is a “drain”, since the phase trajectories (“flow lines”) seem to “enter” this point. Note, however, that the representing point X_t for a non-zero initial position $X_0(a;b)$ only *tends* to the singular point (and very fast, *exponentially*!), but *never reaches it* (similarly to Newton’s differential equation, in which the phase point $(x;v)$ moving along the separatrix in the nondegenerate case never reaches the corresponding “stationary point”; see Ch. V below). A stable node can be compared to an attracting point (*attractor*) for discrete dynamical systems of iterative type. □

Note that this phase portrait (Fig. 63) has also occurred before: in § 4.2 for the differential equation $y' = \frac{y}{4x}$ (Example 5, Fig. 61). *Is this accidental*?

Example 3. The system

$$\begin{cases} x' = x, \\ y' = -y \end{cases}$$

can be solved similarly to the previous ones:

$$x' = x \quad \Leftrightarrow \quad x = a\mathrm{e}^t, \quad a = x(0);$$
$$y' = -y \quad \Leftrightarrow \quad y = b\mathrm{e}^{-t}, \quad b = y(0).$$

The general solution in coordinates is written via exponential functions, *increasing* for x and *decreasing* for y:

$$\mathbf{X}(t) = (a\mathrm{e}^t; b\mathrm{e}^{-t}).$$

Thus, the variables $x = x(t)$ and $y = y(t)$ are related for $a, b \neq 0$ by the *inverse proportionality* relation:

$$x = a\mathrm{e}^t \quad \Rightarrow \quad \mathrm{e}^t = \tfrac{x}{a} \quad \Rightarrow \quad y = b\mathrm{e}^{-t} = b \cdot \tfrac{a}{x} = \tfrac{k}{x}, \quad k = ab \neq 0.$$

Hence, the phase portrait of the system consists of four families of *branches of hyperbolas* $xy = k \neq 0$, one family in each of the four quadrants (Fig. 64). Furthermore, the origin O is a separate trajectory corresponding to the stationary solution. Finally, the semi-axes of the Ox axis are the phase trajectories along which the representing point $X(t) = (x(t); y(t)) = (a\mathrm{e}^t 0)$ is exponentially moving away from the singular point O, and on the semi-axes of the Oy axis, the representing point $X(t) = (x(t); y(t)) = (0; b\mathrm{e}^{-t})$, on the contrary, exponentially tends to the singular point. □

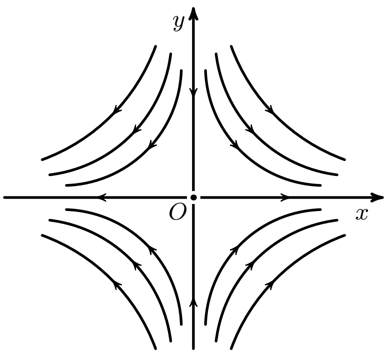


Fig. 64.

The resulting phase portrait has also occurred before, in Example 2 of § 4.2 for the differential equation $y' = -\frac{y}{x}x$ (see Fig. 59). A singular point of the differential equation (or vector field) near which the behaviour of trajectories is approximately the same as that obtained in the example above, i.e., is characterised by the presence (in some neighbourhood of the singular point) of two pairs of *separatrices* (along two of them, the point X_t tends *to* the singular point, and along the other two, it moves away *from* the singular point), is called, as above, a *saddle* (in terms of flows, such a singular point is called a *cross-point*).[1] Note that this special point should be considered *unstable*, because for any initial point $X_0(a; b)$ that does not lie on the Oy-axis (and is arbitrarily close to the point O), the phase trajectory *moves away* from the singular point (but this is *not* a repeller!).

In Examples 1–3 we succeeded because the variables x and y describing the state of the system varied *independently* of each other, i.e., the rate x' of change of the variable x depended only on the value of the variable itself: $x' = \varphi(x)$, and similarly for the variable y: $y' = \psi(y)$. Therefore, each of the equations of the resulting system

$$\begin{cases} x' = \varphi(x), \\ y' = \psi(y) \end{cases} \tag{5}$$

is integrated *separately*, and the solution $\mathbf{X}(t)$ of the system is composed from the obtained dependencies. It is clear that this "trick" will not work if the variables x and y are "tangled" on the right-hand sides of the equations of the system, as, for instance, in the following two examples.

4.3.3. Further examples: more "saddles" and "centers".

Example 4. For the beautiful system

$$\begin{cases} x' = y, \\ y' = x, \end{cases} \tag{6}$$

we will show as many as *three* ways of solution (i.e., finding its solutions).[2]

Solution I (*ingenious*). By adding and subtracting the equations of the system, according to the differentiation rule for the sum and difference of functions, we obtain the equations $(x + y)' = x + y$ and $(x - y)' = y - x$. Now introduce new functions (of time t): $u = x + y$ and $v = x - y$; then

[1] Of the flow lines.

[2] The word *solution*, as in algebraic equations and systems of equations, has *two meanings*: "solutions" that constitute the answer in a problem, and "solution" as a *process* of finding solutions (solving).

instead of the system of equations (6) we obtain a system of the form (5):

$$\begin{cases} u' = u, \\ v' = -v. \end{cases}$$

This system is integrated immediately:

$$u' = u \quad \Leftrightarrow \quad u = a\mathrm{e}^t, \quad v' = -v \quad \Leftrightarrow \quad v = b\mathrm{e}^{-t}.$$

Reversely, the desired functions $x(t)$ and $y(t)$ are expressed through u and v:

$$x = \tfrac{1}{2}(u+v) = \tfrac{1}{2}(a\mathrm{e}^t + b\mathrm{e}^{-t}), \qquad y = \tfrac{1}{2}(u-v) = \tfrac{1}{2}(a\mathrm{e}^t - b\mathrm{e}^{-t}).$$

Note that we will again encounter these functions in Ch. V when considering Newton's equation in the case of a linear repulsive force, $x'' = \lambda^2 x$ with $\lambda = 1$. Holding off with the phase portrait for the moment, let us give another "tempting" approach to the original system.

Solution II (*more ingenious*). From the first equation of the system $y = x'$, by differentiating this equality (which is legitimate, since the function $y = y(t)$, being the coordinate function of the solution of the system, must be differentiable) and taking into account the second equation of the system, we obtain $y' = x'' = x$. As a result, we come to the *differential equation of hyperbolic cosine and sine*: according to the exercises in Ch. III, we obtain

$$x'' = x \quad \Leftrightarrow \quad x = \alpha \cosh t + \beta \sinh t; \quad \alpha = x(0), \quad \beta = x'(0).$$

Moreover, the differential equation $x'' = x$ can be interpreted as Newton's differential equation for a linear repulsive force with "repulsion coefficient" $\lambda = 1$, which, after introducing the velocity variable $v = x'$, leads to a *system of Newton's differential equations*[1]

$$\begin{cases} x' = v, \\ v' = x, \end{cases}$$

which, "up to the notation", *coincides with the original system* (6). Hence, the solutions $\mathbf{X}(t) = (x(t); y(t))$ obey a *conservation law*: for any solution, we have

$$y^2(t) - x^2(t) = C = \text{const}. \tag{7}$$

Without repeating the reasoning and calculations from the comments in the exercises to Ch. III, we conclude that the phase portrait of the system of differential equations (6) consists of equilateral hyperbolas $y^2 - x^2 = C$, $C \neq 0$,[2] whose pair of asymptotes $y = \pm x$ splits into two pairs of "incoming"

[1] See § 1.2 or the beginning of § 5.1.

[2] *Equilateral* hyperbolas are those having the *canonical equation* $\left(\pm \frac{x^2}{a^2} \mp \frac{y^2}{b^2} = 1\right)$ with equal parameters a and b. (An "equilateral ellipse" is a circle; however, an ellipse should more appropriately be called "equiaxial".)

and "outgoing" separatrices and a separate singular point of the *saddle* type (Fig. 65). This example can also explain the origin of the term *saddle*: in this case the phase trajectories lie on the level lines of the function $y^2 - x^2$, whose *graph in the coordinate space* $Oxyz$, i.e., the surface $z = y^2 - x^2$, is "saddle-shaped": when $y = 0$, we obtain a $\cup$-shaped parabola (opening to the top along the Ox axis), and when $x = 0$, we obtain a $\cap$-shaped parabola (opening to the bottom along the Oy axis).

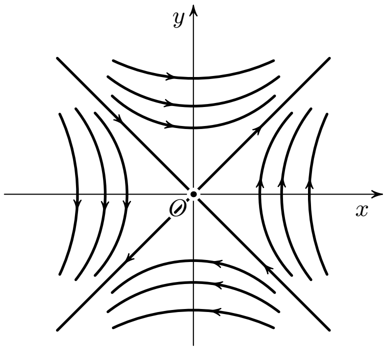


Fig. 65.

Solution III (*even more ingenious*). By comparing the conservation law (7) with the equations of the original system (6), it is not hard to guess this "trick": after multiplying the equations of system (6) "crosswise", we obtain

$$x'x = y'y \quad \Leftrightarrow \quad x'x - y'y = 0 \quad \Leftrightarrow \quad \tfrac{1}{2}(x^2 - y^2)' = 0$$
$$\Leftrightarrow \quad x^2 - y^2 = D = \text{const},$$

i.e., the same conservation law. Using this law we can find by integration the dependence of x and y on time t. □

The next example differs "only" in the sign in one of the equations of the system from Example 4.

Example 5. For the system

$$\begin{cases} x' = y, \\ y' = -x, \end{cases} \tag{8}$$

the solution method I used above does not work (try it!), but the second method, i.e., passing to the differential equation $x'' = y' = -x$, which coincides with the harmonic oscillation equation from § 1.2 with $\omega = 1$, immediately gives a general form of solutions. Namely, according to what was said in Sec. 1.2.5, we obtain

$$x'' = -x \quad \Leftrightarrow \quad x(t) = a\cos t + b\sin t; \quad a = x(0), \quad b = x'(0),$$

or equivalently $x(t) = A\cos(t+\varphi)$, and correspondingly

$$y(t) = x'(t) = b\cos t - a\sin t = -A\sin(t+\varphi).$$

Hence, according to the definition of cosine and sine, the representing point $X_t(x(t); y(t))$ is moving along a circle of radius A with angular velocity 1 in the negative direction (*"clockwise"*). Thus, the phase portrait of this system of differential equations is a family of concentric circles centred at a singular point, the origin $O(0;0)$; under any non-zero initial conditions $X_0 \neq O$, the representing point X_t rotates about the singular point O with unit angular velocity $\omega = -1$, i.e., clockwise (Fig. 66).

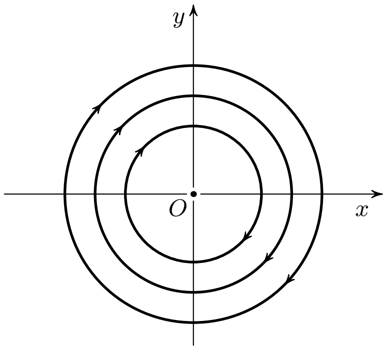


Fig. 66.

It is clear that the third solution method for the system of equations (6) is also suitable for system (8): after cross-multiplying the equations of system (8), we obtain that

$$-x'x = y'y \Leftrightarrow x'x + y'y = 0 \Leftrightarrow \tfrac{1}{2}(x^2+y^2)' = 0 \Leftrightarrow x^2 + y^2 = D = \text{const}; \quad (9)$$

i.e., system (8) obeys its own conservation law. It follows that all phase trajectories lie on circles given by equation (9); using this equation, we can find the dependence of x and y on time t. □

Singular points of differential equations the phase trajectories in the (small enough) neighbourhoods of which are closed lines with periodic motion along them, are called *centers*, positive or negative, depending on the direction of motion along these trajectories. The singular point in Example 5 is a *negative centre*; stable equilibria for one-dimensional dynamical systems corresponding to minima of potential energy (see Sec. 1.2.4 and § 5.1) are also negative centres. Centres are *stable* singular points, but in a slightly different sense — they are *Lyapunov stable*,[1] unlike *stable nodes*, i.e., *attractors*....

Here the trajectories do not tend to the singular point but simply cannot go far from it.

Note that the last solution method for the system of differential equations (8) can easily be generalised to *all* systems of the form

$$\begin{cases} x' = \alpha y, \\ y' = \beta x \end{cases} \tag{10}$$

($\alpha, \beta \neq 0$); it follows that such systems have "conservation laws" (*first integrals*!) $\beta x^2 - \alpha y^2 = C = \text{const}$ and, accordingly, singular points of the type *saddle* or *centre*. But "tangling the variables" is not limited to the type just discussed — in the system of differential equations (1), the right-hand sides $A(x,y)$ and $B(x,y)$ can, of course, depend on both x and y *simultaneously*. We will consider such systems by examples of Newton's differential equations in Ch. V.

§ 4.4. Integrable systems. "Predator–Prey" Biocoenosis Model

4.4.1. Leibniz's formalism for planar systems. Now let us consider a special class of planar systems of differential equations *solvable in quadratures*. The starting point will be Examples 1–3 given in § 4.3, in which we observed not even analogy but *"geometric" identity*, *coincidence* of integral curves of planar systems and first-order differential equations of the form $y'(x) = F(x,y)$:

$$\begin{aligned} x'(t) &= x, & y'(t) &= 3y & &\rightleftarrows & y'(x) &= \frac{3y}{x}; \\ x'(t) &= -4x, & y'(t) &= -y & &\rightleftarrows & y'(x) &= \frac{y}{4x}; \\ x'(t) &= x, & y'(t) &= -y & &\rightleftarrows & y'(x) &= -\frac{y}{x}. \end{aligned}$$

Since in this situation the same letter y denotes quite different functions, namely $y = y(t)$ and $y = y(x)$, and the same letter x denotes both the function $x = x(t)$ and the "independent variable" x (*argument* of the func-

[1]Note that *Lyapunov stability* is a *mathematical term*! Alexander Mikhailovich Lyapunov (1857–1918) was a prominent Russian mathematician, a worthy competitor of the famous Henri Poincaré (1854–1912) in questions of stability; Poincaré, among other things, is famous for having discovered the *special relativity* at the same time as Einstein in 1905.

tion $y=y(x)$), it is quite reasonable to use Leibniz's notation:

$$\begin{aligned} &\frac{dx}{dt}=x, && \frac{dy}{dt}=3y && \rightleftarrows && \frac{dy}{dx}=\frac{3y}{x};\\ &\frac{dx}{dt}=-4x, && \frac{dy}{dt}=-y && \rightleftarrows && \frac{dy}{dx}=\frac{y}{4x};\\ &\frac{dx}{dt}=x, && \frac{dy}{dt}=-y && \rightleftarrows && \frac{dy}{dx}=-\frac{y}{x}. \end{aligned}$$

One can easily notice that the fractions on the right-hand sides of the right-hand equations are fractions obtained by dividing the right-hand sides of the equations of the corresponding systems. Let us try to apply this "trick" to the systems from Examples 4–5 of § 4.3. We obtain

$$\begin{cases} \dfrac{dx}{dt}=y,\\ \dfrac{dy}{dt}=x \end{cases} \rightleftarrows \quad \frac{dy}{dx}=\frac{x}{y} \quad \Leftrightarrow \quad y\,dy=x\,dx \quad \Leftrightarrow \quad y^2=x^2+\text{const.}$$

Similarly,

$$\begin{cases} \dfrac{dx}{dt}=y,\\ \dfrac{dy}{dt}=-x \end{cases} \rightleftarrows \quad \frac{dy}{dx}=-\frac{x}{y} \quad \Leftrightarrow \quad y\,dy=-x\,dx \quad \Leftrightarrow \quad y^2=-x^2+\text{const.}$$

In both cases we have arrived at the right answer. If, *following Leibniz*, we *formally* operate with the equations of the systems, we obtain a very simple "proof" of the legitimacy of the demonstrated approach. Namely, if we consider the expressions dx, dy, and dt as "symbolic" numerators or denominators of the "pseudo-fractions" on the left-hand sides of the equations, then we pass from a planar *autonomous system* to a *single* but *not autonomous* first-order *equation*:

$$\begin{cases} \dfrac{dx}{dt}=A(x,y),\\ \dfrac{dy}{dt}=B(x,y) \end{cases} \rightarrowtail \quad \frac{dy}{dt}:\frac{x}{t}=\frac{y}{x}=\frac{B(x,y)}{A(x,y)} \quad \Leftrightarrow \quad y'(x)=F(x,y),$$

where $F(x,y)=\dfrac{B(x,y)}{A(x,y)}$ and where dt, so to say, has "cancelled out", and thereby the variable t has been "eliminated" from the system. Again, just as in § 4.2, a complete triumph of Leibniz's symbolism! However, it is time to think about the sense.

4.4.2. Solution theorem: justification of Leibniz's formalism.

Theorem 1 (solution theorem). *Let* $\mathbf{X}(t)=(x(t);y(t))$ *be a solution of the autonomous system of differential equations or of the vector equation*

$$\begin{cases} x'(t)=A(x,y),\\ y'(t)=B(x,y) \end{cases} \Leftrightarrow \quad \mathbf{X}'(t)=\mathbf{V}(\mathbf{X}) \tag{1}$$

satisfying the initial conditions $\big(x(t_0); y(t_0)\big) = (x_0; y_0)$, *i.e.,* $\mathbf{X}(t_0) = \mathbf{X}_0$. *Furthermore, let the function* $A(x, y)$ *be non-zero in some neighbourhood of the point* $X_0(x_0; y_0)$. *Then to each value of* $x = x(t)$ *in some neighbourhood* I *of* x_0 *there corresponds a* single *value of* $y = y(t)$,[1] *so the function* $x \mapsto y = y(x)$ *defined on* I *satisfies on this entire interval the non-autonomous differential equation*

$$y'(x) = F(x, y), \quad \textit{or} \quad \frac{dy}{dx} = F(x, y), \quad \textit{where} \quad F(x, y) = \frac{B(x, y)}{A(x, y)}. \tag{2}$$

Proof. Geometrically, the statement of the theorem is *obvious*: any phase trajectory at each of its points X_t (at the end of the radius vector $\overline{OX}_t = \mathbf{X}(t)$) is tangent to the (non-zero) vector $\mathbf{V}(\mathbf{X}) = (A(x, y); B(x, y))$, and hence to the line X_tV_t ($\overline{X_tV_t} = \mathbf{V}(X_t)$) defined by this vector and passing through the point X_t with slope $k = F(x, y) = \dfrac{B(x, y)}{A(x, y)}$ (see Fig. 67). According to the conditions of the theorem, in some time interval $(t_0 - \delta, t_0 + \delta)$ the trajectory is the graph of some function $y = \varphi(x)$,[2] whose derivative, due to the above remark, is equal to the slope of the tangent, i.e., to $k = F(x, y) = F(x, \varphi(x))$, as required. A formal proof is also easy to give.

Since in some neighbourhood U of the point X_0 the function $A(x, y)$ is non-zero, and for values of t in some time interval $J = (t_0 - \delta, t_0 + \delta)$ the point $X(t)$ remains in U, for these t the function $a(t) = A(x(t), y(t))$ takes values of the same sign. Since for $t \in J$ the derivative $x'(t)$ is equal to $a(t)$, the coordinate function $J \to \mathbb{R}\colon t \mapsto x(t)$ is *invertible*, and the inverse function $t = \tau(x)$ is well-defined on some interval $I = (x_0 - \varepsilon, x_0 + \varepsilon)$. Let us define the function $I \to \mathbb{R}\colon x \mapsto y = \varphi(x)$ as $\varphi(x) = y(\tau(x))$; then for any value of $t \in J$ for which $x = x(t) \in I$, the value of $y = \varphi(x)$ coincides with $y(t)$ (since $\tau(x(t)) \equiv t$; Fig. 68).

Now let us find the derivative of $\varphi(x) = y(\tau(x))$. By the chain rule, we have

$$\varphi'(x) = (y(\tau(x)))' = y'(\tau(x)) \cdot \tau'(x),$$

and according to the equations in (1) and by the inverse function differentiation rule,

$$y'(t) = B(x(t), y(t)) \;\Rightarrow\; y'(\tau(x)) = B(x(\tau(x)), y(\tau(x))) = B(x, \varphi(x)),$$
$$x'(t) = A(x(t), y(t)) \;\Rightarrow\; \tau'(x) = \frac{1}{x'(t)}\bigg|_{t=\tau(x)} = \frac{1}{A(x(\tau(x)), y(\tau(x)))} = \frac{1}{A(x, \varphi(x))}.$$

[1] The value of t is *the same* as for the variable $x = x(t)$!

[2] To avoid confusion, we will denote y as a function of x by $y = \varphi(x)$.

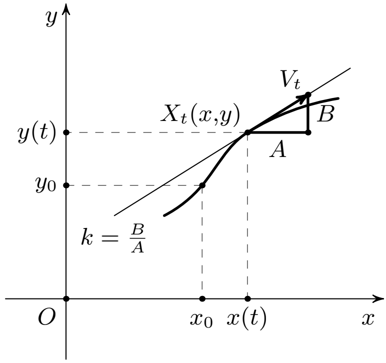


Fig. 67.

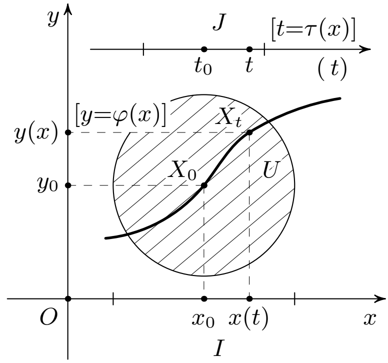


Fig. 68.

Hence, for any $x\in I$ the function $\varphi(x)$ satisfies

$$\varphi'(x)=y'(\tau(x))\cdot\tau'(x)=B(x,\varphi(x))\cdot\frac{1}{A(x,\varphi(x))}=\frac{B(x,\varphi(x))}{A(x,\varphi(x))}=F(x,\varphi(x)),$$

as required. □

4.4.3. Leibniz's solution scheme for planar systems. Thus, we have proved that, roughly speaking, the *integral curves* (phase trajectories) of the *system of differential equations* (1) *coincide* ("geometrically", i.e., as sets of points) *with the integral curves of the differential equation* (2). The question is how to find a solution $\mathbf{X}(t)=(x(t);y(t))$ of system (1) if a solution of the differential equation (2) has already been found? In the case at hand, the following *solution scheme for system* (1) applies. If a solution $y=\varphi(x)$ of the differential equation (2) is found, i.e., in the interval I mentioned in the theorem the relations

$$\frac{dy}{dx}=F(x,y)=\frac{B(x,y)}{A(x,y)}\quad\text{or}\quad y'(x)=\varphi'(x)=F(x,\varphi(x))=\frac{B(x,\varphi(x))}{A(x,\varphi(x))}$$

are satisfied, then, first of all, the dependence $x(t)$ is found. Note that for the desired solution $\mathbf{X}(t)=(x(t);y(t))$ the coordinate functions are related by $y=\varphi(x)$, which, in general, involves an *arbitrary constant*, determined from the initial condition $\varphi(x_0)=y_0$. We substitute this dependence into the first equation of system (1) and solve the resulting differential equation:

$$\frac{dx}{dt}=A(x,\varphi(x))\overset{\text{des}}{=}V(x)\quad\Leftrightarrow\quad\frac{dx}{V(x)}=dt\quad\Leftrightarrow\quad t-t_0=\int\limits_{x_0}^{x}\frac{dz}{V(z)}.$$

We arrive at a well-known expression, the *time integral* as in Sec. 4.1.7, using which, *in principle*, the dependence $t=t(x)$ can be found (if we write this dependence using indefinite integral, it will involve one more arbitrary constant, also determined by the initial condition $x(t_0) = x_0$). From the equation $t = t(x)$, again *in principle*, the inverse dependence $x = x(t)$ can be found, for which the first equation of the system is satisfied: $x'(t) = V(x) = V(x) = A(x, \varphi(x))$. Having specified the second coordinate function $y = y(t)$ as $y(t) = \varphi(x(t))$, we check that it satisfies the second equation of the system:

$$y'(t) = (\varphi(x(t)))' = \varphi'(x(t)) \cdot x'(t) = F(x, \varphi(x)) \cdot A(x, \varphi(x)) = \frac{B(x, \varphi(x))}{A(x, \varphi(x))}.$$

Hence, if the differential equation (2) has been integrated, then finding solutions of system (1) reduces to quadratures (and solving an "algebraic" equation). We may say that, *in a sense*, system (1) and the differential equation (2) are *equivalent*.

We can see how fruitful for mathematical analysis Leibniz's ideas and inventions proved to be in the search for a *"universal symbolic language"*, or, as Leibniz himself put it, *"characteristica universalis"* (*universal characteristic*, or *universal character*).[1] The idea of creating a "common character" dominated Leibniz's thinking and diverse activities from the age of twenty. At the age of 25, while on a diplomatic mission to Paris, he delved (under the influence of Huygens) straight into the *most modern* mathematics (1672), and soon (1675) surpassed his predecessors (but not Newton). Perhaps only mathematical analysis, and partly logic, were the areas in which Leibniz was able to effectively fulfil his dream of the *"universal character"*.

A *universal genius*, in keeping with his intended "character", Gottfried Wilhelm Leibniz was born on 1 July 1646 into a professorial family. At the age of 15 he left school and entered the university, where he studied law. He had a phenomenal memory and exceptional aptitude for languages (some biographers attribute this to Leibniz's "Slavic roots", believing his surname to be a Germanised West Slavic surname Lubenets!). Like Poincaré later, but to an even greater extent, Leibniz was distinguished by his flexibility, his receptivity to everything, his ability to generalise, to synthesise the most diverse ideas, methods, etc. He was an outstanding *mathematician, physicist, historian, diplomat, jurist, theologian* (he was obsessed with the idea of uniting the Catholic and Protestant churches), and, perhaps, *primarily, philosopher*.

Leibniz's philosophical system is based on *"monadology"* and the principle of *"pre-established harmony"*. *Monads*[2] are like atoms of everything — inanimate matter, energy, soul, and spirit; the idea of monads goes back to the Pythagoreans and Plato. The "pre-established" harmony rules the monads, i.e., the *world*, both ideal and phenomenological. It is interesting that the Russian mathematician, professor of the Moscow University Nikolai Vasilievich Bugaev (1880–1934), father of the poet and partly philosopher Andrei Bely (Boris Nikolayevich Bugaev; 1837–1903), was Leibniz's successor in the matter of monads: according to his *"evolutionary monadology"*, the past does not disappear, but accumulates; monads and the whole world become more and more perfect!

Leibniz died on 14 November 1716 in Hanover. The tombstone in the church where he is buried bears only two words: *ossa Leibnitii* (Leibniz's ashes).

---

[1] The term *characteristic* is derived from Greek χαρακτηρ [kharakter], 'stamp', 'mark', 'distinctive quality' in the sense of a feature reflecting the "shape and essence of a thing".

[2] Do not confuse with *maenads*, i.e., *Bacchantes*!

4.4.4. Separation of variables in planar systems. All first-order differential equations of the form $y'(x)=F(x,y(x))$ that we know how to solve belong to or are reducible to *variable separable equations* (see § 4.2). If on the right-hand side of the equations of system (1) the *variables are separable*, i.e., $A(x,y)=P(x)\cdot Q(y)$ and $B(x,y)=R(x)\cdot S(y)$, they are also separable in the "equivalent" equation (2):

$$\begin{cases} \frac{dx}{dt}=P(x)Q(y), \\ \frac{dy}{dt}=R(x)S(y) \end{cases} \rightleftarrows \quad \frac{dy}{dx}=\frac{R(x)S(y)}{P(x)Q(y)}$$

$$\Leftrightarrow \frac{Q(y)}{S(y)}dy=\frac{R(x)}{P(x)}dx \quad \Leftrightarrow \quad \mathcal{F}(y)=\mathcal{G}(x)+C,$$

where C is an arbitrary constant, and the functions $\mathcal{F}(y)$ and $\mathcal{G}(x)$ are antiderivatives of the fractions $f(y)=\frac{Q(y)}{S(y)}$ and $g(x)=\frac{R(x)}{P(x)}$, respectively. One can easily realise that this method can be generalised to systems of the form

$$\begin{cases} x'(t)=P(x)Q(y)k(x,y), \\ y'(t)=R(x)S(y)k(x,y) \end{cases} \rightleftarrows \quad \frac{dy}{dx}=\frac{R(x)S(y)}{P(x)Q(y)},$$

etc. Systems whose right-hand sides can be represented in the above form (note that the factor $k(x,y)$ is *the same* for both equations of the system) are called *variable separable systems.* By the way, among them are also systems of Newton's differential equations for autonomous conservative one-dimensional dynamical systems, which we will deal with in the next chapter.

Let us apply the variable separation method to the study of a very interesting mathematical model related to biology, the *Lotka–Volterra* "predator–prey" *biocoenosis model* discussed in the introductory § 1.1 (Sec. 1.1.8; read it again!).

4.4.5. Qualitative analysis of the Lotka–Volterra model. Recall that this model is reduced to the system of two equations,

$$\begin{cases} x'(t)=ax-bxy, \\ y'(t)=cxy-dy, \end{cases} \Leftrightarrow \quad \begin{cases} x'(t)=x(a-by), \\ y'(t)=y(cx-d), \end{cases} \tag{3}$$

considered in quadrant I ($x>0$, $y>0$) of the Oxy plane, i.e., the "prey–predator" ("carp–pike"!) plane. The vector field of system (3),

$$\mathbf{V}(x,y)=(A(x,y),B(x,y))=(x(a-by),y(cx-d)), \tag{4}$$

is essentially nonlinear, and in quadrant I it has a single singular point $Q(x_0,y_0)$ with $x_0=\frac{d}{c}$ and $y_0=\frac{a}{b}$. The corresponding stationary solution $(x(t);y(t))\equiv(x_0;y_0)$ describes "harmonic" coexistence of population, or "balance", where the gain always equals the loss. Of course, we are interested in other solutions, i.e., *possible dynamics* of the biocoenosis.

First, let us consider a linear approximation of system (3) in the neighbourhood of the singular point, written in "shifted" coordinates $z = x - x_0$, $w = y - y_0$. Clearly, system (3) takes in these coordinates the form

$$\begin{cases} z'(t) = -b(x_0 + z)w, \\ w'(t) = c(y_0 + w)z, \end{cases}$$

and its linear approximation is

$$\begin{cases} z'(t) = -bx_0 w = -\beta w, \\ w'(t) = cy_0 z = \gamma z. \end{cases}$$

It is easily seen that the singular point $(z; w) = (0; 0)$ of the last system is a *centre*, and from this information we cannot judge the behaviour of the trajectories of the original nonlinear system (3).

Some information about phase trajectories can be obtained by "weighing up" how the vector field (4) behaves, i.e., by considering the signs of the components A and B of the vector field $\mathbf{V} = (A; B)$ at various points of the positive quadrant $x > 0$, $y > 0$. Formulae (4) show that in this quadrant we have $\operatorname{sgn} A(x, y) = -\operatorname{sgn}(y - y_0)$, $\operatorname{sgn} B(x, y) = \operatorname{sgn}(x - x_0)$, so the signs of the components of vectors $\mathbf{V}$ and their directions are as shown in Fig. 69.

For example, if $x > x_0$ and $y > y_0$, then the signs of A and B are $(-; +)$, and accordingly, the vector $\mathbf{V}$ is directed up-left, and so on. Note that on the line $x = x_0$, the coordinate $B(x, y)$ of $\mathbf{V}$ is zero and the vector is directed horizontally: to the right if $y < y_0$ and to the left if $y > y_0$. Similarly, on the line $y = y_0$, the coordinate $A(x, y)$ is zero and the vector $\mathbf{V}$ is directed vertically (see Fig. 69).

However, even from this figure it is *not clear* how the trajectories behave: do they wind around the singular point $Q(x_0; y_0)$ as trajectory (A) in Fig. 70, do they spin around Q as trajectory (B), or do they go away from Q as trajectory (C)? *In principle*, any of these is possible. Then, as it sometimes happens, a *"fortunate chance"* helps: the variables x and y in system (3) are *separable*, and the system is integrable!

4.4.6. Integration of the Lotka–Volterra system. Applying the Leibniz formalism to system (3), we obtain a differential equation for the function $y = y(x)$:

$$\frac{dy}{dx} = \frac{y(cx - d)}{x(a - by)} \quad \Leftrightarrow \quad \frac{a - by}{y} dy = \frac{cx - d}{x} dx \quad \Leftrightarrow \quad \left(c - \frac{d}{x}\right) dx = \left(\frac{a}{y} - b\right) dy.$$

Integration gives

$$cx - d \ln x = a \ln y - by + C \quad \Leftrightarrow \quad f(x) = -g(y) + C,$$

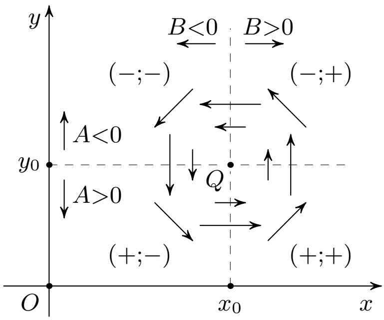


Fig. 69.

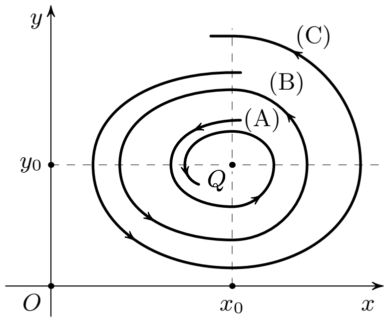


Fig. 70.

where C is a constant and where f and g are the following functions (similar to each other):

$$f(x) = cx - d\ln x, \qquad g(y) = by - a\ln y.$$

Thus, the Lotka–Volterra system of equations in quadrant I has a *first integral*, a "conservation law" of the form

$$f(x) + g(y) = C = \text{const} \quad (x > 0,\ y > 0). \tag{5}$$

Unfortunately, expressing y through x from equation (5) is, to put it mildly, problematic. Let us nevertheless try to obtain from this integral some information on the behaviour of the trajectories of system (3).

Let us explore the function $f(x) = cx - d\ln x$. Its first and second derivatives are

$$f'(x) = c - \frac{d}{x} = \frac{c(x - x_0)}{x}, \qquad f''(x) = \Big(c - \frac{d}{x}\Big)' = \frac{d}{x^2},$$

so on the semi-axis $x > 0$ we have the equality $\operatorname{sgn} f'(x) = \operatorname{sgn}(x - x_0)$; hence, the function f decreases for $x \in (0, x_0)$ and increases for $x \in (x_0, +\infty)$. Thus, x_0 is a strict minimum point; moreover, it is *nondegenerate*, i.e.,

$$f''(x_0) = \frac{d}{x_0^2} \neq 0.$$

The value at the minimum point is

$$f_0 = f(x_0) = c \cdot \frac{d}{c} - d\ln\frac{d}{c} = d\Big(1 - \ln\frac{d}{c}\Big),$$

and the graph of $f = f(x)$ is $\cup$-convex on the semi-axis $(0, +\infty)$ (Fig. 71; recall that we are interested only in positive values of x).

The function $g(y) = by - a\ln y$ behaves similarly: for it $y_0 = \frac{a}{b}$ is a nondegenerate minimum point, and the value at this point is

$$g_0 = g(y_0) = a\big(1 - \ln\frac{d}{c}\big).$$

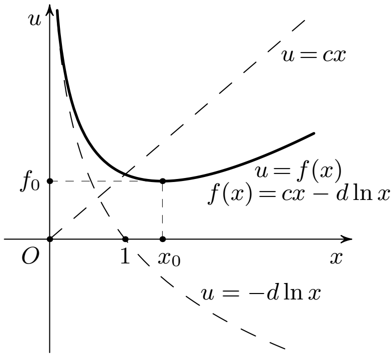


Fig. 71.

Since the function $H(x, y) = f(x) + g(y)$ is a first integral of the Lotka–Volterra system, it follows that *each phase trajectory of the Lotka–Volterra system lies on the level curve of the function* $H(x, y)$, i.e., on the line

$$\Gamma_h = \{(x; y) \mid H(x, y) = f(x) + g(y) = h = \text{const}\}.$$

Clearly, the function $H(x, y)$ is continuously differentiable (arbitrarily many times) in quadrant I, and at the point $Q(x_0; y_0)$ it has a *strict minimum*; i.e., for any $x, y > 0$ whenever the point $(x; y)$ is different from Q, the inequality $H(x, y) > H(x_0, y_0)$ holds. The graph of this function in the $Oxy\zeta$ coordinate space is a "cup-shaped" surface $\zeta = H(x, y)$, and the projections onto the Oxy plane of its sections by "horizontal" planes $\zeta = h$ lying above the minimum point $(x_0; y_0; f_0 + g_0)$ (i.e., for $h > f_0 + g_0$), are *closed lines* (Fig. 72), which are the *level curves of the first integral* $H(x, y)$.

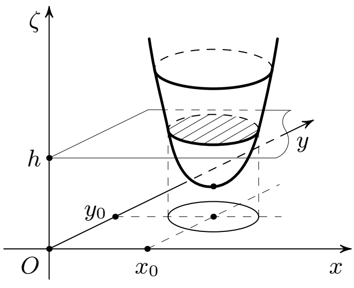


Fig. 72.

Note that the vector field $\mathbf{V}(x, y)$ of the Lotka–Volterra system vanishes (i.e., the vector $\mathbf{V}$ is *zero*) only at the point Q and is non-zero along any level curve Γ_h (fort $h > f_0 + g_0$), so the phase trajectories of the system do not merely lie on these curves but *coincide* with them, and the motion of the phase point $(x;y)=(x(t);y(t))$ along these curves is *periodic* (Fig. 73). Thus, in the Lotka–Volterra model, the "predator–prey" biocenosis evolves in such a way that both population numbers *periodically* increase to maximum values (this is the "upper right corner" of the phase trajectory) and decrease to minimum values (the "lower left corner" of the trajectory), as shown in the plots of the functions $n=x(t)$ and $n=y(t)$ (Fig. 74).

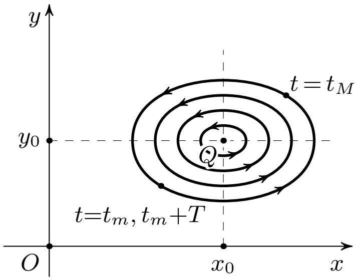


Fig. 73.

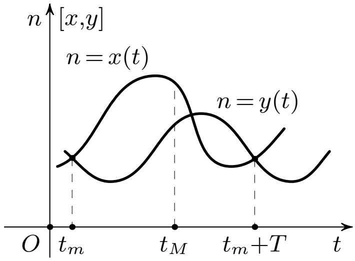


Fig. 74.

This character of evolution can be explained at the "qualitative level" as follows. Let the initial populations of predators and prey be small (this is the "lower left corner" of the phase trajectory). Then a small number of predators allows the x-population to grow, while the y-population only keeps stable so far. When the carp become abundant, the positive summand in the rate of change $y'(t)$ of the pike population, i.e., cxy, becomes dominant — the y-population grows, "stabilising" the x-population (the phase point $(x;y)$ moves from the "lower right corner" to the "upper right corner"). But when the pike population becomes abundant and the carp population does not grow, vice versa, the negative summand in the rate of change $y'(t)$ of the carp population, i.e., $-bxy$, becomes critical in the dynamics of the carp — the pike population is kept more or less stable by eating the carp, whose population is gradually "depleted" (the phase point $(x;y)$ moves from the "upper right corner" to the "upper left corner"). This exhausts the pike population (food is scarce!), and everything returns *"on its circuits"*, to the initial situation.

In this model, it is interesting to find out what changes in the dynamics of the "struggle for existence" are caused by particular external factors, such as periodic feeding of carp or catching of pikes (or carp), etc. However, consideration of further modifications of the Lotka–Volterra model would take us quite far away from our course and require a more thorough delving into the theory.

Exercises, Problems, and Tasks to Chapter IV

1. Find general solutions of the differential equations (including *singular* or *stationary* solutions) and draw in the Oxy plane the family of graphs of these solutions:

(1) $y' = -y^2$; (3) $y' = -y^3$; (5) $y' = -y^4$.
(2) $y' = y^3$; (4) $y' = y^4$;

2. Do the same for the following differential equations:

(1) $y' = \frac{1}{y}$; (3) $y' = \frac{1}{y^2}$; (5) $y' = \frac{1}{y^3}$.
(2) $y' = -\frac{1}{y}$; (4) $y' = -\frac{1}{y^2}$;

3. Do the same for the following differential equations:

(1) $y' = \sqrt{y} \quad (y > 0)$; (3) $y' = \sqrt[3]{y}$; (5) $y' = \frac{1}{2}|y|$.
(2) $y' = \dfrac{1}{\sqrt{y}} \quad (y > 0)$; (4) $y' = |y|$;

4. Do the same for the differential equations (and specify which of the stationary solutions [if any] are stable and which ones are unstable):

(1) $y' = y^2 - 1$; (3) $y' = 1 + y^2$; (5) $y' = y^2 - 3y + 2$.
(2) $y' = 1 - y^2$; (4) $y' = \sqrt{1 - y^2}$;

5. Do the same (as in the preceding exercise) for the following differential equations:

(1) $y' = \cos^2 y$; (3) $y' = \tan y$; (5) $y' = \mathrm{e}^y$.
(2) $y' = \sin^2 y$; (4) $y' = -\cot y$;

6. Find all stationary solutions of the differential equations, and each of these solutions indicate whether it is stable, unstable ("repelling"), or neither:

(1) $y' = y^3 - y$; (3) $y' = y^4 - 1$; (5) $y' = -\sin \pi y$.
(2) $y' = 4y - y^3$; (4) $y' = \sin y$;

7. Let the general solution of the differential equation $y' = g(y)$ be written as $y = \varphi(x, C)$ (C being an arbitrary constant). How to write the general solutions of the following equations using the function (expression) φ:

(1) $y' = -g(y)$; (3) $y' = 2g(y)$; (5) $y' = -3g(2y)$?
(2) $y' = g(-y)$; (4) $y' = g(3y)$;

8. Using the Leibniz formalism, find the general solutions and draw phase portraits for the following differential equations:

(1) $y' = \frac{y}{x}$; (3) $y' = -\frac{x}{y}$; (5) $y' = \frac{x}{2y}$;
(2) $y' = \frac{x}{y}$; (4) $y' = \frac{2y}{x}$; (6) $y' = -\frac{2y}{x}$;

(7) $y' = -\frac{x}{2y}$; (8) $y' = \frac{y}{2x}$; (9) $y' = \frac{2x}{y}$; (10) $y' = -\frac{y}{2x}$; (11) $y' = -\frac{2x}{y}$.

9. Using the Leibniz formalism, find the general solutions (and, when possible, draw phase portraits) of the following differential equations:

(1) $y' = xy$; (2) $y' = -xy$; (3) $y' = 2xy$; (4) $y' = -2xy$; (5) $y' = \frac{1}{2}xy$; (6) $y' = -\frac{1}{2}xy$.

10. Do the same for the following differential equations:

(1) $y' = x^2y$; (2) $y' = xy^2$; (3) $y' = -x^2y$; (4) $y' = -xy^2$; (5) $y' = x^2y^2$; (6) $y' = -x^2y^2$; (7) $y' = x^3y$; (8) $y' = -x^3y$; (9) $y' = xy^3$; (10) $y' = -xy^3$; (11) $y' = x^2y^3$; (12) $y' = x^3y^2$; (13) $y' = -x^2y^3$; (14) $y' = -x^3y^2$.

11. Do the same for the following differential equations:

(1) $y' = \frac{y}{2\sqrt{x}}$ $(x > 0)$;

(2) $y' = -\frac{y}{2\sqrt{x}}$ $(x > 0)$;

(3) $y' = \frac{2y}{\sqrt{x}}$ $(x > 0)$;

(4) $y' = -\frac{2y}{\sqrt{x}}$ $(x > 0)$;

(5) $y' = \frac{x}{2\sqrt{y}}$ $(y > 0)$;

(6) $y' = -\frac{x}{2\sqrt{y}}$ $(y > 0)$;

(7) $y' = \frac{2x}{\sqrt{y}}$ $(y > 0)$;

(8) $y' = -\frac{2x}{\sqrt{y}}$ $(y > 0)$;

(9) $y' = \frac{\sqrt{x}}{\sqrt{y}}$ $(x > 0,\ y > 0)$;

(10) $y' = -\frac{\sqrt{x}}{\sqrt{y}}$ $(x > 0,\ y > 0)$;

(11) $y' = \frac{\sqrt{y}}{\sqrt{x}}$ $(x > 0,\ y > 0)$;

(12) $y' = -\frac{\sqrt{y}}{\sqrt{x}}$ $(x > 0,\ y > 0)$.

12. Do the same for the following differential equations:

(1) $y' = y\sqrt{x}$ $(x > 0)$;

(2) $y' = -y\sqrt{x}$ $(x > 0)$;

(3) $y' = x\sqrt{y}$ $(y > 0)$;

(4) $y' = -x\sqrt{y}$ $(y > 0)$;

(5) $y' = -x\sqrt{y}$ $(y > 0)$;

(6) $y' = \sqrt{x}\sqrt{y}$ $(x > 0,\ y > 0)$;

(7) $y' = -\sqrt{x}\sqrt{y}$ $(x > 0,\ y > 0)$.

13. Analyse the differential equation

$$y' = Ax^{\alpha}y^{\beta},$$

i.e., find the set of its solutions depending on the values of the parameters $A, \alpha, \beta \in \mathbb{R}$.

14. Using the Leibniz formalism, find general solutions and construct phase portraits for the following differential equations:

(1) $y' = \dfrac{y}{x+1}$;
(2) $y' = \dfrac{y+1}{x-2}$;
(3) $y' = \dfrac{x}{y-1}$;
(4) $y' = \dfrac{x-2}{y+1}$.

15. Using the Leibniz formalism, find general solutions of the following differential equations:

(1) $y' = y\mathrm{e}^x$;
(2) $y' = -y\mathrm{e}^x$;
(3) $y' = x\mathrm{e}^y$;
(4) $y' = -x\mathrm{e}^y$;
(5) $y' = y\cos x$;
(6) $y' = y\sin x$;
(7) $y' = y\sin 2x$.

16. Give an example of a function $z = \varphi(x, y)$ of two variables that cannot be represented as

(1) a sum $f(x) + g(y)$;
(2) a product $f(x) \cdot g(y)$

(with a justification!).

17. Prove that a function of two variables $z = \varphi(x, y)$ having derivatives of any order with respect to each of the variables x and y can be represented as a sum $f(x) + g(y)$ of functions of one variable differentiable arbitrarily many times if and only if the so-called mixed (partial) derivative

$$z'_{xy} \overset{\text{des}}{=} (z'_x)'_y,$$

or, in the Leibniz notation,

$$\frac{d^2 z}{dy\, dx} \overset{\text{des}}{=} \frac{d\left(\frac{dz}{dx}\right)}{dy},$$

is identically zero (all the functions are assumed to be *everywhere defined*).

18. Prove that the variables in the following functions are *not seperable*; i.e., these functions cannot be represented as products $f(x) \cdot g(y)$:

(1) $x + y$;
(2) $x - y + 1$;
(3) $x^2 + y^2$;
(4) $x^2 - y^2 + 1$;
(5) $x^2 + y^2 + 1$.

19. Think of some (reasonable) condition for the variables in the function $u = \varphi(x, y)$ to be *separable*, i.e., for the function u to be representable as a product $f(x) \cdot g(y)$ (like the condition in Problem 17).

Separation of variables in homogeneous differential equations

Definition 1. A function of two variables $F(x, y)$ is called a *homogeneous function of degree* k, or *with degree of homogeneity* k) if for all $x, y \in \mathbb{R}$

we have

$$\forall t \neq 0 \quad F(tx, ty) = t^k F(x, y).$$

Definition 2. A first-order differential equation of the form $y' = F(x, y)$ is called a *homogeneous differential equation* if its right-hand side is a *homogeneous function of degree* $k = 0$, i.e., $\forall x, y \in \mathbb{R}$

$$\forall t \neq 0 \quad F(tx, ty) = F(x, y).$$

20. Prove that any homogeneous function $F(x, y)$ of degree $k = 0$ can be represented for $x \neq 0$ as $F(x, y) = \varphi\left(\frac{y}{x}\right)$.

Thus, for $x \neq 0$ a homogeneous differential equation can be written in the Leibniz notation as

$$\frac{dy}{dx} = \varphi\left(\frac{y}{x}\right). \qquad (*)$$

Therefore, *apparently*, it is reasonable to introduce a new *unknown function* $u(x) = \frac{y(x)}{x}$, or in other words, to look for a solution $y(x)$ in the form $y = x \cdot u(x)$.

21. Making in equation $(*)$ the substitution $y = u(x)x$ described above, write the resulting differential equation for the function $u = u(x)$.

Hint: $\frac{d}{dx}(ux) = x\frac{du}{dx} + u$.

Answer: $\frac{du}{dx} = \frac{\varphi(u) - u}{x}$.

Comment. Thus, we obtain a variable separable equation, from which we find the first integral of the above equation

$$\int \frac{du}{\varphi(u) - u} - \ln|x| = C \quad \Leftrightarrow \quad G(u) - \ln|x| = C = \text{const},$$

where $G(u)$ is any antiderivative of the function

$$\psi(u) = \frac{1}{\varphi(u) - u}$$

(and C is a *"pseudo-constant"*). Substituting $u = \frac{y}{x}$, we come to an algebraic[1] equation relating the desired function y and the independent variable x, or the first integral of the original equation:

$$H(x, y) = G\left(\frac{y}{x}\right) - \ln|x| = \text{const}.$$

22. Using homogeneity, find general solutions and draw phase portraits for the following differential equations:

(1) $y' + \frac{4y + x}{y} = 0$; (2) $y' + \frac{y + x}{y} = 0$; (3) $y' = \frac{y - x}{y + x}$.

[1] In a wide sense.

23. Using homogeneity, find first integrals and solutions for the following differential equations:

(1) $y' = \dfrac{2xy}{x^2+y^2}$;
(2) $y' = \dfrac{y}{x}(1+\ln y - \ln x)$;
(3) $y^2 + x^2 y' = xyy'$.

Bernoulli equation

24. By making in the *Bernoulli differential equation*[1]

$$y' = k(x)y + f(x)y^n \tag{$**$}$$

($n \in \mathbb{Z}$; however, n can be any real number) the substitution $y = uv$, reduce it to a linear differential equation of the form (5) and to a variable separable equation.

25. Dividing both sides of the Bernoulli differential equation ($**$) by y^n, show that this equation can be reduced by some substitution (which one?) of the desired function to a linear differential equation.

Hint. Instead of the desired function $y = y(x)$, introduce the function $z = z(x) = y^{1-n}$ (in doing so, the cases $n = 0$ and $n = 1$ should be considered separately).

Planar differential equations

26. On the Oxy plane, sketch vector fields $\mathbf{V}(x,y) = (A(x,y); B(x,y))$ given by their components $(A; B)$:

(1) $(1;-1)$;
(2) $(1;2)$;
(3) $(x;1)$;
(4) $(x;y)$;
(5) $(x;-y)$;
(6) $(-x;y)$;
(7) $(-x;-y)$;
(8) $(y;1)$;
(9) $(-y;1)$;
(10) $(y;x)$;
(11) $(-y;x)$;
(12) $(-y;-x)$;
(13) $(x-1;y+1)$;
(14) $(1-x;y+1)$;
(15) $(x;x)$;
(16) $(y;-y)$.

Draw (exactly or "roughly") phase trajectories of these fields, i.e., of the corresponding systems of differential equations $(x'; y') = (A(x,y); B(x,y))$.

Hint. The following considerations can be used:

(a) An approximate direction of the field vectors is determined by *signs* of the coordinate functions A and B;

(b) In particular, the points at which the field is directed horizontally or vertically can be found, as well as *singular points* of these vector fields;

[1] It was proposed in 1695 by Jacob Bernoulli as a "problem to solve" (a generalisation of a non-homogeneous linear equation). Solutions were given in 1696–97 by Leibniz and Johann Bernoulli; both methods are presented in Exercises 24 and 25.

(c) More exact information on the direction of the field vectors can be obtained using *isoclines* (i.e., curves on which the field vectors form equal angles with, for instance, the Ox axis; see § 2.2 and exercises to Ch. II).

27. Find the general form of solutions $(x(t); y(t))$ of the differential equations with vector fields $\mathbf{V}(x,y) = (A(x,y); B(x,y))$ given by their coordinate functions $(A; B)$:

(1) $(-x, -3y)$;
(2) $(4x, y)$;
(3) $(-x, y)$;
(4) $(-y, x)$;
(5) $(-2x, y)$;
(6) $(-2y, x)$;
(7) $(-2y, -x)$;
(8) $(-x, 2y)$;
(9) $(-y, -2x)$;
(10) $(-y, 2x)$.

28. On the parameter plane Oab, draw sets of points $(a; b)$ for which the singular point $O(0;0)$ of the given systems of differential equations is nondegenerate and has a particular type, namely a node (stable, unstable, degenerate), saddle, centre (left or right), or focus (stable or unstable, left or right):

(1) $x' = ax, \quad y' = by$;
(2) $x' = ay, \quad y' = bx$;
(3) $x' = ax + y, \quad y' = by$;
(4) $x' = ax, \quad y' = x + by$.

29. Find out at what angles the phase trajectories of the systems of differential equations intersect the x- and y-axes:

(1) $x' = y, \quad y' = -\omega^2 x - 2\alpha y$;
(2) $x' = y, \quad y' = -\omega^2 x + 2\alpha y$;
(3) $x' = -y, \quad y' = -\omega^2 x + 2\alpha y$;
(4) $x' = -y, \quad y' = -\omega^2 x - 2\alpha y$,

where ω and α are positive parameters. Consider three cases: $\alpha < \omega$, $\alpha = \omega$, and $\alpha > \omega$.

Remark. For $\alpha < \omega$, singular points of the first three systems are *foci.*

Systems of differential equations in polar coordinates

30. Write in Cartesian coordinates $(x; y)$ the system of differential equations which in polar coordinates $(r; \varphi)$ has the form

$$\begin{cases} r'(t) = r f(r), \\ \varphi'(t) = 1, \end{cases} \qquad (***)$$

where $f(r)$ is a given continuously differentiable function on the interval $\mathbb{R}_+ = [0, +\infty)$.

Hint. Use the formulae $x = r \sin\varphi$, $y = r \sin\varphi$.

Answer: $\begin{cases} x' = -y + x f(r), \\ y' = x + y f(r). \end{cases}$

Comment. Thus, system $(***)$ corresponds to a vector field with components $A(x,y)=-y+xf(r)$, $B(x,y)=x+yf(r)$ $(r=\sqrt{x^2+y^2})$. If the function f is identically zero, then system $(***)$ defines a "standard centre", i.e., the point $O(0;0)$ (which is always singular for system $(***)$) is a centre. If f is *not* identically zero, then system $(***)$ can be viewed as a *"perturbation of the centre"*. The behaviour of the phase trajectories of the *"perturbed centre"* is easier to investigate without passing to Cartesian coordinates, since the differential equation $r'(t)=rf(r)$ is a *standard* one, as discussed in § 4.1.

31. Find all solutions of system $(***)$ in the case where $f(r)=a-r$ $(a>0)$. Draw phase portraits of the system both in the "polar" half-plane $Or\varphi$ (in such portraits, the positive semi-axis Or, $r>0$, is usually directed *to the right*, and the $O\varphi$ axis is directed *upwards*) and in the Cartesian plane Oxy.

Hint. The differential equation $r'=r(a-r)$ is the well-known *logistic equation* (see § 4.1).

Comment. The stationary solution $r(t)\equiv a$ of the logistic equation corresponds to a single *periodic* trajectory of system $(***)$; since $\varphi(t)=t+\varphi_0$, this is a circle of radius a $(x^2+y^2=a^2)$, and motion along it has constant angular velocity $\omega=1$ (in Cartesian coordinates, $x(t)=a\cos(t+\varphi_0)$ and $y(t)=a\sin(t+\varphi_0)$). Closed trajectories of this kind are called *cycles*.

Besides the cycle, the system has a singular point O of the *unstable focus* type: for $0<r<1$, the function $f(r)$ is positive, so $r'(t)>0$ and the point $(r;\varphi)_p$ *moves away* from the pole O. In this case, since for the solution $r(t)$ of the logistic equation we have $r(t)\to a-$ as $t\to+\infty$, while moving away from the pole the trajectory *"winds"* onto the cycle $r=a$ *from the inside*. On the other hand, if the initial point lies outside the cycle (i.e., $r(0)=r_0>a$), then $r(t)\to a+$ as $t\to+\infty$, so such a trajectory also *"winds"* onto the cycle $r=a$, but *from the outside*.

Cycles (periodic trajectories) to which close trajectories *"tend"* as $t\to+\infty$ are called *stable* (attracting) *limit cycles*. The existence of such limit cycles indicates that the corresponding dynamical system as if *"spontaneously"*, *"automatically"* enters the periodic, *"oscillatory"* regime, the regime of so-called *auto-oscillations*. Besides the stable limit cycles, we also consider *unstable* (*repelling*) *limit cycles*, where trajectories close to the cycle do not wind around the cycle but rather move away from it; there are also *semi-stable limit cycles*, where trajectories wind around the cycle on one side and move away from it on the other.

32. Find all solutions of system $(***)$ in the cases where $a>0$ and

(1) $f(r)=r-a$;

(2) $f(r)=(r-a)^2$.

Draw phase portraits of the system in the polar half-plane $Or\varphi$ and in the Cartesian plane Oxy.

33. Analyse the behaviour of solutions of systems of the form $(***)$ for the given functions $f(r)$ and sketch phase portraits of these systems in the polar half-plane and in the Cartesian plane:

(1) $f(r) = (r-1)(r-2)$;
(2) $f(r) = (r-1)(2-r)$;
(3) $f(r) = (r-1)(r-2)(r-3)$;
(4) $f(r) = (r-1)(2-r)(r-3)$;
(5) $f(r) = (r-1)^2(r-2)$;
(6) $f(r) = (r-1)(r-2)^2$;
(7) $f(r) = (r-1)^2(r-2)^2$;
(8) $f(r) = \sin(\pi r)$;
(9) $f(r) = \cos r$;
(10) $f(r) = |\sin(\pi r)|$.