L. ELSGOLTS

Differential Equations and the Calculus of Variations

ИЗДАТЕЛЬСТВО • МИР •

Л. Э. ЭЛЬСГОЛЬЦ

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ И ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

L. ELSGOLTS

Differential equations and the calculus of variations

TRANSLATED FROM THE RUSSIAN
BY
GEORGE YANKOVSKY

MIR PUBLISHERS • MOSCOW

First published 1970

Second printing 1973
Third printing 1977

TO THE READER

Mir Publishers welcome your comments on the content, translation and design of this book.

We would also be pleased to receive any proposals you care to make about our future publications.

Our address is:

USSR, 129820, Moscow I-110, GSP
Pervy Rizhsky Pereulok, 2

MIR PUBLISHERS

На английском языке

The subject of this book is the theory of differential equations and the calculus of variations. It is based on a course of lectures which the author delivered for a number of years at the Physics Department of the Lomonosov State University of Moscow.

Contents

PART TWO

THE CALCULUS OF VARIATIONS

PART ONE

Differential equations

Introduction

In the study of physical phenomena one is frequently unable to find directly the laws relating the quantities that characterize a phenomenon, whereas a relationship between the quantities and their derivatives or differentials can readily be established. One then obtains equations containing the unknown functions or vector functions under the sign of the derivative or differential.

Equations in which the unknown function or the vector function appears under the sign of the derivative or the differential are called *differential equations*. The following are some examples of differential equations:

(1) $\frac{dx}{dt} = -kx$ is the equation of radioactive disintegration (k is the disintegration constant, x is the quantity of undisintegrated substance at time t, and $\frac{dx}{dt}$ is the rate of decay proportional to the quantity of disintegrating substance).

(2) $m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}\left(t, \mathbf{r}, \frac{d\mathbf{r}}{dt}\right)$ is the equation of motion of a particle of mass m under the influence of a force $\mathbf{F}$ dependent on the time, the position of the particle (which is determined by the radius vector $\mathbf{r}$), and its velocity $\frac{d\mathbf{r}}{dt}$ The force is equal to the product of the mass by the acceleration.

(3) $\frac{d^2u}{dx^2} + \frac{d^2u}{dy^2} + \frac{d^2u}{dz^2} = 4\pi\rho(x, y, z)$ is Poisson's equation, which for example is satisfied by the potential $u(x, y, z)$ of an electrostatic field, $\rho(x, y, z)$ is the charge density.

The relation between the sought-for quantities will be found if methods are indicated for finding the unknown functions which are defined by differential equations. The finding of unknown functions defined by differential equations is the principal task of the theory of differential equations.

If in a differential equation the unknown functions or the vector functions are functions of one variable, then the differential equation is called *ordinary* (for example, Eqs. 1 and 2 above). But if the unknown function appearing in the differential equation is a function of two or more independent variables, the differential equation is called a *partial differential equation* (Eq. 3 is an instance).

The *order* of a differential equation is the highest order of the derivative (or differential) of the unknown function.

A *solution* of a differential equation is a function which, when substituted into the differential equation, reduces it to an identity.

To illustrate, the equation of radioactive disintegration

$$\frac{dx}{dt} = -kx \tag{I.1}$$

has the solution

$$x = ce^{-kt}, \tag{I.1$_1$}$$

where c is an arbitrary constant.

It is obvious that the differential equation (I.1) does not yet fully determine the law of disintegration $x = x(t)$. For a full determination, one must know the quantity of disintegrating substance x_0 at some initial instant of time t_0. If x_0 is known, then, taking into account the condition $x(t_0) = x_0$ from (I.1$_1$), we find the law of radioactive disintegration:

$$x = x_0 e^{-k(t-t_0)}.$$

The procedure of finding the solutions of a differential equation is called *integration of the differential equation*. In the above case, it was easy to find an exact solution, but in more complicated cases it is very often necessary to apply approximate methods of integrating differential equations. Just recently these approximate methods still led to arduous calculations. Today, however, high-speed computers are able to accomplish such work at the rate of several hundreds of thousands of operations per second.

Let us now investigate more closely the above-mentioned more complicated problem of finding the law of motion $\mathbf{r} = \mathbf{r}(t)$ of a particle of mass m under the action of a specified force $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$. By Newton's law,

$$m\ddot{\mathbf{r}} = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}}). \tag{I.2}$$

Consequently, the problem reduces to integrating this differential equation. Quite obviously, the law of motion is not yet fully defined by specifying the mass m and the force $\mathbf{F}$; one has also to know the initial position of the particle

$$\mathbf{r}(t_0) = \mathbf{r}_0 \tag{I.2$_1$}$$

and the initial velocity

$$\dot{\mathbf{r}}(t_0) = \dot{\mathbf{r}}_0. \tag{I.2$_2$}$$

We shall indicate an extremely natural approximate method for solving equation (I.2) with initial conditions (I.2$_1$) and (I.2$_2$); the idea of this method can also serve to prove the existence of a solution of the problem at hand.

We take the interval of time $t_0 \leqslant t \leqslant T$ over which it is required to find a solution of the equation (I.2) that will satisfy the initial conditions ($I.2_1$) and ($I.2_2$) and divide it into n equal parts of length $h = \frac{T-t_0}{n}$:

$$[t_0, t_1], [t_1, t_2], \ldots, [t_{n-1}, T],$$

where

$$t_k = t_0 + kh \qquad (k = 1, 2, \ldots, n-1).$$

For large values of n, within the limits of each one of these small intervals of time, the force $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ changes but slightly (the vector function $\mathbf{F}$ is assumed to be continuous); therefore it may be taken, approximately, to be constant over every subinterval $[t_{k-1}, t_k]$, for instance, equal to the value it has at the left-hand boundary point of each subinterval. More exactly, on the subinterval $[t_0, t_1]$ the force $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ is considered constant and equal to $\mathbf{F}(t_0, \mathbf{r}_0, \dot{\mathbf{r}}_0)$. On this assumption, it is easy, from (I.2) and the initial conditions ($I.2_1$) and ($I.2_2$), to determine the law of motion $\mathbf{r}_n(t)$ on the subinterval $[t_0, t_1]$ (the motion will be uniformly variable) and, hence, in particular, one knows the values of $\mathbf{r}_n(t_1)$ and $\dot{\mathbf{r}}_n(t_1)$. By the same method, we approximate the law of motion $\mathbf{r}_n(t)$ on the subinterval $[t_1, t_2]$ considering the force $\mathbf{F}$ as constant on this subinterval and as equal to $\mathbf{F}(t_1, \mathbf{r}_n(t_1), \dot{\mathbf{r}}_n(t_1))$. Continuing this process, we get an approximate solution $\mathbf{r}_n(t)$ to the posed problem with initial conditions for equation (I.2) over the whole interval $[t_0, T]$.

It is intuitively clear that as n tends to infinity, the approximate solution $\mathbf{r}_n(t)$ should approach the exact solution.

Note that the second-order vector equation (I.2) may be replaced by an equivalent system of two first-order vector equations if we regard the velocity $\mathbf{v}$ as the second unknown vector function:

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \quad \frac{d\mathbf{v}}{dt} = \mathbf{F}(t, \mathbf{r}, \mathbf{v}). \tag{I.3}$$

Every vector equation in three-dimensional space may be replaced by three scalar equations by projecting onto the coordinate axes. Thus, equation (I.2) is equivalent to a system of three scalar equations of the second order, and system (I.3) is equivalent to a system of six scalar equations of the first order.

Finally, it is possible to replace one second-order vector equation (I.2) in three-dimensional space by one vector equation of the first order in six-dimensional space, the coordinates here being r_x, r_y, r_z

of the radius vector $\mathbf{r}(t)$ and v_x, v_y, v_z of the velocity vector $\mathbf{v}$. *Phase space* is the term physicists use for this space. The radius vector $\mathbf{R}(t)$ in this space has the coordinates $(r_x, r_y, r_z, v_x, v_y, v_z)$. In this notation, (1.3) has the form

$$\frac{d\mathbf{R}}{dt} = \mathbf{\Phi}(t, \mathbf{R}(t)) \tag{1.4}$$

(the projections of the vector $\mathbf{\Phi}$ in six-dimensional space are the corresponding projections of the right-hand sides of the system (1.3) in three-dimensional space)

With this interpretation, the initial conditions (1.2_1) and (1.2_2) are replaced by the condition

$$\mathbf{R}(t_0) = \mathbf{R}_0 \tag{1.4_1}$$

The solution of (1.4) $\mathbf{R} = \mathbf{R}(t)$ will then be a phase trajectory, to each point of which there will correspond a certain instantaneous state of the moving particle—its position $\mathbf{r}(t)$ and its velocity $\mathbf{v}(t)$.

If we apply the above approximate method to (1.4) with initial condition (1.4_1), then on the first subinterval $[t_0, t_1]$ we must regard the vector function $\mathbf{\Phi}(t, \mathbf{R}(t))$ as constant and equal to $\mathbf{\Phi}(t_0, \mathbf{R}(t_0))$. And so, for $t_0 \leqslant t \leqslant t_0 + h$

$$\frac{d\mathbf{R}}{dt} = \mathbf{\Phi}(t_0, R(t_0));$$

from this, multiplying by dt and integrating between t_0 and t, we get the linear vector function $\mathbf{R}(t)$:

$$\mathbf{R}(t) = \mathbf{R}(t_0) + \mathbf{\Phi}(t_0, \mathbf{R}(t_0))(t - t_0).$$

In particular for $t = t_1$ we will have

$$\mathbf{R}(t_1) = \mathbf{R}(t_0) + h\mathbf{\Phi}(t_0, \mathbf{R}(t_0)).$$

Repeating the same reasoning for the subsequent subintervals, we get

$$\mathbf{R}(t_2) = \mathbf{R}(t_1) + h\mathbf{\Phi}(t_1, \mathbf{R}(t_1)),$$
$$\dots\dots\dots\dots\dots$$
$$\mathbf{R}(t_k) = \mathbf{R}(t_{k-1}) + h\mathbf{\Phi}(t_{k-1}, \mathbf{R}(t_{k-1})),$$
$$\dots\dots\dots\dots\dots$$

Applying these formulas n times we arrive at the value $\mathbf{R}(T)$.

In this method, the desired solution $\mathbf{R}(t)$ is approximately replaced by a piecewise linear vector function, the graph of which is a certain polygonal line called *Euler's polygonal curve.*

In applications, the problem for equation (1.2) is often posed differently: the supplementary conditions are specified at two points instead of one. Such a problem—unlike the problem with the conditions ($I.2_1$) and ($I.2_2$), which is called an initial-value problem or the Cauchy problem — is called a *boundary-value* problem.

For example, let it be required that a particle of mass m, moving under a force $\mathbf{F}(t, \mathbf{r}(t), \dot{\mathbf{r}}(t))$ and located at the initial instant $t=t_0$ in the position $\mathbf{r}=\mathbf{r}_0$, reach the position $\mathbf{r}=\mathbf{r}_1$ at time $t=t_1$. In other words, it is necessary to solve equation (I.2) with the boundary conditions $\mathbf{r}(t_0)=\mathbf{r}_0$, $\mathbf{r}(t_1)=\mathbf{r}_1$. Numerous problems in ballistics reduce to this boundary-value problem. It is obvious that the solution here is frequently not unique, since it is possible to reach the point $\mathbf{r}(t_1)=\mathbf{r}_1$ from the point $\mathbf{r}(t_0)=\mathbf{r}_0$ either via a flat trajectory or a plunging trajectory.

Obtaining an exact or approximate solution of initial-value problems and boundary-value problems is the principal task of the theory of differential equations, however it is often required to determine (or it is necessary to confine oneself to determining) only certain properties of solutions. For instance, one often has to establish whether periodic or oscillating solutions exist, to estimate the rate of increase or decrease of solutions, and to find out whether a solution changes appreciably for small changes in the initial values.

Let us dwell in more detail on the last one of these problems as applied to the equation of motion (I.2). In applied problems, the initial values $\mathbf{r}_0$ and $\dot{\mathbf{r}}_0$ are almost always the result of measurement and, hence, are unavoidably determined with a certain error. This quite naturally brings up the question of the effect of a small change in the initial values on the sought-for solution.

If arbitrarily small changes in the initial values are capable of giving rise to appreciable changes in the solution then the solution determined by inexact initial values $\mathbf{r}_0$ and $\dot{\mathbf{r}}_0$ usually has no applied value at all, since it does not describe the motion of the body under consideration even in an approximate fashion. We thus come to a problem, important in applications, of finding the conditions under which a small change in the initial values $\mathbf{r}_0$ and $\mathbf{r}_0$ gives rise only to a small change in the solution $\mathbf{r}(t)$ which they determine.

A similar question arises in problems in which it is required to find the accuracy with which one must specify the initial values $\mathbf{r}_0$ and $\dot{\mathbf{r}}_0$ so that a moving point should — to within specified accuracy — take up a desired trajectory or arrive in a given region.

Just as important is the problem of the effect, on the solution, of small terms on the right-hand side of equation (I.2) — small but constantly acting forces.

In certain cases, these small forces operating over a large interval of time are capable of distorting the solution drastically, and they must not be neglected. In other cases, the change in the solution due to the action of these forces is inappreciable, and if it does not exceed the required accuracy of computations, such small disturbing forces may be neglected.

We now turn to methods of integrating differential equations and the most elementary ways of investigating their solutions.

CHAPTER 1

First-order differential equations

1. First-Order Differential Equations Solved for the Derivative

An ordinary first-order differential equation of the first degree may, solving for the derivative, be represented as follows:

$$\frac{dy}{dx}=f(x, y).$$

The most elementary case of such an equation

$$\frac{dy}{dx}=f(x)$$

is considered in the course of integral calculus. In this most elementary case, the solution

$$y=\int f(x)\,dx+c$$

contains an arbitrary constant which may be determined if we know the value $y(x_0)=y_0$; then

$$y=y_0+\int_{x_0}^{x} f(x)\,dx.$$

Later on it will be proved that with certain restrictions placed on the function $f(x, y)$, the equation

$$\frac{dy}{dx}=f(x, y)$$

also has a unique solution satisfying the condition $y(x_0)=y_0$, while its *general solution* (that is, the set of solutions containing all solutions without exception) depends on one arbitrary constant.

The differential equation $\frac{dy}{dx}=f(x, y)$ establishes a relation between the coordinates of a point and the slope of the tangent $\frac{dy}{dx}$ to the graph of the solution at that point. Knowing x and y, it is possible to calculate $\frac{dy}{dx}$. Hence, a differential equation of the type under consideration defines a direction field (Fig. 1.1) and the problem of integrating the differential equation consists in finding the curves, called *integral curves*, the direction of the tangents to which at each point coincides with the direction of the field.

Example 1.

$$\frac{dy}{dx}=\frac{y}{x}.$$

At each point different from the point (0, 0), the slope of the tangent to the desired integral curve is equal to the ratio $\frac{y}{x}$, which means it coincides with the slope of a line directed from the coordinate origin to the same point (x, y). The arrows in Fig. 1.2

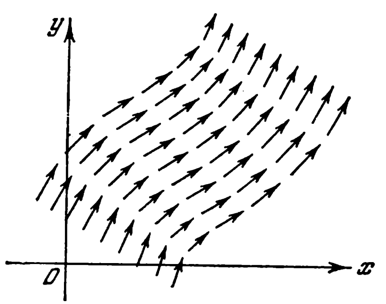


Fig. 1-1

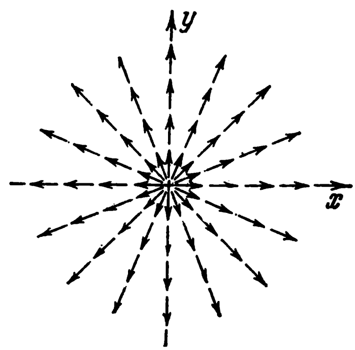


Fig. 1-2

depict a direction field defined by the equation under consideration. Obviously, in this case the straight lines $y=cx$ will be the integral curves, since the directions of these lines coincide everywhere with the direction of the field.

Example 2.

$$\frac{dy}{dx}=-\frac{x}{y}.$$

Note that the slope of the tangent to the desired integral curves $-\frac{x}{y}$ and the slope of the tangent $\frac{y}{x}$ to the integral curves of Example 1 at each point satisfy the orthogonality condition: $-\frac{x}{y}\cdot\frac{y}{x}=-1$. Consequently, a direction field defined by the differential equation under consideration is orthogonal to the direction field given in Fig. 1.2. It is obvious that the integral curves of equation $\frac{dy}{dx}=-\frac{x}{y}$ are circles with centre at the origin

$x^2+y^2=c^2$ (Fig. 1.3) (more precisely, semicircles $y=\sqrt{c^2-x^2}$ and $y=-\sqrt{c^2-x^2}$).

Example 3.

$$\frac{dy}{dx}=\sqrt{x^2+y^2}.$$

To construct a direction field, let us find the locus of points at which tangents to the desired integral curves preserve a constant direction. Such lines are called *isoclines*. We get the isocline equation by taking $\frac{dy}{dx}=k$, where k is a constant; $\sqrt{x^2+y^2}=k$ or $x^2+y^2=k^2$. Thus, in this case the isoclines are circles with centre at the origin of coordinates, and the slope of the tangent to the sought-for integral curves is equal to the radius of these circles. To construct the direction field, we shall assign certain definite values to the constant k (see Fig. 1.4, left). It is now already possible to draw, approximately, the desired integral curves (see Fig. 1.4, right).

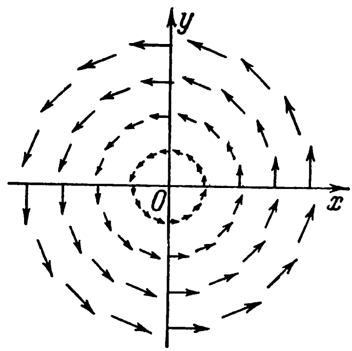


Fig. 1-3

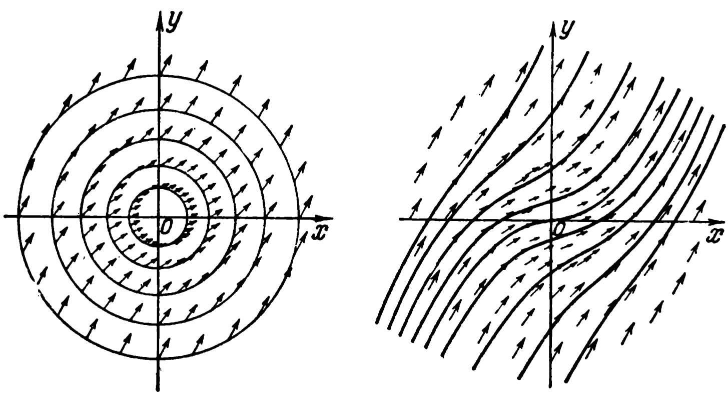


Fig. 1-4

Example 4.

$$y'=1+xy.$$

The isoclines are the hyperbolas $k=xy+1$ or $xy=k-1$; when $k=1$ the hyperbola decomposes into a pair of straight lines $x=0$

and $y=0$ (Fig. 1.5). For $k=0$ we get the isocline $1+xy=0$; this hyperbola partitions the plane into parts, in each of which y' preserves constant sign (Fig. 1.6). The integral curves $y=y(x)$ intersect the hyperbola $1+xy=0$ and pass from the region of increase of the function $y(x)$ to the region of decrease or, conversely, from the region of decrease to that of increase, and consequently, the points of maximum and minimum of the integral curves are located on the branches of this hyperbola.

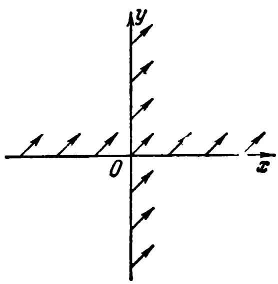


Fig. 1-5

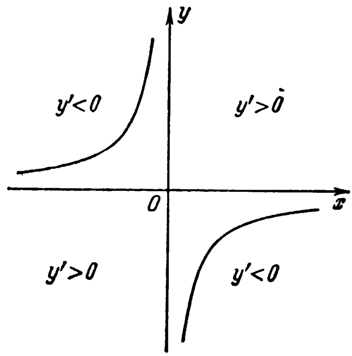


Fig. 1-6

Let us now determine the signs of the second derivative in various regions of the plane:

$$y''=xy'+y \quad \text{or} \quad y''=x(1+xy)+y=x+(x^2+1)\,y.$$

The curve $x+(x^2+1)\,y=0$ or

$$y=-\frac{x}{1+x^2} \tag{1.1}$$

(Fig. 1.7) partitions the plane into two parts, in one of which $y''<0$, and, hence, the integral curves are convex upwards, and

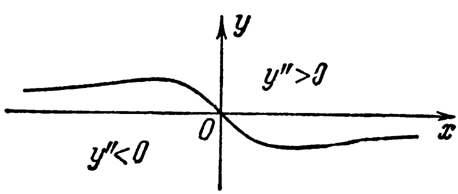


Fig. 1-7

in the other $y''>0$, and thus the integral curves are concave upwards. When passing through the curve (1.1) the integral curves pass from convexity to concavity, and, consequently, it is on this curve that the points of inflection of the integral curves are located.

As a result of this investigation we now know the regions of increase and decrease of the integral curves, the position of the points of maximum and minimum, the regions of convexity and concavity and the location of the inflection points, and also the isocline $k=1$. This information is quite sufficient for us to sketch the locations of the integral curves (Fig. 1.8), but we could draw a few more isoclines, and this would enable us to specify more accurately the location of the integral curves.

In many problems, for instance in almost all problems of a geometrical nature, the variables x and y are absolutely equivalent. It is therefore natural in such problems, if they reduce to solving a differential equation

$$\frac{dy}{dx}=f(x, y), \qquad (1.2)$$

to consider, alongside equation (1.2), also the equation

$$\frac{dx}{dy}=\frac{1}{f(x, y)}. \qquad (1.3)$$

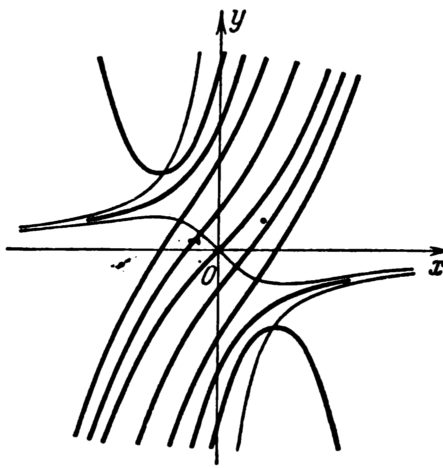


Fig. 1-8

If both of these equations are meaningful, then they are equivalent, because if the function $y=y(x)$ is a solution of the equation (1.2), then the inverse function $x=x(y)$ is a solution of (1.3), and hence, (1.2) and (1.3) have common integral curves.

But if at certain points one of the equations, (1.2) or (1.3), becomes meaningless, then it is natural at such points to replace it by the other equation.

For instance, $\frac{dy}{dx}=\frac{y}{x}$ becomes meaningless at $x=0$. Replacing it by the equation $\frac{dx}{dy}=\frac{x}{y}$, the right side of which is already meaningful at $x=0$, we find another integral curve $x=0$ of this equation in addition to the earlier found solutions $y=cx$ (see page 20).

2. Separable Equations

Differential equations of the form

$$f_2(y)\,dy=f_1(x)\,dx \qquad (1.4)$$

are called *equations with separated variables*. The functions $f_1(x)$ and $f_2(y)$ will be considered continuous.

Assume that $y(x)$ is a solution of this equation; then by substituting $y(x)$ into (1.4) we get an identity, which when integrated yields

$$\int f_2(y)\,dy = \int f_1(x)\,dx + c, \tag{1.5}$$

where c is an arbitrary constant.

We obtained a finite equation (1.5) which is satisfied by all the solutions of (1.4); note that every solution of (1.5) is a solution of (1.4), because if some function $y(x)$ when substituted reduces (1.5) to an identity, then by differentiating this identity we find that $y(x)$ also satisfies the equation (1.4).

The finite equation $\Phi(x, y) = 0$, which defines the solution $y(x)$ of the differential equation as an implicit function of x, is called the *integral* of the differential equation under study.

If the finite equation defines all solutions of a given differential equation without exception, then it is called the *complete (general) integral* of that differential equation. Thus, equation (1.5) is the complete integral of equation (1.4). For (1.5) to define y as an implicit function of x, it is sufficient to require that $f_2(y) \neq 0$.

It is quite possible that in certain problems the indefinite integrals $\int f_1(x)\,dx$ and $\int f_2(y)\,dy$ will not be expressible in terms of elementary functions; nevertheless, in this case as well we shall consider the problem of integrating the differential equation (1.4) as completed in the sense that we have reduced it to a simpler problem, one already studied in the course of integral calculus: the computation of indefinite integrals (quadratures). *

If it is required to isolate a particular solution that satisfies the condition $y(x_0) = y_0$, it will obviously be determined from the equation

$$\int_{y_0}^{y} f_2(y)\,dy = \int_{x_0}^{x} f_1(x)\,dx,$$

which we obtain from

$$\int_{y_0}^{y} f_2(y)\,dy = \int_{x_0}^{x} f_1(x)\,dx + c,$$

taking advantage of the initial condition $y(x_0) = y_0$.

* Since the term 'integral' in the theory of differential equations is often used in the meaning of the integral of a differential equation, the term 'quadrature' is ordinarily used to avoid confusion when dealing with integrals of the functions $\int f(x)\,dx$.

Example 1.

$$x\,dx + y\,dy = 0.$$

The variables are separated since the coefficient of dx is a function of x alone, whereas the coefficient of dy is a function of y alone. Integrating, we obtain

$$\int x\,dx + \int y\,dy = c \quad \text{or} \quad x^2 + y^2 = c_1^2,$$

which is a family of circles with centre at the coordinate origin (compare with Example 2 on page 20).

Example 2.

$$e^{x^2}\,dx = \frac{dy}{\ln y}.$$

Integrating, we get

$$\int e^{x^4}\,dx = \int \frac{dy}{\ln y} + c.$$

The integrals $\int e^{x^2}\,dx$ and $\int \frac{dy}{\ln y}$ are not expressible in terms of elementary functions; nevertheless, the initial equation is considered integrated because the problem has been reduced to quadratures.

Equations of the type

$$\varphi_1(x)\,\psi_1(y)\,dx = \varphi_2(x)\,\psi_2(y)\,dy$$

in which the coefficients of the differentials break up into factors depending solely on x and solely on y are called *differential equations with variables separable*, since by division by $\psi_1(y)\,\varphi_2(x)$ they may be reduced to an equation with separated variables:

$$\frac{\varphi_1(x)}{\varphi_2(x)}\,dx = \frac{\psi_2(y)}{\psi_1(y)}\,dy.$$

Note that division by $\psi_1(y)\,\varphi_2(x)$ may lead to loss of particular solutions that make the product $\psi_1(y)\cdot\varphi_2(x)$ vanish, and if the functions $\psi_1(y)$ and $\varphi_2(x)$ can be discontinuous, then extraneous solutions converting the factor

$$\frac{1}{\psi_1(y)\,\varphi_2(x)}$$

to zero may appear.

Example 3.

$\frac{dy}{dx}=\frac{y}{x}$ (compare with Example 1, page 20). Separate the variables and integrate:

$$\frac{dy}{y}=\frac{dx}{x}, \quad \int\frac{dy}{y}=\int\frac{dx}{x},$$
$$\ln|y|=\ln(x)+\ln c, \quad c>0.$$

Taking antilogarithms, we get $|y|=c|x|$. If we speak only of smooth solutions, the equation $|y|=c|x|$, where $c>0$, is equivalent to the equation $y=\pm cx$ or $y=c_1x$, where c_1 can take on either positive or negative values, but $c_1\neq 0$. If we bear in mind, however, that in dividing by y we lost the solution $y=0$, we can take it that in the solution $y=c_1x$ the constant c_1 also assumes the value $c_1=0$, in which way we obtain the solution $y=0$ that was lost earlier.

Note. If in Example 3 we consider the variables x and y to be equivalent, then equation $\frac{dy}{dx}=\frac{y}{x}$, which is meaningless at $x=0$, must be supplemented by the equation $\frac{dx}{dy}=\frac{x}{y}$ (see page 23), which obviously also has the solution $x=0$ not contained in the solution $y=c_1x$ found above.

Example 4.

$$x(1+y^2)\,dx-y(1+x^2)\,dy=0.$$

Separate the variables and integrate:

$$\frac{y\,dy}{1+y^2}=\frac{x\,dx}{1+x^2}; \quad \int\frac{y\,dy}{1+y^2}=\int\frac{x\,dx}{1+x^2}+c;$$

$$\ln(1+y^2)=\ln(1+x^2)+\ln c_1; \quad 1+y^2=c_1(1+x^2).$$

Example 5.

$$\frac{dx}{dt}=4t\sqrt{x}.$$

Find the solution $x(t)$ that satisfies the condition $x(1)=1$.

Separating variables and integrating, we have

$$\int_1^x\frac{dx}{2\sqrt{x}}=\int_1^t 2t\,dt, \quad \sqrt{x}=t^2, \quad x=t^4.$$

Example 6. As was mentioned in the Introduction, it has been established that the rate of radioactive decay is proportional to the quantity x of substance that has not yet decayed. Find x as a function of the time t if at the initial instant $t=t_0$ we have $x=x_0$.

The constant of proportionality k, called the decay constant, is assumed known. The differential equation of the process will be of the form

$$\frac{dx}{dt}=-kx \tag{1.6}$$

(the minus sign indicates a decrease in x as t increases, $k>0$). Separating the variables and integrating, we get

$$\frac{dx}{x}=-k\,dt; \quad \ln|x|-\ln|x_0|=-k(t-t_0)$$

and then

$$x=x_0e^{-k(t-t_0)}.$$

Let us also determine the half-life τ (that is, the time during which $x_0/2$ decays). Assuming $t-t_0=\tau$, we get $x_0/2=x_0e^{-k\tau}$, whence $\tau=\frac{\ln 2}{k}$.

Not only radioactive disintegration, but also any other monomolecular reaction is described on the basis of the mass action law, by the equation $\frac{dx}{dt}=-kx$, where x is the quantity of substance that has not yet reacted.

The equation

$$\frac{dx}{dt}=kx, \quad k>0, \tag{1.7}$$

which differs from (1.6) only in the sign of the right side, describes many processes of multiplication, like the multiplication of neutrons in nuclear chain reactions or the reproduction of bacteria on the assumption of an extremely favourable environment, in which case the rate of reproduction will be proportional to the number of bacteria present.

The solution of (1.7) that satisfies the condition $x(t_0)=x_0$ is of the form $x=x_0e^{k(t-t_0)}$ and, unlike the solutions of (1.6), $x(t)$ does not diminish but increases exponentially as t increases.

Example 7.

$$\frac{d\rho}{d\varphi}=\rho(\rho-2)(\rho-4).$$

Draw the integral curves without integrating the equation; ρ and φ are polar coordinates.

The equation has the obvious solutions $\rho=0$, $\rho=2$, and $\rho=4$. For $0<\rho<2$, $\frac{d\rho}{d\varphi}>0$; for $2<\rho<4$, $\frac{d\rho}{d\varphi}<0$ and for $\rho>4$, $\frac{d\rho}{d\varphi}>0$.

Consequently, the integral curves are the circles $\rho=2$ and $\rho=4$ and the spirals that wind around the circle $\rho=2$ as φ increases and that unwind from the circle $\rho=4$ as φ increases. The closed integral curves, in sufficiently small neighbourhoods of which the integral curves are spirals, are called *limit cycles*. In our example, the circles $\rho=2$ and $\rho=4$ are limit cycles.

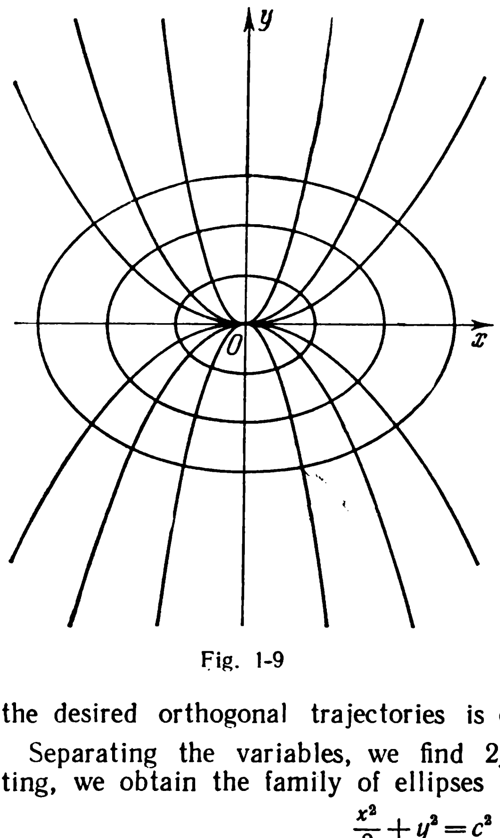


Fig. 1-9

Example 8. Find the orthogonal trajectories of the family of parabolas $y=ax^2$.

The *orthogonal trajectories* of a given family of curves are the lines that cut the given family at right angles. The slopes y_1' and y_2' of the tangents to the curves of the family and to the sought-for orthogonal trajectories must at each point satisfy the orthogonality condition $y_2'=-\frac{1}{y_1'}$. For the family of parabolas $y=ax^2$ we find $y'=2ax$, or since $a=\frac{y}{x^2}$ then $y'=\frac{2y}{x}$. Thus the differential equation of the desired orthogonal trajectories is of the form $y'=-\frac{x}{2y}$.

Separating the variables, we find $2y\,dy+x\,dx=0$ and, integrating, we obtain the family of ellipses

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

(Fig. 1.9).

Example 9. Let $u=xy$ be the potential of velocities of a plane-parallel flow of fluid. Find the equation of the flow lines.

The flow lines are the orthogonal trajectories of a family of equipotential lines $xy=c$. Find the slope of the tangent to the equipotential lines: $xy'+y=0$, $y'=-\frac{y}{x}$. Hence, the differential equation of flow lines is of the form $y'=\frac{x}{y}$ or $y\,dy=x\,dx$; integrating, we obtain $x^2-y^2=c$ or a family of hyperbolas.

Example 10. A homogeneous hollow metallic ball of inner radius r_1 and outer radius r_2 is in a stationary thermal state; the temperature on the inner surface is T_1, on the outer surface T_2. Find the temperature T at a distance r from the centre of the ball, $r_1 \leqslant r \leqslant r_2$.

For reasons of symmetry it follows that T is a function of r alone.

Since the quantity of heat remains invariable between two concentric spheres with centres at the centre of the ball (their radii can vary from r_1 to r_2), the same quantity of heat Q flows through each sphere. Hence, the differential equation describing this process is of the form

$$-4\pi k r^2 \frac{dT}{dr} = Q,$$

where k is the coefficient of thermal conduction.

Separating the variables and integrating, we obtain the desired dependence of T upon r:

$$4\pi k\, dT = -\frac{Q\,dr}{r^2};$$

$$4\pi k \int_{T_1}^{T} dT = -Q \int_{r_1}^{r} \frac{dr}{r^2},$$

$$4\pi k (T - T_1) = Q\left(\frac{1}{r} - \frac{1}{r_1}\right).$$

To determine Q, we use the condition: for $r = r_2$, $T = T_2$

$$Q = \frac{4\pi k (T_2 - T_1)}{\frac{1}{r_2} - \frac{1}{r_1}} = \frac{4\pi k (T_2 - T_1) r_1 r_2}{r_1 - r_2}$$

3. Equations That Lead to Separable Equations

Many differential equations can be reduced to equations with variables separable by changing variables. These include, for example, equations of the type

$$\frac{dy}{dx} = f(ax + by)$$

(where a and b are constants), which by the change of variables $z = ax + by$ are converted into equations with variables separable. Indeed, passing to the new variables x and z, we will have

$$\frac{dz}{dx} = a + b\frac{dy}{dx}, \quad \frac{dz}{dx} = a + bf(z)$$

or

$$\frac{dz}{a+bf(z)}=dx,$$

and the variables are separated. Integrating, we get

$$x=\int\frac{dz}{a+bf(z)}+c.$$

Example 1.

$$\frac{dy}{dx}=2x+y.$$

Setting $z=2x+y$, we have

$$\frac{dy}{dx}=\frac{dz}{dx}-2,\ \frac{dz}{dx}-2=z.$$

Separating the variables and integrating, we get

$$\frac{dz}{z+2}=dx,\ \ln|z+2|=x+\ln c,\ z=-2+ce^{x},$$

$$2x+y=-2+ce^{x},\ y=ce^{x}-2x-2$$

Example 2.

$$\frac{dy}{dx}=\frac{1}{x-y}+1.$$

Putting $x-y=z$, we have

$$\frac{dy}{dx}=1-\frac{dz}{dx},\quad 1-\frac{dz}{dx}=\frac{1}{z}+1;$$

$$\frac{dz}{dx}=-\frac{1}{z},\ z\,dz=-dx,\ z^2=-2x+c,\ (x-y)^2=-2x+c.$$

Also reducible to equations with variables separable are the so-called *homogeneous differential equations of the first order* of the form

$$\frac{dy}{dx}=f\left(\frac{y}{x}\right).$$

Indeed, after the substitution $z=\frac{y}{x}$ or $y=xz$ we get

$$\frac{dy}{dx}=x\frac{dz}{dx}+z,\quad x\frac{dz}{dx}+z=f(z),\quad \frac{dz}{f(z)-z}=\frac{dx}{x},$$

$$\int\frac{dz}{f(z)-z}=\ln|x|+\ln c,\quad x=ce^{\int\frac{dz}{f(z)-z}}.$$

Note that the right side of the homogeneous equation is a homogeneous function of the variables x and y (zero degree of homogeneity) and so an equation of the form

$$M(x, y)\,dx+N(x, y)\,dy=0$$

will be homogeneous if $M(x, y)$ and $N(x, y)$ are homogeneous functions of x and y of the same degree, because in this case

$$\frac{dy}{dx} = -\frac{M(x, y)}{N(x, y)} = f\left(\frac{y}{x}\right).$$

Example 3.

$$\frac{dy}{dx} = \frac{y}{x} + \tan\frac{y}{x}.$$

Putting $y = xz$, $\frac{dy}{dx} = x\frac{dz}{dx} + z$ and substituting into the initial equation, we have

$$x\frac{dz}{dx} + z = z + \tan z, \qquad \frac{\cos z\, dz}{\sin z} = \frac{dx}{x},$$

$$\ln|\sin z| = \ln|x| + \ln c, \ \sin z = cx, \ \sin\frac{y}{x} = cx.$$

Example 4.

$$(x+y)\,dx - (y-x)\,dy = 0.$$

Putting $y = xz$, $dy = xdz + zdx$, we get

$$(x + xz)\,dx - (xz - x)(x\,dz + z\,dx) = 0,$$
$$(1 + 2z - z^2)\,dx + x(1 - z)\,dz = 0,$$
$$\frac{(1-z)\,dz}{1+2z-z^2} + \frac{dx}{x} = 0, \ \frac{1}{2}\ln|1 + 2z - z^2| + \ln|x| = \frac{1}{2}\ln c,$$
$$x^2(1 + 2z - z^2) = c, \ x^2 + 2xy - y^2 = c.$$

Equations of the form

$$\frac{dy}{dx} = f\left(\frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2}\right) \tag{1.8}$$

are converted into homogeneous equations by translating the origin of coordinates to the point of intersection (x_1, y_1) of the straight lines

$$a_1x + b_1y + c_1 = 0 \text{ and } a_2x + b_2y + c_2 = 0.$$

Indeed, the constant term in the equations of these lines in the new coordinates $X = x - x_1$, $Y = y - y_1$ will be zero, the coefficients of the running coordinates remain unchanged, while $\frac{dy}{dx} = \frac{dY}{dX}$ The equation (1.8) is transformed to

$$\frac{dY}{dX} = f\left(\frac{a_1X + b_1Y}{a_2X + b_2Y}\right)$$

or

$$\frac{dY}{dX} = f\left(\frac{a_1 + b_1\frac{Y}{X}}{a_2 + b_2\frac{Y}{X}}\right) = \varphi\left(\frac{Y}{X}\right)$$

and is now a homogeneous equation.

This method cannot be used only when the lines $a_1x+b_1y+c_1=0$ and $a_2x+b_2y+c_2=0$ are parallel. But in this case the coefficients of the running coordinates are proportional: $\frac{a_2}{a_1}=\frac{b_2}{b_1}=k$ and (1.8) may be written as

$$\frac{dy}{dx}=f\left(\frac{a_1x+b_1y+c_1}{k(a_1x+b_1y)+c_2}\right)=F(a_1x+b_1y),$$

and consequently, as indicated on page 29, the change of variables $z=a_1x+b_1y$ transforms the equation under consideration into an equation with variables separable.

Example 5.

$$\frac{dy}{dx}=\frac{x-y+1}{x+y-3}.$$

Solving the system of equations $x-y+1=0$, $x+y-3=0$, we get $x_1=1$, $y_1=2$. Putting $x=X+1$, $y=Y+2$, we will have

$$\frac{dY}{dX}=\frac{X-Y}{X+Y}.$$

The change of variables $z=\frac{Y}{X}$ or $Y=zX$ leads to the separable equation

$$z+X\frac{dz}{dX}=\frac{1-z}{1+z}, \qquad \frac{(1+z)\,dz}{1-2z-z^2}=\frac{dX}{X},$$

$$-\frac{1}{2}\ln|1-2z-z^2|=\ln|X|-\frac{1}{2}\ln c,$$

$$(1-2z-z^2)X^2=c, \quad X^2-2XY-Y^2=c,$$

$$x^2-2xy-y^2+2x+6y=c_1.$$

4. Linear Equations of the First Order

A *first-order linear differential equation* is an equation that is linear in the unknown function and its derivative. A linear equation has the form

$$\frac{dy}{dx}+p(x)\,y=f(x), \tag{1.9}$$

where $p(x)$ and $f(x)$ will henceforward be considered continuous functions of x in the domain in which it is required to integrate equation (1.9).

If $f(x)\equiv 0$, then the equation (1.9) is called homogeneous linear. The variables are separable in a homogeneous linear equation:

$$\frac{dy}{dx}+p(x)\,y=0, \text{ whence } \frac{dy}{y}=-p(x)\,dx,$$

and, integrating, we get

$$\ln|y| = -\int p(x)\,dx + \ln c_1, \quad c_1 > 0,$$

$$y = ce^{-\int p(x)\,dx}, \quad c \neq 0, \tag{1.10}$$

In dividing by y we lost the solution $y \equiv 0$, however it can be included in the set of solutions (1.10) if we assume that c can take the value 0 as well.

The nonhomogeneous linear equation

$$\frac{dy}{dx} + p(x)\,y = f(x) \tag{1.9}$$

may be integrated by the so-called *method of variation of parameters*. In applying this method, one first integrates the appropriate (having the same left-hand member, that is) homogeneous equation

$$\frac{dy}{dx} + p(x)\,y = 0,$$

the general solution of which, as already indicated, is of the form

$$y = ce^{-\int p(x)\,dx}.$$

Given a constant c, the function $ce^{-\int p(x)\,dx}$ is a solution of the homogeneous equation. Let us now try to satisfy the nonhomogeneous equation considering c as a function of x, that is actually performing the change of variables

$$y = c(x)\,e^{-\int p(x)\,dx},$$

where $c(x)$ is a new unknown function of x.

Computing the derivative

$$\frac{dy}{dx} = \frac{dc}{dx}e^{-\int p(x)\,dx} - c(x)\,p(x)\,e^{-\int p(x)\,dx}$$

and substituting it into the original nonhomogeneous equation (1.9), we get

$$\frac{dc}{dx}e^{-\int p(x)\,dx} - c(x)\,p(x)\,e^{-\int p(x)\,dx} + p(x)\,c(x)\,e^{-\int p(x)\,dx} = f(x)$$

or

$$\frac{dc}{dx} = f(x)\,e^{\int p(x)\,dx},$$

whence, integrating we find

$$c(x) = \int f(x)\,e^{\int p(x)\,dx}\,dx + c_1;$$

and consequently

$$y=c(x)e^{-\int p(x)\,dx}=c_1e^{-\int p(x)\,dx}+e^{-\int p(x)\,dx}\int f(x)\,e^{\int p(x)\,dx}dx. \quad (1.11)$$

To summarize: the general solution of a nonhomogeneous linear equation is the sum of the general solution of the corresponding homogeneous equation

$$c_1e^{-\int p(x)\,dx}$$

and of the particular solution of the nonhomogeneous equation

$$e^{-\int p(x)\,dx}\int f(x)\,e^{\int p(x)\,dx}dx$$

obtained from (1.11) for $c_1=0$.

Note that in specific cases it is not advisable to use the cumbersome and involved formula (1.11). It is much easier to repeat each time all the calculations given above.

Example 1.

$$\frac{dy}{dx}-\frac{y}{x}=x^2.$$

Integrate the corresponding homogeneous equation

$$\frac{dy}{dx}-\frac{y}{x}=0 \quad \frac{dy}{y}=\frac{dx}{x}, \qquad \ln|y|=\ln|x|+\ln c, \qquad y=cx.$$

Consider c a function of x, then $y=c(x)x$, $\frac{dy}{dx}=\frac{dc}{dx}x+c(x)$ and, substituting into the original equation and simplifying, we get

$$\frac{dc}{dx}x=x^2 \text{ or } dc=x\,dx, \quad c(x)=\frac{x^2}{2}+c_1.$$

Hence, the general solution is

$$y=c_1x+\frac{x^3}{2}.$$

Example 2.

$$\frac{dy}{dx}-y\cot x=2x\sin x.$$

Integrate the corresponding homogeneous equation

$$\frac{dy}{dx}-y\cot x=0, \quad \frac{dy}{y}=\frac{\cos x}{\sin x}dx,$$
$$\ln|y|=\ln|\sin x|+\ln c, \quad y=c\sin x.$$

We vary the constant

$$y=c(x)\sin x, \quad y'=c'(x)\sin x+c(x)\cos x.$$

Substituting into the original equation, we get

$$c'(x)\sin x + c(x)\cos x - c(x)\cos x = 2x\sin x,$$
$$c'(x) = 2x,\quad c(x) = x^2 + c_1,$$
$$y = x^2\sin x + c_1\sin x.$$

Example 3. In an electric circuit with self-inductance, there occurs a process of establishing alternating electric current. The voltage V is a given function of the time $V = V(t)$; the resistance R and the self-inductance L are constant; the initial current is given: $I(0) = I_0$. Find the dependence of the current $I = I(t)$ on the time.

Using Ohm's law for a circuit with self-inductance, we get

$$V - L\frac{dI}{dt} = RI.$$

The solution of this linear equation that satisfies the initial condition $I(0) = I_0$ has by (1.11) the form

$$I = e^{-\frac{R}{L}t}\left[I_0 + \frac{1}{L}\int_0^t V(t)\,e^{\frac{R}{L}t}\,dt\right]. \tag{1.12}$$

For a constant voltage $V = V_0$, we get

$$I = \frac{V_0}{R} + \left(I_0 - \frac{V_0}{R}\right)e^{-\frac{R}{L}t}.$$

An interesting case is presented by a sinusoidally varying voltage $V = A\sin\omega t$. Here, by (1.12), we get

$$I = e^{-\frac{R}{L}t}\left(I_0 + \frac{A}{L}\int_0^t e^{\frac{R}{L}t}\sin\omega t\,dt\right).$$

The integral on the right side is readily evaluated.

Numerous differential equations can be reduced to linear equations by means of a change of variables. For example, *Bernoulli's equation*, of the form

$$\frac{dy}{dx} + p(x)\,y = f(x)\,y^n,\qquad n \neq 1$$

or

$$y^{-n}\frac{dy}{dx} + p(x)\,y^{1-n} = f(x), \tag{1.13}$$

is reduced to a linear equation by the change of variables $y^{1-n} = z$. Indeed, differentiating $y^{1-n} = z$, we find $(1-n)\,y^{-n}\frac{dy}{dx} = \frac{dz}{dx}$ and, substituting into (1.13), we get the linear equation

$$\frac{1}{1-n}\frac{dz}{dx}+p(x)z=f(x).$$

Example 4.

$$\frac{dy}{dx}=\frac{y}{2x}+\frac{x^2}{2y},$$

$$2y\frac{dy}{dx}=\frac{y^2}{x}+x^2, \quad y^2=z, \quad 2y\frac{dy}{dx}=\frac{dz}{dx},$$

$$\frac{dz}{dx}=\frac{z}{x}+x^2,$$

and further as in Example 1 on page 34.

Equation

$$\frac{dy}{dx}+p(x)y+q(x)y^2=f(x)$$

is called *Riccati's equation* and in the general form is not integrable by quadratures, but may be transformed into *Bernoulli's equation* by a change of variable if a single particular solution $y_1(x)$ of this equation is known. Indeed, assuming $y=y_1+z$, we get

$$y_1'+z'+p(x)(y_1+z)+q(x)(y_1+z)^2=f(x)$$

or, since $y_1'+p(x)y_1+q(x)y_1^2\equiv f(x)$, we will have the Bernoulli equation

$$z'+[p(x)+2q(x)y_1]z+q(x)z^2=0.$$

Example 5.

$$\frac{dy}{dx}=y^2-\frac{2}{x^2}.$$

In this example it is easy to choose a particular solution $y_1=\frac{1}{x}$. Putting $y=z+\frac{1}{x}$, we get $y'=z'-\frac{1}{x^2}$, $z'-\frac{1}{x^2}=\left(z+\frac{1}{x}\right)^2-\frac{2}{x^2}$ or $z'=z^2+2\frac{z}{x}$, which is Bernoulli's equation.

$$\frac{z'}{z^2}=\frac{2}{xz}+1, \quad u=\frac{1}{z}, \quad \frac{du}{dx}=-\frac{z'}{z^2},$$

$$\frac{du}{dx}=-\frac{2u}{x}-1, \quad \frac{du}{u}=-\frac{2dx}{x}, \quad \ln|u|=-2\ln|x|+\ln c,$$

$$u=\frac{c}{x^2} \qquad u=\frac{c(x)}{x^2},$$

$$\frac{c'(x)}{x^2}=-1, \quad c(x)=-\frac{x^3}{3}+c_1, \quad u=\frac{c_1}{x^2}-\frac{x}{3}, \quad \frac{1}{z}=\frac{c_1}{x^2}-\frac{x}{3},$$

$$\frac{1}{y-\frac{1}{x}}=\frac{c_1}{x^2}-\frac{x}{3}, \quad y=\frac{1}{x}+\frac{3x^2}{c_2-x^3}.$$

5. Exact Differential Equations

It may happen that the left-hand side of the differential equation

$$M(x, y)\,dx + N(x, y)\,dy = 0 \tag{1.14}$$

is the total differential of some function $u(x, y)$:

$$du(x, y) = M(x, y)\,dx + N(x, y)\,dy$$

and hence equation (1.14) takes the form

$$du(x, y) = 0.$$

If the function $y(x)$ is a solution of (1.14), then

$$du(x, y(x)) \equiv 0$$

and consequently

$$u(x, y(x)) = c. \tag{1.15}$$

where c is a constant, and conversely, if some function $y(x)$ reduces the finite equation (1.15) to an identity, then, by differentiating the identity, we get $du(x, y(x)) = 0$, and hence $u(x, y) = c$, where c is an arbitrary constant, is the complete integral of the original equation.

If the initial values $y(x_0) = y_0$ are given, then the constant c is determined from (1.15), $c = u(x_0, y_0)$, and

$$u(x, y) = u(x_0, y_0) \tag{1.15$_1$}$$

is the desired particular integral. If $\frac{\partial u}{\partial y} = N(x, y) \neq 0$ at the point (x_0, y_0), then equation (1.15$_1$) defines y as an implicit function of x.

For the left-hand side of (1.14)

$$M(x, y)\,dx + N(x, y)\,dy$$

to be the total differential of some function $u(x, y)$, it is necessary and sufficient, as we know, that

$$\frac{\partial M(x, y)}{\partial y} \equiv \frac{\partial N(x, y)}{\partial x}. \tag{1.16}$$

If this condition, first pointed out by Euler, is fulfilled, then (1.14) is readily integrable. Indeed,

$$du = M\,dx + N\,dy.$$

On the other hand,

$$du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy.$$

Consequently,

$$\frac{\partial u}{\partial x} = M(x, y); \quad \frac{\partial u}{\partial y} = N(x, y),$$

whence

$$u(x, y) = \int M(x, y)\,dx + c(y).$$

When calculating the integral $\int M(x, y)\,dx$, the quantity y is regarded as a constant, and so $c(y)$ is an arbitrary function of y. To determine the function $c(y)$ we differentiate the function $u(x, y)$ with respect to y and, since $\frac{\partial u}{\partial y} = N(x, y)$, we have

$$\frac{\partial}{\partial y}\left(\int M(x, y)\,dx\right) + c'(y) = N(x, y).$$

From this equation we determine $c'(y)$ and, integrating, find $c(y)$.

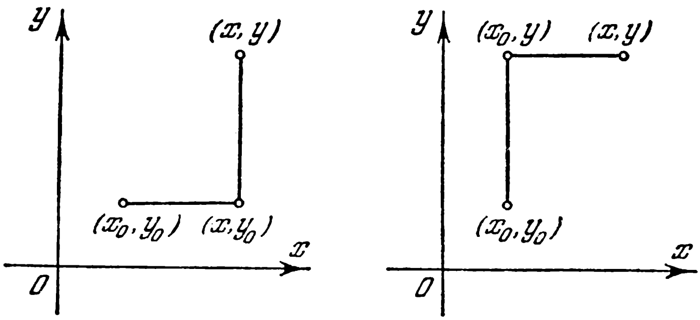


Fig. 1-10

As is known from the course of mathematical analysis, it is still simpler to determine the function $u(x, y)$ from its total differential $du = M(x, y)\,dx + N(x, y)\,dy$, taking the line integral from $M(x, y)\,dx + N(x, y)\,dy$ between some fixed point (x_0, y_0) and a point with variable coordinates (x, y) over any path:

$$u(x, y) = \int_{(x_0, y_0)}^{(x, y)} M(x, y)\,dx + N(x, y)\,dy.$$

In most cases, it is convenient to take for the path of integration a polygonal line consisting of two line segments parallel to the coordinate axes (Fig. 1.10). In this case

$$\int_{(x_0, y_0)}^{(x, y)} M\,dx + N\,dy = \int_{(x_0, y_0)}^{(x, y_0)} M\,dx + \int_{(x, y_0)}^{(x, y)} N\,dy$$

or

$$\int_{(x_0, y_0)}^{(x, y)} M dx + N dy = \int_{(x_0, y_0)}^{(x_0, y)} N dy + \int_{(x_0, y)}^{(x, y)} M dx.$$

Example 1.

$$(x+y+1)\,dx+(x-y^2+3)\,dy=0.$$

The left-hand member of the equation is the total differential of some function $u(x, y)$, since

$$\frac{\partial (x+y+1)}{\partial y} \equiv \frac{\partial (x-y^2+3)}{\partial x},$$

$$\frac{\partial u}{\partial x}=x+y+1, \quad u=\frac{x^2}{2}+xy+x+c(y),$$

$$\frac{\partial u}{\partial y}=x+c'(y), \quad x+c'(y)=x-y^2+3,$$

$$c'(y)=-y^2+3, \quad c(y)=-\frac{y^3}{3}+3y+c_1,$$

$$u=\frac{x^2}{2}+xy+x-\frac{y^3}{3}+3y+c_1.$$

Hence, the complete integral is of the form

$$3x^2+6xy+6x-2y^3+18y=c_2. \qquad (1.17)$$

A different method may also be used to determine the function $u(x, y)$:

$$u(x, y)=\int_{(x_0, y_0)}^{(x, y)} (x+y+1)\,dx+(x-y^2+3)\,dy.$$

For the initial point (x_0, y_0) we choose, for instance, the origin of

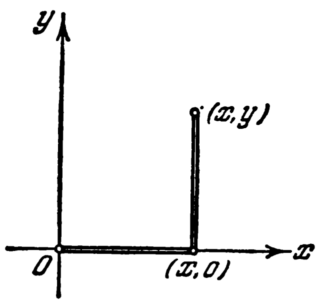


Fig. 1-11

coordinates, and we take the path of integration as shown in Fig. 1.11 (polygonal line). Then

$$u(x, y)=\int_{(0, 0)}^{(x, 0)} (x+1)\,dx+\int_{(x, 0)}^{(x, y)} (x-y^2+3)\,dy=\frac{x^2}{2}+x+xy-\frac{y^3}{3}+3y$$

and the complete integral is of the form

$$\frac{x^2}{2}+x+xy-\frac{y^3}{3}+3y=c$$

or as in (1.17).

In certain cases when the left-hand side of the equation

$$M(x, y)\,dx+N(x, y)\,dy=0 \tag{1.14}$$

is not the total differential, it is easy to choose a function $\mu(x, y)$ such that after multiplying by it the left side of (1.14) is transformed to the total differential

$$du=\mu M\,dx+\mu N\,dy.$$

Such a function μ is called an *integrating factor*. Observe that multiplication by the integrating factor $\mu(x, y)$ can lead to the appearance of extraneous particular solutions that reduce this factor to zero.

Example 2.

$$x\,dx+y\,dy+(x^2+y^2)\,x^2\,dx=0.$$

It is obvious that multiplying by the factor $\mu=\frac{1}{x^2+y^2}$ makes the left-hand member a total differential. Indeed, multiplying by $\mu=\frac{1}{x^2+y^2}$, we get

$$\frac{x\,dx+y\,dy}{x^2+y^2}+x^2\,dx=0$$

or, integrating, $\frac{1}{2}\ln(x^2+y^2)+\frac{x^3}{3}=\ln c_1$. Multiplying by 2 and then taking antilogarithms, we will have

$$(x^2+y^2)\,e^{\frac{2}{3}x^3}=c.$$

Of course it is not always so easy to find the integrating factor. In the general case, to find the integrating factor it is necessary to choose at least one particular solution (not identically zero) of the partial differential equation

$$\frac{\partial \mu M}{\partial y}=\frac{\partial \mu N}{\partial x},$$

or in expanded form

$$\frac{\partial \mu}{\partial y}M+\mu\frac{\partial M}{\partial y}=\frac{\partial \mu}{\partial x}N+\frac{\partial N}{\partial x}\mu,$$

which, when it is divided by μ and certain terms are transposed, yields

$$\frac{\partial \ln \mu}{\partial y}M-\frac{\partial \ln \mu}{\partial x}N=\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}. \tag{1.18}$$

In the general case, integrating this partial differential equation is by no means an easier task than integrating the original equation, though in some cases a particular solution to equation (1.18) may easily be found.

Besides, if we consider that the integrating factor is a function solely of one argument (for example, only of $x+y$ or x^2+y^2, or a function of x alone, or of y only, and so forth), we can then easily integrate the equation (1.18) and indicate the conditions under which an integrating factor of the form under consideration exists. In this way, classes of equations are isolated for which an integrating factor is readily found.

For example, let us find the conditions for which the equation $M\,dx+N\,dy=0$ has an integrating factor dependent solely on x, $\mu=\mu(x)$. Equation (1.18) is then simplified to

$$-\frac{d\ln\mu}{dx}N=\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y},$$

whence, taking $\dfrac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}$ to be a continuous function of x, we get

$$\ln\mu=\int\frac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}\,dx+\ln c,$$

$$\mu=ce^{\int\frac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}\,dx}. \tag{1.19}$$

We can take $c=1$, since it is sufficient to have only one integrating factor.

If $\dfrac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}$ is a function of x alone, then there exists an integrating factor dependent solely on x and equal to (1.19), otherwise there does not exist an integrating factor of the form $\mu(x)$.

The condition for the existence of an integrating factor dependent solely on x is, for example, fulfilled for the linear equation

$$\frac{dy}{dx}+p(x)y=f(x) \quad \text{or} \quad [p(x)y-f(x)]\,dx+dy=0.$$

Indeed, $\dfrac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}=p(x)$ and hence $\mu=e^{\int p(x)\,dx}$. Quite analogously we can find the conditions for the existence of integrating factors of the form

$\mu(y)$, $\mu(x\pm y)$, $\mu(x^2\pm y^2)$, $\mu(x\cdot y)$, $\mu\left(\frac{y}{x}\right)$, and so forth.

Example 3. Does the equation

$$x\,dx + y\,dy + x\,dy - y\,dx = 0 \tag{1.20}$$

have an integrating factor of the form $\mu = \mu(x^2 + y^2)$?

Put $x^2 + y^2 = z$. Equation (1.18), for $\mu = \mu(x^2 + y^2) = \mu(z)$, takes the form

$$2(My - Nx)\frac{d\ln\mu}{\partial z} = \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}$$

and from this

$$\ln|\mu| = \frac{1}{2}\int \varphi(z)\,dz + \ln c$$

or

$$\mu = ce^{\frac{1}{2}\int \varphi(z)\,dz}, \tag{1.21}$$

where

$$\varphi(z) = \frac{\dfrac{\partial N}{\partial x} - \dfrac{\partial M}{\partial y}}{My - Nx}.$$

For the existence of an integrating factor of the given form, it is necessary and, on the assumption of the continuity of $\varphi(z)$, sufficient that $\dfrac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{My - Nx}$ be a function of $x^2 + y^2$ alone. In that case

$$\frac{\dfrac{\partial N}{\partial x} - \dfrac{\partial M}{\partial y}}{My - Nx} = -\frac{2}{x^2 + y^2}$$

and hence the integrating factor $\mu = \mu(x^2 + y^2)$ exists and is equal to (1.21). For $c = 1$ we have

$$\mu = e^{-\int \frac{dz}{z}} = \frac{1}{z} = \frac{1}{x^2 + y^2}.$$

Multiplying (1.20) by $\mu = \frac{1}{x^2 + y^2}$ reduces it to the form

$$\frac{x\,dx + y\,dy}{x^2 + y^2} + \frac{x\,dy - y\,dx}{x^2 + y^2} = 0$$

or

$$\frac{\frac{1}{2}d(x^2 + y^2)}{x^2 + y^2} + \frac{d\left(\frac{y}{x}\right)}{1 + \left(\frac{y}{x}\right)^2} = 0,$$

$$\frac{1}{2}d\ln(x^2 + y^2) + d\arctan\frac{y}{x} = 0.$$

Integrating, we get

$$\ln\sqrt{x^2+y^2} = -\arctan\frac{y}{x} + \ln c$$

and after taking antilogarithms we will have

$$\sqrt{x^2+y^2} = ce^{\arctan\frac{y}{x}}.$$

or in polar coordinates $\rho = ce^{-\varphi}$, i. e. a family of logarithmic spirals.

Example 4. Find the shape of a mirror that reflects, parallel to a given direction, all the rays emanating from a given point.

Locate the origin at a given point and direct the axis of abscissas parallel to the direction given by hypothesis. Let a ray fall on the mirror at the point $M(x, y)$. Consider (Fig. 1.12) the section of the

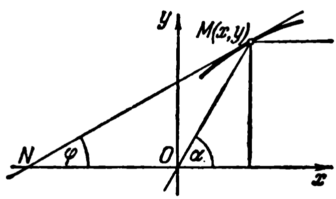


Fig. 1-12

mirror (cut by the xy-plane) that passes through the axis of abscissas and the point M. Draw a tangent MN to the section of the surface of the mirror at the point $M(x, y)$. Since the angle of incidence of the ray is equal to the angle of reflection, the triangle MNO is an isosceles triangle. Therefore,

$$\tan\varphi = y' = \frac{y}{x+\sqrt{x^2+y^2}}.$$

The homogeneous equation thus obtained is readily integrable by the change of variables

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

but a still easier way is to rationalize the denominator and write the equation as

$$x\,dx + y\,dy = \sqrt{x^2+y^2}\,dx.$$

The equation has the obvious integrating factor

$$\mu = \frac{1}{\sqrt{x^2+y^2}}, \quad \frac{x\,dx+y\,dy}{\sqrt{x^2+y^2}} = dx, \quad \sqrt{x^2+y^2} = x+c,$$

$$y^2 = 2cx + c^2$$

(a family of parabolas).

Note. This problem is still more easily solved in the coordinates x and ρ, where $\rho=\sqrt{x^2+y^2}$; here the equation of the section of desired surfaces takes the form

$$dx=d\rho, \quad \rho=x+c.$$

One can prove the existence of an integrating factor, or, what is the same thing, the existence of a nonzero solution of the partial differential equation (1.18) (see page 40) in a certain domain if the functions M and N have continuous derivatives and if at least one of these functions does not vanish. Thus, the integrating-factor method may be regarded as a general method of integrating equations of the form

$$M(x, y)\,dx+N(x, y)\,dy=0;$$

however, because of the difficulty of finding the integrating factor this method is for the most part used only when the integrating factor is obvious.

6. Theorems of the Existence and Uniqueness of Solution of the Equation $\frac{dy}{dx}=f(x, y)$

The class of differential equations that are integrable by quadratures is extremely narrow; for this reason, since Euler's day, approximate methods have become very important in the theory of differential equations.

At the present time, in view of the rapid development of computational technology, approximate methods are of incomparably greater importance.

It is now frequently advisable to apply approximate methods even when an equation may be integrated by quadratures. What is more, even if the solution may be simply expressed in terms of elementary functions, it will often be found that using tables of these functions is more cumbersome than an approximate integration of the equation by computer. However, in order to apply one or another method of aproximate integration of a differential equation, it is first necessary to be sure of the existence of the desired solution and also of the uniqueness of the solution, because in the absence of uniqueness it will not be clear what solution is to be determined.

In most cases, the proof of the theorem of the existence of a solution yields at the same time a method for finding an exact or approximate solution. This elevates still more the significance of existence theorems. For example, Theorem 1.1, which is proved below, substantiates Euler's method of approximate integration of differential equations, which consists in the fact that the desired integral curve of the differential equation $\frac{dy}{dx}=f(x, y)$ that passes through the point (x_0, y_0) is replaced by a polygonal line consisting of straight lines (Fig. 1.13), each segment of which is tangent to the integral curve at one of its boundary points. When applying this method for the approximate calculation of the value of the desired solution of $y(x)$ at the point $x=b$, the segment $x_0 \leqslant x \leqslant b$ (if $b > x_0$) is subdivided into n equal parts by the points $x_0, x_1, x_2, \ldots, x_{n-1}, x_n$, where $x_n=b$. The length of each subdivision $x_{i+1}-x_i=h$ is called the *interval of calculation*, or step. Denote by y_i the approximate values of the desired solution at the points x_i.

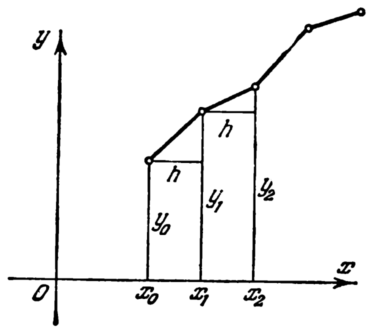


Fig. 1-13

To compute y_1, on the interval $x_0 \leqslant x \leqslant x_1$ replace the desired integral curve by a segment of its tangent at the point (x_0, y_0). Hence, $y_1=y_0+hy'_0$ where $y'_0=f(x_0, y_0)$ (see Fig. 1.13). In similar fashion we calculate:

$$y_2=y_1+hy'_1, \text{ where } y'_1=f(x_1, y_1);$$
$$y_3=y_2+hy'_2, \text{ where } y'_2=f(x_2, y_2);$$
$$\cdots\cdots\cdots\cdots$$
$$y_n=y_{n-1}+hy'_{n-1}, \text{ where } y'_{n-1}=f(x_{n-1}, y_{n-1}).$$

If $b < x_0$, the calculation scheme remains the same, but the interval of calculation h is negative.

It is natural to expect that as $h \to 0$ the *Euler polygonal curves* approach the graph of the desired integral curve and consequently Euler's method yields a more and more precise value of the desired solution at the point b as the calculation interval h decreases. Proof of this assertion at the same time brings us to the following fundamental theorem on the existence and uniqueness of a solution of the equation $\frac{dy}{dx}=f(x, y)$ with the initial condition $y(x_0)=y_0$ in extremely general su'ficient conditions imposed on the function $f(x, y)$.

Theorem 1.1 (on the existence and uniqueness of solution). *If in the equation*

$$\frac{dy}{dx}=f(x, y) \tag{1.22}$$

the function $f(x, y)$ is continuous in the rectangle D:

$$x_0 - a \leqslant x \leqslant x_0 + a, \quad y_0 - b \leqslant y \leqslant y_0 + b,$$

and satisfies, in D, the Lipschitz condition

$$|f(x, y_1) - f(x, y_2)| \leqslant N |y_1 - y_2|,$$

where N is a constant, then there exists a unique solution $y = \bar{y}(x)$, $x_0 - H \leqslant x \leqslant x_0 + H$, of equation (1.22) that satisfies the condition $y(x_0) = y_0$, where

$$H < \min\left(a, \frac{b}{M}, \frac{1}{N}\right),$$

$$M = \max f(x, y) \text{ in } D.$$

The conditions of the theorem require some explanation. It cannot be asserted that the desired solution $y = \bar{y}(x)$ of (1.22) satisfying the condition $y(x_0) = y_0$ will exist at $x_0 - a \leqslant x \leqslant x_0 + a$, since the integral curve $y = \bar{y}(x)$ may leave the rectangle D through the

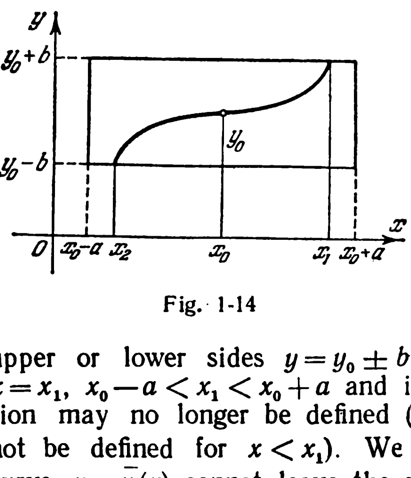


Fig. 1-14

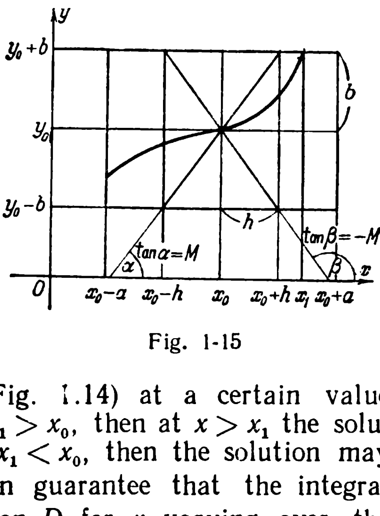


Fig. 1-15

upper or lower sides $y = y_0 \pm b$ (Fig. 1.14) at a certain value $x = x_1$, $x_0 - a < x_1 < x_0 + a$ and if $x_1 > x_0$, then at $x > x_1$ the solution may no longer be defined (if $x_1 < x_0$, then the solution may not be defined for $x < x_1$). We can guarantee that the integral curve $y = \bar{y}(x)$ cannot leave the region D for x varying over the interval $x_0 - H \leqslant x \leqslant x_0 + H$, where H is the least of two numbers a, $\frac{b}{M}$ (Fig. 1.15), since the slope of the tangent to the desired integral curve lies between the slope M and the slope $-M$ of the straight lines depicted in Fig. 1.15. If these straight lines, between which the desired integral curve lies, go beyond the rectangle D through its horizontal sides $y = y_0 \pm b$, then the abscissas of the points of intersection of these sides will be $x_0 \pm \frac{b}{M}$; consequently the abscissa of the point of exit of the integral curve from the rectangle D can only be less than or equal to $x_0+\frac{b}{M}$ and greater than or equal to $x_0-\frac{b}{M}$.

One can prove the existence of the desired solution on the interval $x_0-H\leqslant x\leqslant x_0+H$, where $H=\min\left(a, \frac{b}{M}\right)$, however it is simpler first to prove the existence of a solution on the interval $x_0-H\leqslant x\leqslant x_0+H$, where $H<\min\left(a, \frac{b}{M}, \frac{1}{N}\right)$ and then conditions will be indicated in the future such that with their fulfilment the solution may be continued.

The Lipschitz condition

$$|f(x, y_1)-f(x, y_2)|\leqslant N|y_1-y_2|$$

may be replaced by a somewhat cruder, yet a more easily verifiable condition of the existence of a partial derivative $f'_y(x, y)$ bounded in absolute value in the region D.

Indeed, if in the rectangle D

$$|f'_y(x, y)|\leqslant N,$$

we get, by the mean-value theorem,

$$|f(x, y_1)-f(x, y_2)|=f'_y(x, \xi)|y_1-y_2|,$$

where ξ is a value intermediate between y_1 and y_2. Hence, the point (x, ξ) lies in D and for this reason

$$|f'_y(x, \xi)|\leqslant N \text{ and } |f(x, y_1)-f(x, y_2)|\leqslant N|y_1-y_2|.$$

It is easy to give examples of functions $f(x, y)$ (say, $f(x,y)=|y|$ in the neighbourhood of the points $(x, 0)$) for which the Lipschitz condition holds, but the derivative $\frac{\partial f}{\partial y}$ does not exist at certain points and hence the condition $\left|\frac{\partial f}{\partial y}\right|\leqslant N$ is weaker than the Lipschitz condition.

Proof of the existence and uniqueness theorem. Replace the differential equation

$$\frac{dy}{dx}=f(x, y) \tag{1.22}$$

having the initial condition

$$y(x_0)=y_0 \tag{1.23}$$

by the equivalent integral equation

$$y = y_0 + \int_{x_0}^{x} f(x, y)\,dx. \tag{1.24}$$

Indeed, if a certain function $y = \bar{y}(x)$, when substituted, turns equation (1.22) into an identity and satisfies the condition (1.23), then, integrating the identity (1.22) and taking into account the condition (1.23), we find that $y = \bar{y}(x)$ reduces equation (1.24) to an identity as well. But if some function $y = \bar{y}(x)$, when substituted, reduces (1.24) to an identity, it will obviously also satisfy the condition (1.23); differentiating the identity (1.24), we will find that $y = \bar{y}(x)$ also reduces equation (1.22) to an identity.

Construct Euler's polygonal line $y = y_n(x)$ emanating from the point (x_0, y_0) with calculation interval $h_n = \frac{H}{n}$ on the segment $x_0 \leqslant x \leqslant x_0 + H$, where n is a positive integer (in exactly the same way we prove the existence of a solution on the interval $x_0 - H \leqslant x \leqslant x_0$) Euler's polygonal line that passes through the point (x_0, y_0) cannot leave the region D for $x_0 \leqslant x \leqslant x_0 + H$ (or $x_0 - H \leqslant x \leqslant x_0$), since the slope of each segment of the polygonal line is less than M in absolute value.

We break up the subsequent proof of the theorem into three parts:

(1) *The sequence $y = y_n(x)$ converges uniformly.*

(2) *The function $\bar{y}(x) = \lim_{n \to \infty} y_n(x)$ is a solution of the integral equation* (1.24).

(3) *The solution $\bar{y}(x)$ of the equation* (1.24) *is unique.*

Proof. (1) By the definition of Euler's polygonal line,

$$y'_n(x) = f(x_k, y_k) \quad \text{for} \quad x_k \leqslant x \leqslant x_{k+1}, \quad k = 0, 1, \ldots, n-1$$

(the right-hand derivative is taken at the corner point x_k), or we denote

$$y'_n(x) = f(x, y_n(x)) + [f(x_k, y_k) - f(x, y_n(x))] \tag{1.25}$$

as

$$f(x_k, y_k) - f(x, y_n(x)) = \eta_n(x).$$

By virtue of the uniform continuity of the function $f(x, y)$ in D we have

$$|\eta_n(x)| = |f(x_k, y_k) - f(x, y_n(x))| < \varepsilon_n \tag{1.26}$$

for $n > N(\varepsilon_n)$, where $\varepsilon_n \to 0$ as $n \to \infty$, since $|x - x_k| \leqslant h_n$ and $|y_k - y_n(x)| < Mh_n$ and $h_n = \frac{H}{n} \to 0$ as $n \to \infty$.

Integrating (1.25) with respect to x from x_0 to x and taking into account that $y_n(x_0)=y_0$, we get

$$y_n(x)=y_0+\int_{x_0}^{x} f(t,\ y_n(t))\,dt+\int_{x_0}^{x} \eta_n(t)\,dt. \tag{1.27}$$

Here n may take on any positive integer value and so for integer $m>0$

$$y_{n+m}(x)=y_0+\int_{x_0}^{x} f(t,\ y_{n+m}(t))\,dt+\int_{x_0}^{x} \eta_{n+m}(t)\,dt. \tag{1.28}$$

Subtracting (1.27) from (1.28) termwise and taking the absolute value of the difference, we get

$$\begin{aligned}|y_{n+m}(x)-y_n(x)|&=\left|\int_{x_0}^{x}[f(t,\ y_{n+m}(t))-f(t,\ y_n(t))]\,dt+\right.\\&\left.+\int_{x_0}^{x}\eta_{n+m}(t)\,dt-\int_{x_0}^{x}\eta_n(t)\,dt\right|\leqslant\\&\leqslant\int_{x_0}^{x}|f(t,\ y_{n+m}(t))-f(t,\ y_n(t))|\,dt+\\&+\int_{x_0}^{x}|\eta_{n+m}(t)|\,dt+\int_{x_0}^{x}|\eta_n(t)|\,dt\end{aligned}$$

for $x_0\leqslant x\leqslant x_0+H$ or, taking into consideration (1.26) and the Lipschitz condition:

$$|y_{n+m}(x)-y_n(x)|\leqslant N\int_{x_0}^{x}|y_{n+m}(t)-y_n(t)|\,dt+(\varepsilon_{n+m}+\varepsilon_n)\cdot H.$$

Hence,

$$\begin{aligned}&\max_{x_0\leqslant x\leqslant x_0+H}|y_{n+m}(x)-y_n(x)|\leqslant\\&\qquad\leqslant N\max\int_{x_0}^{x}|y_{n+m}(t)-y_n(t)|\,dt+(\varepsilon_{n+m}+\varepsilon_n)H,\end{aligned}$$

whence

$$\max_{x_0\leqslant x\leqslant x_0+H}|y_{n+m}(x)-y_n(x)|\leqslant\frac{(\varepsilon_{n+m}+\varepsilon_n)H}{1-NH}<\varepsilon$$

for any $\varepsilon>0$ given sufficiently large $n>N_1(\varepsilon)$.

And so

$$\max_{x_0\leqslant x\leqslant x_0+H}|y_{n+m}(x)-y_n(x)|<\varepsilon$$

for $n > N_1(\varepsilon)$, that is, the sequence of continuous functions $y_n(x)$ converges uniformly for $x_0 \leqslant x \leqslant x_0 + H$:

$$y_n(x) \rightrightarrows \bar{y}(x),$$

where $\bar{y}(x)$ is a continuous function.

(2) In equation (1.27) let us pass to the limit as $n \to \infty$:

$$\lim_{n\to\infty} y_n(x) = y_0 + \lim_{n\to\infty} \int_{x_0}^{x} f(x, y_n(x))\,dx + \lim_{n\to\infty} \int_{x_0}^{x} \eta_n(x)\,dx$$

or

$$\bar{y}(x) = y_0 + \lim_{n\to\infty} \int_{x_0}^{x} f(x, y_n(x))\,dx + \lim_{n\to\infty} \int_{x_0}^{x} \eta_n(x)\,dx. \quad (1.29)$$

By virtue of the uniform convergence of $y_n(x)$ to $\bar{y}(x)$ and the uniform continuity of the function $f(x, y)$ in D, the sequence $f(x, y_n(x)) \rightrightarrows f(x, \bar{y}(x))$.

Indeed,

$$|f(x, \bar{y}(x)) - f(x, y_n(x))| < \varepsilon,$$

where $\varepsilon > 0$ if $|\bar{y}(x) - y_n(x)| < \delta(\varepsilon)$, but $|\bar{y}(x) - y_n(x)| < \delta(\varepsilon)$, if $n > N_1(\delta(\varepsilon))$ for all x of the interval $x_0 \leqslant x \leqslant x_0 + H$.

And so $|f(x, \bar{y}(x)) - f(x, y_n(x))| < \varepsilon$ for $n > N_1(\delta(\varepsilon))$, where N_1 is not dependent on x.

By virtue of the uniform convergence of the sequence $f(x, y_n(x))$ to $f(x, \bar{y}(x))$, in (1.29) a passage to the limit is possible under the integral sign. Besides, taking into account that $|\eta_n(\varepsilon)| < \varepsilon_n$, where $\varepsilon_n \to 0$ as $n \to \infty$, in (1.29) we finally get

$$\bar{y}(x) = y_0 + \int_{x_0}^{x} f(x, \bar{y}(x))\,dx.$$

Thus, $\bar{y}(x)$ satisfies equation (1.24).

(3) Assume the existence of two noncoincident solutions $y_1(x)$ and $y_2(x)$ of equation (1.24); thus,

$$\max_{x_0 < x < x_0 + H} |y_1(x) - y_2(x)| \neq 0.$$

A termwise subtraction from the identity

$$y_1(x) \equiv y_0 + \int_{x_0}^{x} f(x, y_1(x))\,dx$$

of the identity

$$y_2(x) \equiv y_0 + \int_{x_0}^{x} f(x, y_2(x))\,dx,$$

yields

$$y_1(x)-y_2(x)\equiv\int_{x_0}^{x}[f(x,\ y_1(x))-f(x,\ y_2(x))]\,dx.$$

Hence

$$\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|=$$

$$=\max_{x_0\leqslant x\leqslant x_0+H}\left|\int_{x_0}^{x}[f(x,\ y_1(x))-f(x,\ y_2(x))]\,dx\right|\leqslant$$

$$\leqslant\max_{x_0\leqslant x\leqslant x_0+H}\left|\int_{x_0}^{x}|f(x,\ y_1(x))-f(x,\ y_2(x))|\,dx\right|.$$

Taking advantage of the Lipschitz condition, we will have

$$\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|\leqslant N\max_{x_0\leqslant x\leqslant x_0+H}\left|\int_{x_0}^{x}|y_1(x)-y_2(x)|\,dx\right|\leqslant$$

$$\leqslant N\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|\max_{x_0\leqslant x\leqslant x_0+H}\left|\int_{x_0}^{x}dx\right|=$$

$$=NH\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|.$$

The inequality obtained

$$\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|\leqslant NH\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|\quad(1.30)$$

is inconsistent if $\max\limits_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|\neq 0$, since by hypothesis $H<\frac{1}{N}$ and from (1.30) it follows that $NH=1$.

The contradiction is eliminated only for

$$\max_{x_0\leqslant x\leqslant x_0+H}|y_1(x)-y_2(x)|=0,$$

that is, if $y_1(x)\equiv y_2(x)$ for $x_0\leqslant x\leqslant x_0+H$.

Note 1. The existence of a solution of equation (1.22) might have been proved by a different method only if continuity of the function $f(x,\ y)$ (without the Lipschitz condition) is assumed; however, continuity alone of the function $f(x,\ y)$ is insufficient for proving the uniqueness of the solution.

Note 2. The existence and uniqueness of the solution $y=y(x)$ are proved only on the interval $x_0-H\leqslant x\leqslant x_0+H$; however, by taking the point $(x_0+H,\ y(x_0+H))$ for the initial point, it is possible, by repeating the reasoning, to extend the solution over an interval of length H_1 if, of course, the conditions of the existence and uniqueness theorem are fulfilled in the neighbourhood of the new initial point. Continuing this process in certain cases, it is possible to extend the solution over the entire semi-axis $x \geqslant x_0$ or even over the entire axis $-\infty < x < \infty$, if the solutions are also extended in the direction of smaller values of x. However, other cases are possible even if the function $f(x, y)$ is defined for any values of x and y.

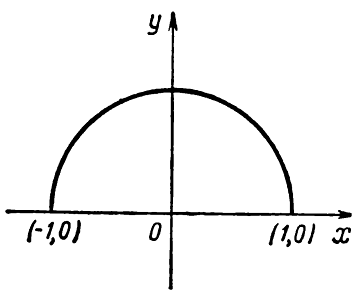


Fig. 1-16

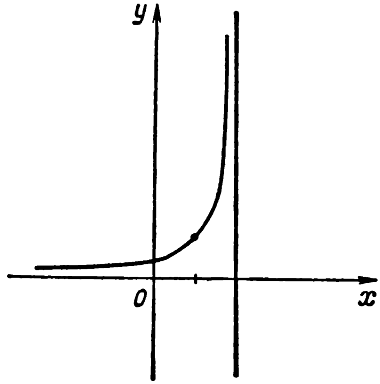


Fig. 1-17

It is possible that the integral curve becomes unextendable due to its approach to a point at which the conditions of the existence and uniqueness theorem are violated, or the integral curve approaches the asymptote parallel to the y-axis.

These possibilities are illustrated in the following examples:

(1) $\frac{dy}{dx} = -\frac{x}{y}$, $y(0) = 1$. Separating variables and integrating, we get

$$x^2 + y^2 = c^2, \qquad y = \pm\sqrt{c^2 - x^2}, \qquad c = 1,\ y = \sqrt{1 - x^2}.$$

The solution cannot be extended beyond the limits of the interval $-1 < x < 1$. At the boundary points $(-1, 0)$ and $(1, 0)$ the right side of the equation $\frac{dy}{dx} = -\frac{x}{y}$ is discontinuous. The conditions of the existence theorem are violated (Fig. 1-16).

(2) $\frac{dy}{dx} = y^2$, $y(1) = 1$. Separating the variables and integrating, we get

$$y = -\frac{1}{x - c}, \qquad c = 2, \qquad y = -\frac{1}{x - 2}$$

and the integral curve is extendable only up to the asymptote $x = 2$ $(-\infty < x < 2)$ (Fig. 1.17).

At the present time the theorem of the existence and uniqueness of solutions not only of differential equations but also of equations of other kinds is very frequently proved by the method of fixed points. The most elementary theorem on fixed points is the contraction-mapping principle.

Contraction-Mapping Principle. *If in a complete metric space* M we specify an operator A that satisfies the conditions:*

(1) the operator A carries the points of the space M into points of the same space: if $y \subset M$ then $A(y) \subset M$;

(2) the operator A brings the points closer together, more precisely,

$$\rho(A[y],\ A[z]) \leqslant \alpha\rho(y,\ z),$$

where y and z are any points of the space M, $\alpha < 1$, and is independent of the choice of y and z, $\rho(y, z)$ is the distance between the points y and z in the M space, then there exists a unique fixed point $\bar{y}$ in the M space, $A[\bar{y}] = \bar{y}$, and this point may be found by the method of successive approximations, i. e. $\bar{y} = \lim\limits_{n \to \infty} y_n$, where $y_n = A[\bar{y}_{n-1}]$ $(n = 1, 2, \ldots)$; the point y_0 is chosen arbitrarily in the M space.

Proof. We shall prove that the sequence

$$y_0,\ y_1,\ y_2,\ \ldots,\ y_n,\ \ldots$$

is fundamental. Evaluate the distance between adjacent terms of this sequence:

$$\left.\begin{aligned}
&\rho(y_2,\ y_1) = \rho(A[y_1],\ A[y_0]) \leqslant \alpha\rho(y_1,\ y_0),\\
&\rho(y_3,\ y_2) = \rho(A[y_2],\ A[y_1]) \leqslant \alpha\rho(y_2,\ y_1) \leqslant \alpha^2\rho(y_1,\ y_0),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\rho(y_{n+1},\ y_n) = \rho(A[y_n],\ A[y_{n-1}]) \leqslant\\
&\qquad\qquad \leqslant \alpha\rho(y_n,\ y_{n-1}) \leqslant \alpha^n\rho(y_1,\ y_0),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}\right\} \quad (1.31)$$

* The space M is called *metric* if a function $\rho(y, z)$ is defined in it of a pair of points of this space satisfying the following conditions for any points of the space y and z:

(1) $\rho(y, z) \geqslant 0$; note that $\rho(y, y) = 0$ and from $\rho(y, z) = 0$ it follows that $y = z$;

(2) $\rho(y, z) = \rho(z, y)$;

(3) $\rho(y, z) \leqslant \rho(y, u) + \rho(u, z)$ (triangle rule). The function ρ is called the *distance* in M space.

A metric space M is *complete* if every fundamental sequence of points of the space M converges in that space. Recall that a sequence $y_1, y_2, \ldots, y_n, \ldots$ is called fundamental if for every $\varepsilon > 0$ there is a number $N(\varepsilon)$ such that for $n \geqslant N(\varepsilon)$ the distance $\rho(y_n, y_{n+m}) < \varepsilon$ for any interval $m > 0$.

If we now apply the triangle rule $m-1$ times and use the inequality (1.31), we get

$$\rho(y_n, y_{n+m}) \leqslant \rho(y_n, y_{n+1}) + \rho(y_{n+1}, y_{n+2}) + \ldots$$
$$\ldots + \rho(y_{n+m-1}, y_{n+m}) \leqslant [\alpha^n + \alpha^{n+1} + \ldots + \alpha^{n+m-1}]\,\rho(y_1, y_0) =$$
$$= \frac{\alpha^n - \alpha^{n+m}}{1-\alpha}\rho(y_1, y_0) < \frac{\alpha^n}{1-\alpha}\rho(y_1, y_0) < \varepsilon$$

for sufficiently large n. Consequently, the sequence $y_0, y_1, y_2, \ldots, y_n, \ldots$ is fundamental and, by virtue of the completeness of the M space, it converges to a certain element of this space: $\lim\limits_{n\to\infty} y_n = \bar{y}$, $\bar{y} \subset M$.

We will now prove that $\bar{y}$ is a fixed point. Let $A[\bar{y}] = \bar{\bar{y}}$. Applying the triangle rule twice, we get

$$\rho(\bar{y}, \bar{\bar{y}}) \leqslant \rho(\bar{y}, y_n) + \rho(y_n, y_{n+1}) + \rho(y_{n+1}, \bar{\bar{y}}).$$

For any $\varepsilon > 0$, we can choose $N(\varepsilon)$ such that for $n \geqslant N(\varepsilon)$

(1) $\rho(\bar{y}, y_n) < \frac{\varepsilon}{3}$, since $\bar{y} = \lim\limits_{n\to\infty} y_n$;

(2) $\rho(y_n, y_{n+1}) < \frac{\varepsilon}{3}$, since the sequence y_n is fundamental;

(3) $\rho(y_{n+1}, \bar{\bar{y}}) = \rho(A[y_n], A[\bar{y}]) \leqslant \alpha\rho(y_n, \bar{y}) < \frac{\varepsilon}{3}$, whence $\rho(\bar{y}, \bar{\bar{y}}) < \varepsilon$, where ε may be chosen arbitrarily small. Hence,

$$\rho(\bar{y}, \bar{\bar{y}}) = 0 \quad \text{and} \quad \bar{y} = \bar{\bar{y}}, \quad A[\bar{y}] = \bar{y}.$$

It remains to prove that the fixed point $\bar{y}$ is unique. If there existed yet another fixed point $\bar{z}$, then $\rho(A[\bar{y}], A[\bar{z}]) = \rho(\bar{y}, \bar{z})$, but this contradicts hypothesis (2) of the theorem.

Let us apply the contraction-mapping principle to the proof of the theorem of the existence and uniqueness of the solution $y(x)$ $(x_0 - h_0 \leqslant x \leqslant x_0 + h_0)$ of the differential equation $\frac{dy}{dx} = f(x, y)$ which satisfies the condition $y(x_0) = y_0$ on the assumption that in the region D

$$x_0 - a \leqslant x \leqslant x_0 + a, \quad y_0 - b \leqslant y \leqslant y_0 + b$$

the function f is continuous and hence is bounded $|f| \leqslant M$ and satisfies the Lipschitz condition

$$|f(x, y) - f(x, z)| \leqslant N|y - z|.$$

The number $h_0 \leqslant \min\left(a, \frac{b}{M}\right)$ will be chosen more precisely below.

Consider a complete metric space C whose points are all possible continuous functions $y(x)$ defined on the interval $x_0 - h_0 \leqslant x \leqslant x_0 + h_0$

whose graphs lie in the region D, and the distance is defined by the equation

$$\rho(y, z) = \max |y - z|,$$

where the maximum is taken for x varying on the interval

$$x_0 - h_0 \leqslant x \leqslant x_0 + h_0.$$

This space is often considered in various problems of mathematical analysis and is called the *space of uniform convergence*, since convergence in the sense of the metric of this space signifies uniform convergence.

Replace the differential equation $\frac{dy}{dx} = f(x, y)$ having initial condition $y(x_0) = y_0$ by the equivalent integral equation

$$y = y_0 + \int_{x_0}^{x} f(x, y)\, dx. \qquad (1.24)$$

Consider the operator

$$A[y] = y_0 + \int_{x_0}^{x} f(x, y(x))\, dx,$$

which associates with every continuous function $y(x)$ that is specified on the interval $x_0 - h_0 \leqslant x \leqslant x_0 + h_0$ and does not go beyond the region D, a continuous function $A[y]$ defined on the same interval, whose graph also lies within D, since $\left| \int_{x_0}^{x} f(x, y)\, dx \right| \leqslant Mh_0 \leqslant b$. The operator $A[y]$ thus satisfies the condition (1) of the contraction-mapping principle.

Here, equation (1.24) is written as $y = A[y]$ and hence to prove the existence and uniqueness theorem it remains to prove the existence, in the space C, of a unique fixed point $\bar{y}(x)$ of the operator A, since in this case $\bar{y} = A[\bar{y}]$ and (1.24) is satisfied.

To prove the theorem it remains to check and see whether the operator A satisfies the condition (2) of the contraction-mapping principle:

$$\rho(A[y], A[z]) \leqslant \alpha\rho(y, z), \ \alpha < 1,$$

where

$$\rho(A[y], A[z]) = \max \left| \int_{x_0}^{x} [f(x, y) - f(x, z)]\, dx \right|.$$

Using the Lipschitz inequality, we get

$$\rho(A[y],\ A[z]) \leqslant N \max \left| \int_{x_0}^{x} |y-z|\,dx \right| \leqslant$$

$$\leqslant N \max |y-z| \max \left| \int_{x_0}^{x} dx \right| = Nh_0 \max |y-z| = Nh_0 \rho(y,\ z).$$

Choosing h_0 so that $Nh_0 \leqslant \alpha < 1$, we find that the operator A satisfies the condition

$$\rho(A[y],\ A[z]) \leqslant \alpha\rho(y,\ z),\ \alpha < 1.$$

Hence, according to the contraction-mapping principle there exists a unique fixed point $\bar{y}(x)$ of the operator A, or, what is the same thing, a unique continuous solution of the equation (1.24), and it may be found by the method of successive approximations.

In quite analogous fashion it is possible to prove the theorem of the existence and uniqueness of the solution $y_1(x),\ y_2(x),\ \ldots,\ y_n(x)$ for the system of equations

$$\frac{dy_i}{dx} = f_i(x,\ y_1,\ y_2,\ \ldots,\ y_n),\ y_i(x_0) = y_{i0} \quad (i = 1,\ 2,\ \ldots,\ n) \qquad (1.32)$$

or

$$y_i = y_{i0} + \int_{x_0}^{x} f(x\ \ y_1,\ y_2,\ \ldots,\ y_n)\,dx \quad (i = 1,\ 2,\ \ldots,\ n) \ (1.33)$$

on the assumption that in the region D defined by the inequalities

$$x_0 - a \leqslant x \leqslant x_0 + a,\ y_{i0} - b_i \leqslant y_i \leqslant y_{i0} + b_i \quad (i = 1,\ 2,\ \ldots,\ n),$$

the right sides of (1.32) satisfy the conditions:

(1) all the functions $f_i(x,\ y_1,\ y_2,\ \ldots,\ y_n)(i = 1,\ 2,\ \ldots,\ n)$ are continuous and hence also bounded, $|f_i| \leqslant M$;

(2) all the functions $f_i (i = 1,\ 2,\ \ldots,\ n)$ satisfy the Lipschitz condition:

$$|f_i(x,\ y_1,\ y_2,\ \ldots,\ y_n) - f_i(x,\ z_1,\ z_2,\ \ldots,\ z_n)| \leqslant N \sum_{i=1}^{n} |y_i - z_i|.$$

Now, a system of n continuous functions $(y_1,\ y_2,\ \ldots,\ y_n)$, i. e. an n-dimensional vector function $Y(x)$ with coordinates $y_1(x)$, $y_2(x),\ \ldots,\ y_n(x)$ defined on the inteval $x_0 - h_0 \leqslant x \leqslant x_0 + h_0$, where $h_0 \leqslant \min\left(a,\ \frac{b_1}{M},\ \ldots,\ \frac{b_n}{M}\right)$, will be a point of C space and will be chosen more exactly below. The distance in C space is defined by the equation

$$\rho(Y(x),\ Z(x)) = \sum_{i=1}^{n} \max |y_i - z_i|,$$

where $z_1, z_2, \ldots, z_n$ are coordinates of the vector function $Z(x)$.

It is not difficult to verify that in such a definition of distance, the set C of n-dimensional vector functions $Y(x)$ turns into a complete metric space. The operator A is defined by the equation

$$A[Y] = \left(y_{10} + \int_{x_0}^{x} f_1(x, y_1, y_2, \ldots, y_n)\,dx,\right.$$

$$\left. y_{20} + \int_{x_0}^{x} f_2(x, y_1, y_2, \ldots, y_n)\,dx, \ldots, y_{n0} + \int_{x_0}^{x} f_n(x, y_1, y_2, \ldots, y_n)\,dx\right),$$

i. e. the action of A on the point $(y_1, y_2, \ldots, y_n)$ yields a point of the same space C with coordinates equal to the right-hand sides of the system (1.33).

The point $A[Y]$ belongs to C since all its coordinates are continuous functions that lie within D if the coordinates of the vector function Y have not gone beyond the region D.

Indeed,

$$\left|\int_{x_0}^{x} f_i(x, y_1, y_2, \ldots, y_n)\,dx\right| \leqslant M \left|\int_{x_0}^{x} dx\right| \leqslant Mh_0 \leqslant b_i,$$

and hence $|y_i - y_{i0}| \leqslant b_i$.

It remains to verify fulfillment of the condition (2) of the contraction-mapping principle:

$\rho\,(A[Y]),\ A[Z]) =$

$$= \sum_{i=1}^{n} \max \left|\int_{x_0}^{x} [f_i(x, y_1, y_2, \ldots, y_n) - f_i(x, z_1, z_2, \ldots, z_n)]\,dx\right| \leqslant$$

$$\leqslant \sum_{i=1}^{n} \max \left|\int_{x_0}^{x} |f_i(x, y_1, y_2, \ldots, y_n) - f_i(x, z_1, z_2, \ldots, z_n)|\,dx\right| \leqslant$$

$$\leqslant N \sum_{i=1}^{n} \max \left|\int_{x_0}^{x} \sum_{i=1}^{n} |y_i - z_i|\,dx\right| \leqslant$$

$$\leqslant N \sum_{i=1}^{n} \max |y_i - z_i| \sum_{i=1}^{n} \max \left|\int_{x_0}^{x} dx\right| = Nnh_0\rho(Y,\ Z).$$

Consequently, if one chooses $h_0 \leqslant \frac{\alpha}{nN}$ where $0 < \alpha < 1$, or $Nnh_0 \leqslant \alpha < 1$, then condition (2) of the contraction-mapping principle will be satisfied and there will be a unique fixed point $\bar{y}$, which may be found by the method of successive approximations. But, by definition of the operator A, the condition $\bar{y}=A(\bar{y})$ is equivalent to the identities

$$\bar{y}_i \equiv y_{i0}+\int_{x_0}^{x} f_i(x, \bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)\,dx \quad (i=1, 2, \ldots, n),$$

where $\bar{y}_i (i=1, 2, \ldots, n)$ are coordinates of the vector function $\bar{y}$, that is, $\bar{y}$ is a unique solution of the system (1.33).

Example 1. Find several successive approximations y_1, y_2, y_3 to the solution of the equation

$$\frac{dy}{dx}=x^2+y^2; \quad y(0)=0, \quad -1 \leqslant x \leqslant 1, \quad -1 \leqslant y \leqslant 1.$$

$$y=\int_0^x (x^2+y^2)\,dx, \quad h_0=\min\left(1, \frac{1}{2}\right)=\frac{1}{2}.$$

Putting $y_0(x) \equiv 0$, we have

$$y_1=\int_0^x x^2\,dx=\frac{x^3}{3}, \quad y_2=\int_0^x \left(x^2+\frac{x^6}{9}\right)dx=\frac{x^3}{3}+\frac{x^7}{63},$$

$$y_3=\int_0^x \left[x^2+\left(\frac{x^3}{3}+\frac{x^7}{63}\right)^2\right]dx=\frac{x^3}{3}+\frac{x^7}{63}\left(1+\frac{2x^4}{33}+\frac{x^8}{945}\right).$$

Example 2. Under what restrictions does the linear equation

$$\frac{dy}{dx}+p(x)\,y=f(x)$$

satisfy the conditions of the existence and uniqueness theorem?

To satisfy the first condition of the theorem it is sufficient that the functions $p(x)$ and $f(x)$ be continuous on the interval of variation x, $a_1 \leqslant x \leqslant a_2$, that we are considering. In this way, the second condition of the existence and uniqueness theorem will be fulfilled, since the partial derivative with respect to y of the right side of the equation $\frac{dy}{dx}=-p(x)\,y+f(x)$ is $-p(x)$ and, due to the continuity of the function $p(x)$ on the interval $a_1 \leqslant x \leqslant a_2$, is bounded in absolute value (see page 47). And so if $p(x)$ and $f(x)$ are continuous on the interval $a_1 \leqslant x \leqslant a_2$, then through every point (x_0, y_0) where $a_1 < x_0 < a_2$, and y_0 is given arbitrarily, there passes a unique integral curve of the linear equation under consideration.

Theorem 1.2 (on the continuous dependence of a solution on a parameter and on the initial values). *If the right side of the differential equation*

$$\frac{dy}{dx}=f(x, y, \mu) \qquad (1.34)$$

is continuous with respect to μ *for* $\mu_0 \leqslant \mu \leqslant \mu_1$ *and satisfies the conditions of the existence and uniqueness theorem, and the Lipschitz constant* N *does not depend on* μ, *then the solution* $y(x, \mu)$ *of this equation that satisfies the condition* $y(x_0) = y_0$ *depends continuously on* μ.

Construct Euler's polygonal lines $y_n = y_n(x, \mu)$ which are continuous functions of μ and, repeating the reasoning on pages 45-51, we find that the sequence $y_n(x, \mu)$ converges uniformly not only in x but also in μ for $x_0 < x \leqslant x_0 + H$, $\mu_0 \leqslant \mu \leqslant \mu_1$, since N and H do not depend on μ if $H < \min\left(a, \frac{b}{M}, \frac{1}{N}\right)$, where $M \geqslant |f(x, y, \mu)|$. Thus, the solution $y = \bar{y}(x, \mu)$ of the equation

$$y = y_0 + \int_{x_0}^{x} f(x, y, \mu)\, dx, \tag{1.35}$$

which is the limit of the sequence $y_n(x, \mu)$, is continuous not only with respect to x, but also with respect to μ.

Note. If one applies to equation (1.35) the method of successive approximations, the successive approximations $y = y_n(x, \mu)$, which are continuous functions of x and μ, uniformly converge to the solution $\bar{y}(x, \mu)$ of the equation (1.35) (since $\alpha = Nh < 1$ is not dependent on μ). Hence, this method can also be used to prove the continuous dependence of the solution on x and μ.

It is obvious that the proof will not change if the right side of the equation (1.34) is a continuous function of several parameters, on the assumption, of course, that the Lipschitz constant N does not depend on them.

Using the same method under similar conditions it might be possible to prove the continuous dependence of the solution $y(x, x_0, y_0)$ of the equation $\frac{dy}{dx} = f(x, y)$ on the initial values x_0 and y_0; it would only be necessary to diminish h_0 somewhat, for otherwise the solutions defined by the initial values close to x_0, y_0 might go beyond the region D for values of x already lying on the interval $x_0 - h_0 < x < x_0 + h_0$.

However, it is still simpler, by a change of variables, to reduce the problem of the dependence of the solution on the initial values to the already considered case of the dependence of the solution on parameters contained on the right side of (1.34). Indeed, putting $z = y(x, x_0, y_0) - y_0$, $t = x - x_0$, we transform the equation $\frac{dy}{dx} = f(x, y)$ with initial condition $y(x_0) = y_0$ into $\frac{dz}{dt} = f(t + x_0, z + y_0)$, $z(0) = 0$, to which one can then apply the theorem on the continuous dependence of the solution on the parameters x_0 and y_0, if the function f is continuous and satisfies the Lipschitz condition.

Analogous theorems may be proved by the same methods for systems of equations.

Observe that the continuous dependence of the solution $y, (x, x_0, y_0)$, $x_0 \leqslant x \leqslant b$ (or $b \leqslant x \leqslant x_0$) upon the initial values x_0 and y_0 signifies that for any $\varepsilon > 0$, a $\delta(\varepsilon, b) > 0$ may be found such that from the inequalities

$$|x_0 - \bar{x}_0| < \delta(\varepsilon, b) \text{ and } |y_0 - \bar{y}_0| < \delta(\varepsilon, b)$$

there will follow the inequality

$$|y(x, x_0, y_0) - y(x, \bar{x}_0, \bar{y}_0)| < \varepsilon \qquad (1.36)$$

for $x_0 \leqslant x \leqslant l$ (Fig. 1.18).

Generally speaking, the number $\delta(\varepsilon, b)$ decreases with increasing b, and as $b \to \infty$ it can approach zero. For this reason, it is by

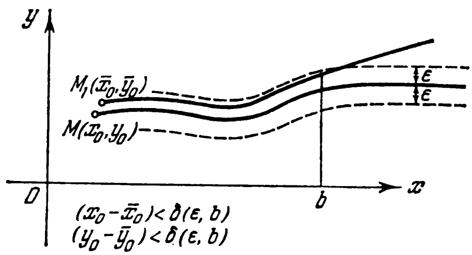


Fig. 1-18

far not always possible to choose a number $\delta(\varepsilon) > 0$ for which the inequality (1.36) would be satisfied for all $x > x_0$, that is, it does not always happen that solutions which have close-lying initial values remain so for arbitrarily large values of the argument.

A solution that changes but slightly for an arbitrary but sufficiently small variation of the initial values, given arbitrarily large values of the argument, is called stable. Stable solutions will be examined in more detail in Chapter 4.

Theorem 1.3 (on the analytical dependence of a solution on a parameter, Poincaré's theorem). *The solution $x(t, \mu)$ of the differential equation $\dot{x} = f(t, x, \mu)$ which satisfies the condition $x(t_0) = x_0$, depends analytically on the parameter μ in the neighbourhood of the value $\mu = \mu_0$, if the function f in the given range of t and x and in some neighbourhood of the point μ_0 is continuous with respect to t and is analytically dependent on μ and x.*

An analogous assertion also holds for the system of equations

$$\dot{x}_i(t) = f_i(t, x_1, x_2, \ldots, x_n, \mu) \quad (i = 1, 2, \ldots, n),$$

and in this case it is assumed that the functions f_i are continuous with respect to the first argument and are analytically dependent on all the other arguments.

We do not give a detailed proof of this theorem (and the same goes for other theorems requiring application of the theory of analytic functions) and refer the reader to a paper by A. Tikhonov [4] which contains the simplest proof of the theorem on the analytic dependence of a solution on a parameter.

The underlying idea of Tikhonov's proof is: assuming that μ can take on complex values as well, the existence is proved of the limit $\lim\limits_{\Delta\mu \to 0} \frac{\Delta_\mu x(t, \mu)}{\Delta\mu} = \frac{\partial x}{\partial \mu}$, which thus signifies analytic dependence of the solution on μ. The existence of this limit follows from the fact that the ratio $\frac{\Delta_\mu x}{\Delta_\mu}$ satisfies the linear differential equation

$$\frac{d}{dt}\frac{\Delta_\mu x}{\Delta\mu} = \frac{f(t, x(t, \mu+\Delta\mu), \mu+\Delta\mu) - f(t, x(t, \mu), \mu+\Delta\mu)}{\Delta_\mu x(t, \mu)} \frac{\Delta_\mu x(t, \mu)}{\Delta\mu} + \frac{f(t, x(t, \mu), \mu+\Delta\mu) - f(t, x(t, \mu), \mu)}{\Delta\mu}, \qquad \left.\frac{\Delta_\mu x}{\Delta\mu}\right|_{t=t_0} = 0,$$

the solution of which is unique and, as the increment $\Delta\mu$ tends to zero by any law whatsoever, approaches the unique solution of the equation

$$\frac{dz}{dt} = \frac{\partial f}{\partial x} z + \frac{\partial f}{\partial \mu}, \qquad z(t_0) = 0.$$

Theorem 1.4 (on the differentiability of solutions). *If in the neighbourhood of a point* (x_0, y_0) *a function* $f(x, y)$ *has continuous derivatives to the order* k *inclusive, the solution* $y(x)$ *of the equation*

$$\frac{dy}{dx} = f(x, y), \tag{1.37}$$

that satisfies the initial condition $y(x_0) = y_0$, *has continuous derivatives to order* $(k+1)$ *inclusive in some neighbourhood of the point* (x_0, y_0).

Proof. Substituting $y(x)$ into equation (1.37), we get the identity

$$\frac{dy}{dx} \equiv f(x, y(x)), \tag{1.37$_1$}$$

and hence the solution $y(x)$ has a continuous derivative $f(x, y(x))$ in some neighbourhood of the point under consideration. Then, by virtue of the existence of continuous derivatives of the function f, there will exist a continuous second derivative of the solution

$$\frac{d^2y}{dx^2} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f(x, y(x)).$$

If $k > 1$, then by virtue of the existence of continuous second-order derivatives of the function f, it is possible, by differentiating the identity (1.37_1) once again, to detect the existence also of a third derivative of the solution:

$$\frac{d^3y}{dx^3} = \frac{\partial^2 f}{\partial x^2} + 2\frac{\partial^2 f}{\partial x\,\partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 + \frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f\right).$$

Repeating this argument k times, we prove the theorem.

Now consider the points (x_0, y_0) in the neighbourhood of which there is no solution of the equation $\frac{dy}{dx} = f(x, y)$ that satisfies the condition $y(x_0) = y_0$, or the solution exists but is not unique. Such points are called *singular points*.

A curve that consists entirely of singular points is called *singular*. If the graph of a certain solution consists entirely of singular points, then the *solution* is called *singular*.

To find singular points or singular curves it is first of all necessary to find a set of points in which the conditions of the existence and uniqueness theorem are violated, since only such points can include singular points. Of course, not every point at which the conditions of the theorem of existence and uniqueness of solution are violated is a singular point, since the conditions of this theorem are sufficient for the existence and uniqueness of the solution but are not necessary.

The first condition of the existence and uniqueness theorem (see page 45) is violated at the discontinuity points of the function $f(x, y)$; note that if the function $f(x, y)$ increases without bound in absolute value upon approaching (by any path) some isolated point of discontinuity (x_0, y_0), then in those problems in which the variables x and y are equivalent (as we have already agreed) the equation $\frac{dy}{dx} = f(x, y)$ must be replaced by the equation $\frac{dx}{dy} = \frac{1}{f(x, y)}$, for which the right-hand side is now continuous at the point (x_0, y_0) if it is taken that $\frac{1}{f(x_0, y_0)} = 0$.

Hence, in problems in which the variables x and y are equivalent the first condition of the existence and uniqueness theorem is violated at those points at which both the function $f(x, y)$ and $\frac{1}{f(x, y)}$ are discontinuous.

One particularly often has to consider equations of the form

$$\frac{dy}{dx} = \frac{M(x, y)}{N(x, y)}, \tag{1.38}$$

where the functions $M(x, y)$ and $N(x, y)$ are continuous. In this case the functions $\frac{M(x, y)}{N(x, y)}$ and $\frac{N(x, y)}{M(x, y)}$ will at the same time be discontinuous only at those points (x_0, y_0) at which $M(x_0, y_0) = N(x_0, y_0) = 0$ and the limits

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} \frac{M(x, y)}{N(x, y)}$$

and

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} \frac{N(x, y)}{M(x, y)}$$

do not exist.

Let us consider several typical singular points of the equation (1.38).

Example 3.

$$\frac{dy}{dx} = \frac{2y}{x}.$$

The right sides of this equation and of the equation $\frac{dx}{dy} = \frac{x}{2y}$ are discontinuous at the point $x = 0$, $y = 0$. Integrating the equation, we get $y = cx^2$, which is a family of parabolas (Fig. 1.19), and

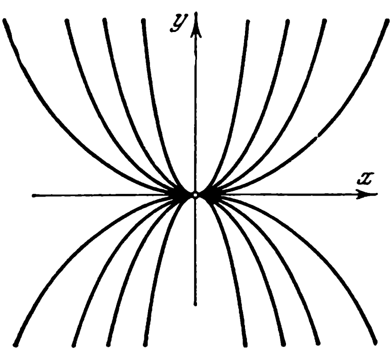


Fig. 1-19.

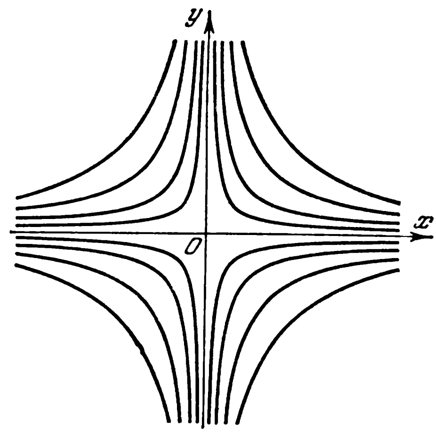


Fig. 1-20

$x = 0$. The singular point at the coordinate origin is called a *nodal point*.

Example 4.

$$\frac{dy}{dx} = -\frac{y}{x}.$$

The right sides of this equation and of equation $\frac{dx}{dy}=-\frac{x}{y}$ are discontinuous at the point $x=0$, $y=0$. Integrating the equation, we get $y=\frac{c}{x}$, which is a family of hyperbolas (Fig. 1.20), and the straight line $x=0$. The singular point at the origin is called a *saddle point*.

Example 5.

$$\frac{dy}{dx}=\frac{x+y}{x-y}.$$

The right sides of this equation and of equation $\frac{dx}{dy}=\frac{x-y}{x+y}$ are discontinuous at the point $x=0$, $y=0$. Integrating the homogeneous

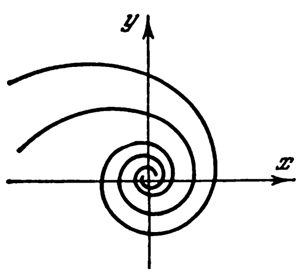


Fig. 1-21

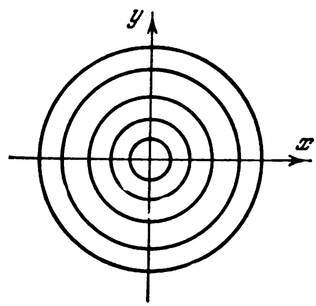


Fig. 1-22

equation (compare with Example 3 on page 42), we have

$$\sqrt{x^2+y^2}=ce^{\arctan\frac{y}{x}}$$

or in polar coordinates $\rho=ce^{\varphi}$, which represents logarithmic spirals (Fig. 1.21). The singular point of this kind is called a *focal point*.

Example 6.

$$\frac{dy}{dx}=-\frac{x}{y}.$$

The right sides of this equation and of the equation $\frac{dx}{dy}=-\frac{y}{x}$ are discontinuous at the point $x=0$, $y=0$. Integrating the equation, we get $x^2+y^2=c^2$, which is a family of circles with centre at the coordinate origin (Fig. 1.22). The singular point of this type, i.e. the singular point whose neighbourhood is filled with a family of closed integral curves, is called a *centre*. In this example there is no solution that satisfies the condition $y(0)=0$.

In Chapter 4 we shall return to the problem of singular points and their classification from a somewhat different point of view.

The second condition of Theorem 1.1 on the existence and uniqueness of solution—the Lipschitz condition, or a cruder condition requiring the existence of a bounded partial derivative $\frac{\partial f}{\partial y}$ is most often violated at points, upon approaching which $\frac{\partial f}{\partial y}$ increases without bound, i.e. at points at which $\frac{1}{\frac{\partial f}{\partial y}}=0$.

Generally speaking, the equation $\frac{1}{\frac{\partial f}{\partial y}}=0$ defines a certain curve at points of which uniqueness may be violated. If at points of this curve uniqueness is violated, then the curve will be singular; if, besides, the curve turns out to be integral, then we have a singular integral curve.

The curve $\frac{1}{\frac{\partial f}{\partial y}}=0$ may have several branches, then with respect to each branch one has to decide whether that branch will be a singular curve and whether it will be an integral curve.

Example 7. Has the equation $\frac{dy}{dx}=y^2+x^2$ a singular solution?

The conditions of the existence and uniqueness theorem are fulfilled in the neighbourhood of any point; hence there is no singular solution.

Example 8. Has the equation $\frac{dy}{dx}=\sqrt[3]{(y-x)^2}+5$ a singular solution?

The right side is continuous but the partial derivative $\frac{\partial f}{\partial y}=\frac{2}{3}(y-x)^{-\frac{1}{3}}$ increases without bound as it approaches the straight line $y=x$; consequently, uniqueness may be violated on the straight line $y=x$. But the function $y=x$ does not satisfy this equation and therefore there is no singular solution.

Example 9. Does the equation $\frac{dy}{dx}=\sqrt[3]{(y-x)^2}+1$ have a singular solution?

As in the previous example, the equation $\frac{1}{\frac{\partial f}{\partial y}}=0$ is of the form $y=x$, but this time the function $y=x$ satisfies the given equation. It remains to find out whether uniqueness has been violated at points of this straight line. By a change of variables $z=y-x$ we reduce the initial equation to an equation with variables separable, after which it is easy to find the solution: $y-x=\frac{(x-c)^3}{27}$. The curves of this family pass through points of the graph of the solution $y=x$ (Fig. 1.23). Thus, uniqueness is violated at every point of the straight line $y=x$ and the function $y=x$ is a singular solution.

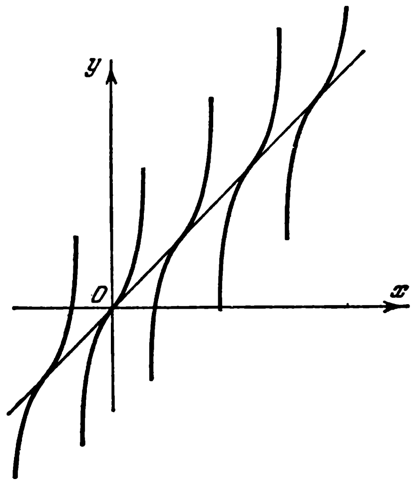


Fig. 1-23

This example shows that continuity alone on the right side of the equation

$$\frac{dy}{dx}=f(x, y), \qquad y(x_0)=y_0$$

is insufficient for uniqueness of solution of the basic initial problem, yet it may be proved that in this case the existence of a solution is already ensured.

7. Approximate Methods of Integrating First-Order Equations

In Sec. 6 we investigated two approximate methods of integrating differential equations: Euler's method and the method of successive approximation. However, both methods have essential drawbacks as a result of which they are comparatively rarely used in actual approximate calculations.

The value of an approximate method depends on the accuracy of the results obtained and the simplicity of calculation. The demerits of the method of successive approximation are a comparatively slow convergence of the approximations to the solution and complexity of the computations. A drawback in Euler's method is its low accuracy. To increase the accuracy one has to take an extremely small interval of computation h, which leads to prolonged computation.

Incidentally, a slight improvement in Euler's method via so-called iteration, yields a rather convenient computational procedure. When applying Euler's method with iteration, one divides the interval $x_0 \leqslant x \leqslant b$, on which the solution of the equation $\frac{dy}{dx} = f(x, y)$ defined by the condition $y(x_0) = y_0$ is to be computed, into equal subintervals of length $h = \frac{b - x_0}{n}$. Denoting $x_0 + kh = x_k$, $y(x_0 + kh) = y_k$, $y'(x_0 + kh) = y'_k$, we compute y_{k+1} (if y_k has already been found), first using Euler's formula:

$$y_{k+1} = y_k + hy'_k \qquad \text{or} \qquad \Delta y_k = y_{k+1} - y_k = hy'_k, \qquad (1.39)$$

i.e. on the interval $x_0 + kh \leqslant x \leqslant x_0 + (k+1)h$ the integral curve passing through the point (x_k, y_k) is replaced by a segment of its tangent at this point (see Fig. 1.13 on page 45). Then the computed value of y_{k+1} is refined, for which purpose the derivative $y'_{k+1} = f(x_{k+1}, y_{k+1})$ is determined and Euler's formula (1.39) is again employed, but in place of y'_k one takes the arithmetic mean of the computed values of the derivatives at the boundary points $\frac{y'_k + y'_{k+1}}{2}$, i.e. one takes

$$\bar{y}_{k+1} = y_k + h\,\frac{y'_k + y'_{k+1}}{2}. \qquad (1.40)$$

The just computed value of $\bar{y}_{k+1}$ permits computing a new value of the derivative $\bar{y}'_{k+1} = f(x_{k+1}, \bar{y}_{k+1})$; then one again computes the arithmetic mean of the values of the derivatives $\frac{y'_k + \bar{y}'_{k+1}}{2}$, again applies formula (1.40)

$$\bar{\bar{y}}_{k+1} = y_k + h\,\frac{y'_k + \bar{y}'_{k+1}}{2}$$

and continues this process until, within the limits of the given accuracy, the results of two successive computations of the values of y_{k+1} coincide. The same method is then used to compute y_{k+2}, and so on.

Euler's method with iteration yields an error of the order of h^3 in each interval and is frequently employed in computational work. However, more precise methods (those of Störmer, Runge, Miln and others) are more often used. They are based on replacing the desired solution by several terms of its Taylor expansion

$$y_{k+1} \approx y_k + hy'_k + \frac{h^2}{2!} y''_k + \dots + \frac{h^n}{n!} y_k^{(n)}, \tag{1.41}$$

i.e., on replacing the desired integral curve by a parabola of order n having nth order tangency with the integral curve at the point $x = x_k$, $y = y_k$.

Direct application of the Taylor formula (1.41) on every interval leads to involved and diversified computations, and so this formula is ordinarily used only when calculating a few values close to $x = x_0$ that are needed for applying more convenient computational schemes, among which *Störmer's method* should be mentioned first. Here, the computation is carried out by one of the following formulas depending on the order of the approximating parabola:

$$y_{k+1} = y_k + q_k + \frac{1}{2}\Delta q_{k-1}, \tag{1.42}$$

$$y_{k+1} = y_k + q_k + \frac{1}{2}\Delta q_{k-1} + \frac{5}{12}\Delta^2 q_{k-2}, \tag{1.43}$$

$$y_{k+1} = y_k + q_k + \frac{1}{2}\Delta q_{k-1} + \frac{5}{12}\Delta^2 q_{k-2} + \frac{3}{8}\Delta^3 q_{k-3}, \tag{1.44}$$

$$y_{k+1} = y_k + q_k + \frac{1}{2}\Delta q_{k-1} + \frac{5}{12}\Delta^2 q_{k-2} + \frac{3}{8}\Delta^3 q_{k-3} + \frac{251}{720}\Delta^4 q_{k-4}, \tag{1.45}$$

. . . .

where

$$q_k = y'_k h, \quad \Delta q_{k-1} = q_k - q_{k-1}, \quad \Delta^2 q_{k-2} = \Delta q_{k-1} - \Delta q_{k-2},$$
$$\Delta^3 q_{k-3} = \Delta^2 q_{k-2} - \Delta^2 q_{k-3}, \quad \Delta^4 q_{k-4} = \Delta^3 q_{k-3} - \Delta^3 q_{k-4}.$$

. . . .

Störmer's formulas may be obtained by integrating the identity $y' \equiv f(x, y(x))$ from x_k to x_{k+1} in which $y(x)$ is the desired solution:

$$y_{k+1} \equiv y_k + \int_{x_k}^{x_{k+1}} f(x, y(x))\, dx,$$

and using the quadrature formula from the course of mathematical analysis:

$$\int_{x_k}^{x_{k+1}} \varphi(x)\, dx \approx h\left[\varphi_k + \frac{1}{2}\Delta\varphi_{k-1} + \frac{5}{12}\Delta^2\varphi_{k-2} + \right.$$
$$\left. + \frac{3}{8}\Delta^3\varphi_{k-3} + \frac{251}{720}\Delta^4\varphi_{k-4} + \dots\right]. \tag{1.46}$$

It will be recalled that this quadrature formula is obtained by replacing the integrand $\varphi(x)$ by the approximating polynomial in

Newton's interpolation formula and by computing the integrals of separate terms.

An estimate of the remainder term of the quadrature formula (1.46) shows that the error in one interval of calculation is of the order of h^3 in (1.42), h^4 in (1.43), h^5 in (1.44) and h^6 in the formula (1.45). Now if we take into account that the errors over several intervals may be cumulative, we have to multiply the evaluations obtained for one interval by $n=\frac{b-x_0}{h}$ in order to estimate the error for n intervals of calculation. This can lead to a change in the above-indicated order of error.

Note. It may be demonstrated by direct expansion of the Taylor formula in the neighbourhood of the point $x=x_k$ that the right-hand side of Störmer's formula (1.42) coincides, to within terms containing h to powers above the second, with the first three terms of the expansion of y_{k+1} in the Taylor series (1.41):

$$y_k+hy'_k+\frac{h^2}{2!}y''_k; \tag{1.47}$$

the right side of the next of the Störmer formulas, (1.43), coincides, to within terms containing h to powers above the third, with

$$y_k+hy'_k+\frac{h^2}{2!}y''_k+\frac{h^3}{3!}y'''_k$$

and so forth. For example, for formula (1.42) we get

$$y_k+hy'_k+\frac{1}{2}h\,\Delta y'_{k-1}=y_k+hy'_k+\frac{1}{2}h\,(y'_k-y'_{k-1}), \tag{1.48}$$

or, expanding

$$y'_{k-1}=y'(x_{k-1})$$

in the Taylor series

$$y'(x_{k-1})=y'(x_k)-hy''(x_k)+\frac{1}{2}h^2y'''(x_k)+\dots$$

and substituting into (1.48), we obtain

$$y_k+hy'_k+\frac{1}{2}h\,(y'_k-y'_{k-1})=y_k+hy'_k+\frac{1}{2}h^2y''_k-\frac{1}{4}h^3y'''_k+\dots,$$

and consequently the first three terms coincide with three terms of the Taylor expansion (1.47).

To begin calculations by Störmer's formulas, we have to know the values of the desired function at several points, not one [using formula (1.42) at two points: x_0 and x_0+h, using (1.43) at three points: x_0, x_0+h, and x_0+2h, and so on]. These first few values may be computed by Euler's method with reduced interval or by using Taylor's formula (1.41), or by the *Runge method* that is briefly described below.

For definiteness take the formula (1.44):

$$y_{k+1}=y_k+q_k+\frac{1}{2}\Delta q_{k-1}+\frac{5}{12}\Delta^2 q_{k-2}+\frac{3}{8}\Delta^3 q_{k-3}$$

and assume that in addition to the given initial value y_0 we have already found y_1, y_2 and y_3. Then we can calculate:

$$q_0=f(x_0, y_0)h,\ q_1=f(x_1, y_1)h,$$
$$q_2=f(x_2, y_2)h,\ q_3=f(x_3, y_3)h,$$

and hence also

$$\Delta q_0=q_1-q_0,\ \Delta q_1=q_2-q_1,\ \Delta q_2=q_3-q_2,$$
$$\Delta^2 q_0=\Delta q_1-\Delta q_0,\ \Delta^2 q_1=\Delta q_2-\Delta q_1,\ \Delta^3 q_0=\Delta^2 q_1-\Delta^2 q_0.$$

Now, using formula (1.44), we compute the value of y_4; with this value we can obtain q_4, Δq_3, $\Delta^2 q_2$, $\Delta^3 q_1$. Then, using (1.44) again, we compute y_5 and so on.

The results of the computation are entered in a table (see below) that gradually fills in.

x	y	q	Δq	$\Delta^2 q$	$\Delta^3 q$
x_0	y_0	q_0	Δq_0	$\Delta^2 q_0$	$\Delta^3 q_0$
x_1	y_1	q_1	Δq_1	$\Delta^2 q_1$	
x_2	y_2	q_2	Δq_2		
x_3	y_3	q_3			
x_4					
x_5					
x_6					

Ordinarily it is required to compute the value of the desired solution of a differential equation at some point $x=b$ with a given accuracy. The question immediately arises: which of Störmer's formulas is most suitable and what interval h guarantees the desired accuracy and at the same time is not too small (this would lead to extra computation). Some idea of the choice of formula to be used in the computation and of the choice of interval is given by the above-indicated orders of error in each interval of calculation. One must also bear in mind that errors may be cumulative over a number of intervals. Given a proper choice of interval h, all differences in the table should vary smoothly, and the last differences in formula (1.44) should only affect the reserve digits. A sharp change of some difference indicates that on the chosen interval h there may be neglected peculiarities in the variation of the function on the interval under consideration, and this may lead to appreciable errors in the calculation of y_{k+1}.

However, all these arguments are not exactly reliable, and more exact evaluations of the error may turn out to be extremely cumbersome and inconvenient. For this reason, the following rather reliable method is employed: starting out with the above inexact reasoning, one chooses some interval h, carries out the computations by one of Störmer's formulas with interval h and $\frac{h}{2}$ and compares the results at common points. If the results coincide to within the given accuracy, then it is decided that the interval h ensures the required accuracy of computation; but if the results do not coincide within the limits of the given accuracy, the interval h is reduced by a factor of 2 again and the computations are carried out with the interval $\frac{h}{2}$ and $\frac{h}{4}$ and the results are again compared, etc.

It is advisable to carry out the computations with interval h and $\frac{h}{2}$ in parallel so as to be able to detect any noncoincidence of results as early as possible and avoid extra work. This method of double computation has the further advantage that it almost completely eliminates errors of calculation, since as a rule they are revealed immediately when the results of computing with intervals h and $\frac{h}{2}$ are compared.

For finding the first few values of y_i that are necessary to begin computations by Störmer's method, one can recommend the Runge method in addition to the methods we have already mentioned (Euler's method with reduced interval with or without iterations, or the method of expansion in a Taylor's series).

In the Runge method, one has to compute four numbers to find y_{k+1}:

$$m_1=f(x_k, y_k), \quad m_2=f\left(x_k+\frac{h}{2}, y_k+\frac{m_1h}{2}\right),$$

$$m_3=f\left(x_k+\frac{h}{2}, y_k+\frac{m_2h}{2}\right), \quad m_4=f(x_k+h, y_k+m_3h), \tag{1.49}$$

and then

$$y_{k+1}=y_k+\frac{h}{6}(m_1+2m_2+2m_3+m_4). \qquad (1.50)$$

Ordinarily, the Runge method is used only to compute the first few values $y_1, y_2, \ldots$, that are needed to begin computations by the Störmer method; however, this method may be used to compute the remaining values as well. Runge's method, like that of Störmer, is based on an approximation of the desired integral curve by a tangent parabola.

If we compare the right side of Runge's formula (1.50) with Taylor's expansion

$$y_{k+1}=y_k+y_k'h+\frac{1}{2!}y_k''h^2+\frac{1}{3!}y_k'''h^3+\frac{1}{4!}y_k^{IV}h^4+\ldots,$$

it will be seen that terms with powers lower than the fifth coincide. For this reason, when computing several initial values by the Runge method, in order to transfer to computations by the Störmer method using formulas (1.42), (1.43) and (1.44), one can use the same interval h; but if Störmer's formula (1.45) is used subsequently, the initial computations by the Runge method must be performed with a reduced interval of calculation, since for one and the same interval of calculation, formula (1.50) does not guarantee the same accuracy as formula (1.45 does. True, rather often the initial computations are performed by the Runge formulas with reduced interval even when employing Störmer's formulas (1.43) and (1.44), since even a slight error in computing the initial values for the Störmer formulas can drastically reduce the accuracy of subsequent computations. *

Modern digital computers perform the above-described computations by Störmer's or Runge's method very rapidly (up to hundreds of thousands of operations per second), and the programming procedure can be substantially simplified by using standard programmes developed for the methods of Störmer and Runge. Here, in the approximate integration of the differential equation $y'=f(x, y)$, $y(x_0)=y_0$, one has only to compile a subprogramme for computing the values of $y_k'=f(x_k, y_k)$ and to include it in the standard programme.

Example.

$$y'=x^2+y^2; \quad y(0)=-1.$$

Find the value $y(0.5)$ to within 0.01.

* A more detailed description of approximate methods of integrating differential equations is given by A. Krylov [6] and I. Berezin and N. Zhidkov [7].

Taking advantage of the Taylor expansion

$$y(x)=y(0)+y'(0)x+\frac{y''(0)x^2}{2!}+\frac{y'''(0)x^3}{3!}+\ldots,$$

we compute the value $y(x)$ at the points $x_1=0.1$ and $x_2=0.2$:

$$y(0.1)=-0.9088 \text{ and } y(0.2)=-0.839$$

(or in place of $y(0.2)$ we compute $y(-0.1)$, which is even to be preferred, since the point $x_1=-0.1$ lies closer to the initial point $x_0=0$ than does the point $x_2=0.2$). The subsequent values are computed by Störmer's formula (1.43) with interval $h=0.1$, and the results are entered in the table (without the differences $\Delta^3 q$). After that, or at the same time, we perform the computations with interval $\frac{h}{2}=0.05$. This yields

$$y(0.5)\approx -0.63.$$

8. Elementary Types of Equations Not Solved for the Derivative

A first-order differential equation not solved for the derivative is of the form

$$F(x, y, y')=0. \tag{1.51}$$

If it is possible to solve this equation for y', we get one or several equations

$$y'=f_i(x, y) \qquad (i=1, 2, \ldots).$$

Integrating these equations that have already been solved for the derivative, we find the solutions of the original equation (1.51).

By way of an example, let us integrate the equation

$$(y')^2-(x+y)y'+xy=0. \tag{1.52}$$

Solving this quadratic equation for y', we get $y'=x$ and $y'=y$. Integrating each one of the equations obtained, we find

$$y=\frac{x^2}{2}+c \tag{1.53}$$

and

$$y=ce^x \tag{1.54}$$

(Fig. 1.24). Both families of solutions satisfy the original equation.

Also, the curves composed of an arc of the integral curve of the family (1.53) and an arc of the integral curve of the family (1.54) will be smooth integral curves of the equation (1.52) if they have a common tangent at a common point. Fig. 1.25 depicts an integral curve of equation (1.52) composed of graphs of the solutions $y=\frac{x^2}{2}+c$ for $c=\frac{1}{2}$, $-\infty<x\leqslant 1$ and $y=ce^x$ for $c=e^{-1}$, $1\leqslant x<\infty$, and Fig. 1.26 depicts an integral curve of equation (1.52), which is composed of the graphs of the solutions $y=\frac{x^2}{2}$ for $x\leqslant 0$ and $y\equiv 0$ for $x>0$.

Thus, equation

$$F(x, y, y')=0 \qquad (1.51)$$

may be integrated by solving for y' and by integration of the thus obtained equations $y'=f_i(x, y)$ $(i=1, 2, \ldots)$ that have already been solved for the derivative.

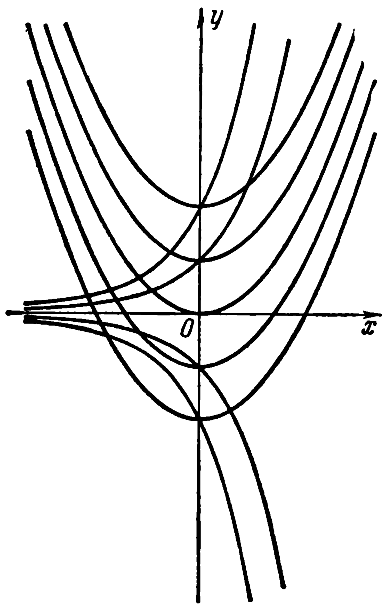


Fig. 1-24

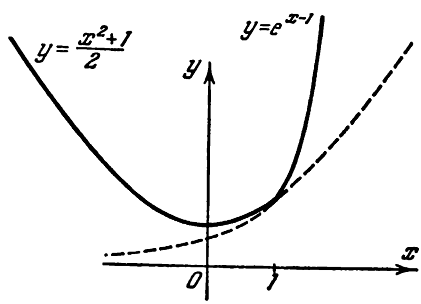


Fig. 1-25

However, equation (1.51) is not always so readily solved for y', and even more infrequently are the equations $y'=f_i(x, y)$ obtained after solving for y' easily integrated. Therefore, one often has to integrate equations of the form (1.51) by other methods. Let us consider a number of cases.

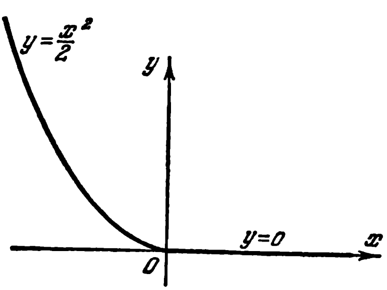


Fig. 1-26

1. *Equation* (1.51) *is of the form*

$$F(y')=0, \qquad (1.55)$$

and there exists at least one real root $y'=k_i$ of this equation.

Since (1.55) does not contain x and y, k_i is a constant. Consequently, integrating the equation $y'=k_i$, we get $y=k_ix+c$ or

$k_i = \frac{y-c}{x}$, but k_i is a root of the equation (1.55); consequently, $F\left(\frac{y-c}{x}\right)=0$ is an integral of the equation in question.

Example 1.

$$(y')^7-(y')^5+y'+3=0.$$

The integral of the equation is

$$\left(\frac{y-c}{x}\right)^7-\left(\frac{y-c}{x}\right)^5+\frac{y-c}{x}+3=0.$$

2. *Equation* (1.51) *is of the form*

$$F(x, y')=0. \tag{1.56}$$

If this equation is hard to solve for y', it is advisable to introduce the parameter t and replace (1.56) by two equations: $x=\varphi(t)$ and $y'=\psi(t)$. Since $dy=y'dx$, in the given case $dy=\psi(t)\varphi'(t)\,dt$, whence $y=\int \psi(t)\varphi'(t)\,dt+c$ and, hence, the integral curves of equation (1.56) are determined in parametric form by the following equations:

$$x=\varphi(t),$$
$$y=\int \psi(t)\varphi'(t)\,dt+c.$$

If equation (1.56) is readily solvable for x, $x=\varphi(y')$, then it is nearly always convenient to introduce $y'=t$ as parameter. Then

$$x=\varphi(t),\quad dy=y'dx=t\varphi'(t)\,dt,\quad y=\int t\varphi'(t)\,dt+c.$$

Example 2.

$$x=(y')^3-y'-1.$$

Put $y'=t$, then

$$x=t^3-t-1, \tag{1.57}$$
$$dy=y'dx=t(3t^2-1)\,dt,$$
$$y=\frac{3t^4}{4}-\frac{t^2}{2}+c_1. \tag{1.58}$$

Equations (1.57) and (1.58) define in parametric form a family of the desired integral curves.

Example 3.

$$x\sqrt{1+y'^2}=y'.$$

Put $y'=\tan t$, $-\frac{\pi}{2}<t<\frac{\pi}{2}$; then

$$x=\sin t, \tag{1.59}$$
$$dy=y'dx=\tan t\,\cos t\,dt=\sin t\,dt,$$
$$y=-\cos t+c_1 \tag{1.60}$$

or, eliminating t from equations (1.59) and (1.60), we get $x^2 + (y-c_1)^2 = 1$, which is a family of circles.

3. *The equation* (1.51) *is of the form*

$$F(y, y') = 0. \tag{1.61}$$

If it is difficult to solve this equation for y', then, as in the preceding case, it is advisable to introduce the parameter t and replace (1.61) by two equations: $y = \varphi(t)$ and $y' = \psi(t)$. Since $dy = y'dx$, then $dx = \frac{dy}{y'} = \frac{\varphi'(t)\,dt}{\psi(t)}$, whence $x = \int \frac{\varphi'(t)\,dt}{\psi(t)} + c$. Thus, in parametric form the desired integral curves are defined by the equations

$$x = \int \frac{\varphi'(t)\,dt}{\psi(t)} + c \quad \text{and} \quad y = \varphi(t).$$

As a particular case, if equation (1.61) is readily solvable for y, it is usually convenient to take y' as the parameter.

Indeed, if $y = \varphi(y')$, then, putting $y' = t$, we get

$$y = \varphi(t),\ dx = \frac{dy}{y'} = \frac{\varphi'(t)\,dt}{t},$$

$$x = \int \frac{\varphi'(t)\,dt}{t} + c.$$

Example 4.

$$y = (y')^5 + (y')^3 + y' + 5.$$

Put $y' = t$; then

$$y = t^5 + t^3 + t + 5, \tag{1.62}$$

$$dx = \frac{dy}{y'} = \frac{(5t^4 + 3t^2 + 1).dt}{t} = \left(5t^3 + 3t + \frac{1}{t}\right)dt,$$

$$x = \frac{5t^4}{4} + \frac{3t^2}{2} + \ln|t| + c. \tag{1.63}$$

Equations (1.62) and (1.63) are parametric equations of a family of integral curves.

Example 5.

$$\frac{y}{\sqrt{1 + y'^2}} = 1$$

Put $y' = \sinh t$; then

$$y = \cosh t \tag{1.64}$$

$$dx = \frac{dy}{y'} = \frac{\sinh t\,dt}{\sinh t} = dt,$$

$$x = t + c \tag{1.65}$$

or, eliminating the parameter t from (1.64) and (1.65), we get $y = \cosh(x - c)$.

Now consider the general case: the left-hand side of equation

$$F(x, y, y')=0 \tag{1.51}$$

depends on all three arguments x, y, y'. Replace (1.51) by its parametric representation:

$$x=\varphi(u, v), \quad y=\psi(u, v), \quad y'=\chi(u, v).$$

Taking advantage of the dependence $dy=y'dx$, we will have

$$\frac{\partial\psi}{\partial u}du+\frac{\partial\psi}{\partial v}dv=\chi(u, v)\left[\frac{\partial\varphi}{\partial u}du+\frac{\partial\varphi}{\partial v}dv\right];$$

from this, solving for the derivative $\frac{dv}{du}$, we get

$$\frac{dv}{du}=\frac{\chi(u, v)\frac{\partial\varphi}{\partial u}-\frac{\partial\psi}{\partial u}}{\frac{\partial\psi}{\partial v}-\chi(u, v)\frac{\partial\varphi}{\partial v}}. \tag{1.66}$$

We have thus obtained a first-order equation that is already *solved for the derivative*, and we have thus reduced the problem to one that has been considered in earlier sections. However, the resulting equation (1.66) is of course by far not always integrable by quadratures.

If the equation

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

is readily solvable for y, it is often convenient to take x and y' for the parameters u and v. Indeed, if equation (1.51) is reduced to the form

$$y=f(x, y'), \tag{1.67}$$

then, considering x and $y'=p$ as parameters, we obtain

$$y=f(x, p), \quad dy=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial p}dp$$

or

$$\frac{dy}{dx}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial p}\frac{dp}{dx},$$

$$p=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial p}\frac{dp}{dx}. \tag{1.68}$$

Integrating (1.68) (of course, it is by far not always integrable by quadratures), we get $\Phi(x, p, c)=0$. The collection of equations $\Phi(x, p, c)=0$ and $y=f(x, p)$, where p is a parameter, defines a family of integral curves.

Note that (1.68) may be obtained by differentiating (1.67) with respect to x. Indeed, differentiating (1.67) with respect to x and assuming $y'=p$, we get $p=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial p}\frac{dp}{dx}$, which coincides with (1.68). For this reason, this method is often called integration of differential equations by means of differentiation.

In quite analogous fashion, the equation

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

can frequently be integrated if it is readily solvable for x:

$$x=f(y, y') \tag{1.69}$$

In this case, taking y and $y'=p$ for the parameters and using the relation $dy=y'dx$, we get

$$dy=p\left[\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial p}dp\right]$$

or

$$\frac{1}{p}=\frac{\partial f}{\partial y}+\frac{\partial f}{\partial p}\frac{dp}{dy}. \tag{1.70}$$

Integrating equation (1.70), we get $\Phi(y, p, c)=0$. This equation together with $x=f(y, p)$ defines the integral curves of the original equation. Equation (1.70) may be obtained from (1.69) by differentiation with respect to y.

To illustrate the use of this method, let us consider the following equation, which is linear in x and y:

$$y=x\varphi(y')+\psi(y'),$$

it is called *Lagrange's equation*. Differentiating with respect to x and putting $y'=p$, we have

$$p=\varphi(p)+x\varphi'(p)\frac{dp}{dx}+\psi'(p)\frac{dp}{dx}, \tag{1.71}$$

or

$$[p-\varphi(p)]\frac{dx}{dp}=x\varphi'(p)+\psi'(p). \tag{1.72}$$

This equation is linear in x and $\frac{dx}{dp}$ and, hence, is readily integrable, for example, by the method of variation of parameters. After obtaining the integral $\Phi(x, p, c)=0$ of equation (1.72) and adjoining to it $y=x\varphi(p)+\psi(p)$, we get equations that define the desired integral curves.

When passing from (1.71) to (1.72) we had to divide by $\frac{dp}{dx}$. But in the process we lose solutions (if they exist) for which p is constant, and hence $\frac{dp}{dx}\equiv 0$. Taking p constant, we note that equation (1.71) is satisfied only if p is a root of the equation $p-\varphi(p)=0$.

Thus, if the equation $p-\varphi(p)=0$ has real roots $p=p_i$, then to the solutions of the Lagrange equation that were found above we have to add $y=x\varphi(p)+\psi(p)$, $p=p_i$ or, eliminating p, $y=x\varphi(p_i)+$ $+\psi(p_i)$ are straight lines.

We have to consider separately the case when $p-\varphi(p)\equiv 0$ and hence, when dividing by $\frac{dp}{dx}$, we lose the solution $p=c$, where c is an arbitrary constant. In this case, $\varphi(y')\equiv y'$ and the equation $y=x\varphi(y')+\psi(y')$ takes the form $y=xy'+\psi(y')$, which is *Clairaut's equation*. Putting $y'=p$, we have $y=xp+\psi(p)$. Differentiating with respect to x, we will have

$$p=p+x\frac{dp}{dx}+\psi'(p)\frac{dp}{dx}$$

or

$$(x+\psi'(p))\frac{dp}{dx}=0,$$

whence either $\frac{dp}{dx}=0$, and hence $p=c$, or $x+\psi'(p)=0$.

In the first case, eliminating p, we get

$$y=cx+\psi(c), \tag{1.73}$$

which is a one-parameter family of integral curves. In the second case, the solution is defined by the equations

$$y=xp+\psi(p) \quad \text{and} \quad x+\psi'(p)=0. \tag{1.74}$$

It can readily be verified that the integral curve defined by the equations (1.74) is the envelope of the family of integral curves (1.73).

Indeed, the *envelope* of a family $\Phi(x, y, c)=0$ is defined by the equations

$$\Phi(x, y, c)=0 \quad \text{and} \quad \frac{\partial\Phi}{\partial c}=0, \tag{1.75}$$

which for the family $y=cx+\psi(c)$ are of the form

$$y=cx+\psi(c), \quad x+\psi'(c)=0$$

and differ from equations (1.74) (Fig. 1.27) only in the designation of the parameter.

Note. As we know, besides the envelope, the equations (1.75) can define the loci of multiple points and sometimes other curves as well; however, if even one of the derivatives $\frac{\partial\Phi}{\partial x}$ and $\frac{\partial\Phi}{\partial y}$ is different from zero and both are bounded at points satisfying the equations (1.75), then these equations define only the envelope. In the given case, these conditions are fulfilled: $\frac{\partial\Phi}{\partial x}=-c$, $\frac{\partial\Phi}{\partial y}=1$. Consequently, equations (1.75) define an envelope that can degenerate into a point if the family (1.73) is a pencil of lines.

Example 6.

$$y=xy'-y'^2 \text{ is Clairaut's equation.}$$

A one-parameter family of integral straight lines has the form

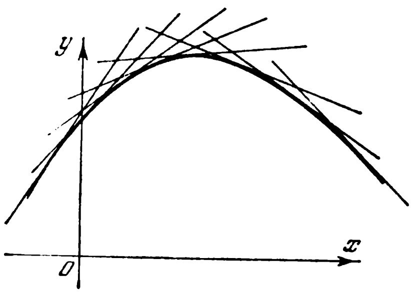


Fig. 1-27

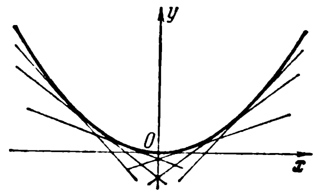


Fig. 1-28

$y=cx-c^2$. Besides, the envelope of this family, defined by the equations $y=cx-c^2$ and $x-2c=0$, is an integral curve. Eliminating c, we get $y=\frac{x^2}{4}$ (Fig. 1.28).

Example 7.

$$y=2xy'-y'^3 \text{ is Lagrange's equation.}$$

$$y'=p,$$

$$y=2xp-p^3. \tag{1.76}$$

Differentiating, we get

$$p=2p+2x\frac{dp}{dx}-3p^2\frac{dp}{dx} \tag{1.77}$$

and, after dividing by $\frac{dp}{dx}$, we arrive at the equation

$$p\frac{dx}{dp}=-2x+3p^2.$$

Integrating this linear equation, we obtain $x=\frac{c_1}{p^2}+\frac{3}{4}p^2$. Hence, the integral curves are defined by the equations $y=2xp-p^3$, $x=\frac{c_1}{p^2}+\frac{3p^2}{4}$.

As mentioned above, in dividing by $\frac{dp}{dx}$ we lose the solutions $p=p_i$, where the p_i are roots of the equation $p-\varphi(p)=0$. Here, we lose the solution $p=0$ of equation (1.77), to which corresponds, by virtue of equation (1.76), the solution of the original equation $y=0$.

9. The Existence and Uniqueness Theorem for Differential Equations Not Solved for the Derivative. Singular Solutions

In Sec. 6 we proved the theorem of existence and uniqueness of the solution $y(x)$ of the equation $\frac{dy}{dx}=f(x, y)$ that satisfies the condition $y(x_0)=y_0$. A similar question arises for equations of the form $F(x, y, y')=0$. It is obvious that for such equations, generally speaking, not one but several integral curves pass through some point (x_0, y_0), since, as a rule, when solving the equation $F(x, y, y')=0$ for y' we get several (not one) real values $y'=f_i(x, y)$ $(i=1, 2, \ldots)$, and if each of the equations $y'=f_i(x, y)$ in the neighbourhood of the point (x_0, y_0) satisfies the conditions of the existence and uniqueness theorem of Sec. 6, then for each one of these equations there will be a unique solution satisfying the condition $y(x_0)=y_0$. Therefore the property of uniqueness of solution of the equation $F(x, y, y')=0$, which satisfies the condition $y(x_0)=y_0$, is usually understood in the sense that not more than one integral curve of the equation $F(x, y, y')=0$ passes through a given point (x_0, y_0) in a given direction.

For example, for the solutions of the equation $\left(\frac{dy}{dx}\right)^2-1=0$, the property of uniqueness is everywhere fulfilled, since through every point (x_0, y_0) there pass two integral curves, but in different directions. Indeed,

$$\frac{dy}{dx}=\pm 1, \quad y=x+c \quad \text{and} \quad y=-x+c.$$

For equation $(y')^2-(x+y)y'+xy=0$ considered on page 73, the property of uniqueness is violated at points of the straight line $y=x$, since the integral curves of the equations $y'=x$ and $y'=y$ pass through points of this line in the same direction (Fig. 1.29).

Theorem 1.5. *There exists a unique solution $y=y(x)$, $x_0-h_0\leqslant x\leqslant x_0+h_0$ (where h_0 is sufficiently small) of the equation*

$$F(x, y, y')=0, \tag{1.78}$$

that satisfies the condition $y(x_0)=y_0$ for which $y'(x_0)=y_0'$, where y_0' is one of the real roots of the equation $F(x_0, y_0, y')=0$, if in a closed neighbourhood of the point (x_0, y_0, y_0') the function $F(x, y, y')$ satisfies the conditions:

(1) *$F(x, y, y')$ is continuous with respect to all arguments;*

(2) *the derivative $\frac{\partial F}{\partial y'}$ exists and is nonzero;*

(3) *there exists a derivative $\frac{\partial F}{\partial y}$ bounded in absolute value:*

$$\left|\frac{\partial F}{\partial y}\right| \leqslant N_1.$$

Proof. According to the familiar theorem on the existence of an implicit function, it may be asserted that conditions (1) and (2)

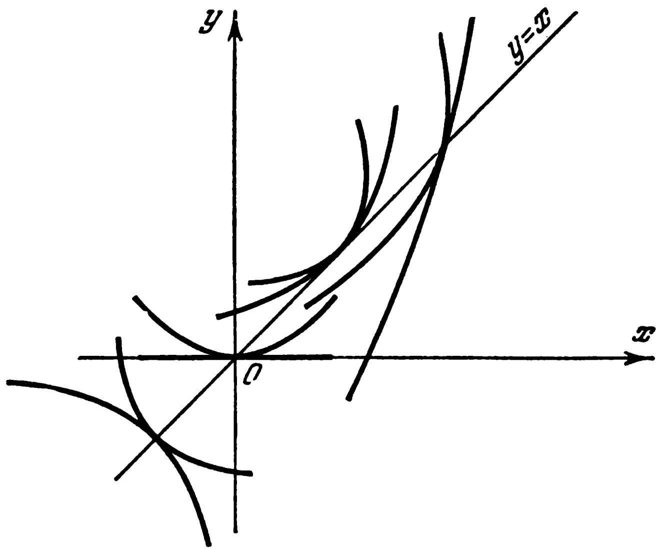


Fig. 1-29

guarantee the existence of a unique function $y' = f(x, y)$, continuous in the neighbourhood of the point (x_0, y_0), that is defined by the equation (1.78) and satisfies the condition $y_0' = f(x_0, y_0)$. It remains to verify whether the function $f(x, y)$ will satisfy the Lipschitz condition or the cruder condition $\left|\frac{\partial f}{\partial y}\right| \leqslant N$ in the neighbourhood of the point (x_0, y_0), for then it will be possible to assert that the equation

$$y' = f(x, y) \tag{1.79}$$

satisfies the conditions of the existence and uniqueness theorem (see Sec. 6, page 46) and, consequently, that there exists a unique solution of the equation (1.79) that satisfies the condition $y(x_0) = y_0$, and there also exists a unique integral curve of the equation (1.78) that passes through the point (x_0, y_0) and has, at that point, the slope of the tangent y'_0.

In accordance with the familiar theorem on implicit functions, it may be asserted that if conditions (1), (2) and (3) are fulfilled, the derivative $\frac{\partial f}{\partial y}$ exists and may be computed by the rule of differentiating implicit functions.

Differentiating the identity $F(x, y, y') = 0$ with respect to y and taking into account that $y' = f(x, y)$, we get

$$\frac{\partial F}{\partial y} + \frac{\partial F}{\partial y'} \frac{\partial f}{\partial y} = 0,$$

$$\frac{\partial f}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial y'}},$$

whence, taking into account the conditions (2) and (3), it follows that $\left| \frac{\partial f}{\partial y} \right| \leqslant N$ in a closed neighbourhood of the point (x_0, y_0).

The set of points (x, y) at which the uniqueness of solutions of the equation

$$F(x, y, y') = 0, \tag{1.78}$$

is violated is called a *singular* set.

At the points of a singular set at least one of the conditions of Theorem 1.5 must be violated. In differential equations encountered in applied problems, conditions (1) and (3) are usually fulfilled but condition (2), $\frac{\partial F}{\partial y'} \neq 0$, is frequently violated.

If conditions (1) and (3) are fulfilled, then at the points of a singular set the equations

$$F(x, y, y') = 0 \quad \text{and} \quad \frac{\partial F}{\partial y'} = 0 \tag{1.80}$$

must be satisfied simultaneously.

Eliminating y' from these equations, we get the equation

$$\Phi(x, y) = 0, \tag{1.81}$$

which must be satisfied by the points of the singular set. However, the uniqueness of solution of the equation (1.78) is not necessarily violated at every point that satisfies the equation (1.81), because the conditions of Theorem 1.5 are only sufficient for uniqueness of solution, but are not necessary, and hence violation of a condition of the theorem does not of necessity imply violation of uniqueness.

Thus, only among points of the curve $\Phi(x, y)=0$, called the *p-discriminant* curve [since the equation (1.80) is most frequently written in the form $F(x, y, p)=0$ and $\frac{\partial F}{\partial p}=0$], can there be points of the singular set.

If some kind of branch $y=\varphi(x)$ of the curve $\Phi(x, y)=0$ belongs to the singular set and at the same time is an integral curve, it is called a *singular integral curve*, and the function $y=\varphi(x)$ is called a *singular solution*.

Thus, in order to find the singular solution of the equation

$$F(x, y, y')=0 \tag{1.78}$$

it is necessary to find the p-discriminant curve defined by the equations

$$F(x, y, p)=0, \ \frac{\partial F}{\partial p}=0,$$

to find out [by direct substitution into equation (1.78)] whether there are integral curves among the branches of the p-discriminant curve and, if there are such curves, to verify whether uniqueness is violated at the points of these curves or not. If uniqueness is violated, then such a branch of the p-discriminant curve is a singular integral curve.

Example 1. Does the Lagrange equation $y=2xy'-(y')^2$ have a singular solution?

Conditions (1) and (3) of the existence and uniqueness theorem are fulfilled. The p-discriminant curve is defined by the equations $y=2xp-p^2$, $2x-2p=0$ or, eliminating p, $y=x^2$. The parabola $y=x^2$ is not an integral curve since the function $y=x^2$ does not satisfy the original equation. There is no singular solution.

Example 2. Find a singular solution of the Lagrange equation

$$x-y=\frac{4}{9}(y')^2-\frac{8}{27}(y')^3. \tag{1.82}$$

Conditions (1) and (3) of the existence and uniqueness theorem are fulfilled. The p-discriminant curve is defined by the equations

$$x-y=\frac{4}{9}p^2-\frac{8}{27}p^3, \quad \frac{8}{9}(p-p^2)=0.$$

From the second equation we find $p=0$ or $p=1$; substituting into the first equation, we obtain

$$y=x \text{ or } y=x-\frac{4}{27}.$$

Only the second of these functions is a solution of the original equation.

To find out whether the solution $y=x-\frac{4}{27}$ is singular, we have to integrate the equation (1.82) and find out whether other integral curves pass through points of the straight line $y=x-\frac{4}{27}$ in the direction of this line. Integrating the Lagrange equation (1.82), we get

$$(y-c)^2=(x-c)^3. \quad (1.83)$$

From (1.83) and Fig. 1.30 it is seen that the straight line $y=x-\frac{4}{27}$ is the envelope of a family of semicubical parabolas $(y-c)^2=(x-c)^3$ and, hence, uniqueness is violated at every point of the straight line $y=x-\frac{4}{27}$: there are two integral curves in one and the same direction—the straight line $y=x-\frac{4}{27}$ and a semicubical parabola that is tangent to this straight line at the point under consideration.

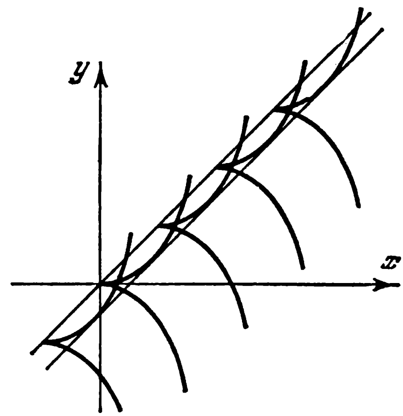


Fig. 1-30

Thus, $y=x-\frac{4}{27}$ is a singular solution.

In this example, the envelope of the family of integral curves is a singular solution.

If the envelope of the family

$$\Phi(x, y, c)=0 \quad (1.84)$$

is a curve which is tangent at each of its points to some curve of the family (1.84), and to each segment of which are tangent an infinite set of curves of this family, then the envelope of a family of integral curves of some equation $F(x, y, y')=0$ will always be a singular integral curve.

Indeed, at points of the envelope the values x, y and y' coincide with the values x, y and y' for the integral curve tangent to the envelope at the point (x, y), and hence at every point of the envelope the values x, y and y' satisfy the equation $F(x, y, y')=0$; that is, the envelope is an integral curve (Fig. 1.31). Uniqueness is violated at every point of the envelope, since at least two integral curves pass through the points of the envelope in the same direction: the envelope and the integral curve of the family (1.84) tangent to it at the point under consideration. Consequently, the envelope is a singular integral curve.

Knowing the family of integral curves $\Phi(x, y, c)=0$ of some differential equation $F(x, y, y')=0$, it is possible to determine its singular solutions by finding the envelope. As we know from the course of differential geometry or mathematical analysis, the envelope is contained in the c-discriminant curve defined by the equations

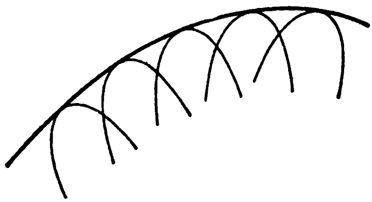


Fig. 1-31

$$\Phi(x, y, c)=0 \quad \text{and} \quad \frac{\partial \Phi}{\partial c}=0,$$

however, the c-discriminant curve can include, besides the envelope, other sets as well, for instance, the set of multiple points of curves of the family under study in which $\frac{\partial \Phi}{\partial x}=\frac{\partial \Phi}{\partial y}=0$. For some branch of the c-discriminant curve definitely to be an envelope, it is sufficient that on it:

(1) there exist the following partial derivatives bounded in absolute value:

$$\left|\frac{\partial \Phi}{\partial x}\right| \leqslant N_1, \quad \left|\frac{\partial \Phi}{\partial y}\right| \leqslant N_2;$$

(2) $\frac{\partial \Phi}{\partial x} \neq 0$ or $\frac{\partial \Phi}{\partial y} \neq 0$.

Note that these conditions are only sufficient, and so curves involving a violation of one of the conditions (1) or (2) can also be envelopes.

Example 3. Given a family of integral curves $(y-c)^2=(x-c)^3$ of some differential equation (see Example 2 on page 84). Find a singular solution of this equation.

Find the c-discriminant curve:

$$(y-c)^2=(x-c)^3 \quad \text{and} \quad 2(y-c)=3(x-c)^2.$$

Eliminating the parameter c, we get

$$y=x \quad \text{and} \quad x-y-\frac{4}{27}=0.$$

The straight line $y=x-\frac{4}{27}$ is an envelope since on it are fulfilled all the conditions of the envelope theorem. The function $y=x$ does not satisfy the differential equation. The straight line $y=x$

is a cusp locus (see Fig. 1.30). The second condition of the envelope theorem is violated at the points of this straight line.

Example 4. Given a family of integral curves

$$y^{\frac{1}{5}}-x+c=0 \tag{1.85}$$

of some differential equation of the first order. Find a singular solution of this equation.

The problem reduces to finding the envelope of the desired family. If one applies directly the above-indicated method for finding the envelope, one gets the contradictory equation $1=0$,

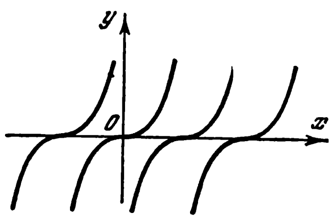


Fig. 1-32

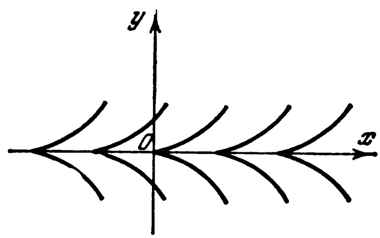


Fig. 1-33

whence it would seem natural to draw the conclusion that the family (1.85) does not have an envelope. However, in this case, the derivative of the left side of (1.85) with respect to y, $\frac{\partial \Phi}{\partial y}=$ $=\frac{1}{5}y^{-\frac{4}{5}}$ becomes infinite when $y=0$ and hence there is a possibility that $y=0$ is the envelope of the family (1.85) that could not be found by the general method since the conditions of the envelope theorem were violated on the straight line $y=0$.

One has to transform equation (1.85) so that the conditions of the envelope theorem are fulfilled for the transformed equation, which is equivalent to the original equation. For example, write (1.85) in the form $y-(x-c)^5=0$. The conditions of the envelope theorem are now fulfilled and, using the general method, we get

$$y=(x-c)^5, \quad 5(x-c)^4=0$$

or, eliminating c, we will have the equation of the envelope $y=$ $=0$ (Fig. 1.32).

Example 5. Given the family of integral curves

$$y^2-(x-c)^3=0 \tag{1.86}$$

of some first-order differential equation. Find a singular solution of this equation.

The c-discriminant curve is defined by the equations

$$y^2-(x-c)^3=0 \quad \text{and} \quad x-c=0$$

or, eliminating c, we get $y=0$. On the straight line $y=0$, both partial derivatives $\frac{\partial\Phi}{\partial x}$ and $\frac{\partial\Phi}{\partial y}$ of the left-hand side of equation (1.86) vanish, hence $y=0$ is the locus of multiple points of curves of the family (1.86), in the given case, the cusp locus. However, the cusp locus in this example is at the same time an envelope. Fig. 1.33 depicts the semicubical parabolas (1.86) and their envelope $y=0$.

PROBLEMS ON CHAPTER 1

1. $\tan y\,dx-\cot x\,dy=0$.
2. $(12x+5y-9)\,dx+(5x+2y-3)\,dy=0$.
3. $x\frac{dy}{dx}=y+\sqrt{x^2+y^2}$
4. $x\frac{dy}{dx}+y=x^3$.
5. $y\,dx-x\,dy=x^2y\,dy$.
6. $\frac{dx}{dt}+3x=e^{2t}$.
7. $y\sin x+y'\cos x=1$.
8. $y'=e^{x-y}$.
9. $\frac{dx}{dt}=x+\sin t$.
10. $x(\ln x-\ln y)\,dy-y\,dx=0$.
11. $xy(y')^2-(x^2+y^2)y'+xy=0$.
12. $(y')^2=9y^4$.
13. $\frac{dx}{dt}=e^{\frac{x}{t}}+\frac{x}{t}$.
14. $x^2+(y')^2=1$.
15. $y=xy'+\frac{1}{y}$.
16. $x=(y')^3-y'+2$.
17. $\frac{dy}{dx}=\frac{y}{x+y^3}$
18. $y=(y')^4-(y')^3-2$.

19. Find the orthogonal trajectories of the family $xy=c$; i.e., find lines that orthogonally intersect the curves of the indicated family.

20. Find the curve whose subtangent is twice the abscissa of the point of tangency.

21. Find the curve whose y-intercept cut by a tangent is equal to the abscissa of the point of tangency.

22. Find the orthogonal trajectories of the family

$$x^2+y^2=2ax.$$

23. Considering that the rate at which a body cools in the air is proportional to the difference between the temperature of the body and that of the air, solve the following problem: if the air temperature is 20° C and the body cools from 100° to 60° C in 20 min, how long will it take the body to reach 30° C?

24. A motor boat is in motion in calm water with a velocity of 10 km/hr. The motor is cut out and in $t=20$ sec the velocity falls to $v_1=6$ km/hr. Determine the speed of the boat 2 minutes after the motor was cut out (assume that the resistance of the water is proportional to the speed of the boat).

25. Find the shape of a mirror that reflects, parallel to a given direction, all the rays emanating from a specified point.

26. $y'^2+y^2=4$.

27. Find a curve whose tangent segment lying between the coordinate axes is divided into equal parts at the point of tangency.

28. $\frac{dy}{dx}=\frac{2y-x-4}{2x-y+5}$. 29. $\frac{dy}{dx}-\frac{y}{1+x}+y^2=0$.

30. Integrate the following equation numerically:

$$\frac{dy}{dx}=x+y^2, \quad y(0)=0.$$

Determine $y(0.5)$ to within 0.01.

31. Integrate numerically the equation

$$\frac{dy}{dx}=xy^3+x^2, \quad y(0)=0.$$

Determine $y(0.6)$ to within 0.01.

32. $y'=1.31x-0.2y^2$, $y(0)=2$.

Form a table of fifteen values of y with interval of computation $h=0.02$.

33. $y=2xy'-y'^2$. 34. $\frac{dy}{dx}=\cos(x-y)$.

35. Using the method of isoclines (see page 21), sketch a family of integral curves of the equation

$$\frac{dy}{dx}=x^2-y^2.$$

36. $(2x+2y-1)\,dx+(x+y-2)\,dy=0.$
37. $y'^3-y'e^{2x}=0.$
38. Find the orthogonal trajectories of the parabolas $y^2+2ax=a^2$.
39. Does the differential equation $y=5xy'-(y')^2$ have a singular solution?
40. Integrate in approximate fashion the equation

$$\frac{dy}{dx}=x-y^2, \quad y(1)=0$$

by the method of successive approximations (determine y_1 and y_2).

41. $y=x^2+\int_1^x \frac{y}{x}\,dx.$
42. Has the equation $y'=\sqrt[3]{x-5y}+2$ a singular solution?
43. $(x-y)\,y\,dx-x^2\,dy=0.$
44. Find the orthogonal trajectories of the family $y^2=cx^3$.
45. $\dot{x}+5x=10t+2$ for $t=1$, $x=2$.
46. $\dot{x}=\frac{x}{t}+\frac{x^2}{t^3}$ for $t=2$, $x=4$
47. $y=xy'+y'^2$ for $x=2$, $y=-1$.
48. $y=xy'+y'^2$ for $x=1$, $y=-1$.
49. $\frac{dy}{dx}=\frac{3x-4y-2}{3x-4y-3}.$
50. $\dot{x}-x\cot t=4\sin t.$
51. $y=x^2+2y'x+\frac{y'^2}{2}.$
52. $y'-\frac{3y}{x}+x^3y^2=0.$
53. $y(1+y'^2)=a.$
54. $(x^2-y)\,dx+(x^2y^2+x)\,dy=0.$
55. Find the integrating factor of the equation

$$(3y^2-x)\,dx+2y(y^2-3x)\,dy=0$$

having the form $\mu=\mu(x+y^2)$.

56. $(x-y)\,y\,dx-x^2\,dy=0.$
57. $y'=\frac{x+y-3}{1-x+y}.$
58. $xy'-y^2\ln x+y=0.$
59. $(x^2-1)\,y'+2xy-\cos x=0.$
60. $(4y+2x+3)\,y'-2y-x-1=0.$
61. $(y^2-x)\,y'-y+x^2=0.$
62. $(y^2-x^2)\,y'+2xy=0.$
63. $3xy^2y'+y^3-2x=0.$
64. $(y')^2+(x+a)\,y'-y=0$, where a is constant.
65. $(y')^2-2xy'+y=0.$
66. $(y')^2+2yy'\cot x-y^2=0.$

Differential equations of the second order and higher

1. The Existence and Uniqueness Theorem for an nth Order Differential Equation

Differential equations of the nth order are of the form

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}) \tag{2.1}$$

or, if they are not solved for the highest derivative,

$$F(x, y, y', \ldots, y^{(n)}) = 0.$$

The theorem of existence and uniqueness for an nth order equation may readily be derived by reducing it to a system of equations for which the existence and uniqueness theorem has already been proved (see page 56).

Indeed, if in the equation $y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})$ we consider as unknown functions not only y but also $y' = y_1$, $y'' = y_2$, ..., $y^{(n-1)} = y_{n-1}$, then equation (2.1) is replaced by the system

$$\left.\begin{aligned} y' &= y_1, \\ y_1' &= y_2, \\ &\cdot \ \cdot \ \cdot \ \cdot \\ y_{n-2}' &= y_{n-1}, \\ y_{n-1}' &= f(x, y, y_1, \ldots, y_{n-1}), \end{aligned}\right\} \tag{2.2}$$

and we can then apply the theorem of the existence and uniqueness of solution of a system of equations (see page 56), according to which if the right sides of all equations of the system (2.2) are continuous in the region under consideration and if they satisfy the Lipschitz condition with respect to all arguments, except x, then there exists a unique solution of the system (2.2) that satisfies the conditions

$$y(x_0) = y_0, \ y_1(x_0) = y_{10}, \ \ldots, \ y_{n-1}(x_0) = y_{n-1,\,0}.$$

The right sides of the first $n-1$ equations (2.2) are continuous and satisfy not only the Lipschitz condition but even the cruder condition of the existence of bounded derivatives with respect to $y, y_1, y_2, \ldots, y_{n-1}$. Hence, the conditions of the existence and uniqueness theorem will be fulfilled if the right side of the last equation $y'_{n-1}=f(x, y, y_1, \ldots, y_{n-1})$ is continuous in the neighbourhood of the initial values and satisfies the Lipschitz condition with respect to all arguments, from the second onwards, or satisfies the cruder condition of the existence of bounded partial derivatives with respect to all arguments from the second onwards.

Thus, reverting to the earlier variables x and y, we finally get the following existence and uniqueness theorem.

Theorem 2.1. *There exists a unique solution of an nth order differential equation $y^{(n)}=f(x, y, y', \ldots, y^{(n-1)})$ that satisfies the conditions*

$$y(x_0)=y_0,\ y'(x_0)=y'_0,\ y''(x_0)=y''_0,\ \ldots,\ y^{(n-1)}(x_0)=y_0^{(n-1)},$$

if in the neighbourhood of the initial values $(x_0, y_0, y'_0, \ldots, y_0^{(n-1)})$ the function f is a continuous function of all its arguments and satisfies the Lipschitz condition with respect to all arguments from the second onwards.

The latter condition may be replaced by the cruder condition of existence in the same neighbourhood of bounded partial derivatives of order one of the function f with respect to all arguments from the second onwards.

The *general solution* of an nth order differential equation is the family of solutions consisting of all possible partial solutions. If the right side of the equation

$$y^{(n)}=f(x, y, y', y'', \ldots, y^{(n-1)}) \tag{2.1}$$

satisfies, in some range of the arguments, the conditions of the existence and uniqueness theorem, then the general solution of equation (2.1) depends on n parameters, which can, for example, be the initial values of the desired function and its derivatives y_0, y'_0, y''_0, $\ldots$, $y_0^{(n-1)}$. In particular, the general solution of the second-order equation $y''=f(x, y, y')$ depends on two parameters, for instance on y_0 and y'_0. Now if y_0 and y'_0 are fixed, that is, if the point (x_0, y_0) is given and also the direction of the tangent to the desired integral curve at this point is given, then a unique integral curve is defined by these conditions when the conditions of the existence and uniqueness theorem are fulfilled.

For example, in the equation of motion of a particle of mass in a straight line under the action of a force $f(t, x, \dot{x})$:

$$m\ddot{x}=f(t, x, x),$$

specification of the initial position of the point $x(t_0)=x_0$ and the initial velocity $\dot{x}(t_0)=\dot{x}_0$ will determine the unique solution, the unique law of motion $x=x(t)$ if, of course, the function f satisfies the conditions of the existence and uniqueness theorem.

The theorem (considered on page 58) on the continuous dependence of the solution on parameters and on the initial values can be extended, without altering the method of proof, to systems of differential equations and hence to equations of the nth order.

2. The Most Elementary Cases of Reducing the Order

In certain cases the order of a differential equation may be reduced. This ordinarily simplifies its integration. The following are some frequently encountered classes of equations that permit reducing the order.

1. *The equation does not contain the desired function and its derivatives up to order $k-1$ inclusive:*

$$F(x, y^{(k)}, y^{(k+1)}, \ldots, y^{(n)})=0. \tag{2.3}$$

In this case the order of the equation may be reduced to $n-k$ by changing the variables: $y^{(k)}=p$.

Indeed after changing the variables the equation (2.3) becomes

$$F(x, p, p', \ldots, p^{(n-k)})=0.$$

From this equation we find $p=p(x, c_1, c_2, \ldots, c_{n-k})$ and y is found from $y^{(k)}=p(x, c_1, c_2, \ldots, c_{n-k})$ by k-fold integration. In particular, if a second-order equation does not contain y, the change of variables $y'=p$ produces an equation of the first order.

Example 1.

$$\frac{d^5y}{dx^5}-\frac{1}{x}\frac{d^4y}{dx^4}=0.$$

Assuming $\frac{d^4y}{dx^4}=p$, we get $\frac{dp}{dx}-\frac{1}{x}p=0$; separating variables and integrating we have: $\ln|p|=\ln|x|+\ln c$, or $p=cx$, $\frac{d^4y}{dx^4}=cx$, whence

$$y=c_1x^5+c_2x^3+c_3x^2+c_4x+c_5.$$

Example 2. Find the law of motion of a body falling in the air without an initial velocity. Take the air resistance to be proportional to the square of the velocity.

The equation of motion is of the form

$$m\frac{d^2s}{dt^2}=mg-k\left(\frac{ds}{dt}\right)^2,$$

where s is the distance covered, m the mass of the body, and t the time. For $t=0$, we have $s=0$ and $\frac{ds}{dt}=0$.

The equation does not explicitly contain the unknown function s, and so the order of the equation may be reduced by taking $\frac{ds}{dt}=v$. Then the equation of motion will be of the form

$$m\frac{dv}{dt}=mg-kv^2.$$

Separating the variables and integrating, we get

$$\frac{m\,dv}{mg-kv^2}=dt;\quad t=m\int_0^v\frac{dv}{mg-kv^2}=\frac{1}{k\sqrt{g}}\tanh^{-1}\frac{kv}{\sqrt{g}};$$

from this we have $v=\frac{\sqrt{g}}{k}\tanh(k\sqrt{g}\,t)$; multiplying by dt 'and integrating again, we find the law of motion:

$$s=\frac{1}{k^2}\ln\cosh(k\sqrt{g}\,t).$$

2. *The equation does not contain an independent variable:*

$$F(y,\ y',\ y'',\ \ldots,\ y^{(n)})=0.$$

Here, the order of the equation may be reduced by unity by the substitution $y'=p$, p being regarded as a new unknown function of y, $p=p(y)$, and, hence, all derivatives $\frac{d^ky}{dx^k}$ have to be expressed in terms of the derivatives of the new unknown function $p(y)$ with respect to y:

$$\frac{dy}{dx}=p,$$

$$\frac{d^2y}{dx^2}=\frac{dp}{dx}=\frac{dp}{dy}\frac{dy}{dx}=\frac{dp}{dy}p,$$

$$\frac{d^3y}{dx^3}=\frac{d}{dx}\left(\frac{dp}{dy}p\right)=\frac{d}{dy}\left(\frac{dp}{dy}p\right)\frac{dy}{dx}=\frac{d^2p}{dy^2}p^2+\left(\frac{dp}{dy}\right)^2p$$

and similarly for derivatives of higher order. It is obvious here that the derivatives $\frac{d^ky}{dx^k}$ are expressed in terms of derivatives of order not higher than $k-1$ of p with respect to y, which is what reduces the order by one.

In particular, if a second-order equation does not contain an independent variable, the indicated change of variable leads to a first-order equation.

Example 3.

$$y\frac{d^2y}{dx^2}-\left(\frac{dy}{dx}\right)^2=0.$$

Putting $\frac{dy}{dx}=p$, $\frac{d^2y}{dx^2}=p\frac{dp}{dy}$, we get the separable equation $yp\frac{dp}{dy}-p^2=0$, the general solution of which is $p=c_1y$ or $\frac{dy}{dx}=c_1y$. Again separating variables and integrating we get $\ln|y|=c_1x+\ln c_2$ or $y=c_2e^{c_1x}$.

Example 4. Integrate the equation of a simple pendulum $\ddot{x}+a^2\sin x=0$ for initial conditions $x(0)=x_0$, $\dot{x}(0)=0$.

We reduce the order by putting

$$\dot{x}=v,\quad \ddot{x}=v\frac{dv}{dx},\qquad v\,dv=-a^2\sin x\,dx,$$

$$\frac{v^2}{2}=a^2(\cos x-\cos x_0),\qquad v=\pm a\sqrt{2(\cos x-\cos x_0)},$$

$$\frac{dx}{dt}=\pm a\sqrt{2(\cos x-\cos x_0)},\qquad t=\pm\frac{1}{a\sqrt{2}}\int_{x}^{x}\frac{dx}{\sqrt{\cos x-\cos x_0}}.$$

The integral on the right side is not expressible by elementary functions but is readily reducible to elliptic functions.

3. *The left-hand side of the equation*

$$F(x,\ y,\ y',\ y'',\dots,y^{(n)})=0 \tag{2.4}$$

is a derivative of some differential expression of the $(n-1)$*th order* $\Phi(x,\ y,\ y',\ \dots,\ y^{(n-1)})$.

Here we readily find a so-called *first integral*, i.e. a differential equation of the $(n-1)$th order containing one arbitrary constant. This equation is equivalent to the given equation of the nth order, so we have reduced the order of the equation by one. Indeed, (2.4) may be rewritten in the form

$$\frac{d}{dx}\Phi(x,\ y,\ y',\ \dots,\ y^{(n-1)})=0. \tag{2.4$_1$}$$

If $y(x)$ is a solution of (2.4$_1$), then the derivative of the function $\Phi(x,\ y,\ y',\ \dots,\ y^{(n-1)})$ is identically zero. Consequently, the function $\Phi(x,\ y,\ y',\ \dots,\ y^{(n-1)})$ is equal to a constant, and we get the first integral

$$\Phi(x,\ y,\ y',\ \dots,\ y^{(n-1)})=c.$$

Example 5.

$$yy''+(y')^2=0.$$

This equation may be written as $d(yy')=0$, whence $yy'=c_1$ or $y\,dy=c_1\,dx$. Thus, $y^2=c_1x+c_2$ is the complete integral.

Sometimes the left-hand side of the equation $F(x,y,y',\dots,y^{(n)})=0$ becomes the derivative of the $(n-1)$th order differential expression $\Phi(x,\ y,\ y',\ \dots,\ y^{(n-1)})$ only after multiplication by some factor $\mu(x,\ y,\ y',\ \dots,\ y^{(n-1)})$.

Example 6.

$$yy'' - (y')^2 = 0.$$

Multiplying by the factor $\mu = \frac{1}{y^2}$, we get $\frac{yy'' - (y')^2}{y^2} = 0$ or $\frac{d}{dx}\left(\frac{y'}{y}\right) = 0$, whence $\frac{y'}{y} = c_1$ or $\frac{d}{dx} \ln |y| = c_1$. Hence, $\ln|y| = c_1 x + \ln c_2$, $c_2 > 0$, whence $y = c_2 e^{c_1 x}$, $c_2 \neq 0$, as in Example 3 of this section.

Note. When multiplying by the factor $\mu(x, y, y', \ldots, y^{(n-1)})$ extraneous solutions may be introduced that make this factor vanish. If the factor μ is discontinuous, a loss of solutions is possible. In Example 6, when multiplying by $\mu = \frac{1}{y^2}$, the solution $y = 0$ was lost. However, this solution can be included in the solution obtained $y = c_2 e^{c_1 x}$ if it is taken that $\bar{c}_2$ can assume the value 0.

4. *The equation $F(x, y, y', \ldots, y^{(n)}) = 0$ is homogeneous in the arguments $y, y', \ldots, y^{(n)}$.*

The order of the following equation (which is homogeneous in $y, y', \ldots, y^{(n)}$)

$$F(x, y, y', \ldots, y^{(n)}) = 0, \tag{2.5}$$

that is, an equation for which the identity

$$F(x, ky, ky', \ldots, ky^{(n)}) = k^p F(x, y, y', \ldots, y^{(n)})$$

holds, may be reduced by unity by the substitution $y = e^{\int z\,dx}$, where z is a new unknown function. Indeed, differentiating, we get

$$\begin{aligned} y' &= e^{\int z\,dx} z, \\ y'' &= e^{\int z\,dx} (z^2 + z'), \\ y''' &= e^{\int z\,dx} (z^3 + 3zz' + z''), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y^{(k)} &= e^{\int z\,dx} \Phi(z, z', z'', \ldots, z^{(k-1)}) \end{aligned}$$

(that this equality is true can be demonstrated by the method of induction).

By substitution into (2.5) and by noting that by virtue of homogeneity the factor $e^{p \int z\,dx}$ may be taken outside the sign of the function F, we get

$$e^{p \int z\,dx} f(x, z, z', \ldots, z^{(n-1)}) = 0$$

or, cancelling $e^{p \int z\,dx}$, we have

$$f(x, z, z', \ldots, z^{(n-1)}) = 0.$$

Example 7.

$$yy'' - (y')^2 = 6xy^2.$$

Putting $y = e^{\int z\,dx}$, we get $z' = 6x$, $z = 3x^2 + c_1$, $y = e^{\int (3x^2 + c_1)\,dx}$ or $y = c_2 e^{(x^3 + c_1 x)}$.

Especially frequently encountered in applications are *second-order* differential equations that allow for reducing the order.

(1) $$F(x, y'') = 0. \qquad (2.6)$$

Here, we can lower the order by the substitution $y' = p$ and we can reduce it to the equation $F\left(x, \frac{dp}{dx}\right) = 0$ considered on page 75.

Equation (2.6) can be solved for the second argument $y'' = f(x)$ and integrated twice, or we can introduce a parameter and replace (2.6) by its parametric representation

$$\frac{d^2y}{dx^2} = \varphi(t), \quad x = \psi(t),$$

whence

$$dy' = y''\,dx = \varphi(t)\,\psi'(t)\,dt, \quad y' = \int \varphi(t)\,\psi'(t)\,dt + c_1,$$

$$dy = y'\,dx, \quad y = \int \left[\int \varphi(t)\,\psi'(t)\,dt + c_1\right]\psi'(t)\,dt + c_2.$$

(2) $$F(y', y'') = 0. \qquad (2.7)$$

Putting $y' = p$, transform (2.7) to equation (1.61), page 76, or represent (2.7) parametrically:

$$y'_x = \varphi(t), \quad y''_{xx} = \psi(t),$$

whence

$$dx = \frac{dy'}{y''} = \frac{\varphi'(t)\,dt}{\psi(t)}, \quad x = \int \frac{\varphi'(t)\,dt}{\psi(t)} + c_1,$$

y is then determined by a quadrature:

$$dy = y'\,dx = \varphi(t)\frac{\varphi'(t)}{\psi(t)}\,dt, \quad y = \int \frac{\varphi(t)\,\varphi'(t)}{\psi(t)}\,dt + c_2.$$

(3) $$F(y, y'') = 0. \qquad (2.8)$$

The order can be reduced by putting

$$\frac{dy}{dx} = p, \qquad \frac{d^2y}{dx^2} = \frac{dp}{dy}\frac{dy}{dx} = p\frac{dp}{dy}.$$

If equation (2.8) is readily solvable for the second argument $y'' = f(y)$, then by multiplying this equation termwise by $2y'\,dx = 2dy$, we get

$d(y')^2 = 2f(y)\,dy$, whence

$$\frac{dy}{dx} = \pm\sqrt{2\int f(y)\,dy + c_1}, \qquad \pm\frac{dy}{\sqrt{2\int f(y)\,dy + c_1}} = dx,$$

$$x + c_2 = \pm\int\frac{dy}{\sqrt{2\int f(y)\,dy + c_1}}.$$

Equation (2.8) may be replaced by its parametric representation $y = \varphi(t)$, $y'' = \psi(t)$; then from $dy' = y''\,dx$ and $dy = y'\,dx$ we get $y'\,dy' = y''\,dy$ or

$$\frac{1}{2}d(y')^2 = \psi(t)\,\varphi'(t)\,dt,$$

$$(y')^2 = 2\int\psi(t)\,\varphi'(t)\,dt + c_1,$$

$$y' = \pm\sqrt{2\int\psi(t)\,\varphi'(t)\,dt + c_1},$$

then from $dy = y'\,dx$ we find dx and then x:

$$dx = \frac{dy}{y'} = \frac{\varphi'(t)\,dt}{\pm\sqrt{2\int\psi(t)\,\varphi'(t)\,dt + c_1}},$$

$$x = \pm\int\frac{\varphi'(t)\,dt}{\sqrt{2\int\psi(t)\,\varphi'(t)\,dt + c_1}} + c_2. \qquad (2.9)$$

It is equation (2.9) and $y = \varphi(t)$ that define in parametric form a family of integral curves.

Example 8.

$$y'' = 2y^3, \quad y(0) = 1, \quad y'(0) = 1.$$

Multiplying both sides of the equation by $2y'\,dx$, we get $d(y')^2 = 4y^3\,dy$, whence $(y')^2 = y^4 + c_1$. Taking the initial conditions into account, we find that $c_1 = 0$ and $y' = y^2$. Hence, $\frac{dy}{y^2} = dx$, $-\frac{1}{y} = x + c_2$, $c_2 = -1$, $y = \frac{1}{1-x}$.

3. Linear Differential Equations of the *n*th Order

An *n*th *order linear differential equation* is an equation linear in the unknown function and its derivatives and, hence, of the form

$$a_0(x)\,y^{(n)} + a_1(x)\,y^{(n-1)} + \ldots + a_{n-1}(x)\,y' + a_n(x)\,y = \varphi(x). \qquad (2.10)$$

If the right side $\varphi(x) \equiv 0$, then the equation is called *homogeneous linear* since it is homogeneous in the unknown function y and its derivatives.

If the coefficient $a_0(x)$ is not zero at any point of some interval $a \leqslant x \leqslant b$, then by dividing by $a_0(x)$ we can reduce the homogeneous linear equation, for x varying on this interval, to the form

$$y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_{n-1}(x) y' + p_n(x) y = 0 \qquad (2.11)$$

or

$$y^{(n)} = -\sum_{i=1}^{n} p_i(x) y^{(n-i)}. \qquad (2.11_1)$$

If the coefficients $p_i(x)$ are continuous on the interval $a \leqslant x \leqslant b$, then in the neighbourhood of any initial values

$$y(x_0) = y_0, \quad y_0'(x_0) = y_0', \quad \ldots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)},$$

where x_0 is any point of the interval $a < x < b$, the conditions of the existence and uniqueness theorem are satisfied.

Indeed, the right side of (2.11_1) is continuous with respect to the collection of all arguments and there exist partial derivatives $\frac{\partial f}{\partial y^{(k)}} = -p_{n-k}(x)$ $(k = 0, 1, \ldots, (n-1))$ bounded in absolute value, since the functions $p_{n-k}(x)$ are continuous on the interval $a \leqslant x \leqslant b$ and, consequently, are bounded in absolute value.

Note that linearity and homogeneity of the equation are retained for any transformation of the independent variable $x = \varphi(t)$, where $\varphi(t)$ is an arbitrary n times differentiable function, the derivative of which, on the interval of variation of t under consideration, is $\varphi'(t) \neq 0$.

Indeed,

$$\frac{dy}{dx} = \frac{dy}{dt} \frac{1}{\varphi'(t)},$$

$$\frac{d^2y}{dx^2} = \frac{d^2y}{dt^2} \frac{1}{[\varphi'(t)]^2} - \frac{dy}{dt} \frac{\varphi''(t)}{[\varphi'(t)]^3},$$

$$\cdots \quad \cdots\cdots\cdots\cdots$$

A derivative of any order $\frac{d^ky}{dx^k}$ is a homogeneous linear function of the derivatives $\frac{dy}{dt}, \frac{d^2y}{dt^2}, \ldots, \frac{d^ky}{dt^k}$ and, consequently, when substituted into equation (2.11) retains its linearity and homogeneity.

Linearity and homogeneity are also retained in a homogeneous linear transformation of the unknown function $y(x) = \alpha(x) z(x)$.

Indeed, by the formula for differentiating a product,

$$y^{(k)} = \alpha(x) z^{(k)} + k\alpha'(x) z^{(k-1)} + \frac{k(k-1)}{2!}\alpha''(x) z^{(k-2)} + \ldots + \alpha^{(k)}(x) z,$$

that is, the derivative $y^{(k)}$ is a homogeneous linear function of z, z', z'', $\ldots$, $z^{(k)}$. Consequently, after change of variables the left side of the homogeneous linear equation

$$a_0(x) y^{(n)} + a_1(x) y^{(n-1)} + \ldots + a_n(x) y = 0$$

will be a homogeneous linear function of $z, z', \ldots, z^{(n)}$.

Let us write the homogeneous linear equation

$$y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_n(x) y = 0$$

in abridged form as

$$L[y] = 0,$$

where

$$L[y] = y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_n(x) y.$$

We shall call $L[y]$ a *linear differential operator.*

A linear differential operator has the following two basic properties:

(1) *A constant factor is taken outside the sign of the linear operator:*

$$L[cy] \equiv cL[y].$$

Indeed,

$$(cy)^{(n)} + p_1(x)(cy)^{(n-1)} + \ldots$$
$$\ldots + p_n(x)(cy) \equiv c[y^{(n)} + p_1(x) y^{(n+1)} + \ldots + p_n(x) y].$$

(2) *A linear differential operator applied to the sum of two functions y_1 and y_2 is equal to the sum of the results of applying the same operator to each function separately:*

$$L[y_1 + y_2] \equiv L[y_1] + L[y_2].$$

Indeed,

$$(y_1 + y_2)^{(n)} + p_1(x)(y_1 + y_2)^{(n-1)} + \ldots + p_n(x)(y_1 + y_2) \equiv$$
$$\equiv [y_1^{(n)} + p_1(x) y_1^{(n-1)} + \ldots + p_n(x) y_1] + [y_2^{(n)} + p_1(x) y_2^{(n-1)} + \ldots$$
$$\ldots + p_n(x) y_2].$$

A corollary of properties (1) and (2) is

$$L\left[\sum_{i=1}^{m} c_i y_i\right] \equiv \sum_{i=1}^{m} c_i L[y_i],$$

where c_i are constants.

Proceeding from the properties of the linear operator L, we shall prove a number of theorems on the solutions of a homogeneous linear equation.

Theorem 2.2. *If y_1 is a solution of a homogeneous linear equation $L[y]=0$, then cy_1, where c is an arbitrary constant, is likewise a solution of that equation.*

Proof. Given $L[y_1]\equiv 0$. It is required to prove that $L[cy_1]\equiv 0$. Taking advantage of property (1) of the operator L, we get

$$L[cy_1]\equiv cL[y_1]\equiv 0.$$

Theorem 2.3. *The sum y_1+y_2 of solutions y_1 and y_2 of a homogeneous linear equation $L[y]=0$ is a solution of that equation.*

Proof. Given $L[y_1]\equiv 0$ and $L[y_2]\equiv 0$. It is required to prove that $L[y_1+y_2]\equiv 0$.

Taking advantage of Property (2) of the operator L, we get

$$L[y_1+y_2]\equiv L[y_1]+L[y_2]\equiv 0.$$

Corollary to Theorems 2.2 and 2.3. *A linear combination with arbitrary constant coefficients $\sum_{i=1}^{m} c_i y_i$ of solutions $y_1, y_2, \ldots, y_m$ of a homogeneous linear equation $L[y]=0$ is a solution of that equation.*

Theorem 2.4. *If a homogeneous linear equation $L[y]=0$ with real coefficients $p_i(x)$ has a complex solution $y(x)=u(x)+iv(x)$, then the real part of this solution $u(x)$ and its imaginary part $v(x)$ are separately solutions of that homogeneous equation.*

Proof. Given $L[u(x)+iv(x)]\equiv 0$. It is required to prove that $L[u]\equiv 0$ and $L[v]\equiv 0$.

Taking advantage of Properties (1) and (2) of the operator L, we get

$$L[u+iv]\equiv L[u]+iL[v]\equiv 0,$$

whence $L[u]\equiv 0$ and $L[v]\equiv 0$, since the complex function of a real variable vanishes identically when, and only when, its real and imaginary parts are identically equal to zero.

Note. We applied Properties (1) and (2) of the operator L to the complex function $u(x)+iv(x)$ of a real variable, which is obviously admissible since in proving Properties (1) and (2) use was made only of the following properties of the derivatives: $(cy)'=cy'$, where c is a constant, and $(y_1+y_2)'=y_1'+y_2'$, which hold true for complex functions of a real variable as well.

The functions $y_1(x), y_2(x), \ldots, y_n(x)$ are called *linearly dependent* over a certain interval of variation of x, $a\leqslant x\leqslant b$, if there exist constant quantities $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that on this interval

$$\alpha_1 y_1+\alpha_2 y_2+\ldots+\alpha_n y_n\equiv 0, \qquad (2.12)$$

and at least one $\alpha_i\neq 0$. Now if the identity (2.12) holds true only for $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, then the functions $y_1, y_2, \ldots, y_n$ are called *linearly independent functions* over the interval $a \leqslant x \leqslant b$.

Example 1. The functions $1, x, x^2, \ldots, x^n$ are linearly independent on any interval $a \leqslant x \leqslant b$, since the identity

$$\alpha_1 + \alpha_2 x + \alpha_3 x^2 + \ldots + \alpha_{n+1} x^n \equiv 0 \qquad (2.13)$$

is only possible if all the $\alpha_i = 0$. If even one $\alpha_i \neq 0$, the left member of the identity (2.13) would be a polynomial of degree not higher than n that can have no more than n distinct roots and, hence, vanishes at no more than n points of the interval under consideration.

Example 2. The functions $e^{k_1 x}, e^{k_2 x}, \ldots, e^{k_n x}$, where $k_i \neq k_j$ for $i \neq j$, are linearly independent on any interval $a \leqslant x \leqslant b$.

Suppose that the functions under consideration are linearly dependent. Then

$$\alpha_1 e^{k_1 x} + \alpha_2 e^{k_2 x} + \ldots + \alpha_n e^{k_n x} \equiv 0, \qquad (2.14)$$

where at least one of the $\alpha_i \neq 0$, for instance, for the sake of definiteness, $\alpha_n \neq 0$. Dividing the identity (2.14) by $e^{k_1 x}$ and differentiating, we get

$$\alpha_2 (k_2 - k_1) e^{(k_2 - k_1) x} + \ldots + \alpha_n (k_n - k_1) e^{(k_n - k_1) x} \equiv 0, \qquad (2.15)$$

which is a linear relation between $n-1$ exponential functions of the form e^{px} with distinct exponents. Dividing the identity (2.15) by $e^{(k_2 - k_1) x}$ and differentiating, we obtain a linear relation between $n-2$ exponential functions with distinct exponents. Continuing this process $n-1$ times, we get

$$\alpha_n (k_n - k_1)(k_n - k_2) \ldots (k_n - k_{n-1}) e^{(k_n - k_{n-1}) x} \equiv 0,$$

which is impossible since α_n, by hypothesis, is different from zero, and $k_i \neq k_j$ when $i \neq j$.

The proof holds true for complex k_i.

Example 3. The functions

$$\begin{matrix} e^{k_1 x}, & x e^{k_1 x}, & \ldots, & x^{n_1} e^{k_1 x}, \\ e^{k_2 x}, & x e^{k_2 x}, & \ldots, & x^{n_2} e^{k_2 x}, \\ \cdot & \cdot & \cdot & \cdot \\ e^{k_p x}, & x e^{k_p x}, & \ldots, & x^{n_p} e^{k_p x}, \end{matrix}$$

where $k_i \neq k_j$ for $i \neq j$, are linearly independent on any interval $a \leqslant x \leqslant b$.

Suppose that these functions are linearly dependent. Then

$$P_1(x) e^{k_1 x} + P_2(x) e^{k_2 x} + \ldots + P_p(x) e^{k_p x} \equiv 0, \qquad (2.16)$$

where $P_i(x)$ is a polynomial of degree not higher than n_i, and at least one polynomial, for instance, $P_p(x)$, is not identically zero. Dividing identity (2.16) by $e^{k_1 x}$ and differentiating $n_1 + 1$ times, we find that the first summand in (2.16) vanishes and we get a linear relation of the same type but with a smaller number of functions:

$$Q_2(x)\, e^{(k_2-k_1)x} + \ldots + Q_p(x)\, e^{(k_p-k_1)x} = 0. \tag{2.17}$$

The degrees of the polynomials Q_i and P_i ($i=2, 3, \ldots, p$) coincide since on differentiating the product $P_i(x)e^{px}$, $p \neq 0$, we get $[P_i(x)p + P_i'(x)]\, e^{px}$, that is, the coefficient of the highest-degree term of the polynomial $P_i(x)$, after differentiating the product $P_i(x)e^{px}$, acquires only the nonzero factor p. In particular, the degrees of the polynomials $P_p(x)$ and $Q_p(x)$ coincide, and hence the polynomial $Q_p(x)$ is not identically zero. Dividing (2.17) by $e^{(k_2-k_1)x}$ and differentiating n_2+1 times, we get a linear relation with a still smaller number of functions. Continuing this process $p-1$ times, we obtain

$$R_p(x)\, e^{(k_p-k_{p-1})x} \equiv 0,$$

which is impossible, since the degree of the polynomial $R_p(x)$ is equal to the degree of the polynomial $P_p(x)$ and, hence, the polynomial $R_p(x)$ is not identically zero.

The proof does not change in the case of complex k_i either.

Theorem 2.5. *If the functions $y_1, y_2, \ldots, y_n$ are linearly dependent on the interval $a \leqslant x \leqslant b$, then on the same interval the determinant*

$$W(x) = W[y_1, y_2, \ldots, y_n] = \begin{vmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ y_1'' & y_2'' & \ldots & y_n'' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} \end{vmatrix},$$

called the Wronskian is identically zero.*

Proof. Given

$$\alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_n y_n \equiv 0 \tag{2.18}$$

on the interval $a \leqslant x \leqslant b$ and not all the α_i are zero. Differentiating the identity (2.18) $n-1$ times, we get

$$\left.\begin{array}{l} \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_n y_n \equiv 0, \\ \alpha_1 y_1' + \alpha_2 y_2' + \ldots + \alpha_n y_n' \equiv 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_1 y_1^{(n-1)} + \alpha_2 y_2^{(n-1)} + \ldots + \alpha_n y_n^{(n-1)} \equiv 0. \end{array}\right\} \tag{2.19}$$

This homogeneous (with respect to all the α_i) linear system of n equations has a nontrivial solution (since not all the α_i are equal to zero) for any value of x on the interval $a \leqslant x \leqslant b$. Consequently,

* Named after the Polish mathematician G. Wronski (1775-1853).

the determinant of the system (2.19), which is the Wronskian $W[y_1, y_2, \ldots, y_n]$, is equal to zero at every point x of the interval $a \leqslant x \leqslant b$.

Theorem 2.6. *If the linearly independent functions $y_1, y_2, \ldots, y_n$ are solutions of the homogeneous linear equation*

$$y^{(n)} + p_n(x) y^{(n-1)} + \ldots + p_n(x) y = 0 \tag{2.20}$$

with continuous coefficients $p_i(x)$ on the interval $a \leqslant x \leqslant b$, then the Wronskian

$$W(x) = \begin{vmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} \end{vmatrix}$$

cannot vanish at any point of the interval $a \leqslant x \leqslant b$.

Proof. Suppose that at some point $x = x_0$ of the interval $a \leqslant x \leqslant b$ the Wronskian $W(x_0) = 0$. Choose constants α_i $(i = 1, 2, \ldots, n)$ so as to satisfy the system of equations

$$\left.\begin{aligned} &\alpha_1 y_1(x_0) + \alpha_2 y_2(x_0) + \ldots + \alpha_n y_n(x_0) = 0, \\ &\alpha_1 y_1'(x_0) + \alpha_2 y_2'(x_0) + \ldots + \alpha_n y_n'(x_0) = 0, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &\alpha_1 y_1^{(n-1)}(x_0) + \alpha_2 y_2^{(n-1)}(x_0) + \ldots + a_n y_n^{(n-1)}(x_0) = 0 \end{aligned}\right\} \tag{2.21}$$

and so that not all the α_i are zero. Such a choice is possible since the determinant of the homogeneous linear system (2.21) of n equations in n unknowns α_i is zero, $W(x_0) = 0$, and, hence, there exist nontrivial solutions of this system. In such a choice of α_i, the linear combination

$$y = \alpha_1 y_1(x) + \alpha_2 y_2(x) + \ldots + \alpha_n y_n(x)$$

will be a solution of the homogeneous linear equation (2.20) that satisfies, by virtue of the equations of the system (2.21), the zero initial conditions

$$y(x_0) = 0, \quad y'(x_0) = 0, \quad \ldots, \quad y^{(n-1)}(x_0) = 0. \tag{2.22}$$

Obviously, such initial conditions are satisfied by the trivial solution $y \equiv 0$ of the equation (2.20) and, by the uniqueness theorem, the initial conditions (2.22) are satisfied only by this solution. Consequently, $\alpha_1 y_1(x) + \alpha_2 y_2(x) + \ldots + \alpha_n y_n(x) \equiv 0$ and the solutions $y_1, y_2, \ldots, y_n$, despite the hypothesis of the theorem, are linearly dependent.

Note. 1. From the Theorems 2.5 and 2.6 it follows that the solutions $y_1, y_2, \ldots, y_n$ of equation (2.20), that are linearly independent on the interval $a \leqslant x \leqslant b$, are also linearly independent on any interval $a_1 \leqslant x \leqslant b_1$ located on the interval $a \leqslant x \leqslant b$.

Note. 2. In Theorem 2.6, in contrast to Theorem 2.5, it was assumed that the functions $y_1, y_2, \ldots, y_n$ are solutions of the homogeneous linear equation (2.20) with continuous coefficients. It is not possible to reject this demand and to consider the functions $y_1, y_2, \ldots, y_n$ arbitrary $n-1$ times continuously differentiable functions. It is easy to give examples of linearly independent functions that are of course not solutions of the equation (2.20) with continuous coefficients for which the Wronskian not only vanishes at separate points but is even identically zero. For instance, let there be defined two functions $y_1(x)$ and $y_2(x)$ on the interval $0 \leqslant x \leqslant 2$:

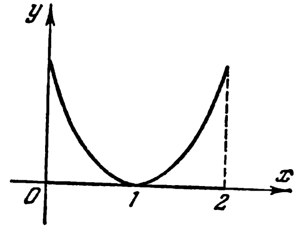


Fig. 2-1

$$y_1(x) = (x-1)^2 \quad \text{for } 0 \leqslant x \leqslant 1$$

and

$$y_1(x) = 0 \quad \text{for } 1 < x \leqslant 2$$

$$y_2(x) = 0 \quad \text{for } 0 \leqslant x \leqslant 1$$

and

$$y_2(x) = (x-1)^2 \quad \text{for } 1 < x \leqslant 2$$

(Fig. 2.1).

Obviously, $\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} \equiv 0$ for $0 \leqslant x \leqslant 2$, since on $0 \leqslant x \leqslant 1$ the second column consists of zeros, while for $1 < x \leqslant 2$ the first column consists of zeros. However, the functions $y_1(x)$ and $y_2(x)$ are linearly independent on the whole interval $0 \leqslant x \leqslant 2$, since, considering the identity $\alpha_1 y_1 + \alpha_2 y_2 \equiv 0$, $0 \leqslant x \leqslant 2$, first on the interval $0 \leqslant x \leqslant 1$, we conclude that $\alpha_1 = 0$ and then, considering this identity on the interval $1 < x \leqslant 2$, we find that $\alpha_2 = 0$ as well.

Theorem 2.7. *The general solution, for $a \leqslant x \leqslant b$, of the homogeneous linear equation*

$$y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_n(x) y = 0 \tag{2.20}$$

with the coefficients $p_i(x)$ $(i = 1, 2, \ldots, n)$, continuous on the interval $a \leqslant x \leqslant b$, is the linear combination $y = \sum_{i=1}^{n} c_i y_i$ of n linearly independent (on the same interval) partial solutions y_i $(i = 1, 2, \ldots, n)$ with arbitrary constant coefficients.

Proof. When $a \leqslant x \leqslant b$, equation (2.20) satisfies the conditions of the theorem of existence and uniqueness. Therefore, when $a \leqslant x \leqslant b$, the solution $y = \sum_{i=1}^{n} c_i y_i$ will be the general solution, that is, it will contain all partial solutions without exception if it is possible to choose arbitrary constants c_i so as to satisfy the arbitrarily specified initial conditions

$$y(x_0)=y_0, \quad y'(x_0)=y'_0, \ \ldots, \ y^{(n-1)}(x_0)=y_0^{(n-1)},$$

where x_0 is any point of the interval $a\leqslant x\leqslant b$.

If we demand that the solution $y=\sum_{i=1}^{n} c_i y_i$ satisfy the initial conditions posed, then we obtain a system of n linear (in c_i, where $i=1, 2, \ldots, n$) equations

$$\left.\begin{array}{l} \sum\limits_{i=1}^{n} c_i y_i(x_0)=y_0, \\ \sum\limits_{i=1}^{n} c_i y'_i(x_0)=y'_0, \\ \cdots\cdots\cdots\cdots \\ \sum\limits_{i=1}^{n} c_i y_i^{(n-1)}(x_0)=y_0^{(n-1)} \end{array}\right\}$$

with n unknowns c_i, with a nonzero determinant of the system, since this determinant is the Wronskian $W(x_0)$ for n linearly independent solutions of equation (2.20). Thus, this system is solvable for c_i, given any choice of x_0 on the interval $a\leqslant x\leqslant b$ and for any kinds of right members.

Corollary to Theorem 2.7. *The maximum number of linearly independent solutions of a homogeneous linear differential equation is equal to its order.*

Note. Any n linearly independent particular solutions of a homogeneous linear equation of nth order are called its *fundamental system of solutions*. Every homogeneous linear equation (2.20) has a fundamental system of solutions. To construct a fundamental system of solutions, arbitrarily specify n^2 numbers

$$y_i^{(k)}(x_0)\,(i=1, 2, \ldots, n;\ k=0, 1, \ldots, n-1)$$

subjecting the choice to the sole restriction that

$$\begin{vmatrix} y_1(x_0) & y_2(x_0) & \ldots & y_n(x_0) \\ y'_1(x_0) & y'_2(x_0) & \ldots & y'_n(x_0) \\ \cdots & \cdots & \cdots & \cdots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \ldots & y_n^{(n-1)}(x_0) \end{vmatrix} \neq 0,$$

where x_0 is any point of the interval $a\leqslant x\leqslant b$. Then the solution $y_i(x)$, defined by the initial values $y_i^{(k)}(x_0)$ $(k=0, 1, \ldots, n-1;$ $i=1, 2, \ldots, n)$, form a fundamental system, since their Wronskian $W(x)$ at the point $x=x_0$ is nonzero and, consequently, on the basis of Theorems 2.5 and 2.6 the solutions $y_1, y_2, \ldots, y_n$ are linearly independent.

Example 4. The equation $y''-y=0$ has obvious linearly independent particular solutions $y_1=e^x$ and $y_2=e^{-x}$ (see page 102, Example 2); consequently the general solution is of the form $y=c_1e^x+c_2e^{-x}$.

Example 5. The solution $y=c_1e^x+c_2\cosh x+c_3\sinh x$ of the equation $y'''-y'=0$ is not the general solution since the solutions e^x, $\cosh x$, $\sinh x$ are linearly dependent. The linearly independent solutions are 1, $\cosh x$, $\sinh x$, and consequently

$$y=c_1+c_2\cosh x+c_3\sinh x,$$

where c_1, c_2 and c_3 are arbitrary constants, will be the general solution of the equation under consideration.

Knowing one nontrivial particular solution y_1 of the homogeneous linear equation

$$y^{(n)}+p_1(x)\,y^{(n-1)}+\ldots+p_n(x)\,y=0, \tag{2.20}$$

it is possible, by the substitution $y=y_1\int u\,dx$, to reduce the order of the equation and retain its linearity and homogeneity.

Indeed, the substitution $y=y_1\int u\,dx$ may be replaced by two substitutions: $y=y_1z$ and $z'=u$. The homogeneous linear transformation

$$y=y_1z \tag{2.23}$$

preserves the linearity and homogeneity of the equation (see pages 99-100) consequently, (2.20) is thus transformed to

$$a_0(x)\,z^{(n)}+a_1(x)\,z^{(n-1)}+\ldots+a_n(x)\,z=0, \tag{2.24}$$

and by virtue of (2.23) the solution $z\equiv 1$ of (2.24) corresponds to the solution $y=y_1$ of (2.20). Substituting $z\equiv 1$ into (2.24), we get $a_n(x)\equiv 0$. Hence, equation (2.24) has the form

$$a_0(x)\,z^{(n)}+a_1(x)\,z^{(n-1)}+\ldots+a_{n-1}(x)\,z'=0,$$

and the substitution $z'=u$ reduces the order by one:

$$a_0(x)\,u^{(n-1)}+a_1(x)\,u^{(n-2)}+\ldots+a_{n-1}(x)\,u=0.$$

Note that the same substitution $y=y_1\int u\,dx$, where y_1 is a solution of the equation $L[y]=0$, also reduces by unity the order of the nonhomogeneous linear equation $L[y]=f(x)$, since this substitution does not affect the right-hand side of the equation.

Knowing k linearly independent (on the interval $a\leqslant x\leqslant b$) solutions $y_1, y_2, \ldots, y_k$ of a homogeneous linear equation, it is possible to reduce the order of the equation to $n-k$ on the same interval $a \leqslant x \leqslant b$.

Indeed, reducing the order of the equation

$$L[y]=0, \tag{2.20}$$

by unity by the substitution $y=y_k \int u\,dx$, we again get a homogeneous linear equation

$$a_0(x)\,u^{(n-1)}+a_1(x)\,u^{(n-2)}+\ldots+a_{n-1}(x)\,u=0 \tag{2.25}$$

of order $n-1$, and we know $k-1$ of its linearly independent solutions,

$$u_1=\left(\frac{y_1}{y_k}\right)', \quad u_2=\left(\frac{y_2}{y_k}\right)', \quad \ldots, \quad u_{k-1}=\left(\frac{y_{k-1}}{y_k}\right)',$$

which we obtain by substituting $y=y_1,\ y=y_2,\ \ldots,\ y=y_{k-1}$ in succession into $y=y_k \int u\,dx$ or $u=\left(\frac{y}{y_k}\right)'$. [Note that the trivial solution $u \equiv 0$ of the equation (2.25) corresponds to the solution $y=y_k$ of (2.20) that we have already used for reducing the order of the equation.]

The solutions $u_1, u_2, \ldots, u_{k-1}$ are linearly independent, since if there existed a linear relation between them on the interval $a \leqslant x \leqslant b$,

$$\alpha_1 u_1+\alpha_2 u_2+\ldots+\alpha_{k-1} u_{k-1} \equiv 0$$

or

$$\alpha_1\left(\frac{y_1}{y_k}\right)'+\alpha_2\left(\frac{y_2}{y_k}\right)'+\ldots+\alpha_{k-1}\left(\frac{y_{k-1}}{y_k}\right)' \equiv 0, \tag{2.26}$$

where at least one $\alpha_i \neq 0$, then, by multiplying by dx and integrating the identity (2.26) from x_0 to x, where $a \leqslant x \leqslant b$, and x_0 is a point of the interval $[a, b]$, we will have

$$\alpha_1 \frac{y_1(x)}{y_k(x)}+\alpha_2 \frac{y_2(x)}{y_k(x)}+\ldots+\alpha_{k-1} \frac{y_{k-1}(x)}{y_k(x)}-$$
$$-\left[\alpha_1 \frac{y_1(x_0)}{y_k(x_0)}+\alpha_2 \frac{y_2(x_0)}{y_k(x_0)}+\ldots+\alpha_{k-1} \frac{y_{k-1}(x_0)}{y_k(x_0)}\right] \equiv 0,$$

or, by multiplying by $y_k(x)$ and putting

$$-\left[\alpha_1 \frac{y_1(x_0)}{y_k(x_0)}+\alpha_2 \frac{y_2(x_0)}{y_k(x_0)}+\ldots+\alpha_{k-1} \frac{y_{k-1}(x_0)}{y_k(x_0)}\right]=\alpha_k,$$

we will get, despite the initial assumption, a linear relation between the solutions $y_1, y_2, \ldots, y_k$:

$$\alpha_1 y_1+\alpha_2 y_2+\ldots+\alpha_k y_k \equiv 0,$$

where at least one $\alpha_i \neq 0$. Thus, by utilizing a single particular solution y_k we reduced the order of the equation by unity and retained its linearity and homogeneity; also, we know $k-1$ linearly independent solutions of the transformed equation Thus, the same method may be used to reduce the order by one more unit; employing still another solution and continuing this process k times, we get a linear equation of order $n-k$.

Example 6.

$$xy''-xy'+y=0. \tag{2.27}$$

This equation has an obvious particular solution $y_1=x$. Reducing the order by the substitution

$$y=x\int u\,dx,\quad y'=xu+\int u\,dx,\quad y''=xu'+2u,$$

we reduce (2.27) to the form

$$x^2u'+(2-x)\,xu=0,$$

whence

$$\frac{du}{u}=\frac{x-2}{x}dx,\quad u=c_1\frac{e^x}{x^2},\quad y=x\int u\,dx=x\left[c_1\int\frac{e^x}{x^2}dx+c_2\right].$$

Lemma. *Two equations of the form*

$$y^{(n)}+p_1(x)\,y^{(n-1)}+\ldots+p_n(x)\,y=0, \tag{2.28}$$

$$y^{(n)}+q_1(x)\,y^{(n-1)}+\ldots+q_n(x)\,y=0, \tag{2.29}$$

where the functions $p_i(x)$ and $q_i(x)$ $(i=1, 2, \ldots, n)$, continuous on the interval $a\leqslant x\leqslant b$ and having a common fundamental system of solutions $y_1, y_2, \ldots, y_n$, coincide, that is $p_i(x)\equiv q_i(x)$ $(i=1, 2, \ldots, n)$ on the interval $a\leqslant x\leqslant b$.

Proof. Subtracting (2.29) termwise from (2.28), we get a new equation:

$$[p_1(x)-q_1(x)]\,y^{(n-1)}+[p_2(x)-q_2(x)]\,y^{(n-2)}+\ldots$$
$$\ldots+[p_n(x)-q_n(x)]\,y=0, \tag{2.30}$$

the solutions of which are the functions $y_1, y_2, \ldots, y_n$ that satisfy the equations (2.28) and (2.29) simultaneously.

Assume that at least one of the coefficients of equation (2.30) $[p_i(x)-q_i(x)]$ is different from zero at least at one point x_0 of the interval $a\leqslant x\leqslant b$. Then, by virtue of the continuity of the functions $p_i(x)$ and $q_i(x)$, this coefficient is different from zero in a certain neighbourhood of the point x_0 and, consequently, in this neighbourhood the functions $y_1, y_2, \ldots, y_n$ are linearly independent solutions of the homogeneous linear equation (2.30) of order not higher than $n-1$, which contradicts the corollary to Theorem 2.7. Hence, all the coefficients of the equation (2.30)

$$p_i(x)-q_i(x)\equiv 0\qquad (i=1, 2, \ldots, n),$$

that is $p_i(x) \equiv q_i(x)$ $(i=1, 2, \ldots, n)$ on the interval $a \leqslant x \leqslant b$.

Thus, the fundamental system of solutions $y_1, y_2, \ldots, y_n$ fully determines the homogeneous linear equation

$$y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_n(x) y = 0, \tag{2.28}$$

and, consequently, we can pose the problem of finding the equation (2.28) that has the specified fundamental system of solutions

$$y_1, y_2, \ldots, y_n.$$

Since any solution y of the desired equation (2.28) must be linearly dependent on the solutions $y_1, y_2, \ldots, y_n$, the Wronskian $W[y_1, y_2, \ldots, y_n, y] = 0$. We write this equation in expanded form:

$$\begin{vmatrix} y_1 & y_2 & \ldots & y_n & y \\ y_1' & y_2' & \ldots & y_n' & y' \\ y_1'' & y_2'' & \ldots & y_n'' & y'' \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} & y^{(n-1)} \\ y_1^{(n)} & y_2^{(n)} & \ldots & y_n^{(n)} & y^{(n)} \end{vmatrix} = 0,$$

or, expanding the elements of the last column,

$$W[y_1, y_2, \ldots, y_n]\, y^{(n)} - \begin{vmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-2)} & y_2^{(n-2)} & \ldots & y_n^{(n-2)} \\ y_1^{(n)} & y_2^{(n)} & \ldots & y_n^{(n)} \end{vmatrix} y^{(n-1)} + \ldots = 0. \tag{2.31}$$

The equation obtained, (2.31), is the desired homogeneous linear equation having the specified fundamental system of solutions $y_1, y_2, \ldots, y_n$ (since for $y = y_i$ $(i=1, 2, \ldots, n)$ $W[y_1, y_2, \ldots, y_n, y] \equiv 0$). Dividing both sides of (2.31) by the nonzero coefficient $W[y_1, y_2, \ldots, y_n]$ of the highest derivative, we reduce it to the form (2.28).

From this it follows, in particular, that

$$p_1(x) = -\frac{\begin{vmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-2)} & y_2^{(n-2)} & \ldots & y_n^{(n-2)} \\ y_1^{(n)} & y_2^{(n)} & \ldots & y_n^{(n)} \end{vmatrix}}{W[y_1, y_2, \ldots, y_n]}.$$

Note that the determinant

$$\begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \dots & \dots & \dots & \dots \\ y_1^{(n-2)} & y_2^{(n-2)} & \dots & y_n^{(n-2)} \\ y_1^{(n)} & y_2^{(n)} & \dots & y_n^{(n)} \end{vmatrix} \tag{2.32}$$

is equal to the derivative of the Wronskian $W[y_1, y_2, \dots, y_n]$. Indeed, by the rule for differentiating a determinant, the derivative

$$\frac{d}{dx}\begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \dots & \dots & \dots & \dots \\ y_1^{(n-2)} & y_2^{(n-2)} & \dots & y_n^{(n-2)} \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix}$$

is equal to the sum over i from 1 to n determinants differing from the Wronskian in that in them the elements of the ith row have been differentiated while the remaining rows of the Wronskian are preserved without change. In this sum, only the last determinant, for $i=n$, which coincides with the determinant (2.32), can be different from zero. The other determinants are zero, since their ith and $i+1$st rows coincide.

Consequently, $p_1(x) = -\frac{W'}{W}$. Whence, by multiplying by dx and integrating, we get

$$\ln|W| = -\int p_1(x)\,dx + \ln c, \qquad W = ce^{-\int p_1(x)\,dx}$$

or

$$W = ce^{-\int_{x_0}^{x} p_1(x)\,dx}. \tag{2.33}$$

For $x = x_0$ we get $c = W(x_0)$, whence

$$W(x) = W(x_0)\,e^{-\int_{x_0}^{x} p_1(x)\,dx}. \tag{2.34}$$

Formulas (2.33) or (2.34), which were first derived by M. Ostrogradsky and, independently, by Liouville, are called *Ostrogradsky-Liouville formulas.*

The Ostrogradsky-Liouville formula (2.34) may be employed for integrating a second-order homogeneous linear equation

$$y'' + p_1(x)\,y' + p_2(x)\,y = 0, \tag{2.35}$$

if a single nontrivial solution of this equation, y_1, is known. According to the Ostrogradsky-Liouville formula (2.34), any solution of (2.35) must also be a solution of the equation

$$\begin{vmatrix} y_1 & y \\ y_1' & y' \end{vmatrix} = c_1 e^{-\int p_1(x)\,dx}$$

or

$$y_1 y' - y y_1' = c_1 e^{-\int p_1(x)\,dx}.$$

This linear equation of the first order is most easily integrated by the integrating factor method.

Multiplying by $\mu = \frac{1}{y_1^2}$, we get

$$\frac{d}{dx}\left(\frac{y}{y_1}\right) = \frac{c_1}{y_1^2} e^{-\int p_1(x)\,dx},$$

whence

$$\frac{y}{y_1} = \int \frac{c_1 e^{-\int p_1(x)\,dx}}{y_1^2}\,dx + c_2,$$

or

$$y = c_2 y_1 + c_1 y_1 \int \frac{e^{-\int p_1(x)\,dx}}{y_1^2}\,dx.$$

4. Homogeneous Linear Equations with Constant Coefficients and Euler's Equations

1. *Homogeneous linear equations with constant coefficients.* If in a homogeneous linear equation

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y = 0 \qquad (2.36)$$

all the coefficients a_i are constant, then its particular solutions may be found in the form $y = e^{kx}$, where k is a constant. Indeed, putting into (2.36) $y = e^{kx}$ and $y^{(p)} = k^p e^{kx}$ $(p = 1, 2, \ldots, n)$, we will have

$$a_0 k^n e^{kx} + a_1 k^{n-1} e^{kx} + \ldots + a_n e^{kx} = 0.$$

Cancelling the nonvanishing factor e^{kx}, we get the so-called *characteristic equation*

$$a_0 k^n + a_1 k^{n-1} + \ldots + a_{n-1} k + a_n = 0. \qquad (2.37)$$

This equation of degree n determines those values of k for which $y = e^{kx}$ is a solution of the original homogeneous linear equation with constant coefficients (2.36). If all the roots $k_1, k_2, \ldots, k_n$ of the characteristic equation are distinct, then we have thus found n

linearly independent solutions $e^{k_1 x}$, $e^{k_2 x}$, ..., $e^{k_n x}$ of the equation (2.36) (see page 102, Example 2). Consequently

$$y = c_1 e^{k_1 x} + c_2 e^{k_2 x} + \ldots + c_n e^{k_n x},$$

where c_i are arbitrary constants, is the general solution of the original equation (2.36). This method of integrating linear equations with constant coefficients was first employed by Euler.

Example 1.

$$y'' - 3y' + 2y = 0.$$

The characteristic equation is of the form $k^2 - 3k + 2 = 0$, its roots are $k_1 = 1$, $k_2 = 2$. Hence, the general solution of the original equation is of the form $y = c_1 e^x + c_2 e^{2x}$.

Example 2.

$$y''' - y' = 0.$$

The characteristic equation $k^3 - k = 0$ has the roots $k_1 = 0$, $k_2 = 1$, $k_3 = -1$. The general solution of the equation under discussion is $y = c_1 + c_2 e^x + c_3 e^{-x}$.

Since the coefficients of (2.36) are assumed real, the complex roots of the characteristic equation can appear only as conjugate pairs. The complex solutions $e^{(\alpha+\beta i)x}$ and $e^{(\alpha-\beta i)x}$ that correspond to the pair of complex conjugate roots

$$k_1 = \alpha + \beta i \text{ and } k_2 = \alpha - \beta i,$$

can be replaced by two real solutions: the real and imaginary parts (see page 101) of one of the solutions

$$e^{(\alpha+\beta i)x} = e^{\alpha x} (\cos \beta x + i \sin \beta x),$$

or

$$e^{(\alpha-\beta i)x} = e^{\alpha x} (\cos \beta x - i \sin \beta x).$$

Thus, to the pair of complex conjugate roots $k_{1,2} = \alpha \pm \beta i$ there correspond two real solutions: $e^{\alpha x} \cos \beta x$ and $e^{\alpha x} \sin \beta x$.

Example 3.

$$y'' + 4y' + 5y = 0.$$

The characteristic equation is of the form $k^2 + 4k + 5 = 0$; its roots are $k_{1,2} = -2 \pm i$. The general solution is

$$y = e^{-2x} (c_1 \cos x + c_2 \sin x).$$

Example 4.

$$y'' + a^2 y = 0.$$

The characteristic equation $k^2 + a^2 = 0$ has the roots $k_{1,2} = \pm ai$. The general solution is

$$y = c_1 \cos ax + c_2 \sin ax.$$

8 - 378

If there are multiple roots among the roots of the characteristic equation, then the number of different solutions of the form e^{kx} is less than n and, hence, the lacking linearly independent solutions have to be sought in a different form.

We shall prove that if a characteristic equation has a root k_i of multiplicity α_i, then the solutions of the original equation will be not only $e^{k_i x}$, but also $xe^{k_i x}$, $x^2e^{k_i x}, \ldots$, $x^{\alpha_i - 1}e^{k_i x}$.

First suppose that the characteristic equation has a root $k_i = 0$ of multiplicity α_i. Hence, the left-hand side of the characteristic equation (2.37) has a common factor k^{α_i} in this case, i.e. the coefficients $a_n = a_{n-1} = \ldots = a_{n-\alpha_i+1} = 0$ and the characteristic equation is of the form

$$a_0k^n + a_1k^{n-1} + \ldots + a_{n-\alpha_i}k^{\alpha_i} = 0.$$

The corresponding homogeneous linear differential equation

$$a_0y^{(n)} + a_1y^{(n-1)} + \ldots + a_{n-\alpha_i}y^{(\alpha_i)} = 0$$

obviously has the particular solutions $1, x, x^2, \ldots, x^{\alpha_i - 1}$ since the equation does not have derivatives of order lower than α_i. Thus, to the multiple root $k_i = 0$ of multiplicity α_i there correspond α_i linearly independent (see page 102, Example 1) solutions

$$1, x, x^2, \ldots, x^{\alpha_i - 1}.$$

If the characteristic equation has a root $k_i \neq 0$ of multiplicity α_i, then a change of variables

$$y = e^{k_i x} z \tag{2.38}$$

reduces the problem to the already considered case of a zero multiple root.

Indeed, as was pointed out on pages 99-100, a homogeneous linear transformation of the unknown function (2.38) preserves linearity and homogeneity of the equation. In the change of variables (2.38) the coefficients are also held constant, since

$$y^{(p)} = (ze^{k_i x})^{(p)} = e^{k_i x}\left(z^{(p)} + pz^{(p-1)}k_i + \frac{p(p-1)}{2!}z^{(p-2)}k_i^2 + \ldots + zk_i^p\right),$$

and after substitution into equation (2.36) and cancelling of $e^{k_i x}$ only constant coefficients remain in $z, z', \ldots, z^{(n)}$.

And so the transformed equation will be a homogeneous linear equation of the nth order with constant coefficients

$$b_0z^{(n)} + b_1z^{(n-1)} + \ldots + b_nz = 0, \tag{2.39}$$

and the roots of the characteristic equation

$$a_0k^n + a_1k^{n-1} + \ldots + a_n = 0 \tag{2.37}$$

differ from the roots of the characteristic equation for the transformed equation (2.39)

$$b_0p^n + b_1p^{n-1} + \dots + b_n = 0 \tag{2.40}$$

by the summand k_i, since between the solutions $y = e^{kx}$ of (2.36) and $z = e^{px}$ of (2.39) there must be a relation $y = ze^{k_ix}$ or $e^{kx} = e^{px}e^{k_ix}$, whence $k = p + k_i$. Therefore, to the root $k = k_i$ of (2.37) there corresponds the root $p_i = 0$ of equation (2.40).

It is easy to verify that in this correspondence the multiplicity of the root will be preserved as well, i.e. the root $p_i = 0$ will have multiplicity α_i.

Indeed, the multiple root k_i of equation (2.37) may be regarded as the result of the coincidence of different roots of this equation when its coefficients are changed; but then, by virtue of the relation $k = p + k_i$ the α_i roots of (2.40) will coincide with $p = 0$.

To the root $p = 0$ of multiplicity α_i there correspond particular solutions $z = 1$, $z = x$, $\dots$, $z = x^{\alpha_i - 1}$. Hence, by virtue of the relations $y = ze^{k_ix}$, to the root k_i of multiplicity α_i of (2.37) there will correspond α_i particular solutions

$$y = e^{k_ix}, \quad y = xe^{k_ix}, \quad \dots, \quad y = x^{\alpha_i - 1}e^{k_ix}. \tag{2.41}$$

It remains to show that the solutions

$$e^{k_ix}, \; xe^{k_ix}, \; \dots, \; x^{\alpha_i - 1}e^{k_ix} \quad (i = 1, 2, \dots, m), \tag{2.42}$$

where m is the number of distinct roots k_i of the characteristic equation, are linearly independent, but this was already proved in Example 3, page 102.

Thus, the general solution of equation (2.36) is of the form

$$y = \sum_{i=1}^{m} (c_{0i} + c_{1i}x + c_{2i}x^2 + \dots + c_{\alpha_i - 1, i}x^{\alpha_i - 1})e^{k_ix},$$

where c_{si} are arbitrary constants.

Example 5.

$$y''' - 3y'' + 3y' - y = 0.$$

The characteristic equation $k^3 - 3k^2 + 3k - 1 = 0$ or $(k-1)^3 = 0$ has the triple root $k_{1,2,3} = 1$. Hence, the general solution is of the form

$$y = (c_1 + c_2x + c_3x^2)\, e^x.$$

If a characteristic equation has a multiple complex root $p + qi$ of multiplicity α, the solutions

$$e^{(p+qi)x}, \; xe^{(p+qi)x}, \; x^2e^{(p+qi)x}, \; \dots, \; x^{\alpha-1}e^{(p+qi)x}$$

that correspond to it may be transformed by means of Euler's formulas

$$e^{(p+qi)x} = e^{px}(\cos qx + i \sin qx)$$

and, separating the real and imaginary parts, one can obtain 2α real solutions:

$$\left.\begin{array}{l} e^{px}\cos qx,\ xe^{px}\cos qx,\ x^2e^{px}\cos qx,\ \ldots,\ x^{\alpha-1}e^{px}\cos qx, \\ e^{px}\sin qx,\ xe^{px}\sin qx,\ x^2e^{px}\sin qx,\ \ldots,\ x^{\alpha-1}e^{px}\sin qx. \end{array}\right\} \quad (2.43)$$

Taking the real parts and the imaginary parts of the solutions corresponding to the conjugate root $p-qi$ of the characteristic equation, we will not get any new linearly independent solutions. Thus, to the pair of complex conjugate roots $p \pm qi$ of multiplicity α there correspond 2α linearly independent real solutions (2.43).

Example 6.

$$y^{\text{IV}} + 2y'' + y = 0.$$

The characteristic equation $k^4 + 2k^2 + 1 = 0$ or $(k^2+1)^2 = 0$ has double roots $\pm i$. Hence, the general solution is of the form

$$y = (c_1 + c_2x)\cos x + (c_3 + c_4x)\sin x.$$

2. *Euler's equations.* Equations of the form

$$a_0x^ny^{(n)} + a_1x^{(n-1)}y^{(n-1)} + \ldots + a_{n-1}xy' + a_ny = 0, \qquad (2.44)$$

where all the a_i are constants, are called *Euler's equations*. An Euler equation can be transformed, by changing the independent variable $x = e^t$ *, into a homogeneous linear equation with constant coefficients.

Indeed, as was shown on page 99, the linearity and homogeneity of an equation are preserved during transformation of an independent variable, and the coefficients become constant because

$$\frac{dy}{dx} = \frac{dy}{dt}e^{-t},$$

$$\frac{d^2y}{dt^2} = e^{-2t}\left(\frac{d^2y}{dt^2} - \frac{dy}{dt}\right),$$

$$\cdots\cdots\cdots$$

$$\frac{d^ky}{dx^k} = e^{-kt}\left(\beta_1\frac{dy}{dt} + \beta_2\frac{d^2y}{dt^2} + \ldots + \beta_k\frac{d^ky}{dt^k}\right), \qquad (2.45)$$

where all β_i are constants; upon substitution into equation (2.44) the factors e^{-kt} and $x^k = e^{kt}$ cancel.

The validity of (2.45) can readily be proved by the method of induction. Indeed, assuming that (2.45) is true and differentiating it once again with respect to x, we prove the truth of equality

* Or $x = -e^t$ if $x < 0$; for the sake of definiteness, we will henceforth consider $x > 0$.

(2.45) for $\frac{d^{k+1}y}{dx^{k+1}}$ as well:

$$\frac{d^{k+1}y}{dx^{k+1}} = e^{-(k+1)t}\left(\beta_1 \frac{d^2y}{dt^2} + \beta_2 \frac{d^3y}{dt^3} + \ldots + \beta_k \frac{d^{k+1}y}{dt^{k+1}}\right) -$$
$$- ke^{-(k+1)t}\left(\beta_1 \frac{dy}{dt} + \beta_2 \frac{d^2y}{dt^2} + \ldots + \beta_k \frac{d^ky}{dt^k}\right) =$$
$$= e^{-(k+1)t}\left(\gamma_1 \frac{dy}{dt} + \gamma_2 \frac{d^2y}{dt^2} + \ldots + \gamma_{k+1} \frac{d^{k+1}y}{dt^{k+1}}\right),$$

where all the γ_i are constants.

Thus, the validity of the formula (2.45) is proved and, consequently, the products with constant coefficients

$$x^k \frac{d^ky}{dx^k} = \beta_1 \frac{dy}{dt} + \beta_2 \frac{d^2y}{dt^2} + \ldots + \beta_k \frac{d^ky}{dt^k}$$

that linearly enter into the Euler equation

$$\sum_{k=0}^{n} a_{n-k} x^k \frac{d^ky}{dx^k} = 0 \tag{2.44'}$$

are linearly (with constant coefficients) expressed in terms of the derivatives of the function y with respect to the new independent variable t. From this it follows that the transformed equation will be a homogeneous linear equation with constant coefficients:

$$b_0 \frac{d^ny}{dt^n} + b_1 \frac{d^{n-1}y}{dt^{n-1}} + \ldots + b_{n-1} \frac{dy}{dt} + b_n y = 0. \tag{2.46}$$

Instead of transforming the Euler equation into a linear equation with constant coefficients, the particular solutions of which are of the form $y = e^{kt}$, it is possible immediately to seek the solution of the original equation in the form $y = x^k$, since

$$e^{kt} = x^k.$$

The resulting equation (after cancelling x^k)

$$a_0 k(k-1)\ldots(k-n+1) + a_1 k(k-1)\ldots(k-n+2) + \ldots$$
$$\ldots + a_n = 0 \tag{2.47}$$

for a determination of k should coincide with the characteristic equation of the transformed equation (2.46). Hence, to the roots k_i of (2.47) of multiplicity α_i there correspond the solutions

$$e^{k_i t},\ te^{k_i t},\ t^2e^{k_i t},\ \ldots,\ t^{\alpha_i - 1}e^{k_i t}$$

of the transformed equation or

$$x^{k_i},\ x^{k_i}\ln x,\ x^{k_i}\ln^2 x,\ \ldots,\ x^{k_i}\ln^{\alpha_i - 1} x$$

of the original equation, and to the complex conjugate roots $p \pm qi$ of (2.47) of multiplicity α there correspond the solutions

$$e^{pt}\cos qt,\ te^{pt}\cos qt,\ \ldots,\ t^{\alpha-1}e^{pt}\cos qt,$$
$$e^{pt}\sin qt,\ te^{pt}\sin qt,\ \ldots,\ t^{\alpha-1}e^{pt}\sin qt$$

of the transformed equation or

$$x^p\cos(q\ln x),\ x^p\ln x\cos(q\ln x),\ \ldots,\ x^p\ln^{\alpha-1}x\cos(q\ln x),$$
$$x^p\sin(q\ln x),\ x^p\ln x\sin(q\ln x),\ \ldots,\ x^p\ln^{\alpha-1}x\sin(q\ln x)$$

of the original Euler equation.

Example 7.

$$x^2y''+\frac{5}{2}xy'-y=0.$$

We seek a solution in the form $y=x^k$; $k(k-1)+\frac{5}{2}k-1=0$, whence $k_1=\frac{1}{2}$, $k_2=-2$. Hence, the general solution for $x>0$ is of the form

$$y=c_1x^{\frac{1}{2}}+c_2x^{-2}.$$

Example 8.

$$x^2y''-xy'+y=0.$$

We seek a solution in the form $y=x^k$; $k(k-1)-k+1=0$, or $(k-1)^2=0$, $k_{1,2}=1$. Hence, the general solution for $x>0$ will be

$$y=(c_1+c_2\ln x)x.$$

Example 9.

$$x^2y''+xy'+y=0.$$

We seek a solution in the form $y=x^k$; $k(k-1)+k+1=0$, whence $k_{1,2}=\pm i$. Hence, the general solution for $x>0$ is of the form

$$y=c_1\cos\ln x+c_2\sin\ln x.$$

Equations of the form

$$a_0(ax+b)^n y^{(n)}+a_1(ax+b)^{n-1}y^{(n-1)}+\ldots$$
$$\ldots+a_{n-1}(ax+b)y'+a_ny=0 \qquad (2.48)$$

are also called *Euler's equations* and may be reduced to the equation (2.44) by a change of the independent variable $ax+b=x_1$. Hence, particular solutions of this equation may be sought in the form $y=(ax+b)^k$, or the equation (2.48) may be transformed to a homogeneous linear equation with constant coefficients by changing the variables $ax+b=e^t$ (or $ax+b=-e^t$, if $ax+b<0$).

5. Nonhomogeneous Linear Equations

A *nonhomogeneous linear equation* is of the form

$$a_0(x)\, y^{(n)} + a_1(x)\, y^{(n-1)} + \ldots + a_n(x)\, y = \varphi(x).$$

If $a_0(x) \neq 0$ on the interval of variation of x, then after division by $a_0(x)$ we obtain

$$y^{(n)} + p_1(x)\, y^{(n-1)} + \ldots + p_n(x)\, y = f(x). \tag{2.49}$$

We write this equation briefly as (retaining earlier notation)

$$L[y] = f(x).$$

If for $a \leqslant x \leqslant b$ all the coefficients $p_i(x)$ in equation (2.49) and the right-hand side $f(x)$ are continuous, then it has a unique solution that satisfies the conditions

$$y^{(k)}(x_0) = y_0^{(k)} \qquad (k = 0,\ 1,\ \ldots,\ n-1),$$

where $y_0^{(k)}$ are any real numbers and x_0 is any point of the interval $a < x < b$.

Indeed, the right side of the equation

$$y^{(n)} = -p_1(x)\, y^{(n-1)} - p_2(x)\, y^{(n-2)} - \ldots - p_n(x)\, y + f(x) \quad (2.49_1)$$

in the neighbourhood of the initial values under consideration satisfies the conditions of the existence and uniqueness theorem:

(1) the right-hand side is continuous with respect to all arguments;

(2) it has bounded partial derivatives with respect to all $y^{(k)}$ $(k = 0,\ 1,\ \ldots,\ n-1)$, since these derivatives are equal to the coefficients $-p_{n-k}(x)$ which by assumption are continuous on the interval $a \leqslant x \leqslant b$. Once again we observe that no restrictions are imposed on the initial values $y_0^{(k)}$.

From the two basic properties of a linear operator

$$L[cy] = cL[y],$$
$$L[y_1 + y_2] = L[y_1] + L[y_2],$$

where c is a constant, there immediately follows:

1. *The sum of* $\tilde{y} + y_1$ *of the solution* $\tilde{y}$ *of the nonhomogeneous equation*

$$L[y] = f(x) \tag{2.49}$$

and of the solution y_1 *of the corresponding homogeneous equation* $L[y] = 0$ *is a solution of the nonhomogeneous equation* (2.49).

Proof.

$$L[\tilde{y} + y_1] = L[\tilde{y}] + L[y_1],$$

but $L[\tilde{y}] \equiv f(x)$, *and* $L[y_1] \equiv 0$; *hence,*

$$L[\tilde{y}+y_1] \equiv f(x).$$

2. *If* y_i *is a solution of the equation* $L[y]=f_i(x)$ $(i=1, 2, \ldots, m)$, *then* $y=\sum_{i=1}^{m}\alpha_i y_i$ *is a solution of the equation*

$$L[y]=\sum_{i=1}^{m}\alpha_i f_i(x),$$

where the α_i *are constants.*

Proof.

$$L\left[\sum_{i=1}^{m}\alpha_i y_i\right] \equiv \sum_{i=1}^{m} L[\alpha_i y_i] \equiv \sum_{i=1}^{m}\alpha_i L[y_i], \tag{2.50}$$

but $L[y_i] \equiv f_i(x)$, *hence,*

$$L\left[\sum_{i=1}^{m}\alpha_i y_i\right] \equiv \sum_{i=1}^{m}\alpha_i f_i(x).$$

This property is often called *the principle of superposition* and obviously holds true also for $m \to \infty$ if the series $\sum_{i=1}^{\infty}\alpha_i y_i$ converges and admits of an n-fold termwise differentiation, since in this case a passage to the limit is possible in identities (2.50).

3. *If equation* $L[y]=U(x)+iV(x)$, *where all the coefficients* $p_i(x)$ *and functions* $U(x)$ *and* $V(x)$ *are real, has a solution* $y=u(x)+ iv(x)$, *then the real part of the solution* $u(x)$ *and the imaginary part* $v(x)$ *are, respectively, solutions of the equations*

$$L[y]=U(x), \quad L[y]=V(x).$$

Proof.

$$L[u+iv] \equiv U(x)+iV(x)$$

or

$$L[u]+iL[v] \equiv U(x)+iV(x).$$

Hence, separately the real parts $L[u] \equiv U(x)$ and the imaginary parts $L[v] \equiv V(x)$ are equal.

Theorem 2.8. *The general solution, on the interval* $a \leqslant x \leqslant b$, *of equation* $L[y]=f(x)$ *with continuous (on the same interval) coefficients* $p_i(x)$ *and with right side* $f(x)$ *is equal to the sum of the general solution* $\sum_{i=1}^{n}c_i y_i$ *of the corresponding homogeneous equation and of some particular solution* $\tilde{y}$ *of the nonhomogeneous equation.*

Proof. We have to prove that

$$y=\sum_{i=1}^{n}c_i y_i+\tilde{y}, \tag{2.51}$$

where c_i are arbitrary constants and y_i $(i=1, 2, \ldots, n)$ are linearly independent solutions of the corresponding homogeneous equation, is the general solution of the nonhomogeneous equation $L[y]=f(x)$. Taking into account Item (1) (see pages 119-120) and that the existence and uniqueness theorem holds with regard to this equation, it is necessary to prove that by choosing the constants c_i in (2.51) it is possible to satisfy the arbitrarily specified initial conditions

$$y^{(k)}(x_0)=y_0^{(k)} \qquad (k=0, 1, 2, \ldots, n-1), \tag{2.52}$$

where $a \leqslant x_0 \leqslant b$. By requiring that the solution (2.51) should satisfy the initial conditions (2.52), we arrive at the following system of equations

$$\left.\begin{aligned} &\sum_{i=1}^{n} c_i y_i(x_0)+\tilde{y}(x_0)=y_0, \\ &\sum_{i=1}^{n} c_i y_i'(x_0)+\tilde{y}'(x_0)=y_0', \\ &\sum_{i=1}^{n} c_i y_i''(x_0)+\tilde{y}''(x_0)=y_0'', \\ &\cdots\cdots\cdots\cdots\cdots \\ &\sum_{i=1}^{n} c_i y_i^{(n-1)}(x_0)+\tilde{y}^{(n-1)}(x_0)=y_0^{(n-1)}. \end{aligned}\right\} \tag{2.53}$$

This system of n equations, which is linear with respect to the constants c_i, in n unknowns with arbitrary right-hand sides admits of a unique solution for the c_i $(i=1, 2, \ldots, n)$, since the determinant of the system (2.53), being the Wronskian $W[y_1, y_2, \ldots, y_n]$ for the linearly independent system of solutions of the corresponding homogeneous equation, is different from zero for any values of x on the interval $a \leqslant x \leqslant b$ and, in particular, for $x=x_0$.

Hence, integration of a nonhomogeneous linear equation reduces to finding one particular solution of the equation and to integrating the corresponding homogeneous linear equation.

Example 1.

$$y''+y=x.$$

A particular solution of this equation, $y=x$, which is obviously the general solution of the corresponding homogeneous equation, is of the form

$$y=c_1\cos x+c_2\sin x \qquad \text{(see page 113, Example 4).}$$

Thus, the general solution of the original nonhomogeneous equation is

$$y=c_1\cos x+c_2\sin x+x.$$

If choice of a particular solution of the nonhomogeneous equation is difficult, but the general solution of the corresponding homogeneous equation $y=\sum_{i=1}^{n} c_i y_i$ *is found, then it is possible to integrate the nonhomogeneous linear equation by the method of variation of parameters.*

In applying this method, we seek the solution of the nonhomogeneous equation in the form $y=\sum_{i=1}^{n} c_i(x) y_i$; that is, in place of the unknown function y we actually introduce n unknown functions $c_i(x)$. Since the choice of functions $c_i(x)$ $(i=1, 2, \ldots, n)$ has to satisfy only one equation

$$y^{(n)}+p_1(x) y^{(n-1)}+\ldots+p_n(x) y=f(x), \tag{2.49}$$

we can demand that these n functions $c_i(x)$ should satisfy some other $n-1$ equations, which we choose so that the derivatives of the function $y=\sum_{i=1}^{n} c_i(x) y_i(x)$ should be, as far as possible, of the form that they have in the case of constant c_i. Choose the $c_i(x)$ so that the second sum on the right of

$$y'=\sum_{i=1}^{n} c_i(x) y_i'(x)+\sum_{i=1}^{n} c_i'(x) y_i(x)$$

should be equal to zero,

$$\sum_{i=1}^{n} c_i'(x) y_i(x)=0,$$

and, consequently,

$$y'=\sum_{i=1}^{n} c_i(x) y_i'(x),$$

that is, y' is of the same form as in the case of constant c_i. In the same fashion, we demand that the second sum in the second derivative

$$y''=\sum_{i=1}^{n} c_i(x) y''+\sum_{i=1}^{n} c_i'(x) y_i'$$

vanish and we thus subject $c_i(x)$ to the second condition:

$$\sum_{i=1}^{n} c_i'(x) y_i'=0.$$

Continuing to evaluate the derivatives of the function $y=\sum_{i=1}^{n} c_i(x) y_i$

up to order $n-1$ inclusive and demanding each time that the sum $\sum_{i=1}^{n} c_i'(x)\, y_i^{(k)}(x)$ vanish,

$$\sum_{i=1}^{n} c_i'(x)\, y_i^{(k)}(x) = 0 \qquad (k = 0, 1, 2, \ldots, n-2), \tag{2.54}$$

we get

$$\left.\begin{aligned}
y &= \sum_{i=1}^{n} c_i(x)\, y_i, \\
y' &= \sum_{i=1}^{n} c_i(x)\, y_i', \\
y'' &= \sum_{i=1}^{n} c_i(x)\, y_i'', \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
y^{(n-1)} &= \sum_{i=1}^{n} c_i(x)\, y_i^{(n-1)}, \\
y^{(n)} &= \sum_{i=1}^{n} c_i(x)\, y_i^{(n)} + \sum_{i=1}^{n} c_i'(x)\, y_i^{(n-1)}.
\end{aligned}\right\} \tag{2.55}$$

In the last equation, we cannot demand that $\sum_{i=1}^{n} c_i' y_i^{(n-1)} = 0$ since the functions $c_i(x)$ are already subject to the $n-1$ conditions (2.54), and we still have to satisfy the original equation (2.49). Substituting $y, y', \ldots, y^{(n)}$ of (2.55) into the equation

$$y^{(n)} + p_1(x)\, y^{(n-1)} + \ldots + p_n(x)\, y = f(x), \tag{2.49}$$

we get the equation we need to determine $c_i(x)\,(i = 1, 2, \ldots, n)$. And it is obvious here that in the left-hand member of (2.49) only the sum $\sum_{i=1}^{n} c_i'(x)\, y_i^{(n-1)}$ will remain, since all the other terms are of the same form as in the case of constant c_i, and when the c_i are constant the function $y = \sum_{i=1}^{n} c_i y_i$ satisfies the corresponding homogeneous equation.

We can also convince ourselves of this by means of direct calculation:

$$\sum_{i=1}^{n} c_i' y_i^{(n-1)} + \sum_{i=1}^{n} c_i y_i^{(n)} + p_1(x) \sum_{i=1}^{n} c_i y_i^{(n-1)} + p_2(x) \sum_{i=1}^{n} c_i y_i^{(n-2)} + \ldots$$
$$\ldots + p_n(x) \sum_{i=1}^{n} c_i y_i = f(x)$$

or

$$\sum_{i=1}^{n} c_i' y_i^{(n-1)} + \sum_{i=1}^{n} c_i \left[y_i^{(n)} + p_1(x) y_i^{(n-1)} + \ldots + p_n(x) y_i \right] = f(x). \quad (2.56)$$

All the y_i are particular solutions of the corresponding homogeneous equation, consequently, $y_i^{(n)} + p_1(x) y_i^{(n-1)} + \ldots + p_n(x) y_i \equiv 0$ $(i = 1, 2, \ldots, n)$ and equation (2.56) takes the form

$$\sum_{i=1}^{n} c_i' y_i^{(n-1)} = f(x).$$

To summarize, then, the functions $c_i(x)$ $(i = 1, 2, \ldots, n)$ are determined from the system of n linear equations

$$\left.\begin{aligned} &\sum_{i=1}^{n} c_i'(x) y_i = 0, \\ &\sum_{i=1}^{n} c_i'(x) y_i' = 0, \\ &\sum_{i=1}^{n} c_i'(x) y_i'' = 0, \\ &\quad \cdots\cdots\cdots \\ &\sum_{i=1}^{n} c_i'(x) y_i^{(n-2)} = 0, \\ &\sum_{i=1}^{n} c_i'(x) y_i^{(n-1)} = f(x) \end{aligned}\right\} \quad (2.57)$$

with a nonzero determinant of the system, since this determinant

$$\begin{vmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ \ldots & \ldots & \ldots & \ldots \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} \end{vmatrix}$$

is the Wronskian for linearly independent solutions of the corresponding homogeneous equation. Having determined all the $c_i'(x) = \varphi_i(x)$ from (2.57), we find, using quadratures,

$$c_i(x) = \int \varphi_i(x)\, dx + \bar{c}_i.$$

Example 2.

$$y'' + y = \frac{1}{\cos x}.$$

The general solution of the corresponding homogeneous equation is of the form $y=c_1\cos x+c_2\sin x$. We vary c_1 and c_2:

$$y=c_1(x)\cos x+c_2(x)\sin x.$$

$c_1(x)$ and $c_2(x)$ are found from the system of equations (2.57)

$$c_1'(x)\cos x+c_2'(x)\sin x=0,$$
$$-c_1'(x)\sin x+c_2'(x)\cos x=\frac{1}{\cos x},$$

whence

$$c_1'(x)=-\frac{\sin x}{\cos x},\qquad c_1(x)=\ln|\cos x|+\bar{c}_1;$$
$$c_2'(x)=1,\qquad c_2(x)=x+\bar{c}_2.$$

The general solution of the original equation is

$$y=\bar{c}_1\cos x+\bar{c}_2\sin x+\cos x\ln|\cos x|+x\sin x.$$

Example 3.

$$\ddot{x}+a^2x=f(t).$$

The general solution of the corresponding homogeneous equation is of the form $x=c_1\cos at+c_2\sin at$. Varying the constants $x=c_1(t)\cos at+c_2(t)\sin at$, we obtain

$$c_1'(t)\cos at+c_2'(t)\sin at=0,$$
$$-ac_1'(t)\sin at+ac_2'(t)\cos at=f(t),$$

and from this

$$c_1'(t)=-\frac{1}{a}f(t)\sin at,\quad c_1(t)=-\frac{1}{a}\int_0^t f(u)\sin au\,du+\bar{c}_1,$$

$$c_2'(t)=\frac{1}{a}f(t)\cos at,\quad c_2(t)=\frac{1}{a}\int_0^t f(u)\cos au\,du+\bar{c}_2,$$

$$x(t)=-\frac{\cos at}{a}\int_0^t f(u)\sin au\,du+\frac{\sin at}{a}\int_0^t f(u)\cos au\,du+$$
$$+\bar{c}_1\cos at+\bar{c}_2\sin at,$$

or

$$x(t)=\frac{1}{a}\int_0^t f(u)\,[\cos au\sin at-\sin au\cos at]\,du+\bar{c}_1\cos at+\bar{c}_2\sin at,$$

whence we finally get

$$x(t)=\frac{1}{a}\int_0^t f(u)\sin a(t-u)\,du+\bar{c}_1\cos at+\bar{c}_2\sin at.$$

Note that the first summand on the right is a particular solution of the original equation that satisfies the initial conditions $x(0)=0$, $\dot{x}(0)=0$.

Thus, a knowledge of n linearly independent particular solutions of the corresponding homogeneous equation permits us, by using the method of variation of parameters, to integrate the nonhomogeneous equation

$$L[y]=f(x).$$

Now if only k, where $k<n$, linearly independent solutions $y_1, y_2, \ldots, y_k$ of the corresponding homogeneous equation are known, then, as pointed out on pages 107-108, a change of variables permits reducing the order of the equation to $n-k$ while retaining its linearity. Observe that if $k=n-1$, then the order of the equation is reduced to the first, and a first-order linear equation can always be integrated by quadratures.

In similar fashion we can utilize the k solutions $\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_k$ of the nonhomogeneous equation, since their differences are already solutions of the corresponding homogeneous equation. Indeed,

$$L[\tilde{y}_j]\equiv f(x), \qquad L[\tilde{y}_p]\equiv f(x);$$

consequently

$$L[\tilde{y}_j-\tilde{y}_p]\equiv L[\tilde{y}_j]-L[\tilde{y}_p]\equiv f(x)-f(x)\equiv 0.$$

If the particular solutions of the corresponding homogeneous equation

$$(\tilde{y}_1-\tilde{y}_k),\quad (\tilde{y}_2-\tilde{y}_k),\quad \ldots,\quad (\tilde{y}_{k-1}-\tilde{y}_k) \tag{2.58}$$

are linearly independent, then the order of the equation $L(y)=f(x)$ may be reduced to $n-(k-1)$. Obviously, the other differences $\tilde{y}_j-\tilde{y}_p$ are linear combinations of the solutions (2.58)

$$\tilde{y}_j-\tilde{y}_p=(\tilde{y}_j-\tilde{y}_k)-(\tilde{y}_p-\tilde{y}_k)$$

and consequently cannot be employed for further reduction of the order.

There is also the so-called *Cauchy method* for finding a particular solution of a nonhomogeneous linear equation

$$L[y(x)]=f(x). \tag{2.59}$$

In this method, it is assumed that we know the solution $K(x, s)$

(dependent on a single parameter) of the corresponding homogeneous equation $L[y(x)]=0$, which solution satisfies the conditions

$$K(s, s)=K'(s, s)=\dots=K^{(n-2)}(s, s)=0; \qquad (2.60)$$

$$K^{(n-1)}(s, s)=1. \qquad (2.61)$$

It is easy to verify that in this case

$$y(x)=\int_{x_0}^{x} K(x, s) f(s)\, ds \qquad (2.62)$$

will be a particular solution of the equation (2.59), which solution satisfies the zero initial conditions

$$y(x_0)=y'(x_0)=\dots=y^{(n-1)}(x_0)=0.$$

Indeed, differentiating (2.62) and taking into account the conditions (2.60) and (2.61), we obtain

$$\left.\begin{aligned} y'(x)&=\int_{x_0}^{x} K'_x(x, s) f(s)\, ds,\\ y''(x)&=\int_{x_0}^{x} K''_x(x, s) f(s)\, ds,\\ &\dots\dots\dots\dots\dots\dots\\ y^{(n-1)}(x)&=\int_{x_0}^{x} K_x^{(n-1)}(x, s) f(s)\, ds,\\ y^{(n)}(x)&=\int_{x_0}^{x} K_x^{(n)}(x, s) f(s)\, ds+f(x) \end{aligned}\right\} \qquad (2.63)$$

Putting (2.62) and (2.63) into (2.59), we get

$$\int_{x_0}^{x} L[K(x, s)] f(s)\, ds+f(x)\equiv f(x),$$

since $K(x, s)$ is a solution of the corresponding homogeneous equation and $L[K(x, s)]\equiv 0$.

The solution $K(x, s)$ may be isolated from the general solution $y=\sum_{i=1}^{n} c_i y_i(x)$ of the homogeneous equation if the arbitrary constants c_i are chosen so as to satisfy the conditions (2.60) and (2.61).

Example 4. The general solution of the equation

$$y''+a^2y=f(x) \qquad (2.64)$$

is $y = c_1 \cos ax + c_2 \sin ax$; the conditions (2.60) and (2.61) lead to the following equations:

$$c_1 \cos as + c_2 \sin as = 0,$$
$$-ac_1 \sin as + ac_2 \cos as = 1.$$

Hence,

$$c_1 = -\frac{\sin as}{a}, \quad c_2 = \frac{\cos as}{a}$$

and the sought-for solution $K(x, s)$ is of the form

$$K(x, s) = \frac{1}{a} \sin a(x-s).$$

According to (2.62), the solution of equation (2.64) that satisfies zero initial conditions is representable as

$$y(x) = \frac{1}{a} \int_{x_0}^{x} \sin a(x-s) f(s)\, ds.$$

For $x_0 = 0$, this solution coincides with the one obtained earlier (see pages 125-126) by a different method.

We can give a physical interpretation to the function $K(x, s)$ and to the solution of the linear equation with right-hand side in the form (2.62). It will be more convenient here to denote the independent variable by t.

In many problems the solution $y(t)$ of the equation

$$y^{(n)} + p_1(t) y^{(n-1)} + \ldots + p_n(t) y = f(t) \tag{2.65}$$

describes the displacement of some system, while the function $f(t)$ describes a force acting on the system, and t is the time.

First suppose that when $t < s$ the system was at rest and its displacement is caused by a force $f_\varepsilon(t)$ that differs from zero only in the interval $s < t < s+\varepsilon$, and the momentum of this force is unity:

$$\int_{s}^{s+\varepsilon} f_\varepsilon(\tau)\, d\tau = 1.$$

Denote by $y_\varepsilon(t)$ the solution of the equation

$$y^{(n)} + p_1(t) y^{(n-1)} + \ldots + p_n(t) y = f_\varepsilon(t).$$

It is easy to verify that there exists a limit $y_\varepsilon(t)$ as $\varepsilon \to 0$ that does not depend on the choice of the function $f_\varepsilon(t)$ on the assumption that it does not change sign. Indeed,

$$y_\varepsilon(t) = \int_{t_0}^{t} K(t, s) f_\varepsilon(s)\, ds.$$

Applying the mean-value theorem for $t > s+\varepsilon$, we get

$$y_\varepsilon(t) = K(t,\ s+\varepsilon^*) \int_s^{s+\varepsilon} f_\varepsilon(\tau)\, d\tau = K(t,\ s+\varepsilon^*),$$

where $0 < \varepsilon^* < \varepsilon$; hence,

$$\lim_{\varepsilon \to 0} y_\varepsilon(t) = K(t,\ s).$$

It is therefore natural to call the function $K(t, s)$ the *influence function* of instantaneous momentum at time $t = s$.

Partitioning the interval $(t_0,\ t)$ by points s_i $(i = 0,\ 1,\ \ldots,\ m)$ into m equal parts of length $\Delta s = \frac{t-t_0}{m}$, we represent the function $f(t)$ in (2.65) as a sum of the functions $f_i(t)$, where $f_i(t)$ is different from zero only on the ith interval $s_{i-1} < t < s_i$, on which it coincides with the function $f(t)$:

$$f(t) = \sum_{i=1}^{m} f_i(t).$$

By virtue of the superposition principle (page 120) the solution of the equation (2.65) is of the form

$$y(t) = \sum_{i=1}^{m} y_i(t),$$

where $y_i(t)$ are solutions of the equations

$$y^{(n)} + p_1(t)\, y^{(n-1)} + \ldots + p_n(t)\, y = f_i(t)$$

with zero initial values. If m is sufficiently great, the solution $y_i(t)$ may be regarded as the influence function of instantaneous momentum of intensity $f_i(s_i)\,\Delta s$. Consequently,

$$y(t) \cong \sum_{i=1}^{m} K(t,\ s_i)\, f(s_i)\, \Delta s.$$

Passing to the limit as $m \to \infty$, we get the solution of the equation (2.65) with zero initial conditions in the form

$$y = \int_{t_0}^{t} K(t,\ s)\, f(s)\, ds,$$

which indicates that the effect of a constantly acting force may be regarded as the superposition of the influences of instantaneous momenta.

6. Nonhomogeneous Linear Equations with Constant Coefficients and Euler's Equations

In many cases, when solving nonhomogeneous linear equations with constant coefficients, it is possible to select without difficulty certain particular solutions and thus to reduce the problem to integration of the appropriate homogeneous equation.

For example, let the right-hand side be a polynomial of degree s and hence the equation will be of the form

$$a_0y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-1}y'+a_ny= \\ =A_0x^s+A_1x^{s-1}+\ldots+A_s, \qquad (2.66)$$

where all the a_i and A_i are constants.

If $a_n\neq 0$, then there exists a particular solution of the equation (2.66) that also has the form of a polynomial of degree s. Indeed, putting

$$y=B_0x^s+B_1x^{s-1}+\ldots+B_s$$

into equation (2.66) and comparing the coefficients of identical degrees of x in the left and right members, we get, for a determination of the coefficients B_i, the following system of linear equations which is always solvable if $a_n\neq 0$:

$$a_nB_0=A_0, \qquad B_0=\frac{A_0}{a_n},$$

$$a_nB_1+sa_{n-1}B_0=A_1,$$

whence B_1 is determined,

$$a_nB_2+(s-1)a_{n-1}B_1+s(s-1)a_{n-2}B_0=A_2,$$

whence B_2 is determined,

. . . .

$$a_nB_s+\ldots=A_s,$$

whence B_s is determined.

Thus, if $a_n\neq 0$, then there exists a particular solution in the form of a polynomial, the degree of which is equal to the degree of the polynomial in the right-hand member.

Now suppose that $a_n=0$; for the sake of generality, suppose also that $a_{n-1}=a_{n-2}=\ldots=a_{n-\alpha+1}=0$ but $a_{n-\alpha}\neq 0$, that is, $k=0$ is the α-fold root of the characteristic equation, not excepting the case of $\alpha=1$. Then equation (2.66) takes the form

$$a_0y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-\alpha}y^{(\alpha)}=A_0x^s+A_1x^{s-1}+\ldots+A_s. \qquad (2.67)$$

Assuming $y^{(\alpha)}=z$, we arrive at the preceding case and hence there exists a particular solution of the equation (2.67) for which

$$y^{(\alpha)}=B_0x_s+B_1x^{s-1}+\ldots+B_s,$$

and so y is a polynomial of degree $s+\alpha$; also, the terms beginning with degree $\alpha-1$ and lower in this polynomial will have arbitrary constant coefficients which, in the particular case, can be chosen equal to zero. Then a particular solution will have the form

$$y=x^{\alpha}(B_0x^s+B_1x^{s-1}+\ldots+B_s).$$

Example 1.

$$y''+y=x^2+x. \tag{2.68}$$

A particular solution is of the form

$$y=B_0x^2+B_1x+B_2.$$

Putting this into (2.68) and comparing coefficients of identical degrees of x, we obtain

$$B_0=1,\ B_1=1,\ B_2=-2,\ \tilde{y}=x^2+x-2.$$

The general solution is

$$y=c_1\cos x+c_2\sin x+x^2+x-2.$$

Example 2.

$$y''+y'=x-2.$$

We seek a particular solution in the form

$$y=x(B_0x+B_1).$$

Putting that in the equation and comparing the coefficients of identical degrees of x in the left-hand and right-hand members of the resulting identity, we find

$$B_0=\frac{1}{2},\quad B_1=-3,\quad \tilde{y}=x\left(\frac{1}{2}x-3\right).$$

The general solution is

$$y=c_1+c_2e^{-x}+x\left(\frac{1}{2}x-3\right).$$

Let us now consider a nonhomogeneous linear equation of the form

$$a_0y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=e^{px}(A_0x^s+A_1x^{s-1}+\ldots+A_s), \tag{2.69}$$

where all the a_j and A_j are constants. As has been shown above (see page 114), the change of variables $y=e^{px}z$ transforms equation (2.69) to the form

$$e^{px}[b_0z^{(n)}+b_1z^{(n-1)}+\ldots+b_nz]=e^{px}\ (A_0x^s+A_1x^{s-1}+\ldots+A_s)$$

or

$$b_0z^{(n)}+b_1z^{(n-1)}+\ldots+b_nz=A_0x^s+A_1x^{s-1}+\ldots+A_s, \tag{2.70}$$

where all the b_j are constants.

A particular solution of (2.70), if $b_n \neq 0$, is of the form

$$\tilde{z} = B_0x^s + B_1x^{s-1} + \ldots + B_s,$$

and, hence, a particular solution of (2.69) is

$$\tilde{y} = e^{px}(B_0x^s + B_1x^{s-1} + \ldots + B_s).$$

The condition $b_n \neq 0$ means that $\bar{k} = 0$ is not a root of the characteristic equation

$$b_0\bar{k}^n + b_1\bar{k}^{n-1} + \ldots + b_n = 0, \tag{2.71}$$

and hence $k = p$ is not a root of the characteristic equation

$$a_0k^n + a_1k^{n-1} + \ldots + a_n = 0, \tag{2.72}$$

since the roots of these characteristic equations are connected by the relationship $k = \bar{k} + p$ (see page 115).

Now if $\bar{k} = 0$ is a root of multiplicity α of the characteristic equation (2.71), in other words, if $k = p$ is a root of the same multiplicity α of the characteristic equation (2.72), then the particular solutions of the equations (2.70) and (2.69) are, respectively, of the form

$$\tilde{z} = x^\alpha (B_0x^s + B_1x^{s-1} + \ldots + B_s),$$
$$\tilde{y} = x^\alpha e^{px} (B_0x^s + B_1x^{s-1} + \ldots + B_s).$$

To summarize, then: *if the right-hand member of a linear differential equation with constant coefficients is of the form*

$$e^{px}(A_0x^s + A_1x^{s-1} + \ldots + A_s),$$

then, if p is not a root of the characteristic equation, a particular solution is to be sought in the same form:

$$\tilde{y} = e^{px}(B_0x^s + B_1x^{s-1} + \ldots + B_s).$$

But if p is a root of multiplicity α of the characteristic equation (this case is called singular or resonance), then a particular solution has to be sought in the form

$$\tilde{y} = x^\alpha e^{px}(B_0x^s + B_1x^{s-1} + \ldots + B_s).$$

Example 3.

$$y'' + 9y = e^{5x}.$$

A particular solution has to be sought in the form

$$\tilde{y} = Be^{5x}$$

Example 4.

$$y'' + y = e^{3x}(x-2).$$

A particular solution has to be sought in the form

$$\tilde{y} = e^{3x}(B_0x + B_1).$$

Example 5.

$$y'' - y = e^x(x^2 - 1).$$

A particular solution has to be sought in the form

$$\tilde{y} = xe^x(B_0x^2 + B_1x + B_2),$$

since $k = 1$ is a simple root of the characteristic equation.

Example 6.

$$y''' + 3y'' + 3y' + y = e^{-x}(x - 5).$$

A particular solution has to be sought in the form

$$\tilde{y} = x^3e^{-x}(B_0x + B_1),$$

since $k = -1$ is a triple root of the characteristic equation.

Observe that our arguments hold true for a complex p as well, therefore if the right-hand member of a linear differential equation is of the form

$$e^{px}[P_s(x)\cos qx + Q_s(x)\sin qx], \tag{2.73}$$

where one of the polynomials $P_s(x)$ or $Q_s(x)$ is of degree s, and the other is of degree not higher than s, then, transforming the trigonometric functions by Euler's formulas to the exponential form, we obtain on the right

$$e^{(p+qi)x}R_s(x) + e^{(p-qi)x}T_s(x), \tag{2.74}$$

where $R_s(x)$ and $T_s(x)$ are polynomials of degree s.

The rule mentioned above can now be applied to each term on the right, namely, if $p \pm qi$ are not roots of the characteristic equation, then a particular solution may be sought in the same form as the right-hand side of (2.74); but if $p \pm qi$ are roots of multiplicity α of the characteristic equation, then a particular solution acquires the factor x^α as well.

If we again return to trigonometric functions, this rule may be formulated as follows:

(a) *If $p \pm qi$ are not roots of the characteristic equation, then a particular solution has to be sought in the form*

$$\tilde{y} = e^{px}[\bar{P}_s(x)\cos qx + \bar{Q}_s(x)\sin qx],$$

where $\bar{P}_s(x)$ and $\bar{Q}_s(x)$ are polynomials of degree s with undetermined coefficients.

Note that if one of the *polynomials* $P_s(x)$ or $Q_s(x)$ is of degree lower than s or even, in particular, is identically zero, then, still, both the *polynomials* $\bar{P}_s(x)$ and $\bar{Q}_s(x)$ will, generally speaking, be of degree s.

(b) *If* $p \pm qi$ *are* α*-fold roots of the characteristic equation (the resonance case), a particular solution must be sought in the form*

$$\tilde{y} = x^{\alpha} e^{px} [\bar{P}_s(x) \cos qx + \bar{Q}_s(x) \sin qx].$$

Example 7.

$$y'' + 4y' + 4y = \cos 2x.$$

Since the numbers $\pm 2i$ are not roots of the characteristic equation, we seek a particular solution in the form

$$\tilde{y} = A \cos 2x + B \sin 2x.$$

Example 8.

$$y'' + 4y = \cos 2x.$$

Since the numbers $\pm 2i$ are simple roots of the characteristic equation, we seek a particular solution in the form

$$\tilde{y} = x(A \cos 2x + B \sin 2x).$$

Example 9.

$$y^{\mathrm{IV}} + 2y'' + y = \sin x.$$

Since the numbers $\pm i$ are double roots of the characteristic equation, we seek a particular solution in the form

$$\tilde{y} = x^2 (A \cos x + B \sin x).$$

Example 10.

$$y'' + 2y' + 2y = e^{-x} (x \cos x + 3 \sin x).$$

Since the numbers $-1 \pm i$ are simple roots of the characteristic equation, we seek a particular solution in the form

$$\tilde{y} = xe^{-x}[(A_0 x + A_1) \cos x + (B_0 x + B_1) \sin x].$$

In many cases it is advisable to pass to exponential functions when finding particular solutions of linear equations with constant coefficients with right-hand members of the form (2.73).

For example, in the equation

$$y'' - 2y' + y = \cos x$$

we can transform $\cos x$ by Euler's formula or, more simply, we can consider the equation

$$y'' - 2y' + y = e^{ix}, \tag{2.75}$$

the real part of the solution of which must satisfy the original equation (see page 120).

A particular solution of equation (2.75) may be sought in the form

$$y = Ae^{ix}.$$

Then

$$A = \frac{i}{2}, \quad y = \frac{i}{2}(\cos x + i \sin x).$$

A particular solution of the original equation is

$$\tilde{y}_1 = \operatorname{Re} y = -\frac{1}{2} \sin x.$$

In many cases, the operator method is very convenient for finding particular solutions of nonhomogeneous linear equations with constant coefficients.

The operator method of solving linear differential equations with constant coefficients. For derivatives of order k we introduce the notation

$$\frac{d^k y}{dx^k} = D^k y.$$

Using this notation, we write the equation

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y = f(x)$$

as

$$a_0 D^n y + a_1 D^{n-1} y + \ldots + a_n y = f(x)$$

or

$$(a_0 D^n + a_1 D^{n-1} + \ldots + a_{n-1} D + a_n)\, y = f(x). \qquad (2.76)$$

The expression

$$a_0 D^n + a_1 D^{n-1} + \ldots + a_{n-1} D + a_n$$

is called the *operator polynomial*. We denote this operator polynomial briefly as $F(D)$; the equation (2.76) can be written in the form

$$F(D)\, y = f(x).$$

It is easy to establish the truth of the following identities by direct verification:

1. $F(D)\, e^{kx} \equiv e^{kx} F(k)$,
2. $F(D^2) \sin ax \equiv \sin ax\, F(-a^2)$,
3. $F(D^2) \cos ax \equiv \cos ax\, F(-a^2)$,
4. $F(D)\, e^{kx} v(x) \equiv e^{kx} F(D+k)\, v(x)$.

Indeed:

1. $F(D)\, e^{kx} = (a_0 D^n + a_1 D^{n-1} + \ldots + a_n)\, e^{kx} =$
$= e^{kx}(a_0 k^n + a_1 k^{n-1} + \ldots + a_n) = e^{kx} F(k)$.

2. $F(D^2) \sin ax = (a_0 D^{2n} + a_1 D^{2n-2} + \ldots + a_{n-1} D^2 + a_n) \sin ax =$
$= [a_0(-a^2)^n + a_1(-a^2)^{n-1} + \ldots + a_{n-1}(-a^2) + a_n] \sin ax =$
$= \sin ax\, F(-a^2)$.

Identity (3) is proved in analogous fashion:

$$F(D^2)\cos ax = \cos ax\, F(-a^2).$$

4. $$F(D)\,e^{kx}v(x) = \sum_{p=0}^{n} a_{n-p} D^p (e^{kx} v(x)) =$$
$$= e^{kx}\sum_{p=0}^{n} a_{n-p}\Big[k^p v(x) + pk^{p-1}Dv + \frac{p(p-1)}{2!}k^{p-2}D^2 v + \ldots + D^p v\Big] =$$
$$= e^{kx}\sum_{p=0}^{n} a_{n-p}(D+k)^p v = e^{kx}F(D+k)\,v(x).$$

The sum of the operators $F_1(D)$ and $F_2(D)$ is the operator $[F_1(D)+F_2(D)]$, whose operation on a certain function $f(x)$ is determined by the equality

$$[F_1(D)+F_2(D)]\,f(x) = F_1(D)\,f(x) + F_2(D)\,f(x).$$

From this definition it follows that

$$\sum_{p=0}^{n} a_{n-p}D^p + \sum_{p=0}^{n} b_{n-p}D^p = \sum_{p=0}^{n}(a_{n-p}+b_{n-p})\,D^p,$$

since the operation of the left and right members of this equality on a certain n times differentiable function $f(x)$ leads to one and the same result, that is, the rule of adding operator polynomials does not differ from the rule of adding ordinary (nonoperator) polynomials.

The product of two operators $F_1(D)\cdot F_2(D)$ is an operator whose operation on a certain function $f(x)$ differentiable a sufficiently large number of times is determined by the equality

$$[F_1(D)\cdot F_2(D)[f(x) = F_1(D)\,[F_2(D)\,f(x)],$$

that is, the function $f(x)$ is first operated on by the right-hand factor and then the result of the operation of the right-hand factor on the function $f(x)$ is operated on by the left-hand factor.

On the basis of this definition, it is easy to see that the rule for multiplication of operator polynomials does not differ from that of ordinary (nonoperator) polynomials. Indeed,

$$\sum_{p=0}^{n} a_{n-p}D^p \sum_{q=0}^{m} b_{m-q}D^q = \sum_{p=0}^{n}\sum_{q=0}^{m} a_{n-p}b_{m-q}D^{p+q}, \qquad (2.77)$$

since

$$\sum_{p=0}^{n} a_{n-p}D^p \sum_{q=0}^{m} b_{m-q}D^q f(x) =$$

$$= \sum_{p=0}^{n} a_{n-p}D^p \left[\sum_{q=0}^{m} b_{m-q} f^{(q)}(x)\right] = \sum_{p=0}^{n} \sum_{q=0}^{m} a_{n-p}b_{m-q}f^{(p+q)}(x),$$

which coincides with the result of operating on $f(x)$ with the operator

$$\sum_{p=0}^{n} \sum_{q=0}^{m} a_{n-p}b_{m-q}D^{p+q}.$$

From (2.77), in particular, it follows that **multiplication of operators is commutative:**

$$F_1(D)\, F_2(D) = F_2(D)\, F_1(D).$$

The validity of the distributive law

$$F(D)\,[F_1(D) + F_2(D)] = F(D)\, F_1(D) + F(D)\, F_2(D)$$

follows directly from the rule of differentiating a sum. Hence, the operations of addition and multiplication of operator polynomials do not differ from the same operations involving ordinary (non-operator) polynomials.

Now let us define the operator $\frac{1}{F(D)}$.

The result of the operation of operator $\frac{1}{F(D)}$ on a certain continuous function $f(x)$ is the solution $y = \frac{1}{F(D)} f(x)$ of the equation

$$F(D)\, y = f(x), \tag{2.78}$$

Consequently,

$$F(D) \left[\frac{1}{F(D)} f(x)\right] \equiv f(x). \tag{2.79}$$

It might be considered that $\frac{1}{F(D)} f(x)$ is the solution of equation (2.78) defined by some specific, say zero, initial conditions; however, for our purposes it is more convenient to consider that $\frac{1}{F(D)} f(x)$ is one of the solutions (which one is immaterial) of equation (2.78) and, hence, the operation of the operator $\frac{1}{F(D)}$ on a certain function $f(x)$ is defined only up to a summand equal to the solution of the corresponding homogeneous equation.

In that meaning of the operation of the operator $\frac{1}{F(D)}$ the equation

$$\frac{1}{F(D)}[F(D)f(x)]=f(x) \tag{2.80}$$

will be valid, since $f(x)$ is obviously a solution of the equation

$$F(D)y=F(D)f(x).$$

The product of the operators $\Phi(D)$ by $\frac{1}{F(D)}$ is determined by the equation

$$\Phi(D)\frac{1}{F(D)}f(x)=\Phi(D)\left[\frac{1}{F(D)}f(x)\right].$$

Similarly

$$\frac{1}{F(D)}\Phi(D)f(x)=\frac{1}{F(D)}[\Phi(D)f(x)].$$

Therefore, in formulas (2.79) and (2.80) the brackets may be dropped. Also observe that

$$\frac{1}{D^p}f(x)=\int\int\ldots\int f(x)\,dx^p,$$

since $\frac{1}{D^p}f(x)$ is, by definition of the operator $\frac{1}{F(D)}$, a solution of the equation $D^py=f(x)$.

Let us verify the following properties of the operator $\frac{1}{F(D)}$:

(1) $$\frac{1}{F(D)}kf(x)=k\frac{1}{F(D)}f(x),$$

where k is a constant factor, since

$$F(D)k\frac{1}{F(D)}f(x)=kF(D)\frac{1}{F(D)}f(x)=kf(x).$$

(2) $$\frac{1}{F(D)}e^{kx}=\frac{e^{kx}}{F(k)}, \text{ if } F(k)\neq 0.$$

Indeed, $\frac{e^{kx}}{F(k)}$ is a solution of the equation $F(D)y=e^{kx}$, since by formula (1), page 135,

$$F(D)\frac{e^{kx}}{F(k)}\equiv\frac{F(k)e^{kx}}{F(k)}\equiv e^{kx}.$$

(3) $$\frac{1}{F(D^2)}\sin ax=\frac{\sin ax}{F(-a^2)}, \text{ if } F(-a^2)\neq 0.$$

Indeed, $\frac{\sin ax}{F(-a^2)}$ is a solution of the equation $F(D^2)y=\sin ax$, since by formula (2), page 135,

$$F(D^2)\frac{\sin ax}{F(-a^2)}\equiv\frac{1}{F(-a^2)}F(-a^2)\sin ax\equiv\sin ax.$$

(4) $$\frac{1}{F(D^2)}\cos ax=\frac{\cos ax}{F(-a^2)}, \quad \text{if } F(-a^2)\neq 0,$$

since by formula (3), page 135,

$$F(D^2)\frac{\cos ax}{F(-a^2)}\equiv\frac{1}{F(-a^2)}F(-a^2)\cos ax\equiv\cos ax.$$

(5) $$\frac{1}{F(D)}e^{kx}v(x)=e^{kx}\frac{1}{F(D+k)}v(x).$$

Indeed, $e^{kx}\frac{1}{F(D+k)}v(x)$ is a solution of the equation $F(D)\,y=e^{kx}v(x)$, since by formula (4), page 135,

$$F(D)\,e^{kx}\frac{1}{F(D+k)}v(x)=e^{kx}F(D+k)\frac{1}{F(D+k)}v(x)\equiv e^{kx}v(x).$$

(6) $$\frac{1}{F(D)}[f_1(x)+f_2(x)]=\frac{1}{F(D)}f_1(x)+\frac{1}{F(D)}f_2(x).$$

This equality is a corollary to the principle of superposition (page 120).

(7) $$\frac{1}{F_1(D)\cdot F_2(D)}f(x)=\frac{1}{F_1(D)}\frac{1}{F_2(D)}f(x),$$

that is,

$$y=\frac{1}{F_1(D)}\left[\frac{1}{F_2(D)}f(x)\right] \tag{2.81}$$

is a solution of the equation

$$F_1(D)\,F_2(D)\,y=f(x). \tag{2.82}$$

Indeed, substituting (2.81) into (2.82), we get

$$F_2(D)\,F_1(D)\frac{1}{F_1(D)}\left[\frac{1}{F_2(D)}f(x)\right]\equiv F_2(D)\frac{1}{F_2(D)}f(x)\equiv f(x).$$

Some examples of finding particular solutions of nonhomogeneous linear equations with constant coefficients by the operator method are given below:

(1) $y''+4y=e^x$, or $(D^2+4)\,y=e^x$, whence

$$y=\frac{1}{D^2+4}e^x=\frac{e^x}{5}.$$

(2) $y^{\text{IV}}+y=2\cos 3x$, or $(D^4+1)\,y=2\cos 3x$,

$$y=\frac{1}{D^4+1}2\cos 3x=\frac{2\cos 3x}{(-9)^2+1}=\frac{1}{41}\cos 3x.$$

(3) $y''+9y=5\sin x$, $(D^2+9)\,y=5\sin x$,

$$y=\frac{1}{D^2+9}5\sin x=\frac{5\sin x}{-1+9}=\frac{5}{8}\sin x.$$

(4) $y'' - 4y' + 4y = x^2e^{2x}, \qquad (D-2)^2 y = x^2e^{2x},$

$$y = \frac{1}{(D-2)^2} e^{2x}x^2 = e^{2x}\frac{1}{D^2}x^2 = e^{2x}\frac{x^4}{12}.$$

(5) $y''' - 3y'' + 3y' - y = e^x, \qquad (D-1)^3 y = e^x,$

$$y = \frac{1}{(D-1)^3} e^x.$$

$F(k) = 0$, and so in place of the second formula we use formula (5) (page 139), regarding e^x as the product $e^x \cdot 1$:

$$y = \frac{1}{(D-1)^3} e^x \cdot 1 = e^x \frac{1}{D^3} 1 = e^x \frac{x^3}{6}.$$

(6) $y''' - y = \sin x,$

$$(D^3 - 1) y = \sin x. \qquad (2.83)$$

$y = \frac{1}{D^3 - 1} \sin x$. Since the operator contains odd degrees of D, formula (4) cannot be employed. Therefore in place of the original equation we consider the equation $(D^3 - 1) y = e^{ix}$ or

$$(D^3 - 1) y = \cos x + i \sin x. \qquad (2.84)$$

The imaginary part of the solution of (2.84) will be the solution of the original equation (see page 120):

$$y = \frac{1}{D^3 - 1} e^{ix} = \frac{e^{ix}}{i^3 - 1} = \frac{-e^{ix}}{1+i} = \frac{(-1+i)(\cos x + i \sin x)}{2} =$$

$$= -\frac{1}{2}(\cos x + \sin x) + \frac{i}{2}(\cos x - \sin x).$$

The imaginary part of the solution $\frac{\cos x - \sin x}{2}$ of equation (2.83) is a solution of the equation (2.83).

(7) $\quad y'' + y = \cos x, \quad (D^2 + 1) y = \cos x, \quad y = \frac{1}{D^2 + 1} \cos x.$

Formula (3) (page 138) cannot be applied since $F(-a^2) = 0$; and so once again in place of the given equation we consider the equation

$$y'' + y = e^{ix} \quad \text{or} \quad y'' + y = \cos x + i \sin x$$

and take the real part of its solution

$$(D^2 + 1) y = e^{ix}, \qquad y = \frac{1}{D^2 + 1} e^{ix} = \frac{1}{(D-i)}\frac{1}{(D+i)} e^{ix} =$$

$$= \frac{1}{D-i}\frac{e^{ix}}{2i} = \frac{e^{ix}}{2i}\frac{1}{D} \cdot 1 = \frac{e^{ix}x}{2i} = \frac{x(\cos x + i \sin x)}{2i}.$$

Taking the real part of the thus found solution of the auxiliary equation $\frac{x \sin x}{2}$ we obtain the solution of the original equation,

(8) $$y^{\text{IV}} - y = e^x, \quad (D^4 - 1)\, y = e^x, \quad y = \frac{1}{D^4 - 1} e^x =$$
$$= \frac{1}{D-1} \frac{1}{(D+1)(D^2+1)} e^x = \frac{1}{D-1} \frac{e^x}{4} = \frac{1}{4} e^x \frac{1}{D} 1 = \frac{xe^x}{4}.$$

Now let us find out how the operator $\frac{1}{F(D)}$ operates on the polynomial

$$P_p(x) = A_0 x^p + A_1 x^{p-1} + \ldots + A_p.$$

We formally divide 1 by the polynomial

$$F(D) = a_n + a_{n-1} D + \ldots + a_0 D^n, \quad a_n \neq 0,$$

arranged in increasing powers of D, by the rule of division of ordinary (nonoperator) polynomials. We stop the process of division when the quotient is an operator polynomial of degree p:

$$b_0 + b_1 D + \ldots + b_p D^p = Q_p(D).$$

Then the remainder will be the polynomial

$$R(D) = c_{p+1} D^{p+1} + c_{p+2} D^{p+2} + \ldots + c_{p+n} D^{p+n},$$

which contains the operator D to powers not lower than $p+1$. By virtue of the relationship between the dividend, divisor, quotient and remainder, we get

$$F(D)\, Q_p(D) + R(D) \equiv 1. \tag{2.85}$$

This identity holds true for ordinary (nonoperator) polynomials, but since the rules of addition and multiplication of operator polynomials do not differ from the rules of addition and multiplication of ordinary polynomials, the identity also holds true for operator polynomials. Operating with the right side and left side of the identity (2.85) on the polynomial $A_0 x^p + A_1 x^{p-1} + \ldots + A_p$, we get

$$[F(D)\, Q_p(D) + R(D)]\, (A_0 x^p + A_1 x^{p-1} + \ldots + A_p) \equiv$$
$$\equiv A_0 x^p + A_1 x^{p-1} + \ldots + A_p$$

or, taking into account that

$$R(D)\, (A_0 x^p + A_1 x^{p-1} + \ldots + A_p) \equiv 0,$$

since $R(D)$ contains D to powers not lower than $p+1$, we will have

$$F(D)\, [Q_p(D)\, (A_0 x^p + A_1 x^{p-1} + \ldots + A_p)] \equiv$$
$$\equiv A_0 x^p + A_1 x^{p-1} + \ldots + A_p,$$

that is, $Q_p(D)(A_0x^p + A_1x^{p-1} + \ldots + A_p)$ is a solution of the equation

$$F(D)\,y = A_0x^p + A_1x^{p-1} + \ldots + A_p.$$

And so

$$\frac{1}{F(D)}(A_0x^p + A_1x^{p-1} + \ldots + A_p) = Q_p(D)(A_0x^p + A_1x^{p-1} + \ldots + A_p).$$

For example:

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

Dividing 1 by $1 + D^2$, we get $Q_2(D) = 1 - D^2$. Hence,

$$y = (1 - D^2)(x^2 - x + 2) = x^2 - x.$$

(10) $$y'' + 2y' + 2y = x^2e^{-x}, \qquad (D^2 + 2D + 2)\,y = x^2e^{-x},$$
$$y = \frac{1}{D^2+2D+2}x^2e^{-x} = e^{-x}\frac{1}{D^2+1}\,x^2 = e^{-x}(1 - D^2)\,x^2 = e^{-x}(x^2 - 2).$$

(11) $$y'' + y = x\cos x, \qquad (D^2+1)\,y = x\cos x.$$

Let us pass over to the equation $(D^2+1)\,y = xe^{ix}$ and then take the real part of the solution

$$y = \frac{1}{D^2+1}\,xe^{ix} = e^{ix}\frac{1}{D(D+2i)}\,x = e^{ix}\frac{1}{D}\left(\frac{1}{2i} + \frac{D}{4}\right)x =$$
$$= e^{ix}\frac{1}{D}\left(\frac{x}{2i} + \frac{1}{4}\right) = e^{ix}\left(\frac{x^2}{4i} + \frac{x}{4}\right) = (\cos x + i\sin x)\left(\frac{x^2}{4i} + \frac{x}{4}\right).$$

Taking the real part $\frac{x^2}{4}\sin x + \frac{x}{4}\cos x$, we get the desired solution.

Note. The last example indicates how one should operate on the polynomial with the operator $\frac{1}{F(D)}$ if $a_n = 0$. Representing $F(D)$ in the form $D^s\Phi(D)$, where the absolute term of the polynomial $\Phi(D)$ is no longer zero, we operate on the polynomial first with the operator $\frac{1}{\Phi(D)}$ and then with the operator $\frac{1}{D^s}$.

The nonhomogeneous *Euler equations*

$$a_0x^ny^{(n)} + a_1x^{n-1}y^{(n-1)} + \ldots + a_ny = f(x) \tag{2.86}$$

or

$$a_0(ax+b)^n\,y^{(n)} + a_1(ax+b)^{n-1}\,y^{(n-1)} + \ldots + a_ny = f(x) \tag{2.87}$$

may be integrated by solving the corresponding homogeneous equations (see page 120) and by choosing one particular solution of the nonhomogeneous equation, or by applying the method of variation of parameters. However, it is ordinarily simpler at first to integrate the homogeneous equation and, for choice of a particular solution, to transform the Euler equation (2.86) by the change of variable $x = \pm e^t$ [for equation (2.87) $ax + b = \pm e^t$] to an equation with constant coefficients for which methods of finding particular solutions have been thoroughly developed.

Example 11.

$$x^2 y''(x) - xy'(x) + y(x) = x \ln^3 x. \tag{2.88}$$

We seek the solution of the corresponding homogeneous equation in the form $y = x^k$:

$$k^2 - 2k + 1 = 0; \tag{2.89}$$

$k_{1,2} = 1$; hence, the general solution of the homogeneous equation is of the form $y = (c_1 + c_2 \ln x) x$. The change of variables $x = e^t$ transforms equation (2.88) to an equation with constant coefficients $\ddot{y}(t) - 2\dot{y}(t) + y = t^3 e^t$ [the left-hand side of this equation can straightway be written in accordance with the characteristic equation (2.89)]. Using the operator method, it is easy to find a particular solution of the transformed equation:

$$y = \frac{1}{(D-1)^2} e^t t^3 = e^t \frac{1}{D^2} t^3 = \frac{e^t t^5}{20}, \qquad y = \frac{x \ln^5 x}{20}.$$

Consequently, the general solution of equation (2.88) is of the form

$$y = \left(c_1 + c_2 \ln x + \frac{\ln^5 x}{20}\right) x.$$

7. Integration of Differential Equations by Means of Series

The problem of integrating homogeneous linear equations of the nth order

$$p_0(x) y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_n(x) y = 0 \tag{2.90}$$

reduces to choosing n or at least $n-1$ linearly independent particular solutions. However, particular solutions are readily selected only in exceptional cases. In more involved cases, particular solutions are sought in the form of a sum of a certain series $\sum_{i=1}^{\infty} \alpha_i \varphi_i(x)$, especially often in the form of the sum of a power series or a generalized power series.

The conditions under which there exist solutions in the form of the sum of a power series or a generalized power series are ordinarily established by methods of the theory of functions of a complex variable, with which we do not assume the reader is familiar, and so the basic theorems of this section are given without proof as applied to second-order equations most frequently encountered in applications.

Theorem 2.9 (on the analyticity of a solution). *If* $p_0(x)$, $p_1(x)$, $p_2(x)$ *are analytic functions of* x *in the neighbourhood of the point* $x=x_0$ *and* $p_0(x_0)\neq 0$, *then the solutions of the equation*

$$p_0(x)y''+p_1(x)y'+p_2(x)y=0 \qquad (2.91)$$

are also analytic functions in a certain neighbourhood of the same point, and, hence, the solutions of (2.91) *may be sought in the form*

$$y=a_0+a_1(x-x_0)+a_2(x-x_0)^2+\ldots+a_n(x-x_0)^n+\ldots$$

Theorem 2.10 (on the expansibility of a solution in a generalized power series). *If equation* (2.91) *satisfies the conditions of the preceding theorem, but* $x=x_0$ *is a zero of finite order* s *of the function* $p_0(x)$, *a zero of order* $s-1$ *or higher of the function* $p_1(x)$ *(if* $s>1$*) and a zero of order not lower than* $s-2$ *of the coefficient of* $p_2(x)$ *(if* $s>2$*), then there exists at least one nontrivial solution of the equation* (2.91) *in the form of a sum of the generalized power series*

$$y=a_0(x-x_0)^k+a_1(x-x_0)^{k+1}+\ldots+a_n(x-x_0)^{k+n}+\ldots, \qquad (2.92)$$

where k *is some real number that may be either integral or fractional, either positive or negative.*

The second solution, linearly independent with respect to (2.92), is also as a rule in the form of a sum of a generalized power series, but sometimes may also contain the product of the generalized power series by $\ln(x-x_0)$.

However, in specific examples one can dispense with the two theorems just formulated, all the more so since these theorems (as they are stated) do not establish the domains of convergence of the series under consideration. In concrete problems the most used procedure is to choose a power series or a generalized power series that formally satisfies the differential equation; that is, such that if substituted turns the equation (2.90) of order n into an identity if one assumes convergence of the series and the possibility of n-fold termwise differentiation. Having formally obtained a solution in the form of a series, the next step is to investigate the convergence and the possibility of n-fold termwise differentiation. In the region where the series converges and admits an n-fold termwise differentiation, it not only formally satisfies the equation, but its sum is indeed the desired solution.

Example 1.

$$y''-xy=0. \qquad (2.93)$$

We seek the solution in the form of a power series:

$$y=\sum_{n=0}^{\infty}a_nx^n.$$

Proceeding from Theorem 2.9 or formally differentiating this series termwise twice and substituting into (2.93), we get

$$\sum_{n=2}^{\infty} a_n n(n-1) x^{n-2} - x \sum_{n=0}^{\infty} a_n x^n = 0.$$

Comparing the coefficients of identical powers of x in the left-hand and right-hand members of the identity, we get: $a_2 = 0$, $3 \cdot 2a_3 - a_0 = 0$, whence $a_3 = \frac{a_0}{2 \cdot 3}$; $4 \cdot 3a_4 - a_1 = 0$, whence $a_4 = \frac{a_1}{3 \cdot 4}$; $5 \cdot 4a_5 - a_2 = 0$, whence $a_5 = \frac{a_2}{4 \cdot 5}, \ldots, n(n-1)a_n - a_{n-3} = 0$, whence $a_n = \frac{a_{n-3}}{(n-1)n}, \ldots$. Consequently

$$a_{3n-1} = 0, \quad a_{3n} = \frac{a_0}{2 \cdot 3 \cdot 5 \cdot 6 \ldots (3n-1)3n},$$

$$a_{3n+1} = \frac{a_1}{3 \cdot 4 \cdot 6 \cdot 7 \ldots 3n(3n+1)} \qquad (n = 1, 2, \ldots),$$

a_0 and a_1 remain arbitrary. Thus,

$$y = a_0 \left[1 + \frac{x^3}{2 \cdot 3} + \frac{x^6}{2 \cdot 3 \cdot 5 \cdot 6} + \ldots + \frac{x^{3n}}{2 \cdot 3 \cdot 5 \cdot 6 \ldots (3n-1)3n} + \ldots\right] +$$
$$+ a_1 \left[x + \frac{x^4}{3 \cdot 4} + \frac{x^7}{3 \cdot 4 \cdot 6 \cdot 7} + \ldots + \frac{x^{3n+1}}{3 \cdot 4 \cdot 6 \cdot 7 \ldots 3n(3n+1)} + \ldots\right]. \quad (2.94)$$

The radius of convergence of this power series is equal to infinity. Therefore, the sum of the series (2.94) is, for any values of x, a solution of the equation under consideration.

Example 2.

$$x^2 y'' + xy' + (x^2 - n^2) y = 0. \quad (2.95)$$

This equation is called *Bessel's equation* of order n, although it first appeared in the works of Euler and Bernoulli. Many problems of mathematical physics reduce to the Bessel equation, and so we shall investigate it in somewhat more detail.

By Theorem 2.10, at least one nontrivial solution of the Bessel equation can be found in the form of the sum of the generalized power series

$$y = \sum_{p=0}^{\infty} a_p x^{k+p}.$$

Differentiating this series twice term-by-term and substituting into equation (2.95), we get

$$x^2 \sum_{p=0}^{\infty} a_p (k+p)(k+p-1) x^{k+p-2} +$$
$$+ x \sum_{p=0}^{\infty} a_p (k+p) x^{k+p-1} + (x^2 - n^2) \sum_{p=0}^{\infty} a_p x^{k+p} = 0.$$

Comparing the coefficients of identical powers of x in the left and right members of the equation, we obtain

$$a_0 [k^2 - n^2] = 0,$$
$$a_1 [(k+1)^2 - n^2] = 0,$$
$$[(k+2)^2 - n^2] a_2 + a_0 = 0,$$
$$[(k+3)^2 - n^2] a_3 + a_1 = 0,$$
$$\cdots\cdots\cdots$$
$$[(k+p)^2 - n^2] a_p + a_{p-2} = 0.$$

Since the coefficient a_0 of the lowest power of x may be considered nonzero, the first equation reduces to

$$k^2 - n^2 = 0, \text{ whence } k = \pm n.$$

For definiteness we will meanwhile regard $k = n \geqslant 0$; then from the second equation $a_1 [(n+1)^2 - n^2] = 0$ we get $a_1 = 0$ and, hence, all the $a_{2p+1} = 0$,

$$a_2 = -\frac{a_0}{(n+2)^2 - n^2} = -\frac{a_0}{2^2 (n+1)},$$
$$a_4 = -\frac{a_2}{(n+4)^2 - n^2} = -\frac{a_2}{2^2 (n+2) 2} = \frac{a_0}{2^4 (n+1)(n+2) 1 \cdot 2},$$
$$\cdots\cdots\cdots$$
$$a_{2p} = \frac{(-1)^p a_0}{2^{2p} \cdot p! (n+1)(n+2) \ldots (n+p)},$$
$$\cdots\cdots\cdots$$

For $k = -n$ we get, in quite analogous fashion,

$$a_{2p+1} = 0, \quad a_{2p} = \frac{(-1)^p a_0}{2^{2p} p! (-n+1)(-n+2) \ldots (-n+p)}.$$

For $k = n$ we have the solution

$$y = a_0 \sum_{p=0}^{\infty} \frac{(-1)^p x^{2p+n}}{2^{2p} p! (n+1)(n+2) \ldots (n+p)}.$$

This solution may be written more conveniently if one takes the arbitrary constant $a_0 = \frac{1}{2^n \Gamma(n+1)}$, where Γ is Euler's *gamma function*. Recall that

$$\Gamma(p) = \int_0^{\infty} e^{-x} x^{p-1} dx \text{ for } p > 0, \quad \Gamma(p+1) = p\Gamma(p).$$

Then

$$y = \sum_{p=0}^{\infty} \frac{(-1)^p \left(\frac{x}{2}\right)^{2p+n}}{p! \, \Gamma(n+p+1)}. \tag{2.96}$$

This solution is ordinarily written as $J_n(x)$ and is called *Bessel's function of the first kind of order n.*

For $k=-n$, and for a_0 chosen as $a_0=\frac{1}{2^{-n}\Gamma(-n+1)}$, we similarly get *Bessel's function of the first kind of order* $-n$:

$$J_{-n}(x)=\sum_{p=0}^{\infty}\frac{(-1)^p\left(\frac{x}{2}\right)^{2p-n}}{p!\,\Gamma(-n+p+1)}. \tag{2.97}$$

The series (2.96) and (2.97) converge for any values of x [in (2.97) $x\neq 0$] and admit two-fold termwise differentiation; hence, $J_n(x)$, $J_{-n}(x)$ are solutions of Bessel's equation (2.95).

For nonintegral n the solutions $J_n(x)$ and $J_{-n}(x)$ are obviously linearly independent, since their expansions in series begin with different powers of x and, consequently, the linear combination $\alpha_1 J_n(x)+\alpha_2 J_{-n}(x)$ can be identically zero only when $\alpha_1=\alpha_2=0$.

Now if n is some integer, then, since for integral negative values of p and for $p=0$ the function $\Gamma(p)$ becomes infinite, expansions in series of the functions $J_n(x)$ and $J_{-n}(x)$ begin with the same powers of x and, as is readily verifiable, the functions $J_n(x)$ and $J_{-n}(x)$ will exhibit the following linear relation:

$$J_{-n}(x)=(-1)^n J_n(x).$$

Hence, when n is integer, one must seek, in place of $J_{-n}(x)$, another solution that would be linearly independent of $J_n(x)$. Such a solution may be obtained by various methods; for instance, it is possible, knowing a single particular solution of $J_n(x)$, to reduce the order of the equation (2.95) by the substitution indicated on page 107, or to seek straightway a solution in the form of the sum of a generalized power series and the product of a generalized power series into $\ln x$. The solution [linearly independent of $J_n(x)$] obtained by any one of these procedures in the case of a completely definite choice of the arbitrary constant factor is called *Bessel's function of the second kind* and is denoted as $Y_n(x)$.

However, $Y_n(x)$ is most often defined as follows: taking n nonintegral, consider the solution $Y_n(x)$ of the Bessel equation, which solution is a linear combination of the solutions $J_n(x)$ and $J_{-n}(x)$:

$$Y_n(x)=\frac{J_n(x)\cos n\pi-J_{-n}(x)}{\sin n\pi};$$

then, passing to the limit for n approaching an integer, we get a particular solution [which is linearly independent of $J_n(x)$] of the Bessel equation $Y_n(x)$, which is now defined for integral values of n as well.

Thus, the general solution of Bessel's equation for n a noninteger is of the form

$$y = c_1 J_n(x) + c_2 J_{-n}(x)$$

and for n an integer,

$$y = c_1 J_n(x) + c_2 Y_n(x),$$

where c_1 and c_2 are arbitrary constants.

Bessel's functions of the first and second kind have been studied in great detail; in particular, detailed tables of their values have been compiled. For this reason, if a problem has been reduced to Bessel's functions, then it may be considered solved to the same extent that we consider as solved a problem in which the answer is given for example in trigonometric functions.

The following equation is frequently encountered in applications:

$$x^2 y'' + xy' + (m^2 x^2 - n^2) y = 0. \tag{2.98}$$

This equation may be reduced to Bessel's equation by a change of variables $x_1 = mx$. Indeed, given such a change of variables,

$$\frac{dy}{dx} = \frac{dy}{dx_1} \frac{dx_1}{dx} = \frac{dy}{dx_1} m, \quad \frac{d^2y}{dx^2} = \frac{d^2y}{dx_1^2} m^2$$

and equation (2.98) turns into the Bessel equation:

$$x_1^2 \frac{d^2y}{dx_1^2} + x_1 \frac{dy}{dx_1} + (x_1^2 - n^2) y = 0.$$

Thus, the general solution of (2.98) for n a noninteger is of the form

$$y = c_1 J_n(mx) + c_2 J_{-n}(mx),$$

and for n an integer,

$$y = c_1 J_n(mx) + c_2 Y_n(mx).$$

Example 3.

$$x^2 y'' + xy' + \left(4x^2 - \frac{9}{25}\right) y = 0.$$

The general solution of the equation is of the form

$$y = c_1 J_{\frac{3}{5}}(2x) + c_2 J_{-\frac{3}{5}}(2x).$$

Example 4.

$$x^2 y'' + xy' + (3x^2 - 4) y = 0.$$

The general solution is

$$y = c_1 J_2(x\sqrt{3}) + c_2 Y_2(x\sqrt{3}).$$

Example 5. Integrate the equation

$$x^2y'' + xy' + \left(4x^2 - \frac{1}{9}\right)y = 0$$

provided that the solution must be continuous at the point $x=0$ and $y(0.3)=2$.

The general solution has the form

$$y = c_1 J_{\frac{1}{3}}(2x) + c_2 J_{-\frac{1}{3}}(2x).$$

The function $J_{-\frac{1}{3}}(2x)$ is discontinuous at $x=0$ since the series (2.97) begins with negative powers of x. Hence, the solution y is continuous at the point $x=0$ only for $c_2=0$:

$$y = c_1 J_{\frac{1}{3}}(2x).$$

Satisfying the second condition, $y(0.3)=2$, we get

$$c_1 = \frac{2}{J_{\frac{1}{3}}(0.6)}$$

Tables of Bessel's functions yield $J_{\frac{1}{3}}(0.6)=0.700$; thus, $c_1 \approx 2.857$ and

$$y \approx 2.857 J_{\frac{1}{3}}(2x).$$

Applied problems frequently demand finding *periodic solutions* of some differential equation. In this case it is ordinarily advisable to seek the solution in the form of the sum of some Fourier series:

$$x(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} \left(A_n \cos \frac{n\pi}{l} t + B_n \sin \frac{n\pi}{l} t\right).$$

Observe that if the equation

$$x^{(n)} = F(t,\ x,\ \dot{x},\ \ldots,\ x^{(n-1)}) \tag{2.99}$$

has a periodic solution $x_0(t)$ with period T, then the right side of (2.99) along the integral curve under consideration is a periodic function of the period T with respect to the first argument. Indeed, substituting into equation (2.99) the periodic solution $x=x_0(t)$, we get the identity

$$x_0^{(n)}(t) = F(t,\ x_0(t),\ \dot{x}_0(t),\ \ldots,\ x_0^{(n-1)}(t)).$$

If in this identity we replace t by $t+T$, we will not — by virtue of the periodicity of the function $x_0(t)$ and its derivatives — alter the left-hand member of the equation and will not change the arguments of the right-hand member beginning with the second; hence

$$F(t,\ x_0(t),\ \dot{x}_0(t),\ \ldots,\ x_0^{(n-1)}(t)) \equiv$$
$$\equiv F(t+T,\ x_0(t),\ \dot{x}_0(t),\ \ldots,\ x_0^{(n-1)}(t)),$$

that is, the function F along the integral curve $x=x_0(t)$ has a period T with respect to the explicitly appearing argument t.

Therefore, if the right side of equation (2.99) is not a periodic function [for any choice of $x_0(t)$] with respect to the first argument, then no periodic solutions exist either. If the function F is not explicitly dependent on t, that is, it is constant with respect to the argument t, then F may be regarded as a periodic function (with respect to t) of any period and therefore the existence of periodic solutions of any period whatsoever is not excluded.

For example, let it be required to find the periodic solutions of the equation

$$\ddot{x}+a^2x=f(t). \tag{2.100}$$

For a periodic solution to exist we have to assume that f is a periodic function. Without any essential loss of generality it may be taken that $f(t)$ is a periodic function with period 2π, since if the function $f(t)$ had a period T, then after transformation of the independent variable $t_1=\frac{2\pi}{T}t$ the right side would become a function with period 2π with respect to the new independent variable t_1.

Further suppose that the function $f(t)$ is continuous and can be expanded in a Fourier series:

$$f(t)=\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kt+b_k\sin kt). \tag{2.101}$$

We seek the periodic solution in the form

$$x(t)=\frac{A_0}{2}+\sum_{k=1}^{\infty}(A_k\cos kt+B_k\sin kt). \tag{2.102}$$

Formally differentiating the series (2.102) termwise twice and substituting into equation (2.100), we get

$$-\sum_{k=1}^{\infty}k^2(A_k\cos kt+B_k\sin kt)+$$
$$+a^2\left[\frac{A_0}{2}+\sum_{k=1}^{\infty}(A_k\cos kt+B_k\sin kt)\right]\equiv$$
$$\equiv\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kt+b_k\sin kt),$$

whence, if a is not an integer, we determine the coefficients of the series (2.102):

$$\left.\begin{aligned} \frac{a^2A_0}{2}=\frac{a_0}{2}, \qquad & A_0=\frac{a_0}{a^2}, \\ (a^2-k^2)\,A_k=a_k, \qquad & A_k=\frac{a_k}{a^2-k^2}, \\ (a^2-k^2)\,B_k=b_k, \qquad & B_k=\frac{b_k}{a^2-k^2}. \end{aligned}\right\} \tag{2.103}$$

Consequently, the equation (2.100) is formally satisfied by the series

$$\frac{a_0}{2a^2}+\sum_{k=1}^{\infty}\frac{a_k\cos kt+b_k\sin kt}{a^2-k^2}. \tag{2.104}$$

Obviously, series (2.104) converges and admits of twofold termwise differentiation since the series (2.101) converges uniformly by virtue of the continuity of $f(t)$, and the coefficients of the series

$$-\sum_{k=1}^{\infty}\frac{k^2(a_k\cos kt+b_k\sin kt)}{a^2-k^2}, \tag{2.105}$$

made up of the second derivatives of the terms of the series (2.104), differ from the coefficients a_k and b_k of the series (2.101) solely in the factor $-\frac{k^2}{a^2-k^2}$, which does not depend on t and monotonically approaches 1 as $k \to \infty$ (this proof is not sufficiently rigorous). Consequently, the series (2.105) converges uniformly and, hence, the series (2.104) may be differentiated termwise twice. And so the series (2.104) not only formally satisfies the equation (2.100), but its sum $x(t)$ exists and is a periodic solution of the equation (2.100).

If a differs but slightly from an integer n and $a_n \neq 0$ or $b_n \neq 0$, then *resonance* sets in; this consists in a sharp rise, as a approaches n, of at least one of the coefficients

$$A_n=\frac{a_n}{a^2-n^2}, \qquad B_n=\frac{b_n}{a^2-n^2}.$$

But if $a=n$ and at least one of the coefficients a_n or b_n is not zero, then no periodic solutions exist since to the resonance terms

$$a_n\cos nt+b_n\sin nt$$

in the right-hand member of equation (2.100), as indicated on page 134, there corresponds, in accordance with the principle of superposition, a nonperiodic term of the form

$$t\,(A_n\cos nt+B_n\sin nt),$$

in the general solution of the equation (2.100), whereas the other summands in the general solution of the equation will be periodic functions. Consequently, for $a=n$, a periodic solution of the equation (2.100) exists only if there are no resonance terms $a_n \cos nt + b_n \sin nt$ in the right-hand member, that is in the case of

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos nt\, dt = 0, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin nt\, dt = 0. \tag{2.106}$$

In the latter case, i.e. when $a=n$, $a_n=b_n=0$, a periodic solution of the equation (2.100) exists, and for $k \neq n$, the coefficients are determined from the formulas (2.103), while the coefficients A_n and B_n remain arbitrary, since $A_n \cos nt + B_n \sin nt$ is, for arbitrary A_n and B_n, a solution of the corresponding homogeneous equation.

Example 6. Determine the periodic solution of the equation

$$x + 2x = \sum_{k=1}^{\infty} \frac{\sin kt}{k^4}.$$

We seek the solution in the form of a series,

$$x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} (A_k \cos kt + B_k \sin kt)$$

and, determining the coefficients A_k and B_k from formulas (2.103), we get

$$x(t) = \sum_{k=1}^{\infty} \frac{\sin kt}{k^4(2-k^2)}.$$

Example 7. Determine the periodic solution of the equation

$$\ddot{x} + 4x = \sin^2 t.$$

Since the conditions of the existence of the periodic solution (2.106) are not satisfied,

$$\int_0^{2\pi} \sin^2 t \sin 2t\, dt = 0,$$

but

$$\int_0^{2\pi} \sin^2 t \cos 2t\, dt \neq 0,$$

no periodic solution exists.

Example 8. Determine the periodic solution of the equation

$$\ddot{x}+x=\sum_{k=2}^{\infty}\frac{\cos kt}{k^2}.$$

The resonance terms $a_1 \cos t + b_1 \sin t$ are absent in the right-hand member. Therefore, a periodic solution exists and is determined by the formulas (2.103):

$$x(t)=\sum_{k=2}^{\alpha}\frac{\cos kt}{k^2(1-k^2)}+c_1\cos t+c_2\sin t,$$

where c_1 and c_2 are arbitrary constants.

8. The Small Parameter Method and Its Application in the Theory of Quasilinear Oscillations

In the preceding section we indicated a method for finding periodic solutions of linear equations of the torm

$$\ddot{x}+a^2x=f(t).$$

In many practical problems it is necessary to find the periodic solution of an analogous equation, but having a small nonlinear term in the right-hand member:

$$\ddot{x}+a^2x=f(t)+\mu F(t,\ x,\ \dot{x},\ \mu), \tag{2.107}$$

where μ is a small parameter.

If we discard the term $\mu F(t, x, \dot{x}, \mu)$, i.e. if we consider $\mu=0$ in the equation (2.107), then we have a linear equation

$$\ddot{x}+a^2x=f(t),$$

which is known as the generating equation for (2.107).

One of the most effective methods of finding periodic solutions of an equation of nonlinear oscillations with a small nonlinearity (2.107) is that devised by Poincaré and Lyapunov—a method of expanding the solution in a series of powers of a small parameter μ, which is widely used at present in solving a great diversity of problems.

Proceeding from the theorem on the analytic dependence of a solution on a parameter (see page 60), which theorem is readily generalized to equations of the second and higher orders, it may be asserted that the solutions $x(t, \mu)$ of equation (2.107) will be analytic functions of the parameter μ for sufficiently small absolute values of μ, if the function $f(t)$ is continuous and the function

$F(t, x, \dot{x}, \mu)$, continuous with respect to t, is analytically dependent on the remaining arguments: on x and $\dot{x}$ in the region in which these variables will in future continue to vary, and on μ for sufficiently small absolute values of μ.

Assuming that these conditions are fulfilled, we seek the periodic solution $x(t, \mu)$ in the form of the sum of the series

$$x(t, \mu) = x_0(t) + \mu x_1(t) + \mu^2 x_2(t) + \ldots + \mu^n x_n(t) + \ldots .$$

We differentiate this series twice term by term:

$$\dot{x}(t, \mu) = \dot{x}_0(t) + \mu \dot{x}_1(t) + \ldots + \mu^n \dot{x}_n(t) + \ldots,$$
$$\ddot{x}(t, \mu) = \ddot{x}_0(t) + \mu \ddot{x}_1(t) + \ldots + \mu^n \ddot{x}_n(t) + \ldots,$$

and substitute into equation (2.107), in which the function $F(t, x, \dot{x}, \mu)$ has first been expanded in a series of powers of $x - x_0$, $\dot{x} - \dot{x}_0$ and μ:

$$\ddot{x} + a^2 x = f(t) + \mu \left[F(t, x_0, \dot{x}_0, 0) + \left(\frac{\partial F}{\partial x}\right)_{\substack{x = x_0 \\ \dot{x} = \dot{x}_0 \\ \mu = 0}} (x - x_0) + \right.$$
$$\left. + \left(\frac{\partial F}{\partial \dot{x}}\right)_{\substack{x = x_0 \\ \dot{x} = \dot{x}_0 \\ \mu = 0}} (\dot{x} - \dot{x}_0) + \left(\frac{\partial F}{\partial \mu}\right)_{\substack{x = x_0 \\ \dot{x} = \dot{x}_0 \\ \mu = 0}} \mu + \ldots \right]. \quad (2.108)$$

Comparing the coefficients of identical powers of μ in the left and right members of (2.108), we get

$$\left.\begin{aligned} &\ddot{x}_0 + a^2 x_0 = f(t), \\ &\ddot{x}_1 + a^2 x_1 = F(t, x_0, \dot{x}_0, 0), \\ &\ddot{x}_2 + a^2 x_2 = \left(\frac{\partial F}{\partial x}\right)_{\substack{x = x_0 \\ \dot{x} = \dot{x}_0 \\ \mu = 0}} x_1 + \left(\frac{\partial F}{\partial \dot{x}}\right)_{\substack{x = x_0 \\ \dot{x} = \dot{x}_0 \\ \mu = 0}} \dot{x}_1 + \left(\frac{\partial F}{\partial \mu}\right)_{\substack{x = x_0 \\ \dot{x} = \dot{x}_0 \\ \mu = 0}} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\}. \quad (2.109)$$

The first of these linear equations coincides with the generating equation. Integrating it and substituting the obtained solution $x_0(t)$ into the second equation, we again get a linear equation for determining $x_1(t)$, and so forth.

For the determination of $x_n(t)$ we also get a linear equation, since the right-hand member of this equation will contain only x_j and $\dot{x}_j$ with indices less than n because, due to the presence of the factor μ in F, the terms containing x_n and $\dot{x}_n$ on the right, and all the more so x_k and $\dot{x}_k$ with large indices, will have the factor μ to a power of at least $n+1$.

In this section we consider only the problem of finding periodic solutions, and so it is natural to impose on the right-hand side of the equation

$$\ddot{x}+a^2x=f(t)+\mu F(t, \dot{x}, x, \mu),$$

in accordance with the note on pages 149-150, yet another restriction; we require that the right-hand side be a periodic function with respect to the explicitly occurring argument t. Without any essential loss of generality, it may be taken that the smallest period of the right side, if the right side is explicitly dependent on t, is equal to 2π; then if $f(t)$ is not equal to a constant quantity, the periodic solutions of the equation (2.107), if they exist, can only have periods equal to 2π or that are multiples of 2π, given sufficiently small μ (see page 150).

To find the periodic solution of equation (2.108) in the form

$$x(t, \mu)=x_0(t)+\mu x_1(t)+\ldots+\mu^n x_n(t)+\ldots \qquad (2.110)$$

it is necessary to determine the periodic solutions $x_k(t)$ of equations (2.109). Indeed, if the solution $x(t, \mu)$ has a constant period 2π (or $2m\pi$, where m is an integer) for any sufficiently small absolute value of μ, then

$$x_0(t)+\mu x_1(t)+\ldots+\mu^n x_n(t)+\ldots \equiv x_0(t+2\pi)+$$
$$+\mu x_1(t+2\pi)+\ldots+\mu^n x_n(t+2\pi)+\ldots . \qquad (2.111)$$

Hence, the coefficients of identical powers of μ in the right and left members of the identity (2.111) must be equal, that is,

$$x_n(t)\equiv x_n(t+2\pi),$$

and this signifies the periodicity of the functions $x_n(t)$ ($n=0, 1, 2, \ldots$). That the coefficients of identical powers of μ in the left and right members of the identity (2.110) coincide may be seen, for example, by differentiating the identity (2.110) n times with respect to μ; then, assuming $\mu=0$, we get

$$x_n(2\pi+t)=x_n(t) \qquad (n=0, 1, 2, \ldots).$$

Thus we have to find the periodic solutions of the equations (2.109). Here it is advisable to consider the following cases separately.

1. *Nonresonance case: a is not an integer*. If a is a noninteger, the first of the equations (2.109) has a unique periodic solution $x_0=\varphi_0(t)$, which is found by the method of the preceding section (see page 150). Then find $x_1(t)$, $x_2(t)$, etc. by the same method.

If, using this method, we found the general term of the series (2.110), and if we established the convergence of the series and the validity of its twofold term-by-term differentiation, then the sum of the series (2.110) would be the desired periodic solution with period 2π. However, finding the general term of the series (2.110) is usually an exceedingly complicated problem, and so one has to confine oneself to computing only the first few terms of the series, which would be sufficient for an approximation of the periodic solution if we were confident that the series converges and its sum is a periodic solution.

Of great importance in this connection are the theorems of Poincaré on the existence of periodic solutions. In particular, these theorems permit finding the conditions under which there definitely exists a unique periodic solution of the equation (2.107) that approaches the periodic solution of the generating equation as $\mu \to 0$.

If the conditions of Poincaré's theorem are fulfilled and, hence, there exists a unique periodic solution of equation (2.107) that approaches a periodic solution of the generating equation as $\mu \to 0$, then the sum of the unique series with periodic coefficients (2.110), which series formally satisfies the equation (2.107), must exist and must coincide with the sought-for periodic solution. It is not necessary then to seek the general term of the series (2.110) for investigating the series for convergence and it is possible, after finding the first few terms of the series (2.110), to state that, given a small μ, their sum is approximately equal to the desired periodic solution.

Poincaré's theorems are based on information from the theory of analytic functions and are rather involved. We therefore give only the most elementary of these theorems at the end of this section, but even so it will permit us to assert that in the nonresonance case under consideration equation (2.107) always has a unique periodic solution for sufficiently small μ.

Example 1. Determine approximately the periodic solution of the equation

$$\ddot{x} + 2x = \sin t + \mu x^2,$$

where μ is a small parameter [determine two terms of the series (2.110)]. We seek the solution in the form

$$x(t, \mu) = x_0(t) + \mu x_1(t) + \ldots + \mu^n x_n(t) + \ldots .$$

We find the periodic solution of the generating equation

$$\ddot{x}_0 + 2x_0 = \sin t, \quad x_0(t) = \sin t.$$

The periodic solution of the equation

$$\ddot{x}_1 + 2x_1 = \sin^2 t \quad \text{or} \quad \ddot{x}_1 + 2x_1 = \frac{1 - \cos 2t}{2}$$

is of the form

$$x_1 = \frac{1}{4} + \frac{\cos 2t}{4}.$$

Hence, the periodic solution

$$x(t, \mu) \approx \sin t + \frac{1}{4}(1 + \cos 2t)\mu.$$

2. *Resonance case.* The small-parameter method may also be employed in the resonance case, i.e. in the case when in equation (2.107) a is equal to an integer n or tends to an integer n as $\mu \to 0$.

If in the equation (2.107) a differs but slightly from an integer n, more precisely, the difference $a^2 - n^2$ is of an order not lower than μ,

$$a^2 - n^2 = a_1\mu, \tag{2.112}$$

where a_1 is bounded as $\mu \to 0$, then the equation

$$\ddot{x} + a^2x = f(t) + \mu F(t, x, \dot{x}, \mu)$$

may be rewritten in the form

$$\ddot{x} + n^2x = f(t) + (n^2 - a^2)x + \mu F(t, x, \dot{x}, \mu),$$

whence, by virtue of (2.112),

$$\ddot{x} + n^2x = f(t) + \mu F_1(t, x, \dot{x}, \mu),$$

where the function F_1 satisfies the same conditions which by assumption are satisfied by the function F.

Hence, in the resonance case we can henceforth consider a equal to an integer:

$$\ddot{x} + n^2x = f(t) + \mu F(t, x, \dot{x}, \mu).$$

Applying the small-parameter method, we seek the periodic solution in the form of a series

$$x(t, \mu) = x_0(t) + \mu x_1(t) + \ldots + \mu^k x_k(t) + \ldots.$$

To determine the functions $x_k(t)$ we again get equations (2.109), in which $a^2 = n^2$, but in the given case the generating equation

$$\ddot{x}_0 + n^2x_0 = f(t) \tag{2.113}$$

has a periodic solution only if there are no resonance terms in the right-hand member, that is, if the conditions (see page 152)

$$\left.\begin{aligned}\int_0^{2\pi} f(t)\cos nt\,dt = 0,\\ \int_0^{2\pi} f(t)\sin nt\,dt = 0\end{aligned}\right\}\qquad(2.106)$$

are fulfilled.

If these conditions are fulfilled, then all solutions of the equation (2.113) will be periodic with a period 2π (see page 152):

$$x_0(t) = c_{10}\cos nt + c_{20}\sin nt + \varphi_0(t).$$

The function $x_1(t)$ is determined from the equation

$$\ddot{x}_1 + n^2 x_1 = F(t,\ x_0,\ x_0,\ 0).\qquad(2.114)$$

This equation also has periodic solutions only if resonance terms in the right-hand member are absent, i.e. if the following conditions are fulfilled:

$$\left.\begin{aligned}\int_0^{2\pi} F(t,\ x_0,\ x_0,\ 0)\cos nt\,dt = 0,\\ \int_0^{2\pi} F(t,\ x_0,\ \dot{x}_0,\ 0)\sin nt\,dt = 0.\end{aligned}\right\}\qquad(2.115)$$

Equations (2.115) contain c_{10} and c_{20} which, generally speaking, are determined from this system.

Let c_{10} and c_{20} satisfy the system (2.115); then all solutions of (2.114) have the period 2π:

$$x_1(t) = c_{11}\cos nt + c_{21}\sin nt + \varphi_1(t),\qquad(2.116)$$

and c_{11} and c_{21} are again determined from the two conditions of the absence of resonance terms in the following equation of (2.109):

$$\ddot{x}_2 + n^2 x_2 = \left(\frac{\partial F}{\partial x}\right)_{\substack{x=x_0\\ \dot{x}=\dot{x}_0\\ \mu=0}} x_1 + \left(\frac{\partial F}{\partial \dot{x}}\right)_{\substack{x=x_0\\ \dot{x}=\dot{x}_0\\ \mu=0}} \dot{x}_1 + \left(\frac{\partial F}{\partial \mu}\right)_{\substack{x=x_0\\ \dot{x}=\dot{x}_0\\ \mu=0}},$$

and so forth.

Consequently, not to every periodic solution

$$x_0 = c_{10}\cos nt + c_{20}\sin nt + \varphi_0(t)$$

of the generating equation but only to some [the values c_{10} and c_{20} of which satisfy the equations (2.115)] do there correspond periodic solutions of equation (2.107) for small μ. Of course, in the resonance case as well, in order to be sure [without finding the general term of the series (2.110)] that a periodic solution will be found by the indicated process it is first necessary to prove the existence theorem of periodic solutions. This observation also refers to the cases described in Items 3 and 4.

3. *Resonance of the* nth *kind*. Occasionally in systems described by the equation

$$\ddot{x}+a^2x=f(t)+\mu F(t, x, \dot{x}, \mu), \tag{2.107}$$

which satisfies the conditions given above, intensive oscillations are observed when the proper frequency differs but slightly from $1/n$, where n is an integer. This phenomenon became known as resonance of the nth kind.

From the mathematical viewpoint, this means that for a differing only slightly from $1/n$, where n is an integer greater than unity, equation (2.107) may have periodic solutions with a period $2\pi n$, which are not periodic solutions with a period 2π.

Let

$$\ddot{x}+\frac{1}{n^2}x=f(t)+\mu F(t, x, \dot{x}, \mu) \tag{2.117}$$

[if a differs only slightly from $1/n$, more precisely $a^2-\frac{1}{n^2}=\mu a_1$, where a_1 remains bounded as $\mu\to 0$, then, transposing the term $\left(a^2-\frac{1}{n^2}\right)x$ to the right side and including it in $\mu F(t, x, \dot{x}, \mu)$, we get an equation of the form of (2.117)].

We seek the periodic solution of equation (2.117) with a period $2\pi n$ in the form of a series

$$x(t, \mu)=x_0(t)+\mu x_1(t)+\ldots+\mu^n x_n(t)+\ldots. \tag{2.110}$$

Putting (2.110) into (2.117) and comparing coefficients of identical powers of μ, we get the equations (2.109) in which $a=\frac{1}{n}$. To determine $x_0(t)$ we obtain the generating equation

$$\ddot{x}_0+\frac{1}{n^2}x_0=f(t), \tag{2.118}$$

which has a periodic solution with a period $2\pi n$ only in the absence of resonance terms in the right-hand member, i.e. for

$$\int_0^{2\pi n} f(t)\cos\frac{t}{n}\,dt=0 \text{ and} \int_0^{2\pi n} f(t)\sin\frac{t}{n}\,dt=0.$$

If these conditions are fulfilled, then all solutions of the equation

(2.118) have the period $2\pi n$:

$$x_0 = c_{10} \cos \frac{t}{n} + c_{20} \sin \frac{t}{n} + \varphi_0 (t),$$

where c_{10} and c_{20} are arbitrary constants.

The equation which determines x_1,

$$\ddot{x}_1 + \frac{1}{n^2} x_1 = F(t, x_0, \dot{x}_0, \mu), \tag{2.119}$$

will have periodic solutions with a period $2\pi n$ only in the absence of resonance terms in the right members, i.e. when the conditions

$$\left. \begin{aligned} \int_0^{2\pi n} F(t, x_0, x_0, \mu) \cos \frac{t}{n}\, dt = 0, \\ \int_0^{2\pi n} F(t, x_0, \dot{x}_0, \mu) \sin \frac{t}{n}\, dt = 0 \end{aligned} \right\} \tag{2.120}$$

are fulfilled. Generally speaking, c_{10} and c_{20} are determined from these conditions.

If conditions (2.120) are satisfied, then all solutions of equation (2.119) have the period $2\pi n$:

$$x_1 = c_{11} \cos \frac{t}{n} + c_{21} \sin \frac{t}{n} + \varphi_1 (t).$$

To determine the arbitrary constants c_{11} and c_{21}, we make use of the two conditions of the absence of resonance terms in the following equation of (2.109):

$$\ddot{x}_2 + \frac{1}{n^2} x_2 = \left(\frac{\partial F}{\partial x} \right)_{\substack{x=x_0 \\ \dot{x}=\dot{x}_0 \\ \mu=0}} x_1 + \left(\frac{\partial F}{\partial x} \right)_{\substack{x=x_0 \\ \dot{x}=\dot{x}_0 \\ \mu=0}} x_1 + \left(\frac{\partial F}{\partial \mu} \right)_{\substack{x=x_0 \\ \dot{x}=\dot{x}_0 \\ \mu=0}},$$

and so forth.

4. *Autonomous case.* Let us assume that the right-hand side of equation (2.107) is not explicitly dependent on t, and the equation is of the form

$$\ddot{x} + a^2 x = \mu F(x, x, \mu), \tag{2.121}$$

where the function F satisfies the conditions posed above. At first glance it would seem that an investigation of (2.121) should be simpler than an investigation of equation (2.107), in which the right side depends on the argument t, however the absence of t in the right member of the equation actually complicates the problem.

If the right side is explicitly dependent on t, then, as has already been pointed out, the possible periods of the solutions are known, since the periods of the solutions can be only equal to, or multiples of, the period of the right-hand member along the solutions with respect to the argument t that occurs explicitly.

Now if the right-hand side does not contain t, it may be regarded as a periodic function of an arbitrary period, and, hence, the possibility remains of the existence of solutions of any period; the period of solutions, generally speaking, will be a function of the parameter μ. In view of the fact that the period of the solution $x(t, \mu)$ is, generally speaking, a function of μ, it would not be advisable to seek the solution in the form of the series

$$x(t, \mu)=x_0(t)+\mu x_1(t)+\ldots+\mu^n x_n(t)+\ldots, \qquad (2.110)$$

since each one of the functions $x_n(t)$ taken separately does not necessarily have to be a periodic function and, hence, the functions $x_n(t)$ could not be found by the methods given above. It is therefore necessary to transform equation (2.121) to a new independent variable so that the equation now has a constant period with respect to the new variable, and then seek the solution in the form of the series (2.110).

First, for purposes of simplification, we transform equation (2.121) by a change of the independent variable $t_1=at$ to the form

$$\frac{d^2x}{dt_1^2}+x=\mu F_1(x, \dot{x}, \mu). \qquad (2.122)$$

Each solution of the generating equation $x_0(t_1)=c_1\cos(t_1-t_0)$ will have a period 2π and the periodic solutions of the equation (2.122), when $\mu\neq 0$, if they exist, will have the period $2\pi+\alpha(\mu)$, and it may be proved that $\alpha(\mu)$ is an analytic function of μ for sufficiently small μ.

Expand $\alpha(\mu)$ in a series of powers of μ; then

$$2\pi+\alpha(\mu)=2\pi(1+h_1\mu+h_2\mu^2+\ldots+h_n\mu^n+\ldots), \qquad (2.123)$$

where h_j are certain constant quantities that we do not yet know.

Transform the variables so that the periodic solution $x(t, \mu)$ of the equation (2.122) has a constant period 2π and not the period $2\pi+\alpha(\mu)$. This is attained by the change of variables

$$t_1=t_2(1+h_1\mu+h_2\mu^2+\ldots+h_n\mu^n+\ldots), \qquad (2.124)$$

since, by virtue of the relationship (2.123), the new variable t_2 varies from 0 to 2π when t_1 varies from 0 to $2\pi+\alpha(\mu)$. In the process. equation (2.122) is transformed to

$$\ddot{x}_{t_2t_2}+(1+h_1\mu+\ldots+h_n\mu^n+\ldots)^2 x=$$
$$=\mu(1+h_1\mu+\ldots+h_n\mu^n+\ldots)^2 F_1(x, (1+h_1\mu+\ldots$$
$$\ldots+h_n\mu^n+\ldots)^{-1}\dot{x}_{t_2}, \mu). \qquad (2\ 125)$$

We seek the periodic solution of this equation in the form

$$x(t_2, \mu) = x_0(t_2) + \mu x_1(t_2) + \ldots + \mu^n x_n(t_2) + \ldots, \qquad (2.126)$$

where $x_n(t_2)$ are periodic functions of the argument t_2 of period 2π. Putting (2.126) into equation (2.125) and comparing the coefficients of identical powers of μ in the left and right members of the equality, we get

$$\ddot{x}_0 + x_0 = 0, \quad \text{whence} \quad x_0 = c\cos(t_2 - t_0),$$
$$\ddot{x}_1 + x_1 = -2h_1 x_0 + F_1(x_0, \dot{x}_0, 0)$$

or

$$\ddot{x}_1 + x_1 = -2h_1 c\cos(t_2 - t_0) + F_1(c\cos(t_2 - t_0), -c\sin(t_2 - t_0), 0) \qquad (2.127)$$

. . . .

For equation (2.127) to have periodic solutions, it is necessary and sufficient that resonance terms [see (2.106)] be absent in its right-hand member, that is, that

$$\left.\begin{aligned} &\int_0^{2\pi} F_1(c\cos(t_2 - t_0), -c\sin(t_2 - t_0), 0)\sin(t_2 - t_0)\,dt_2 = 0, \\ &-2h_1 c + \frac{1}{\pi}\int_0^{2\pi} F_1(c\cos(t_2 - t_0), -c\sin(t_2 - t_0), 0) \times \\ &\qquad\qquad \times \cos(t_2 - t_0)\,dt_2 = 0. \end{aligned}\right\} \qquad (2.128)$$

The first of these equations permits finding the values of c, and the second, those of h_1; having determined them, we find those solutions of the generating equation $x_0 = c\cos(t_2 - t_0)$ in the neighbourhood of which, for small μ, there appear periodic solutions of the equation (2.122); and we approximately determine the period of the desired solution

$$2\pi + \alpha(\mu) \approx 2\pi(1 + h_1\mu).$$

Knowing c and h_1, it is possible to determine $x_1(t_2)$ and, if necessary, to compute by the same method $x_2(t_2)$, $x_3(t_2)$, and so forth.

Example 2.

$$\ddot{x} + x = \mu\dot{x}(9 - x^2). \qquad (2.129)$$

Determine the solutions of the generating equation to which the periodic solutions of the equation (2.129) approach as $\mu \to 0$.

The solutions of the generating equation are of the form $x = c\cos(t - t_0)$. To determine the desired values of c, we take advantage of the first equation of (2.128):

$$\int_0^{2\pi} c\,(9-c^2\cos^2(t-t_0))\sin^2(t-t_0)\,dt=0$$

or $\pi c\left(9-\frac{c^2}{4}\right)=0$, whence $c_1=0$, $c_{2,3}=\pm 6$.

For $c_1=0$ we get the trivial solution $x\equiv 0$ of the generating equation, which remains a solution of the equation (2.129) for any μ.

For $c_{2,3}=\pm 6$, we get $x=\pm 6\cos(t-t_0)$.

Let us prove the most elementary of Poincaré's theorems on the *existence and uniqueness of a periodic solution tending* to a periodic solution of the generating equation as $\mu\to 0$, as applied to an equation of the form

$$\ddot{x}=f(t,\ x,\ \dot{x},\ \mu), \tag{2.130}$$

where the function f satisfies the conditions of the theorem on the analytic dependence of the solution upon the parameter μ for sufficiently small absolute values of μ. Besides, let us assume that the function f is explicitly dependent on t and has a period 2π with respect to t. Also assume that the generating equation $\ddot{x}=f(t,\ x,\ \dot{x},\ 0)$ has a unique periodic solution $x=\varphi_0(t)$ with a period 2π.

The solution of the equation (2.130) which satisfies the initial conditions

$$x(t_0)=\varphi_0(t_0)+\beta_0,\quad \dot{x}(t_0)=\dot{\varphi}_0(t_0)+\beta_1,$$

will be denoted by $x(t,\ \mu,\ \beta_0,\ \beta_1)$. Thus, β_0 and β_1 are deviations of the initial values of the solution $x(t,\ \mu,\ \beta_0,\ \beta_1)$ and its derivative $\dot{x}(t,\ \mu,\ \beta_0,\ \beta_1)$ from the initial values $\varphi_0(t_0)$ and $\dot{\varphi}_0(t_0)$ of the periodic solution of the generating equation.

The problem is to indicate the conditions under which for each sufficiently small absolute value of μ there exists a unique periodic solution $x(t,\ \mu,\ \beta_0,\ \beta_1)$ of the equation (2.130), which approaches a periodic solution $\varphi_0(t)$ of the generating equation as $\mu\to 0$.

If the solution $x(t,\ \mu,\ \beta_0,\ \beta_1)$ is periodic with a period 2π, then the following conditions should obviously be satisfied:

$$\left.\begin{aligned} x(2\pi,\ \mu,\ \beta_0,\ \beta_1)-x(0,\ \mu,\ \beta_0,\ \beta_1)=0,\\ \dot{x}(2\pi,\ \mu,\ \beta_0,\ \beta_1)-\dot{x}(0,\ \mu,\ \beta_0,\ \beta_1)=0.\end{aligned}\right\} \tag{2.131}$$

Denoting the left members of these equations by $\Phi_0(\mu,\ \beta_0,\ \beta_1)$ and $\Phi_1(\mu,\ \beta_0,\ \beta_1)$, respectively, we write the system (2-131) in the form

$$\left.\begin{aligned} \Phi_0(\mu,\ \beta_0,\ \beta_1)=0,\\ \Phi_1(\mu,\ \beta_0,\ \beta_1)=0.\end{aligned}\right\} \tag{2.132}$$

The conditions (2.132), called the *conditions of periodicity*, are not only necessary but also sufficient for periodicity of the solution $x(t, \mu, \beta_0, \beta_1)$ of the equation (2.130). Indeed, by virtue of the periodicity of the right side of (2.130) with respect to t, this right side takes on identical values at the points $(t, x, \dot{x})$, $(t+2\pi, x, \dot{x})$, ... in the intervals $0 \leqslant t \leqslant 2\pi$, $2\pi \leqslant t \leqslant 4\pi$, ... Thus, if at the points $t=0$ and $t=2\pi$ we specify identical initial values x_0 and $\dot{x}_0$, then they determine, in the intervals $0 \leqslant t \leqslant 2\pi$ and $2\pi \leqslant t \leqslant 4\pi$, absolutely identical integral curves (Fig. 2.2); more precisely, curves that are periodic continuations of one another.

By the theorem on implicit functions it may be asserted that if the Jacobian

$$\frac{D(\Phi_0, \Phi_1)}{D(\beta_0, \beta_1)} \neq 0$$

at the point $\mu=0$, $\beta_0=\beta_1=0$, then for every sufficiently small absolute value of μ there exists a unique pair of functions $\beta_0(\mu)$ and $\beta_1(\mu)$ that satisfy the conditions of periodicity (2.132) and that

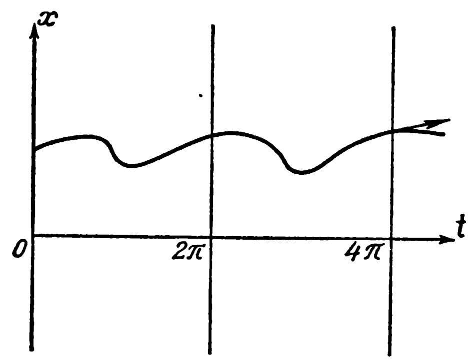


Fig. 2-2

tend to zero as $\mu \to 0$, i.e., under the indicated conditions for every sufficiently small μ there exists a unique periodic solution of the equation (2.130) that tends to the periodic solution of the generating equation as $\mu \to 0$*. It is this assertion that is the essence of Poincaré's theorem.

Example 3. Prove that in the nonresonance case, the requirements of the theorem on the existence and uniqueness of a periodic solution are fulfilled for the equation

$$\ddot{x}+a^2x=f(t)+\mu F(t, x, \dot{x}, \mu), \qquad (2.107)$$

where f and F satisfy the above-indicated conditions (see pages 153-154).

* See I. Malkin [3] for more details on existence theorems of periodic solutions.

We seek the solution $x(t, \mu, \beta_0, \beta_1)$, which is an analytic function of the last three arguments for sufficiently small values of these arguments, in the form

$$x(t, \mu, \beta_0, \beta_1) = x_0(t) + x_{11}(t)\beta_0 + x_{12}(t)\beta_1 + x_{13}(t)\mu + \dots \quad (2.133)$$

Putting (2.133) into equation (2.107) and comparing coefficients of identical powers of μ, β_0 and β_1, we get the following equations for determining x_{11} and x_{12}:

$$\left.\begin{array}{lll} \ddot{x}_{11} + a^2 x_{11} = 0, & x_{11}(0) = 1, & \dot{x}_{11}(0) = 0, \\ \ddot{x}_{12} + a^2 x_{12} = 0, & x_{12}(0) = 0, & \dot{x}_{12}(0) = 1, \end{array}\right\} \quad (2.134)$$

(the initial values are obtained from the conditions

$$x(t_0, \mu, \beta_0, \beta_1) = x_0(t_0) + \beta_0,$$
$$\dot{x}(t_0, \mu, \beta_0, \beta_1) = \dot{x}_0(t_0) + \beta_1,$$

whence

$$x_{11} = \cos at, \quad x_{12} = \frac{1}{a}\sin at.$$

The conditions of periodicity (2.132) have the form

$$(\cos 2a\pi - 1)\beta_0 + \frac{1}{a}\sin 2a\pi\beta_1 + \dots = 0,$$

$$-a\sin 2a\pi\beta_0 + (\cos 2a\pi - 1)\beta_1 + \dots = 0,$$

where the unwritten terms do not affect the magnitude of the determinant

$$\frac{D(\Phi_0, \Phi_1)}{D(\beta_0, \beta_1)} \quad \text{for} \quad \mu = \beta_0 = \beta_1 = 0.$$

The determinant

$$\left|\frac{D(\Phi_0, \Phi_1)}{D(\beta_0, \beta_1)}\right|_{\mu=\beta_0=\beta_1=0} = (\cos 2a\pi - 1)^2 + \sin^2 2a\pi$$

is not zero, since a is not an integer.

9. Boundary-Value Problems. Essentials

As was mentioned in the introduction, in addition to the basic initial-value problem, one often has to solve so-called *boundary-value problems.* In these problems the value of the sought-for function is given not at one but at two points bounding the interval on which it is required to determine the solution. For example, in the problem of the motion of a particle of mass m under the action of a given force $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ it is frequently necessary to find the law of motion if at the initial time $t = t_0$ the particle was located in a position characterized by the radius vector $\mathbf{r}_0$ and at time $t = t_1$ it has to reach point $\mathbf{r} = \mathbf{r}_1$.

The problem reduces to integrating the differential equation of motion

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}(t, \mathbf{r}, \mathbf{r})$$

with the boundary conditions $\mathbf{r}(t_0) = \mathbf{r}_0$; $\mathbf{r}(t_1) = \mathbf{r}_1$.

Note that this problem, generally speaking, does not have a unique solution; if one is speaking of a ballistic problem and about points on the earth's surface, then one and the same point may be reached by a plunging trajectory and a flat trajectory (Fig. 2.3); what is more, given very large initial velocities it is possible to reach the same point even after a single or a multiple orbiting of the globe.

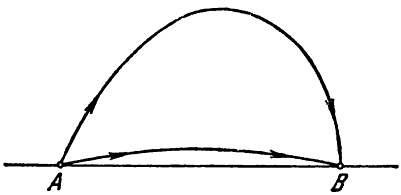


Fig. 2-3

A similar boundary-value problem may be posed for a ray of light passing through a refracting medium: to find the direction in which the ray of light must emanate from point A in order to reach another specified point B.

It is obvious that such a problem does not always have a solution and if solutions exist, then there may be several or even an infinity of solutions (for example, if the rays emanating from A are focussed at B).

If it is possible to find the general solution of the differential equation of a boundary-value problem, then to solve the problem one has to determine the arbitrary constants contained in the general solution proceeding from the boundary conditions. Of course, a real solution does not always exist, and if it does, it need not be the only one.

To illustrate the possibilities that arise here, let us consider the following boundary-value problem:

Find the solution of the equation

$$y'' + y = 0 \qquad (2.135)$$

that satisfies the conditions: $y(0) = 0$, $y(x_1) = y_1$.

The general solution of (2.135) is of the form

$$y = c_1 \cos x + c_2 \sin x.$$

The first boundary condition is satisfied for $c_1 = 0$; here $y = c_2 \sin x$.

If $x_1 \neq n\pi$, where n is an integer, then from the second boundary condition we find $y_1 = c_2 \sin x_1$, $c_2 = \frac{y_1}{\sin x_1}$. Hence, in this case the solution of the boundary-value problem is unique:

$$y = \frac{y_1}{\sin x_1} \sin x.$$

But if $x_1 = n\pi$ and $y_1 = 0$, then all the curves of the bundle $y = c_2 \sin x$ are graphs of solutions of the boundary-value problem.

For $x_1 = n\pi$, $y_1 \neq 0$, the boundary-value problem has no solutions, since not a single curve of the bundle $y = c_2 \sin x$ passes through the point (x_1, y_1) where

$$x_1 = n\pi, \quad y_1 \neq 0.$$

We consider in somewhat more detail the boundary-value problems for second-order linear equations:

$$y'' + p_1(x) y' + p_2(x) y = \varphi(x), \tag{2.136}$$

$$y(x_0) = y_0, \quad y(x_1) = y_1. \tag{2.137}$$

By the linear change of variables

$$z = y - \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) - y_0$$

the boundary conditions (2.137) reduce to zero conditions $z(x_0) = z(x_1) = 0$; the linearity of the equation (2.136) is not violated.

By multiplying by $e^{\int p_1(x)\,dx}$ the linear equation (2.136) is reduced to

$$\frac{d}{dx}(p(x) y') + q(x) y = f(x), \tag{2.138}$$

where $p(x) = e^{\int p_1(x)\,dx}$. Therefore, without any essential loss of generality it is possible to replace the study of the boundary-value problem (2.136), (2.137) by the study of the boundary-value problem for equation (2.138) with the boundary conditions

$$y(x_0) = y(x_1) = 0. \tag{2.139}$$

First consider the boundary-value problem (2.138), (2.139), where $f(x)$ is a function, localized at the point $x = s$, with unit momentum. More precisely, we consider the equation

$$\frac{d}{dx}(p(x) y') + q(x) y = f_\varepsilon(x, s) \tag{2.140}$$

with the boundary conditions $y(x_0) = y(x_1) = 0$, where the function $f_\varepsilon(x, s)$ is zero over the entire interval $[x_0, x_1]$ with the exception of the ε-neighbourhood of the point $x=s$, $s-\varepsilon<x<s+\varepsilon$, and

$$\int_{s-\varepsilon}^{s+\varepsilon} f_\varepsilon(x, s)\,dx=1.$$

Denote by $G_\varepsilon(x, s)$ the continuous solution of this boundary-value problem and pass to the limit as $\varepsilon \to 0$:

$$\lim_{\varepsilon\to 0} G_\varepsilon(x, s)=G(x, s). \tag{2.141}$$

It would not be difficult to prove the existence of this limit, which does not depend on the choice of the function $f_\varepsilon(x, s)$, but this is not necessary, since so far our reasoning has been of a heuristic nature, and on page 169 we will give a precise definition of the function $G(x, s)$.

The function $G(x, s)$ is called the *influence function* or *Green's function* of the boundary-value problem under consideration. Just as on pages 128-129 the solution of the boundary-value problem (2.138), (2.139) with continuous right side in (2.138) may be regarded as a superposition of the solutions of the boundary-value problems that correspond to functions, localized in a point, with momenta $f(s_i)\Delta s$, where the s_i are points of division of the interval $[x_0, x_1]$ into m equal parts, $\Delta s=\frac{x_1-x_0}{m}$. More precisely, an approximate solution of the boundary-value problem (2.138), (2.139) is equal to the integral sum

$$\sum_{i=1}^{m} G(x, s_i)\,f(s_i)\,\Delta s,$$

and the limit of this sum as $m\to\infty$,

$$y(x)=\int_{x_0}^{x_1} G(x, s)\,f(s)\,ds, \tag{2.142}$$

is the solution of the boundary-value problem (2.138), (2.139) at hand.

The physical meaning of the influence function $G(x, s)$ and of the solution (2.142) will become still clearer if in equation (2.140) one regards $y(x)$ as the displacement of some system under the influence of a force $f(x)$ continuously distributed over the interval $[x_0, x_1]$ [say, the deviation of a string from the equilibrium position under the effect of a distributed load with density $f(x)$]. Hence, $G(x, s)$ describes the displacement caused by a unit concentrated force applied at the point $x=s$, and the solution (2.142) is regarded as the limit of the sum of solutions corresponding to the concentrated forces.

Green's function has the following properties that follow from its definition (2.141):

1. $G(x, s)$ is continuous with respect to x for fixed s and $x_0 \leqslant x \leqslant x_1$, $x_0 < s < x_1$.

2. $G(x, s)$ is a solution of the corresponding homogeneous equation

$$\frac{d}{dx}(p(x)y') + q(x)y = 0$$

over the entire interval $[x_0, x_1]$ with the exception of the point $x = s$ (since outside this point, in the case of a function localized in the point $x = s$, the right side is zero).

3. $G(x, s)$ satisfies the boundary conditions:

$$G(x_0, s) = G(x_1, s) = 0.$$

4. At the point $x = s$ the derivative $G'_x(x, s)$ must have a discontinuity of the first kind with a jump $\frac{1}{p(s)}$. Indeed, one should expect a discontinuity only at the point of localization of the function, that is, at $x = s$. Multiplying the identity

$$\frac{d}{dx}(p(x)G'_\varepsilon(x, s)) + q(x)G_\varepsilon(x, s) \equiv f_\varepsilon(x, s)$$

by dx and integrating from $s - \varepsilon$ to $s + \varepsilon$, we get

$$p(x)G'_\varepsilon(x, s)\Big|_{s-\varepsilon}^{s+\varepsilon} + \int_{s-\varepsilon}^{s+\varepsilon} q(x)G_\varepsilon(x, s)\,dx = 1$$

and, passing to the limit as $\varepsilon \to 0$, we have

$$[G'(s+0, s) - G'(s-0, s)] = \frac{1}{p(s)}.$$

All these arguments regarding Green's function have been of a heuristic nature. Let us now invest them with the necessary rigour.

Definition. *Green's function* $G(x, s)$ *of the boundary-value problem* (2.138), (2.139) *is a function that satisfies the above-indicated conditions* (1), (2), (3), (4).

Direct substitution into equation (2.138) verifies that

$$y(x) = \int_{x_0}^{x_1} G(x, s) f(s)\,ds \qquad (2.142)$$

is a solution of this equation [the boundary conditions (2.139) are obviously satisfied by virtue of Property (3)].

Indeed,

$$y'(x)=\int_{x_0}^{x_1} G'_x(x,\ s)\,f(s)\,ds=\int_{x_0}^{x} G'_x(x,\ s)\,f(s)\,ds+\int_{x}^{x_1} G'_x(x,\ s)\,f(s)\,ds;$$

$$y''(x)=\int_{x_0}^{x} G''_{xx}(x,\ s)\,f(s)\,ds+G'_x(x,\ x-0)\,f(x)+$$

$$+\int_{x}^{x_1} G''_{xx}(x,\ s)\,f(s)\,ds-G'_x(x,\ x+0)\,f(x)=$$

$$=\int_{x_0}^{x_1} G''_{xx}(x,\ s)\,f(s)\,ds+[G'_x(x+0,\ x)-G'_x(x-0,\ x)]\,f(x).$$

Putting (2.142) into (2.138), we get

$$\int_{x_0}^{x_1} [p(x)\,G''_{xx}(x,\ s)+p'(x)\,G'_x(x,\ s)+q(x)\,G(x,\ s)]\,dx+$$

$$+p(x)\,[G'_x(x+0,\ x)-G'(x-0,\ x)]\,f(x)\equiv f(x)$$

by virtue of the conditions (2) and (4).

We now consider a method for constructing Green's function, from which we will also obtain a sufficient condition for its existence.

Consider the solution $y_1(x)$ of the equation

$$\frac{d}{dx}(p(x)\,y')+q(x)\,y=0 \tag{2.143}$$

defined by the initial conditions

$$y(x_0)=0, \qquad y'(x_0)=y'_0\neq 0.$$

This solution, generally speaking, does not satisfy the second boundary condition $y(x_1)=0$. The case $y_1(x_0)=y_1(x_1)=0$ is exceptional and we shall not consider it here.

It is obvious that the solutions $c_1y_1(x)$, where c_1 is an arbitrary constant, likewise satisfy the boundary condition $y(x_0)=0$. Similarly we find the nontrivial solution $y_2(x)$ of the equation (2.143) that satisfies the second boundary condition $y_2(x_1)=0$; this same condition is satisfied by all the solutions of the family $c_2y_2(x)$, where c_2 is an arbitrary constant.

We seek Green's function in the form

$$G(x,\ s)=\begin{cases} c_1y_1(x) & \text{for} \quad x_0\leqslant x\leqslant s, \\ c_2y_2(x) & \text{for} \quad s<x<x_1, \end{cases}$$

and we choose the constants c_1 and c_2 so that conditions (1) and (4) are fulfilled, i.e., so that the function $G(x, s)$ is continuous with respect to x for fixed s and, in particular, is continuous at the point $x=s$:

$$c_1 y_1(s) = c_2 y_2(s), \tag{2.144}$$

and so that $G'_x(x, s)$ at the point $x=s$ has a jump $\frac{1}{p(s)}$:

$$c_2 y'_2(s) - c_1 y'_1(s) = \frac{1}{p(s)}. \tag{2.145}$$

By the hypothesis that $y_1(x_1) \neq 0$, the solutions $y_1(x)$ and $y_2(x)$ are linearly independent since all solutions linearly dependent upon $y_1(x)$ are of the form $c_1 y_1(x)$ and, hence, for $c_1 \neq 0$ do not vanish at the point x_1 at which the solution $y_2(x)$ vanishes. Therefore, the determinant of the system (2.144) and (2.145), which is the Wronskian $W(y_1(x), y_2(x)) = W(x)$ at the point $x=s$, is not zero and the constants c_1 and c_2, which satisfy the system (2.144) and (2.145), are readily determined:

$$c_1 = \frac{y_2(s)\, y_1(x)}{W(s)\, p(s)}, \qquad c_2 = \frac{y_1(s)\, y_2(x)}{W(s)\, p(s)},$$

whence

$$G(x, s) = \begin{cases} \dfrac{y_2(s)\, y_1(x)}{W(s)\, p(s)} & \text{for} \quad x_0 \leqslant x \leqslant s, \\ \dfrac{y_1(s)\, y_2(x)}{W(s)\, p(s)} & \text{for} \quad s < x \leqslant x_1. \end{cases} \tag{2.146}$$

Example. Find Green's function of the boundary-value problem

$$y''(x) + y(x) = f(x), \quad y(0) = 0, \quad y\left(\frac{\pi}{2}\right) = 0.$$

The solutions of the corresponding homogeneous equation which satisfy the conditions $y(0)=0$ and $y\left(\frac{\pi}{2}\right) = 0$ have the form $y_1 = c_1 \sin x$ and $y_2 = c_2 \cos x$, respectively, and so, according to (2.146)

$$G(x, s) = \begin{cases} -\cos s \sin x & \text{for} \quad 0 \leqslant x \leqslant s, \\ -\sin s \cos x & \text{for} \quad s < x \leqslant \frac{\pi}{2}. \end{cases}$$

Note. We presumed (page 170) that there does not exist a nontrivial solution $y(x)$ of the homogeneous equation (2.143) satisfying the zero boundary conditions $y(x_0) = y(x_1) = 0$. This condition guarantees not only the existence and uniqueness of the boundary-value problem (2.138), (2.139), but also the uniqueness of Green's function.

Indeed, if one assumes the existence of two different Green's functions $G_1(x, s)$ and $G_2(x, s)$ for the boundary-value problem (2.138), (2.139), then we get two different solutions of this problem:

$$y_1(x) = \int_{x_0}^{x_1} G_1(x, s) f(s)\, ds$$

and

$$y_2(x) = \int_{x_0}^{x_1} G_2(x, s) f(s)\, ds,$$

the difference of which

$$\int_{x_0}^{x_1} [G_1(x, s) - G_2(x, s)] f(s)\, ds,$$

contrary to hypothesis, will be a nontrivial solution of the corresponding homogeneous equation, which solution satisfies the zero boundary conditions.

PROBLEMS ON CHAPTER 2

1. $y'' - 6y' + 10y = 100$, for $x = 0$, $y = 10$, $y' = 5$.
2. $\ddot{x} + x = \sin t - \cos 2t$.
3. $y'y''' - 3(y'')^2 = 0$.
4. $y'' + y = \dfrac{1}{\sin^3 x}$.
5. $x^2 y'' - 4xy' + 6y = 2$.
6. $y'' + y = \cosh x$.
7. $y'' + \dfrac{2}{1-y}(y')^2 = 0$.
8. $\dfrac{d^2x}{dt^2} - 4\dfrac{dx}{dt} + 4x = e^t + e^{2t} + 1$.
9. $(1 + x^2) y'' + (y')^2 + 1 = 0$.
10. $x^3 \dfrac{d^2x}{dt^2} + 1 = 0$.
11. $y^{IV} - 16y = x^2 - e^x$.
12. $(y''')^2 + (y'')^2 = 1$.
13. $\dfrac{d^6x}{dt^6} - \dfrac{d^4x}{dt^4} = 1$.
14. $\dfrac{d^4x}{dt^4} - 2\dfrac{d^2x}{dt^2} + x = t^2 - 3$.

15. $y''+4xy=0$; integrate by means of power series.

16. $x^2y''+xy'+\left(9x^2-\frac{1}{25}\right)y=0$; integrate by reducing to Bessel's equation.

17. $y''+(y')^2=1,\ y(0)=0,\ y'(0)=1.$

18. $y''=3\sqrt{y},\ y(0)=1,\ y'(0)=2.$

19. $y''+y=1-\frac{1}{\sin x}$.

20. $\frac{d^2u}{dr^2}+\frac{2}{r}\frac{du}{dr}=0.$

21. Find the velocity of a body falling to the earth's surface from an infinitely great height, assuming the motion is due solely to the earth's gravity. Consider the radius of the earth 6,400 km.

22. Find the law of motion of a body falling without initial velocity, assuming that the resistance of the air is proportional to the square of the velocity and that the velocity has as its limit 75 m/sec for $t\to\infty$.

23. A chain of length 6 metres slides off a table. At the initial instant of motion 1 metre of the chain was hanging from the table. How long will it take the whole chain to slide off? (Disregard friction.)

24. A chain is thrown over a smooth nail. At start of motion, one side is hanging down 8 metres, the other side, 10 metres. How long will it take the whole chain to slide off the nail? (Ignore friction.)

25. A train is in motion on a horizontal track. The train weighs P, the thrust of the locomotive is F, the force of resistance when in motion $W=a+bv$, where a and b are constants and v is the speed of the train; s is the path traversed. Determine the law of motion of the train assuming that $s=0$ and $v=0$ for $t=0$.

26. A load of p kg is suspended on a spring and has stretched it a cm. The spring is then stretched another A cm and is released without initial velocity. Find the law of motion of the spring disregarding the resistance of the medium.

27. Two identical loads are suspended from the end of a spring. Find the law of motion of one of the loads if the other breaks. It is given that the elongation of the spring under the effect of one of the loads is a cm.

28. A particle of mass m is repulsed from a centre O with a force proportional to the distance. The resistance of the medium is proportional to the velocity of motion. Find the law of motion.

29. Find the periodic solution, with period 2π, of the equation

$$\ddot{x}+2x=f(t),$$

where the function $f(t)=\pi^2 t - t^3$ for $-\pi < t \leqslant \pi$ and is continued periodically.

30. $yy'' + (y')^2 = \frac{yy'}{\sqrt{1+x^2}}$.

31. $yy'y'' = (y')^3 + (y'')^2$.

32. $\ddot{x} + 9x = t \sin 3t$.

33. $y'' + 2y' + y = \sinh x$.

34. $y''' - y = e^x$.

35. $y'' - 2y' + 2y = xe^x \cos x$.

36. $(x^2 - 1)y'' - 6y = 1$. A particular solution of the corresponding homogeneous equation has the form of a polynomial.

37. Find the solution $u = u(x^2 + y^2)$ of the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

that depends solely on $x^2 + y^2$.

38. Find the solution $u = u(x^2 + y^2 + z^2)$ of the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

which is a function of $x^2 + y^2 + z^2$.

39. A material particle is slowly sinking into a liquid. Find the law of motion on the assumption that in slow submersion the resistance of the liquid is proportional to the speed of submersion.

40. Integrate the equation of motion $m\ddot{x} = f(t\ \ x,\ \dot{x})$ on the assumption that the right side is a function only of x or only of $\dot{x}$:

(a) $m\ddot{x} = f(x)$,

(b) $m\ddot{x} = f(\dot{x})$.

41. $y^{\mathrm{VI}} - 3y^{\mathrm{V}} + 3y^{\mathrm{IV}} - y''' = x$.

42. $x^{\mathrm{IV}} + 2x'' + x = \cos t$.

43. $(1+x)^2 y'' + (1+x)y' + y = 2\cos \ln(1+x)$.

44. Determine the periodic solution of the equation

$$\ddot{x} + 2x + 2x = \sum_{n=1}^{\infty} \frac{\sin nt}{n^4}.$$

45. Find the periodic solution of the equation

$$\ddot{x} + a_1 x + a_2 x = f(t),$$

where a_1 and a_2 are constants and $f(t)$ is a continuous periodic function with period 2π that can be expanded in a Fourier series, $a_1 \neq 0$ and $a_2 \neq 0$.

46. $\ddot{x}+3x=\cos t+\mu x^2$; μ is a small parameter. Give an approximation of the periodic solution.

47. $x^3y''-xy'+y=0$; integrate the equation if $y_1=x$ is a particular solution.

48. Find the homogeneous linear equation with the following fundamental system of solutions: $y_1=x$, $y_2=\frac{1}{x}$.

49. $x^{IV}+x=t^3$.

50. $x=(y'')^3+y''+1$.

51. $\ddot{x}+10\dot{x}+25x=2^t+te^{-5t}$.

52. $xyy''-x(y')^2-yy'=0$.

53. $y^{VI}-y=e^{2x}$.

54. $y^{VI}+2y^{IV}+y''=x+e^x$.

55. $6y''y^{IV}-5(y''')^2=0$.

56. $xy''=y'\ln\frac{y'}{x}$.

57. $y''+y=\sin 3x\cos x$.

58. $y''=2y^3$, $y(1)=1$, $y'(1)=1$.

59. $yy''-(y')^2=y'$.

CHAPTER 3

Systems of differential equations

1. Fundamentals

The equation of motion of a particle of mass m under the action of a force $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$$

can be replaced by a system of three scalar equations of second order by projection on the coordinate axes:

$$m\frac{d^2x}{dt^2} = X(t, x, y, z, \dot{x}, \dot{y}, \dot{z}),$$

$$m\frac{d^2y}{dt^2} = Y(t, x, y, z, \dot{x}, \dot{y}, \dot{z}),$$

$$m\frac{d^2z}{dt^2} = Z(t, x, y, z, \dot{x}, \dot{y}, \dot{z}),$$

or a system of six equations of the first order, if for the unknown functions we take not only the coordinates x, y, z of the moving particle, but also the projections $\dot{x}, \dot{y}, \dot{z}$ of its velocity $\frac{dr}{dt}$:

$$\dot{x} = u,$$

$$\dot{y} = v,$$

$$\dot{z} = w;$$

$$m\dot{u} = X(t, x, y, z, u, v, w),$$

$$m\dot{v} = Y(t, x, y, z, u, v, w),$$

$$m\dot{w} = Z(t, x, y, z, u, v, w).$$

It is then usual to specify the unitial position of the point $x(t_0) = x_0$, $y(t_0) = y_0$, $z(t_0) = z_0$ and the initial velocity $u(t_0) = u_0$, $v(t_0) = v_0$, $w(t_0) = w_0$.

This basic problem with initial values has already been considered in Sec. 6, Chapter 1 (page 56). There, proof was given of the theorem of existence and uniqueness of the solution of a system of differential equations

$$\left.\begin{aligned}\frac{dx_1}{dt}&=f_1(t, x_1, x_2, \ldots, x_n),\\ \frac{dx_2}{dt}&=f_2(t, x_1, x_2, \ldots, x_n),\\ &\cdots\cdots\cdots\cdots\\ \frac{dx_n}{dt}&=f_n(t, x_1, x_2, \ldots, x_n),\end{aligned}\right\}\tag{3.1}$$

which satisfies the initial conditions

$$x_i(t_0)=x_{i0} \qquad (i=1, 2, \ldots, n). \tag{3.2}$$

We recall that the sufficient conditions for the existence and uniqueness of solution of a system (3.1), given initial conditions (3.2), are:

(1) continuity of all functions f_i in the neighbourhood of the initial values;

(2) fulfillment of the Lipschitz condition for all functions f_i with respect to all arguments, beginning with the second one in the same neighbourhood.

Condition (2) may be replaced by a cruder condition by requiring the existence of partial derivatives bounded in absolute value:

$$\frac{\partial f_i}{\partial x_j} \qquad (i, j=1, 2, \ldots, n).$$

The *solution* of the system of differential equations $x_1(t)$, $x_2(t)$, $\ldots$, $x_n(t)$ is an n-dimensional vector function which we will briefly designate as $X(t)$. In this notation, system (3.1) may be written as

$$\frac{dX}{dt}=F(t, X),$$

where F is a vector function with coordinates $(f_1, f_2, \ldots, f_n)$ and the initial conditions are in the form $X(t_0)=X_0$, where X_0 is an n-dimensional vector with coordinates $(x_{10}, x_{20}, \ldots, x_{n0})$.

The solutions of the system of equations

$$x_1=x_1(t), \; x_2=x_2(t), \; \ldots, \; x_n=x_n(t)$$

or, briefly, $X=X(t)$ defines in Euclidean space with coordinates $t, x_1, x_2, \ldots, x_n$ a certain curve called the *integral curve*. Upon fulfillment of the conditions (1) and (2) of the theorem of existence and uniqueness, a unique integral curve passes through every point of this space and the assemblage of such curves forms an n-parameter family. As parameters of this family, one can, for example, take the initial values $x_{10}, x_{20}, \ldots, x_{n0}$.

A different interpretation of the solutions

$$x_1 = x_1(t), \qquad x_2 = x_2(t), \ \ldots, \ x_n = x_n(t),$$

or, briefly $X = X(t)$, is possible; it is particularly convenient if the right sides of (3.1) do not depend explicitly on t.

In Euclidean space with rectangular coordinates $x_1, x_2, \ldots, x_n$ the solution $x_1 = x_1(t)$, $x_2 = x_2(t)$, $\ldots$, $x_n = x_n(t)$ defines a law of motion of some trajectory depending on the variation of the parameter t, which in this interpretation will be called the time. In such an interpretation, the derivative $\frac{dX}{dt}$ will be the velocity of motion of a point, and $\frac{dx_1}{dt}, \frac{dx_2}{dt}, \ldots, \frac{dx_n}{dt}$ will be the coordinates of the velocity of that point. Given this interpretation, which is extremely convenient and natural in many physical and mechanical problems, the system

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \ldots, x_n) \quad (i = 1, 2, \ldots, n) \tag{3.1}$$

or

$$\frac{dX}{dt} = F(t, X)$$

is ordinarily called *dynamical*, the space with coordinates x_1, $x_2, \ldots, x_n$ is called the *phase space*, and the curve $X = X(t)$ is called the *phase trajectory*.

At a specified instant of time t, the dynamical system (3.1) defines a field of velocities in the space $x_1, x_2, \ldots, x_n$. If the vector function F is explicitly dependent on t, then the field of velocities varies with time and the phase trajectories can intersect. But if the vector function F or, what is the same thing, all the functions f_i, are not dependent explicitly on t, then the field of velocities is stationary, that is to say, it does not vary with time, and the motion will be steady.

In the latter case, if the conditions of the theorem of existence and uniqueness are fulfilled, then only one trajectory will pass through each point of the phase space $(x_1, x_2, \ldots, x_n)$. Indeed, in this case an infinite number of different motions $X = X(t + c)$, where c is an arbitrary constant, occur along each trajectory $X = X(t)$; this is easy to see if we make a change of variables $t_1 = t + c$ after which the dynamical system does not change form:

$$\frac{dX}{dt_1} = F(X)$$

and consequently $X = X(t_1)$ will be its solution, or, in the old variables, $X = X(t + c)$.

If in the case at hand two trajectories passed through a certain point X_0 of the phase space,

$$X = X_1(t) \quad \text{and} \quad X = X_2(t), \quad X_1(\bar{t}_0) = X_2(\bar{\bar{t}}_0) = X_0,$$

then, taking on each of them that motion for which the point X_0 is reached at time $t = t_0$, i.e., considering the solutions

$$X = X_1(t - t_0 + \bar{t}_0) \quad \text{and} \quad X = X_2(t - t_0 + \bar{\bar{t}}_0),$$

we obtain a contradiction with the existence and uniqueness theorem, since two different solutions $X_1(t - t_0 + \bar{t}_0)$ and $X_2(t - t_0 + \bar{\bar{t}}_0)$ satisfy one and the same initial condition $X(t_0) = X_0$.

Example. The system of equations

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -x \tag{3.3}$$

has the following family of solutions (as may readily be verified by direct substitution):

$$x = c_1 \cos(t - c_2);$$
$$y = -c_1 \sin(t - c_2).$$

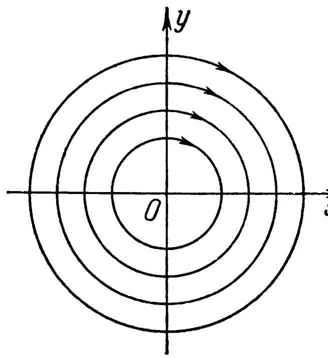


Fig. 3-1

Regarding t as a parameter, we get a family of circles on the phase plane x, y with centre at the origin of coordinates (Fig. 3.1). The right member of (3.3) is not dependent on t and satisfies the conditions of the existence and uniqueness theorem, and so the trajectories do not intersect. Fixing c_1, we get a definite trajectory, and to different c_2 there will correspond different motions along this trajectory. The equation of the trajectory $x^2 + y^2 = c_1^2$ does not depend on c_2 so that all the motions for fixed c_1 are executed along one and the same trajectory. When $c_1 = 0$ the phase trajectory consists of a single point called in this case the *rest point* of the system (3.3).

2. Integrating a System of Differential Equations by Reducing It to a Single Equation of Higher Order

One of the main methods of integrating a system of differential equations consists in the following: all unknown functions (except one) are eliminated from the equations of the system (3.1) and from the equations obtained by differentiation of the equations that make up the system; to determine this one function, a single differential equation of higher order is obtained. Integrating the equation of higher order, we find one of the unknown functions; the other unknown functions are determined (if possible without integrations) from the original equations and from the equations obtained as a result of their differentiation.

The following examples will serve as an illustration.

Example 1.

$$\frac{dx}{dt}=y, \quad \frac{dy}{dt}=x.$$

Differentiate one of these equations, for instance the first, $\frac{d^2x}{dt^2}=\frac{dy}{dt}$; eliminating $\frac{dy}{dt}$ by means of the second equation, we get $\frac{d^2x}{dt^2}-x=0$, whence $x=c_1e^t+c_2e^{-t}$. Utilizing the first equation, we get $y=\frac{dx}{dt}=c_1e^t-c_2e^{-t}$.

We determined y without integrations by means of the first equation. If we had determined y from the second equation

$$\frac{dy}{dt}=x=c_1e^t+c_2e^{-t}, \quad y=c_1e^t-c_2e^{-t}+c_3,$$

we would have introduced extraneous solutions, since direct substitution into the original system of equations shows that the system fs satisfied by the functions $x=c_1e^t+c_2e^{-t}$, $y=c_1e^t-c_2e^{-t}+c_3$ not ior an arbitrary c_3 but only for $c_3=0$.

Example 2.

$$\frac{dx}{dt}=3x-2y, \tag{3.4$_1$}$$

$$\frac{dy}{dt}=2x-y. \tag{3.4$_2$}$$

Differentiate the second equation:

$$\frac{d^2y}{dt^2}=2\frac{dx}{dt}-\frac{dy}{dt}. \tag{3.5}$$

From (3.4$_2$) and (3.5) we determine x and $\frac{dx}{dt}$:

$$x=\frac{1}{2}\left(\frac{dy}{dt}+y\right), \tag{3.6}$$

$$\frac{dx}{dt}=\frac{1}{2}\left(\frac{d^2y}{dt^2}+\frac{dy}{dt}\right).$$

Substituting into (3.4$_1$) we get

$$\frac{d^2y}{dt^2}-2\frac{dy}{dt}+y=0.$$

Integrate the resulting homogeneous linear equation with constant coefficients $y = e^t (c_1 + c_2 t)$; substituting into (3.6), we find $x(t)$:

$$x = \frac{1}{2} e^t (2c_1 + c_2 + 2c_2 t).$$

Example 3.

$$\frac{d^2x}{dt^2} = y, \quad \frac{d^2y}{dt^2} = x.$$

Differentiating the first equation, we get $\frac{d^2y}{dt^2} = \frac{d^4x}{dt^4}$ and, substituting into the second equation, we have $\frac{d^4x}{dt^4} = x$. Integrating this homogeneous linear equation with constant coefficients, we obtain

$$x = c_1 e^t + c_2 e^{-t} + c_3 \cos t + c_4 \sin t$$

and, substituting into the first equation, we find

$$y = c_1 e^t + c_2 e^{-t} - c_3 \cos t - c_4 \sin t.$$

Now we shall describe more exactly the process of eliminating all unknown functions except one from the system of equations.

First we will show that one of the unknown functions, say, $x_1(t)$, which is a component of the solution $x_1(t), x_2(t), \ldots, x_n(t)$ of the system of differential equations:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= f_1(t, x_1, x_2, \ldots, x_n), \\ \frac{dx_2}{dt} &= f_2(t, x_1, x_2, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ \frac{dx_n}{dt} &= f_n(t, x_1, x_2, \ldots, x_n), \end{aligned}\right\} \tag{3.1}$$

satisfies a certain nth order equation; here we assume that all the functions f_i have continuous partial derivatives up to order $n-1$ inclusive with respect to all arguments. Substituting some solution $x_1(t), x_2(t), \ldots, x_n(t)$ into the system (3.1), we reduce all the equations of the system to identities. In particular, the first equation of the system,

$$\frac{dx_1}{dt} = f_1(t, x_1, x_2, \ldots, x_n),$$

will be reduced to an identity.

Differentiate this identity with respect to t:

$$\frac{d^2x_1}{dt^2} = \frac{\partial f_1}{\partial t} + \sum_{i=1}^{n} \frac{\partial f_1}{\partial x_i} \frac{dx_i}{dt}$$

or

$$\frac{d^2x_1}{dt^2} = \frac{\partial f_1}{\partial t} + \sum_{i=1}^{n} \frac{\partial f_1}{\partial x_i} f_i, \qquad (3.7_2^a)$$

and designating the right side of the latter identity by $F_2(t, x_1, x_2, \ldots, x_n)$, we get

$$\frac{d^2x_1}{dt^2} = F_2(t, x_1, x_2, \ldots, x_n). \qquad (3.7_2^b)$$

Again differentiating this identity:

$$\frac{d^3x_1}{dt^3} = \frac{\partial F_2}{\partial t} + \sum_{i=1}^{n} \frac{\partial F_2}{\partial x_i} \frac{dx_i}{dt}$$

or

$$\frac{d^3x_1}{dt^3} = \frac{\partial F_2}{\partial t} + \sum_{i=1}^{n} \frac{\partial F_2}{\partial x_i} f_i, \qquad (3.7_3^a)$$

and denoting the right side of the latter identity by $F_3(t, x_1, x_2, \ldots, x_n)$, we get

$$\frac{d^3x_1}{dt^3} = F_3(t, x_1, x_2, \ldots, x_n). \qquad (3.7_3^b)$$

We again differentiate this identity; continuing the process $n-2$ times, we finally get the identity

$$\frac{d^{n-1}x_1}{dt^{n-1}} = F_{n-1}(t, x_1, x_2, \ldots, x_n), \qquad (3.7_{n-1}^b)$$

which, when differentiated once again and when the identities (3.1) are taken advantage of, yields

$$\frac{d^nx_1}{dt^n} = F_n(t, x_1, x_2, \ldots, x_n).$$

We thus have $n-1$ identities

$$\left.\begin{aligned} \frac{dx_1}{dt} &= f_1(t, x_1, x_2, \ldots, x_n), && (3.7_1) \\ \frac{d^2x_1}{dt^2} &= F_2(t, x_1, x_2, \ldots, x_n), && (3.7_2) \\ &\cdots\cdots\cdots\cdots\cdots \\ \frac{d^{n-1}x_1}{dt^{n-1}} &= F_{n-1}(t, x_1, x_2, \ldots, x_n) && (3.7_{n-1}) \end{aligned}\right\} \qquad (3.7)$$

and one more identity

$$\frac{d^nx_1}{dt^n} = F_n(t, x_1, x_2, \ldots, x_n). \qquad (3.8)$$

Suppose that in the given range of the variables the determinant

$$\frac{D(f_1, F_2, F_3, \ldots, F_{n-1})}{D(x_2, x_3, x_4, \ldots, x_n)} \neq 0.$$

Then the system (3.7) may be solved for $x_2, x_3, \ldots, x_n$ by expressing them in terms of the variables $t, x_1, \frac{dx_1}{dt}, \ldots, \frac{d^{n-1}x_1}{dt^{n-1}}$. Putting into the last equation (3.8) the variables $x_2, x_3, \ldots, x_n$ found from the system (3.7), we get an equation of the nth order:

$$\frac{d^n x_1}{dt^n} = \Phi\left(t, x_1, \frac{dx_1}{dt}, \ldots, \frac{d^{n-1}x_1}{dt^{n-1}}\right), \qquad (3.8_1)$$

which is satisfied by the function $x_1(t)$, which by hypothesis was a function $x_1(t)$ of the solution $x_1(t), x_2(t), \ldots, x_n(t)$ of the system (3.1).

Now let us prove that if we take any solution $x_1(t)$ of this nth order equation (3.8_1), put it into the system (3.7) and determine from this system $x_2(t), x_3(t), \ldots, x_n(t)$, then the system of functions

$$x_1(t), x_2(t), \ldots, x_n(t) \qquad (3.9)$$

will be the solution of the system (3.1).

We put this system of functions (3.9) into (3.7) and thus reduce all the equations of the system to identities; in particular we obtain the identity

$$\frac{dx_1}{dt} \equiv f_1(t, x_1, x_2, \ldots, x_n). \qquad (3.7_1)$$

Differentiating this identity with respect to t, we will have

$$\frac{d^2x_1}{dt^2} = \frac{\partial f_1}{\partial t} + \sum_{i=1}^{n} \frac{\partial f_1}{\partial x_i}\frac{dx_i}{dt}. \qquad (3.10)$$

In this identity it is not yet possible to replace $\frac{dx_i}{dt}$ with the functions f_i, since we have not yet proved that the functions $x_1, x_2, \ldots, x_n$ obtained by the above-mentioned method from the equation (3.8) and the system (3.7) satisfy the system (3.1); what is more, it is precisely this assertion that is the aim of our proof.

Subtracting identity (3.7_2), taken in the expanded form (3.7_2^a), termwise from the identity (3.10), we get

$$\sum_{i=1}^{n} \frac{\partial f_1}{\partial x_1}\left(\frac{dx_i}{dt} - f_i\right) \equiv 0$$

or, by virtue of (3.7_1),

$$\sum_{i=2}^{n} \frac{\partial f_1}{\partial x_i}\left(\frac{dx_i}{dt} - f_i\right) \equiv 0.$$

In analogous fashion, differentiating identity (3.7_2) and subtracting (3.7_3^a), and then differentiating identities (3.7_3) and subtracting (3.7_4^a), and so on, we obtain

$$\sum_{i=2}^{n} \frac{\partial F_2}{\partial x_i}\left(\frac{dx_i}{dt}-f_i\right)=0,$$
$$\dots\dots\dots\dots$$
$$\sum_{i=2}^{n} \frac{\partial F_{n-1}}{\partial x_i}\left(\frac{dx_i}{dt}-f_i\right)=0.$$

Since the determinant of the homogeneous linear system of equations

$$\left.\begin{aligned} &\sum_{i=2}^{n} \frac{\partial f_1}{\partial x_i}\left(\frac{dx_i}{dt}-f_i\right)=0,\\ &\sum_{i=2}^{n} \frac{\partial F_2}{\partial x_i}\left(\frac{dx_i}{dt}-f_i\right)=0,\\ &\dots\dots\dots\dots\\ &\sum_{i=2}^{n} \frac{\partial F_{n-1}}{\partial x_i}\left(\frac{dx_i}{dt}-f_i\right)=0, \end{aligned}\right\} \tag{3.11}$$

consisting of $(n-1)$ equations in $n-1$ unknowns $\left(\frac{dx_i}{dt}-f_i\right)$ $(i=2, 3, \dots, n)$ coincides with the nonzero functional determinant

$$\frac{D(f_1, F_2, \dots, F_{n-1})}{D(x_2, x_3, \dots, x_n)} \neq 0,$$

the system (3.11) has only trivial solutions at each point of the region under consideration:

$$\frac{dx_i}{dt}-f_i\equiv 0 \qquad (i=2, 3, \dots, n).$$

Taking into account also (3.7_1), we find that the n functions x_1, x_2, $\dots$, x_n are the solution of the system of equations

$$\frac{dx_i}{dt}=f_i(t, x_1, x_2, \dots, x_n) \qquad (i=1, 2, \dots, n).$$

Note. 1. This process of eliminating all functions except one presupposes that

$$\frac{D(f_1, F_2, \dots, F_{n-1})}{D(x_2, x_3, \dots, x_n)} \neq 0. \tag{3.12}$$

If this condition is not fulfilled, then the same process may be employed, but in place of the function x_1 take some other one of the functions $x_2, x_3, \ldots, x_n$ that make up the solution of the system (3.1). Now if the condition (3.12) is not fulfilled for any choice of some function $x_2, x_3, \ldots, x_n$ in place of x_1, then various exceptional cases are possible. We illustrate them in the following examples.

Example 4.

$$\frac{dx_1}{dt}=f_1(t, x_1),$$
$$\frac{dx_2}{dt}=f_2(t, x_2),$$
$$\frac{dx_3}{dt}=f_3(t, x_3).$$

The system has disintegrated into quite independent equations, each of which has to be integrated separately.

Example 5.

$$\frac{dx_1}{dt}=f_1(t, x_1),$$
$$\frac{dx_2}{dt}=f_2(t, x_2, x_3), \quad \frac{\partial f_2}{\partial x_3}\neq 0,$$
$$\frac{dx_3}{dt}=f_3(t, x_2, x_3).$$

The latter two equations may be reduced to one equation of the second order in the manner indicated above, but the first equation, which contains the unknown function x_1 that does not appear in the other equation, has to be integrated separately.

Note. 2. If we apply the above-indicated process of eliminating all unknown functions except one to the system

$$\frac{dx_i}{dt}=\sum_{j=1}^{n} a_{ij}(t)\,x_j \quad (i=1, 2, \ldots, n),$$

called a homogeneous linear system, then, as is readily verifiable, the nth order equation

$$\frac{d^n x_1}{dt^n}=\Phi\left(t, x_1, \frac{dx_1}{dt}, \ldots, \frac{d^{n-1}x_1}{dt^{n-1}}\right) \tag{3.8$_1$}$$

will also be homogeneous linear, and if all the coefficients a_{ij} were constant, then the equation (3.8$_1$) as well will be a homogeneous linear equation with constant coefficients. A similar remark holds true for the nonhomogeneous linear system

$$\frac{dx_i}{dt}=\sum_{j=1}^{n} a_{ij}(t)\,x_j+f_i(t) \quad (i=1, 2, \ldots, n),$$

for which the equation (3.8_1) will be a nonhomogeneous linear equation of the nth order.

3. Finding Integrable Combinations

Integration of the system of differential equations

$$\frac{dx_i}{dt}=f_i(t,\ x_1,\ x_2,\ \ldots,\ x_n) \quad (i=1,\ 2,\ \ldots,\ n) \tag{3.1}$$

is often accomplished by choosing so-called integrable combinations.

An *integrable combination* is a differential equation which is a consequence of the equations (3.1) but one which is readily integrable, for example an equation of the type

$$d\Phi(t,\ x_1,\ x_2,\ \ldots,\ x_n)=0$$

or an equation that, by a change of variables, reduces to some kind of integrable type of equation in one unknown function.

Example 1.

$$\frac{dx}{dt}=y, \qquad \frac{dy}{dt}=x.$$

Adding the equations term by term, we find a single integrable combination

$$\frac{d(x+y)}{dt}=x+y \quad \text{or} \quad \frac{d(x+y)}{x+y}=dt,$$

whence

$$\ln|x+y|=t+\ln c_1, \quad x+y=c_1e^t.$$

Subtracting termwise the second equation of the system from the first, we get a second integrable combination

$$\frac{d(x-y)}{dt}=-(x-y) \quad \text{or} \quad \frac{d(x-y)}{x-y}=-dt,$$

$$\ln|x-y|=-t+\ln c_2, \quad x-y=c_2e^{-t}.$$

We have thus found two finite equations

$$x+y=c_1e^t \quad \text{and} \quad x-y=c_2e^{-t},$$

from which the solution of the original system can be determined,

$$x=\frac{1}{2}(c_1e^t+c_2e^{-t}), \qquad y=\frac{1}{2}(c_1e^t-c_2e^{-t})$$

or

$$x=\bar{c}_1e^t+\bar{c}_2e^{-t}, \qquad y=\bar{c}_1e^t-\bar{c}_2e^{-t}.$$

One integrable combination permits obtaining one finite equation

$$\Phi_1(t,\ x_1,\ x_2,\ \ldots,\ x_n)=c_1,$$

which relates the unknown functions and the independent variable; such a finite equation is called a *first integral of the system* (3.1).

Thus, the first integral

$$\Phi(t, x_1, x_2, \ldots, x_n) = c \tag{3.13}$$

of the system of equations (3.1) is a finite equation that is converted into an identity for some value c if the solution of the system (3.1) is substituted in place of $x_i(t)$ $(i = 1, 2, \ldots, n)$.

The left-hand member $\Phi(t, x_1 x_2, \ldots, x_n)$ of equation (3.13) is often also called a first integral, and then the first integral is defined as a function not identically equal to a constant, but retaining a constant value along the integral curves of the system (1).

Geometrically, the first integral $\Phi(t, x_1, x_2, \ldots, x_n) = c$ for fixed c may be interpreted as an n-dimensional surface in $(n+1)$-dimensional space with coordinates $t, x_1, x_2, \ldots, x_n$, the surface possessing the property that each integral curve having a common point with the surface lies entirely within the surface. For variable c, we get a family of nonintersecting surfaces possessing the same property, that is consisting of the points of some $(n-1)$-parameter family of integral curves of the system (3.1).

If k integrable combinations have been found, we get k first integrals:

$$\left.\begin{array}{l} \Phi_1(t, x_1, x_2, \ldots, x_n) = c_1, \\ \Phi_2(t, x_1, x_2, \ldots, x_n) = c_2, \\ \cdots\cdots\cdots\cdots \\ \Phi_k(t, x_1, x_2, \ldots, x_n) = c_k. \end{array}\right\} \tag{3.14}$$

If all these integrals are independent, i.e., if at least one determinant

$$\frac{D(\Phi_1, \Phi_2, \ldots, \Phi_k)}{D(x_{j_1}, x_{j_2}, \ldots, x_{j_k})} \neq 0,$$

where $x_{j_1}, x_{j_2}, \ldots, x_{j_k}$ are certain k functions made up of $x_1, x_2, \ldots, x_n$, then it is possible to express the k unknown functions of the system (3.14) in terms of the others; substituting into (3.1), we reduce the problem to integration of a system of equations with a smaller number of unknowns. If $k = n$ and all the integrals are independent, then all the unknown functions are determined from the system (3.14).

Example 2.

$$\frac{dx}{dt} = y - z, \qquad \frac{dy}{dt} = z - x, \qquad \frac{dz}{dt} = x - y.$$

Adding the equations of this system term by term, we get

$$\frac{dx}{dt} + \frac{dy}{dt} + \frac{dz}{dt} = 0 \quad \text{or} \quad \frac{d}{dt}(x + y + z) = 0$$

and from this we have

$$x+y+z=c_1.$$

The first integral that has been found permits expressing one of the unknown functions in terms of the remaining variables and this enables us to reduce the problem to integration of a system of two equations in two unknown functions. However, in the given case it is easy to find one more first integral. Multiply the first equation termwise by x, the second by y, the third by z and add:

$$x\frac{dx}{dt}+y\frac{dy}{dt}+z\frac{dz}{dt}=0,$$

or, multiplying by 2, we get

$$\frac{d}{dt}(x^2+y^2+z^2)=0,$$

whence

$$x^2+y^2+z^2=c_2.$$

Using the two first integrals thus found it is possible to express the two unknown functions in terms of the other variables and thus to reduce the problem to integration of one equation in one unknown function.

Example 3.

$$A\frac{dp}{dt}=(B-C)qr,\quad B\frac{dq}{dt}=(C-A)rp,\quad C\frac{dr}{dt}=(A-B)pq,$$

where A, B and C are constants (this system is encountered in the theory of motion of rigid bodies). Multiplying the first equation by p, the second by q, the third by r and adding, we get

$$Ap\frac{dp}{dt}+Bq\frac{dq}{dt}+Cr\frac{dr}{dt}=0,$$

whence we find the first integral

$$Ap^2+Bq^2+Cr^2=c_1.$$

Multiplying the first equation by Ap, the second by Bq, the third by Cr and adding, we have

$$A^2p\frac{dp}{dt}+B^2q\frac{dq}{dt}+C^2r\frac{dr}{dt}=0,$$

and then integrating we get yet another first integral,

$$A^2p^2+B^2q^2+C^2r^2=c_2.$$

If we exclude the case of $A=B=C$, in which the system is integrated directly, the first integrals that have been found are independent and, hence, it is possible to eliminate the two unknown functions by taking advantage of these first integrals; to determine the third function we get one equation with variables separable.

When finding integrable combinations, it is often convenient to use the so-called symmetric form of writing the system of equations (3.1):

$$\frac{dx_1}{\varphi_1(t, x_1, x_2, \ldots, x_n)} = \frac{dx_2}{\varphi_2(t, x_1, x_2, \ldots, x_n)} = \ldots$$
$$\ldots = \frac{dx_n}{\varphi_n(t, x_1, x_2, \ldots, x_n)} = \frac{dt}{\varphi_0(t, x_1, x_2, \ldots, x_n)}, \quad (3.15)$$

where

$$f_i(t, x_1, x_2, \ldots, x_n) = \frac{\varphi_i(t, x_1, x_2, \ldots, x_n)}{\varphi_0(t, x_1, x_2, \ldots, x_n)} \qquad (i = 1, 2, \ldots, n).$$

The variables are involved equivalently in a system given in symmetric form, and this sometimes simplifies finding integrable combinations.

Example 4.

$$\frac{dx}{x^2 - y^2 - z^2} = \frac{dy}{2xy} = \frac{dz}{2xz}. \quad (3.16)$$

Integrating the equation

$$\frac{dy}{2xy} = \frac{dz}{2xz},$$

we find $\frac{y}{z} = c_1$. Multiplying numerators and denominators of the first of the relations of the system (3.16) by x, the second by y, the third by z and forming a derived proportion, we get

$$\frac{x\,dx + y\,dy + z\,dz}{x(x^2 + y^2 + z^2)} = \frac{dy}{2xy}$$

and from this

$$\ln(x^2 + y^2 + z^2) = \ln|y| + \ln c_2$$

or

$$\frac{x^2 + y^2 + z^2}{y} = c_2.$$

The independent first integrals thus found,

$$\frac{y}{z} = c_1 \quad \text{and} \quad \frac{x^2 + y^2 + z^2}{y} = c_2,$$

determine the desired integral curves.

4. Systems of Linear Differential Equations

A system of differential equations is called *linear* if it is linear in all unknown functions and their derivatives. A system of n linear equations of the first order written in normal form looks like

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)\, x_j + f_i(t), \quad (i = 1, 2, \ldots, n), \qquad (3.17)$$

or, in vector form,

$$\frac{dX}{dt} = AX + F, \tag{3.18}$$

where X is an n-dimensional vector with coordinates $x_1(t)$, $x_2(t)$, ..., $x_n(t)$, F is an n-dimensional vector with coordinates $f_1(t)$, $f_2(t)$, ..., $f_n(t)$, which it will be convenient to regard in future as one-column matrices:

$$X = \begin{Vmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{Vmatrix}, \qquad F = \begin{Vmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{Vmatrix},$$

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}, \qquad \frac{dX}{dt} = \begin{Vmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \cdot \\ \cdot \\ \cdot \\ \frac{dx_n}{dt} \end{Vmatrix}.$$

According to the rule of matrix multiplication, the rows of the first factor must be multiplied by the column of the second; thus,

$$AX = \begin{Vmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \cdot \;\cdot \;\cdot \;\cdot \\ \sum_{j=1}^{n} a_{nj}x_j \end{Vmatrix}, \qquad AX + F = \begin{Vmatrix} \sum_{j=1}^{n} a_{1j}x_j + f_1 \\ \sum_{j=1}^{n} a_{2j}x_j + f_2 \\ \cdot \;\cdot \;\cdot \;\cdot \;\cdot \;\cdot \\ \sum_{j=1}^{n} a_{nj}x_j + f_n \end{Vmatrix}.$$

The equality of any matrices implies the equality of all their elements, hence the one matrix equation (3.18) or

$$\begin{Vmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \cdot \\ \cdot \\ \cdot \\ \frac{dx_n}{dt} \end{Vmatrix} = \begin{Vmatrix} \sum_{j=1}^{n} a_{1j}x_j + f_1 \\ \sum_{j=1}^{n} a_{2j}x_j + f_2 \\ \cdot \;\cdot \;\cdot \;\cdot \;\cdot \;\cdot \\ \sum_{j=1}^{n} a_{nj}x_j + f_n \end{Vmatrix}$$

is equivalent to the system (3.17).

If all the functions $a_{ij}(t)$ and $f_i(t)$ in (3.17) are continuous on the interval $a \leqslant t \leqslant b$, then in a sufficiently small neighbourhood of every point $(t_0, x_{10}, x_{20}, \ldots, x_{n0})$, where $a \leqslant t_0 \leqslant b$, the conditions of the existence and uniqueness theorem are fulfilled (see page 177) and, hence, a unique integral curve of the system (3.17) passes through every such point.

Indeed, in the case at hand the right-hand members of the system (3.17) are continuous and their partial derivatives with respect to any x_j are bounded, since these partial derivatives are equal to the coefficients $a_{ij}(t)$ which are continuous on the interval $a \leqslant t \leqslant b$.

We define the *linear operator* L by the equality

$$L[X] = \frac{dX}{dt} - AX,$$

then equation (3.18) may be written still more concisely as

$$L[X] = F. \tag{3.19}$$

If all the $f_i(t) \equiv 0$ $(i = 1, 2, \ldots, n)$ or, what is the same thing, the matrix $F = 0$, then the system (3.17) is called *homogeneous linear*. In abridged notation, a homogeneous linear system is of the form

$$L[X] = 0. \tag{3.20}$$

The operator L has the following two properties:

(1) $L[cX] = cL[X]$,

where c is an arbitrary constant.

(2) $L[X_1 + X_2] = L[X_1] + L[X_2]$.

Indeed,

$$\frac{d(cX)}{dt} - A(cX) \equiv c\left[\frac{dX}{dt} - AX\right],$$

$$\frac{d(X_1 + X_2)}{dt} - A(X_1 + X_2) \equiv \left(\frac{dX_1}{dt} - AX_1\right) + \left(\frac{dX_2}{dt} - AX_2\right).$$

A consequence of Properties (1) and (2) is

$$L\left[\sum_{i=1}^{m} c_i X_i\right] \equiv \sum_{i=1}^{m} c_i L[X_i],$$

where c_i are arbitrary constants.

Theorem 3.1. *If X is a solution of a homogeneous linear system $L[X] = 0$, then cX, where c is an arbitrary constant, is a solution of the same system.*

Proof. Given $L[X] \equiv 0$; it is required to prove that $L[cX] \equiv 0$. Taking advantage of Property (1) of the operator L, we have

$$L[cX] \equiv cL[X] \equiv 0.$$

Theorem 3.2. *The sum X_1+X_2 of two solutions X_1 and X_2 of a homogeneous linear system of equations is a solution of that system.*

Proof. Given $L[X_1]\equiv 0$ and $L[X_2]\equiv 0$.
It is required to prove that $L[X_1+X_2]\equiv 0$.
Using Property (2) of the operator L, we obtain

$$L[X_1+X_2]\equiv L[X_1]+L[X_2]\equiv 0.$$

Corollary to Theorems 3.1 and 3.2. *A linear combination $\sum_{i=1}^{m} c_iX_i$ with arbitrary constant coefficients of the solutions $X_1, X_2, \ldots, X_m$ of a homogeneous linear system $L[X]=0$ is a solution of that system.*

Theorem 3.3. *If the homogeneous linear system* (3.20) *with real coefficients $a_{ij}(t)$ has a complex solution $X=U+iV$, then the real and imaginary parts*

$$U=\begin{Vmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{Vmatrix} \quad \text{and} \quad V=\begin{Vmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{Vmatrix}$$

separately are solutions of that system.

Proof. Given $L[U+iV]\equiv 0$. It is required to prove that

$$L[U]\equiv 0 \quad \text{and} \quad L[V]\equiv 0.$$

Using Properties (1) and (2) of the operator L, we get

$$L[U+iV]\equiv L[U]+iL[V]\equiv 0.$$

Hence, $L[U]\equiv 0$ and $L[V]\equiv 0$.

The vectors $X_1, X_2, \ldots, X_n$, where

$$X_i=\begin{Vmatrix} x_{1i}(t) \\ x_{2i}(t) \\ \cdot \\ \cdot \\ \cdot \\ x_{ni}(t) \end{Vmatrix},$$

are called *linearly dependent* on the interval $a\leqslant t\leqslant b$ if there exist constants $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\alpha_1X_1+\alpha_2X_2+\ldots+\alpha_nX_n\equiv 0 \tag{3.21}$$

for $a\leqslant t\leqslant b$; and at least one $\alpha_i\neq 0$. But if the identity (3.21) holds true only for $\alpha_1=\alpha_2=\ldots=\alpha_n=0$, then the vectors $X_1, X_2, \ldots, X_n$ are termed *linearly independent*.

We note that the single vector identity (3.21) is equivalent to n identities:

$$\left.\begin{aligned} &\sum_{i=1}^{n} \alpha_i x_{1i}(t) \equiv 0, \\ &\sum_{i=1}^{n} \alpha_i x_{2i}(t) \equiv 0, \\ &\dots\dots\dots \\ &\sum_{i=1}^{n} \alpha_i x_{ni}(t) \equiv 0. \end{aligned}\right\} \qquad (3.21_1)$$

If the vectors X_i $(i = 1, 2, \ldots, n)$ are linearly dependent and thus there exists a nontrivial system α_i (i.e., not all α_i are zero) that satisfies the system of n homogeneous (with respect to α_i) linear equations (3.21_1), then the determinant of the system (3.21_1)

$$W = \begin{vmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \ldots & x_{nn} \end{vmatrix}$$

must be zero for all values of t of the interval $a \leqslant t \leqslant b$. This determinant of the system is called the *Wronskian determinant* of the system of vectors $X_1, X_2, \ldots, X_n$.

Theorem 3.4. *If the Wronskian W of the solutions $X_1, X_2, \ldots, X_n$ of the homogeneous linear system of equations* (3.20) *with coefficients $a_{ij}(t)$ continuous on the interval $a \leqslant t \leqslant b$ is zero at least in one point $t = t_0$ of the interval $a \leqslant t \leqslant b$, then the solutions $X_1, X_2, \ldots, X_n$ are linearly dependent on that interval, and, hence, $W \equiv 0$ on that interval.*

Proof. Since the coefficients $a_{ij}(t)$ $(i, j = 1, 2, \ldots, n)$ are continuous, the system (3.20) satisfies the conditions of the existence and uniqueness theorem. Hence, the initial value $X(t_0) = 0$ [or in more detail, $x_1(t_0) = 0$, $x_2(t_0) = 0$, $\ldots$, $x_n(t_0) = 0$] determines the unique solution of this system, and this solution is obviously the trivial solution of the system (3.20) $X(t) \equiv 0$ [or, in more detail, $x_1(t) \equiv 0$, $x_2(t) \equiv 0$, $\ldots$, $x_n(t) \equiv 0$]. The determinant $W(t_0) = 0$. Hence, there exists a nontrivial system $c_1, c_2, \ldots, c_n$ that satisfies the equation

$$c_1 X_1(t_0) + c_2 X_2(t_0) + \ldots + c_n X_n(t_0) \equiv 0,$$

since this single vector equation is equivalent to a system of n homogeneous (with respect to c_i) linear equations with the zero determinant

$$\sum_{i=1}^{n} c_i x_{1i}(t_0) = 0,$$
$$\sum_{i=1}^{n} c_i x_{2i}(t_0) = 0,$$
$$\cdots\cdots\cdots$$
$$\sum_{i=1}^{n} c_i x_{ni}(t_0) = 0.$$

The solution of the equation (3.20) $X(t) = \sum_{i=1}^{n} c_i X_i(t)$ that corresponds to this nontrivial system $c_1, c_2, \ldots, c_n$ satisfies the zero initial conditions $X(t_0) = 0$ and, consequently, coincides with the trivial solution of the system (3.20):

$$\sum_{i=1}^{n} c_i X_i(t) \equiv 0,$$

i.e., the X_i are linearly dependent

Note. As the most elementary examples show, this theorem does not extend to the arbitrary vectors $X_1, X_2, \ldots, X_n$ which are not solutions of the system (3.20) with continuous coefficients.

Example 1. The system of vectors

$$X_1 = \left\| \begin{matrix} t \\ t \end{matrix} \right\| \quad \text{and} \quad X_2 = \left\| \begin{matrix} t^2 \\ t^2 \end{matrix} \right\|$$

is linearly independent, since from

$$\alpha_1 X_1 + \alpha_2 X_2 \equiv 0$$

or

$$\begin{cases} \alpha_1 t + \alpha_2 t^2 \equiv 0, \\ \alpha_1 t + \alpha_2 t^2 \equiv 0 \end{cases}$$

it follows that $\alpha_1 = \alpha_2 = 0$ (see page 102, Example 1). At the same time the Wronskian $\begin{vmatrix} t & t^2 \\ t & t^2 \end{vmatrix}$ is identically zero. Hence, the vectors X_1 and X_2 cannot be solutions of one and the same homogeneous linear system (3.20) with continuous coefficients $a_{ij}(t)$ $(i, j = 1, 2)$

Theorem 3.5. *The linear combination $\sum_{i=1}^{n} c_i X_i$ of n linearly independent solutions $X_1, X_2, \ldots, X_n$ of the homogeneous linear system* (3.20) *with coefficients $a_{ij}(t)$ continuous on the interval $a \leqslant t \leqslant b$ is the general solution of the system* (3.20) *on that interval.*

Proof. Since the coefficients $a_{ij}(t)$ are continuous on the interval $a \leqslant t \leqslant b$, the system satisfies the conditions of the existence and uniqueness theorem and, hence, to prove the theorem it is sufficient to notice that by proper choice of the constants c_i in the solution $\sum_{i=1}^{n} c_i X_i$ it is possible to satisfy arbitrarily chosen initial conditions $X(t_0) = X_0$,

$$X_0 = \begin{vmatrix} x_{10} \\ x_{20} \\ \cdot \\ \cdot \\ \cdot \\ x_{n0} \end{vmatrix},$$

where t_0 is one of the values of t on the interval $a \leqslant t \leqslant b$, i.e., it is possible to satisfy the single vector equation

$$\sum_{i=1}^{n} c_i X_i(t_0) = X_0$$

or the equivalent system of n scalar equations:

$$\begin{gathered} \sum_{i=1}^{n} c_i x_{1i}(t_0) = x_{10}, \\ \sum_{i=1}^{n} c_i x_{2i}(t_0) = x_{20}, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \sum_{i=1}^{n} c_i x_{ni}(t_0) = x_{n0}. \end{gathered}$$

This system is solvable for c_i for any x_{i0}, since the determinant of the system is the Wronskian determinant for a linearly independent system of solutions $X_1, X_2, \ldots, X_n$ and, hence, does not vanish at any point of the interval $a \leqslant t \leqslant b$.

Example 2.

$$\left.\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= -x. \end{aligned}\right\} \tag{3.22}$$

It may readily be verified that the system (3.22) is satisfied by the solutions

$$x_1 = \cos t,\ y_1 = -\sin t \text{ and } x_2 = \sin t,\ y_2 = \cos t.$$

These solutions are linearly independent since the Wronskian

$$\begin{vmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{vmatrix} = 1$$

is not zero. And so the general solution is of the form

$$x = c_1 \cos t + c_2 \sin t,$$
$$y = -c_1 \sin t + c_2 \cos t,$$

where c_1 and c_2 are arbitrary constants.

Theorem 3.6. *If $\tilde{X}$ is a solution of the nonhomogeneous linear system*

$$L[X] = F, \tag{3.19}$$

and X_1 is a solution of the corresponding homogeneous system $L[X] = 0$ then the sum $X_1 + \tilde{X}$ is also a solution of the nonhomogeneous system $L[X] = F$.

Proof. Given $L[\tilde{X}] \equiv F$ and $L[X_1] \equiv 0$. Prove that $L[X_1 + \tilde{X}] \equiv F$. Using Property (2) of the operator L, we get

$$L[X_1 + \tilde{X}] \equiv L[X_1] + L[\tilde{X}] \equiv F.$$

Theorem 3.7. *The general solution, on the interval $a \leqslant t \leqslant b$, of the nonhomogeneous system* (3.19) *with coefficients $a_{ij}(t)$ continuous on that interval and with right sides $f_i(t)$ is equal to the sum of the general solution $\sum_{i=1}^{n} c_i X_i$ of the corresponding homogeneous system and the particular solution $\tilde{X}$ of the nonhomogeneous system under consideration.*

Proof. Since the conditions of the existence and uniqueness theorem are fulfilled (see page 191), to prove the theorem it will suffice to notice that by means of a proper choice of arbitrary constants c_i in the solution $X = \sum_{i=1}^{n} c_i X_i + \tilde{X}$ it is possible to satisfy arbitrary specified initial conditions

$$X(t_0) = X_0 = \begin{Vmatrix} x_{10} \\ x_{20} \\ \cdot \\ \cdot \\ \cdot \\ x_{n0} \end{Vmatrix},$$

i.e., we have to prove that the one matrix equation

$$\sum_{i=1}^{n} c_i X_i(t_0) + \tilde{X}(t_0) = X_0$$

or an equivalent system of equations

$$\left.\begin{aligned}&\sum_{i=1}^{n} c_i x_{1i}(t_0)+\tilde{x}_1(t_0)=x_{10},\\&\sum_{i=1}^{n} c_i x_{2i}(t_0)+\tilde{x}_2(t_0)=x_{20},\\&\cdots\cdots\cdots\cdots\\&\sum_{i=1}^{n} c_i x_{ni}(t_0)+\tilde{x}_n(t_0)=x_{n0}\end{aligned}\right\}\qquad(3.23)$$

always has the solution $c_1, c_2, \ldots, c_n$, no matter what the right-hand members. But in this form this assertion is obvious, since the determinant of the system (3.23) is the Wronskian at the point $t=t_0$ for the linearly independent solutions $X_1, X_2, \ldots, X_n$ of the corresponding homogeneous system and, by Theorem 3.4, is not zero. Hence, the system (3.23) has the solution $c_1, c_2, \ldots, c_n$ for any right-hand members.

Theorem 3.8 (the principle of superposition). *The solution of the system of linear equations*

$$L[X]=\sum_{i=1}^{m} F_i, \qquad F_i=\left\|\begin{matrix}f_{1i}(t)\\f_{2i}(t)\\\cdot\\\cdot\\f_{ni}(t)\end{matrix}\right\|$$

is the sum $\sum_{i=1}^{m} X_i$ *of the solutions* X_i *of the equations*

$$L[X_i]=F_i \quad (i=1, 2, \ldots, m).$$

Proof. Given $L[X_i]\equiv F_i$ $(i=1, 2, \ldots, m)$. Prove that

$$L\left[\sum_{i=1}^{m} X_i\right]\equiv\sum_{i=1}^{m} F_i.$$

Taking advantage of Property (2) of the operator L, we obtain

$$L\left[\sum_{i=1}^{m} X_i\right]\equiv\sum_{i=1}^{m} L[X_i]\equiv\sum_{i=1}^{m} F_i.$$

Note. Theorem 3.8 can, without any change in the proof, be obviously extended to the case when $m\to\infty$, if the series $\sum_{i=1}^{\infty} X_i$ converges and admits term-by-term differentiation.

Theorem 3.9. *If the system of linear equations*

$$L[X]=U+iV,$$

where

$$U=\begin{Vmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{Vmatrix}, \qquad V=\begin{Vmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{Vmatrix},$$

with real functions $a_{ij}(t)$, $u_i(t)$, $v_i(t)$ (i, $j=1, 2, \ldots, n$) has the solution

$$X=\tilde{U}+i\tilde{V}, \quad \tilde{U}=\begin{Vmatrix} \tilde{u}_1 \\ \tilde{u}_2 \\ \cdot \\ \cdot \\ \cdot \\ \tilde{u}_n \end{Vmatrix}, \qquad \tilde{V}\begin{Vmatrix} \tilde{v}_1 \\ \tilde{v}_2 \\ \cdot \\ \cdot \\ \cdot \\ \tilde{v}_n \end{Vmatrix},$$

then the real part of the solution $\tilde{U}$ and its imaginary part $\tilde{V}$ are respectively solutions of the equations

$$L[X]=U \quad \text{and} \quad L[X]=V.$$

Proof. Given $L[\tilde{U}+i\tilde{V}]\equiv U+iV$; prove that $L[\tilde{U}]\equiv U$, $L[\tilde{V}]\equiv V$.

Using Properties (1) and (2) of the operator L, we get

$$L[\tilde{U}+i\tilde{V}]\equiv L[\tilde{U}]+iL[\tilde{V}]\equiv U+iV.$$

Hence, $L[\tilde{U}]\equiv U$ and $L[\tilde{V}]\equiv V$.

If the general solution of the corresponding homogeneous system $L(X)=0$ is known, and one cannot choose a particular solution of the nonhomogeneous system $L(X)=F$ and, consequently, one cannot take advantage of Theorem 3.7, then the method of variation of parameters may be applied.

Let $X=\sum_{i=1}^{n} c_iX_i$ be the general solution of the corresponding homogeneous system

$$\frac{dX}{dt}-AX=0$$

for arbitrary constants c_i, and, hence, X_i ($i=1, 2, \ldots, n$) are linearly independent particular solutions of the same homogeneous system. We seek the solution of the nonhomogeneous system

$$\frac{dX}{dt} - AX = F$$

in the form

$$X = \sum_{i=1}^{n} c_i(t) X_i,$$

where the $c_i(t)$ are new unknown functions. Substitution into the nonhomogeneous equation yields

$$\sum_{i=1}^{n} c_i'(t) X_i + \sum_{i=1}^{n} c_i(t) \frac{dX_i}{dt} = A \sum_{i=1}^{n} c_i(t) X_i + F,$$

or, since $\frac{dX_i}{dt} \equiv AX_i$, we have

$$\sum_{i=1}^{n} c_i'(t) X_i = F.$$

This vector equation is equivalent to a system of n equations:

$$\left.\begin{array}{l} \sum_{i=1}^{n} c_i'(t) x_{1i} = f_1(t) \\ \sum_{i=1}^{n} c_i'(t) x_{2i} = f_2(t), \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \sum_{i=1}^{n} c_i'(t) x_{ni} = f_n(t). \end{array}\right\} \qquad (3.24)$$

All the $c_i'(t)$ are determined from this system of n equations in n unknowns $c_i'(t)$ $(i = 1, 2, \ldots, n)$ with the determinant of the system W coinciding with the Wronskian for the linearly independent solutions $X_1, X_2, \ldots, X_n$ and, consequently, not zero:

$$c_i'(t) = \varphi_i(t) \qquad (i = 1, 2, \ldots, n),$$

whence, by integrating, we find the unknown functions $c_i(t)$:

$$c_i(t) = \int \varphi_i(t)\,dt + \bar{c}_i \qquad (i = 1, 2, \ldots, n).$$

Example 3.

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x + \frac{1}{\cos t}.$$

The general solution of the corresponding homogeneous system

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x$$

is of the form $x = c_1 \cos t + c_2 \sin t$, $y = -c_1 \sin t + c_2 \cos t$ (see page 195, Example 2). We vary the constants

$$x = c_1(t) \cos t + c_2(t) \sin t,$$
$$y = -c_1(t) \sin t + c_2(t) \cos t.$$

$c_1'(t)$ and $c_2'(t)$ are determined from the system (3.24), which, in this case, is of the form

$$c_1'(t) \cos t + c_2'(t) \sin t = 0,$$
$$-c_1'(t) \sin t + c_2'(t) \cos t = \frac{1}{\cos t},$$

whence

$$c_1'(t) = -\frac{\sin t}{\cos t}, \quad c_2'(t) = 1.$$

Therefore,

$$c_1(t) = \ln|\cos t| + \bar{c}_1,$$
$$c_2(t) = t + \bar{c}_2$$

and we finally get

$$x = \bar{c}_1 \cos t + \bar{c}_2 \sin t + \cos t \ln|\cos t| + t \sin t,$$
$$y = -\bar{c}_1 \sin t + \bar{c}_2 \cos t - \sin t \ln|\cos t| + t \cos t.$$

5. *Systems of Linear Differential Equations with Constant Coefficients*

The linear system of equations

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij} x_j + f_i(t) \qquad (i = 1, 2, \ldots, n)$$

is a *linear system with constant coefficients*. In the vector form it is

$$\frac{dX}{dt} = AX + F,$$

in which all the coefficients a_{ij} are constant or, what is the same thing, the matrix A is constant.

A system of homogeneous or nonhomogeneous linear equations with constant coefficients is most simply integrated by reducing it to a single equation of higher order. As was noted on page 185, the resulting equation of higher order will be linear with constant coefficients.

However, it is possible directly to find the fundamental system of solutions of a homogeneous linear system with constant coefficients.

We will seek the solutions of the system

$$\left.\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n, \end{aligned}\right\} \tag{3.25}$$

where all the a_{ij} are constant, in the form

$$x_1 = \alpha_1 e^{kt}, \quad x_2 = \alpha_2 e^{kt}, \quad \ldots, \quad x_n = \alpha_n e^{kt},$$

with constants α_j $(j = 1, 2, \ldots, n)$. Substituting into the system (3.25), cancelling e^{kt} and transposing all terms to one side, we get

$$\left.\begin{aligned} (a_{11} - k)\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n &= 0, \\ a_{21}\alpha_1 + (a_{22} - k)\alpha_2 + \ldots + a_{2n}\alpha_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\alpha_1 + a_{n2}\alpha_2 + \ldots + (a_{nn} - k)\alpha_n &= 0. \end{aligned}\right\} \tag{3.26}$$

For this system of n homogeneous linear equations in n unknowns α_j $(j = 1, 2, \ldots, n)$ to have a nontrivial solution, it is necessary and sufficient that the determinant of the system (3.26) be zero:

$$\begin{vmatrix} a_{11} - k & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - k & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - k \end{vmatrix} = 0. \tag{3.27}$$

From this equation of degree n we determine the values of k at which the system (3.26) has nontrivial solutions α_j $(j = 1, 2, \ldots, n)$. Equation (3.27) is called the *characteristic equation*. If all the roots k_i $(i = 1, 2, \ldots, n)$ of the characteristic equation are distinct, then, by putting them into the system (3.26) in succession, we determine the corresponding nontrivial values $\alpha_j^{(i)}$ $(i, j = 1, 2, \ldots, n)$ and, consequently, we find the n solutions of the original system (3.25) in the form

$$x_1^{(i)} = \alpha_1^{(i)} e^{k_i t}, \quad x_2^{(i)} = \alpha_2^{(i)} e^{k_i t}, \quad \ldots, \quad x_n^{(i)} = \alpha_n^{(i)} e^{k_i t} \quad (i = 1, 2, \ldots, n), \tag{3.28}$$

where the upper index indicates the number of the solution and the lower index, the number of the unknown function.

Using vector notation, we get the same result more compactly:

$$\frac{dX}{dt} = AX; \tag{3.25$_1$}$$

we seek the solution in the form

$$X=\tilde{A}e^{kt}, \quad \text{where} \quad A=\begin{Vmatrix}\alpha_1\\ \alpha_2\\ \cdot\\ \cdot\\ \cdot\\ \alpha_n\end{Vmatrix},$$

$$\tilde{A}ke^{kt}=A\tilde{A}e^{kt}$$

or

$$(A-kE)\tilde{A}=0, \tag{3.29}$$

where E is a unit matrix:

$$E=\begin{Vmatrix}1 & 0 & 0 & \dots & 0\\ 0 & 1 & 0 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \dots & 1\end{Vmatrix}.$$

So that the equation (3.29) is satisfied by the nontrivial matrix $\tilde{A}$

$$\tilde{A}\neq\begin{Vmatrix}0\\ 0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{Vmatrix},$$

it is necessary and sufficient that the matrix $A-kE$ should be singular, i.e., that its determinant be zero: $|A-kE|=0$. For each root k_i of this characteristic equation $|A-kE|=0$ we determine, from (3.29), the nonzero matrix $\tilde{A}^{(i)}$ and, if all roots k_i of the characteristic equation are distinct, we get n solutions:

$$X_1=\tilde{A}^{(1)}e^{k_1t}, \quad X_2=\tilde{A}^{(2)}e^{k_2t}, \quad \dots, \quad X_n=\tilde{A}^{(n)}e^{k_nt},$$

where

$$\tilde{A}^{(i)}=\begin{Vmatrix}\alpha_1^{(i)}\\ \alpha_2^{(i)}\\ \cdot\\ \cdot\\ \cdot\\ \alpha_n^{(i)}\end{Vmatrix}.$$

It is easy to show that these solutions are linearly independent.

Indeed, if there were a linear dependence

$$\sum_{i=1}^{n} \beta_i \bar{A}^{(i)} e^{k_i t} = 0$$

or, in expanded form,

$$\left.\begin{aligned}
&\sum_{i=1}^{n} \beta_i \alpha_1^{(i)} e^{k_i t} \equiv 0, \\
&\sum_{i=1}^{n} \beta_i \alpha_2^{(i)} e^{k_i t} \equiv 0, \\
&\cdots\cdots\cdots \\
&\sum_{i=1}^{n} \beta_i \alpha_n^{(i)} e^{k_i t} \equiv 0,
\end{aligned}\right\} \tag{3.30}$$

then, by virtue of the linear independence of the functions $e^{k_i t}$ $(i=1, 2, \ldots, n)$ (see pages 101-102) it would follow from (3.30) that

$$\left.\begin{aligned}
&\beta_i \alpha_1^{(i)} = 0, \\
&\beta_i \alpha_2^{(i)} = 0, \\
&\cdots\cdots \\
&\beta_i \alpha_n^{(i)} = 0.
\end{aligned}\right\} \quad (i=1, 2, \ldots, n). \tag{3.31}$$

But since for every i, at least one of the $\alpha_1^{(i)}, \alpha_2^{(i)}, \ldots, \alpha_n^{(i)}$ $(i=1, 2, \ldots, n)$ is different from zero, it follows from (3.31) that $\beta_i = 0$ $(i=1, 2, \ldots, n)$.

And so the solutions $\bar{A}^{(i)} e^{k_i t}$ $(i=1, 2, \ldots, n)$ are linearly independent and the general solution of the system (3.25) is of the form

$$X = \sum_{i=1}^{n} c_i \bar{A}^{(i)} e^{k_i t}$$

or

$$x_j = \sum_{i=1}^{n} c_i \alpha_j^{(i)} e^{k_i t} \quad (j=1, 2, \ldots, n),$$

where the c_i are arbitrary constants.

The constants $\alpha_j^{(i)}$ $(j=1, 2, \ldots, n)$ are ambiguously determined from the system (3.26) for $k=k_i$, since the determinant of the system is zero and, hence, at least one equation is a consequence of the others. The ambiguity in the determination of $\alpha_j^{(i)}$ is due to the fact that a solution of the system of homogeneous linear equations remains a solution of the same system when it is multiplied by an arbitrary constant factor.

To the complex root $k_j = p + qi$ of the characteristic equation (3.27) there corresponds the solution

$$X_j = \tilde{A}^{(j)} e^{k_j t} \tag{3.32}$$

which, if all the coefficients a_{ij} are real, can be replaced by two real solutions: the real part and imaginary part of the solution (3.32) (see page 192). A complex conjugate root $k_{j+1} = p - qi$ of the characteristic equation will not yield any new linearly independent real solutions.

If the characteristic equation has a multiple root k_s of multiplicity γ, then, taking into account that the system of equations (3.25) can be reduced (by a process indicated on pages 181-183) to a single homogeneous linear equation with constant coefficients of order n or less (see Note 2 on page 185), it is possible to assert that the solution of the system (3.25) is of the form

$$X(t) = (\tilde{A}_0^{(s)} + \tilde{A}_1^{(s)} t + \ldots + \tilde{A}_{\gamma-1}^{(s)} t^{\gamma-1}) e^{k_s t}, \tag{3.33}$$

where

$$\tilde{A}_i^{(s)} = \left\| \begin{matrix} \alpha_{1i}^{(s)} \\ \alpha_{2i}^{(s)} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{ni}^{(s)} \end{matrix} \right\|,$$

$\alpha_{ji}^{(s)}$ are constants.

It should be noted that even in cases when the system of n equations (3.25) is reduced to an equation of order lower than n (see Note 1 on page 184), the characteristic equation of the latter necessarily has roots that coincide with the roots of the equation (3.27) [since the equation to which the system was reduced has to have solutions of the form $e^{k_s t}$ where the k_s are the roots of the equation (3.27)]. But it may be that the multiplicities of these roots, if the order of the equation obtained is less than n, will be lower than the multiplicities of the roots of the equation (3.27) and, hence, it may be that in the solution (3.33) the degree of the first factor will be lower than $\gamma - 1$, i.e., if we seek the solution in the form of (3.33), it may turn out that some of the coefficients $\tilde{A}_i^{(s)}$, including the coefficient of the highest-degree term, vanish.

Thus, we have to seek a solution of the system (3.25), which solution corresponds to a multiple root of the characteristic equation, in the form of (3.33). Putting (3.33) into (3.25_1) and demanding that it become an identity, we define the matrices $A_i^{(s)}$; some of them, including $A_{\gamma-1}^{(s)}$ as well, may turn out equal to zero.

Note. There is a more precise way of indicating the type of solution of the system (3.25) that corresponds to a multiple root of the characteristic equation (3.27). Transforming the system (3.25) by means of a nonsingular linear transformation to a system in which the matrix $\| A-kE \|$ has the Jordan normal form and then integrating the resulting readily integrable system of equations, we find that the solution which corresponds to the multiple root k_s of multiplicity γ of the characteristic equation (3.27) is of the form

$$X(t) = (\tilde{A}_0^{(s)} + \tilde{A}_1^{(s)} t + \ldots + \tilde{A}_{\beta-1}^{(s)} t^{\beta-1}) e^{k_s t},$$

where β is the highest degree of the elementary divisor of the matrix $\| A-kE \|$ corresponding to the root k_s.

Example 1.

$$\frac{dx}{dt} = x + 2y, \qquad \frac{dy}{dt} = 4x + 3y.$$

The characteristic equation

$$\begin{vmatrix} 1-k & 2 \\ 4 & 3-k \end{vmatrix} = 0 \quad \text{or} \quad k^2 - 4k - 5 = 0$$

has roots $k_1 = 5$, $k_2 = -1$. Hence, we seek the solution in the form

$$\begin{aligned} x_1 &= \alpha_1^{(1)} e^{5t}, & y_1 &= \alpha_2^{(1)} e^{5t}, \\ x_2 &= \alpha_1^{(2)} e^{-t}, & y_2 &= \alpha_2^{(2)} e^{-t}. \end{aligned} \tag{3.34}$$

Putting (3.34) in the original system, we get $-4\alpha_1^{(1)} + 2\alpha_2^{(1)} = 0$, whence $\alpha_2^{(1)} = 2\alpha_1^{(1)}$; $\alpha_1^{(1)}$ remains arbitrary. Consequently,

$$x_1 = c_1 e^{5t}, \ y_1 = 2c_1 e^{5t}, \ c_1 = \alpha_1^{(1)}.$$

For determining the coefficients $\alpha_1^{(2)}$ and $\alpha_2^{(2)}$ we get the equation $2\alpha_1^{(2)} + 2\alpha_2^{(2)} = 0$, whence $\alpha_2^{(2)} = -\alpha_1^{(2)}$; the coefficient $\alpha_1^{(2)}$ remains arbitrary.

Hence,

$$x_2 = c_2 e^{-t}, \ y_2 = -c_2 e^{-t}, \ c_2 = \alpha_1^{(2)}.$$

The general solution

$$\begin{aligned} x &= c_1 e^{5t} + c_2 e^{-t}, \\ y &= 2c_1 e^{5t} - c_2 e^{-t}. \end{aligned}$$

Example 2.

$$\frac{dx}{dt} = x - 5y,$$

$$\frac{dy}{dt} = 2x - y.$$

The characteristic equation

$$\begin{vmatrix} 1-k & -5 \\ 2 & -1-k \end{vmatrix} = 0 \quad \text{or} \quad k^2 + 9 = 0$$

has roots $k_{1,2}=\pm 3i$, $x_1=\alpha_1 e^{3it}$, $y_1=\alpha_2 e^{3it}$, $(1-3i)\alpha_1-5\alpha_2=0$. This equation is satisfied, for example, by $\alpha_1=5$, $\alpha_2=1-3i$. Therefore,

$$x_1=5e^{3it}=5(\cos 3t+i\sin 3t),$$
$$y_1=(1-3i)e^{3it}=(1-3i)(\cos 3t+i\sin 3t).$$

The real part and the imaginary part of this solution are likewise solutions of the system at hand, and their linear combination with arbitrary constant coefficients is the general solution:

$$x=5c_1\cos 3t+5c_2\sin 3t,$$
$$y=c_1(\cos 3t+3\sin 3t)+c_2(\sin 3t-3\cos 3t).$$

Example 3.

$$\left.\begin{aligned}\frac{dx}{dt}&=x-y,\\ \frac{dy}{dt}&=x+3y.\end{aligned}\right\}\qquad(3.35)$$

The characteristic equation

$$\begin{vmatrix}1-k & -1\\ 1 & 3-k\end{vmatrix}=0 \quad\text{or}\quad k^2-4k+4=0$$

has a multiple root $k_{1,2}=2$. Hence, the solution is to be sought in the form

$$\left.\begin{aligned}x&=(\alpha_1+\beta_1 t)e^{2t},\\ y&=(\alpha_2+\beta_2 t)e^{2t}\end{aligned}\right\}\qquad(3.36)$$

Putting (3.36) into (3.35), we get

$$2\alpha_1+\beta_1+2\beta_1 t\equiv\alpha_1+\beta_1 t-\alpha_2-\beta_2 t,$$

whence

$$\beta_2=-\beta_1.$$
$$\alpha_2=-\alpha_1-\beta_1,$$

α_1 and β_1 remain arbitrary. Denoting these arbitrary constants by c_1 and c_2, respectively, we get the general solution in the form

$$x=(c_1+c_2 t)e^{2t},$$
$$y=-(c_1+c_2+c_2 t)e^{2t}.$$

6. Approximate Methods of Integrating Systems of Differential Equations and Equations of Order n

All the methods (given in Sec. 7, Ch. 1) of approximate integration of differential equations of the first order may be carried over, without essential changes, to systems of first-order equations and also to equations of order two and higher, which are reduced in the usual manner to systems of first-order equations (see page 91).

1. *Successive approximation*. As was pointed out on page 56, the method of successive approximations is applicable to systems of equations

$$\frac{dy_i}{dx}=f_i(x,\ y_1,\ y_2,\ \ldots,\ y_n)\qquad (i=1,\ 2,\ \ldots,\ n)\qquad (3.37)$$

with initial conditions $y_i(x_0)=y_{i0}$ $(i=1,\ 2,\ \ldots,\ n)$ if the functions f_i are continuous in all arguments and satisfy the Lipschitz conditions with respect to all arguments, from the second onwards.

The zero approximation $y_{i0}(x)$ $(i=1,\ 2,\ \ldots,\ n)$ may be chosen arbitrarily as long as the initial conditions are satisfied, and further approximations are computed from the formula

$$y_{i,\,k+1}(x)=y_{i0}+\int_{x_0}^{x} f_i(x,\ y_{1k},\ y_{2k},\ \ldots,\ y_{nk})\,dx\quad (i=1,\ 2,\ \ldots,\ n).$$

Just as for one equation of the first order, this method is rarely applied in practical calculations due to the relatively slow convergence of the approximations and the complexity and diversity of the computations.

2. *Euler's method*. An integral curve of the system of differential equations

$$\frac{dy_i}{dx}=f_i(x,\ y_1,\ y_2,\ \ldots,\ y_n)$$
$$(i=1,\ 2,\ \ldots,\ n),$$

defined by the initial conditions $y_i(x_0)=$ $=y_{i0}$ $(i=1,\ 2,\ \ldots,\ n)$ is replaced by a polygonal line that is tangent, at one of the boundary points of each segment, to the integral curve passing through the same point (Fig 3.2 depicts Euler's polygonal line and its projection only on the xy_1-plane). The interval $x_0\leqslant x\leqslant b$ on which the solution has to be computed is partitioned into segments of length h, and the computation is performed using the formulas

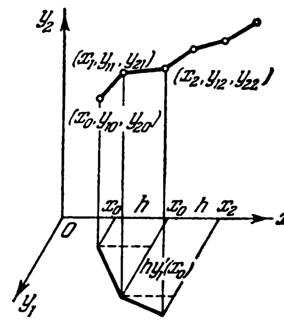


Fig. 3-2

$$y_i(x_{k+1})=y_i(x_k)+hy_i'(x_k)\qquad (i=1,\ 2,\ \ldots,\ n).$$

The convergence of Euler's polygonal lines to the integral curve as $h\to 0$ is proved in the same way as for a single first-order equation (see page 48). Iteration may be employed to increase the precision.

3. *Expansion by means of Taylor's formula.* Assuming that the right-hand members of the system of equations (3.37) are differentiable k times (in order to ensure differentiability of the solutions $k+1$ times), we replace the desired solutions by the first few terms of their Taylor expansions:

$$y_i(x) \approx y_i(x_0) + y_i'(x_0)(x-x_0) + y_i''(x_0)\frac{(x-x_0)^2}{2!} + \dots$$
$$\dots + y_i^{(k)}(x_0)\frac{(x-x_0)^k}{k!} \quad (i = 1, 2, \dots, n).$$

The error may be estimated by estimating the remainder term in Taylor's formula

$$R_{ik} = y_i^{(k+1)}[x_0 + \theta(x-x_0)]\frac{(x-x_0)^{k+1}}{(k+1)!}, \quad \text{where} \quad 0 < \theta < 1.$$

This method yields good results only in a small neighbourhood of the point x_0.

4. *Störmer's method.* The interval $x_0 \leqslant x \leqslant b$ is partitioned into subintervals of length h, and the computation of the solution of the system (3.37) is performed on the basis of one of the following formulas:

$$y_{i,k+1} = y_{ik} + q_{ik} + \frac{1}{2}\Delta q_{i,k-1}, \tag{3.38}$$

$$y_{i,k+1} = y_{ik} + q_{ik} + \frac{1}{2}\Delta q_{i,k-1} + \frac{5}{12}\Delta^2 q_{i,k-2}, \tag{3.39}$$

$$y_{i,k+1} = y_{ik} + q_{ik} + \frac{1}{2}\Delta q_{i,k-1} + \frac{5}{12}\Delta^2 q_{i,k-2} + \frac{3}{8}\Delta^3 q_{i,k-3}, \tag{3.40}$$

. . . .

where

$$(i = 1, 2, \dots, n), \qquad y_{ik} = y_i(x_k),$$
$$x_k = x_0 + kh, \qquad q_{ik} = y_i'(x_k)h,$$
$$\Delta q_{i,k-1} = q_{ik} - q_{i,k-1}, \qquad \Delta^2 q_{i,k-2} = \Delta q_{i,k-1} - \Delta q_{i,k-2},$$
$$\Delta^3 q_{i,k-3} = \Delta^2 q_{i,k-2} - \Delta^2 q_{i,k-3}.$$

The formulas (3.38), (3.39), and (3.40) may be obtained in exactly the same way as for one equation of the first order (see page 68). When using these formulas the order of error remains the same as for one equation.

To start computations by means of Störmer's formula, one has to know the first few values $y_i(x_k)$ which may be found by a Taylor expansion or by Euler's method with a diminished interval; just as in the case of one equation, the precision may be enhanced by applying iteration (see pages 67-68) or by Runge's method.

5. *Runge's method*. The following numbers are computed

$$m_{i1}=f_i(x_k, y_{1k}, y_{2k}, \dots, y_{nk}),$$
$$m_{i2}=f_i\left(x_k+\frac{h}{2}, y_{1k}+\frac{hm_{11}}{2}, y_{2k}+\frac{hm_{21}}{2}, \dots, y_{nk}+\frac{hm_{n1}}{2}\right),$$
$$m_{i3}=f_i\left(x_k+\frac{h}{2}, y_{1k}+\frac{hm_{12}}{2}, y_{2k}+\frac{hm_{22}}{2}, \dots, y_{nk}+\frac{hm_{n2}}{2}\right),$$
$$m_{i4}=f_i(x_k+h, y_{1k}+hm_{13}, y_{2k}+hm_{23}, \dots, y_{nk}+hm_{n3}).$$

Knowing these numbers, we find $y_{i,\,k+1}$ from the formula

$$y_{i,\,k+1}=y_{ik}+\frac{h}{6}(m_{i1}+2m_{i2}+2m_{i3}+m_{i4}) \qquad (i=1, 2, \dots, n).$$

The order of the error is the same as for one equation.

Depending on the required precision of the result, the interval (step) h is roughly chosen with account taken of the order of errors in the formulas used and is improved by means of trial computations with step h and $\frac{h}{2}$. The most reliable approach is to perform the computations with h and $\frac{h}{2}$ of all required values of $y_i(x_k)$, and if the results all coincide within the limits of the given precision, then h is considered to ensure the given accuracy of computations, otherwise the step is again reduced and the computations are performed with step $\frac{h}{2}$ and $\frac{h}{4}$, etc. With a properly chosen step h, the differences Δq_{ik}, $\Delta^2 q_{ik}$, ... should vary smoothly, and the last differences in Störmer's formulas should affect only reserve decimals.

PROBLEMS ON CHAPTER 3

1. $\frac{dx}{dt}=y, \quad \frac{dy}{dt}=-x, \quad x(0)=0, \quad y(0)=1.$
2. $\frac{d^2x_1}{dt^2}=x_2, \quad \frac{d^2x_2}{dt^2}=x_1, \quad x_1(0)=2, \quad \dot{x}_1(0)=2,$
   $x_2(0)=2, \quad \dot{x}_2(0)=2.$
3. $\frac{dx}{dt}+5x+y=e^t, \quad \frac{dy}{dt}-x-3y=e^{2t}.$
4. $\frac{dx}{dt}=y, \quad \frac{dy}{dt}=z, \quad \frac{dz}{dt}=x.$
5. $\frac{dx}{dt}=y, \quad \frac{dy}{dt}=\frac{y^2}{x}.$
6. $\frac{dx}{dt}+\frac{dy}{dt}=-x+y+3, \quad \frac{dx}{dt}-\frac{ay}{dt}=x+y-3.$

7. $\frac{dy}{dx}=\frac{z}{x}, \quad \frac{dz}{dx}=-xy.$

8. $\frac{dx}{z-y}=\frac{dy}{x-z}=\frac{dz}{y-x}.$

9. $\frac{dx}{dt}=-x+y+z, \quad \frac{dy}{dt}=x-y+z, \quad \frac{dz}{dt}=x+y-z.$

10. $t\frac{dx}{dt}+y=0, \quad t\frac{dy}{dt}+x=0.$

11. $\frac{dx}{dt}=y+1, \quad \frac{dy}{dt}=-x+\frac{1}{\sin t}.$

12. $\frac{dx}{dt}=\frac{y}{x-y}, \quad \frac{dy}{dt}=\frac{x}{x-y}.$

13. $\dot{x}+y=\cos t, \quad \dot{y}+x=\sin t$

14. $\dot{x}+3x-y=0, \quad \dot{y}-8x+y=0, \quad x(0)=1, \quad y(0)=4.$

15. $\frac{d^2\theta}{dt^2}+\sin\theta=0$ for $t=0, \quad \theta=\frac{\pi}{36}, \quad \frac{d\theta}{dt}=0.$

Determine $\theta(1)$ to within 0.001.

16. $\dot{x}(t)=ax-y, \quad \dot{y}(t)=x+ay;$ a is a constant.

17. $\dot{x}+3x+4y=0, \quad \dot{y}+2x+5y=0.$

18. $\dot{x}=-5x-2y, \quad \dot{y}=x-7y$

19. $\dot{x}=y-z, \quad \dot{y}=x+y, \quad \dot{z}=x+z$

20. $\dot{x}-y+z=0, \quad \dot{y}-x-y=t, \quad \dot{z}-x-z=t.$

21. $\frac{dx}{x(y-z)}=\frac{dy}{y(z-x)}=\frac{dz}{z(x-y)}.$

22. $\frac{dx}{x(y^2-z^2)}=\frac{dy}{y(z^2-x^2)}=\frac{dz}{z(x^2-y^2)}.$

23. $\dot{X}=AX,$ where $X=\begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix}$ and $A=\begin{Vmatrix} 2 & -1 \\ 3 & -2 \end{Vmatrix}.$

CHAPTER 4

Theory of stability

1. Fundamentals

In order to make a real phenomenon amenable to mathematical description, one unavoidably has to simplify and idealize it, isolating and taking into account only the more essential of the influencing factors and rejecting all other less essential factors. Inescapably, one is confronted by the question of how well the simplifying suppositions have been chosen. It may very well be that ignored factors produce a strong effect on the phenomenon under consideration, substantially altering its quantitative and even qualitative characteristics. The question is finally resolved by practice — the correspondence between the conclusions obtained and experimental findings — yet in many cases it is possible to indicate the conditions under which certain simplifications are definitely impossible.

If a phenomenon is described by the following system of differential equations

$$\frac{dy_i}{dt} = \Phi_i(t, y_1, y_2, \ldots, y_n) \quad (i = 1, 2, \ldots, n) \qquad (4.1)$$

with initial conditions $y_i(t_0) = y_{i0}$ $(i = 1, 2, \ldots, n)$, which are ordinarily the results of measurements and, hence, are inevitably obtained with a certain error, the question naturally arises as to the effect of small changes in the initial values on the desired solution.

If it turns out that arbitrarily small changes in the initial data are capable of producing a substantial change in the solution, then the solution defined by the chosen inaccurate initial data is ordinarily devoid of any practical meaning and cannot describe the given phenomenon even approximately.

This brings us to the important question of finding the conditions under which a sufficiently small change in the initial values brings about an arbitrarily small change in the solution.

If t varies on a finite interval $t_0 \leqslant t \leqslant T$, the answer to this question is given by the theorem on the continuous dependence of solutions on the initial values (see pages 58-59). But if t can take on arbitrarily large values, then the problem is dealt with by the theory of stability.

The solution $\varphi_i(t)\,(i=1, 2, \ldots, n)$ of the system (4.1) is called stable, or, more precisely, *Lyapunov stable*, if for any $\varepsilon > 0$ we can choose a $\delta(\varepsilon) > 0$ such that for any solution $y_i(t)\,(i=1, 2, \ldots, n)$ of that system the initial values of which satisfy the inequalities

$$|y_i(t_0)-\varphi_i(t_0)| < \delta(\varepsilon) \qquad (i=1, 2, \ldots, n),$$

for all $t \geqslant t_0$, the inequalities

$$|y_i(t)-\varphi_i(t)| < \varepsilon \quad (i=1, 2, \ldots, n) \tag{4.2}$$

hold true; that is, solutions that are close for initial values remain close for all $t \geqslant t_0$.

Note. If the system (4.1) satisfies the conditions of the theory on the continuous dependence of solutions upon the initial values, then in the definition of stability we can write $t \geqslant T \geqslant t_0$ in place of $t \geqslant t_0$, since by virtue of this theorem the solutions on the interval $t_0 \leqslant t \leqslant T$ remain close for sufficiently close initial values.

If, given an arbitrarily small $\delta > 0$, the inequalities (4.2) are not fulfilled for at least one solution $y_i(t)\,(i=1, 2, \ldots, n)$, then the solution $\varphi_i(t)$ is called *unstable*. Unstable solutions are rarely of interest in practical problems.

If a solution $\varphi_i(t)\,(i=1, 2, \ldots, n)$ is not only stable but, in addition, satisfies the condition

$$\lim_{t\to\infty} |y_i(t)-\varphi_i(t)| = 0, \tag{4.3}$$

if $|y_i(t_0)-\varphi_i(t_0)| < \delta_1$, $\delta_1 > 0$, then the solution $\varphi_i(t)\,(i=1, 2, \ldots, n)$ is called *asymptotically stable*.

Note that the stability of a solution $\varphi_i(t)\,(i=1, 2, \ldots, n)$ does not yet follow from the single condition (4.3).

Example 1. Test for stability the solution of the differential equation $\frac{dy}{dt} = -a^2y$, $a \neq 0$, defined by the initial condition $y(t_0) = y_0$.

The solution

$$y = y_0 e^{-a^2(t-t_0)}$$

is asymptotically stable since

$$|y_0 e^{-a^2(t-t_0)} - \bar{y}_0 e^{-a^2(t-t_0)}| = e^{-a^2(t-t_0)} |y_0 - \bar{y}_0| < \varepsilon$$

for $t \geqslant t_0$ if $|y_0-\bar{y}_0| < \varepsilon e^{-a^2 t_0}$ and

$$\lim_{t\to\infty} e^{-a^2(t-t_0)} |y_0-\bar{y}_0| = 0.$$

Example 2. Test for stability the solution of the equation $\frac{dy}{dt} = a^2y$, $a \neq 0$, defined by the condition $y(t_0) = y_0$.

The solution $y = y_0 e^{a^2(t-t_0)}$ is unstable since it is impossible to choose a $\delta > 0$ so small that there should follow from the inequality $|\bar{y}_0 - y_0| < \delta(\varepsilon)$

$$|\bar{y}_0 e^{a^2(t-t_0)} - y_0 e^{a^2(t-t_0)}| < \varepsilon$$

or

$$e^{a^2(t-t_0)} |\bar{y}_0 - y_0| < \varepsilon$$

for all $t \geqslant t_0$.

Investigating the stability of some solution

$$y_i = \bar{y}_i(t) \qquad (i = 1, 2, \ldots, n)$$

of the system of equations

$$\frac{dy_i}{dt} = \Phi_i(t, y_1, y_2, \ldots, y_n) \qquad (i = 1, 2, \ldots, n) \tag{4.1}$$

may be reduced to investigating the stability of a trivial solution: a *rest point* located at the coordinate origin.

Indeed, transform the system of equations (4.1) to new variables, putting

$$x_i = y_i - \bar{y}_i(t) \qquad (i = 1, 2, \ldots, n). \tag{4.4}$$

The new unknown functions x_i are the deviations $y_i - \bar{y}_i(t)$ of the earlier unknown functions from the functions $\bar{y}_i(t)$ that define the solution being tested for stability.

By virtue of (4.4), the system (4.1) in new variables takes the form

$$\frac{dx_i}{dt} = -\frac{d\bar{y}_i}{dt} + \Phi_i(t, x_1 + \bar{y}_1(t), x_2 + \bar{y}_2(t), \ldots, x_n + \bar{y}_n(t))$$
$$(i = 1, 2, \ldots, n). \tag{4.5}$$

It is obvious that to the solution (being tested for stability) $y_i = \bar{y}_i t$ $(i = 1, 2, \ldots, n)$ of the system (4.1) there corresponds the trivial solution $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system (4.5) by virtue of the dependence $x_i = y_i - \bar{y}_i(t)$; investigating the stability of the solution $y_i = \bar{y}_i(t)$ $(i = 1, 2, \ldots, n)$ of the system (4.1) may be replaced by the stability testing of the trivial solution of the system (4.5). Therefore, henceforward we can take it, without loss of generality, that we test for stability the trivial solution or, what is the same thing, the rest point of the system of equations located at the coordinate origin.

Let us formulate the conditions of stability as applied to the rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$.

The rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system (4.5) is stable in the sense of Lyapunov if for each $\varepsilon > 0$ it is possible to choose a $\delta(\varepsilon) > 0$ such that from the inequality

$$|x_i(t_0)| < \delta(\varepsilon) \quad (i = 1, 2, \ldots, n)$$

there follows

$$|x_i(t)| < \varepsilon \qquad (i = 1, 2, \ldots, n) \quad \text{for} \quad t \geqslant T \geqslant t_0.$$

Or, somewhat differently: the rest point $x_i \equiv 0\,(i = 1, 2, \ldots, n)$ is stable in the sense of Lyapunov if for every $\varepsilon > 0$ it is possible to choose a $\delta_1(\varepsilon) > 0$ such that from the inequality

$$\sum_{i=1}^{n} x_i^2(t_0) < \delta_1^2(\varepsilon)$$

there follows

$$\sum_{i=1}^{n} x_i^2(t) < \varepsilon^2$$

for $t \geqslant T$; that is, the trajectory, whose initial point lies in the δ_1-neighbourhood of the coordinate origin for $t \geqslant T$, does not go beyond the ε-neighbourhood of the origin.

2. Elementary Types of Rest Points

We investigate the position of trajectories in the neighbourhood of the rest point $x = 0$, $y = 0$ of a system of two homogeneous linear equations with constant coefficients:

$$\left.\begin{aligned} \frac{dx}{dt} &= a_{11}x + a_{12}y, \\ \frac{dy}{dt} &= a_{21}x + a_{22}y, \end{aligned}\right\} \tag{4.6}$$

where

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

We seek the solution in the form $x = \alpha_1 e^{kt}$, $y = \alpha_2 e^{kt}$ (see page 201). To determine k we get the characteristic equation

$$\begin{vmatrix} a_{11} - k & a_{12} \\ a_{21} & a_{22} - k \end{vmatrix} = 0$$

or

$$k^2 - (a_{11} + a_{22})\,k + (a_{11}a_{22} - a_{21}a_{12}) = 0.$$

To within a constant factor, α_1 and α_2 are determined from one of the equations:

$$\left.\begin{aligned} (a_{11} - k)\,\alpha_1 + a_{12}\alpha_2 &= 0, \\ a_{21}\alpha_1 + (a_{22} - k)\,\alpha_2 &= 0. \end{aligned}\right\} \tag{4.7}$$

We consider the following cases:

(a) *The roots k_1 and k_2 of the characteristic equation are real and distinct.*

The general solution is of the form

$$\left.\begin{aligned} x &= c_1\alpha_1 e^{k_1 t} + c_2\beta_1 e^{k_2 t}, \\ y &= c_1\alpha_2 e^{k_1 t} + c_2\beta_2 e^{k_2 t}, \end{aligned}\right\} \tag{4.8}$$

where α_i and β_i are constants determined from equations (4.7) for $k=k_1$ and for $k=k_2$, respectively, and c_1 and c_2 are arbitrary constants.

We then have the following cases:

(1) If $k_1<0$ and $k_2<0$, then the rest point $x=0$, $y=0$ is asymptotically stable, since due to the presence of factors $e^{k_1 t}$ and $e^{k_2 t}$ in (4.8) all the points lying in any δ-neighbourhood of the

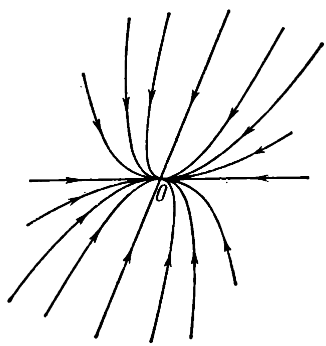


Fig. 4-1

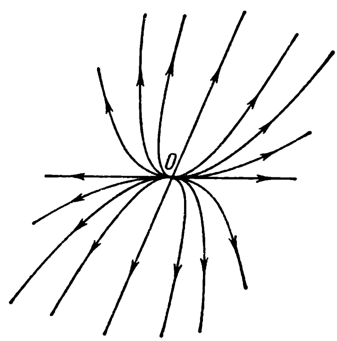


Fig. 4-2

origin at the initial time $t=t_0$ pass into points lying in an arbitrarily small ε-neighbourhood of the origin (given sufficiently large t), and as $t \to \infty$ they tend to the origin. Fig. 4.1 depicts the arrangement of trajectories about a rest point of the type under consideration; it is called a *stable nodal point*. The arrows indicate the direction of motion along the trajectories as t increases.

(2) Let $k_1>0$, $k_2>0$. This case passes into the preceding one when t is replaced by $-t$. Hence, the trajectories have the same shape as in the preceding case, but the point moves along the trajectories in the opposite direction (Fig. 4.2). It is obvious that as t increases, points that are arbitrarily close to the origin recede from the ε-neighbourhood of the origin—the rest point is unstable in the sense of Lyapunov. This type of rest point is called an *unstable nodal point*.

(3) If $k_1 > 0$, $k_2 < 0$, then the rest point is also unstable, since a point moving along the trajectory

$$x = c_1\alpha_1 e^{k_1 t}, \quad y = c_1\alpha_2 e^{k_1 t} \tag{4.9}$$

for arbitrarily small values of c_1 goes out of the ε-neighbourhood of the origin as t increases.

Note that in the case under consideration there are motions which approach the origin, namely:

$$x = c_2\beta_1 e^{k_2 t}, \qquad y = c_2\beta_2 e^{k_2 t}.$$

Given different values of c_2, we get different motions along one and the same straight line $y = \frac{\beta_2}{\beta_1} x$. As t increases, the points on this straight line move in the direction of the origin (Fig. 4.3). Note

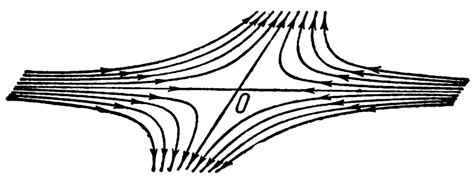


Fig. 4-3

further that the points of the trajectory (4.9) move, as t increases, along the straight line $y = \frac{\alpha_2}{\alpha_1} x$ receding from the origin. But if $c_1 \neq 0$ and $c_2 \neq 0$, then both as $t \to \infty$ and as $t \to -\infty$, the trajectory leaves the neighbourhood of the rest point.

A rest point of this type is called a *saddle point* (Fig. 4.3) because the arrangement of trajectories in the neighbourhood of such a point resembles the arrangement of level lines in the neighbourhood of a saddle point of some surface

$$z = f(x, y).$$

(b) *The roots of the characteristic equation are complex.*

$$k_{1,2} = p \pm qi, \qquad q \neq 0.$$

The general solution of this system may be represented in the form (see page 204)

$$\left.\begin{aligned} x &= e^{pt}(c_1 \cos qt + c_2 \sin qt), \\ y &= e^{pt}(c_1^* \cos qt + c_2^* \sin qt), \end{aligned}\right\} \tag{4.10}$$

where c_1 and c_2 are arbitrary constants and c_1^* and c_2^* are certain linear combinations of these constants.

The following cases are then possible:

(1) $k_{1,2} = p \pm qi, \quad p < 0, \quad q \neq 0.$

The factor e^{pt}, $p < 0$, tends to zero with increasing t, and the second one, the periodic factor in equations (4.10), remains bounded.

If p were equal to zero, then the trajectories would, by virtue of the periodicity of the second factors on the right side of (4.10), be closed curves circling round the rest point $x = 0$, $y = 0$ (Fig. 4.4). The presence of the factor e^{pt}, $p < 0$, tending to zero as t increases

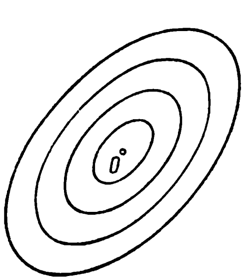


Fig. 4-4

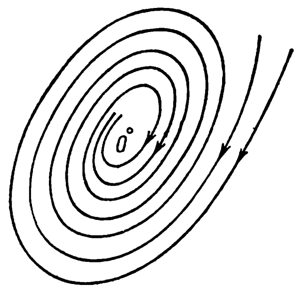


Fig. 4-5

converts the closed curves into spirals, which, as $t \to \infty$, asymptotically approach the coordinate origin (Fig. 4.5); given a sufficiently large t, the points which at $t = t_0$ were located in any δ-neighbourhood of the origin go into a specified ε-neighbourhood of the rest point $x = 0$, $y = 0$, and tend to the rest point as t increases further. Hence, the rest point is asymptotically stable; it is called a *stable focal point*. A focal point differs from a nodal point in that a tangent to the trajectories does not tend to a specific limit as the point of tangency approaches the rest point.

(2) $k_{1,2} = p \pm qi, \quad p > 0, \quad q \neq 0.$

This case passes into the preceding one when t is replaced by $-t$. Consequently, the trajectories do not differ from the trajectories of the preceding case, but motion along them occurs in the opposite direction as t increases (Fig. 4.6). Because of the presence of the increasing factor e^{pt}, points which at the initial instant were arbitrarily close to the origin leave the ε-neighbourhood of the origin as t increases; the rest point is unstable. It is termed an *unstable focal point*.

(3) $k_{1,2} = \pm qi, \quad q \neq 0.$

As has already been mentioned, due to the periodicity of the solutions, the trajectories are closed curves containing within them the rest point (Fig. 4.4), which in this case is called a *centre*. The centre is a stable rest point, since for a given $\varepsilon > 0$ it is possible to choose a $\delta > 0$ such that closed trajectories whose initial points lie in the δ-neighbourhood of the origin do not go out of the ε-neighbourhood of the origin or, what is the same thing, it is possible to choose such small c_1 and c_2 that the solutions

$$\left.\begin{aligned} x &= c_1 \cos qt + c_2 \sin qt, \\ y &= c_1^* \cos qt + c_2^* \sin qt \end{aligned}\right\} \tag{4.11}$$

will satisfy the inequality

$$x^2(t) + y^2(t) < \varepsilon^2.$$

Note, however, that there is no asymptotic stability in this case, since $x(t)$ and $y(t)$ in (4.11) do not tend to zero as $t \to \infty$.

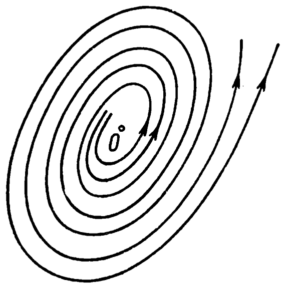


Fig. 4-6

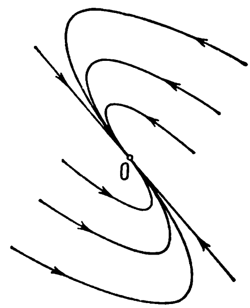


Fig. 4-7

(c) *Roots that are multiples of* $k_1 = k_2$.

(1) $k_1 = k_2 < 0$.

The general solution is of the form

$$x(t) = (c_1\alpha_1 + c_2\beta_1 t)\, e^{k_1 t},$$
$$y(t) = (c_1\alpha_2 + c_2\beta_2 t)\, e^{k_1 t};$$

the possibility of $\beta_1 = \beta_2 = 0$ is not excluded, but then α_1 and α_2 will be arbitrary constants.

Due to the presence of the factor $e^{k_1 t}$ rapidly tending to zero as $t \to \infty$, the product $(c_1\alpha_i + c_2\beta_i t)\, e^{k_1 t}$ $(i = 1, 2)$ tends to zero as $t \to \infty$, and for a sufficiently large t all the points of any δ-neighbourhood of the origin enter the given ε-neighbourhood of the origin, and, hence, the rest point is asymptotically stable. Fig. 4.7 depicts this type of rest point, which [like in Case (a) of (1)] is called a *stable nodal point*. This nodal point occupies an intermediate position between the nodal point (a) (1) and the focal point (b) (1), since for an arbitrarily small change in the real coefficients a_{11}, a_{12}, a_{21}, a_{22} it can turn either into a stable focal point or into a stable nodal point of type (a) (1) because in the case of an arbitrarily small change in the coefficients the multiple root can pass either into a pair of complex conjugate roots or into a pair of real distinct roots. If $\beta_1 = \beta_2 = 0$, we again get a stable nodal point (the so-called proper node) depicted in Fig. 4.8.

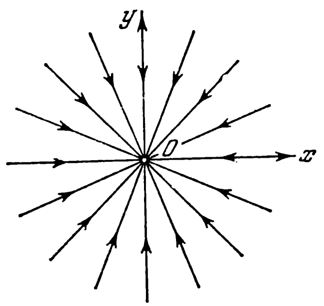


Fig. 4-8

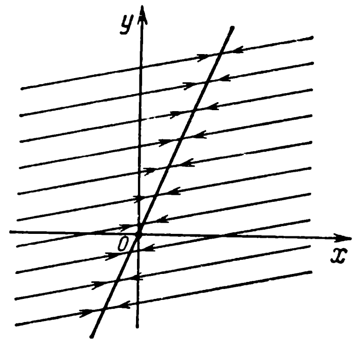


Fig. 4-9

(2) If $k_1 = k_2 > 0$, then changing t to $-t$ leads to the preceding case. Hence, the trajectories do not differ from the trajectories of the preceding case shown in Figs. 4.7 and 4.8, but motion along them occurs in the opposite direction. In this case the rest point is called an *unstable nodal point*, like in case (a) (2).

Thus all the possibilities have been exhausted, since the case $k_1 = 0$ (or $k_2 = 0$) is excluded by the condition

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

Note 1. If

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 0,$$

then the characteristic equation

$$\begin{vmatrix} a_{11}-k & a_{12} \\ a_{21} & a_{22}-k \end{vmatrix} = 0$$

has a zero root $k_1 = 0$. Assume that $k_1 = 0$, but $k_2 \neq 0$. Then the general solution of the system (4.6) will be of the form

$$x = c_1\alpha_1 + c_2\beta_1 e^{k_2 t},$$
$$y = c_1\alpha_2 + c_2\beta_2 e^{k_2 t}.$$

Eliminating t, we get a family of parallel straight lines $\beta_1(y-c_1\alpha_2)=\beta_2(x-c_1\alpha_1)$. When $c_2=0$ we get a one-parameter family of rest points located on the straight line $\alpha_1 y=\alpha_2 x$. If $k_2<0$, then as $t\to\infty$ the points on every trajectory approach the rest point $x=c_1\alpha_1$, $y=c_1\alpha_2$ lying on this trajectory (Fig. 4.9). The rest point $x\equiv 0$, $y\equiv 0$ is stable, but there is no asymptotic stability.

But if $k_2>0$, then the trajectories are arranged in the same way, but points on the trajectories move in the opposite direction, and the rest point $x\equiv 0$, $y\equiv 0$ is unstable.

But if $k_1=k_2=0$, then two cases are possible:

1. The general solution of the system (4.6) is of the form $x=c_1$, $y=c_2$—all the points are rest points and all the solutions are stable.
2. The general solution is of the form

$$x=c_1+c_2t, \quad y=c_1^*+c_2^*t,$$

where c_1^* and c_2^* are linear combinations of arbitrary constants c_1 and c_2. The rest point $x\equiv 0$, $y\equiv 0$ is unstable.

Note 2. The classification of rest points is closely associated with the classification of singular points (see pages 62-64).

Indeed, in the case under consideration the system

$$\left.\begin{aligned}\frac{dx}{dt}&=a_{11}x+a_{12}y,\\ \frac{dy}{dt}&=a_{21}x+a_{22}y,\end{aligned}\right\} \tag{4.6}$$

where

$$\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}\neq 0,$$

may be reduced, by eliminating t, to the equation

$$\frac{dy}{dx}=\frac{a_{21}x+a_{22}y}{a_{11}x+a_{12}y}, \tag{4.12}$$

the integral curves of which coincide with the trajectories of motion of the system (4.6). Here, the rest point $x=0$, $y=0$ of the system (4.6) is a singular point of the equation (4.12).

It will be noted that if both roots of the characteristic equation have a negative real part [cases (a) (1); (b) (1); (c) (1)], then the rest point is asymptotically stable. But if at least one root of the characteristic equation has a positive real part [cases (a) (2); (a) (3); (b) (2); (c) (2)], then the rest point is unstable.

Similar assertions hold true also for a system of n homogeneous linear equations with constant coefficients:

$$\frac{dx_i}{dt}=\sum_{j=1}^{n}a_{ij}x_j \qquad (i=1, 2, \ldots, n). \tag{4.13}$$

If the real parts of all roots of the characteristic equation of the system (4.13) are negative, then the trivial solution $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ is asymptotically stable.

Indeed, the particular solutions corresponding to a certain root k_s of the characteristic equation are of the form (pages 201 and 203)

$$x_i = \alpha_i e^{k_s t} \qquad (i = 1, 2, \ldots, n)$$

if k_s are real,

$$x_j = e^{p_s t} (\beta_j \cos q_s t + \gamma_j \sin q_s t)$$

if $k_s = p_s + q_s i$ and, finally, in the case of multiple roots, the solutions are of the same kind, but multiplied by certain polynomials $P_j(t)$. It is obvious that all solutions of this kind, if the real parts of the roots are negative ($p_s < 0$, or if k_s is real, then $k_s < 0$) tend to zero as $t \to \infty$ not more slowly than ce^{-mt}, where c is a constant factor and $-m < 0$ and greater than the greatest real part of the roots of the characteristic equation. Consequently, given sufficiently large t, the trajectory points whose initial values lie in any δ-neighbourhood of the origin enter an arbitrarily small ε-neighbourhood of the origin and, as $t \to \infty$, approach the origin without bound: the rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ is asymptotically stable.

But if the real part of at least one root of the characteristic equation is positive, $\operatorname{Re} k_i = p_i > 0$, then a solution of the form $x_j = c\alpha_j e^{k_i t}$ corresponding to this root, or in the case of a complex k_i, its real (or imaginary) part $ce^{p_i t} (\beta_j \cos q_i t + \gamma_j \sin q_i t)$ $(j = 1, 2, \ldots, n)$, no matter how small the absolute values of c, increases in absolute value without bound as t increases; and, consequently, the points located at the initial moment on these trajectories in an arbitrarily small δ-neighbourhood of the origin leave any specified ε-neighbourhood of the origin as t increases. Hence, if the real part of at least one root of the characteristic equation is positive, then the rest point $x_j \equiv 0$ $(j = 1, 2, \ldots, n)$ of the system (4.13) is unstable.

Example 1. What type of rest point does the following system of equations have?

$$\frac{dx}{dt} = x - y,$$

$$\frac{dy}{dt} = 2x + 3y.$$

The characteristic equation

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

or

$$k^2 - 4k + 5 = 0$$

has the roots $k_{1,2}=2\pm i$, hence, the rest point $x=0$, $y=0$ is an unstable focal point.

Example 2. $\ddot{x}=-a^2x-2b\dot{x}$ is the equation of elastic oscillations with account taken of friction or resistance of the medium (for $b>0$). Going over to an equivalent system of equations, we have

$$\dot{x}=y,$$
$$\dot{y}=-a^2x-2by.$$

The characteristic equation is of the form

$$\begin{vmatrix} -k & 1 \\ -a^2 & -2b-k \end{vmatrix}=0 \quad \text{or} \quad k^2+2bk+a^2=0,$$

whence $k_{1,2}=-b\pm\sqrt{b^2-a^2}$.

Consider the following cases:

(1) $b=0$, i.e., the resistance of the medium is ignored. All motions are periodic. The rest point at the origin is a centre.

(2) $b^2-a^2<0$, $b>0$. The rest point is a stable focal point. The oscillations die out.

(3) $b^2-a^2\geqslant 0$, $b>0$. The rest point is a stable nodal point. All solutions are damped and nonoscillating. This case sets in if the resistance of the medium is great ($b\geqslant a$).

(4) $b<0$ (the case of negative friction), $b^2-a^2<0$. The rest point is an unstable focal point.

(5) $b<0$, $b^2-a^2\geqslant 0$ (the case of large negative friction). The rest point is an unstable nodal point.

Example 3. Test for stability the rest point of the system of equations

$$\frac{dx}{dt}=2y-z,$$
$$\frac{dy}{dt}=3x-2z,$$
$$\frac{dz}{dt}=5x-4y.$$

The characteristic equation is of the form

$$\begin{vmatrix} -k & 2 & -1 \\ 3 & -k & -2 \\ 5 & -4 & -k \end{vmatrix}=0$$

or

$$k^3-9k+8=0.$$

In the general case it is rather difficult to determine the roots of a cubic equation; however, in the given case one root $k_1=1$ is readily found, and since this root has a positive real part, we may assert that the rest point $x=0$, $y=0$, $z=0$ is unstable.

3. Lyapunov's Second Method

At the end of last century, the celebrated Russian mathematician Aleksandr Mikhailovich Lyapunov elaborated an extremely general method for investigating the solutions of a system of differential equations for stability:

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n). \qquad (4.14)$$

It became known as *Lyapunov's second method*.

Theorem 4.1 (Lyapunov's stability theorem). *If there exists a differentiable function $v(x_1, x_2, \ldots, x_n)$, called Lyapunov's function, that satisfies the following conditions in the neighbourhood of the coordinate origin:*

(1) $v(x_1, x_2, \ldots, x_n) \geqslant 0$ and $v = 0$ only for $x_i = 0$ $(i = 1, 2, \ldots, n)$, i.e., the function v has a strict minimum at the origin;

(2) $\frac{dv}{dt} = \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i(t, x_1, \ldots, x_n) \leqslant 0$ for $t \geqslant t_0$,

then the rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ is stable.

In condition (2), the derivative $\frac{dv}{dt}$ is taken along the integral curve, i.e. it is computed on the assumption that the arguments x_i $(i = 1, 2, \ldots, n)$ of the function $v(x_1, x_2, \ldots, x_n)$ are replaced by the solution $x_i(t)$ $(i = 1, 2, \ldots, n)$ of the system of differential equations (4.14).

Indeed, on this assumption $\frac{dv}{dt} = \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} \frac{dx_i}{dt}$ or, replacing $\frac{dx_i}{dt}$ by the right sides of the system (4.14), we finally get

$$\frac{dv}{dt} = \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i(t, x_1, x_2, \ldots, x_n).$$

Proof of Lyapunov's stability theorem. In the neighbourhood of the origin, as also in the neighbourhood of any point of a strict minimum (Fig. 4.10), the level surfaces $v(x_1, x_2, \ldots, x_n) = c$ of the function $v(x_1, x_2, \ldots, x_n)$ are closed surfaces, inside of which is located a minimum point, the coordinate origin. Given $\varepsilon > 0$ For a sufficiently small $c > 0$ the surface of the level $v = c$ lies completely in the ε-neighbourhood of the origin,* but does not pass through the coordinate origin; hence, a $\delta > 0$ may be chosen such that the δ-neighbourhood of the origin completely lies inside the

* More precisely, at least one closed component of the surface level $v = c$ lies in the ε-neighbourhood of the origin.

surface $v=c$; and in this neighbourhood $v<c$. If the initial point with coordinates $x_i(t_0)$ $(i=1, 2, \ldots, n)$ is chosen in the δ-neighbourhood of the origin (Fig. 4.11) and, hence, $v(x_1(t_0), x_2(t_0), \ldots$ $\ldots, x_n(t_0))=c_1<c$, then for $t>t_0$ the point of the trajectory defined by these initial conditions cannot go beyond the limits of the ε-neighbourhood of the origin and even beyond the limits of the

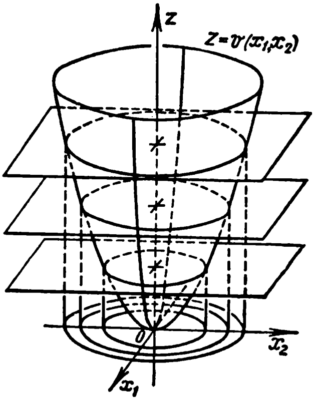


Fig. 4-10

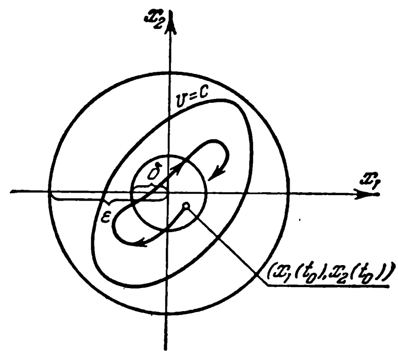


Fig. 4-11

level surface $v=c$, since, by virtue of condition (2) of the theorem, the function v does not increase along the trajectory and, hence, for $t \geqslant t_0$

$$v(x_1(t), x_2(t), \ldots, x_n(t)) \leqslant c_1 < c.$$

Note. Lyapunov proved the stability theorem on more general assumptions; in particular, he assumed that the function v can be dependent on t as well: $v=v(t, x_1, x_2, \ldots, x_n)$. Then, to make the stability theorem hold, the first condition has to be replaced by the following:

$$v(t, x_1, x_2, \ldots, x_n) \geqslant w(x_1, x_2, \ldots, x_n) \geqslant 0$$

in the neighbourhood of the origin for $t \geqslant t_0$, where the continuous function w has a strict minimum at the coordinate origin, $v(t, 0, 0, \ldots, 0)=w(0, 0, \ldots, 0)=0$, and the second condition remains unchanged, $\frac{dv}{dt} \leqslant 0$; however, in this case

$$\frac{dv}{dt}=\frac{\partial v}{\partial t}+\sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f(t, x_1, x_2, \ldots, x_n).$$

The scheme of the proof remains the same, except that now we must take into account that by virtue of condition (1), the level surface $v(t, x_1, x_2, \ldots, x_n) = c$, which moves as t varies, remains within the surface of the level $w(x_1, x_2, \ldots, x_n) = c$ for all variations of $t \geqslant t_0$ (Fig. 4.12).

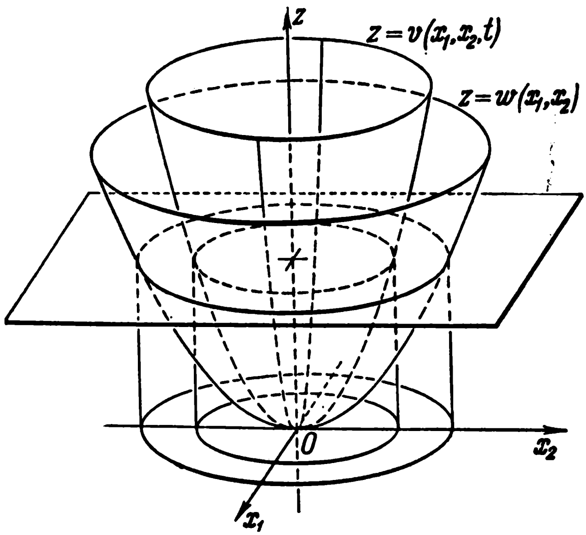


Fig. 4-12

Theorem 4.2 (Lyapunov's theorem on asymptotic stability). *If there exists a differentiable Lyapunov function $v(x_1, x_2, \ldots, x_n)$ that satisfies the conditions:*

(1) *$v(x_1, x_2, \ldots, x_n)$ has a strict minimum at the origin: $v(0, 0, \ldots, 0) = 0$;*

(2) *the derivative of the function v, computed along the integral curves of the system* (4.14)

$$\frac{dv}{dt} = \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i(t, x_1, x_2, \ldots, x_n) \leqslant 0,$$

and outside an arbitrarily small neighbourhood of the origin, that is, for $\sum_{i=1}^{n} x_i^2 \geqslant \delta_1^2 > 0$, $t \geqslant T_0 \geqslant t_0$, the derivative $\frac{dv}{dt} \leqslant -\beta < 0$, where β is a constant, then the rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system (4.14) *is asymptotically stable.*

Proof. Since the conditions of the stability theorem are fulfilled, it follows that for every $\varepsilon > 0$ a $\delta(\varepsilon) > 0$ may be chosen such that the trajectory, the initial point of which lies in the δ-neighbourhood of the origin, does not, for $t \geqslant t_0$, leave the ε-neighbourhood of the origin. Hence, in particular, condition (2) is fulfilled along such a trajectory for $t > T_0$, and for this reason, the function v monotonically decreases along the trajectory as t increases, and along the trajectory there is a limit to the function v as $t \to \infty$:

$$\lim_{t \to \infty} v(t, x_1(t), x_2(t), \ldots, x_n(t)) = \alpha \geqslant 0.$$

We have to prove that $\alpha = 0$, since if $\alpha = 0$, then from condition (1) it follows that $\lim_{t \to \infty} x_i(t) = 0$ $(i = 1, 2, \ldots, n)$, i.e. the rest point $x_i = 0$ $(i = 1, 2, \ldots, n)$ is asymptotically stable. Suppose $\alpha > 0$; then the trajectory, for $t > t_0$, lies in the region $v \geqslant \alpha$, and therefore outside a certain δ_1-neighbourhood of the origin; that is, where we have, by condition (2), $\frac{dv}{dt} \leqslant -\beta < 0$ for $t \geqslant T_0$. Multiplying inequality $\frac{dv}{dt} \leqslant -\beta$ by dt and integrating along the trajectory from T_0 to t, we get

$$v(x_1(t), x_2(t), \ldots, x_n(t)) - v(x_1(T_0), x_2(T_0), \ldots, x_n(T_0)) \leqslant \\ \leqslant -\beta(t - T_0)$$

or

$$v(x_1(t), x_2(t), \ldots, x_n(t)) \leqslant \\ \leqslant v(x_1(T_0), x_2(T_0), \ldots, x_n(T_0)) - \beta(t - T_0).$$

Given a sufficiently large t, the right side is negative, and hence also $v(x_1(t), x_2(t), \ldots, x_n(t)) < 0$, which contradicts condition (1).

Note. The theorem on asymptotic stability is generalized to the case of the function v dependent on $t, x_1, x_2, \ldots, x_n$ if the first condition, as in the preceding theorem, is replaced by the following:

$$v(t, x_1, x_2, \ldots, x_n) \geqslant w(x_1, x_2, \ldots, x_n) \geqslant 0,$$

where the function w has a strict minimum at the origin and, besides, if we ask that the function $v(t, x_1, x_2, \ldots, x_n)$ uniformly approach zero in t as $\sum_{i=1}^{n} x_i^2 \to 0$.

Theorem 4.3 (Chetayev's instability theorem). *If there is a differentiable function $v(x_1, x_2, \ldots, x_n)$ that satisfies the following conditions in a certain closed h-neighbourhood of the coordinate origin (1) in an arbitrarily small neighbourhood U of the origin there exists a region $(v > 0)$ in which $v > 0$, and $v = 0$ on a part* *of the boundary of the region* $(v>0)$ *lying in* U; (2) *in the region* $(v>0)$ *the derivative*

$$\frac{dv}{dt}=\sum_{i=1}^{n}\frac{\partial v}{\partial x_i}f_i(t,\ x_1,\ x_2,\ \ldots,\ x_n)>0,$$

and in the region $(v\geqslant\alpha)$, $\alpha>0$, *the derivative* $\frac{dv}{dt}\geqslant\beta>0$, *then the rest point* $x_i\equiv 0$ $(i=1,\ 2,\ \ldots,\ n)$ *of the system* (4.14) *is unstable.*

Proof. We take the initial point $x_1(t_0),\ x_2(t_0),\ \ldots,\ x_n(t_0)$ in an arbitrarily small neighbourhood of the coordinate origin in the region $(v>0)$, $v(x_1(t_0), x_2(t_0), \ldots$ $\ldots,\ x_n(t_0))=\alpha>0$ (Fig. 4.13). Since $\frac{dv}{dt}\geqslant 0$ along the trajectory, the function v does not diminish along the trajectory and hence as long as the trajectory lies inside the h-neighbourhood of the origin we are interested in (where the conditions of the theorem are fulfilled), the trajectory must lie in the region $(v\geqslant\alpha)$. Assume that the trajectory does not leave the h-neighbourhood of the origin. Then, by virtue of condition (2), the derivative $\frac{dv}{dt}\geqslant\beta>0$ along the trajectory for $t\geqslant t_0$. Multiplying this inequality by dt and integrating, we get

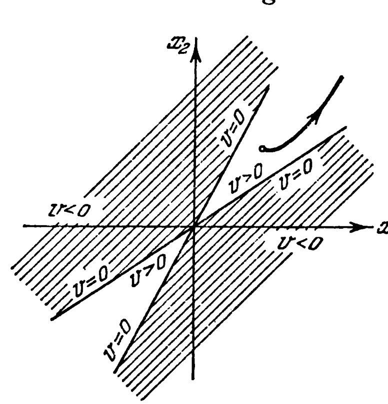


Fig. 4-13

$$v(x_1(t),\ x_2(t),\ \ldots,\ x_n(t))-v(x_1(t_0),\ x_2(t_0),\ \ldots,\ x_n(t_0))\geqslant\beta(t-t_0),$$

whence it follows that as $t\to\infty$ the function v increases without bound along the trajectory; but this runs counter to the assumption that the trajectory does not leave the closed h-neighbourhood of the origin, since the continuous function v is bounded in this h-neighbourhood.

Note. N. G. Chetayev proved the instability theorem on the assumption that v can also depend on t; then the hypothesis is somewhat modified; in particular, we have to demand that the function v be bounded in the region $(v\geqslant 0)$ in the h-neighbourhood under consideration of the origin.

Example 1. Test for stability the trivial solution of the system:

$$\frac{dx}{dt}=-y-x^3,\quad \frac{dy}{dt}=x-y^3.$$

The function $v(x, y)=x^2+y^2$ satisfies the conditions of Lyapunov's theorem on asymptotic stability:

(1) $v(x, y)\geqslant 0,\ v(0, 0)=0$;

(2) $\frac{dv}{dt}=2x(-y-x^3)+2y(x-y^3)=-2(x^4+y^4)\leqslant 0$.

Outside the neighbourhood of the origin, $\frac{dv}{dt}\leqslant-\beta<0$. Consequently, the solution $x\equiv 0,\ y\equiv 0$ is asymptotically stable.

Example 2. Test for stability the trivial solution $x\equiv 0,\ y\equiv 0$ of the system

$$\frac{dx}{dt}=-xy^4; \qquad \frac{dy}{dt}=yx^4.$$

The function $v(x, y)=x^4+y^4$ satisfies the conditions of the Lyapunov stability theorem:

(1) $v(x, y)=x^4+y^4\geqslant 0, \quad v(0, 0)=0$;

(2) $\frac{dv}{dt}=-4x^4y^4+4x^4y^4\equiv 0$.

Therefore, the trivial solution $x\equiv 0,\ y\equiv 0$ is stable.

Example 3. Test for stability the rest point $x\equiv 0,\ y\equiv 0$ of the system of equations

$$\frac{dx}{dt}=y^3+x^5,$$

$$\frac{dy}{dt}=x^3+y^5.$$

The function $v=x^4-y^4$ satisfies the conditions of Chetayev's theorem:

(1) $v>0$ for $|x|>|y|$;

(2) $\frac{dv}{dt}=4x^3(y^3+x^5)-4y^3(x^3+y^5)=4(x^8-y^8)>0$

for $|x|>|y|$; and for $v\geqslant\alpha>0,\ \frac{dv}{dt}\geqslant\beta>0$. Hence, the rest point $x\equiv 0,\ y\equiv 0$ is unstable.

Example 4. Test for stability the trivial solution $x_i\equiv\equiv 0\ (i=1, 2, \ldots, n)$ of the system of equations

$$\frac{dx_i}{dt}=\frac{\partial u(x_1, x_2, \ldots, x_n)}{\partial x_i} \qquad (i=1, 2, \ldots, n)$$

if it is given that the function $u(x_1, x_2, \ldots, x_n)$ has a strict maximum at the coordinate origin.

For the Lyapunov function we take the difference

$$v(x_1, x_2, \ldots, x_n)=u(0, 0, \ldots, 0)-u(x_1, x_2, \ldots, x_n),$$

which obviously vanishes when $x_i = 0$ $(i = 1, 2, \ldots, n)$, has a strict minimum at the origin, and, hence, satisfies condition (1) of Lyapunov's stability theorem. The derivative along the integral curves

$$\frac{dv}{dt} = -\sum_{i=1}^{n} \frac{\partial u}{\partial x_i} \frac{dx_i}{dt} = -\sum_{i=1}^{n} \left(\frac{\partial u}{\partial x_i}\right)^2 \leqslant 0.$$

Thus, the conditions of Lyapunov's stability theorem are fulfilled and so the trivial solution is stable.

Example 5. Test for stability the trivial solution $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system of equations

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)\, x_j, \quad \text{where} \quad a_{ij}(t) = -a_{ji}(t) \text{ for } i \neq j$$
$$\text{and all} \quad a_{ii}(t) \leqslant 0.$$

The trivial solution is stable since the function $v = \sum_{i=1}^{n} x_i^2$ satisfies the conditions of Lyapunov's stability theorem:

(1) $v \geqslant 0$ and $v(0, 0, \ldots, 0) = 0$;

(2) $\frac{dv}{dt} = 2\sum_{i=1}^{n} x_i \frac{dx_i}{dt} = 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(t)\, x_i x_j = 2\sum_{i=1}^{n} a_{ii}(t)\, x_i^2 \leqslant 0.$

4. Test for Stability Based on First Approximation

When testing for stability the rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system of differential equations

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n), \qquad (4.14)$$

where f_i are functions differentiable in the neighbourhood of the coordinate origin, frequent use is made of the following method: taking advantage of the differentiability of the functions $f_i(t, x_1, x_2, \ldots, x_n)$, we represent the system (4.14) in the neighbourhood of the origin $x_i = 0$ $(i = 1, 2, \ldots, n)$ in the form

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)\, x_j + R_i(t, x_1, x_2, \ldots, x_n) \quad (i = 1, 2, \ldots, n), \qquad (4.15)$$

where the R_i are of an order higher than first with respect to $\sqrt{\sum_{i=1}^{n} x_i^2}$, and in place of the rest point $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system (4.15) we test for stability the same rest point of the linear system

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)\, x_j \quad (i = 1, 2, \ldots, n), \tag{4.16}$$

which is called a *system of equations of first approximation* for the system (4.15). The conditions for the applicability of this method, which was used for a long time without any substantiation, were investigated in detail by A. Lyapunov and were later extended in the works of many mathematicians, particularly by O. Perron, I. Malkin, K. Persidsky, N. Chetayev.

Testing for stability a system of equations of first approximation is of course a much easier task than testing the original, generally speaking, nonlinear system; however, even testing the linear system (4.16) with variable coefficients $a_{ij}(t)$ is an extremely complicated problem. But if all the a_{ij} are constant, that is, the system is stationary to a first approximation, then the stability test of the linear system (4.16) does not encounter any fundamental difficulties (see pages 220-221).

Theorem 4.4. *If the system of equations* (4.15) *is stationary to a first approximation, all the terms R_i in a sufficiently small neighbourhood of the coordinate origin for $t \geqslant T \geqslant t_0$ satisfy the inequalities* $|R_i| \leqslant N\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}+\alpha}$, *where N and α are constants, and* $\alpha > 0$ $\Bigl($*i.e., if the R_i do not depend on t, then their order is greater than first with respect to* $\sqrt{\sum_{i=1}^{n} x_i^2}\Bigr)$ *and all the roots of the characteristic equation*

$$\begin{vmatrix} a_{11}-k & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-k & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn}-k \end{vmatrix} = 0 \tag{4.17}$$

have negative real parts, then the trivial solutions $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system of equations (4.15) *and the system of equations* (4.16) *are asymptotically stable; hence in this case a test for stability based on a first approximation is permissible.*

Theorem 4.5. *If the system of equations* (4.15) *is stationary to a first approximation, all the functions R_i satisfy the conditions of the preceding theorem, and at least one root of the characteristic* *equation* (4.17) *has a positive real part, then the rest points* $x_i \equiv 0\ (i=1,\ 2,\ \ldots,\ n)$ *of the system* (4.15) *and the system* (4.16) *are unstable; consequently, this case too permits testing for stability on the basis of a first approximation.*

Theorems 4.4 and 4.5 fail to embrace only the so-called critical case with respect to restrictions imposed on the roots of the characteristic equation: all real parts of the roots of the characteristic equation are nonpositive, and the real part of at least one root is zero.

In the critical case, the nonlinear terms R_i begin to influence the stability of the trivial solution of the system (4.15) and, generally speaking, it is impossible to test for stability on the basis of a first approximation.

Theorems 4.4 and 4.5 are proved in Malkin's book [2].

To give the reader an idea of the methods of proof of such theorems, we prove Theorem 4.4 on the assumption that all the roots k_i of the characteristic equation are real and distinct

$$k_i < 0 \qquad (i=1,\ 2,\ \ldots,\ n), \qquad k_i \neq k_j \quad \text{for} \quad i \neq j.$$

In vector notation, the system (4.15) and the system (4.16) are respectively of the form

$$\frac{dX}{dt} = AX + R, \tag{4.15$_1$}$$

$$\frac{dX}{dt} = AX, \tag{4.16$_1$}$$

where

$$X = \left\| \begin{matrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{matrix} \right\|, \qquad A = \left\| \begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{matrix} \right\|, \qquad R = \left\| \begin{matrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ \cdot \\ R_n \end{matrix} \right\|.$$

With the aid of a nondegenerate linear transformation with constant coefficients $X = BY$, where

$$B = \left\| \begin{matrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{matrix} \right\|, \qquad Y = \left\| \begin{matrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{matrix} \right\|,$$

we transform the system (4.16$_1$) to $B\frac{dY}{dt} = ABY$ or $\frac{dY}{dt} = B^{-1}ABY$.

We choose the matrix B so that the matrix $B^{-1}AB$ is diagonal:

$$B^{-1}AB=\begin{Vmatrix} k_1 & 0 & 0 & \dots & 0 \\ 0 & k_2 & 0 & \dots & 0 \\ 0 & 0 & k_3 & \dots & 0 \\ 0 & 0 & 0 & \dots & k_n \end{Vmatrix}.$$

Then the system (4.16) becomes

$$\frac{dy_i}{dt}=k_iy_i \qquad (i=1,\ 2,\ \dots,\ n)$$

and the system (4.15) passes, in the same transformation, into

$$\frac{dy_i}{dt}=k_iy_i+\bar{R}_i(t,\ y_1,\ y_2,\ \dots,\ y_n) \qquad (i=1,\ 2,\ \dots,\ n), \tag{4.18}$$

where $|\bar{R}_i|\leqslant\bar{N}\left(\sum_{i=1}^{n}y_i^2\right)^{\frac{1}{2}+\alpha}$, $\bar{N}$ is a constant, $\alpha>0$, $t\geqslant T$.

Relative to the system (4.18), the Lyapunov function that satisfies the conditions of the asymptotic stability theorem is

$$v=\sum_{i=1}^{n}y_i^2.$$

Indeed,

(1) $v(y_1,\ y_2,\ \dots,\ y_n)\geqslant 0, \qquad v(0,\ 0,\ \dots,\ 0)=0;$

(2) $\frac{dv}{dt}=2\sum_{i=1}^{n}y_i\frac{dy_i}{dt}=2\sum_{i=1}^{n}k_iy_i^2+2\sum_{i=1}^{n}k_iy_iR_i\leqslant\sum_{i=1}^{n}k_iy_i^2\leqslant 0$

for sufficiently small y_i, since all $k_i<0$, and the double sum $2\sum_{i=1}^{n}k_iy_iR_i$ may, for sufficiently small y_i, be made less, in absolute value, than the sum $\sum_{i=1}^{n}k_iy_i^2$.

Finally, outside the neighbourhood of the coordinate origin

$$\frac{dv}{dt}\leqslant-\beta<0.$$

Example 1. Test for stability the rest point $x=0$, $y=0$ of the system

$$\left.\begin{aligned} \frac{dx}{dt}&=x-y+x^2+y^2\sin t, \\ \frac{dy}{dt}&=x+y-y^2. \end{aligned}\right\} \tag{4.19}$$

The nonlinear terms satisfy the conditions of Theorems 4.4 and 4.5. Test for stability the rest point $x=0$, $y=0$ of the first-approximation system

$$\left.\begin{aligned}\frac{dx}{dt}&=x-y,\\ \frac{dy}{dt}&=x+y.\end{aligned}\right\} \tag{4.20}$$

The characteristic equation $\begin{vmatrix}1-k & -1\\ 1 & 1-k\end{vmatrix}=0$ has the roots $k_{1,2}=$ $=1\pm i$, and so, by virtue of Theorem 4.5, the rest point of the systems (4.19) and (4.20) is unstable.

Example 2. Test for stability the rest point $x=0$, $y=0$ of the system

$$\left.\begin{aligned}\frac{dx}{dt}&=2x+8\sin y,\\ \frac{dy}{dt}&=2-e^x-3y-\cos y.\end{aligned}\right\} \tag{4.21}$$

Expanding $\sin y$, e^x and $\cos y$ by Taylor's formula, we represent the system as

$$\frac{dx}{dt}=2x+8y+R_1,\qquad \frac{dy}{dt}=-x-3y+R_2,$$

where R_1 and R_2 satisfy the conditions of the Theorems 4.4 and 4.5.

The characteristic equation $\begin{vmatrix}2-k & 8\\ -1 & -3-k\end{vmatrix}=0$ for the first-approximation system

$$\frac{dx}{dt}=2x+8y,\quad \frac{dy}{dt}=-x-3y \tag{4.22}$$

has roots with negative real parts. Consequently, the rest point $x=0$, $y=0$ of the systems (4.21) and (4.22) is asymptotically stable.

Example 3. Test for stability the rest point $x=0$, $y=0$ of the system

$$\left.\begin{aligned}\frac{dx}{dt}&=-4y-x^3,\\ \frac{dy}{dt}&=3x-y^3.\end{aligned}\right\} \tag{4.23}$$

The characteristic equation $\begin{vmatrix}-k & -4\\ 3 & -k\end{vmatrix}=0$ for the first-approximation system has pure imaginary roots, which is the critical case.

A first-approximation investigation is impossible. In this case it is easy to choose a Lyapunov function

$$v = 3x^2 + 4y^2.$$

(1) $v(x, y) \geqslant 0$, $v(0, 0) = 0$;

(2) $\frac{dv}{dt} = 6x(-4y - x^3) + 8y(3x - y^3) = -(6x^4 + 8y^4) \leqslant 0$;

note that outside a certain neighbourhood of the origin $\frac{dv}{dt} \leqslant -\beta < 0$, hence the rest point $x = 0$, $y = 0$ is asymptotically stable by the theorem of the preceding section.

Let us examine this example in somewhat more detail. The first-approximation system of equations

$$\frac{dx}{dt} = -4y, \quad \frac{dy}{dt} = 3x \tag{4.24}$$

had the centre at the origin. The nonlinear terms in the system (4.23) converted this centre into a stable focal point.

The general case too exhibits a similar but somewhat more complicated geometrical pattern. Let the first-approximation system of the system

$$\left.\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + R_1(x_1, x_2), \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + R_2(x_1, x_2) \end{aligned}\right\} \tag{4.25}$$

have a rest point of the centre type at the origin. As on page 229, assume that the nonlinear terms $R_1(x_1, x_2)$ and $R_2(x_1, x_2)$ are of order higher than first with respect to $\sqrt{x_1^2 + x_2^2}$. These nonlinear terms are small, in a sufficiently small neighbourhood of the origin, in comparison with linear terms, but still they somewhat distort the direction field defined by the first-approximation linear system. For this reason, a trajectory emanating from some point (x_0, y_0) is slightly displaced (after a circuit of the origin) from the linear-system trajectory passing through the same point, and, generally speaking, does not arrive at the point (x_0, y_0). The trajectory is not closed.

If after such a circuit of the origin all the trajectories approach the origin, then a stable focal point arises at the origin; but if the trajectories recede from the origin, an unstable focal point develops.

An exceptional case is possible in which all the trajectories of the nonlinear system which are located in the neighbourhood of the origin remain closed; however, the most typical case is that in which only certain closed curves (or, possibly, none) remain closed, while the others are converted into spirals.

Such closed trajectories in the neighbourhood of which all trajectories are spirals are called *limit cycles*.

If the trajectories close to the limit cycle are spirals that approach the limit cycle as $t \to \infty$, then the limit cycle is called *stable* (Fig. 4.14). If the trajectories close to the limit cycle are spirals receding from the limit cycle as $t \to \infty$, then the limit cycle is called *unstable*. And if the spirals approach the limit cycle from one side as $t \to \infty$, and recede from it on the other side (Fig. 4.15), then the limit cycle is called *half-stable*.

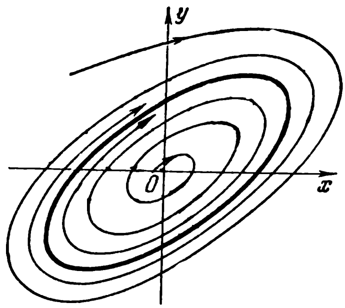


Fig. 4-14

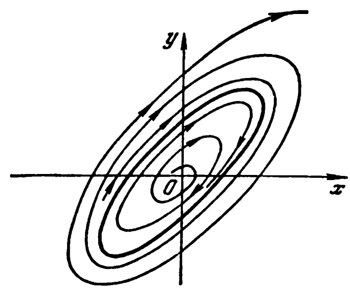


Fig. 4-15

Thus, transition from the first-approximation system (4.16) to the system (4.25), generally speaking, leads to a transformation of the centre into a focal point surrounded by p (the case $p=0$ is not excluded) limit cycles.

On pages 160-161, when studying the periodic solutions of the autonomous quasi-linear system

$$\ddot{x} + a^2 x = \mu f(x, \dot{x}, \mu), \tag{4.26}$$

we encountered a similar instance. Indeed, replacing (4.26) by an equivalent system, we get

$$\left.\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -a^2 x + \mu f(x, y, \mu). \end{aligned}\right\} \tag{4.27}$$

The corresponding linear system:

$$\dot{x} = y, \qquad \dot{y} = -a^2 x$$

has a rest point of the centre type at the origin; the addition of small (for small μ) nonlinear terms converts the centre, generally speaking, into a focal point surrounded by several limit cycles whose radii were found from the equation (2.128), page 162.

The only difference between the cases (4.25) and (4.27) consists in the fact that the terms R_1 and R_2 are small only in a sufficiently small neighbourhood of the origin, whereas in the case (4.27) the term $\mu f(x, y, \mu)$ can be made small, for a sufficiently small μ, not only in a sufficiently small neighbourhood of the coordinate origin.

In Example 2 (page 162), a limit cycle appears, for small μ, in the neighbourhood of a circle of radius 6 with centre at the coordinate origin, which circle is the trajectory of the generating equation.

In applications, stable limit cycles are usually found to correspond to auto-oscillatory processes, i.e. periodic processes in which small perturbations practically do not alter the amplitude and frequency of the oscillations.

5. Criteria of Negativity of the Real Parts of All Roots of a Polynomial

In the preceding section, the problem of the stability of a trivial solution of a broad class of systems of differential equations was reduced to investigating the signs of the real parts of the roots of the characteristic equation.

If the characteristic equation has a high degree, then its solution is complicated; for this reason, very important are methods which permit establishing (without solving the equation) whether all its roots have a negative real part or not.

Theorem 4.6 (Hurwitz's theorem)*. *A necessary and sufficient condition for the negativity of the real parts of all the roots of the polynomial*

$$z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n$$

with real coefficients is the positivity of all the principal diagonals of the minors of the Hurwitz matrix

$$\left\| \begin{matrix} a_1 & 1 & 0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & 1 & \dots & 0 \\ a_5 & a_4 & a_3 & a_2 & \dots & \\ a_7 & a_6 & a_5 & a_4 & \dots & \\ . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & \dots & a_n \end{matrix} \right\| .$$

The principal diagonal of the Hurwitz matrix exhibits the coefficients of the polynomial under consideration in the order of their numbers from a_1 to a_n. The columns alternately consist of coeffici-

* The proof of the Hurwitz theorem may be found in courses of higher algebra (for example, A. Kurosh *Course of Higher Algebra*).

ents with odd only or even only indices, including the coefficient $a_0 = 1$; hence the matrix element $b_{ik} = a_{2i-k}$. All missing coefficients (that is, coefficients with indices greater than n or less than 0) are replaced by zeros.

Denote the principal diagonal minors of the Hurwitz matrix:

$$\Delta_1 = |a_1|, \ \Delta_2 = \begin{vmatrix} a_1 & 1 \\ a_3 & a_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \ \ldots$$

$$\ldots, \ \Delta_n = \begin{vmatrix} a_1 & 1 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & \ldots & \ldots \\ a_5 & a_4 & a_3 & \ldots & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & a_n \end{vmatrix}$$

Observe that since $\Delta_n = \Delta_{n-1} a_n$, the last of the Hurwitz conditions $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_n > 0$ may be replaced by the demand that $a_n > 0$*.

Let us apply the Hurwitz theorem to polynomials of second, third, and fourth degree.

(a) $z^2 + a_1 z + a_2$.

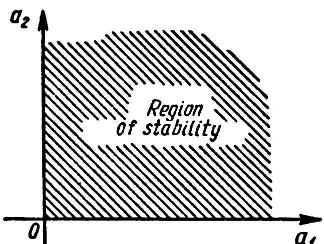


Fig. 4-16

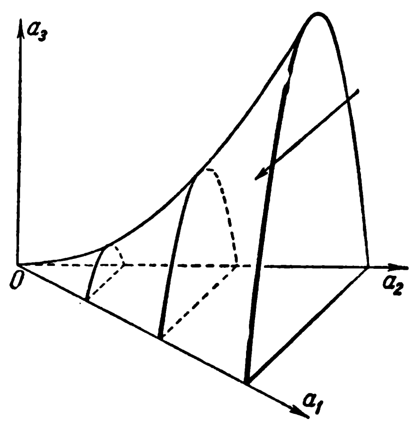


Fig. 4-17

The Hurwitz conditions reduce to $a_1 > 0$, $a_2 > 0$. These inequalities define the first quadrant in the space of the coefficients a_1 and a_2 (Fig. 4.16). Fig. 4-16 depicts the region of asymptotic stability of a trivial solution of some system of differential equations

* Note that from the Hurwitz conditions it follows that all the $a_i > 0$; however, the positivity of all coefficients is not enough for the real parts of all roots to be negative.

that satisfies the conditions of Theorem 4.1, provided that z^2+a_1z+ $+a_2$ is its characteristic polynomial.

(b) $z^3+a_1z^2+a_2z+a_3$.

The Hurwitz conditions reduce to $a_1>0$, $a_1a_2-a_3>0$, $a_3>0$. The region defined by this inequality in the coefficient space is depicted in Fig. 4-17.

(c) $z^4+a_1z^3+a_2z^2+a_3z+a_4$.

The Hurwitz conditions reduce to

$$a_1>0, \quad a_1a_2-a_3>0, \quad (a_1a_2-a_3)a_3-a_1^2a_4>0, \quad a_4>0.$$

The Hurwitz conditions are very convenient and readily verifiable for the polynomials we have just considered. But the Hurwitz conditions rapidly become complicated as the degree of the polynomial increases, and it is often more convenient to apply other criteria for the negativity of the real parts of the roots of a polynomial.

Example. For what values of the parameter α is the trivial solution $x_1=0$, $x_2=0$, $x_3=0$ of the system of differential equations

$$\frac{dx_1}{dt}=x_3, \quad \frac{dx_2}{dt}=-3x_1, \quad \frac{dx_3}{dt}=\alpha x_1+2x_2-x_3$$

asymptotically stable?

The characteristic equation is of the form

$$\begin{vmatrix} -k & 0 & 1 \\ -3 & -k & 0 \\ \alpha & 2 & -1-k \end{vmatrix}=0 \text{ or } k^3+k^2-\alpha k+6=0.$$

By the Hurwitz criterion, $a_1>0$, $a_1a_2-a_3>0$, $a_3>0$ will be the conditions of asymptotic stability. In the given case, these conditions reduce to $-\alpha-6>0$, whence $\alpha<-6$.

6. The Case of a Small Coefficient of a Higher-Order Derivative

The theorem on the continuous dependence of a solution upon a parameter (see pages 58-59) asserts that the solution of the differential equation $\dot{x}(t)=f(t, x(t), \mu)$ is continuously dependent on the parameter μ if in the closed range of t, x and μ under consideration, the function f is continuous with respect to the collection of arguments and satisfies the Lipschitz condition with respect to x:

$$|f(t, \bar{x}, \mu)-f(t \ x, \mu)|\leqslant N|\bar{x}-x|,$$

where N does not depend on t, x and μ.

The conditions of this theorem are ordinarily fulfilled in problems of physics and mechanics, but one case of the discontinuous dependence of the right side on a parameter is comparatively often encountered in applications. This section is devoted to a study of the given case.

Consider the equation

$$\mu \frac{dx}{dt} = f(t, x), \tag{4.28}$$

where μ is a small parameter. The problem is to find out whether for small values of $|\mu|$ it is possible to ignore the term $\mu \frac{dx}{dt}$, i.e. whether it is possible to replace approximately a solution of the equation $\mu \frac{dx}{dt} = f(t, x)$ by a solution of the so-called degenerate equation

$$f(t, x) = 0. \tag{4.29}$$

We cannot take advantage here of the theorem on the continuous dependence of a solution on a parameter, since the right side of the equation

$$\frac{dx}{dt} = \frac{1}{\mu} f(t, x) \tag{4.28$_1$}$$

is discontinuous when $\mu = 0$.

For the time being, let us assume for the sake of simplicity that the degenerate equation (4.29) has only one solution $x = \varphi(t)$; also assume for definiteness that $\mu > 0$. As the parameter μ tends to zero, the derivative $\frac{dx}{dt}$ of the solutions of the equation $\frac{dx}{dt} = \frac{1}{\mu} f(t, x)$ will, at every point at which $f(t, x) \neq 0$, increase without bound in absolute value, having a sign that coincides with the sign of the function $f(t, x)$. Consequently, at all points at which $f(t, x) \neq 0$, tangents to the integral curves tend to a direction parallel to the x-axis as $\mu \to 0$; and if $f(t, x) > 0$, then the solution $x(t, \mu)$ of the equation (4.28$_1$) increases with increasing t, since $\frac{dx}{dt} > 0$, and if $f(t, x) < 0$, then the solution $x(t, \mu)$ diminishes with increasing t, since $\frac{dx}{dt} < 0$.

Let us consider the case (a) shown in Fig. 4-18 in which the sign of the function $f(t, x)$ changes (with x increasing and t fixed) from + to — when crossing the graph of the solution $x = \varphi(t)$ of the degenerate equation.

The arrows indicate the direction-field of tangents to the integral curves for sufficiently small μ The direction-field is in the direction of the graph of the root of the degenerate equation. Therefore, no matter what the initial values $x(t_0)=x_0$, the integral curve defined by these initial values, while almost parallel to the x-axis, tends to the graph of the root of the degenerate equation, and as t increases can no longer leave the neighbourhood of this graph. Hence, in this case, given $t \geqslant t_1 > t_0$ and a sufficiently small μ, we can, approximately, replace the solution $x(t, \mu)$ of the equation (4.28) by the solution of the degenerate equation. In the case at hand, the solution $x=\varphi(t)$ of the degenerate equation is called stable.

Consider case (b): the sign of the function $f(t, x)$ changes from $-$ to $+$ when passing across the graph of the solution $x=\varphi(t)$ of the degenerate equation for increasing x and fixed t. Fig. 4-19 shows

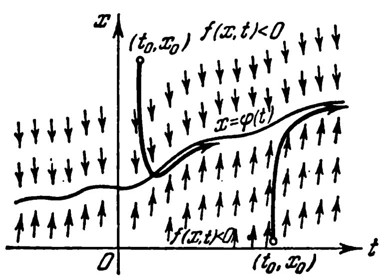


Fig. 4-18

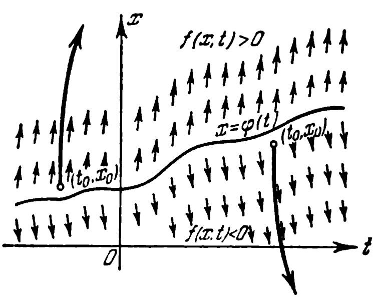


Fig. 4-19

the field of directions tangent to the integral curves for a sufficiently small μ. It is obvious in this case that no matter what the initial values $x(t_0)=x_0$ which satisfy the condition $f(t_0, x_0) \neq 0$ alone, the integral curve defined by these values (for a sufficiently small μ) having a tangent almost parallel to the x-axis recedes from the graph of the solution $x=\varphi(t)$ of the degenerate equation. In this case, the solution $x=\varphi(t)$ of equation (4.29) is called unstable. In the unstable case, one cannot replace the solution $x=x(t, \mu)$ of the original equation by the solution of the degenerate equation. In other words, one cannot ignore the term $\mu \frac{dx}{dt}$ in the equation $\mu \frac{dx}{dt}=f(t, x)$, no matter how small μ is.

A third, so-called half-stable, case is possible: case (c). The sign of the function $f(t, x)$ does not change when passing across the graph of the solution of the degenerate equation. Fig. 4.20 shows the direction-field in the case of a half-stable solution $x=\varphi(t)$.

As a rule, in the half-stable case it is also impossible to approximately replace the solution of the original equation, $x=x(t, \mu)$ by the solution of the degenerate equation, since, firstly, the integral curves defined by the initial values lying on one side of the graph of the solution $x=\varphi(t)$ recede from this graph, secondly, the integral curves approaching the graph of the solution $x=\varphi(t)$ can cross it to the unstable side (Fig. 4.20) and then recede from the graph of the solution $x=\varphi(t)$. Finally, even if the integral curve $x=x(t, \mu)$ remains in the neighbourhood of the graph of the solution on the stable side, the inevitable perturbations that occur in practical problems can throw the graph of the solution $x=x(t, \mu)$ to the unstable side of the graph of the solution of the degenerate equation, after which the integral curve $x=x(t, \mu)$ will recede from the graph of solution $x=\varphi(t)$.

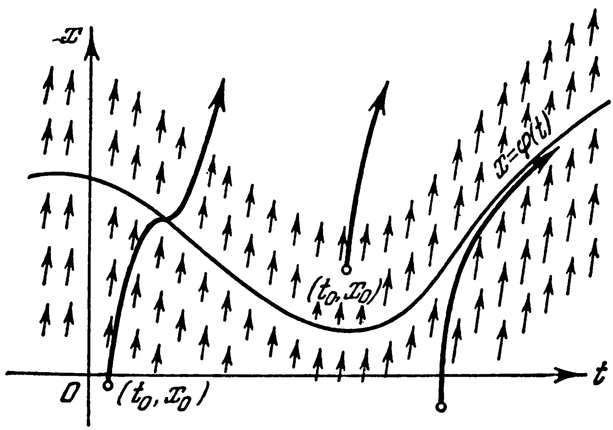


Fig. 4-20

Note that if on the graph of the solution of the degenerate equation $\frac{\partial f}{\partial x}<0$, then the solution $x=\varphi(t)$ is definitely stable; but if $\frac{\partial f}{\partial x}>0$, then the solution $x=\varphi(t)$ is unstable, since in the first case the function f decreases with increasing x in the neighbourhood of the curve $x=\varphi(t)$ and, hence, changes sign from + to —, whereas in the second case it increases with increasing x and, hence, the function f changes sign from — to + when crossing the graph of the solution $x=\varphi(t)$.

If the degenerate equation has several solutions $x=\varphi_i(t)$, then each one of them has to be investigated for stability; depending on the choice of the initial values, the integral curves of the initial equation may behave differently as $\mu \to 0$. For example, in the case depicted in Fig. 4.21 of three solutions $x=\varphi_i(t)$ $(i=1, 2, 3)$ of the degenerate equation, the graphs of which do not intersect, the solutions $x=x(t, \mu)$, $\mu>0$, of the original equation defined by the initial points lying above the graph of the function $x=\varphi_2(t)$ tend to a stable solution of the degenerate equation $x=\varphi_2(t)$ as $\mu\to 0$ for $t>t_0$, while the solutions $x=x(t, \mu)$ defined by the initial points lying below the graph of the function $x=\varphi_2(t)$ tend to the stable solution $x=\varphi_3(t)$ of the degenerate equation as $\mu\to 0$ for $t>t_0$ (Fig. 4.21).

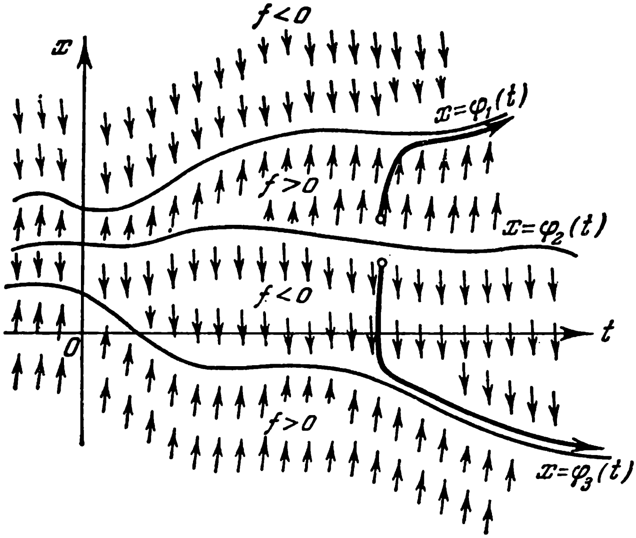


Fig. 4-21

Example 1. Find out whether the solution $x=x(t, \mu)$ of the equation $\mu\frac{dx}{dt}=x-t$, $\mu>0$ satisfying the initial conditions $x(t_0)=x_0$ tends to the solution of the degenerate equation $x-t=0$ as $\mu\to 0$ for $t>t_0$.

The solution $x=x(t, \mu)$ does not tend to the solution of the degenerate equation $x=t$, since the solution of the degenerate equation is unstable because $\frac{\partial(x-t)}{\partial x}=1>0$ (Fig. 4.22).

Example 2. The same with respect to the equation

$$\mu\frac{dx}{dt}=\sin^2 t-3e^x.$$

The solution of the degenerate equation $x=2\ln|\sin t|-\ln 3$ is stable since $\frac{\partial(\sin^2 t-3e^x)}{\partial x}=-3e^x<0$. Hence, the solution of the original equation $x=x(t, \mu)$ tends to the solution of the degenerate equation as $\mu\to 0$ for $t>t_0$.

Example 3. The same with respect to the solution of the equation

$$\mu \frac{dx}{dt} = x(t^2 - x + 1), \quad \mu > 0, \quad x(t_0) = x_0.$$

Of the two solutions $x=0$ and $x=t^2+1$ of the degenerate equation $x(t^2-x+1)=0$, the first is unstable, since $\left.\frac{\partial x(t^2-x+1)}{\partial x}\right|_{x=0} = t^2+1>0$ and the second is stable, since $\left.\frac{\partial x(t^2-x+1)}{\partial x}\right|_{x=t^2+1} = -t^2-1<0$.

If the initial point (t_0, x_0) lies in the upper half-plane $x>0$, then the integral curve of the original equation as $\mu \to 0$ approaches the graph of the solution $x=t^2+1$ of the degenerate equation (Fig. 4.23) and remains in its neighbourhood.

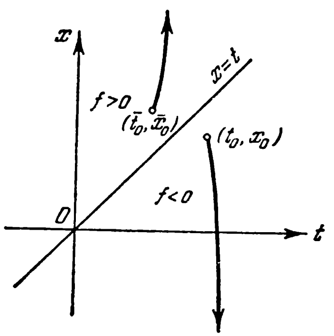


Fig. 4-22

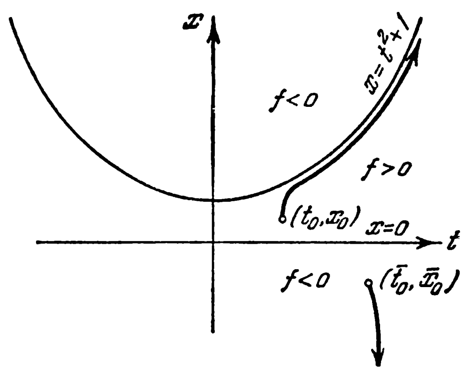


Fig. 4-23

But if the initial point lies in the lower half-plane, $x<0$, the $\lim_{\mu\to 0} x(t, \mu) = -\infty$ for $t>t_0$ (Fig. 4.23).

The problem of the dependence of a solution on a small coefficient μ of the highest derivative also arises in regard to nth-order equations

$$\mu x^{(n)}(t) = f(t, x, \dot{x}, \ddot{x}, \ldots, x^{(n-1)}),$$

and systems of differential equations.

An equation of the nth order may, in the usual manner (see page 91), be reduced to a system of first-order equations, and hence the principal problem is to investigate the system of first-order equations with one or several small coefficients of the derivatives. This problem has been studied in detail by A. Tikhonov [4] and A. Vasilieva.

7. Stability Under Constantly Operating Perturbations

If the system of equations being investigated

$$\frac{dx_i}{dt}=f_i(t,\ x_1,\ x_2,\ \dots,\ x_n),\qquad x_i(t_0)=x_{i0} \qquad (4.30)$$
$$(i=1,\ 2,\ \dots,\ n)$$

is subjected to small brief perturbations, then, for a small range of t, $\bar{t}_0\leqslant t\leqslant\bar{\bar{t}}_0$, the system (4.30) should be replaced by a perturbed system:

$$\left.\begin{aligned}\frac{dx_i}{dt}=f_i(t,\ x_1,\ x_2,\ \dots,\ x_n)+R_i(t,\ x_1,\ x_2,\ \dots,\ x_n),\\ x_i(\bar{t}_0)=\tilde{x}_i(\bar{t}_0)\qquad (i=1,\ 2,\ \dots,\ n),\end{aligned}\right\}\quad (4.31)$$

where all the $R_i(t,\ x_1,\ x_2,\ \dots,\ x_n)$ are small in absolute value; when $t\geqslant\bar{\bar{t}}_0$ the perturbations cease and we again revert to (4.30), but with somewhat altered initial values at the point $\bar{\bar{t}}_0$, $x_i(\bar{\bar{t}}_0)=$ $=\tilde{x}_i(\bar{\bar{t}}_0)+\delta_i\ (i=1,\ 2,\ \dots,\ n)$, where $\tilde{x}_i(t)$ $(i=1,\ 2,\ \dots,\ n)$ is the solution under investigation of the system (4.30), and all the δ_i are small in absolute value for small $|R_i|$ by virtue of the theorem on the continuous dependence of a solution upon a parameter (Fig. 4.24).

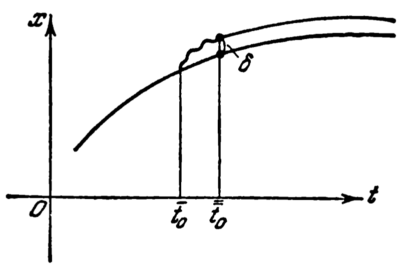


Fig. 4-24

Consequently, the operation of short-time perturbations ultimately reduces to perturbations of the initial values, and the problem of stability with respect to such short-time or, as they are often called, instantaneous perturbations reduces to the above-considered problem of stability in the sense of Lyapunov.

But if perturbations operate constantly, then the system (4.30) must be replaced by the system (4.31) for all $t\geqslant t_0$ and a completely new problem of stability arises under constantly acting perturbations. This problem has been investigated by I. Malkin and G. Duboshin.

As in the investigation of stability in the sense of Lyapunov, it is possible, by a change of variables $x_i=y_i-\varphi_i(t)$ $(i=1,\ 2,\ \dots,\ n)$, to transform the solution under investigation $y_i=\varphi_i(t)$ $(i=1, 2,\ \dots,\ n)$ of the system $\frac{dy_i}{dt}=\Phi_i(t,\ y_1,\ y_2,\ \dots,\ y_n)$ $(i=1,\ 2,\ \dots,\ n)$ into the trivial solution $x_i\equiv 0$ $(i=1,\ 2,\ \dots,\ n)$ of the transformed system. Therefore, from now on it may be assumed that, given constantly operating perturbations, we test for stability the trivial solution $x_i \equiv 0$ $(i = 1, 2, \ldots, n)$ of the system of equations (4.30).

The trivial solution of the system (4.30) is called *stable* with respect to constantly operating perturbations if for every $\varepsilon > 0$ it is possible to choose $\delta_1 > 0$ and $\delta_2 > 0$ such that from the inequalities $\sum_{i=1}^{n} R_i^2 < \delta_1^2$ for $t \geqslant t_0$ and $\sum_{i=1}^{n} x_{i0}^2 < \delta_2^2$ it follows that

$$\sum_{i=1}^{n} x_i^2(t) < \varepsilon^2 \quad \text{for} \quad t \geqslant t_0,$$

where $x_i(t)$ $(i = 1, 2, \ldots, n)$ is the solution of the system (4.31) defined by the initial conditions $x_i(t_0) = x_{i0}$ $(i = 1, 2, \ldots, n)$.

Theorem 4.7 (Malkin's theorem). *If for the system of equations* (4.30) *there exists a differentiable Lyapunov function* $v(t, x_1, x_2, \ldots x_n)$ *which satisfies the following conditions in the neighbourhood of the coordinate origin for* $t \geqslant t_0$:

(1) $v(t, x_1, x_2, \ldots, x_n) \geqslant w_1(x_1, x_2, \ldots, x_n) \geqslant 0$, $v(t, 0, 0, \ldots, 0) = 0$, *where* w_1 *is a continuous function that only vanishes at the origin;*

(2) *the derivatives* $\frac{\partial v}{\partial x_s}$ $(s = 1, 2, \ldots, n)$ *are bounded in absolute value;*

(3) *the derivative* $\frac{dv}{dt} = \frac{\partial v}{\partial t} + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i \leqslant -w_2(x_1, x_2, \ldots, x_n) \leqslant 0$,

where the continuous function $w_2(x_1, x_2, \ldots, x_n)$ *can vanish only at the origin, then the trivial solution of the system* (4.30) *is stable with respect to constantly operating perturbations.*

Proof. Observe that by virtue of the boundedness of the derivatives $\frac{\partial v}{\partial x_s}$ $(s = 1, 2, \ldots, n)$ the function v tends to zero uniformly in t for $t \geqslant t_0$ as $\sum_{i=1}^{n} x_i^2 \to 0$, since by the mean-value theorem $v(t, x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \left(\frac{\partial v}{\partial x_i}\right) x_i$, where $\left(\frac{\partial v}{\partial x_i}\right)$ are derivatives computed for certain intermediate [between 0 and x_i $(i = 1, 2, \ldots, n)$] values of the arguments $x_1, x_2, \ldots, x_n$.

Note also that outside a certain δ-neighbourhood of the origin, (i. e., for $\sum_{i=1}^{n} x_i^2 > \delta^2 > 0$) and for $t \geqslant t_0$, by virtue of conditions (2) and (3), the derivative

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} R_i \leqslant -k < 0$$

for sufficiently small absolute values of R_i ($i = 1, 2, \ldots, n$).

Let us specify $\varepsilon > 0$ and choose some level surface (or one of its components) $w_1 = l$, $l > 0$, lying wholly in the ε-neighbourhood of the coordinate origin.

By virtue of condition (1), the moving (given a variable $t \geqslant t_0$) level surface $v(t, x_1, x_2, \ldots, x_n) = l$ lies inside the level surface $w_1 = l$, whereas, by virtue of the fact that the function v tends to zero uniformly in t as $\sum_{i=1}^{n} x_i^2 \to 0$, it lies outside a certain δ_2-neighbourhood of the origin in which $v < l$ and, hence, on the level surface $v(t, x_1, x_2, \ldots, x_n) = l$, given any $t \geqslant t_0$, the derivative

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} R_i \leqslant -k < 0,$$

if $\sum_{i=1}^{n} R_i^2 < \delta_1$, $\delta_1 > 0$, where δ_1 is sufficiently small. The trajectory defined by the initial point $x_i(t_0) = x_{i0}$ ($i = 1, 2, \ldots, n$) lying in the above-indicated δ_2-neighbourhood of the origin cannot, given $t \geqslant t_0$, go beyond the ε-neighbourhood of the origin, since, by virtue of the choice of δ_2, $v(t_0, x_{10}, x_{20}, \ldots, x_{n0}) < l$ and, hence, if for $t \geqslant t_0$ the trajectory went beyond the ε-neighbourhood of the origin or at least beyond the level surface $w_1 = l$, then it should, at some value $t = T$, cross the level surface $v(t, x_1, x_2, \ldots, x_n) = l$ for the first time; and the function v should, in the neighbourhood of the point of intersection, increase along the trajectory but this contradicts the condition $\frac{dv}{dt} \leqslant -k < 0$ along the trajectory at points of the level surface $v(t, x_1, x_2, \ldots, x_n) = l$.

Comparing the conditions of the Malkin theorem with those of the Lyapunov theorem on asymptotic stability (see note on page 226), we see that they nearly coincide; the additional feature of Malkin's theorem is the demand of boundedness of the derivatives $\frac{\partial v}{\partial x_s}$ ($s = 1, 2, \ldots, n$) so that asymptotic stability and stability with respect to constant perturbations are extremely close properties, though they do not coincide.

Example 1. Is the trivial solution $x=0$, $y=0$ of the system of equations

$$\frac{dx}{dt}=a^2y-x^3,$$
$$\frac{dy}{dt}=-b^2x-y^3,$$

where a and b are constants, stable with respect to constantly operating perturbations?

The Lyapunov function that satisfies all the conditions of the Malkin theorem is $v=b^2x^2+a^2y^2$.

Thus the rest point $x=0$, $y=0$ is stable with respect to constantly acting perturbations.

Example 2. Do we have a stable rest point $x_i \equiv 0$ $(i=1, 2, \ldots, n)$ of the system

$$\frac{dx_i}{dt}=\sum_{j=1}^{n} a_{ij}x_j+R_i(t, x_1, x_2, \ldots, x_n) \qquad (i=1, 2, \ldots, n) \quad (4.32)$$

with respect to constantly operating perturbations if all the a_{ij} are constant and the R_i satisfy the conditions of the Lyapunov theorem, page 230, that is $|R_i| \leqslant N\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}+\alpha}$, $\alpha>0$, N is constant, and all the roots of the characteristic equation of the first-approximation system are distinct and negative?

On page 231, after a change of variables that reduced the linear parts of equation (4.32) to the canonical form, a Lyapunov function $v=\sum_{i=1}^{n} y_i^2$ was indicated that satisfied all the conditions of Malkin's theorem; hence, the rest point $x_i=0$ $(i=1, 2, \ldots, n)$ is stable with respect to constantly operating perturbations.

The same result may be obtained on the assumption that the real parts of all the roots of the characteristic equation (multiple ones may also occur among them) are negative; only in this case the choice of the Lyapunov function is substantially more involved.

PROBLEMS ON CHAPTER 4

1. Test for stability the rest point $x=0$, $y=0$ of the system

$$\frac{dx}{dt}=-2x-3y+x^5,$$

$$\frac{dy}{dt}=x+y-y^2.$$

2. Test for stability the rest point $x=0$, $y=0$, $z=0$ of the system

$$\frac{dx}{dt}=x-y-z, \quad \frac{dy}{dt}=x+y-3z \quad \frac{dz}{dt}=x-5y-3z.$$

3. For what values of α is the rest point $x=0$, $y=0$, $z=0$ of the system $\frac{dx}{dt}=\alpha x-y$, $\frac{dy}{dt}=\alpha y-z$, $\frac{dz}{dt}=\alpha z-x$ stable?

4. For what values of α does the system

$$\frac{dx}{dt}=y+\alpha x-x^5,$$
$$\frac{dy}{dt}=-x-y^5$$

have a stable rest point $x=0$, $y=0$?

5. To what limit does the solution of the differential equation

$$\mu\frac{dx}{dt}=(x^2+t^2-4)(x^2+t^2-9), \quad x(1)=1$$

tend as $\mu\to 0$, for $\mu>0$, $t>1$?

6. To what limit does the solution of the differential equation $\mu\frac{dx}{dt}=x-t+5$, $x(2)=5$ tend as $\mu\to 0$, for $\mu>0$, $t>2$?

7. Test for stability the rest point $x=0$, $y=0$ of the system of equations

$$\frac{dx}{dt}=x+e^y-\cos y,$$
$$\frac{dy}{dt}=x-y-\sin y.$$

8. Is the solution $x=0$, $y=0$ of the following system of equations stable with respect to constantly operating perturbations:

$$\frac{dx}{dt}=-2y-x^5,$$
$$\frac{dy}{dt}=5x-y^5?$$

9. Is the solution $x\equiv 0$ of the equation

$$\dddot{x}+5\ddot{x}+2x+20=0$$

stable?

10. Is the solution $x\equiv 0$ of the equation

$$\dddot{x}+5\ddot{x}+6x+x=0$$

stable?

11. What type of rest point $x=0$, $y=0$ does the system of equations

$$\frac{dx}{dt}=x+3y, \qquad \frac{dy}{dt}=5x-y$$

have?

12. Determine the periodic solution of the equation $\ddot{x}+2\dot{x}+2x=\sin t$ and test it for stability.

13. $\dddot{x}+2\ddot{x}+5\dot{x}+3x=\cos t$. Is the periodic solution of this equation stable?

14. Test for stability the rest point $x\equiv 0$, $y\equiv 0$ of the system

$$\dot{x}=y^3+x^5, \quad \dot{y}=x^3+y^5.$$

15. Test for stability the solutions of the system of equations

$$\dot{x}=3y-2x+e^t,$$
$$\dot{y}=5x-4y+2.$$

16. Test for stability the trivial solution of the equation

$$\dddot{x}+2\ddot{x}+3\dot{x}+7\sinh x=0.$$

17. Test for stability the trivial solution of the equation

$$\ddot{x}+(\alpha-1)\dot{x}+(4-\alpha^2)x=0,$$

where α is a parameter.

18. Is the solution $x\equiv 0$, $y\equiv 0$ of the system

$$\dot{x}=3y-x^3, \quad \dot{y}=-4x-3y^5$$

stable for constantly operating perturbations?

19. Is the trivial solution stable of the system

$$\dot{X}(t)=AX(t),$$

where $X(t)$ is a vector in three-dimensional space, and

$$A=\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 1\\ 1 & 3 & 1\end{pmatrix}?$$

20. Test the solutions of the equation

$$\dddot{x}+4\ddot{x}+5\dot{x}=t$$

for stability.

21. Test the solutions of the equation

$$\ddot{x} + 9x = \sin t$$

for stability.

22. $\dddot{x} + x = \cos t$. Find the periodic solution and test it for stability.

23. Find the region of stability of

$$\ddot{x} + \alpha\dot{x} + (1 - \alpha)x = 0.$$

24. $\dddot{x} + \ddot{x} + \alpha^2\dot{x} + 5\alpha x = 0$. Find the region of stability.

CHAPTER 5

First-order partial differential equations

1. Fundamentals

As was pointed out in the introduction (page 13), *partial differential equations* are differential equations in which the unknown functions are functions of more than one independent variable.

Very many physical phenomena are described by partial differential equations. The equation

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2=n(x,\ y,\ z)$$

describes the propagation of light rays in an inhomogeneous medium with refractive index $n(x,\ y,\ z)$; the equation

$$\frac{\partial u}{\partial t}=a^2\frac{\partial^2 u}{\partial x^2}$$

describes the temperature variation of a rod; the equation

$$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}$$

is the equation of the vibration of a string; the Laplace equation

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0$$

is satisfied by a field potential in regions devoid of charges, and so forth.

In this chapter we will deal briefly only with methods of integrating first-order partial differential equations, the theory of which is closely associated with the integration of certain systems of ordinary equations.

Higher-order partial differential equations, which are integrated by quite different methods, are taken up in another book of this series.

We consider a few elementary examples.

Example 1.

$$\frac{\partial z(x,\ y)}{\partial x}=y+x.$$

Integrating with respect to x, we get

$$z(x, y) = xy + \frac{x^2}{2} + \varphi(y),$$

where $\varphi(y)$ is an arbitrary function of y.

Example 2.

$$\frac{\partial^2 z(x, y)}{\partial x\,\partial y} = 0 \quad \text{or} \quad \frac{\partial}{\partial x}\left\{\frac{\partial z}{\partial y}\right\} = 0.$$

Integrating with respect to x, we get $\frac{\partial z}{\partial y} = \varphi(y)$, where $\varphi(y)$ is an arbitrary function of y. Then integrating with respect to y, we have

$$z = \int \varphi(y)\,dy + \varphi_1(x),$$

where $\varphi_1(x)$ is an arbitrary function of x. Or, writing

$$\int \varphi(y)\,dy = \varphi_2(y),$$

we finally get

$$z(x, y) = \varphi_1(x) + \varphi_2(y),$$

where $\varphi_2(y)$, by virtue of the arbitrariness of the function $\varphi(y)$, is likewise an arbitrary differentiable function of y.

The foregoing examples suggest that the general solution of a partial differential equation of the first order depends on one arbitrary function; the general solution of a second-order equation depends on two arbitrary functions, and the general solution of a pth-order equation most likely depends on p arbitrary functions.

These assumptions are true, but they require a more precise statement. To refine them, we formulate a theorem of S. Kovalevskaya on the existence and uniqueness of the solution of a partial differential equation.

Theorem 5.1 (Kovalevskaya's theorem). *There is a unique analytic (in the neighbourhood of the point $x_{10}, x_{20}, \ldots, x_{n0}$) solution of an equation solved for one of the derivatives of maximum order*

$$\frac{\partial^p z}{\partial x_1^p} = f\left(x_1, x_2, \ldots, x_n, z, \frac{\partial z}{\partial x_1}, \frac{\partial^2 z}{\partial x_1^2}, \ldots, \frac{\partial^{p-1} z}{\partial x_1^{p-1}}, \frac{\partial z}{\partial x_2}, \frac{\partial^2 z}{\partial x_1 \partial x_2}, \frac{\partial^2 z}{\partial x_2^2}, \ldots, \frac{\partial^p z}{\partial x_n^p}\right) \qquad \text{(A)}$$

that satisfies the conditions, for $x = x_{10}$,

$$z = \varphi_0(x_2, x_3, \ldots, x_n), \ \frac{\partial z}{\partial x_1} = \varphi_1(x_2, x_3, \ldots, x_n), \ \ldots, \ \frac{\partial^{p-1} z}{\partial x_1^{p-1}} = \varphi_{p-1}(x_2, x_3, \ldots, x_n),$$

if the functions $\varphi_0, \varphi_1, \ldots, \varphi_{p-1}$ *are analytic functions in the neighbourhood of the initial point* $x_{20}, x_{30}, \ldots, x_{n0}$, *and* f *is an analytic function in the neighbourhood of the initial values of its arguments* $x_{10}, x_{20}, \ldots, x_{n0}, z_0 = \varphi_0(x_{20}, x_{30}, \ldots, x_{n0})$,

$$\left(\frac{\partial z}{\partial x_1}\right)_0 = \varphi_1(x_{20}, \ldots, x_{n0}), \quad \ldots, \quad \left(\frac{\partial^p z}{\partial x_n^p}\right)_0 = \left(\frac{\partial^p \varphi_0}{\partial x_n^p}\right)_{x_i = x_{i0}}.$$

The solution is given by specifying the initial functions $\varphi_0, \varphi_1, \ldots, \varphi_{p-1}$; by arbitrarily varying them in the class of analytic functions, we get a collection of analytic solutions of the original equation (A) that depends on p arbitrary functions.

We do not give the proof of this theorem which requires invoking the theory of analytic functions.

2. Linear and Quasilinear First-Order Partial Differential Equations

A *nonhomogeneous linear equation* or a *quasilinear first-order partial differential equation* is an equation of the type

$$X_1(x_1, x_2, \ldots, x_n, z)\frac{\partial z}{\partial x_1} + X_2(x_1, x_2, \ldots, x_n, z)\frac{\partial z}{\partial x_2} + \ldots$$
$$\ldots + X_n(x_1, x_2, \ldots, x_n, z)\frac{\partial z}{\partial x_n} = Z(x_1, x_2, \ldots, x_n, z). \tag{5.1}$$

This equation is linear in the derivatives but can be nonlinear in the unknown function z.

If the right side is identically zero and the coefficients X_i are not dependent on z, then the equation (5.1) is called *homogeneous linear*.

To make the geometrical interpretation more pictorial, we first consider a quasilinear equation in two independent variables:

$$P(x, y, z)\frac{\partial z}{\partial x} + Q(x, y, z)\frac{\partial z}{\partial y} = R(x, y, z). \tag{5.1$_1$}$$

The functions P, Q and R will be considered continuous in the range under consideration of the variables and not vanishing simultaneously.

Consider the continuous vector field

$$\mathbf{F} = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k},$$

where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are unit vectors directed along the coordinate axes.

The vector lines of this field (i.e., lines, the tangent to which at each point is in a direction coinciding with the direction of the vector $\mathbf{F}$ at that point) are determined from the condition of collinearity of the vector $\mathbf{t} = \mathbf{i}\,dx + \mathbf{j}\,dy + \mathbf{k}\,dz$ directed along the

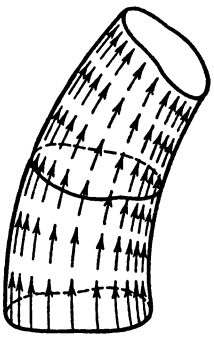


Fig. 5-1

tangent to the desired lines, and of the field vector **F**:

$$\frac{dx}{P(x, y, z)} = \frac{dy}{Q(x, y, z)} = \frac{dz}{R(x, y, z)}.$$

Surfaces made up of vector lines, or, to be more precise, surfaces completely containing vector lines which have at least one point in common with the surface, are called *vector surfaces* (Fig. 5.1).

It is obvious that vector surfaces may be obtained by considering the set of points lying on an arbitrarily chosen (continuously parameter-dependent) one-parameter family of vector lines. A vector surface is characterized by the fact that a vector N directed normally to the surface is, at any point of the surface, orthogonal to the field vector **F**:

$$(\mathbf{N}\cdot\mathbf{F}) = 0. \qquad (5.2)$$

If a vector surface is given by the equation $z = f(x, y)$, then the vector

$$\mathbf{N} = \frac{\partial z}{\partial x}\mathbf{i} + \frac{\partial z}{\partial y}\mathbf{j} - \mathbf{k}$$

and condition (5.2) takes the form

$$P(x, y, z)\frac{\partial z}{\partial x} + Q(x, y, z)\frac{\partial z}{\partial y} = R(x, y, z). \qquad (5.3)$$

If a vector surface is given by the equation $u(x, y, z) = 0$ and hence the vector $\mathbf{N} = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}$, then the equation (5.2) assumes the form

$$P(x, y, z)\frac{\partial u}{\partial x} + Q(x, y, z)\frac{\partial u}{\partial y} + R(x, y, z)\frac{\partial u}{\partial z} = 0. \qquad (5.4)$$

Consequently, in order to find the vector surfaces, one has to integrate the quasilinear equation (5.3) or the homogeneous linear equation (5.4), depending on whether one seeks the equation of the desired vector surfaces explicitly or implicitly.

Since vector surfaces may be made up of *vector lines*, integration of equations (5.3) [or (5.4)] reduces to integration of a system of ordinary differential equations of the vector lines.

Form a system of differential equations of the vector lines:

$$\frac{dx}{P(x, y, z)} = \frac{dy}{Q(x, y, z)} = \frac{dz}{R(x, y, z)}. \qquad (5.5)$$

Let $\psi_1(x, y, z)=c_1$ and $\psi_2(x, y, z)=c_2$ be two independent first integrals of the system (5.5). In arbitrary fashion we isolate from the two-parameter family of vector lines $\psi_1(x, y, z)=c_1$, $\psi_2(x, y, z)=c_2$, which are called *characteristics* of the equation (5.3) [or (5.4)], a one-parameter family, thus establishing some kind of continuous relation $\Phi(c_1, c_2)=0$ between the parameters c_1 and c_2. Eliminating the parameters c_1 and c_2 from the system

$$\psi_1(x, y, z)=c_1, \quad \psi_2(x, y, z)=c_2, \quad \Phi(c_1, c_2)=0,$$

we get the desired equation of vector surfaces:

$$\Phi(\psi_1(x, y, z), \psi_2(x, y, z))=0, \tag{5.6}$$

where Φ is an arbitrary function. We have thus found the integral of the quasilinear equation (5.3) that depends on an arbitrary function.

If it is required to find not an arbitrary vector surface of a field

$$\mathbf{F}=P(x, y, z)\,\mathbf{i}+Q(x, y, z)\,\mathbf{j}+R(x, y, z)\,\mathbf{k},$$

but a surface passing through a given line defined by the equations $\Phi_1(x, y, z)=0$ and $\Phi_2(x, y, z)=0$, then the function Φ in (5.6) will no longer be arbitrary but will be determined by eliminating the variables x, y, z from the system of equations

$$\Phi_1(x, y, z)=0, \quad \Phi_2(x, y, z)=0.$$
$$\psi_1(x, y, z)=c_1, \quad \psi_2(x, y, z)=c_2,$$

which must simultaneously be satisfied at the points of the given line $\Phi_1=0$ and $\Phi_2=0$, through which we draw the characteristics defined by the equations $\psi_1(x, y, z)=c_1$, $\psi_2(x, y, z)=c_2$.

Note that the problem becomes indeterminate if the given line $\Phi_1(x, y, z)=0$, $\Phi_2(x, y, z)=0$ is a characteristic, since in this case one may include the line in various one-parameter families of characteristics and thus obtain various integral surfaces passing through this line.

And so the integral of the quasilinear equation

$$P(x, y, z)\frac{\partial z}{\partial x}+Q(x, y, z)\frac{\partial z}{\partial y}=R(x, y, z),$$

which depends on an arbitrary function, may be obtained in the following manner: we integrate the auxiliary system of equations

$$\frac{dx}{P(x, y, z)}=\frac{dy}{Q(x, y, z)}=\frac{dz}{R(x, y, z)}$$

and, finding two independent first integrals of this system

$$\psi_1(x, y, z)=c_1, \qquad \psi_2(x, y, z)=c_2,$$

we obtain the desired integral in the form $\Phi(\psi_1(x, y, z), \psi_2(x, y, z))=0$, where Φ is an arbitrary function.

The equation of the integral surface of the same quasilinear equation which passes through the given line defined by the equations $\Phi_1(x, y, z)=0$ and $\Phi_2(x, y, z)=0$ may be found by taking the above-mentioned function Φ not arbitrarily, but by determining the function $\Phi(c_1, c_2)$ via elimination of x, y, z from the equations

$$\Phi_1(x, y, z)=0, \qquad \Phi_2(x, y, z)=0,$$
$$\psi_1(x, y, z)=c_1, \qquad \psi_2(x, y, z)=c_2,$$

as a result of which we obtain the equation $\Phi(c_1, c_2)=0$, and the desired integral will be $\Phi(\psi_1(x, y, z), \psi_2(x, y, z))=0$.

Example 1. Determine the integral, dependent on an arbitrary function, of the equation

$$\frac{\partial z}{\partial x}+\frac{\partial z}{\partial y}=1.$$

Form an auxiliary system of equations:

$$dx=dy=dz.$$

Its first integrals will be of the form $x-y=c_1$, $z-x=c_2$. The integral of the original equation $\Phi(x-y, z-x)=0$, where Φ is an arbitrary function, or as solved for z: $z=x+\varphi(x-y)$, where φ is an arbitrary differentiable function.

Example 2. Find the integral surface of the equation

$$x\frac{\partial z}{\partial y}-y\frac{\partial z}{\partial x}=0$$

passing through the curve $x=0$, $z=y^2$.

Integrate the system of equations

$$\frac{\partial x}{-y}=\frac{\partial y}{x}=\frac{dz}{0},$$

whence $z=c_1$, $x^2+y^2=c_2$. Eliminating x, y and z from the equations

$$x^2+y^2=c_2, \quad z=c_1, \quad x=0, \quad z=y^2,$$

we get $c_1=c_2$, whence $z=x^2+y^2$.

Example 3. Find the integral surface of the same equation

$$x\frac{\partial z}{\partial y}-y\frac{\partial z}{\partial x}=0$$

passing through the circle

$$z=1, \quad x^2+y^2=4. \tag{5.7}$$

Since the given line (5.7) is a vector line (characteristic), the problem is indeterminate. Indeed, the integral surfaces of the equations are all possible surfaces of rotation $z=\Phi(x^2+y^2)$, the axes of rotation of which coincide with the z-axis. There obviously exist an infinity of such surfaces of rotation passing through the circle (5.7); for example, paraboloids of revolution $z=x^2+y^2-3$, $4z=x^2+y^2$, $z=-x^2-y^2+5$, the sphere $x^2+y^2+z^2=5$, and so forth.

If the equation of the curve through which it is required to pass an integral surface of the equation (5.1_1) is given in parametric form:

$$x_0=x_0(s), \quad y_0=y_0(s), \quad z_0=z_0(s), \tag{B}$$

then ordinarily the solution is conveniently sought in the parametric form:

$$x=x(t, s), \quad y=y(t, s), \quad z=z(t, s).$$

We introduce a parameter t into the system (5.5) that defines the characteristics, assuming

$$\frac{dx}{P(x, y, z)}=\frac{dy}{Q(x, y, z)}=\frac{dz}{R(x, y, z)}=dt. \tag{5.5$_1$}$$

So that the characteristics should pass through the given curve, we seek the solution of the system (5.5_1) that satisfies, for $t=0$ (or $t=t_0$), the initial conditions

$$x=x_0(s), \quad y=y_0(s), \quad z=z_0(s).$$

For such initial conditions and given s fixed, we get a characteristic that passes through the fixed point of the curve (B). Given a variable s, we get a family of characteristics

$$x=x(t, s), \quad y=y(t, s), \quad z=z(t, s) \tag{C}$$

that pass through points of the given curve (B) (it is assumed here that the given curve (B) is not a characteristic). The set of points lying on this family of characteristics (C) is what forms the desired integral surface.

Example 4.

$$\frac{\partial z}{\partial x}-\frac{\partial z}{\partial y}=1.$$

Find the integral surface that passes through the curve $x_0=s$, $y_0=s^2$, $z_0=s^3$.

The system of equations that determines the characteristics is of the form

$$dx=-dy=dz=dt.$$

Its general solution is

$$x=t+c_1, \quad y=-t+c_2, \quad z=t+c_3.$$

Using the initial conditions, we determine the arbitrary constants and finally get

$$x = t + s, \quad y = -t + s^2, \quad z = t + s^3.$$

Let us now pass to the case of n independent variables. It is natural to expect that the above-indicated solution scheme for the three-dimensional case may be extended to the $(n+1)$-dimensional case as well.

Let us begin with a study of the homogeneous linear equation

$$X_1(x_1, x_2, \ldots, x_n)\frac{\partial z}{\partial x_1} + X_2(x_1, x_2, \ldots, x_n)\frac{\partial z}{\partial x_2} + \ldots \\ \ldots + X_n(x_1, x_2, \ldots, x_n)\frac{\partial z}{\partial x_n} = 0, \tag{5.8}$$

where the continuous functions $X_i(x_1, x_2, \ldots, x_n)$ do not vanish simultaneously at any point of the region under investigation and in that region have bounded partial derivatives.

We form an auxiliary system of equations

$$\frac{dx_1}{X_1(x_1, x_2, \ldots, x_n)} = \frac{dx_2}{X_2(x_1, x_2, \ldots, x_n)} = \ldots = \frac{dx_n}{X_n(x_1, x_2, \ldots, x_n)}, \tag{5.9}$$

which, given the restrictions indicated above, satisfies the conditions of the existence and uniqueness theorem.

We find the $n-1$ independent first integral of the system (5.9)

$$\psi_1(x_1, x_2, \ldots, x_n) = c_1,$$
$$\psi_2(x_1, x_2, \ldots, x_n) = c_2,$$
$$\ldots\ldots\ldots\ldots$$
$$\psi_{n-1}(x_1, x_2, \ldots, x_n) = c_{n-1}.$$

In a space with the coordinates $x_1, x_2, \ldots, x_n$, this system of integrals defines an $(n-1)$-parameter family of lines called the characteristics of the equation (5.8) We shall prove that the left side of any first integral $\psi(x_1, x_2, \ldots, x_n) = c$ of the system (5.9) is a solution of the original homogeneous linear partial differential equation (5.8).

Indeed, the function $\psi \equiv c$ along any integral curve of the system (5.9). Hence,

$$d\psi = \sum_{i=1}^{n} \frac{\partial \psi}{\partial x_i} dx_i \equiv 0. \tag{5.10}$$

along any integral curve But along the integral curve of the system (5.9), the differentials dx_i are proportional to the functions X_i;

hence, since the left member of the identity

$$\sum_{i=1}^{n} \frac{\partial \psi}{\partial x_i} dx_i \equiv 0$$

is homogeneous in dx_i, the differentials dx_i may be replaced by the quantities X_i, which are proportional to them; and we then find that along the integral curves of the system (5.9)

$$\sum_{i=1}^{n} \frac{\partial \psi}{\partial x_i} X_i \equiv 0. \tag{5.11}$$

The integral curves of the system (5.9) pass through every point of the given range of variables $x_1, x_2, \ldots, x_n$ and the left-hand member of the identity (5.11) is not dependent on the constants $c_1, c_2, \ldots, c_{n-1}$ and, consequently, does not vary from one integral curve to the next; hence the identity (5.11) holds true not only along a certain integral curve but throughout the whole considered range of the variables $x_1, x_2, \ldots, x_n$; and this means that the function ψ is a solution of the original equation

$$\sum_{i=1}^{n} X_i \frac{\partial z}{\partial x_i} = 0.$$

It is obvious that $\Phi(\psi_1, \psi_2, \ldots, \psi_{n-1}) = c$, where Φ is an arbitrary function, is a first integral of the system (5.9), since all the functions $\psi_1, \psi_2, \ldots, \psi_{n-1}$ are turned into constants along the integral curve of the system (5.9); consequently, $\Phi(\psi_1, \psi_2, \ldots, \psi_{n-1})$ too becomes a constant along the integral curve of the system (5.9). Hence, $z = \Phi(\psi_1, \psi_2, \ldots, \psi_{n-1})$, where Φ is an arbitrary differentiable function, is a solution of the homogeneous linear equation (5.8).

We will prove that

$$z = \Phi(\psi_1(x_1, \ldots, x_n), \quad \psi_2(x_1, \ldots, x_n), \quad \psi_{n-1}(x_1, \ldots, x_n))$$

is the general solution of the equation (5.8).

Theorem 5.2. *$z = \Phi(\psi_1, \psi_2, \ldots, \psi_{n-1})$, where Φ is an arbitrary function, is the general solution of the equation*

$$\sum_{i=1}^{n} X_i(x_1, x_2, \ldots, x_n) \frac{\partial z}{\partial x_i} = 0, \tag{5.8}$$

that is, a solution containing all the solutions of the equation without exception.

Proof. Assume that $z=\psi(x_1, x_2, \ldots, x_n)$ is some solution of the equation (5.8) and prove that there is a function Φ such that $\psi=\Phi(\psi_1, \psi_2, \ldots, \psi_{n-1})$.

Since ψ and $\psi_1, \psi_2, \ldots, \psi_{n-1}$ are solutions of the equation (5.8), it follows that

$$\left.\begin{array}{l} \sum_{i=1}^{n} X_i \frac{\partial\psi}{\partial x_i} \equiv 0, \\ \sum_{i=1}^{n} X_i \frac{\partial\psi_1}{\partial x_i} \equiv 0, \\ \sum_{i=1}^{n} X_i \frac{\partial\psi_2}{\partial x_i} \equiv 0, \\ \cdots\cdots\cdots \\ \sum_{i=1}^{n} X_i \frac{\partial\psi_{n-1}}{\partial x_i} \equiv 0. \end{array}\right\} \quad (5.12)$$

Regarding the system (5.12) as a homogeneous linear system of n equations in X_i $(i=1, 2, \ldots, n)$ and noting that this homogeneous system at every point $x_1, x_2, \ldots, x_n$ of the given interval has a nontrivial solution, since $X_i(x_1, x_2, \ldots, x_n)$, by hypothesis, do not vanish simultaneously, we come to the conclusion that the determinant of this system

$$\begin{vmatrix} \frac{\partial\psi}{\partial x_1} & \frac{\partial\psi}{\partial x_2} & \cdots & \frac{\partial\psi}{\partial x_n} \\ \frac{\partial\psi_1}{\partial x_1} & \frac{\partial\psi_1}{\partial x_2} & \cdots & \frac{\partial\psi_1}{\partial x_n} \\ \frac{\partial\psi_2}{\partial x_1} & \frac{\partial\psi_2}{\partial x_2} & \cdots & \frac{\partial\psi_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial\psi_{n-1}}{\partial x_1} & \frac{\partial\psi_{n-1}}{\partial x_2} & \cdots & \frac{\partial\psi_{n-1}}{\partial x_n} \end{vmatrix}$$

is identically zero in the interval under consideration. However, the fact that the Jacobian of the functions $\psi, \psi_1, \psi_2, \ldots, \psi_{n-1}$ vanishes identically shows that there is a functional relation between these functions:

$$F(\psi, \psi_1, \psi_2, \ldots, \psi_{n-1})=0. \quad (5.13)$$

By virtue of the independence of the first integrals $\psi_i(x_1, x_2, \ldots, x_n)=c_i$ $(i=1, 2, \ldots, n-1)$ of the system (5.9), at least one of the minors of the $(n-1)$th order of the Jacobian

$$\frac{D(\psi, \psi_1, \psi_2, \ldots, \psi_{n-1})}{D(x_1, x_2, x_3, \ldots, x_n)}$$

of the form

$$\frac{D(\psi_1, \psi_2, \ldots, \psi_n)}{D(x_{\alpha_1}, x_{\alpha_2}, \ldots, x_{\alpha n-1})}$$

is nonzero. Hence, equation (5.13) may be given in the form

$$\psi = \Phi(\psi_1, \psi_2, \ldots, \psi_{n-1}).$$

Example 5. Integrate the equation

$$\sum_{i=1}^{n} x_i \frac{\partial z}{\partial x_i} = 0. \tag{5.14}$$

The system of equations defining the characteristics is of the form

$$\frac{dx_1}{x_1} = \frac{dx_2}{x_2} = \ldots = \frac{dx_n}{x_n}.$$

The independent first integrals of this system are

$$\frac{x_1}{x_n} = c_1, \quad \frac{x_2}{x_n} = c_2, \quad \ldots, \quad \frac{x_{n-1}}{x_n} = c_{n-1}.$$

The general solution of the original equation

$$z = \Phi\left(\frac{x_1}{x_n}, \frac{x_2}{x_n}, \ldots, \frac{x_{n-1}}{x_n}\right)$$

is an arbitrary homogeneous function of degree zero.

Euler's theorem on homogeneous functions asserts that homogeneous functions of degree zero satisfy the given equation (5.14); we have now proved that only homogeneous functions of degree zero possess this property.

A nonhomogeneous linear equation of the first order

$$\sum_{i=1}^{n} X_i(x_1, x_2, \ldots, x_n, z) \frac{\partial z}{\partial x_i} = Z(x_1, x_2, \ldots, x_n, z), \tag{5.15}$$

where all the X_i and Z are continuously differentiable functions that do not vanish simultaneously in the given range of the variables $x_1, x_2, \ldots, x_n, z$, is integrated by reducing it to a homogeneous linear equation.

For this purpose, as in the case of three variables, it suffices to seek the solution z of the equation (5.15) in implicit form:

$$u(x_1, x_2, \ldots, x_n, z) = 0, \tag{5.16}$$

where $\frac{\partial u}{\partial z} \neq 0$.

Indeed, assuming that the function $z = z(x_1, x_2, \ldots, x_n)$ is determined from the equation (5.16), and differentiating the identity

$$u(x_1, x_2, \ldots, x_n, z(x_1, x_2, \ldots, x_n)) \equiv 0$$

with respect to x_i, we get

$$\frac{\partial u}{\partial x_i}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial x_i}=0,$$

whence

$$\frac{\partial z}{\partial x_i}=-\frac{\frac{\partial u}{\partial x_i}}{\frac{\partial u}{\partial z}}.$$

Substituting $\frac{\partial z}{\partial x_i}$ into (5.15), multiplying by $-\frac{\partial u}{\partial z}$ and transposing all terms to the left-hand side of the equation, we get the homogeneous linear equation

$$\sum_{i=1}^{n} X_i(x_1, x_2, \ldots, x_n, z)\frac{\partial u}{\partial x_i}+Z(x_1, x_2, \ldots, x_n, z)\frac{\partial u}{\partial z}=0, \quad (5.17)$$

which the function u must satisfy; however, this only on the assumption that z is a function of $x_1, x_2, \ldots, x_n$ defined by the equation $u(x_1, x_2, \ldots, x_n, z)=0$.

Thus, we have to find the functions u that reduce the homogeneous linear equation (5.17) to an identity by virtue of the equation

$$u(x_1, x_2, \ldots, x_n, z)=0.$$

First find the functions u that reduce the equation (5.17) to an identity, given independently varying $x_1, x_2, \ldots, x_n, z$. All such functions u are solutions of the homogeneous equation (5.17) and may be found by a method we already know: we form a system of equations that defines the characteristics

$$\frac{dx_1}{X_1(x_1, x_2, \ldots, x_n, z)}=\frac{dx_2}{X_2(x_1, x_2, \ldots, x_n, z)}=\ldots$$
$$\ldots=\frac{dx_n}{X_n(x_1, x_2, \ldots, x_n, z)}=\frac{dz}{Z(x_1, x_2, \ldots, x_n, z)}; \quad (5.18)$$

we find n independent first integrals of this system:

$$\psi_1(x_1, x_2, \ldots, x_n, z)=c_1,$$
$$\psi_2(x_1, x_2, \ldots, x_n, z)=c_2,$$
$$\cdots\cdots\cdots\cdots$$
$$\psi_n(x_1, x_2, \ldots, x_n, z)=c_n;$$

then the general solution of (5.17) is of the form

$$u=\Phi(\psi_1, \psi_2, \ldots, \psi_n),$$

where Φ is an arbitrary function.

The solution z of equation (5.15), which depends on an arbitrary function, is determined from the equation

$$u(x_1, x_2, \ldots, x_n, z) = 0 \quad \text{or} \quad \Phi(\psi_1, \psi_2, \ldots, \psi_n) = 0.$$

However, besides the solutions found by this method there may be solutions z which are determined from the equations $u(x_1, x_2, \ldots, x_n, z) = 0$, where the function u is not a solution of the equation (5.17), but reduces this equation to an identity only by virtue of the equation $u(x_1, x_2, \ldots, x_n, z) = 0$. Such solutions are called *special*.

In a certain sense, there are not very many special solutions; they cannot even form one-parameter families.

Indeed, if the special solutions formed a one-parameter family and were defined by the equation

$$u(x_1, x_2, \ldots, x_n, z) = c, \tag{5.19}$$

where c is a parameter, $c_0 \leqslant c \leqslant c_1$, then the equation (5.17) should reduce to an identity by virtue of the equation (5.19) for any c. But since (5.17) does not contain c, it cannot reduce to an identity by virtue of (5.19), which contains c, and, hence, must be an identity with respect to all the variables $x_1, x_2, \ldots, x_n, z$ which vary independently.

The last statement admits of a simple geometrical interpretation. When we say that the equation (5.17) reduces to an identity by virtue of the equation $u(x_1, x_2, \ldots, x_n, z) = 0$, we assert that (5.17) reduces to an identity at points of the surface $u = 0$, but cannot reduce to an identity at other points of the space $x_1, x_2, \ldots, x_n, z$ But if equation (5.17), which does not contain c, reduces to an identity by virtue of the equation $u = c$, where c is a continuously varying parameter, then this means that the equation (5.17) reduces to an identity on all surfaces $u = c$, $c_0 \leqslant c \leqslant c_1$ that do not intersect and that fill a certain part D of the space $x_1, x_2, \ldots, x_n, z$; and, hence, (5.17) reduces to an identity in the domain D for independently varying $x_1, x_2, \ldots, x_n, z$.

In concrete problems it is ordinarily required to find the solution of the equation (5.15) that satisfies some other initial conditions as well, and since there are comparatively few special solutions in the above-indicated meaning, only in quite exceptional cases will they satisfy the initial conditions; for this reason it is only in rare cases that they need to be allowed for.

Example 6. Integrate the equation

$$\sum_{i=1}^{n} x_i \frac{\partial z}{\partial x_i} = pz, \tag{5.20}$$

where p is a constant.

The system of equations

$$\frac{dx_1}{x_1}=\frac{dx_2}{x_2}=\ldots=\frac{dx_n}{x_n}=\frac{dz}{pz}$$

has the following independent integrals:

$$\frac{x_1}{x_n}=c_1,\ \frac{x_2}{x_n}=c_2,\ \ldots,\ \frac{x_{n-1}}{x_n}=c_{n-1},\ \frac{z}{x_n^p}=c_n.$$

Hence, solution z of the original equation is determined from the equation

$$\Phi\left(\frac{x_1}{x_n},\ \frac{x_2}{x_n},\ \ldots,\ \frac{x_{n-1}}{x_n},\ \frac{z}{x_n^p}\right)=0,$$

whence

$$z=x_n^p\,\psi\left(\frac{x_1}{x_n},\ \frac{x_2}{x_n},\ \ldots,\ \frac{x_{n-1}}{x_n}\right).$$

And so the solution is an arbitrary homogeneous function of degree p.

It may be proved that the equation (5.20) does not have special integrals and, hence, Euler's theorem on homogeneous functions is invertible, that is, only homogeneous functions of degree p satisfy the equation (5.20).

The concept of a characteristic can be extended to systems of quasilinear equations of the following special type:

$$\begin{aligned}P(x,\ y,\ u,\ v)\frac{\partial u}{\partial x}+Q(x,\ y,\ u,\ v)\frac{\partial u}{\partial y}&=R_1(x,\ y,\ u,\ v),\\ P(x,\ y,\ u,\ v)\frac{\partial v}{\partial x}+Q(x,\ y,\ u,\ v)\frac{\partial v}{\partial y}&=R_2(x,\ y,\ u,\ v).\end{aligned}\qquad\text{(D)}$$

The *characteristics* of this system are the vector lines of a vector field in the four-dimensional space

$\mathbf{F}=P(x,\ y,\ u,\ v)\,\mathbf{i}+Q(x,\ y,\ u,\ v)\,\mathbf{j}+R_1(x,\ y,\ u,\ v)\,\mathbf{k}_1+R_2(x,\ y,\ u,v)\,\mathbf{k}_2$, where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}_1$, $\mathbf{k}_2$ are unit vectors along the respective axes (x, y u and v).

The characterisctics are defined by the system of equations

$$\frac{dx}{P(x,\ y,\ u,\ v)}=\frac{dy}{Q(x,\ y,\ u,\ v)}=\frac{du}{R_1(x,\ y,\ u,\ v)}=\frac{dv}{R_2(x,\ y,\ u,\ v)}.\qquad\text{(E)}$$

The system of equations (D) is written as follows in vector notation:

$$(\mathbf{F}\cdot\mathbf{N}_1)=0\ \text{and}\ (\mathbf{F}\cdot\mathbf{N}_2)=0,$$

where $\mathbf{N}_1$ and $\mathbf{N}_2$ are vectors with the coordinates $\left(\frac{\partial u}{\partial x},\ \frac{\partial u}{\partial y},\ -1,0\right)$ and $\left(\frac{\partial v}{\partial x},\ \frac{\partial v}{\partial y},\ 0,\ -1\right)$ and directed normally to the desired three-dimensional cylindrical surfaces $u=u\ (x,\ y)$ and $v=v(x,\ y)$, respectively.

Hence, from the geometrical point of view, integration of the system (D) reduces to finding two three-dimensional cylindrical surfaces $u=u(x, y)$ and $v=v(x, y)$, the normals to which are orthogonal to the vector lines at the points of intersection of the surfaces.

It is obvious that this condition will be fulfilled if the two-dimensional surface S, on which, generally speaking, the three-dimensional cylindrical surfaces $u=u(x, y)$ and $v=v(x, y)$ intersect, consists of vector lines, since these vector lines will lie simultaneously on the surfaces $u=u(x, y)$ and $v=v(x, y)$ and, consequently, will be orthogonal to the vectors $\mathbf{N}_1$ and $\mathbf{N}_2$. If we take any two—independent of u and v—first integrals $\Phi_1(x, y, u, v)=0$ and $\Phi_2(x, y, u, v)=0$ of the system (E), in other words, if we take two three-dimensional vector surfaces, we will, generally speaking, get, at their intersection, the two-dimensional surface S, which consists of vector lines; for if a certain point belongs simultaneously to the vector surfaces $\Phi_1(x, y, u, v)=0$ and $\Phi_2(x, y, u, v)=0$, then the vector line passing through this point also lies in each of these surfaces.

Solving the system of equations $\Phi_1(x, y, u, v)=0$ and $\Phi_2(x, y, u, v)=0$ for u and v, we get the equations of two three-dimensional cylindrical surfaces $u=u(x, y)$ and $v=v(x, y)$ that intersect along the same two-dimensional surface S which consists of vector lines. Hence, the functions thus found, $u=u(x, y)$ and $v=v(x, y)$, will be solutions of the original system.

The solution of the system (D) which is dependent on two arbitrary functions may be found by applying the same method, but taking the first integrals of the system (E) in the most general form:

$$\left.\begin{aligned}\Phi_1(\psi_1(x, y, u, v), \psi_2(x, y, u, v), \psi_3(x, y, u, v))=0,\\ \Phi_2(\psi_1(x, y, u, v), \psi_2(x, y, u, v), \psi_3(x, y, u, v))=0,\end{aligned}\right\} \quad \text{(F)}$$

where $\psi_1(x, y, u, v)$, $\psi_2(x, y, u, v)$ and $\psi_3(x, y, u, v)$ are independent first integrals of the system (E), and Φ_1 and Φ_2 are arbitrary functions (see page 259).

The equations (F) define the solutions $u(x, y)$ and $v(x, y)$ of the system D as implicit functions of x and y, which are dependent on the choice of arbitrary functions Φ_1 and Φ_2, if the composite functions Φ_1 and Φ_2 are independent with respect to u and v.

3. Pfaffian Equations

In Sec. 2 we considered two problems that arose naturally in the study of the continuous vector field

$$\mathbf{F}=P(x, y, z)\mathbf{i}+Q(x, y, z)\mathbf{j}+R(x, y, z)\mathbf{k}.$$

These were the problems of finding vector lines and vector surfaces.

A problem that arises almost as frequently is that of finding the family of surfaces $U(x, y, z) = c$ that are orthogonal to the vector lines. The equation of such surfaces is of the form $(\mathbf{F}\cdot\mathbf{t}) = 0$, where $\mathbf{t}$ is a vector lying in the plane tangential to the desired surfaces:

$$\mathbf{t} = \mathbf{i}\,dx + \mathbf{j}\,dy + \mathbf{k}\,dz,$$

or, in expanded form,

$$P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz = 0. \tag{5.21}$$

Equations of the form (5.21) are called *Pfaffian equations.*

If the field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ is potential:

$$\mathbf{F} = \operatorname{grad} U, \quad \text{i.e.} \quad P = \frac{\partial U}{\partial x}, \qquad Q = \frac{\partial U}{\partial y}, \qquad R = \frac{\partial U}{\partial z},$$

then the desired surfaces are level surfaces $U(x, y, z) = c$ of the potential function U. Here it is not difficult to find the desired surfaces, since

$$U = \int_{(x_0, y_0, z_0)}^{(x, y, z)} P\,dx + Q\,dy + R\,dz,$$

where the line integral is taken along any path between the chosen fixed point (x_0, y_0, z_0) and the point with variable coordinates (x, y, z), for example, along a polygonal curve consisting of straight-line segments parallel to the coordinate axes.

But if the field $\mathbf{F}$ is not potential, then in certain cases it is possible to choose a scalar factor $\mu(x, y, z)$ such that when it is multiplied by the vector $\mathbf{F}$ the field becomes potential.

If such a factor exists, then $\mu\mathbf{F} = \operatorname{grad} U$ or

$$\mu P = \frac{\partial U}{\partial x}, \quad \mu Q = \frac{\partial U}{\partial y}, \quad \mu R = \frac{\partial U}{\partial z}$$

and, consequently,

$$\frac{\partial(\mu P)}{\partial y} = \frac{\partial(\mu Q)}{\partial x}, \qquad \frac{\partial(\mu Q)}{\partial z} = \frac{\partial(\mu R)}{\partial y}, \quad \frac{\partial(\mu R)}{\partial x} = \frac{\partial(\mu P)}{\partial z},$$

or

$$\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} = \frac{1}{\mu}\left(Q\frac{\partial \mu}{\partial x} - P\frac{\partial \mu}{\partial y}\right),$$
$$\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y} = \frac{1}{\mu}\left(R\frac{\partial \mu}{\partial y} - Q\frac{\partial \mu}{\partial z}\right),$$
$$\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z} = \frac{1}{\mu}\left(P\frac{\partial \mu}{\partial z} - R\frac{\partial \mu}{\partial x}\right).$$

Multiplying the first of these identities by R, the second by P and the third by Q and adding all three identities termwise, we get the necessary condition for the existence of the integrating factor μ:

$$R\left(\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}\right)+P\left(\frac{\partial Q}{\partial z}-\frac{\partial R}{\partial y}\right)+Q\left(\frac{\partial R}{\partial x}-\frac{\partial P}{\partial z}\right)=0$$

or $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})=0$, where the vector rot $\mathbf{F}$—the field rotation—is defined by the equality

$$\operatorname{rot}\ \mathbf{F}=\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathbf{i}+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathbf{j}+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathbf{k}.$$

If this condition, known as the condition of *total integrability* of the equation (5.21), is not fulfilled, then there does not exist a family of surfaces $U(x, y, z)=c$ that are orthogonal to the vector lines of the field $\mathbf{F}(x, y, z)$.

Indeed, if such a family $U(x, y, z)=c$ existed, the left side of (5.21) could differ from

$$\frac{\partial U}{\partial x}dx+\frac{\partial U}{\partial y}dy+\frac{\partial U}{\partial z}dz$$

solely in a certain factor $\mu(x, y, z)$, which would then be an integrating factor for the equation (5.21).

Thus, for a family of surfaces $U(x, y, z)=c$ orthogonal to the vector lines of a vector field $\mathbf{F}$ to exist, it is necessary that the vectors $\mathbf{F}$ and rot $\mathbf{F}$ should be orthogonal, that is $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})\equiv 0$.

Note. The condition $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})=0$ is also called the condition for integrability of the Pfaffian equation $P\,dx+Q\,dy+R\,dz=0$ by one relation $U(x, y, z)=c$.

Sometimes it is not the surfaces orthogonal to the vector lines of the field $\mathbf{F}$ that have to be determined, but the lines that have the same property; in other words, it is necessary to integrate the Pfaffian equation by two relations, not one:

$$U_1(x, y, z)=0 \quad \text{and} \quad U_2(x, y, z)=0. \qquad (5.22)$$

To find such lines, one can arbitrarily specify one of the equations (5.22), for example,

$$U_1(x, y, z)=0, \qquad (5.23)$$

and, with the aid of (5.23), eliminating from (5.21) one of the variables, say z, we get a differential equation of the form

$$M(x, y)\,dx+N(x, y)\,dy=0;$$

integrating this equation we find the desired lines on an arbitrarily chosen surface $U_1(x, y, z)=0$.

We shall show that the condition $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})=0$ is not only necessary but also sufficient for the existence of a family of surfaces that are orthogonal to the vector lines.

Note that on the desired surfaces $U(x, y, z)=c$, the equation

$$P\,dx+Q\,dy+R\,dz=0$$

must reduce to an identity or, what is the same thing, on these surfaces the line integral

$$\int_L P\,dx+Q\,dy+R\,dz \tag{5.24}$$

must be zero over any path (including open paths).

Let us consider all possible rotation surfaces, that is, vector surfaces of the field rot **F**. Quite obviously, by virtue of the Stokes theorem

$$\int_C \mathbf{F}\,d\mathbf{r}=\iint_D \operatorname{rot}\mathbf{F}\cdot\mathbf{n}\,d\sigma,$$

where $d\mathbf{r}=\mathbf{i}\,dx+\mathbf{j}\,dy+\mathbf{k}\,dz$, and the integral (5.24) is zero over any closed path on the rotation surface (since the scalar product of the unit vector of the normal to the surface **n** and the vector rot **F** is zero) Now, from among the rotation surfaces let us choose those on which all the integrals

$$\int_L \mathbf{F}\,d\mathbf{r}=\int_L P\,dx+Q\,dy+R\,dz$$

are likewise zero over open paths. To construct such a surface that will pass through a given point $M(x_0, y_0, z_0)$, draw through this point M a line that is orthogonal to the vector lines of the field **F**. Such lines are defined by the equation

$$P\,dx+Q\,dy+R\,dz=0, \tag{5.21}$$

to which is added the equation of an arbitrary surface $z=f(x\ y)$, passing through the point M (the equation of this surface is most often taken in the form of $z=f_1(x)$ or $z=f_2(y)$ or even in the form of $z=a$, where a is a constant). Putting $z=f(x, y)$ into (5.21), we get an ordinary equation of the form

$$M(x, y)\,dx+N(x, y)\,dy=0,$$

by integrating this equation and taking into consideration the initial condition $y(x_0)=y_0$, we get the desired curve l which passes through the point $M(x_0, y_0, z_0)$ and which is orthogonal to the vector lines (Fig. 5.2).

If this line is not a rotation line, then by drawing a rotation line through each point of the line l, we get the desired surface S, which is orthogonal to the vector lines of the field **F**.

Indeed, taking any unclosed curve L on the surface S (Fig. 5.2) and drawing rotation lines through its boundary points to intersection with the curve l at the points p_1 and p_2, we get a closed contour consisting of the segment of line l between p_1 and p_2, the curve L and two rotation lines.

The line integral $\int_C P\,dz + Q\,dy + R\,dz$ taken along this closed path C is zero, since the path lies on the rotation surface while the same integral taken in a segment of the arc l and over segments of the rotation lines is zero, since the arc l and the rotation lines are orthogonal to the vector lines of the field $\mathbf{F}$ (the rotation lines are orthogonal to the vector lines of the field $\mathbf{F}$ by virtue of the condition $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})=0$). Hence the integral $\int_L P\,dx + Q\,dy + R\,dz$ along an arbitrarily chosen open path L is zero, i.e., the surface S is an integral surface of the equation (5.21) passing through the given point M.

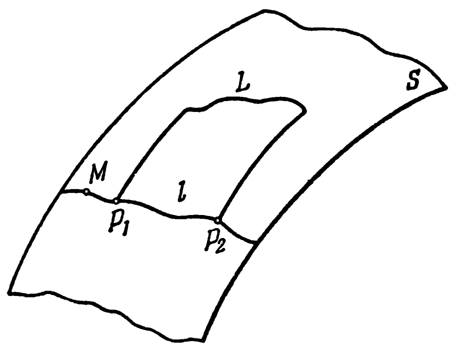


Fig. 5-2

This method for proving the sufficiency of the condition $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})=0$ for the existence of a family of surfaces orthogonal to the vector lines of the field $\mathbf{F}$ also points a way (true, not the shortest) to find these surfaces.

Example 1.

$$z\,dx + (x-y)\,dy + zy\,dz = 0.$$

The condition $(\mathbf{F}\cdot\operatorname{rot}\mathbf{F})=0$, where $\mathbf{F}=z\mathbf{i}+(x-y)\mathbf{j}+yz\mathbf{k}$, is not fulfilled, hence, this equation cannot be integrated by means of one relation.

Example 2.

$$(6x+yz)\,dx + (xz-2y)\,dy + (xy+2z)\,dz = 0.$$

Since $\operatorname{rot}\mathbf{F}\equiv 0$, where $\mathbf{F}=(6x+yz)\,\mathbf{i}+(xz-2y)\,\mathbf{j}+(xy+2z)\,\mathbf{k}$, it follows that $\mathbf{F}=\operatorname{grad} U$, where

$$U=\int_{(0,\,0,\,0)}^{(x,\,y,\,z)} (6x+yz)\,dx + (xz-2y)\,dy + (xy+2z)\,dz.$$

For the path of integration we choose a polygonal line the segments of which are parallel to the axes of coordinates. Integrating, we get $U = 3x^2 - y^2 + z^2 + xyz$ and, hence, the desired integral is

$$3x^2 - y^2 + z^2 + xyz = c.$$

Example 3.

$$yz\,dx + 2\,xz\,dp + xy\,dz = 0,$$

$$\mathbf{F} = yz\mathbf{i} + 2xz\mathbf{j} + xy\mathbf{k}, \quad \operatorname{rot} \mathbf{F} = -x\mathbf{i} + z\mathbf{k}.$$

The integrability condition $(\mathbf{F} \cdot \operatorname{rot} \mathbf{F}) = 0$ is fulfilled. On some surface, say on the plane $z = 1$, we find curves that are orthogonal to the vector lines:

$$z = 1, \quad y\,dx = 2x\,dy = 0, \quad xy^2 = a.$$

Through the curves of the family $z = 1$, $xy^2 = a$ we draw rotation surfaces; to do this we integrate the system of equations of rotation lines:

$$\frac{dx}{-x} = \frac{dy}{0} = \frac{dz}{z}, \quad y = c_1, \quad xz = c_2$$

Eliminating x, y and z from the equations $z = 1$, $xy^2 = a$, $y = c_1$, $xz = c_2$, we get $c_1^2 c_2 = a$. Hence, the desired integral of the original equation is of the form $xy^2 z = a$.

Note. Another common way of integrating the Pfaffian equation

$$P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz = 0 \qquad (5.21)$$

is to temporarily consider z (or some other variable) fixed and integrate the ordinary equation

$$P(x, y, z)\,dx + Q(x, y, z)\,dy = 0, \qquad (5.25)$$

in which z plays the role of a parameter.

Having obtained the integral of the equation (5.25)

$$U(x, y, z) = c(z), \qquad (5.26)$$

in which the arbitrary constant may be a function of the parameter z, we choose this function $c(z)$ so that equation (5.21) is satisfied. Differentiating (5.26), we get

$$\frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy + \left[\frac{\partial U}{\partial z} - c'(z)\right] dz = 0 \qquad (5.27)$$

The coefficiients of the differentials of the variables in the equations (5.21) and (5.27) must be proportional:

$$\frac{\frac{\partial U}{\partial x}}{P} = \frac{\frac{\partial U}{\partial y}}{Q} = \frac{\frac{\partial U}{\partial z} - c(z)}{R}.$$

From the equation $\frac{\frac{\partial U}{\partial x}}{P}=\frac{\frac{\partial U}{\partial z}-c'(z)}{R}$ it is possible to determine $c'(z)$ since one can prove that, provided the condition $(\mathbf{F}\cdot\text{rot}\,\mathbf{F})=0$ is fulfilled, this equation contains only z, $c'(z)$ and $U(x, y, z)=c(z)$.

4. First-Order Nonlinear Equations

First let us consider a case where the desired function depends on two independent variables. First-order partial differential equations in three variables are of the form

$$F(x, y, z, p, q)=0, \tag{5.28}$$

where

$$p=\frac{\partial z}{\partial x},\quad q=\frac{\partial z}{\partial y}.$$

The differential equation (5.28) establishes, at every point (x, y, z) of the region in which the first three arguments vary, the relation $\varphi(p, q)=0$ between the numbers p and q which define the direction of the normal $\mathbf{N}(p, q, -1)$ to the desired *integral surfaces* $z=z(x,y)$ of equation (5.28).

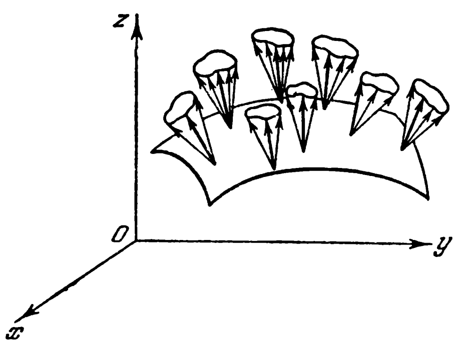


Fig. 5-3

Thus, the direction of the normal to the desired integral surfaces at a certain point (x, y, z) is not defined exactly; there is only isolated a one-parameter family of possible directions of normals—a certain cone of admissible directions of the normals $\mathbf{N}(p, q, -1)$, where p and q satisfy the equation $\varphi(p, q)=0$ (Fig. 5.3).

Thus, the problem of integrating the equation (5.28) reduces to finding the surfaces $z=z(x, y)$, the normals to which would, at every point, be directed along one of the permissible directions of the cone of normals at that point.

Proceeding from this geometric interpretation, we shall indicate a method of finding the integral of the equation (5.28) which depends on an arbitrary function, if its integral $\Phi(x, y, z, a, b)=0$ that depends on two parameters a and b is known.

The integral $\Phi(x, y, z, a, b)=0$ of (5.28), which depends on two essential arbitrary constants a and b, is called the *complete integral*.

Since the original differential equation (5.28) imposes restrictions solely on the direction of the normals to the desired integral surfaces, every surface, the normals to which coincide with the normals to the integral surfaces at the same points, will be an integral surface. Consequently, the envelopes of a two-parameter or one-parameter family of integral surfaces will be integral surfaces, since the normal to the envelope coincides with the normal to one of the integral surfaces of the family passing through the same point.

The envelope of a two-parameter family of integral surfaces, on the assumption of the existence of bounded partial derivatives $\frac{\partial\Phi}{\partial x}$, $\frac{\partial\Phi}{\partial y}$, $\frac{\partial\Phi}{\partial z}$ not vanishing simultaneously, and the existence of the derivatives $\frac{\partial\Phi}{\partial a}$ and $\frac{\partial\Phi}{\partial b}$, is defined by the equations

$$\Phi(x, y, z, a, b) = 0, \quad \frac{\partial\Phi}{\partial a} = 0, \quad \frac{\partial\Phi}{\partial b} = 0. \tag{5.29}$$

We also obtain an integral surface by isolating from the two-parameter family of integral surfaces $\Phi(x, y, z, a, b) = 0$, in arbitrary fashion, a one-parameter family (to do this, we take b as an arbitrary differentiable function of the parameter a); and in finding the envelope of the one-parameter family $\Phi(x, y, z, a, b(a)) = 0$, we also obtain an integral surface. Assuming that bounded derivatives of the function Φ exist with respect to all arguments and that the derivatives $\frac{\partial\Phi}{\partial x}$, $\frac{\partial\Phi}{\partial y}$, $\frac{\partial\Phi}{\partial z}$ do not vanish simultaneously, the envelope of this one-parameter family is given by the equation

$$\Phi(x, y, z, a, b(a)) = 0 \quad \text{and} \quad \frac{\partial}{\partial a}\{\Phi(x, y, z, a, b(a)\} = 0$$

or

$$\Phi(x, y, z, a, b(a)) = 0 \quad \text{and} \quad \frac{\partial\Phi}{\partial a} + \frac{\partial\Phi}{\partial b} b'(a) = 0. \tag{5.30}$$

These two equations define a set of integral surfaces that depends on the choice of an arbitrary function $b = b(a)$. Of course, the presence in equation (5.30) of an arbitrary function does not permit us to assert that the equations (5.30) define the set of all integral surfaces of the original equation (5.28) without exception. For example, this set, generally speaking, does not contain the integral surface defined by the equations (5.29), but still the presence of an arbitrary function in (5.30) is usually sufficient for one to isolate an integral surface that satisfies the given initial Cauchy conditions (see page 252).

Thus, knowing the complete integral, it is possible to construct an integral that depends on an arbitrary function.

In many cases, it is not at all difficult to find the complete integral, for instance:

(1) If equation (5.28) is of the form $F(p, q)=0$ or $p=\varphi(q)$, then by putting $q=a$, where a is an arbitrary constant, we get

$$p=\varphi(a), \quad dz=p\,dx+q\,dy=\varphi(a)\,dx+a\,dy,$$

whence

$$z=\varphi(a)\,x+ay+b$$

is the complete integral.

(2) If equation (5.28) can be reduced to the form $\varphi_1(x, p)=\varphi_2(y, q)$, then, putting $\varphi_1(x, p)=\varphi_2(y, q)=a$, where a is an arbitrary constant, and solving (if this is possible) for p and q, we get $p=\psi_1(x, a)$, $q=\psi_2(y, a)$,

$$dz=p\,dx+q\,dy=\psi_1(x, a)\,dx+\psi_2(y, a)\,dy,$$

$$z=\int \psi_1(x, a)\,dx+\int \psi_2(y, a)\,dy+b$$

is the complete integral.

(3) If the equation (5.28) is of the form $F(z, p, q)=0$, then by putting $z=z(u)$, where $u=ax+y$, we obtain

$$F\left(z, a, \frac{dz}{du}, \frac{dz}{du}\right)=0.$$

Integrating this ordinary equation, we get $z=\Phi(u, a, b)$, where b is an arbitrary constant, or

$$z=\Phi(ax+y, a, b)$$

is the complete integral.

(4) If the equation (5.28) is of a form resembling the Clairaut equation:

$$z=px+qy+\varphi(p, q),$$

then, as may be readily verified by direct substitution, the complete integral is

$$z=ax+by+\varphi(a, b).$$

Example 1. Find the complete integral of the equation $p=3q^3$.

$$q=a, \quad p=3a^3, \quad dz=3a^3\,dx+a\,dy,$$
$$z=3a^3x+ay+b.$$

Example 2. Find the complete integral of the equation $pq=2xy$.

$$\frac{p}{x}=\frac{2y}{q}=a, \quad p=ax, \quad q=\frac{2y}{a}, \quad dz=ax\,dx+\frac{2y}{a}\,dy,$$
$$z=\frac{ax^2}{2}+\frac{y^2}{a}+b.$$

Example 3. Find the complete integral of the equation $z^3 = pq^2$.

$$z = z(u), \text{ where } u = ax + y, \ p = a\frac{dz}{du}, \ q = \frac{dz}{du},$$

$$z^3 = a\left(\frac{dz}{du}\right)^3 \text{ or } \frac{dz}{du} = a_1 z, \text{ where } a_1 = a^{-\frac{1}{3}},$$

$$\ln|z| = a_1 u + \ln b, \ z = be^{a_1 u},$$

$$z = be^{a_1\left(\frac{x}{a_1^3} + y\right)}.$$

Example 4. Find the complete integral of the equation

$$z = px + qy + p^2 + q^2.$$

The complete integral is

$$z = ax + by + a^2 + b^2.$$

In more complicated cases, the complete integral of the equation

$$F(x, y, z, p, q) = 0$$

is found by one of the general methods.

The simplest idea is that underlying the *method of Lagrange and Charpit*. In this method, an equation

$$U(x, y, z, p, q) = a \qquad (5.31)$$

is chosen for the equation

$$F(x, y, z, p, q) = 0 \qquad (5.28)$$

so that the functions $p = p(x, y, z, a)$ and $q = q(x, y, z, a)$, which are determined from the system of equations (5.28) and (5.31), should lead to the Pfaffian equation that is integrable by a single relation

$$dz = p(x, y, z, a)\,dx + q(x, y, z, a)\,dy. \qquad (5.32)$$

Then the integral of the Pfaffian equation $\Phi(x, y, z, a, b) = 0$, where b is an arbitrary constant appearing during integration of (5.32), will be the complete integral of the equation (5.28). The function U is determined from the integrability condition of the equation (5.32) by single relation:

$$(\mathbf{F} \cdot \operatorname{rot} \mathbf{F}) = 0, \quad \text{where} \quad \mathbf{F} = p(x, y, z, a)\,\mathbf{i} + q(x, y, z, a)\,\mathbf{j} - \mathbf{k},$$

that is, in expanded form, from the equation

$$p\frac{\partial q}{\partial z} - q\frac{\partial p}{\partial z} - \frac{\partial p}{\partial y} + \frac{\partial q}{\partial x} = 0. \qquad (5.33)$$

The derivatives $\frac{\partial q}{\partial x}$, $\frac{\partial p}{\partial y}$, $\frac{\partial p}{\partial z}$, $\frac{\partial q}{\partial z}$ are computed by differentiating the identities

$$\left.\begin{array}{l} F(x, y, z, p, q)=0, \\ U(x, y, z, p, q)=a, \end{array}\right\} \tag{5.34}$$

in which p and q are regarded as functions of x, y and z, which are defined by the system (5.34).

Differentiating with respect to x, we get

$$\frac{\partial F}{\partial x}+\frac{\partial F}{\partial p}\frac{\partial p}{\partial x}+\frac{\partial F}{\partial q}\frac{\partial q}{\partial x}=0,$$
$$\frac{\partial U}{\partial x}+\frac{\partial U}{\partial p}\frac{\partial p}{\partial x}+\frac{\partial U}{\partial q}\frac{\partial q}{\partial x}=0,$$

whence

$$\frac{\partial q}{\partial x}=-\frac{\frac{D(F, U)}{D(p, x)}}{\frac{D(F, U)}{D(p, q)}}.$$

Similarly, defferentiating (5.34) with respect to y and determining $\frac{\partial p}{\partial y}$, we have

$$\frac{\partial p}{\partial y}=-\frac{\frac{D(F, U)}{D(y, q)}}{\frac{D(F, U)}{D(p, q)}}.$$

Differentiating (5.34) with respect to z and solving for $\frac{\partial p}{\partial z}$, $\frac{\partial q}{\partial z}$, we will have

$$\frac{\partial p}{\partial z}=-\frac{\frac{D(F, U)}{D(z, q)}}{\frac{D(F, U)}{D(p, q)}},$$

$$\frac{\partial q}{\partial z}=-\frac{\frac{D(F, U)}{D(p, z)}}{\frac{D(F, U)}{D(p, q)}}.$$

Substituting the computed derivatives into the integrability condition (5.33) and multiplying by the determinant $\frac{D(F, U)}{D(p, q)}$, which we assume to be different from zero, we get

$$p\left(\frac{\partial F}{\partial z}\frac{\partial U}{\partial p}-\frac{\partial F}{\partial p}\frac{\partial U}{\partial z}\right)+q\left(\frac{\partial F}{\partial z}\frac{\partial U}{\partial q}-\frac{\partial F}{\partial q}\frac{\partial U}{\partial z}\right)+$$
$$+\left(\frac{\partial F}{\partial y}\frac{\partial U}{\partial q}-\frac{\partial F}{\partial q}\frac{\partial U}{\partial y}\right)+\left(\frac{\partial F}{\partial x}\frac{\partial U}{\partial p}-\frac{\partial F}{\partial p}\frac{\partial U}{\partial x}\right)=0$$

or

$$\frac{\partial F}{\partial p}\frac{\partial U}{\partial x}+\frac{\partial F}{\partial q}\frac{\partial U}{\partial y}+\left(p\frac{\partial F}{\partial p}+q\frac{\partial F}{\partial q}\right)\frac{\partial U}{\partial z}-$$
$$-\left(\frac{\partial F}{\partial x}+p\frac{\partial F}{\partial z}\right)\frac{\partial U}{\partial p}-\left(\frac{\partial F}{\partial y}+q\frac{\partial F}{\partial z}\right)\frac{\partial U}{\partial q}=0. \quad (5\ 35)$$

To determine the function U we obtained a homogeneous linear equation (5.35), which may be integrated by the method indicated in Sec. 2 of this chapter: an equation of the characteristics is formed:

$$\frac{dx}{\frac{\partial F}{\partial p}}=\frac{dy}{\frac{\partial F}{\partial q}}=\frac{dz}{p\frac{\partial F}{\partial p}+q\frac{\partial F}{\partial q}}=-\frac{dp}{\frac{\partial F}{\partial x}+p\frac{\partial F}{\partial z}}=-\frac{dq}{\frac{\partial F}{\partial y}+q\frac{\partial F}{\partial z}}, \quad (5.36)$$

then at least one first integral of the system (5.36) is found,

$$U_1(x,\ y,\ z,\ p,\ q)=a,$$

and if the functions F and U_1 are independent of p and q, i. e. $\frac{D(F,U_1)}{D(p,q)}\neq 0$, then the first integral $U_1(x, y, z, p, q)$ will be the desired solution of the equation (5.35).

Thus, by determining $p=p(x, y, z, a)$ and $q=q(x, y, z, a)$ from the system of equations

$$F(x,\ y,\ z,\ p,\ q)=0,$$
$$U_1(x,\ y,\ z,\ p,\ q)=a$$

and substituting into

$$dz=p(x,\ y,\ z,\ a)\,dx+q(x,\ y,\ z,\ a)\,dy,$$

we get the Pfaffian equation integrable by a single relation, which, when solved, yields the complete integral of the original equation:

$$\Phi(x,\ y,\ z,\ a,\ b)=0.$$

Example 5. Find the complete integral of the equation

$$yzp^2-q=0. \quad (5.37)$$

The system (5.36) is of the form

$$\frac{dx}{2pyz}=-dy=\frac{dz}{2p^2yz-q}=-\frac{dp}{yp^3}=-\frac{dp}{zp^2+yp^2q}.$$

Taking advantage of the original equation, we simplify the denominator of the third ratio and obtain the integrable combination $\frac{dz}{p^2yz}=-\frac{dp}{p^3y}$, whence

$$p=\frac{a}{z}. \quad (5.38)$$

From equations (5.37) and (5.38) we find $p=\frac{a}{z}$, $q=\frac{a^2y}{z}$, whence $dz=\frac{a}{z}dx+\frac{a^2y}{z}dy$. Multiplying by $2z$ and integrating, we find the complete integral of the original equation $z^2=2ax+a^2y^2+b$.

Knowing the complete integral $\Phi(x, y, z, a, b)=0$ of the equation

$$F(x, y, z, p, q)=0,$$

it is, generally speaking, possible to solve the basic initial problem (see p. 252) or even the more general problem of determining the integral surface that passes through a given curve,

$$x=x(t), \quad y=y(t), \quad z=z(t). \tag{5.39}$$

Define the function $b=b(a)$ so that the envelope of the one-parameter family

$$\Phi(x, y, z, a, b(a))=0, \tag{5.40}$$

defined by the equations (5.40) and

$$\frac{\partial\Phi}{\partial a}+\frac{\partial\Phi}{\partial b}b'(a)=0, \tag{5.41}$$

should pass through the given curve (5.39).

At points of the given curve, both equations (5.40) and (5.41) with respect to t reduce to identities:

$$\Phi(x(t), y(t), z(t), a, b(a))=0 \tag{5.42}$$

and

$$\frac{\partial\Phi(x(t), y(t), a, b(a))}{\partial a}+\frac{\partial\Phi(x(t), y(t), z(t), a, b(a))}{\partial b}b'(a)=0. \tag{5.43}$$

However, it would be rather complicated to determine the function $b=b(a)$ from these equations. It is much easier to determine this function from the system of equations (5.42) and

$$\frac{\partial\Phi}{\partial x}x'(t)+\frac{\partial\Phi}{\partial y}y'(t)+\frac{\partial\Phi}{\partial z}z'(t)=0. \tag{5.44}$$

or in abbreviated notation

$$(\mathbf{N}\cdot\mathbf{t})=0.$$

where $\mathbf{t}$ is the vector of the tangent to the given curve

$$x=x(t), \quad y=y(t), \quad z=z(t), \tag{5.39}$$

and $\mathbf{N}$ is the vector of the normal to the surface $\Phi=0$, and hence also to the desired envelope at the appropriate points. The condition (5.44) is geometrically obvious, since the desired surface must pass through the given curve and, consequently, the tangent to this curve must lie in the plane tangential to the desired surface.

Example 6. Find the integral surface of the equation $z=px+ +qy+\frac{pq}{4}$ that passes through the curve $y=0$, $z=x^2$.

The complete integral of this equation (see case (4) on page 273) is of the form $z=ax+by+\frac{ab}{4}$. The equation of the given curve may be written in parametric form $x=t$, $y=0$, $z=t^2$.

To determine the function $b=b(a)$, we form a system of equations (5.42) and (5.44), which in the given case are of the form $t^2=at+\frac{ab}{4}$ and $2t=a$, whence $b=-a$, $z=a(x-y)-\frac{a^2}{4}$. The envelope of this family is determined by the equations

$$z=a(x-y)-\frac{a^2}{4}$$

and

$$x-y-\frac{a}{2}=0.$$

Eliminating a, we get $z=(x-y)^2$.

If the system (5.36) (page 276) is easy to integrate, then the *method of characteristics (Cauchy's method*—see below) is very convenient for solving the generalized Cauchy problem that has been posed.

The *integral surface* $z=z(x, y)$ of the equation

$$F(x, y, z, p, q)=0$$

that passes through the given curve

$$x_0=x_0(s), \; y_0=y_0(s), \; z_0=z_0(s)$$

may [as in the case of the quasilinear equation (see page 256)] be pictured as consisting of points lying on a certain one-parameter family of curves

$$x=x(t, s), \qquad y=y(t, s), \qquad z=z(t, s),$$

where s is the parameter of the family, called *characteristics*.

First we find the family of characteristics that depends on several parameters, and then, drawing the characteristics through points of the curve

$$x_0=x_0(s), \qquad y_0=y_0(s), \qquad z_0=z_0(s)$$

and satisfying certain other conditions as well, we isolate the one-parameter family of curves in which s may be taken as the parameter:

$$x=x(t, s), \qquad y=y(t, s), \qquad z=z(t, s)$$

(Fig. 5.4). The set of points lying on these curves is what forms the desired integral surface. That in brief is the underlying idea of Cauchy's method.

Let $z=z(x, y)$ be the integral surface of the equation

$$F(x, y, z, p, q)=0. \quad (5.45)$$

Then, by differentiating the identity (5.45) with respect to x and with respect to y, we get

$$F_x+pF_z+F_p\frac{\partial p}{\partial x}+F_q\frac{\partial q}{\partial x}=0,$$

$$F_y+qF_z+F_p\frac{\partial p}{\partial y}+F_q\frac{\partial q}{\partial y}=0,$$

or, since $\frac{\partial q}{\partial x}=\frac{\partial p}{\partial y}$, we have

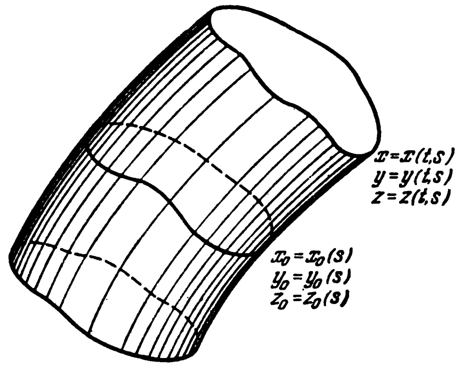


Fig. 5-4

$$\left.\begin{aligned} F_x+F_zp+F_p\frac{\partial p}{\partial x}+F_q\frac{\partial p}{\partial y}=0,\\ F_y+F_zq+F_p\frac{\partial q}{\partial x}+F_q\frac{\partial q}{\partial y}=0. \end{aligned}\right\} \quad (5.46)$$

The equations of the characteristics for the system of equations (5.46), which is quasilinear in p and q, and z is considered a known function of x and y, are of the form (see p. 264)

$$\frac{dx}{F_p}=\frac{dy}{F_q}=-\frac{dp}{F_x+pF_z}=-\frac{dq}{F_y+qF_z}=dt. \quad (5.47)$$

Since z is connected with p and q by the equation

$$dz=p\,dx+q\,dy, \quad (5.48)$$

it follows that, along the characteristic,

$$\frac{dz}{dt}=p\frac{dx}{dt}+q\frac{dy}{dt}=pF_p+qF_q$$

or

$$\frac{dz}{pF_p+qF_q}=dt, \quad (5.49)$$

which enables us to supplement the system (5.47) with yet another equation (5.49).

Thus, assuming that $z=z(x, y)$ is the solution of equation (5.45) we arrive at the system

$$\frac{dx}{F_p}=\frac{dy}{F_q}=\frac{dz}{pF_p+qF_q}=-\frac{dp}{F_x+pF_z}=-\frac{dq}{F_y+qF_z}=dt. \qquad (5.50)$$

From (5.50) it is possible, without knowing the solution $z=z(x, y)$ of the equation (5.45), to find the functions $x=x(t)$, $y=y(t)$, $z=z(t)$, $p=p(t)$, $q=q(t)$; that is, we can find the curves

$$x=x(t), \qquad y=y(t), \qquad z=z(t),$$

called *characteristics*, and in each point of a characteristic we can find the numbers $p=p(t)$ and $q=q(t)$ that determine the direction of the plane

$$Z-z=p(X-x)+q(Y-y). \qquad (5.51)$$

The characteristic, together with the plane (5.51) referred to each of its points, is called a *characteristic strip*.

We shall show that it is possible, from characteristics, to form the desired integral surface of the equation $F(x, y, z, p, q)=0$.

First of all note that the function F retains a constant value along the integral curve of the system (5.50):

$$F(x, y, z, p, q)=c,$$

in other words, the function $F(x, y, z, p, q)$ is a first integral of the system (5.50).

Indeed, along the integral curve of the system (5.50).

$$\frac{d}{dt}F(x, y, z, p, q)=F_x\frac{dx}{dt}+F_y\frac{dy}{dt}+F_z\frac{dz}{dt}+F_p\frac{dp}{dt}+F_q\frac{dq}{dt}=$$
$$=F_xF_p+F_yF_q+F_z(pF_p+qF_q)-F_p(F_x+pF_z)-F_q(F_y+qF_z)\equiv 0;$$

consequently, along the integral curve of the system (5.50),

$$F(x, y, z, p, q)=c, \quad \text{where} \quad c=F(x_0, y_0, z_0, p_0, q_0).$$

In order that the equation $F(x, y, z, p, q)=0$ should be satisfied along the integral curves of the system (5.50), the initial values $x_0(s)$, $y_0(s)$, $z_0(s)$, $p_0(s)$, $q_0(s)$ must be chosen so that they will satisfy the equation

$$F(x_0, y_0, z_0, p_0, q_0)=0.$$

Integrating the system (5.50) for initial values $x_0=x_0(s)$, $y_0=y_0(s)$, $z_0=z_0(s)$, $p_0=p_0(s)$, $q_0=q_0(s)$ that satisfy the equation $F(x_0, y_0, z_0, p_0, q_0)=0$, we get $x=x(t, s)$, $y=y(t, s)$, $z=z(t, s)$, $p=p(t, s)$, $q=q(t, s)$.

For a fixed s we will have one of the characteristics

$$x=x(t, s), \quad y=y(t, s), \quad z=z(t, s);$$

varying s, we get a certain surface, in each point of which, for $p=p(t, s)$, $q=q(t, s)$, the equation $F(x, y, z, p, q)=0$ is satisfied, but it is also necessary to find out whether, in the process, $p=\frac{\partial z}{\partial x}$ and $q=\frac{\partial z}{\partial y}$ or, what is the same thing, whether $dz=p\,dx+$ $+q\,dy$ or

$$dz=p\left(\frac{\partial x}{\partial s}ds+\frac{\partial x}{\partial t}dt\right)+q\left(\frac{\partial y}{\partial s}ds+\frac{\partial y}{\partial t}dt\right)=\frac{\partial z}{\partial s}ds+\frac{\partial z}{\partial t}dt,$$

which is equivalent to the two conditions

$$p\frac{\partial x}{\partial s}+q\frac{\partial y}{\partial s}-\frac{\partial z}{\partial s}=0, \tag{5.52}$$

$$p\frac{\partial x}{\partial t}+q\frac{\partial y}{\partial t}-\frac{\partial z}{\partial t}=0. \tag{5.53}$$

The latter of these equations obviously reduces to an identity, since we have already required, when forming the system (5.50), that $dz=p\,dx+q\,dy$ along the characteristic. Incidentally, this is quite evident from direct inspection, if one takes into account that by virtue of the system (5.50)

$$\frac{\partial x}{\partial t}=F_p, \quad \frac{\partial y}{\partial t}=F_q, \quad \frac{\partial z}{\partial t}=pF_p+qF_q$$

$\Big($in (5.50) in place of $\frac{\partial x}{\partial t}$, $\frac{\partial y}{\partial t}$, $\frac{\partial z}{\partial t}$ we wrote $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$, since we considered s fixed$\Big)$.

In order that the equation (5.52) may be satisfied, it is necessary to impose certain other restrictions on the choice of the initial values $x_0(s)$, $y_0(s)$, $z_0(s)$, $p_0(s)$, $q_0(s)$. Indeed, put

$$p\frac{\partial x}{\partial s}+q\frac{\partial y}{\partial s}-\frac{\partial z}{\partial s}=U \tag{5.54}$$

and prove that $U\equiv 0$ if the initial value $U|_{t=0}=0$, whence it will follow that if the initial functions

$$x_0(s), \quad y_0(s), \quad z_0(s), \quad p_0(s), \; q_0(s)$$

are so chosen that

$$p_0(s)\,x_0'(s)+q_0(s)\,y_0'(s)-z_0'(s)=0,$$

then $U\equiv 0$ for all t.

Differentiating (5.54) with respect to t, we get

$$\frac{\partial U}{\partial t}=\frac{\partial p}{\partial t}\frac{\partial x}{\partial s}=p\frac{\partial^2 x}{\partial t\,\partial s}+\frac{\partial q}{\partial t}\frac{\partial y}{\partial s}+q\frac{\partial^2 y}{\partial t\,\partial s}-\frac{\partial^2 z}{\partial t\,\partial s}$$

and, taking into account the result of differentiating the identity (5.53) with respect to s:

$$\frac{\partial p}{\partial s}\frac{\partial x}{\partial t}+p\frac{\partial^2 x}{\partial s\,\partial t}+\frac{\partial q}{\partial s}\frac{\partial y}{\partial t}+q\frac{\partial^2 y}{\partial s\,\partial t}-\frac{\partial^2 z}{\partial s\,\partial t}=0,$$

we will have

$$\frac{\partial U}{\partial t}=\frac{\partial p}{\partial t}\frac{\partial x}{\partial s}+\frac{\partial q}{\partial t}\frac{\partial y}{\partial s}-\frac{\partial p}{\partial s}\frac{\partial x}{\partial t}-\frac{\partial q}{\partial s}\frac{\partial y}{\partial t}$$

or, by virtue of the equations (5.50),

$$\frac{\partial U}{\partial t}=-(F_x+pF_z)\frac{\partial x}{\partial s}-(F_y+qF_z)\frac{\partial y}{\partial s}-F_p\frac{\partial p}{\partial s}-F_q\frac{\partial q}{\partial s}=$$
$$=-\left(F_x\frac{\partial x}{\partial s}+F_y\frac{\partial y}{\partial s}+F_z\frac{\partial z}{\partial s}+F_p\frac{\partial p}{\partial s}+F_q\frac{\partial q}{\partial s}\right)-$$
$$-F_z\left(p\frac{\partial x}{\partial s}+q\frac{\partial y}{\partial s}-\frac{\partial z}{\partial s}\right)=-\frac{\partial}{\partial s}\{F\}-F_zU=-F_zU,$$

since $F\equiv 0$, and hence the total partial derivative $\frac{\partial}{\partial s}\{F\}=0$. From the equation

$$\frac{\partial U}{\partial t}=-F_zU \tag{5.55}$$

we find $U=U_0e^{-\int_0^t F_z\,dt}$. Hence if $U_0=0$, then $U\equiv 0$, which incidentally follows also from the uniqueness of the solution $U\equiv 0$ of the linear equation (5.55), which satisfies the condition $U|_{t=0}=0$.

Thus, when integrating the equation

$$F(x,\ y,\ z,\ p,\ q)=0 \tag{5.45}$$

with initial conditions $x_0=x_0(s)$, $y_0=y_0(s)$, $z_0=z_0(s)$, use the Cauchy method to determine the functions $p_0=p_0(s)$ and $q_0=q_0(s)$ from the equations

$$F(x_0(s),\ y_0(s),\ z_0(s),\ p_0(s),\ q_0(s))=0$$

and

$$p_0(s)x_0'(s)+q_0(s)y_0'(s)-z_0'(s)=0$$

and then integrate the system of equations

$$\frac{dx}{F_p}=\frac{dy}{F_q}=\frac{dz}{pF_p+qF_q}=-\frac{dp}{F_x+pF_z}=-\frac{dq}{F_y+qF_z}=dt \tag{5.50}$$

with the initial conditions: for $t=0$

$$x=x_0(s),\ y=y_0(s),\ z=z_0(s),\ p=p_0(s),\ q=q_0(s).$$

The three functions

$$x=x(t,\ s),\ y=y(t,\ s),\ z=z(t,\ s)$$

of the solution of the system (5.50) yield in parametric form the equation of the desired integral surface of the equation (5.45).

The foregoing is readily generalized to nonlinear partial differential equations with an arbitrary number of independent variables

$$F(x_1, x_2, \ldots, x_n, z, p_1, p_2, \ldots, p_n)=0, \qquad (5.56)$$

where the

$$p_i=\frac{\partial z}{\partial x_i} \quad (i=1, 2, \ldots, n).$$

It is required to determine the integral n-dimensional surface $z=z(x_1, x_2, \ldots, x_n)$ of the equation (5.56) that passes through a given $(n-1)$-dimensional surface:

$$x_{i0}=x_{i0}(s_1, s_2, \ldots, s_{n-1}) \quad (i=1, 2, \ldots, n), \qquad (5.57)$$
$$z_0=z_0(s_1, s_2, \ldots, s_{n-1}).$$

For the time being suppose that we know the initial values of the functions

$$p_{i0}=p_{i0}(s_1, s_2, \ldots, s_{n-1}) \quad (i=1, 2, \ldots, n); \qquad (5.58)$$

then, integrating the auxiliary system of equations

$$\frac{dx_1}{F_{p_1}}=\frac{dx_2}{F_{p_2}}=\ldots=\frac{dx_n}{F_{p_n}}=\frac{dz}{\sum_{i=1}^{n} p_i F_{p_i}}=$$
$$=-\frac{dp_1}{F_{x_1}+p_1F_z}=\ldots=-\frac{dp_n}{F_{x_n}+p_nF_z}=dt \qquad (5.59)$$

with initial conditions (5.57) and (5.58), we get

$$\left.\begin{aligned} x_i&=x_i(t, s_1, s_2, \ldots, s_{n-1}),\\ z&=z(t, s_1, s_2, \ldots, s_{n-1}),\\ p_i&=p_i(t, s_1, s_2, \ldots, s_{n-1}). \end{aligned}\right\} \quad (i=1, 2, \ldots, n). \qquad (5.60)$$

For fixed $s_1, s_2, \ldots, s_{n-1}$, the equations (5.60) define, in a space with coordinates $x_1, x_2, \ldots, x_n, z$, curves called *characteristics*, to each point of which are also referred the numbers $p_i=p_i(t, s_1, s_2, \ldots, s_{n-1})$ that define the direction of certain planes

$$Z-z=\sum_{i=1}^{n} p_i(X_i-x_i). \qquad (5.61)$$

The characteristics together with the planes (5.61) form so-called *characteristic strips*.

Changing the parameters $s_1, s_2, \ldots, s_{n-1}$ yields an $(n-1)$-parameter family of characteristics

$$x_i=x_i(t, s_1, \ldots, s_{n-1}), \qquad z=z(t, s_1, \ldots, s_{n-1})$$

that pass through the given $(n-1)$-dimensional surface (5.57).

We shall show that for a definite choice of the functions

$$p_{i0}=p_{i0}(s_1, s_2, \dots, s_{n-1}) \qquad (i=1, 2, \dots, n)$$

points lying on the characteristics of the family (5.60) form the desired n-dimensional integral surface. Hence we have to prove that for a definite choice of the functions $p_{i0}(s_1, s_2, \dots, s_{n-1})$:

(1) $F(x_1(t, s_1, \dots, s_{n-1}), \dots, x_n(t, s_1, \dots, s_{n-1}),$
$z(t, s_1, \dots, s_{n-1}), p_1(t, s_1, \dots, s_{n-1}), \dots, p_n(t, s_1, \dots, s_{n-1}))\equiv 0,$

(2) $p_i=\frac{\partial z}{\partial x_i}$ $(i=1, 2, \dots, n)$ or, what is the same thing,

$$dz=\sum_{i=1}^{n} p_i\, dx_i.$$

It can be readily verified that the function $F(x_1, x_2, \dots, x_n, z, p_1, p_2, \dots, p_n)$ is a first integral of the system of equations (5 59). Indeed, along the integral curves of the system (5.59),

$$\frac{d}{dt}F(x_1, x_2, \dots, x_n, z, p_1, p_2, \dots, p_n)\equiv$$
$$\equiv\sum_{i=1}^{n} F_{x_i}\frac{dx_i}{dt}+F_z\frac{dz}{dt}+\sum_{i=1}^{n} F_{p_i}\frac{dp_i}{dt}\equiv$$
$$\equiv\sum_{i=1}^{n} F_{x_i}F_{p_i}+F_z\sum_{i=1}^{n} p_iF_{p_i}-\sum_{i=1}^{n} F_{p_i}(F_{x_i}+p_iF_z)\equiv 0$$

and, hence, along the integral curves of the system (5.59)

$$F(x_1, x_2, \dots, x_n, z, p_1, p_2, \dots, p_n)=c,$$

where c is a constant equal to $F(x_{10}, x_{20}, \dots, x_{n0}, z_0, p_{10}, p_{20}, \dots, p_{n0})$.

In order that the functions (5.60) should satisfy the equation (5 56) along the integral curves of the system (5.59), one has to choose the initial values $p_{i0}(s_1, s_2, \dots, s_{n-1})$ so that

$$F(x_{10}(s_1, \dots, s_{n-1}), \dots, x_{n0}(s_1, \dots, s_{n-1}), z(s_1, \dots, s_{n-1}),$$
$$p_1(s_1, \dots, s_{n-1}), \dots, p_n(s_1, \dots, s_{n-1}))=0.$$

It remains to verify that $dz=\sum_{i=1}^{n} p_i\, dx_i$ or

$$\frac{\partial z}{\partial t}dt+\sum_{j=1}^{n-1}\frac{\partial z}{\partial s_j}ds_j\equiv\sum_{i=1}^{n} p_i\left(\frac{\partial x_i}{\partial t}dt+\sum_{j=1}^{n-1}\frac{\partial x_i}{\partial s_j}ds_j\right).$$

This identity is equivalent to the following:

$$\frac{\partial z}{\partial t}-\sum_{i=1}^{n} p_i\frac{\partial x_i}{\partial t}\equiv 0 \tag{5.62}$$

and

$$\frac{\partial z}{\partial s_j}-\sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial s_j} \equiv 0 \qquad (j=1, 2, \ldots, n-1). \tag{5.63}$$

The validity of identity (5.62) becomes obvious if one takes into account that by virtue of the system (5.59)

$$\frac{\partial z}{\partial t}=\sum_{i=1}^{n} p_i F_{p_i} \quad \text{and} \quad \frac{\partial x_i}{\partial t}=F_{p_i} \qquad (i=1, 2, \ldots, n)$$

(in place of $\frac{dz}{dt}$ and $\frac{dx_i}{dt}$ we write the partial derivatives since all the s_i in the system (5.59) were assumed to be fixed).

To prove the indentities (5.63), which are true only for a definite choice of the initial values $p_{i0}(s_1, s_2, \ldots, s_{n-1})$ we put

$$U_j=\frac{\partial z}{\partial s_j}-\sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial s_j} \qquad (j=1, 2, \ldots, n-1)$$

and, differentiating U_j with respect to t, we get

$$\frac{\partial U_j}{\partial t}=\frac{\partial^2 z}{\partial t\, \partial s_j}-\sum_{i=0}^{n} p_i \frac{\partial^2 x_i}{\partial t\, \partial s_j}-\sum_{i=0}^{n} \frac{\partial p_i}{\partial t} \frac{\partial x_i}{\partial s_j}. \tag{5.64}$$

Taking into account the result of differentiating the identity (5.62) with respect to s_j

$$\frac{\partial^2 z}{\partial t\, \partial s_j}-\sum_{i=0}^{n} p_i \frac{\partial^2 x_i}{\partial t\, \partial s_j}-\sum_{i=0}^{n} \frac{\partial p_i}{\partial s_j} \frac{\partial x_i}{\partial t} \equiv 0,$$

we can rewrite the equation (5.64) as

$$\frac{\partial U_j}{\partial t}=\sum_{i=0}^{n} \frac{\partial p_i}{\partial s_j} \frac{\partial x_i}{\partial t}-\sum_{i=0}^{n} \frac{\partial p_i}{\partial t} \frac{\partial x_i}{\partial s_j}.$$

Taking advantage of the system (5.59), we have

$$\frac{\partial U_j}{\partial t}=\sum_{i=1}^{n} \frac{\partial p_i}{\partial s_j} F_{p_i}+\sum_{i=1}^{n} (F_{x_i}+p_i F_z) \frac{\partial x_i}{\partial s_j}=$$

$$=\sum_{i=1}^{n} \left(\frac{\partial F}{\partial x_i} \frac{\partial x_i}{\partial s_j}+\frac{\partial F}{\partial p_i} \frac{\partial p_i}{\partial s_j}\right)+\frac{\partial F}{\partial z} \frac{\partial z}{\partial s_j}-\frac{\partial F}{\partial z}\left(\frac{\partial z}{\partial s_j}-\sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial s_j}\right)=$$

$$=\frac{\partial}{\partial s_j}\{F\}-F_z U_j.$$

The total partial derivative $\frac{\partial}{\partial s_j}\{F\}=0$, since $F\equiv 0$ and, hence, the functions U_j are solutions of the homogeneous linear equations $\frac{\partial U_j}{\partial t}=-F_z U_j$, which have the unique solution $U_j\equiv 0$ if $U_j|_{t=0}=0$. Consequently, if the initial values $p_{i0}(s_1, s_2, \ldots, s_{n-1})$ $(i=1, 2, \ldots, n)$ are chosen so that $U_j|_{t=0}=0$ or $\left(\frac{\partial z}{\partial s_j}-\sum_{i=1}^{n} p_i\frac{\partial x_i}{\partial s_j}\right)_{t=0}=0$ $(j=1, 2, \ldots, n-1)$, then

$$\frac{\partial z}{\partial s_j}-\sum_{i=1}^{n} p_i\frac{\partial x_i}{\partial s_j}\equiv 0 \qquad (j=1, 2, \ldots, n-1)$$

and, hence, on the surface (5.60) $dz=\sum_{i=1}^{n} p_i\,dx_i$, i.e., $p_i=\frac{\partial z}{\partial x_i}$ $(i=1, 2, \ldots, n)$.

Thus, to find the integral surface of the equation $F(x_1, x_2, \ldots, x_n, z, p_1, p_2, \ldots, p_n)=0$ passing through the $(n-1)$-dimensional surface

$$x_{i0}=x_{i0}(s_1, s_2, \ldots, s_{n-1}) \qquad (i=1, 2, \ldots, n),$$
$$z_0=z_0(s_1, s_2, \ldots s_{n-1}),$$

it is necessary to determine the initial values $p_{i0}(s_1, s_2, \ldots, s_{n-1})$ from the equations

$$\left.\begin{aligned} &F(x_{10}, x_{20}, \ldots, x_{n0}, z, p_{10}, p_{20}, \ldots, p_{n0})=0,\\ &\frac{\partial z_0}{\partial s_j}-\sum_{i=1}^{n} p_{i0}\frac{\partial x_{i0}}{\partial s_j}=0 \qquad (j=1, 2, \ldots, n-1);\end{aligned}\right\} \tag{5.65}$$

then, integrating the system (5.59) (page 283) with initial conditions

$$\left.\begin{aligned} x_{i0}&=x_{i0}(s_1, s_2, \ldots, s_{n-1}),\\ z_0&=z_0(s_1, s_2, \ldots, s_{n-1}),\\ p_{i0}&=p_{i0}(s_1, s_2, \ldots, s_{n-1})\end{aligned}\right\} \quad (i=1, 2, \ldots, n),$$

we get

$$x_i=x_i(t, s_1, s_2, \ldots, s_{n-1}) \qquad (i=1, 2, \ldots, n),$$
$$z=z(t, s_1, s_2, \ldots, s_{n-1}), \tag{5.66}$$
$$p_i=p_i(t, s_1, s_2, \ldots, s_{n-1}) \qquad (i=1, 2, \ldots, n). \tag{5.67}$$

The equations (5.66) and (5.67) are the parametric equations of the desired integral surface.

Note. We assumed that the system of equations (5.65) was solvable for p_{i0} and also that the system (5.59) satisfies the conditions of the existence and uniqueness theorem.

Example 1. Find the integral surface of the equation $z=pq$ passing through the straight line $x=1$, $z=y$.

Write the equation of the straight line $x=1$, $z=y$ in parametric form, $x_0=1$, $y_0=s$, $z_0=s$. Find $p_0(s)$ and $q_0(s)$ from the equations (5.65): $s=p_0q_0$, $1-q_0=0$, whence $p_0=s$, $q_0=1$. Integrate the system (5.59):

$$\frac{dx}{q}=\frac{dy}{p}=\frac{dz}{2pq}=\frac{dp}{p}=\frac{dq}{q}=dt,$$

$$p=c_1e^t,\quad q=c_2e^t,\quad x=c_2e^t+c_3,\quad y=c_1e^t+c_4,\quad z=c_1c_2e^{2t}+c_5.$$

Taking into account that for $t=0$

$$x=1,\ y=s,\ z=s,\ p=s,\ q=1,$$

we have

$$p=se^t,\quad q=e^t,\quad x=e^t,\quad y=se^t,\quad z=se^{2t}.$$

Consequently, the desired integral surface is

$$x=e^t,\quad y=se^t,\quad z=se^{2t}\quad \text{or}\quad z=xy.$$

Example 2. Integrate the equation $\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2=2$ provided that for $x=0$, $z=y$, or in parametric form $x_0=0$, $y_0=s$, $z_0=s$.

Determine $p_0(s)$ and $q_0(s)$:

$$p_0^2+q_0^2=2,\quad 1-q_0=0,$$

whence $q_0=1$, $p_0=\pm 1$

Integrate the system of equations (5.59):

$$\frac{dx}{2p}=\frac{dy}{2q}=\frac{dz}{4}=\frac{dp}{0}=\frac{dq}{0}=dt,$$

$$p=c_1,\quad q=c_2,\quad x=2c_1t+c_3,\quad y=2c_2t+c_4,\quad z=4t+c_5;$$

using the initial conditions $p_0=\pm 1$, $q_0=1$, $x_0=0$, $y_0=s$, $z_0=s$, we obtain $p=\pm 1$, $q=1$, $x=\pm 2t$, $y=2t+s$, $z=4t+s$. The last three equations are the parametric equations of the desired integral surface. Eliminating the parameters t and s, we get $z=y\pm x$.

In problems of mechanics one often has to solve the Cauchy problem for the equation

$$\frac{\partial v}{\partial t}+H(t, x_1, x_2, \ldots, x_n, p_1, p_2, \ldots, p_n)=0, \qquad (5.68)$$

where the $p_i=\frac{\partial v}{\partial x_i}$, which is a special case of the equation (5.56) (p. 283). The Cauchy method, which, as applied to equation (5.68), is frequently called *Jacobi's first method*, yields the system of equations

$$dt = \frac{\partial x_1}{\frac{\partial H}{\partial p_1}} = \frac{dx_2}{\frac{\partial H}{\partial p_2}} = \ldots = \frac{dx_n}{\frac{\partial H}{\partial p_n}} = -\frac{dp_1}{\frac{\partial H}{\partial x_1}} =$$
$$= -\frac{dp_2}{\frac{\partial H}{\partial x_2}} = \ldots = -\frac{dp_n}{\frac{\partial H}{\partial x_n}} = \frac{dv}{\sum_{i=0}^{n} p_i \frac{\partial H}{\partial p_i} + \frac{\partial v}{\partial t}},$$

whence

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{\partial t} = -\frac{\partial H}{\partial x_i}, \qquad (i = 1, 2, \ldots, n) \tag{5.69}$$

and

$$\frac{dv}{dt} = \sum_{i=1}^{n} p_i \frac{\partial H}{\partial p_i} + \frac{\partial v}{\partial t}$$

or

$$\frac{dv}{dt} = \sum_{i=1}^{n} p_i \frac{\partial H}{\partial p_i} - H. \tag{5.70}$$

The system of $2n$ equations (5.69) does not contain v and may be integrated independently of equation (5.70); then the function v is found from (5.70) by a quadrature. Therein lies the specificity of applying Cauchy's method to equation (5.68). Besides, in this case there is no necessity of introducing an auxiliary parameter into the system (5.50), since that role can be successfully played by the independent variable t.

PROBLEMS ON CHAPTER 5

1. $\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y} = 0.$

2 $\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} = 2z.$

3 $x\frac{\partial z}{\partial y} = z.$

4. $z\frac{\partial z}{\partial x} - y\frac{\partial z}{\partial y} = 0.$

5. $y\frac{\partial z}{\partial x} = z$ for $x = 2,\ z = y.$

6. $x\frac{\partial z}{\partial x} - y\frac{\partial z}{\partial y} = z$ for $y = 1,\ z = 3x.$

7. $yz\frac{\partial z}{\partial x}+\frac{\partial z}{\partial y}=0$ for $x=0,\ z=y^2$.

8. Find surfaces that are orthogonal to the surfaces of the family $z=axy$.

9. Find surfaces that are orthogonal to the surfaces of the family $xyz=a$.

10. $\frac{x}{3}\frac{\partial z}{\partial x}-\frac{y}{5}\frac{\partial z}{\partial y}=z-5$.

11. $\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}+\frac{du}{\partial z}=0$.

12. $x\frac{\partial u}{\partial x}+2y\frac{\partial u}{\partial y}+3z\frac{\partial u}{\partial z}=4u$.

13. $\frac{\partial^2 z(x,\ y)}{\partial x^2}=0$.

14. $\frac{\partial z}{\partial x}-2x\frac{\partial z}{\partial y}=0$ for $x=1,\ z=y^2$.

15. Can the equation

$$(y^2+z^2-x^2)\,dx+xz\,dy+xy\,dz=0$$

be integrated by one relation?

16. Integrate by one relation the equation

$$(y+3z^2)\,dx+(x+y)dy+6xz\,dz=0.$$

17. Find the complete integral of the equation

$$pq=x^2y^2.$$

18. Find the complete integral of the equation

$$z=px+qy+p^2q^2.$$

19. Find the complete integral of the equation

$$pq=9z^2.$$

20. Find the complete integral of the equation

$$p=\sin q.$$

21. Find the surfaces that are orthogonal to the vector lines of the vector field

$$\mathbf{F}=(2xy-3yz)\,\mathbf{i}+(x^2-3xz)\,\mathbf{j}-3xu\,\mathbf{k}.$$

22. Find a family of surfaces that are orthogonal to the vector lines of the vector field

$$\mathbf{F} = (2x - y)\,\mathbf{i} + (3y - z)\,\mathbf{j} + (x - 2y)\,\mathbf{k}.$$

23. Find the vector lines, the vector surfaces and surfaces that are orthogonal to the vector lines of the field

$$\mathbf{F} = x\mathbf{i} + y\mathbf{j} - z\mathbf{k}.$$

24. $z = pq + 1$ for $y = 2$, $z = 2x + 1$.

25. $2z = pq - 3xy$ for $x = 5$, $z = 15y$.

26. $4z = p^2 + q^2$ for $x = 0$, $z = y^2$.

PART TWO

The calculus of variations

Introduction

Besides problems in which it is necessary to determine the maximal and minimal values of a certain function $z = f(x)$, in physics there are often encountered problems where one has to find the maximal and minimal values of special quantities called functionals.

Functionals are variable quantities whose values are determined by the choice of one or several functions.

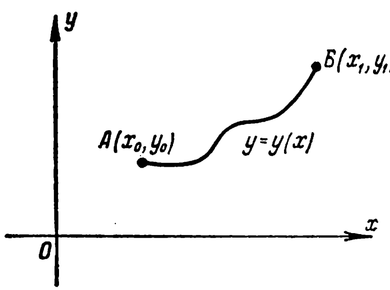


Fig. A

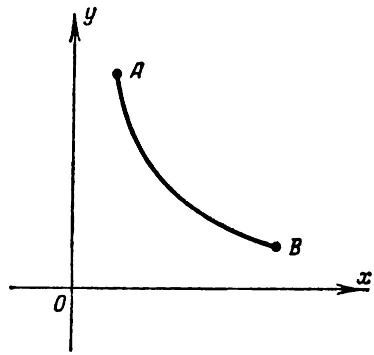


Fig. B

For example, the arc length l of a plane (or space) curve connecting two given points $A(x_0, y_0)$ and $B(x_1, y_1)$ (see Fig. A) is a functional. The quantity l may be computed if the equation of the curve $y = y(x)$ is given; then

$$l[y(x)] = \int_{x_0}^{x_1} \sqrt{1 + (y')^2}\, dx.$$

The area S of a surface is also a functional, since it is determined by the choice of surface, i.e., by the choice of the function $z(x, y)$ that enters into the equation of the surface $z = z(x, y)$. As is known,

$$S[z(x, y)] = \iint_D \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy,$$

where D is a projection of the surface on the xy-plane.

Moments of inertia, static moments, the coordinates of the centre of gravity of a homogeneous curve or surface are also functionals, since their values are determined by the choice of the curve or surface, i. e. by the choice of functions that enter into the equation of the curve or surface.

In all these examples we have a relationship that is characteristic of functionals: to a function (or vector function) there corresponds a number, whereas when we specify a function $z=f(x)$, to a number there corresponds a number.

The calculus of variations investigates methods that permit finding maximal and minimal values of functionals. Problems in which it is required to investigate a function for a maximum or a minimum are called *variational problems*.

Numerous laws of mechanics and physics reduce to the statement that a certain functional in a given process has to reach a minimum or a maximum. Thus stated, such laws are termed *variational principles* of mechanics or physics. The following are some variational principles or elementary consequences of them: the principle of least action, the law of conservation of energy, the law of conservation of momentum, the law of conservation of angular momentum, various variational principles of classical and relativistic field theory, Fermat's principle in optics, the principle of Castigliano in the theory of elasticity, and so forth.

The calculus of variations began to develop in 1696 and became an independent mathematical discipline with its own methods of investigation after the fundamental works of Euler (1707-1783), whom we may justifiably consider the founder of the calculus of variations.

Three problems exerted a considerable influence on the development of the calculus of variations:

The problem of the brachistochrone. In 1696 Johann Bernoulli published a letter in which he advanced the problem of the line of quickest descent *(brachistochrone)*. In this problem it is required to find the line connecting two specified points A and B that do not lie on a vertical line and possessing the property that a moving particle slides down this line from A to B in the shortest time (Fig. B).

It is easy to see that the line of quickest descent will not be the straight line connecting A and B, though that is the shortest distance between the two points, because the velocity of motion in a straight line will build up comparatively slowly; whereas if we take a curve that is steeper near A, even though the path becomes longer, a considerable portion of the distance will be covered at a greater speed. The problem of the brachistochrone was solved by Johann Bernoulli, Jacob Bernoulli, Leibnitz, Newton, L'Hospital.

It turned out that the line of quickest descent is a cycloid (see pages 316-317).

The problem of geodesics. It is required to determine the line of minimum length connecting two given points on a surface $\varphi(x, y, z) = 0$ (Fig. C). Such shortest lines are termed *geodesics*. This is a typical variational problem involving the so-called *connected* or *conditional extremum*. We have to find the minimum of the functional

$$l = \int_{x_0}^{x_1} \sqrt{1 + y'^2 + z'^2}\, dx,$$

and the functions $y(x)$ and $z(x)$ are subject to the condition $\varphi(x, y, z) = 0$. This problem was solved in 1698 by Jacob Bernoulli, but a general method for solving such problems was only given in the works of Euler and Lagrange.

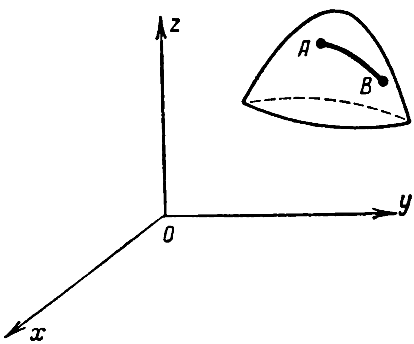


Fig. C

The isoperimetric problem. It is required to find a closed line of given length l bounding a maximum area S. This is the circle, as was known even in ancient Greece. In this problem one has to find the extremum of the functional S with the auxiliary peculiar condition that the length of the curve must be constant; that is, the functional

$$l = \int_{t_0}^{t_1} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}\, dt$$

retains a constant value. Conditions of this kind are called isoperimetric. General methods for solving problems with isoperimetric conditions were elaborated by Euler.

Methods will now be presented for solving a variety of variational problems; in the main the following functionals which are frequently encountered in applications are investigated for extrema:

$$\int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx,$$

$$\int_{x_0}^{x_1} F(x, y(x), y'(x), \ldots, y^{(n)}(x))\, dx,$$

$$\int_{x_0}^{x_1} F(x, y_1(x), \ldots, y_n(x), y_1'(x), \ldots, y_n'(x))\, dx,$$

$$\iint_D F\left(x, y, z(x, y), \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dx, dy,$$

in which the functions F are given, while the functions $y(x)$, $y_1(x), \ldots$ $\ldots, y_n(x)$, $z(x, y)$ are arguments of functionals.

The method of variations in problems with fixed boundaries

1. Variation and Its Properties

Methods of solving variational problems, i. e. problems involving the investigation of functionals for maxima and minima, are extremely similar to the methods of investigating functions for maxima and minima. It is therefore worth while recalling briefly the theory of maxima and minima of functions and in parallel introduce analogous concepts and prove similar theorems for functionals.

1. A variable z is a *function* of a variable quantity x [written $z=f(x)$] if to every value of x over a certain range of x there corresponds a value of z; i.e., we have a correspondence: to the number x there corresponds a number z.

Functions of several variables are defined in similar fashion.

1. A variable quantity v is a *functional* dependent on a function $y(x)$ [written $v=v[y(x)]$] if to each function $y(x)$ of a certain class of functions $y(x)$ there corresponds a value v, i.e. we have a correspondence: to the function $y(x)$ there corresponds a number v.

Functionals dependent on several functions, and functionals dependent on functions of several independent variables are similarly defined.

2. The *increment* Δx of the argument x of a function $f(x)$ is the difference between two values of the variable $\Delta x=x-x_1$. If x is the independent variable, then the differential x coincides with the increment, $dx=\Delta x$.

2. The *increment*, or *variation*, δy of the argument $y(x)$ of a functional $v[y(x)]$ is the difference between two functions $\delta y = y(x)-y_1(x)$. Here it is assumed that $y(x)$ varies in arbitrary fashion in some class of functions.

3. A function $f(x)$ is called *continuous* if to a small change of x there corresponds a small change in the function $f(x)$.

3. A functional $v[y(x)]$ is called *continuous* if to a small change of $y(x)$ there corresponds a small change in the functional $v[y(x)]$.

The latter definition requires some explanation, for the question immediately arises as to what changes of the function $y(x)$, which is the argument of the functional, are called small or, what is the same, what curves $y=y(x)$ and $y=y_1(x)$ are considered close or only slightly different.

It may be taken that the functions $y(x)$ and $y_1(x)$ are close if the absolute value of their difference $y(x)-y_1(x)$ is small for all values of x for which the functions $y(x)$ and $y_1(x)$ are prescribed; that is, we can consider as close such curves as have close-lying ordinates.

However, for such a definition of proximity of curves, the functionals of the kind

$$v[y(x)]=\int_{x_0}^{x_1} F(x, y, y')\,dx$$

that frequently occur in applications will be continuous only in exceptional cases due to the presence of the argument y' in the integrand function. For this reason, in many cases it is more natural to consider as close only those curves which have close-lying ordinates and are close as regards the directions of tangents at the respective points; that is, to require that, for close curves, not only should the absolute value of the difference $y(x)-y_1(x)$ be small, but also the absolute value of the difference $y'(x)-y_1'(x)$.

It is sometimes necessary to consider as close only those functions for which the absolute values of each of the following differences are small:

$$y(x)-y_1(x), \quad y'(x)-y_1'(x),$$
$$y''(x)-y_1''(x), \ \ldots, \ y^{(k)}(x)-y_1^{(k)}(x).$$

This compels us to introduce the following definitions of proximity of the curves $y=y(x)$ and $y=y_1(x)$.

The curves $y=y(x)$ and $y=y_1(x)$ are close in the sense of zero-order proximity if the absolute value of the difference $y(x)-y_1(x)$ is small.

The curves $y=y(x)$ and $y=y_1(x)$ are close in the sense of first-order proximity if the absolute values of the differences $y(x)-y_1(x)$ and $y'(x)-y_1'(x)$ are small.

The curves

$$y=y(x) \text{ and } y=y_1(x)$$

are close in the sense of kth order proximity if the absolute values of the differences

$$y(x)-y_1(x),$$
$$y'(x)-y_1'(x),$$
$$\cdots\cdots\cdots$$
$$y^{(k)}(x)-y_1^{(k)}(x)$$

are small.

Fig. 6.1 exhibits curves close in the sense of zero-order proximity but not close in the sense of first-order proximity, since the ordinates are close but the directions of the tangents are not. In Fig. 6.2. are depicted curves close in the sense of first-order proximity.

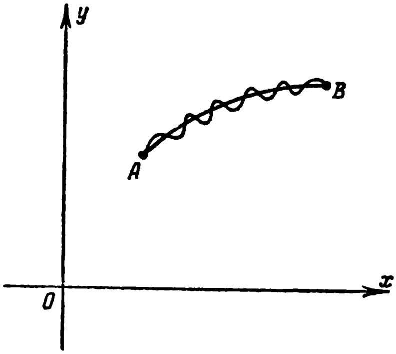


Fig. 6-1

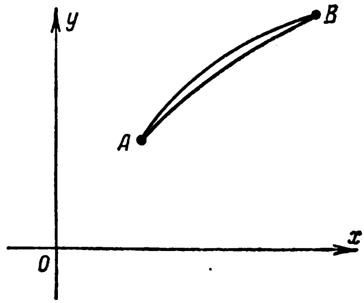


Fig. 6-2

From these definitions it follows that if the curves are close in the sense of kth order proximity, then they are definitely close in the sense of any lesser order of proximity.

We can now refine the concept of continuity of a functional.

3′. A function $f(x)$ is *continuous* at $x=x_0$ if for any positive ε there is a $\delta>0$ such that $|f(x)-f(x_0)|<\varepsilon$ for $|x-x_0|<\delta$.

It is assumed here that x takes on values at which the function $f(x)$ is defined.

3′. The functional $v[y(x)]$ is *continuous* at $y=y_0(x)$ in the sense of kth order proximity if for any positive ε there is a $\delta>0$ such that $|v[y(x)]-v[y_0(x)]|<\varepsilon$ for

$$|y(x)-y_0(x)|<\delta,$$
$$|y'(x)-y_0'(x)|<\delta,$$
$$\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots$$
$$|y^{(k)}(x)-y_0^{(k)}(x)|<\delta.$$

It is assumed here that the function $y(x)$ is taken from a class of functions on which the functional $v[y(x)]$ is defined.

One might also define the notion of distance $\rho(y_1, y_2)$ between the curves $y=y_1(x)$ and $y=y_2(x)$ $(x_0\leqslant x\leqslant x_1)$ and then close-lying curves would be curves with small separation.

If we assume that

$$\rho(y_1, y_2) = \max_{x_0 \leqslant x \leqslant x_1} |y_1(x) - y_2(x)|,$$

that is if we introduce the space metric C_0 (see pages 54-55), we have the concept of zero-order proximity. If it is taken that

$$\rho(y_1, y_2) = \sum_{p=1}^{k} \max_{x_0 \leqslant x \leqslant x_1} |y_1^{(p)}(x) - y_2^{(p)}(x)|$$

(it is assumed that y_1 and y_2 have continuous derivatives up to order k inclusive), then the proximity of the curves is understood in the sense of kth order proximity.

4. A *linear function* is a function $l(x)$ that satisfies the following conditions:

$$l(cx) = cl(x),$$

where c is an arbitrary constant, and

$$l(x_1 + x_2) = l(x_1) + l(x_2).$$

A linear function of one variable is of the form

$$l(x) = kx,$$

where k is constant.

4. A *linear functional* is a functional $L[y(x)]$ that satisfies the following conditions

$$L[cy(x)] = cL[y(x)],$$

where c is an arbitrary constant and

$$L[y_1(x) + y_2(x)] = L[y_1(x)] + L[y_2(x)].$$

The following is an instance of a linear functional:

$$L[y(x)] = \int_{x_0}^{x_1} (p(x)y + q(x)y')\,dx.$$

5. If the increment of a function

$$\Delta f = f(x + \Delta x) - f(x)$$

may be represented in the form

$$\Delta f = A(x)\Delta x + \beta(x, \Delta x)\cdot\Delta x,$$

where $A(x)$ does not depend on Δx, and $\beta(x, \Delta x) \to 0$ as $\Delta x \to 0$, then the function is called *differentiable*, while the part of the increment that is linear with respect to Δx—$A(x)\Delta x$—is called the *differential* of the function and is denoted by df. Dividing

5. If the increment of a functional

$$\Delta v = v[y(x) + \delta y] - v[y(x)]$$

may be represented in the form

$$\Delta v = L[y(x), \delta y] + \beta(y(x), \delta y)\max|\delta y|,$$

where $L[y(x), \delta y]$ is a functional linear with respect to δy, and $\max|\delta y|$ is the maximum value of $|\delta y|$ and $\beta(y(x), \delta y) \to 0$ as $\max|\delta y| \to 0$, then the part of the increment of the functional that is linear with respect to by Δx and passing to the limit as $\Delta x \to 0$, we find that $A(x) = f'(x)$ and, hence.

$$df = f'(x)\,\Delta x.$$

δy, i.e. $L[y(x), \delta y]$, is called the *variation* of the functional and is denoted by δv.

Thus the *variation of a functional is the principal part of the increment of the functional, which part is linear in* δy.

In the examination of functionals, the variation plays the same role as the differential does in investigations of functions.

Another, almost equivalent definition of the differential of a function and the variation of a functional may be given. Consider the value of the function $f(x+\alpha\Delta x)$ for fixed x and Δx and varying values of the parameter α. For $\alpha = 1$ we get an increased value of the function $f(x+\Delta x)$, for $\alpha = 0$ we get the initial value of the function $f(x)$. It may be readily verified that the derivative of $f(x+\alpha\Delta x)$ with respect to α for $\alpha = 0$ is equal to the differential of the function $f(x)$ at the point x. Indeed, by the rule for differentiating a composite function

$$\frac{\partial}{\partial\alpha} f(x+\alpha\,\Delta x)\,|_{\alpha=0} = f'(x+\alpha\Delta x)\,|_{x=0} = f'(x)\,\Delta x = df(x).$$

In the same way, for a function of several variables

$$z = f(x_1, x_2, \ldots, x_n)$$

one can obtain the differential by differentiating

$$f(x_1+\alpha\Delta x_1,\ x_2+\alpha\Delta x_2,\ \ldots,\ x_n+\alpha\Delta x_n)$$

with respect to α, assuming $\alpha = 0$. Indeed,

$$\frac{\partial}{\partial\alpha} f(x_1+\alpha\,\Delta x_1,\ x_2+\alpha\Delta x_2,\ \ldots, x_n+\alpha\Delta x_n)\,|_{\alpha=0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\Delta x_i = df.$$

Likewise, for functionals of the form $v[y(x)]$ or more complicated ones depending on several unknown functions or on functions of several variables, it is possible to determine the variation as a derivative of the functional $v[y(x)+\alpha\delta y]$ with respect to α for $\alpha = 0$. Indeed, if the functional has a variation in the sense of the principal linear part of the increment, then its increment will be of the form

$$\Delta v = v[y(x)+\alpha\delta y] - v[y(x)] = L(y,\ \alpha\,\delta y) + \beta(y,\ \alpha\,\delta y)\,|\alpha|\max|\delta y|.$$

The derivative of $v[y+\alpha\delta y]$ with respect to α at $\alpha = 0$ is

$$\lim_{\Delta\alpha\to 0}\frac{\Delta v}{\Delta\alpha} = \lim_{\alpha\to 0}\frac{\Delta v}{\alpha} = \lim_{\alpha\to 0}\frac{L(y,\ \alpha\,\delta y) + \beta[y(x),\ \alpha\,\delta y]\,|\alpha|\max|\delta y|}{\alpha} =$$

$$= \lim_{\alpha\to 0}\frac{L(y,\ \alpha\,\delta y)}{\alpha} + \lim_{\alpha\to 0}\frac{\beta[y(x),\alpha\,\delta y]\,|\alpha|\max|\delta y|}{\alpha} = L(y,\ \delta y)$$

since by virtue of linearity

$$L(y, \alpha\,\delta y) = \alpha L(y, \delta y)$$

and

$$\lim_{\alpha \to 0} \frac{\beta\,[y(x), \alpha\,\delta y]\,|\alpha|\max|\delta y|}{\alpha} = \lim_{\alpha \to 0} \beta\,[y(x), \alpha\,\delta y]\max|\delta y| = 0$$

because $\beta\,[y(x), \alpha\,\delta y] \longrightarrow 0$ as $\alpha \longrightarrow 0$. Thus, if there exists a variation in the sense of the principal linear part of the increment of the functional, then there also exists a variation in the sense of the derivative with respect to the parameter for the initial value of the parameter, and both of these definitions are equivalent.

The latter definition of a variation is somewhat broader than the former, since there are instances of functionals, from the increments of which it is impossible to isolate the principal linear part, but the variation exists in the meaning of the second definition.

6. The *differential* of a function $f(x)$ is equal to

$$\frac{\partial}{\partial \alpha} f(x + \alpha\,\Delta x)|_{\alpha=0}.$$

6. The *variation* of a functional $v[y(x)]$ is equal to

$$\frac{\partial}{\partial \alpha} v\,[y(x) + \alpha\,\delta y]\,|_{\alpha=0}.$$

Definition. *A functional* $v[y(x)]$ *reaches a maximum on a curve* $y = y_0(x)$ *if the values of the functional* $v[y(x)]$ *on any curve close to* $y = y_0(x)$ *do not exceed* $v\,[y_0(x)]$; that is $\Delta v = v\,[y(x)] - v\,[y_0(x)] \leqslant 0$.

If $\Delta v \leqslant 0$, and $\Delta v = 0$ only for $y(x) = y_0(x)$, then it is said that a *strict maximum* is reached on the curve $y = y_0(x)$. The curve $y = y_0(x)$, on which a *minimum* is achieved, is defined in similar fashion. In this case, $\Delta v \geqslant 0$ for all curves close to the curve $y = y_0(x)$.

7. ***Theorem.*** *If a differentiable function* $f(x)$ *achieves a maximum or a minimum at an interior point* $x = x_0$ *of the domain of definition of the function, then at this point*

$$df = 0.$$

7. ***Theorem.*** *If a functional* $v\,[y(x)]$ *having a variation achieves a maximum or a minimum at* $y = y_0(x)$, *where* $y(x)$ *is an interior point of the domain of definition of the functional, then at* $y = y_0(x)$,

$$\delta v = 0.$$

Proof of the theorem for functionals. For fixed $y_0(x)$ and δy $v\,[y_0(x) + \alpha\,\delta y] = \varphi(\alpha)$ is a function of α, which for $\alpha = 0$, by hypothesis, reaches a maximum or a minimum; hence, the derivative

$$\varphi'(0) = 0^{*}, \quad \text{and} \quad \frac{\partial}{\partial \alpha} v\,[y_0(x) + \alpha\,\delta y]\Big|_{\alpha=0} = 0,$$

* α can take on either positive or negative values in the neighbourhood of the point $\alpha = 0$, since $y_0(x)$ is an interior point of the domain of definition of the functional.

i. e. $\delta v = 0$. Thus, the variation of a functional is zero on curves on which an extremum of the functional is achieved.

The concept of the *extremum* of a functional must be made more specific. When speaking of a maximum or a minimum, more precisely, of a relative maximum or minimum, we had in view the largest or smallest value of the functional only relative to values of the functional on close-lying curves. But, as has already been pointed out, the proximity of curves may be understood in different ways, and for this reason it is necessary, in the definition of a maximum or minimum, to indicate the order of proximity.

If a functional $v[y(x)]$ reaches a maximum or a minimum on a curve $y = y_0(x)$ with respect to all curves for which the absolute value of the difference $y(x) - y_0(x)$ is small, i.e. with respect to curves close to $y = y_0(x)$ in the sense of zero-order proximity, then the maximum of minimum is called *strong*.

However, if a functional $v[y(x)]$ attains, on the curve $y = y_0(x)$, a maximum or minimum only with respect to curves $y = y(x)$ close to $y = y_0(x)$ in the sense of first-order proximity, i.e. with respect to curves close to $y = y_0(x)$ not only as regards ordinates but also as regards the tangent directions, then the maximum or the minimum is termed *weak*.

Quite obviously, if a strong maximum (or minimum) is attained on a curve $y = y_0(x)$, then most definitely a weak one has been attained as well, since if the curve is close to $y = y_0(x)$ in the sense of first-order proximity, then it is also close in the sense of zero-order proximity. It is possible, however, that on the curve $y = y_0(x)$ a weak maximum (minimum) has been attained, yet a strong maximum (minimum) is not achieved; in other words, among the curves $y = y(x)$ close to $y = y_1(x)$ both as to ordinates and as to the tangent directions, there may not be any curves for which $v[y(x)] > v[y_0(x)]$ (in the case of a minimum $v[y(x)] < v[y_0(x)]$), and among the curves $y = y(x)$ that are close as regards ordinates but not close as regards the tangent directions there may be those for which $v[y(x)] > v[y_0(x)]$ (in the case of a minimum $v[y(x)] < < v[y_0(x)]$). The difference between a strong and weak extremum will not have essential meaning in the derivation of the basic necessary condition for an extremum, but it will be extremely essential in Chapter 8 in studying the sufficient conditions for an extremum.

Note also that if on a curve $y = y_0(x)$ an extremum is attained, then not only $\frac{\partial}{\partial \alpha} v[y_0(x) + \alpha \delta y]\Big|_{\alpha=0} = 0$, but also $\frac{\partial}{\partial \alpha} v[y(x, \alpha)]\Big|_{\alpha=0} = 0$, where $y(x, \alpha)$ is any family of admissible curves, and for $\alpha = 0$ and $\alpha = 1$ the function $y(x, \alpha)$ must, respectively, transform to $y_0(x)$ and $y_0(x) + \delta y$. Indeed, $v[y(x, \alpha)]$ is a function of α since specifying α determines a curve of the family $y=y(x, \alpha)$, and this means that it also defines the value of the functional $v[y(x, \alpha)]$.

It is assumed that this function achieves an extremum at $\alpha=0$, hence, the derivative of this function vanishes at $\alpha=0$.*

Thus, $\frac{\partial}{\partial\alpha} v[y(x, \alpha)]\Big|_{\alpha=0}=0$, however, this derivative generally speaking will no longer coincide with the variation of the function but will, as has been shown above, vanish simultaneously with δv on curves that achieve an extremum of the functional.

All definitions of this section and the fundamental theorem (page 302) can be extended almost without any change to functionals dependent on several unknown functions:

$$v[y_1(x), y_2(x), \ldots, y_n(x)]$$

or dependent on one or several functions of many variables:

$$v[z(x_1, x_2, \ldots, x_n)],$$

$$v[z_1(x_1, x_2, \ldots, x_n), z_2(x_1, x_2, \ldots, x_n), \ldots, z_m(x_1, x_2, \ldots, x_n)].$$

For example, the variation δv of the functional $v[z(x, y)]$ may be defined either as the principal part of the increment

$$\Delta v=v[z(x, y)+\delta z]-v[z(x, y)],$$

linear in δz, or as a derivative with respect to the parameters for the initial value of the parameter

$$\frac{\partial}{\partial\alpha} v[z(x, y)+\alpha\,\delta z]\Big|_{\alpha=0},$$

and if for $z=z(x, y)$ the functional v attains an extremum, then for $z=z(x, y)$ the variation $\delta v=0$, since $v[z(x, y)+\alpha\delta z]$ is a function of α, which for $\alpha=0$, by hypothesis, attains an extremum and, hence, the derivative of this function with respect to α for $\alpha=0$ vanishes, $\frac{\partial}{\partial\alpha} v[z(x, y)+\alpha\delta z]\Big|_{\alpha=0}=0$ or $\delta v=0$.

2. Euler's Equation

Let us investigate the functional

$$v[y(x)]=\int_{x_0}^{x_1} F(x, y(x), y'(x))\,dx \tag{6.1}$$

---

* It is assumed that α can take on any values close to $\alpha=0$ and $\frac{\partial v[y(x, \alpha)]}{\partial\alpha}\Big|_{\alpha=0}$ exists.

for an extreme value, the boundary points of the admissible curves being fixed: $y(x_0)=y_0$ and $y(x_1)=y_1$ (Fig. 6.3). We will consider the function $F(x, y, y')$ three times differentiable.

We already know that a necessary condition for an extremum is that the variation of the functional vanish. We will now show how this basic theorem is applied to the functional under consideration, and we will repeat the earlier argument as applied to the functional (6.1). Assume that the extremum is attained on a twice-differentiable curve $y=y(x)$ (by only requiring that admissible

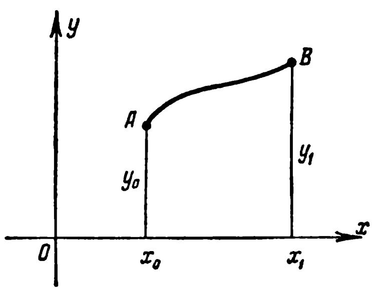


Fig. 6-3

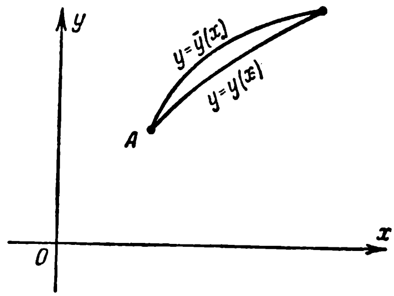


Fig. 6-4

curves have first-order derivatives, we can prove by a different method that the curve which achieves the extremum has a second derivative as well).

Take some admissible curve $y=\bar{y}(x)$ close to $y=y(x)$ and include the curves $y=y(x)$ and $y=\bar{y}(x)$ in a one-parameter family of curves

$$y(x, \alpha)=y(x)+\alpha(\bar{y}(x)-y(x));$$

for $\alpha=0$ we get the curve $y=y(x)$, for $\alpha=1$ we have $y=\bar{y}(x)$ (Fig. 6.4). As we already know, the difference $\bar{y}(x)-y(x)$ is called the variation of the function $y(x)$ and is symbolized as δy.

In variational problems, the variation δy plays a role similar to that of the increment of the independent variable Δx in problems involving investigating functions $f(x)$ for extreme values. The variation $\delta y=\bar{y}(x)-y(x)$ of the function is a function of x. This function may be differentiated once or several times; $(\delta y)'=\bar{y}'(x)-y'(x)=\delta y'$, that is, the derivative of the variation is equal to the variation of the derivative, and similarly

$$(\delta y)''=\bar{y}''(x)-y''(x)=\delta y'',$$

$$\cdots\cdots\cdots\cdots$$

$$(\delta y)^{(k)}=\bar{y}^{(k)}(x)-y^{(k)}(x)=\delta y^{(k)}.$$

We consider the family $y=y(x, \alpha)$, where $y(x, \alpha)=y(x)+\alpha\delta y$, which for $\alpha=0$ contains a curve on which an extreme value is achieved, and which for $\alpha=1$ contains a certain close-lying admissible curve, the so-called comparison curve.

If one considers the values of the functional

$$v[y(x)]=\int_{x_0}^{x_1} F(x, y, y')\,dx$$

only on curves of the family $y=y(x, \alpha)$, then the functional becomes a function of α:

$$v[y(x, \alpha)]=\varphi(\alpha),$$

since the value of the parameter α determines the curve of the family $y=y(x, \alpha)$ and thus determines also the value of the functional $v[y(x, \alpha)]$. This function $\varphi(\alpha)$ is extremized for $\alpha=0$ since for $\alpha=0$ we have $y=y(x)$, and the functional is assumed to have achieved an extremum in comparison with any neighbouring admissible curve and, in particular, with respect to curves of the family $y=y(x, \alpha)$ in the neighbourhood. A necessary condition for the extremum of the function $\varphi(\alpha)$ for $\alpha=0$ is, as we know, that its derivative for $\alpha=0$ vanish:

$$\varphi'(0)=0.$$

Since

$$\varphi(\alpha)=\int_{x_0}^{x_1} F(x, y(x, \alpha), y'_x(x, \alpha))\,dx,$$

it follows that

$$\varphi'(\alpha)=\int_{x_0}^{x_1}\left[F_y\frac{\partial}{\partial\alpha}y(x, \alpha)+F_{y'}\frac{\partial}{\partial\alpha}y'(x, \alpha)\right]dx,$$

where

$$F_y=\frac{\partial}{\partial y}F(x, y(x, \alpha), y'(x, \alpha)),$$

$$F_{y'}=\frac{\partial}{\partial y'}F(x, y(x, \alpha), y'(x, \alpha)),$$

or since

$$\frac{\partial}{\partial\alpha}y(x, \alpha)=\frac{\partial}{\partial\alpha}[y(x)+\alpha\delta y]=\delta y$$

and

$$\frac{\partial}{\partial\alpha}y'(x, \alpha)=\frac{\partial}{\partial\alpha}[y'(x)+\alpha\delta y']=\delta y',$$

we get

$$\varphi'(\alpha)=\int_{x_0}^{x_1}[F_y(x,\ y(x,\ \alpha),\ y'(x,\ \alpha))\,\delta y+$$
$$+F_{y'}(x,\ y(x,\ \alpha),\ y'(x,\ \alpha))\,\delta y']\,dx;$$

$$\varphi'(0)=\int_{x_0}^{x_1}[F_y(x,\ y(x),\ y'(x))\,\delta y+F_{y'}(x,\ y(x),\ y'(x))\,\delta y']\,dx.$$

As we already know, $\varphi'(0)$ is called the variation of the functional and is denoted by δv. A necessary condition for the extremum of a functional v is that its variation vanish: $\delta v=0$. For the functional

$$v[y(x)]=\int_{x_0}^{x_1}F(x,\ y,\ y')\,dx$$

this condition has the form

$$\int_{x_0}^{x_1}[F_y\,\delta y+F_{y'}\,\delta y']\,dx=0.$$

We integrate the second term by parts and, taking into account that $\delta y'=(\delta y)'$, we get

$$\delta v=[F_{y'}\,\delta y]_{x_0}^{x_1}+\int_{x_0}^{x_1}\left(F_y-\frac{d}{dx}F_{y'}\right)\delta y\,dx.$$

But

$$\delta y\,|_{x=x_0}=\bar{y}(x_0)-y(x_0)=0 \quad\text{and}\quad \delta y\,|_{x=x_1}=\bar{y}(x_1)-y(x_1)=0,$$

because all admissible curves in the elementary problem under consideration pass through fixed boundary points and, hence,

$$\delta v=\int_{x_0}^{x_1}\left(F_y-\frac{d}{dx}F_{y'}\right)\delta y\,dx.$$

Thus, the necessary condition for an extremum takes the form

$$\int_{x_0}^{x_1}\left(F_y-\frac{d}{dx}F_{y'}\right)\delta y\,dx=0, \tag{6.2}$$

the first factor $F_y-\frac{d}{dx}F_{y'}$ on the extremizing curve $y=y(x)$ is a given continuous function, while the second factor δy, because of the arbitrary choice of the comparison curve $y=\bar{y}(x)$, is an arbitrary function that satisfies only certain very general conditions, namely: at the boundary points $x=x_0$ and $x=x_1$ the function δy vanishes, it is continuous and differentiable once or several times; δy or δy and $\delta y'$ are small in absolute value.

To simplify the condition obtained, (6.2), let us take advantage of the following lemma.

The fundamental lemma of the calculus of variations. *If for every continuous function* $\eta(x)$

$$\int_{x_0}^{x_1} \Phi(x)\,\eta(x)\,dx=0,$$

where the function $\Phi(x)$ *is continuous on the interval* $[x_0, x_1]$, *then*

$$\Phi(x)\equiv 0$$

on that interval.

Note. The statement of the lemma and its proof do not change if the following restrictions are imposed on the functions: $\eta(x_0)=\eta(x_1)=0$; $\eta(x)$ has continuous derivatives to order p, $|\eta^{(s)}(x)|<\varepsilon$ $(s=0, 1, \ldots, q;\ q\leqslant p)$.

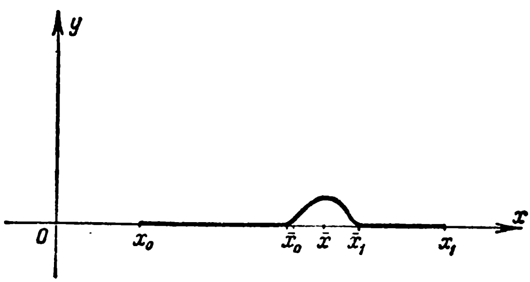


Fig. 6-5

Proof. Assuming that at the point $x=\bar{x}$ lying on the interval $x_0\leqslant x\leqslant x_1$, $\Phi(x)\neq 0$, we arrive at a contradiction. Indeed, from the continuity of the function $\Phi(x)$ it follows that if $\Phi(\bar{x})\neq 0$, then $\Phi(x)$ maintains its sign in a certain neighbourhood $(\bar{x}_0\leqslant x\leqslant\bar{x}_1)$ of the point $\bar{x}$; but then, having chosen a function $\eta(x)$, which also maintains its sign in this neighbourhood and is equal to zero outside this neighbourhood (Fig. 6.5), we get

$$\int_{x_0}^{x_1} \Phi(x)\,\eta(x)\,dx=\int_{\bar{x}_0}^{\bar{x}_1} \Phi(x)\,\eta(x)\,dx\neq 0,$$

since the product $\Phi(x)\eta(x)$ does not change sign on the interval $(\bar{x}_0 \leqslant x \leqslant \bar{x}_1)$ and vanishes outside this interval. We have thus arrived at a contradiction; hence, $\Phi(x) \equiv 0$. The function $\eta(x)$ may for example be chosen thus: $\eta(x) \equiv 0$ outside the interval $(\bar{x}_0 \leqslant x \leqslant \bar{x}_1)$; $\eta(x) = k(x-\bar{x}_0)^{2n}(x-\bar{x}_1)^{2n}$ on the interval $(\bar{x}_0 \leqslant x \leqslant \bar{x}_1)$, where n is a positive integer and k is a constant factor. It is obvious that the function $\eta(x)$ satisfies the above conditions: it is continuous, has continuous derivatives up to order $2n-1$, vanishes at the points x_0 and x_1 and may be made arbitrarily small in absolute value together with its derivatives by reducing the absolute value of the constant k.

Note. Repeating this argument word for word, one can prove that if the function $\Phi(x, y)$ is continuous in the region D on the plane (x, y) and $\iint\limits_D \Phi(x, y)\eta(x, y)\,dx\,dy = 0$ for an arbitrary choice of the function $\eta(x, y)$ satisfying only certain general conditions (continuity, differentiability once or several times, and vanishing at the boundaries of the region D, $|\eta| < \varepsilon$, $|\eta'_x| < \varepsilon$, $|\eta'_y| < \varepsilon$), then $\Phi(x, y) \equiv 0$ in the region D. When proving the fundamental lemma, the function $\eta(x, y)$ may be chosen, for example, as follows: $\eta(x, y) \equiv 0$ outside a circular neighbourhood of sufficiently small radius ε_1 of the point $(\bar{x}, \bar{y})$ in which $\Phi(\bar{x}, \bar{y}) \neq 0$, and in this neighbourhood of the point $(\bar{x}, \bar{y})$ the function $\eta(x, y) = k[(x-\bar{x})^2 + (y-\bar{y})^2 - \varepsilon_1^2]^{2n}$ (Fig. 6.6). An analogous lemma holds true for n-fold multiple integrals.

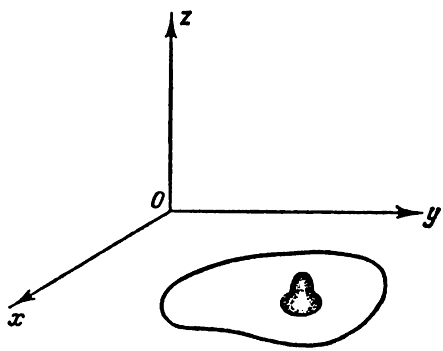


Fig. 6-6

Now let us use the fundamental lemma to simplify the above-obtained condition (6.2) for the extremum of the elementary functional (6.1)

$$\int_{x_0}^{x_1} \left(F_y - \frac{d}{dx} F_{y'}\right) \delta y\,dx = 0. \tag{6.2}$$

All conditions of the lemma are fulfilled: on the extremizing curve the factor $\left(F_y - \frac{d}{dx} F_{y'}\right)$ is a continuous function, and the variation δy is an arbitrary function on which only restrictions of a general nature that are provided for by the fundamental lemma have been imposed; hence, $F_y - \frac{d}{dx} F_{y'} \equiv 0$ on the curve $y = y(x)$ which extremizes the functional under consideration, i.e. $y = y(x)$ is a solution of the second-order differential equation

$$F_y - \frac{d}{dx} F_{y'} = 0,$$

or in expanded form

$$F_y - F_{xy'} - F_{yy'} y' - F_{y'y'} y'' = 0.$$

This equation is called *Euler's equation* (it was first published in 1744). The integral curves of Euler's equation $y = y(x, C_1, C_2)$ are called *extremals*. It is only on extremals that the functional

$$v[y(x)] = \int_{x_0}^{x_1} F(x, y, y')\, dx$$

can be extremized. To find the curve that extremizes the functional (6.1), integrate the Euler equation and determine both arbitrary constants that enter into the general solution of this equation, proceeding from the conditions on the boundary $y(x_0) = y_0$, $y(x_1) = y_1$. Only on extremals that satisfy these conditions can the functional be extremized. However, in order to establish whether indeed an extremum (and whether it is a maximum or a minimum) is achieved on them, one has to take advantage of the sufficient conditions for an extremum given in Chapter 8.

Recall that the boundary-value problem

$$F_y - \frac{d}{dx} F_{y'} = 0, \qquad y(x_0) = y_0, \qquad y(x_1) = y_1$$

does not always have a solution and if the solution exists, it may not be unique (see page 166).

Note that in many variational problems the existence of a solution is obvious from the physical or geometrical meaning of the problem and if the solution of Euler's equation satisfying the boundary conditions is unique, then this unique extremal will be the solution of the given variational problem.

Example 1. On what curves can the functional

$$v[y(x)] = \int_0^{\frac{\pi}{2}} [(y')^2 - y^2]\, dx; \; y(0) = 0, \; y\left(\frac{\pi}{2}\right) = 1$$

be extremized? The Euler equation is of the form $y'' + y = 0$; its general solution is $y = C_1 \cos x + C_2 \sin x$. Utilizing the boundary conditions we get $C_1 = 0$, $C_2 = 1$; hence, only on the curve $y = \sin x$ can an extremum be achieved.

Example 2. On what curves can the functional

$$v[y(x)] = \int_0^1 [(y')^2 + 12xy]\,dx, \quad y(0) = 0, \quad y(1) = 1$$

be extremized? Euler's equation is of the form $y'' - 6x = 0$, whence $y = x^3 + C_1 x + C_2$. Using the boundary conditions, we get $C_1 = 0$, $C_2 = 0$; therefore, an extremum can be achieved only on the curve $y = x^3$.

In these two examples, Euler's equation was readily integrable, but that is by far not always so, since only in exceptional cases can second-order differential equations be integrated in closed form. We consider some elementary cases of the integrability of the Euler equation.

(1) *F is independent of* y':

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

The Euler equation has the form $F_y(x, y) = 0$, since $F_{y'} \equiv 0$. The solution of the finite equation $F_y(x, y) = 0$ thus obtained does not contain any arbitrary elements and therefore, generally speaking, does not satisfy the boundary conditions $y(x_0) = y_0$ and $y(x_1) = y_1$.

Consequently, there does not, generally speaking, exist a solution of this variational problem. Only in exceptional cases when the curve

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

passes through the boundary points (x_0, y_0) and (x_1, y_1) does there exist a curve on which an extremum can be attained.

Example 3.

$$v[y(x)] = \int_{x_0}^{x_1} y^2\,dx; \quad y(x_0) = y_0,$$
$$y(x_1) = y_1.$$

Euler's equation has the form

$$F_y = 0 \text{ or } y = 0.$$

The extremal $y = 0$ passes through the boundary points only for $y_0 = 0$ and $y_1 = 0$ (Fig. 6.7). If $y_0 = 0$ and $y_1 = 0$, then, obviously, the function $y = 0$ minimizes the functional $v = \int_{x_0}^{x_1} y^2\,dx$ since $v[y(x)] \geqslant 0$, and $v = 0$ for $y = 0$. But if at least one of the y_0 and y_1 is not zero, then the functional is not minimized on continuous functions, which is understandable since it is possible to choose a sequence of continuous functions $y_n(x)$, whose graphs consist of an arc of a curve more and more steeply descending from the point (x_0, y_0) to the axis of abscissas, then of a segment of

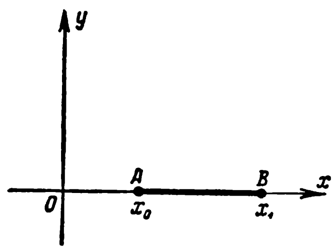


Fig. 6-7

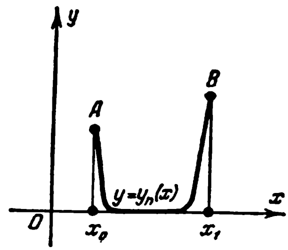


Fig. 6-8

the axis of abscissas that almost coincides with the entire segment (x_0, x_1) and, finally, of the arc of the curve which near the point x_1 rises steeply to the point (x_1, y_1) (Fig. 6.8). It is obvious that on the curves of such a sequence the values of the functional differ from zero by arbitrarily small values and hence the lower bound of the values of the functional is zero; this lower bound, however, cannot be attained on a continuous curve, since for any continuous curve $y=y(x)$ different from an identical zero, the integral $\int_{x_0}^{x_1} y^2\,dx>0$. This lower bound of values of the functional is attained on the discontinuous function (Fig. 6.9)

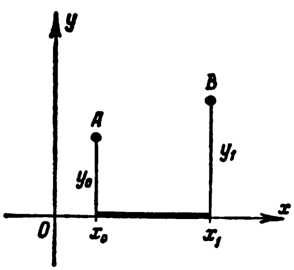


Fig. 6-9

$$y(x_0)=y_0,$$
$$y(x)=0 \text{ for } x_0<x<x_1,$$
$$y(x_1)=y_1.$$

(2) *The function F is linearly dependent on y′*:

$$F(x, y, y')=M(x, y)+N(x, y)\,y';$$

$$v[y(x)]=\int_{x_0}^{x_1}\left[M(x, y)+N(x, y)\frac{dy}{dx}\right]dx.$$

The Euler equation is of the form

$$\frac{\partial M}{\partial y}+\frac{\partial N}{\partial y}y'-\frac{d}{dx}N(x, y)=0,$$

or

$$\frac{\partial M}{\partial y}+\frac{\partial N}{\partial y}y'-\frac{\partial N}{\partial x}-\frac{\partial N}{\partial y}y'=0,$$

or

$$\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}=0;$$

but again, as in the preceding case, this is a finite, not a differential, equation. Generally speaking, the curve $\frac{\partial M}{\partial y}-\frac{\partial N}{\partial y}=0$ does not satisfy the boundary conditions; consequently, the variational problem does not, as a rule, have any solutions in the class of continuous functions. But if $\frac{\partial M}{\partial y}-\frac{\partial N}{\partial y}\equiv 0$, then the expression $M\,dx+N\,dy$ is an exact differential and

$$v=\int_{x_0}^{x_1}\left(M+N\frac{dy}{dx}\right)dx=\int_{x_0}^{x_1}(M\,dx+N\,dy)$$

is independent of the integration path, the value of the functional v is constant on admissible curves. The variational problem becomes meaningless.

Example 4.

$$v[y(x)]=\int_0^1 (y^2+x^2y')\,dx;\quad y(0)=0,\quad y(1)=a.$$

Euler's equation is of the form $\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}=0$ or $y-x=0$. The first boundary condition $y(0)=0$ is satisfied, but the second boundary condition is satisfied only for $a=1$. But if $a\neq 1$, then there is no extremal that satisfies the boundary conditions.

Example 5.

$$v[y(x)]=\int_{x_0}^{x_1}(y+xy')\,dx \text{ or } v[y(x)]=\int_{x_0}^{x_1}(y\,dx+x\,dy);$$

$$y(x_0)=y_0,\quad y(x_1)=y_1.$$

Euler's equation reduces to the identity $1\equiv 1$. The integrand is an exact differential and the integral does not depend on the path of integration:

$$v[y(x)]=\int_{x_0}^{x_1}d(xy)=x_1y_1-x_0y_0,$$

no matter which curve we integrate along. The variational problem is meaningless.

(3) *F is dependent solely on* y':

$$F = F(y').$$

Euler's equation is of the form $F_{y'y'}y'' = 0$, since $F_y = F_{xy'} = F_{yy'} = 0$. Whence $y'' = 0$ or $F_{y'y'} = 0$. If $y'' = 0$, then $y = C_1x + C_2$ is a two-parameter family of straight lines. But if the equation $F_{y'y'}(y') = 0$ has one or several real roots $y' = k_i$, then $y = k_ix + C$ and we get a one-parameter family of straight lines contained in the above-obtained two-parameter family $y = C_1x + C_2$. Thus, in the case of $F = F(y')$, the extremals are all possible straight lines $y = C_1x + C_2$.

Example 6. The arc length of the curve

$$l[y(x)] = \int_{x_0}^{x_1} \sqrt{1 + y'^2}\, dx$$

has for extremals the straight lines $y = C_1x + C_2$.

Example 7. The time $t[y(x)]$ spent on displacement along a certain curve $y = y(x)$ from the point $A(x_0, y_0)$ to the point $B(x_1, y_1)$ if the velocity $\frac{ds}{dt} = v(y')$ is dependent solely on y', is a functional of the form

$$t[y(x)] = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{v(y')}\, dx$$

$$\left(\frac{ds}{dt} = v(y'); \quad dt = \frac{ds}{v(y')} = \frac{\sqrt{1 + y'^2}\, dx}{v(y')}; \qquad t = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{v(y')}\, dx\right).$$

Hence the extremals of this functional are straight lines.

(4) *F is dependent solely on x and* y':

$$F = F(x, y').$$

Euler's equation takes the form $\frac{d}{dx}F_{y'}(x, y') = 0$ and, hence, has a first integral $F_{y'}(x, y') = C_1$; and since the first-order equation $F_y(x, y') = C_1$ thus obtained does not contain y, the equation may be integrated either by direct solution for y' and integration, or by means of introducing a properly chosen parameter (see p. 75).

Example 8. The functional

$$t[y(x)] = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{x}\, dx$$

(t is the time spent on translation along the curve $y = y(x)$ from one point to another if the rate of motion $v=x$, since if $\frac{ds}{dt}=x$, then $dt=\frac{ds}{x}$ and $t=\int_{x_0}^{x_1}\frac{\sqrt{1+y'^2}}{x}dx\Big)$. The first integral of the Fuler equation $F_{y'}=C_1$ is of the form $\frac{y'}{x\sqrt{1+y'^2}}=C_1$. This equation is most readily integrated if one introduces a parameter putting $y'=\tan t$; then

$$x=\frac{1}{C_1}\frac{y'}{\sqrt{1+y'^2}}=\frac{1}{C_1}\sin t$$

or $x=\overline{C_1}\sin t$, where $\overline{C_1}=\frac{1}{C_1}$;

$\frac{dy}{dx}=\tan t$; $dy=\tan t\,dx=\tan t.\ \overline{C_1}\cos t\,dt=\overline{C_1}\sin t\,dt$;

integrating, we get $y=-\overline{C_1}\cos t+C_2$.

Thus, $x=\overline{C}_1\sin t,\ y-C_2=-\overline{C}_1\cos t$

or, eliminating t, we have $x^2+(y-C_2)^2=\overline{C}_1^2$, which is a family of circles with centres on the axis of ordinates.

(5) *F is dependent on y and y′ alone*:

$$F=F(y,y').$$

Euler's equation is of the form: $F_y-F_{yy'}y'-F_{y'y'}y''=0$, since $F_{xy'}=0$. If we multiply this equation termwise by y', then, as is readily verifiable, the left-hand side becomes an exact derivative $\frac{d}{dx}(F-y'F_{y'})$.

Indeed,

$$\frac{d}{dx}(F-y'F_{y'})=F_yy'+F_{y'}y''-y''F_{y'}-F_{yy'}y'^2-F_{y'y'}y'y''=$$
$$=y'(F_y-F_{yy'}y'-F_{y'y'}y'').$$

Hence, Euler's equation has the first integral

$$F-y'F_{y'}=C_1,$$

and since this first-order equation does not contain x explicitly, it may be integrated by solving for y' and separating the variables, or by introducing a parameter.

Example 9. The minimum-surface-of-revolution problem: find a curve with specified boundary points whose rotation about the axis of abscissas generates a surface of minimum area (Fig. 6.10).

As we know, the area of a surface of revolution is

$$S[y(x)]=2\pi\int_{x_0}^{x_1}y\sqrt{1+y'^2}\,dx.$$

The integrand is dependent solely on y and y' and, hence, a first integral of Euler's equation will have the form

$$F-y'F_{y'}=C_1$$

or in the given case

$$y\sqrt{1+y'^2}-\frac{yy'^2}{\sqrt{1+y'^2}}=C_1.$$

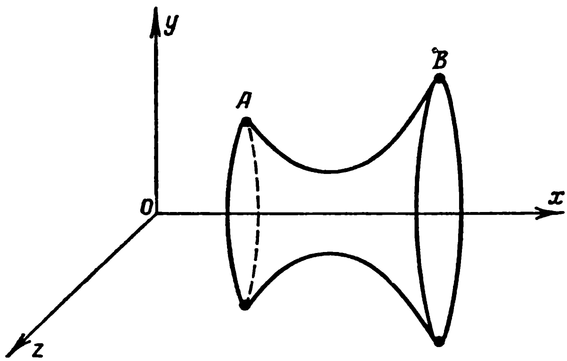


Fig. 6-10

After simplification we have $\frac{y}{\sqrt{1+y'^2}}=C_1$. The simplest way to integrate this equation is by the substitution $y'=\sinh t$, then $y=C_1\cosh t$, and

$$dx=\frac{dy}{y'}=\frac{C_1 \sinh t\, dt}{\sinh t}=C_1\,dt;\quad x=C_1t+C_2.$$

And so the desired surface is formed by revolution of a line, the equation of which, in parametric form, is

$$x=C_1t+C_2,$$
$$y=C_1\cosh t.$$

Eliminating the parameter t, we get $y=C_1\cosh\frac{x-C_2}{C_1}$, a family of catenaries, the revolution of which forms surfaces called catenoids. The constants C_1 and C_2 are found from the condition of the passage of the desired line through given boundary points (depending on the position of the points A and B, there may be one, two or zero solutions).

Example 10. The problem of the brachistochrone (see page 294): find the curve connecting given points A and B which is traversed by a particle sliding from A to B in the shortest time (friction and resistance of the medium are ignored).

Put the coordinate origin at A, make the x-axis horizontal and the y-axis vertical. The speed of the particle is $\frac{ds}{dt}=\sqrt{2gy}$, whence we find the time spent in moving from $A(0, 0)$ to $B(x_1, y_1)$:

$$t[y(x)]=\frac{1}{\sqrt{2g}}\int_0^{x_1}\frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx; \qquad y(0)=0, \qquad y(x_1)=y_1.$$

Although in this case the integral is not proper, it is easy to establish that here as well we can take advantage of the preceding theory. Since this functional also belongs to the most elementary type and its integrand does not contain x explicitly, the Euler equation has a first integral $F-y'F_{y'}=C$ or in the given case

$$\frac{\sqrt{1+y'^2}}{\sqrt{y}}-\frac{y'^2}{\sqrt{y(1+y'^2)}}=C,$$

whence, after simplification, we get $\frac{1}{\sqrt{y(1+y'^2)}}=C$ or $y(1+y'^2)=C_1$. Introduce the parameter t, putting $y'=\cot t$; then we have

$$y=\frac{C_1}{1+\cot^2 t}=C_1\sin^2 t=\frac{C_1}{2}(1-\cos 2t);$$

$$dx=\frac{dy}{y'}=\frac{2C_1\sin t\cos t\,dt}{\cot t}=2C_1\sin^2 t\,dt=C_1(1-\cos 2t)\,dt;$$

$$x=C_1\left(t-\frac{\sin 2t}{2}\right)+C_2=\frac{C_1}{2}(2t-\sin 2t)+C_2.$$

Consequently, in parametric form the equation of the desired line is

$$x-C_2=\frac{C_1}{2}(2t-\sin 2t), \qquad y=\frac{C_1}{2}(1-\cos 2t).$$

If the parameter is transformed by the substitution $2t=t_1$ and if we take into account that $C_2=0$, since $x=0$ for $y=0$, then we get the equation of a family of cycloids in the ordinary form:

$$x=\frac{C_1}{2}(t_1-\sin t_1),$$

$$y=\frac{C_1}{2}(1-\cos t_1),$$

where $\frac{C_1}{2}$ is the radius of a rolling circle, which is found from the condition of passage of the cycloid through the point $B(x_1, y_1)$. Thus, the brachistochrone is a cycloid.

3. Functionals of the Form

$$\int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx$$

In order to obtain the necessary conditions for the extremum of a functional v of a more general type

$$v[y_1, y_2, \ldots, y_n] = \int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx$$

for the given boundary conditions of all functions

$$y_1(x_0) = y_{10}, \quad y_2(x_0) = y_{20}, \ldots, y_n(x_0) = y_{n0},$$
$$y_1(x_1) = y_{11}, \quad y_2(x_1) = y_{21}, \ldots, y_n(x_1) = y_{n1},$$

we shall vary only one of the functions

$$y_j(x) \quad (j = 1, 2, \ldots, n),$$

holding the other functions unchanged. Then the functional $v[y_1, y_2, \ldots, y_n]$ will reduce to a functional dependent only on a single varied function, for example, on $y_i(x)$,

$$v[y_1, y_2, \ldots, y_n] = \tilde{v}[y_i]$$

of the form considered in Sec. 2, and, hence, the extremizing function must satisfy Euler's equation

$$F_{y_i} - \frac{d}{dx} F_{y_i'} = 0.$$

Since this argument is applicable to any function y_i $(i = 1, 2, \ldots, n)$, we get a system of second-order differential equations

$$F_{y_i} - \frac{d}{dx} F_{y_i'} = 0 \qquad (i = 1, 2, \ldots, n),$$

which, generally speaking, define a $2n$-parameter family of integral curves in the space $x, y_1, y_2, \ldots, y_n$—which is the family of extremals of the given variational problem.

If, for example, the functional depends only on two functions $y(x)$ and $z(x)$:

$$v[y(x), \quad z(x)] = \int_{x_0}^{x_1} F(x, y, z, y', z')\,dx;$$

$$y(x_0) = y_0, \qquad z(x_0) = z_0, \qquad y(x_1) = y_1, \qquad z(x_1) = z_1,$$

that is to say, it is defined by the choice of space curve $y = y(x)$, $z = z(x)$ (Fig. 6.11), then by varying $y(x)$ alone and holding $z(x)$

constant we can change our curve so that its projection on the xz-plane does not change, i.e. the curve all the time remains on the projecting cylinder $z=z(x)$ (Fig. 6.12).

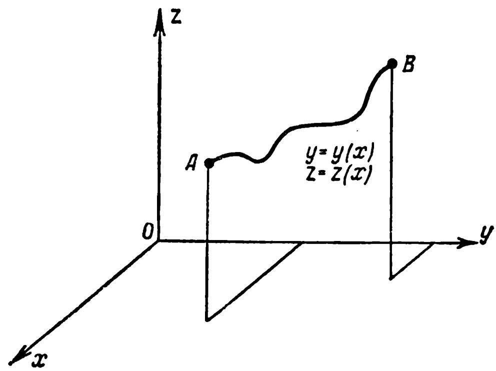
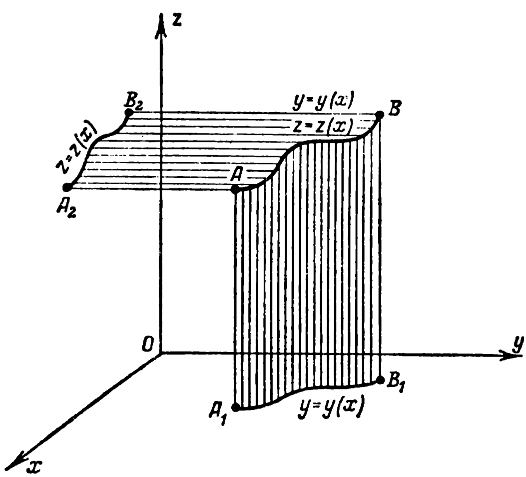


Fig. 6-11



· Fig. 6-12

Similarly, by fixing $y(x)$ and varying $z(x)$, we vary the curve so that all the time it lies on the projecting cylinder $y=y(x)$. We then obtain a system of two Euler's equations:

$$F_y - \frac{d}{dx} F_{y'} = 0 \quad \text{and} \quad F_z - \frac{d}{dx} F_{z'} = 0.$$

Example 1. Find the extremals of the functional

$$v[y(x), z(x)] = \int_0^{\frac{\pi}{2}} [y'^2 + z'^2 + 2yz]\, dx, \qquad y(0)=0, \qquad y\left(\frac{\pi}{2}\right)=1,$$

$$z(0)=0, \qquad z\left(\frac{\pi}{2}\right) = -1.$$

The system of Euler's differential equations is of the form

$$y'' - z = 0,$$
$$z'' - y = 0.$$

Eliminating one of the unknown functions, say z, we get $y^{IV} - y = 0$. Integrating this linear equation with constant coefficients, we obtain

$$y = C_1e^x + C_2e^{-x} + C_3 \cos x + C_4 \sin x;$$
$$z = y''; \quad z = C_1e^x + C_2e^{-x} - C_3 \cos x - C_4 \sin x.$$

Using the boundary conditions, we find

$$C_1 = 0, \quad C_2 = 0, \quad C_3 = 0, \quad C_4 = 1;$$

hence, $y = \sin x$, $z = -\sin x$.

Example 2. Find the extremals of the functional

$$v[y(x), z(x)] = \int_{x_0}^{x_1} F(y', z')\, dx.$$

The system of Euler's equations is of the form

$$F_{y'y'}y'' + F_{y'z'}z'' = 0; \quad F_{y'z'}y'' + F_{z'z'}z'' = 0,$$

whence, assuming $F_{y'y'}F_{z'z'} - (F_{y'z'})^2 \neq 0$, we get $y'' = 0$ and $z'' = 0$ or $y = C_1x + C_2$, $z = C_3x + C_4$ are a family of straight lines in space.

Example 3. Find the differential equations of the lines of propagation of light in an optically nonhomogeneous medium in which the speed of light is $v(x, y, z)$.

According to Fermat's principle, light is propagated from one point $A(x_0, y_0)$ to another $B(x_1, y_1)$ along a curve for which the time T of passage of light will be least. If the equation of the desired curve $y = y(x)$ and $z = z(x)$, then

$$T = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2 + z'^2}}{v(x, y, z)}\, dx.$$

For this functional, the system of Euler's equations

$$\frac{\partial v}{\partial y}\frac{\sqrt{1+y'^2+z'^2}}{v^2}+\frac{d}{dx}\frac{y'}{v\sqrt{1+y'^2+z'^2}}=0,$$

$$\frac{\partial v}{\partial z}\frac{\sqrt{1+y'^2+z'^2}}{v^2}+\frac{d}{dx}\frac{z'}{v\sqrt{1+y'^2+z'^2}}=0$$

will be a system that defines the lines of light propagation.

4. Functionals Dependent on Higher-Order Derivatives

Let us investigate the extreme value of the functional

$$v[y(x)]=\int_{x_0}^{x_1} F(x,\ y(x),\ y'(x),\ \ldots,\ y^{(n)}(x))\,dx,$$

where we consider the function F differentiable $n+2$ times with respect to all arguments and we assume that the boundary conditions are of the form

$$y(x_0)=y_0,\ y'(x_0)=y_0',\ \ldots,\ y^{(n-1)}(x_0)=y_0^{(n-1)};$$
$$y(x_1)=y_1,\ y'(x_1)=y_1',\ \ldots,\ y^{(n-1)}(x_1)=y_1^{(n-1)},$$

i.e. at the boundary points are given the values not only of the function but also of its derivatives up to the order $n-1$ inclusive. Suppose that an extremum is attained on the curve $y=y(x)$, which is $2n$ times differentiable, and let $y=\bar{y}(x)$ be the equation of some comparison curve, which is also $2n$ times differentiable.

Consider the one-parameter family of functions

$$y(x,\ \alpha)=y(x)+\alpha[\bar{y}(x)-y(x)] \text{ or } y(x,\ \alpha)=y(x)+\alpha\delta y.$$

For $\alpha=0$, $y(x,\ \alpha)=y(x)$ and for $\alpha=1$, $y(x,\ \alpha)=\bar{y}(x)$. If one considers the value of the functional $v[y(x)]$ only on curves of the family $y=y(x,\ \alpha)$, then the functional reduces to a function of the parameter α, which is extremized for $\alpha=0$; hence, $\frac{d}{d\alpha}v[y(x,\alpha)]\Big|_{\alpha=0}=0$. According to Sec. 1, this derivative is called the *variation of the functional* v and is symbolized by δv:

$$\delta v=\left[\frac{d}{d\alpha}\int_{x_0}^{x_1} F(x,\ y(x,\ \alpha),\ y'(x,\ \alpha),\ \ldots,\ y^{(n)}(x,\ \alpha))\,dx\right]_{\alpha=0}=$$
$$=\int_{x_0}^{x_1}(F_y\,\delta y+F_{y'}\,\delta y'+F_{y''}\,\delta y''+\ldots+F_{y^{(n)}}\,\delta y^{(n)})\,dx.$$

Integrate the second summand on the right once term-by-term

$$\int_{x_0}^{x_1} F_{y'}\,\delta y'\,dx = [F_{y'}\,\delta y]_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} F_{y'}\,\delta y\,dx,$$

the third summand twice:

$$\int_{x_0}^{x_1} F_{y''}\,\delta y''\,dx = [F_{y''}\,\delta y']_{x_0}^{x_1} - \left[\frac{d}{dx} F_{y''}\,\delta y\right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \frac{d^2}{dx^2} F_{y''}\,\delta y\,dx,$$

and so forth; the last summand, n times:

$$\int_{x_0}^{x_1} F_{y^{(n)}}\,\delta y^{(n)}\,dx = [F_{y^{(n)}}\,\delta y^{(n-1)}]_{x_0}^{x_1} - \left[\frac{d}{dx} F_{y^{(n)}}\,\delta y^{(n-2)}\right]_{x_0}^{x_1} + \ldots$$

$$\ldots + (-1)^n \int_{x_0}^{x_1} \frac{d^n}{dx^n} F_{y^{(n)}}\,\delta y\,dx.$$

Taking into account the boundary conditions, by virtue of which for $x = x_0$ and for $x = x_1$, the variations $\delta y = \delta y' = \delta y'' = \ldots = \delta y^{(n-1)} = 0$, we finally get

$$\delta v = \int_{x_0}^{x_1} \left(F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} + \ldots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}}\right) \delta y\,dx.$$

Since on the extremizing curve we have

$$\delta v = \int_{x_0}^{x_1} \left(F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} + \ldots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}}\right) \delta y\,dx = 0$$

for an arbitrary choice of the function δy and since the first factor under the integral sign is a continuous function of x on the same curve $y = y(x)$, it follows that by virtue of the fundamental lemma the first factor is identically zero:

$$F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} + \ldots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}} \equiv 0.$$

Thus, the function $y = y(x)$, which extremizes the functional

$$v[y(x)] = \int_{x_0}^{x_1} F(x,\ y,\ y',\ y'',\ \ldots,\ y^{(n)})\,dx,$$

must be a solution of the equation

$$F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} + \ldots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}} = 0.$$

This differential equation of order $2n$ is called the *Euler-Poisson equation*, and its integral curves are termed *extremals* of the variational problem under consideration. The general solution of this equation contains $2n$ arbitrary constants, which, generally speaking, may be determined from the $2n$ boundary conditions:

$$y(x_0)=y_0,\ y'(x_0)=y'_0,\ \ldots,\ y^{(n-1)}(x_0)=y_0^{(n-1)};$$
$$y(x_1)=y_1,\ y'(x_1)=y'_1,\ \ldots,\ y^{(n-1)}(x_1)=y_1^{(n-1)}.$$

Example 1. Find the extremal of the functional

$$v[y(x)]=\int_0^1 (1+y''^2)\,dx;$$
$$y(0)=0,\ y'(0)=1,\ y(1)=1,\ y'(1)=1.$$

The Euler-Poisson equation is of the form $\frac{d^2}{dx^2}(2y'')=0$ or $y^{\mathrm{IV}}=0$; its general solution is $y=C_1x^3+C_2x^2+C_3x+C_4$. Using the boundary conditions, we get

$$C_1=0,\ C_2=0,\ C_3=1,\ C_4=0.$$

And so the extremum can be attained only on the straight line $y=x$.

Example 2. Determine the extremal of the functional

$$v[y(x)]=\int_0^{\frac{\pi}{2}} (y''^2-y^2+x^2)\,dx,$$

that satisfies the conditions

$$y(0)=1,\ y'(0)=0,\ y\left(\frac{\pi}{2}\right)=0,\ y'\left(\frac{\pi}{2}\right)=-1.$$

The Euler-Poisson equation is of the form $y^{\mathrm{IV}}-y=0$; its general solution is $y=C_1e^x+C_2e^{-x}+C_3\cos x+C_4\sin x$. Using the boundary conditions, we get $C_1=0$, $C_2=0$, $C_3=1$, $C_4=0$. And so the extremum can be achieved only on the curve $y=\cos x$.

Example 3. Determine the extremal of the functional

$$v[y(x)]=\int_{-l}^{l}\left(\frac{1}{2}\mu y''^2+\rho y\right)dx,$$

that satisfies the boundary conditions

$$y(-l)=0,\qquad y'(-l)=0,\qquad y(l)=0,\qquad y'(l)=0.$$

This is the variational problem to which is reduced the problem of finding the axis of a flexible bent cylindrical beam fixed at the ends. If the beam is homogeneous, then ρ and μ are constants and the Euler-Poisson equation has the form

$$\rho+\frac{d^2}{dx^2}(\mu y'')=0 \text{ or } y^{\text{IV}}=-\frac{\rho}{\mu},$$

whence

$$y=-\frac{\rho x^4}{24\mu}+C_1x^3+C_2x^2+C_3x+C_4.$$

Using the boundary conditions, we finally get

$$y=-\frac{\rho}{24\mu}(x^4-2l^2x^2+l^4) \text{ or } y=-\frac{\rho}{24\mu}(x^2-l^2)^2.$$

If the functional v is of the form

$$v[y(x),\ z(x)]=\int_{x_0}^{x_1} F(x,\ y,\ y',\ \ldots,\ y^{(n)},\ z,\ z'\ \ldots,\ z^{(m)})\,dx,$$

then by varying only $y(x)$ and assuming $z(x)$ to be fixed, we find that the extremizing functions $y(x)$ and $z(x)$ must satisfy the Euler-Poisson equation

$$F_y-\frac{d}{dx}F_{y'}+\ldots+(-1)^n\frac{d^n}{dx^n}F_{y^{(n)}}=0,$$

whereas by varying $z(x)$ and holding $y(x)$ fixed we find that the very same functions must satisfy the equation

$$F_z-\frac{d}{dx}F_{z'}+\ldots+(-1)^m\frac{d^m}{dx^m}F_{z^{(m)}}=0.$$

Thus, the functions $z(x)$ and $y(x)$ must satisfy a system of two equations:

$$F_y-\frac{d}{dx}F_{y'}+\ldots+(-1)^n\frac{d^n}{dx^n}F_{y^{(n)}}=0,$$

$$F_z-\frac{d}{dx}F_{z'}+\ldots+(-1)^m\frac{d^m}{dx^m}F_{z^{(m)}}=0.$$

We can argue in the same fashion when investigating for the extremum of a functional dependent on any number of functions:

$$v[y_1,\ y_2,\ \ldots,\ y_n]=$$
$$=\int_{x_0}^{x_1} F(x,\ y_1,\ y_1',\ \ldots,\ y_1^{(n_1)},\ y_2,\ y_2',\ \ldots,\ y_2^{(n_2)},\ \ldots$$
$$\ldots,\ y_m,\ y_m',\ \ldots,\ y_m^{(n_m)})\,dx.$$

Varying some one function $y_i(x)$ and holding the others fixed, we get the basic necessary condition for an extremum in the form

$$F_{y_i}-\frac{d}{dx}F_{y_i'}+\ldots+(-1)^{n_i}\frac{d^{n_i}}{dx^{n_i}}F_{y_i^{(n_i)}}=0 \qquad (i=1,\ 2,\ \ldots,\ m).$$

5. Functionals Dependent on the Functions of Several Independent Variables

Let us investigate the following functional for an extremum:

$$v[z(x, y)] = \iint_D F\left(x, y, z, \frac{dz}{dx}, \frac{dz}{dy}\right) dx\, dy;$$

the values of the function $z(x, y)$ are given on the boundary C of domain D, that is, a spatial path (or contour) $\tilde{C}$ is given, through

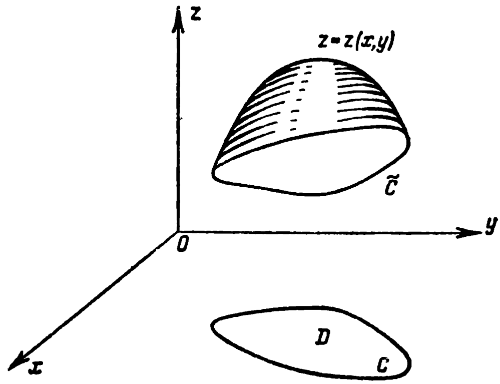


Fig. 6-13

which all permissible surfaces have to pass (Fig. 6.13). To abbreviate notation, put $\frac{\partial z}{\partial x} = p$, $\frac{\partial z}{\partial y} = q$. We will consider the function F as three times differentiable. We assume the extremizing surface $z = z(x, y)$ to be twice differentiable.

Let us consider a one-parameter family of surfaces $z = z(x, y, \alpha) = z(x, y) + \alpha \delta x$, where $\delta z = \bar{z}(x, y) - z(x, y)$, including for $\alpha = 0$ the surface $z = z(x, y)$ on which the extremum is achieved, and for $\alpha = 1$, a certain permissible surface $z = z(x, y)$. On functions of the family $z(x, y, \alpha)$, the functional v reduces to the function α, which has to have an extremum for $\alpha = 0$; consequently, $\frac{\partial}{\partial \alpha} v[z(x, y, \alpha)]|_{\alpha=0} = 0$. If, in accordance with Sec. 1, we call the derivative of $v[z(x, y, \alpha)]$ with respect to α, for $\alpha = 0$, the *variation of the functional* and symbolize it by δv, we will have

$$\delta v = \left\{\frac{\partial}{\partial \alpha} \iint_D F(x, y, z(x, y, \alpha), p(x, y, \alpha), q(x, y, \alpha))\, dx\, dy\right\}_{\alpha=0} = \iint_D [F_z \delta z + F_p \delta p + F_q \delta q]\, dx\, dy,$$

where

$$z(x, y, \alpha) = z(x, y) + \alpha\,\delta z,$$
$$p(x, y, \alpha) = \frac{\partial z(x, y, \alpha)}{\partial x} = p(x, y) + \alpha\,\delta p,$$
$$q(x, y, \alpha) = \frac{\partial z(x, y, \alpha)}{\partial y} = q(x, y) + \alpha\,\delta q.$$

Since

$$\frac{\partial}{\partial x}\{F_p\,\delta z\} = \frac{\partial}{\partial x}\{F_p\}\,\delta z + F_p\,\delta p,$$
$$\frac{\partial}{\partial y}\{F_q\,\delta z\} = \frac{\partial}{\partial y}\{F_q\}\,\delta z + F_q\,\delta q,$$

it follows that

$$\iint_D (F_p\,\delta p + F_q\,\delta q)\,dx\,dy =$$
$$= \iint_D \left[\frac{\partial}{\partial x}\{F_p\,\delta z\} + \frac{\partial}{\partial y}\{F_q\,\delta z\}\right] dx\,dy -$$
$$- \iint_D \left[\frac{\partial}{\partial x}\{F_p\} + \frac{\partial}{\partial y}\{F_q\}\right]\delta z\,dx\,dy,$$

where $\frac{\partial}{\partial x}\{F_p\}$ is the so-called total partial derivative with respect to x. When calculating it, y is assumed to be fixed, but the dependence of z, p and q upon x is taken into account:

$$\frac{\partial}{\partial x}\{F_p\} = F_{px} + F_{pz}\frac{\partial z}{\partial x} + F_{pp}\frac{\partial p}{\partial x} + F_{pq}\frac{\partial q}{\partial x}$$

and similarly

$$\frac{\partial}{\partial y}\{F_q\} = F_{qy} + F_{qz}\frac{\partial z}{\partial x} + F_{qp}\frac{\partial p}{\partial y} + F_{qq}\frac{\partial q}{\partial y}.$$

Using the familiar Green's function

$$\iint_D \left(\frac{\partial N}{\partial x} + \frac{\partial M}{\partial y}\right) dx\,dy = \int_C (N\,dy - M\,dx)$$

we get

$$\iint_D \left[\frac{\partial}{\partial x}\{F_p\,\delta z\} + \frac{\partial}{\partial y}\{F_q\,\delta z\}\right] dx\,dy = \int_C (F_p\,dy - F_q\,dx)\,\delta z = 0.$$

The last integral is equal to zero, since on the contour C the variation $\delta z = 0$ because all permissible surfaces pass through one and the same spatial contour $\tilde{C}$. Consequently,

$$\iint_D [F_p \delta p + F_q \delta q]\, dx\, dy = -\iint_D \left[\frac{\partial}{\partial x}\{F_p\} + \frac{\partial}{\partial y}\{F_q\}\right] \delta z\, dx\, dy,$$

and the necessary condition for an extremum,

$$\iint_D (F_z \delta z + F_p \delta p + F_q \delta q)\, dx\, dy = 0,$$

takes the form

$$\iint_D \left(F_z - \frac{\partial}{\partial x}\{F_p\} - \frac{\partial}{\partial y}\{F_q\}\right) \delta z\, dx\, dy = 0.$$

Since the variation δz is arbitrary (only restrictions of a general nature are imposed on δz that have to do with continuity and differentiability, vanishing on the contour C, etc.) and the first factor is continuous, it follows from the fundamental lemma (page 308) that on the extremizing surface $z = z(x, y)$

$$F_z - \frac{\partial}{\partial x}\{F_p\} - \frac{\partial}{\partial y}\{F_q\} \equiv 0.$$

Consequently, $z(x, y)$ is a solution of the equation

$$F_z - \frac{\partial}{\partial x}\{F_p\} - \frac{\partial}{\partial y}\{F_q\} = 0.$$

This second-order partial differential equation that must be satisfied by the extremizing function $z(x, y)$ is called the *Ostrogradsky equation* after the celebrated Russian mathematician M. Ostrogradsky who in 1834 first obtained the equation (for rectangular domains D it had already appeared in the works of Euler).

Example 1.

$$v[z(x, y)] = \iint_D \left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2\right] dx dy,$$

the values of the function z are given on the boundary C of the domain D: $z = f(x, y)$. Here the Ostrogradsky equation is of the form

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$

or, in abbreviated notation,

$$\Delta z = 0,$$

which is the familiar *Laplace equation*; we have to find a solution, continuous in D, of this equation that takes on specified values on the boundary of the domain D. This is one of the basic problems of mathematical physics, called the *Dirichlet problem*.

Example 2.

$$v[z(x, y)] = \iint_D \left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 2zf(x, y)\right] dx\,dy,$$

a function z is given on the boundary of a domain D. Here, the Ostrogradsky equation is of the form

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = f(x, y),$$

or, in abbreviated notation,

$$\Delta z = f(x, y).$$

This equation is called *Poisson's equation* and is also frequently encountered in problems of mathematical physics.

Example 3. The problem of finding a surface of minimal area stretched over a given contour C reduces to investigating the functional

$$S[z(x, y)] = \iint_D \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy$$

for a minimum. Here the Ostrogradsky equation has the form

$$\frac{\partial}{\partial x}\left\{\frac{p}{\sqrt{1+p^2+q^2}}\right\} + \frac{\partial}{\partial y}\left\{\frac{q}{\sqrt{1+p^2+q^2}}\right\} = 0$$

or

$$\frac{\partial^2 z}{\partial x^2}\left[1 + \left(\frac{\partial z}{\partial y}\right)^2\right] - 2\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}\frac{\partial^2 z}{\partial x\,\partial y} + \frac{\partial^2 z}{\partial y^2}\left[1 + \left(\frac{\partial z}{dx}\right)^2\right] = 0,$$

i.e. at every point the mean curvature is zero. It is known that soap bubbles stretched on a given contour C are a physical realization of minimal surfaces.

For the functional

$$v[z(x_1, x_2, \ldots, x_n)] =$$
$$= \iint_D \ldots \int F(x_1, x_2, \ldots, x_n, z, p_1, p_2, \ldots, p_n)\, dx_1\, dx_2 \ldots dx_n,$$

where $p_i = \frac{\partial z}{\partial x_i}$, in quite the same fashion we obtain, from the basic necessary condition for an extremum, $\delta v = 0$, the following Ostrogradsky equation:

$$F_z - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\{F_{p_i}\} = 0,$$

which the function

$$z = z(x_1, x_2, \dots x_n)$$

extremizing the functional v must satisfy.

For example, for the functional

$$v = \iiint\limits_D \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2\right] dx\, dy\, dz$$

the Ostrogradsky equation is of the form

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

If the integrand of the functional v depends on derivatives of higher order, then, by applying several times the transformations used in deriving the Ostrogradsky equation, we find, as the necessary condition for an extremum, that the extremizing function has to satisfy an equation similar to the Euler-Poisson equation (pages 322-333).

For example, for the functional

$$v[z(x, y)] = \iint\limits_D F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial x \partial y}, \frac{\partial^2 z}{\partial y^2}\right) dx\, dy$$

we get the equation

$$F_z - \frac{\partial}{\partial x}\{F_p\} - \frac{\partial}{\partial y}\{F_q\} + \frac{\partial^2}{\partial x^2}\{F_r\} + \frac{\partial^2}{\partial x\, \partial y}\{F_s\} + \frac{\partial^2}{\partial y^2}\{F_t\} = 0,$$

where

$$p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y}, \quad r = \frac{\partial^2 z}{\partial x^2}, \quad s = \frac{\partial^2 z}{\partial x\, \partial y}, \quad t = \frac{\partial^2 z}{\partial y^2}.$$

This fourth-order partial differential equation must be satisfied by the function extremizing the functional v.

For example, for the functional

$$v = \iint\limits_D \left[\left(\frac{\partial^2 z}{\partial x^2}\right)^2 + \left(\frac{\partial^2 z}{\partial y^2}\right)^2 + 2\left(\frac{\partial^2 z}{\partial x\, \partial y}\right)^2\right] dx\, dy$$

the extremizing function z must satisfy the so-called *biharmonic equation*

$$\frac{\partial^4 z}{\partial x^4} + 2\frac{\partial^4 z}{\partial x^2\, \partial y^2} + \frac{\partial^4 z}{\partial y^4} = 0,$$

which is ordinarily written briefly as $\Delta\Delta z = 0$.

For the functional

$$v = \iint\limits_D \left[\left(\frac{\partial^2 z}{\partial x^2}\right)^2 + \left(\frac{\partial^2 z}{\partial y^2}\right)^2 + 2\left(\frac{\partial^2 z}{\partial x\, \partial y}\right)^2 - 2zf(x, y)\right] dx\, dy$$

the extremizing function $z(x, y)$ must satisfy the equation $\Delta\Delta z = f(x, y)$.

The biharmonic equation is also involved in extremum problems of the functional

$$v = \iint_D \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right)^2 dx\, dy$$

or the more general functional

$$v = \iint_D \left\{ \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right)^2 - 2(1-\mu) \left[\frac{\partial^2 z}{\partial x^2} \frac{\partial^2 z}{\partial y^2} - \left(\frac{\partial^2 z}{\partial x \partial y} \right)^2 \right] \right\} dx\, dy,$$

where μ is the parameter.

6. Variational Problems in Parametric Form

In many variational problems the solution is more conveniently sought in parametric form. For example, in the isoperimetric problem (see page 295) of finding a closed curve of given length l bounding a maximum area S, it is inconvenient to seek the solution in the form $y = y(x)$, since by the very meaning of the problem the function $y(x)$ is ambiguous (Fig. 6.14). Therefore, in this problem it is advisable to seek the solution in parametric form: $x = x(t)$, $y = y(t)$. Hence, in the given case we have to seek the extremum of the functional

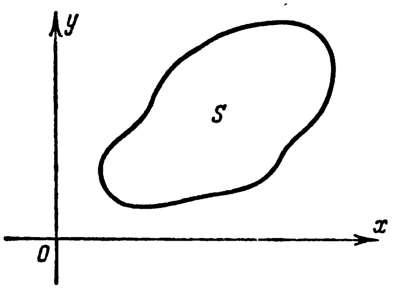


Fig. 6-14

$$S[x(t), y(t)] = \frac{1}{2} \int_0^T (x\dot{y} - y\dot{x})\, dt$$

provided that $l = \int_0^T \sqrt{\dot{x}^2 + \dot{y}^2}\, dt$, where l is a constant.

In the investigation of a certain functional

$$v[y(x)] = \int_{x_0}^{x_1} F(x, y, y')\, dx$$

for an extremum let it be more advisable to seek the solution in the parametric form $x = x(t)$, $y = y(t)$; then the functional will be reduced to the following form:

$$v[x(t), y(t)] = \int_{t_0}^{t_1} F\left(x(t), y(t), \frac{\dot{y}(t)}{\dot{x}(t)} \right) \dot{x}(t)\, dt.$$

Note that after transformation of the variables, the integrand

$$F\left(x(t),\ y(t),\ \frac{\dot{y}(t)}{\dot{x}(t)}\right)\dot{x}(t)$$

does not contain t explicitly and, with respect to the variables $\dot{x}$ and $\dot{y}$, is a homogeneous function of the first degree.

Thus, the functional $v[x(t),\ y(t)]$ is not an arbitrary functional of the form

$$\int_{t_0}^{t_1} \Phi(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t))\,dt$$

that depends on two functions $x(t)$ and $y(t)$, but only an extremely particular case of such a functional, since its integrand does not contain t explicitly and is a homogeneous function of the first degree in the variables $\dot{x}$ and $\dot{y}$.

If we were to go over to any other parametric representation of the desired curve $x=x(\tau)$, $y=y(\tau)$, then the functional $v[x,\ y]$ would be reduced to the form $\int_{\tau_0}^{\tau_1} F\left(x,\ y,\ \frac{\dot{y}_\tau}{\dot{x}_\tau}\right)\dot{x}_\tau\,d\tau$. Hence, the integrand of the functional v does not change its form when the parametric representation of the curve is changed. Thus, the functional v depends on the type of curve and not on its parametric representation.

It is easy to see the truth of the following assertion: if the integrand of the functional

$$v[x(t),\ y(t)] = \int_{t_0}^{t_1} \Phi(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t))\,dt$$

does not contain t explicitly and is a homogeneous function of the first degree in $\dot{x}$ and $\dot{y}$, then the functional $v[x(t),\ y(t)]$ depends solely on the kind of curve $x=x(t)$, $y=y(t)$, and not on its parametric representation. Indeed, let

$$v[x(t),\ y(t)] = \int_{t_0}^{t_1} \Phi(x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t))\,dt,$$

where

$$\Phi(x,\ y,\ k\dot{x},\ k\dot{y}) = k\Phi(x,\ y,\ \dot{x},\ \dot{y}).$$

Let us pass to a new parametric representation putting

$$\tau=\varphi(t)\quad (\dot{\varphi}(t)\neq 0),\quad x=x(\tau),\quad y=y(\tau).$$

Then

$$\int_{t_0}^{t_1} \Phi(x(t), y(t), \dot{x}(t), \dot{y}(t))\,dt = \int_{\tau_0}^{\tau_1} \Phi(x(\tau), y(\tau), \dot{x}_\tau(\tau)\dot{\varphi}(t), \dot{y}_\tau(\tau)\dot{\varphi}(t)) \frac{d\tau}{\dot{\varphi}(t)}.$$

By virtue of the fact that Φ is a homogeneous function of the first degree in $\dot{x}$ and $\dot{y}$, we have

$$\Phi(x, y, \dot{x}_\tau\dot{\varphi}, \dot{y}_\tau\dot{\varphi}) = \dot{\varphi}\Phi(x, y, \dot{x}_\tau, \dot{y}_\tau),$$

whence

$$\int_{t_0}^{t_1} \Phi(x, y, \dot{x}_t, \dot{y}_t)\,dt = \int_{\tau_0}^{\tau_1} \Phi(x, y, \dot{x}_\tau, \dot{y}_\tau)\,d\tau,$$

that is, the integrand has not changed with a change in the parametric representation.

The arc length $\int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2}\,dt$ * and an area bounded by a certain curve $\frac{1}{2}\int_{t_0}^{t_1} (x\dot{y} - y\dot{x})\,dt$ are examples of such functionals.

In order to find the extremals of functionals of this kind,

$$v[x(t), y(t)] = \int_{t_0}^{t_1} \Phi(x, y, \dot{x}, \dot{y})\,dt,$$

where Φ is a homogeneous function of the first degree in $\dot{x}$ and $\dot{y}$, and also for functionals with an arbitrary integrand function $\Phi(t, x, y, \dot{x}, \dot{y})$, one has to solve a system of Euler's equations:

$$\Phi_x - \frac{d}{dt}\Phi_{\dot{x}} = 0; \quad \Phi_y - \frac{d}{dt}\Phi_{\dot{y}} = 0.$$

However, in the case under consideration, these equations are not independent, since they must be satisfied by a certain solution $x = x(t)$, $y = y(t)$ and also by any other pairs of functions that yield a different parametric representation of the same curve, which, in the case of Euler's equations being independent, would lead to a contradiction with the theorem of the existence and uniqueness

* The function $\sqrt{\dot{x}^2 + \dot{y}^2}$ is a *positive* homogeneous function of the first degree; i.e., for it the condition $F(kx, ky) = kF(x, y)$ is satisfied only for positive k. However, this is quite sufficient for the theory described in this section to hold true, since upon changing the variables $\tau = \varphi(t)$ we can assume $\dot{\varphi}(t) > 0$.

of a solution of a system of differential equations. This is an indication that for functionals of the form

$$v[x(t), y(t)] = \int_{t_0}^{t_1} \Phi(x, y, \dot{x}, \dot{y})\, dt,$$

where Φ is a homogeneous function of the first degree in $\dot{x}$ and $\dot{y}$, one of the Euler equations is a consequence of the other. To find the extremals, we have to take one of the Euler equations and integrate it together with the equation defining the choice of parameter. For example, to the equation $\Phi_x - \frac{d}{dt}\Phi_{\dot{x}} = 0$ we can adjoin the equation $\dot{x}^2 + \dot{y}^2 = 1$, which indicates that the arc length of the curve is taken as the parameter.

7. Some Applications

The basic variational principle in mechanics is the principle of least action of Ostrogradsky and Hamilton, which states that from among the possible motions of a system of particles (i. e. those consistent with constraints) that motion is accomplished which gives a stationary value (i.e. a value corresponding to an argument for which the variation of the function is zero) to the integral

$$\int_{t_0}^{t_1} (T - U)\, dt,$$

where T is the kinetic and U the potential energy of the system.

Let us apply this principle to a few problems in mechanics.

Example 1. Given a system of particles with masses m_i ($i = 1, 2, \ldots, n$) and coordinates (x_i, y_i, z_i) acted upon by forces $\bar{F}_i$ that possess the force function (potential) $-U$, which is dependent solely on the coordinates:

$$F_{ix} = -\frac{\partial U}{\partial x_i}; \qquad F_{iy} = -\frac{\partial U}{\partial y_i}; \qquad F_{iz} = -\frac{\partial U}{\partial z_i},$$

where F_{ix}, F_{iy}, F_{iz} are the coordinates of the vector $\bar{F}_i$ acting on the point (x_i, y_i, z_i). Find the differential equations of motion of the system. In this case the kinetic energy

$$T = \frac{1}{2} \sum_{i=1}^{n} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2),$$

and the potential energy of the system is equal to U. The system of Euler's equations for the integral

$$\int_{t_0}^{t_1} (T-U)\,dt$$

is of the form

$$-\frac{\partial U}{\partial x_i}-\frac{d}{dt}\frac{\partial T}{\partial \dot{x}_i}=0;\quad -\frac{\partial U}{\partial y_i}-\frac{d}{dt}\frac{\partial T}{\partial \dot{y}_i}=0;\quad -\frac{\partial U}{\partial z_i}-\frac{d}{dt}\frac{\partial T}{\partial \dot{z}_i}=0,$$

or

$$m_i\ddot{x}_i-F_{ix}=0;\quad m_i\ddot{y}_i-F_{iy}=0;\quad m_i\ddot{z}_i-F_{iz}=0$$
$$(i=1,\ 2,\ \ldots,\ n).$$

If the motion were subject to still another system of independent constraints

$$\varphi_j(t,\ x_1,\ x_2,\ \ldots,\ x_n,\ y_1,\ y_2,\ \ldots,\ y_n,\ z_1,\ z_2,\ \ldots,\ z_n)=0$$
$$(j=1,\ 2,\ \ldots,\ m,\ m<3n),$$

then from the constraint equations it would be possible to express m variables in terms of $3n-m$ independent variables (not counting the time t) or express all $3n$ variables in terms of $3n-m$ new (already independent) coordinates

$$q_1,\ q_2,\ \ldots,\ q_{3n-m}.$$

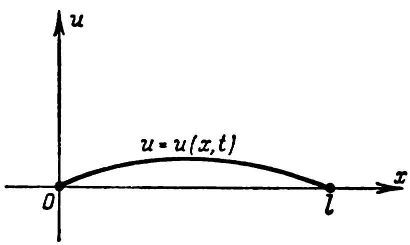


Fig. 6-15

Then T and U might also be regarded as functions of $q_1,\ q_2,\ \ldots$ $\ldots,\ q_{3n-m}$ and t:

$$T=T(q_1,\ q_2,\ \ldots,\ q_{3n-m},\ \dot{q}_1,\ \dot{q}_2,\ \ldots,\ \dot{q}_{3n-m},\ t),$$
$$U=U(q_1,\ q_2,\ \ldots,\ q_{3n-m},\ t),$$

and the system of Euler's equations would have the form

$$\frac{\partial (T-U)}{\partial q_i}-\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i}=0\quad (i=1,\ 2,\ \ldots,\ 3n-m).$$

Example 2. Let us derive the differential equation of free vibrations of a string.

Put the coordinate origin at one of the ends of the string. When in a state of rest under tension, the string lies on a certain straight line along which we shall direct the axis of abscissas (Fig. 6.15). Any deviation from the equilibrium position $u(x,\ t)$ will be a function of the abscissa x and the time t.

The potential energy U of an element of the absolutely flexible string is proportional to the extension of the string. A segment dx of the string in the deformed state is of length $ds=\sqrt{1+u_x'^2}\,dx$ to within higher order infinitesimals, and hence, the elongation of an element is $\left(\sqrt{1+u_x'^2}-1\right)dx$. By Taylor's formula, $\sqrt{1+u_x'^2}\approx$ $\approx 1+\frac{1}{2}u_x'^2$.

Assuming u_x' to be small and neglecting higher powers of u_x', we find that the potential energy of an element is equal to $\frac{1}{2}ku_x'^2\,dx$, where k is a proportionality factor, and the potential energy of the whole string is

$$\frac{1}{2}\int_0^l ku_x'^2\,dx.$$

The kinetic energy of the string is

$$\frac{1}{2}\int_0^l \rho u_t'^2\,dx,$$

where ρ is the density. The integral $\int_{t_0}^{t_1}(T-U)\,dt$ is, in this case, of the form

$$v=\int_{t_0}^{t_1}\int_0^l\left[\frac{1}{2}\rho u_t'^2-\frac{1}{2}ku_x'^2\right]dx\,dt.$$

The equation of motion of the string will be the Ostrogradsky equation for the functional v. Thus, the equation of motion of the string has the form

$$\frac{\partial}{\partial t}(\rho u_t')-\frac{\partial}{\partial x}(ku_x')=0.$$

If the string is homogeneous, then ρ and k are constants, and the equation of a vibrating string is simplified:

$$\rho\frac{\partial^2 u}{\partial t^2}-k\frac{\partial^2 u}{\partial x^2}=0.$$

Now assume that the string is also acted upon by an external force $f(t, x)$ which is perpendicular to the string when the latter is in the equilibrium position and is calculated per unit mass. As is readily verifiable, the force function of this external force acting on an element of the string is $\rho f(t, x)\,u\,dx$; consequently, the

Ostrogradsky-Hamilton integral $\int_{t_0}^{t_1}(T-U)\,dt$ is of the form

$$\int_{t_0}^{t_1}\int_0^l \left[\frac{1}{2}\rho u_t'^2-\frac{1}{2}ku_x'^2+\rho f(t,\ x)u\right]dx\,dt,$$

and the equation of forced vibrations of the string is

$$\frac{\partial}{\partial t}\left(\rho u_t'\right)-\frac{\partial}{\partial x}\left(ku_x'\right)-\rho f(t,\ x)=0,$$

or, if the string is homogeneous,

$$\frac{\partial^2 u}{\partial t^2}-\frac{k}{\rho}\frac{\partial^2 u}{\partial x^2}=f(t,\ x).$$

We can similarly obtain the equation of a vibrating membrane.

Example 3. Let us derive the equation of vibrations of a rectilinear bar. Direct the axis of abscissas along the axis of the bar in the equilibrium position. A deviation from the equilibrium position $u(x,\ t)$ will be a function of x and of the time t, the kinetic energy of a bar of length l is

$$T=\frac{1}{2}\int_0^l \rho u_t'^2\,dx.$$

We assume the bar to be inextensible. The potential energy of an elastic bar with constant curvature is proportional to the square of the curvature. Consequently, the differential dU of the potential energy of the bar is

$$dU=\frac{1}{2}k\left\{\frac{\frac{\partial^2 u}{\partial x^2}}{\left[1+\left(\frac{\partial u}{\partial x}\right)^2\right]^{3/2}}\right\}^2,$$

and the potential energy of the whole bar, the curvature of the axis of which, generally speaking, is variable, will be

$$U=\frac{1}{2}\int_0^l\left\{\frac{\frac{\partial^2 u}{\partial x^2}}{\left[1+\left(\frac{\partial u}{\partial x}\right)^2\right]^{3/2}}\right\}^2 dx.$$

Suppose that the deviations of the bar from the equilibrium position are small and the term $\left(\frac{\partial u}{\partial x}\right)^2$ in the denominator may be neglected; then

$$U=\frac{1}{2}\int_0^l k\left(\frac{\partial^2 u}{\partial x^2}\right)^2 dx.$$

The Ostrogradsky-Hamilton integral is of the form

$$\int_{t_0}^{t_1}\int_{0}^{l}\left[\frac{1}{2}\rho u_t'^2-\frac{1}{2}ku_{xx}''^2\right]dx\,dt.$$

Consequently, in the case of the free vibrations of an elastic bar we will have the following equation of motion:

$$\frac{\partial}{\partial t}(\rho u_t')+\frac{\partial^2}{\partial x^2}(ku_{xx}'')=0.$$

If the bar is homogeneous, then ρ and k are constants, and the equation of vibrations of the bar is transformed to

$$\rho\frac{\partial^2 u}{\partial t^2}+k\frac{\partial^4 u}{\partial x^4}=0.$$

If the bar is acted upon by an external force $f(t, x)$, then we also have to take into account the potential of this force (see preceding example).

The principle of least action may be applied in deriving field equations. Consider the scalar, vector or tensor field $w=w(x, y, z, t)$. The integral $\int_{t_0}^{t_1}(T-U)\,dt$ here will, generally speaking, be equal to the quadruple integral taken over the space coordinates x, y, z and the time t of some function L, called the *density of the Lagrange function* or the *Lagrangian*.

Ordinarily, the Lagrangian is a function of w, $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$, $\frac{\partial w}{\partial z}$, $\frac{\partial w}{\partial t}$:

$$L=L\left(w, \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z}, \frac{\partial w}{\partial t}\right),$$

and, therefore, the action is of the form

$$\iiiint_D L\left(w, \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z}, \frac{\partial w}{\partial t}\right)dx\,dy\,dz\,dt. \tag{6.3}$$

According to the principle of least action, the field equation is the Ostrogradsky equation for the functional (6.3):

$$L_w-\frac{\partial}{\partial x}\{L_{p_1}\}-\frac{\partial}{\partial y}\{L_{p_2}\}-\frac{\partial}{\partial z}\{L_{p_3}\}-\frac{\partial}{\partial t}\{L_{p_4}\}=0,$$

where

$$p_1=\frac{\partial w}{\partial x},\quad p_2=\frac{\partial w}{\partial y},\quad p_3=\frac{\partial w}{\partial z},\quad p_4=\frac{\partial w}{\partial t}.$$

PROBLEMS ON CHAPTER 6

1. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{y}\, dx.$$

2. Test for an extremum the functional

$$v[y(x)] = \int (y^2 + 2xyy')\, dx; \quad y(x_0) = y_0; \quad y(x_1) = y_1.$$

3. Test for an extremum the functional

$$v[y(x)] = \int_0^1 (xy + y^2 - 2y^2y')\, dx; \quad y(0) = 1; \quad y(1) = 2.$$

4. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} y'(1 + x^2y')\, dx.$$

5. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (y'^2 + 2yy' - 16y^2)\, dx.$$

6. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (xy' + y'^2)\, dx.$$

7. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} \frac{1+y^2}{y'^2}\, dx.$$

8. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (y^2 + y'^2 - 2y \sin x)\, dx.$$

9. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (16y^2 - y''^2 + x^2)\, dx.$$

10. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (2xy + y'''^2)\,dx.$$

11. Find the extremals of the functional

$$v[y(x),\ z(x)] = \int_{x_0}^{x_1} (2yz - 2y^2 + y'^2 - z'^2)\,dx.$$

12. Write the Ostrogradsky equation for the functional

$$v[z(x,\ y)] = \int\int_D \left[\left(\frac{\partial z}{\partial x}\right)^2 - \left(\frac{\partial z}{\partial y}\right)^2\right] dx\,dy.$$

13. Write the Ostrogradsky equation for the functional

$$v[u(x,\ y,\ z)] = \int\int\int_D \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 + 2uf(x,\ y,\ z)\right] dx\,dy\,dz.$$

14. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} \frac{y'^2}{x^3}\,dx.$$

15. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (y^2 + y'^2 + 2ye^x)\,dx.$$

16. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} (y^2 - y'^2 - 2y\sin x)\,dx.$$

17. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} \left[y^2 + (y')^2 + \frac{2y}{\cosh x}\right] dx.$$

18. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} [x^2(y')^2 + 2y^2 + 2xy]\,dx.$$

19. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} [(y'')^2 - 2(y')^2 + y^2 - 2y \sin x]\, dx.$$

20. Find the extremals of the functional

$$v[y(x)] = \int_{x_0}^{x_1} [(y''')^2 + y^2 - 2yx^3]\, dx.$$

Variational problems with moving boundaries and certain other problems

1. An Elementary Problem with Moving Boundaries

When we investigated the functional

$$v=\int_{x_0}^{x_1} F(x, y, y')\,dx$$

in Chapter 6 it was assumed that the boundary points (x_0, y_0) and (x_1, y_1) are given. Now let us suppose that one or both of the boundary points can move. Then the class of permissible curves is extended: in addition to the comparison curves that have common boundary points with the curve under investigation, we can now also take curves with displaced boundary points.

Therefore, if on a curve $y=y(x)$ an extremum is reached in a problem with moving boundary points, then the extremum is all the more attained relative to a narrower class of curves having common boundary points with the curve $y=y(x)$ and, hence, the basic condition for achieving an extremum in a problem with fixed boundaries must be fulfilled—the function $y(x)$ must be a solution of the Euler equation

$$F_y-\frac{d}{dx}F_{y'}=0.$$

Thus, the curves $y=y(x)$ on which an extremum is achieved in the moving-boundary problem must be extremals.

The general solution of the Euler equation contains two arbitrary constants, for the determination of which we need two conditions. In the problem with fixed boundary points, these conditions were

$$y(x_0)=y_0 \quad \text{and} \quad y(x_1)=y_1.$$

In the moving-boundary problem, one or both of these conditions are absent and the missing conditions for a determination of the arbitrary constants of the general solution of Euler's equation have to be obtained from the basic necessary condition for an extremum which is that the variation δv be equal to zero.

Since in the moving-boundary problem an extremum is attained only on solutions $y=y(x, C_1, C_2)$ of the Euler equation, from now on we can consider the value of the functional only on functions of this family Then the functional $v[y(x, C_1, C_2)]$ is reduced to a function of the parameters C_1 and C_2 and of the limits of integration x_0 and x_1, while the variation of the functional coincides with the differential of this function. For the purpose of simplification, we shall assume that one of the boundary points, say (x_0, y_0), is fixed, and the other (x_1, y_1) can be moved and passes to point $(x_1+\Delta x_1, y_1+\Delta y_1)$ or, as ordinarily denoted in the calculus of variations, $(x_1+\delta x_1, y_1+\delta y_1)$.

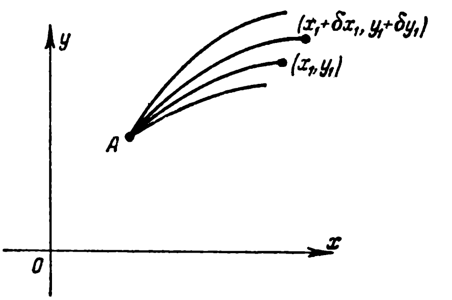


Fig. 7-1

We will call the permissible curves $y=y(x)$ and $y=y(x)+\delta y$ neighbouring if the absolute values of the variations δy and $\delta y'$ are small, and the absolute values of the increments δx_1 and δy_1 are also small (the increments δx_1 and δy_1 are ordinarily called variations of the limit values x_1 and y_1).

The extremals passing through the point (x_0, y_0) form a pencil of extremals $y=y(x, C_1)$. The functional $v[y(x, C_1)]$ on the curves of this pencil is reduced to a function of C_1 and x_1 If the curves of the pencil $y=y(x, C_1)$ in the neighbourhood of the extremal under consideration do not intersect, then $v[y(x, C_1)]$ may be regarded as a one-valued function of x_1 and y_1 since specification of x_1 and y_1 determines the extremal of the pencil (Fig. 7.1) and thus determines the value of the functional.

Let us compute the variation of the functional $v[y(x, C_1)]$ on the extremals of the pencil $y=y(x, C_1)$ when the boundary point is displaced from the position (x_1, y_1) to the position $(x_1+\delta x_1, y_1+\delta y_1)$. Since the functional v on the curves of the pencil is reduced to a function of x_1 and y_1, its variation coincides with the differential of this function. From the increment Δv, remove the principal part that is linear in δx_1 and δy_1:

$$\Delta v = \int_{x_0}^{x_1+\delta x_1} F(x, y+\delta y, y'+\delta y')\,dx - \int_{x_0}^{x_1} F(x, y, y')\,dx =$$

$$= \int_{x_1}^{x_1+\delta x_1} F(x, y+\delta y, y'+\delta y')\,dx +$$

$$+ \int_{x_0}^{x_1} [F(x, y+\delta y, y'+\delta y') - F(x, y, y')]\,dx \qquad (7.1)$$

Using the mean-value theorem, transform the first term on the right:

$$\int_{x_1}^{x_1+\delta x_1} F(x,\ y+\delta y,\ y'+\delta y')\,dx = F\,|_{x=x_1+\theta\,\delta x_1}\,\delta x_1, \quad \text{where} \quad 0<\theta<1,$$

by virtue of the continuity of the function F we will have

$$F\,|_{x=x_1+\theta\,\delta x_1} = F(x,\ y,\ y')\,|_{x=x_1} + \varepsilon_1,$$

where $\varepsilon_1 \to 0$ as $\delta x_1 \to 0$ and $\delta y_1 \to 0$.

Thus

$$\int_{x_1}^{x_1+\delta x_1} F(x,\ y+\delta y,\ y'+\delta y')\,dx = F(x,\ y,\ y')\,|_{x=x_1}\,\delta x_1 + \varepsilon_1\,\delta x_1.$$

We transform the second term on the right-hand side of (7.1) by means of Taylor's expansion of the integrand:

$$\int_{x_0}^{x_1} [F(x,\ y+\delta y,\ y'+\delta y') - F(x,\ y,\ y')]\,dx =$$
$$= \int_{x_0}^{x_1} [F_y(x,\ y,\ y')\,\delta y + F_{y'}(x,\ y,\ y')\,\delta y']\,dx + R_1.$$

where R_1 is an infinitesimal of higher order than δy or $\delta y'$. In turn, the linear part

$$\int_{x_0}^{x_1} (F_y\delta y + F_{y'}\delta y')\,dx$$

may be transformed, by integrating by parts the second summand of the integrand, to the form

$$[F_{y'}\,\delta y]_{x_0}^{x_1} + \int_{x_0}^{x_1} \left(F_y - \frac{d}{dx}F_{y'}\right)\delta y\,dx.$$

The values of the functional are only taken on extremals, hence $F_y - \frac{d}{dx}F_{y'} \equiv 0$. Since the boundary point $(x_0,\ y_0)$ is fixed, it follows that $\delta y\,|_{x=x_0} = 0$. Hence

$$\int_{x_0}^{x_1} (F_y\,\delta y + F_{y'}\,\delta y')\,dx = [F_{y'}\,\delta y]\,|_{x=x_1}.$$

Note that $\delta y\,|_{x=x_1}$ is not equal to δy_1, the increment of y_1, since δy_1 is the increment of y_1 when the boundary point is displaced to the position $(x_1+\delta x_1, y_1+\delta y_1)$, and $\delta y|_{x=x_1}$ is the increment of the ordinate at the point x_1 when going from the extremal passing through the points (x_0, y_0) and (x_1, y_1) to the extremal passing through the points (x_0, y_0) and $(x_1+\delta x_1, y_1+\delta y_1)$ (Fig. 7.2).

From the figure it is clear that $BD=\delta y|_{x=x_1}$; $FC=\delta y_1$;

$$EC \approx y'(x_1)\,\delta x_1; \quad BD = FC - EC$$

or

$$\delta y|_{x=x_1} \approx \delta y_1 - y'(x_1)\,\delta x_1.$$

Here, the approximate equality holds true to within infinitesimals of higher order.

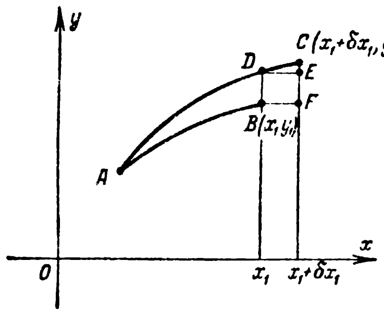


Fig. 7-2

And so we finally have

$$\int_{x_1}^{x_1+\delta x_1} F\,dx \approx F|_{x=x_1}\,\delta x_1;$$

$$\int_{x_0}^{x_1} [F(x, y+\delta y, y'+\delta y') - F(x, y, y')]\,dx \approx F_{y'}|_{x=x_1}\cdot(\delta y_1 - y'(x_1)\,\delta x_1),$$

where the approximate equalities likewise hold true to within terms higher than the first in δx_1 and δy_1. Hence, from (7.1) we get

$$\delta v = F|_{x=x_1}\,\delta x_1 + F_{y'}|_{x=x_1}(\delta y_1 - y'(x_1)\,\delta x_1) = (F - y'F_{y'})|_{x=x_1}\,\delta x_1 + F_{y'}|_{x=x_1}\,\delta y_1,$$

or

$$d\bar{v}(x_1, y_1) = (F - y'F_{y'})|_{x=x_1}\,dx_1 + F_{y'}|_{x=x_1}\,dy_1,$$

where $\bar{v}(x_1, y_1)$ is a function into which the functional v was transformed on the extremals $y=y(x, C_1)$, and $dx_1=\Delta x_1=\delta x_1$, $dy_1=\Delta y_1=\delta y_1$ are increments of the coordinates of the boundary point. The basic necessary condition for an extremum $\delta v=0$ takes the form

$$(F - y'F_{y'})|_{x=x_1}\,\delta x_1 + F_{y'}|_{x=x_1}\,\delta y_1 = 0. \qquad (7.2)$$

If the variations δx_1 and δy_1 are independent, then it follows that

$$(F - y'F_{y'})|_{x=x_1} = 0 \quad \text{and} \quad F_{y'}|_{x=x_1} = 0.$$

However, more often one has to consider the case when the variations δx_1 and δy_1 are dependent.

For instance, allow the right boundary point (x_1, y_1) to move along a certain curve

$$y_1 = \varphi(x_1).$$

Then $\delta y_1 \approx \varphi'(x_1)\delta x_1$ and, consequently, the condition (7.2) takes the form $[F+(\varphi'-y')F_{y'}]\delta x_1=0$ or, since δx_1 varies arbitrarily, it follows that $[F+(\varphi'-y')F_{y'}]_{x=x_1}=0$. This condition establishes a relationship between the slopes of φ' and y' at the boundary point. It is called the *transversality condition*.

The transversality condition together with the condition $y_1=\varphi(x_1)$ generally speaking enables us to determine one or several extremals of the pencil $y=y(x, C_1)$ on which an extremum may be achieved. If the boundary point (x_0, y_0) can move along some curve $y_0=\psi(x_0)$, then in the very same way we will find that the transversality condition

$$[F+(\psi'-y')F_{y'}]_{x=x_0}=0$$

must be satisfied also at the point (x_0, y_0).

Example 1. Find the transversality condition for functionals of the form

$$v=\int_{x_0}^{x_1} A(x, y)\sqrt{1+y'^2}\,dx.$$

The transversality condition $F+F_{y'}(\varphi'-y')=0$ is in this case of the form

$$A(x, y)\sqrt{1+y'^2}+\frac{A(x, y)\,y'}{\sqrt{1+y'^2}}(\varphi'-y')=0 \text{ or } \frac{A(x, y)(1+\varphi'y')}{\sqrt{1+y'^2}}=0;$$

assuming that $A(x, y)\neq 0$ at the boundary point, we get $1+y'\varphi'=0$ or $y'=-\frac{1}{\varphi'}$; that is, the transversality condition is in this case reduced to the orthogonality condition.

Example 2. Test for an extremum the functional $\int_0^{x_1}\frac{\sqrt{1+y'^2}}{y}dx$ given that $y(0)=0$ and $y_1=x_1-5$ (Fig. 7.3). The integral curves of the Euler equation (Problem 1, page 338) are the circles $(x-C_1)^2+y^2=C_2^2$. The first boundary condition yields $C_1=C_2$. Since for the given functional the transversality condition reduces to the orthogonality condition (see preceding example), it follows that the straight line $y_1=x_1-5$ must be a diameter of the circle, and, hence, the centre of the desired circle lies at the point (0, 5), where the straight line $y_1=x_1-5$ meets the axis of abscissas. Consequently, $(x-5)^2+y^2=25$ or $y=\pm\sqrt{10x-x^2}$. And so, an extremum can be achieved only on arcs of the circle $y=\sqrt{10x-x^2}$ and $y=-\sqrt{10x-x^2}$.

If the boundary point (x_1, y_1) can move only along a vertical straight line (Fig. 7.4) and hence $\delta x_1 = 0$, then the condition (7.2) passes into $F_{y'}|_{x=x_1} = 0$.

For example, in the problem of the brachistochrone (see pages 316-317) let the left boundary point be fixed and the right allowed to

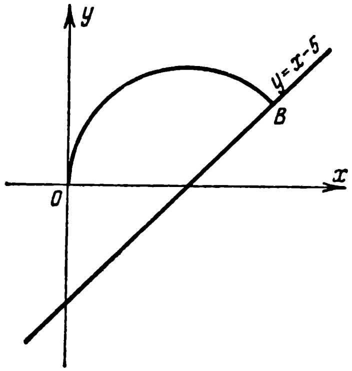


Fig. 7-3.

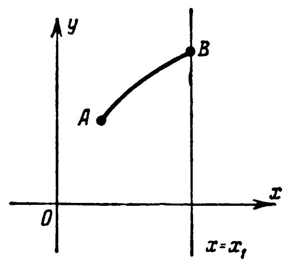


Fig. 7-4

move along a vertical straight line. The extremals of the functional $v = \int_0^{x_1} \frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx$ are cycloids, the equations of which, if the condition $y(0) = 0$ is taken into consideration, will be of the form

$$x = C_1 (t - \sin t),$$
$$y = C_1 (1 - \cos t).$$

To determine C_1, use the condition $F_{y'}|_{x=x_1} = 0$, which in the given case is of the form

$$\left. \frac{y'}{\sqrt{y}\sqrt{1+y'^2}} \right|_{x=x_1} = 0,$$

whence $y'(x_1) = 0$; that is, the desired cycloid must intersect the straight line $x = x_1$ at right angles and, hence, the point $x = x_1$, $y = y_1$ must be a cusp of the cycloid (Fig. 7.5). Since to the cusp there corresponds the value $t = \pi$, it follows that $x_1 = C_1\pi$, $C_1 = \frac{x_1}{\pi}$. Hence, an extremum can be achieved only on the cycloid

$$x = \frac{x_1}{\pi}(t - \sin t); \quad y = \frac{x_1}{\pi}(1 - \cos t).$$

If the boundary point (x_1, y_1) in the problem of the extremum of the functional $v=\int_{x_0}^{x_1} F(x, y, y')\,dx$ is permitted to move along

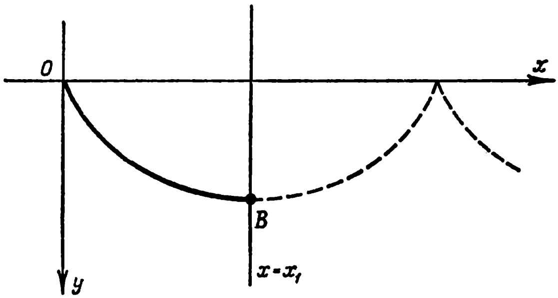


Fig. 7-5

the horizontal straight line $y=y_1$, then $\delta y_1=0$ and the condition (7.2), or the transversality condition, takes the form

$$[F-y'F_{y'}]_{x=x_1}=0.$$

2. The Moving-Boundary Problem for a Functional of the Form $\int_{x_0}^{x_1} F(x, y, z, y', z')\,dx$

If in investigating the functional

$$v=\int_{x_0}^{x_1} F(x, y, z, y', z')\,dx$$

for an extremum, one of the boundary points, say $B(x_1, y_1, z_1)$ is moved, and the other, $A(x_0, y_0, z_0)$, is fixed (or both boundary points are movable), then it is obvious that an extremum may be achieved only on the integral curves of the system of Euler's equations

$$F_y-\frac{d}{dx}F_{y'}=0;\quad F_z-\frac{d}{dx}F_{z'}=0.$$

Indeed, if on a certain curve C an extremum is achieved in the moving-boundary problem, i.e. a maximum or minimum value of v is obtained compared with the values of v on all close-lying permissible curves, which include both those curves that have common boundary points with the extremizing curve C and also those curves whose boundary points do not coincide with the boundary points of C, then an extremum is surely achieved on the curve

C with respect to a narrower class of neighbouring curves having common boundary points with the curve C.

Hence, the necessary conditions for an extremum in the problem with fixed boundary points must be satisfied on the curve C, and in particular, the curve C must be an integral curve of a system of Euler's equations.

The general solution of the system of Euler's equations contains four arbitrary constants. Knowing the coordinates of the boundary point $A(x_0, y_0, z_0)$, which we consider to be fixed, it is possible, generally speaking, to eliminate two arbitrary constants.

To determine the other two arbitrary constants we have to have another two equations, which will be obtained from the condition $\delta v = 0$; note that in computing the variation we will now assume that the functional is specified only on solutions of the system of Euler's equations, for an extremum is attainable only on them. Then the functional v is reduced to the function $\Phi(x_1, y_1, z_1)$ of the coordinates x_1, y_1, z_1 of the point $B(x_1, y_1, z_1)$, and the variation of the functional is reduced to the differential of this function. *

The calculation of the variation of v may be performed exactly as indicated on pages 342-344:

$$\Delta v = \int_{x_0}^{x_1+\delta x_1} F(x, y+\delta y, z+\delta z, y'+\delta y', z'+\delta z')\,dx - \\ - \int_{x_0}^{x_1} F(x, y, z, y', z')\,dx = \\ = \int_{x_1}^{x_1+\delta x_1} F(x, y+\delta y, z+\delta z, y'+\delta y', z'+\delta z')\,dx + \\ + \int_{x_0}^{x_1} [F(x, y+\delta y, z+\delta z, y'+\delta y', z'+\delta z') - \\ - F(x, y, z, y', z')]\,dx.$$

Apply the mean-value theorem to the first integral and take advantage of the continuity of the function F; in the second integral isolate the principal linear part by means of Taylor's formula. These transformations yield

$$\delta v = F|_{x=x_1}\,\delta x_1 + \int_{x_0}^{x_1} [F_y\,\delta y + F_z\,\delta z + F_{y'}\,\delta y' + F_{z'}\,\delta z']\,dx$$

* The function Φ will be single-valued if the extremals of the pencil centred at A do not intersect, for then the point $B(x_1, y_1, z_1)$ uniquely defines an extremal.

Integrating by parts the last two terms under the integral sign yields

$$\delta v = F|_{x=x_1}\,\delta x_1 + [F_{y'}\,\delta y]_{x=x_1} + [F_{z'}\,\delta z]_{x=x_1} + \\ + \int_{x_0}^{x_1}\left[\left(F_y - \frac{d}{dx}F_{y'}\right)\delta y + \left(F_z - \frac{d}{dx}F_{z'}\right)\delta z'\right] dx.$$

Since the values of v are calculated only on extremals, it follows that

$$F_y - \frac{d}{dx}F_{y'} \equiv 0; \qquad F_z - \frac{d}{dx}F_{z'} \equiv 0$$

and hence,

$$\delta v = F|_{x=x_1}\,\delta x_1 + [F_{y'}\,\delta y]_{x=x_1} + [F_{z'}\,\delta z]_{x=x_1}.$$

Arguing in the same way as on page 344, we get

$$\delta y|_{x=x_1} \approx \delta y_1 - y'(x_1)\,\delta x_1 \quad \text{and} \quad \delta z|_{x=x_1} \approx \delta z_1 - z'(x_1)\,\delta x_1,$$

and, consequently,

$$\delta v = [F - y'F_{y'} - z'F_{z'}]_{x=x_1}\,\delta x_1 + F_{y'}|_{x=x_1}\,\delta y_1 + F_{z'}|_{x=x_1}\,\delta z_1 = 0.$$

If the variations δx_1, δy_1, δz_1 are independent, then from the condition $\delta v = 0$ we have

$$[F - y'F_{y'} - z'F_{z'}]_{x=x_1} = 0; \quad F_{y'}|_{x=x_1} = 0 \quad \text{and} \quad F_{z'}|_{x=x_1} = 0.$$

If the boundary point $B(x_1, y_1, z_1)$ can move along some curve $y_1 = \varphi(x_1)$; $z = \psi(x_1)$, then $\delta y_1 = \varphi'(x_1)\,\delta x_1$, and $\delta z_1 = \psi'(x_1)\,\delta x_1$ and the condition $\delta v = 0$ or

$$[F - y'F_{y'} - z'F_{z'}]_{x=x_1}\,\delta x_1 + F_{y'}|_{x=x_1}\,\delta y_1 + F_{z'}|_{x=x_1}\,\delta z_1 = 0$$

passes into the condition

$$[F + (\varphi' - y')F_{y'} + (\psi' - z')F_{z'}]_{x=x_1}\,\delta x_1 = 0,$$

whence, by virtue of the arbitrariness of δx_1, we have

$$[F + (\varphi' - y')F_{y'} + (\psi' - z')F_{z'}]_{x=x_1} = 0.$$

This condition is called the *transversality condition* in the problem of investigating the functional

$$v = \int_{x_0}^{x_1} F(x, y, z, y', z')\,dx$$

for an extremum. Together with the equations $y_1 = \varphi(x_1)$, $z_1 = \psi(x_1)$, the transversality condition yields the equations needed to determine the arbitrary constants in the general solution of the system of Euler's equations.

If the boundary point $B(x_1, y_1, z_1)$ can move over some surfaces $z_1 = \varphi(x_1, y_1)$, then $\delta z_1 = \varphi_{x_1}\delta x_1 + \varphi_{y_1}\delta y_1$ and the variations δx_1 and δy_1 are arbitrary. Consequently, the condition $\delta v = 0$ or, in expanded form,

$$[F - y'F_{y'} - z'F_{z'}]_{x=x_1}\delta x_1 + F_{y'}|_{x=x_1}\delta y_1 + F_{z'}|_{x=x_1}\delta z_1 = 0$$

is transformed to the condition

$$[F - y'F_{y'} - z'F_{z'} + \varphi'_x F_{z'}]_{x=x_1}\delta x_1 + [F_{y'} + F_{z'}\varphi'_y]_{x=x_1}\delta y_1 = 0.$$

Whence, since δx_1 and δy_1 are independent, we get

$$[F - y'F_{y'} + (\varphi'_x - z')F_{z'}]_{x=x_1} = 0, \quad [F_{y'} + F_{z'}\varphi'_y]_{x=x_1} = 0.$$

These two conditions, together with the equation $z_1 = \varphi(x_1, y_1)$, generally speaking, enable one to determine two arbitrary constants in the general solution of the system of Euler's equations

If the boundary point $A(x_0, y_0, z_0)$ is moving, then by the same method at this point we get similar conditions.

If we consider the functional

$$v = \int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y'_1, y'_2, \ldots, y'_n)\,dx,$$

then, without changing the proof procedure, we find that in the case of the moving point $B(x_1, y_{11}, y_{21}, \ldots, y_{n1})$ at this point

$$\left(F - \sum_{i=1}^{n} y'_i F_{y'_i}\right)\Bigg|_{x=x_1}\delta x_1 + \sum_{i=1}^{n} F_{y'_i}\Bigg|_{x=x_1}\delta y_{i1} = 0.$$

Example 1. Find the transversality condition for the functional

$$v = \int_{x_0}^{x_1} A(x, y, z)\sqrt{1 + y'^2 + z'^2}\,dx \text{ if } z_1 = \varphi(x_1, y_1).$$

The transversality conditions

$$[F - y'F_{y'} + (\varphi'_x - z')F_{z'}]_{x=x_1} = 0 \quad \text{and} \quad [F_{y'} + F_{z'}\varphi'_y]_{x=x_1} = 0$$

in this case are of the form $1 + \varphi'_x z' = 0$ and $y' + \varphi'_y z' = 0$ for $x = x_1$ or $\frac{1}{\varphi'_x} = \frac{y'}{\varphi'_y} = \frac{z'}{-1}$ for $x = x_1$, that is, they are the condition for parallelism of the vector of the tangent $\bar{t}(1, y', z')$ to the desired extremal at the point (x_1, y_1, z_1) and the vector of the normal $\bar{N}(\varphi'_x, \varphi'_y, -1)$ to the surface $z = \varphi(x, y)$ at the same point. Hence, in this case, the transversality condition becomes an orthogonality condition of the extremal to the surface $z = \varphi(x, y)$.

Example 2. Find the extremal distance between two surfaces

$$z=\varphi(x, y) \quad \text{and} \quad z=\psi(x, y).$$

In other words, find the extremum of the integral $l=\int_{x_0}^{x_1}\sqrt{1+y'^2+z'^2}\,dx$, provided that the coordinates of one of the boundary points (x_0, y_0, z_0) satisfy the equation $z_0=\varphi(x_0, y_0)$ and the coordinates of the other boundary point (x_1, y_1, z_1) satisfy the equation $z_1=\psi(x_1, y_1)$.

Since the integrand depends solely on y' and z', the extremals are straight lines (Example 2, page 320). Since the functional $\int_{x_0}^{x_1}\sqrt{1+y'^2+z'^2}\,dx$ is a particular case of the functional $\int_{x_0}^{x_1}A(x, y, z)\sqrt{1+y'^2+z'^2}\,dx$ considered in the preceding example, the transversality conditions both at the point (x_0, y_0, z_0) and at the point (x_1, y_1, z_1) pass into orthogonality conditions. Hence, an extremum can be achieved only on straight lines that are orthogonal both to the surface $z=\varphi(x, y)$ at the point (x_0, y_0, z_0) and to the surface $z=\psi(x, y)$ at the point (x_1, y_1, z_1) (Fig. 7.6).

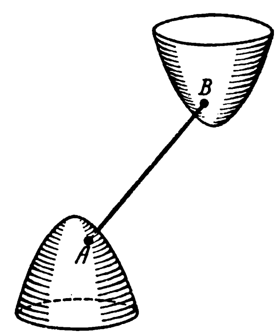


Fig. 7-6

Example 3. Test for an extremum the functional $v=\int_{x_0}^{x_1}(y'^2+z'^2+2yz)\,dx$ given that $y(0)=0$; $z(0)=0$, and the point (x_1, y_1, z_1) can move over the plane $x=x_1$.

The system of Euler's equations has the form $z''-y=0$; $y''-z=0$, whence $y^{IV}-y=0$; $y=C_1\cosh x+C_2\sinh x+C_3\cos x+C_4\sin x$, $z=y''$; $z=C_1\cosh x+C_2\sinh x-C_3\cos x-C_4\sin x$. From the conditions $y(0)=0$, and $z(0)=0$ we get: $C_1+C_3=0$ and $C_1-C_3=0$, whence $C_1=C_3=0$. The condition at the moving boundary point

$$(F-y'F_{y'}-z'F_{z'})_{x=x_1}\delta x_1+F_{y'}|_{x=x_1}\delta y_1+F_{z'}|_{x=x_1}\delta z_1=0$$

passes into the conditions

$$F_{y'}|_{x=x_1}=0 \quad \text{and} \quad F_{z'}|_{x=x_1}=0,$$

since $\delta x_1=0$ and δy_1 and δz_1 are arbitrary. In the example under consideration, $F_{y'}=2y'$; $F_{z'}=2z'$, hence

$$y'(x_1)=0 \quad \text{and} \quad z'(x_1)=0$$

or

$$C_2 \cosh x + C_4 \cos x_1 = 0 \quad \text{and} \quad C_2 \cosh x - C_4 \cos x_1 = 0.$$

If $\cos x_1 \neq 0$, then $C_2 = C_4 = 0$ and an extremum may be achieved only on the straight line $y = 0$; $z = 0$; but if $\cos x_1 = 0$, i.e., $x_1 = \frac{\pi}{2} + n\pi$, where n is an integer, then $C_2 = 0$, C_4 is an arbitrary constant, $y = C_4 \sin x$, $z = -C_4 \sin x$. It is easy to verify that in the last case the functional $v = 0$ for any C_4.

3. Extremals with Corners

Up till now we have considered variational problems in which the desired function $y = y(x)$ was assumed continuous with a continuous derivative. However, in many problems the latter requirement is not natural; what is more, in certain classes of variational problems the solution is, as a rule, attained on extremals having corner points. Such problems include for instance problems involving the reflection and refraction of extremals, which are a generalization of the corresponding problems involving the reflection and refraction of light.

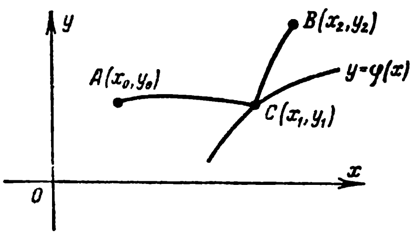


Fig. 7-7

The reflection-of-extremals problem. Find the curve that extremizes the functional $v = \int_{x_0}^{x_2} F(x, y, y')\, dx$ and passes through the given points $A(x_0, y_0)$ and $B(x_2, y_2)$; the curve must arrive at B only after being reflected from a given line $y = \varphi(x)$ (Fig. 7.7).

It is natural to assume that at the point of reflection $C(x_1, y_1)$ there can be a corner point of the desired extremal and, consequently, at this point the left-hand derivative $y'(x_1 - 0)$ and the right-hand derivative $y'(x_1 + 0)$ are, generally speaking, distinct. It is therefore more convenient to represent the functional $v[y(x)]$ in the form

$$v[y(x)] = \int_{x_0}^{x_1} F(x, y, y')\, dx + \int_{x_1}^{x_2} F(x, y, y')\, dx;$$

here, on each of the intervals $x_0 < x \leqslant x_1$ and $x_1 \leqslant x \leqslant x_2$ the derivative $y'(x)$ is assumed continuous and, hence, we can take advantage of the results given above.

The basic necessary condition for an extremum, $\delta v=0$, takes the form

$$\delta v=\delta\int_{x_0}^{x_1} F(x,\ y,\ y')\,dx+\delta\int_{x_1}^{x_2} F(x,\ y,\ y')\,dx=0.$$

Since the point $(x_1,\ y_1)$ can move along the curve $y=\varphi(x)$, it follows that in calculating the variations $\delta\int_{x_0}^{x_1} F(x,\ y,\ y')\,dx$ and $\delta\int_{x_1}^{x_2} F(x,\ y,\ y')\,dx$ we are involved in the conditions of the problem with a boundary point moving along a given curve and we can make use of the results of Sec. 1 (page 341). It is obvious that the curves AC and CB are extremals. Indeed, on these segments $y=y(x)$ is a solution of the Euler equation, since if we assume that one of these curves has already been found and if we vary the other one alone, then the problem reduces to finding the extremum of the functional $\int_{x_0}^{x_1} F\,dx \left(\text{or } \int_{x_1}^{x_2} F\,dx\right)$ in the problem with fixed boundary points. For this reason, when calculating the variation of the functional we will assume that the functional is considered only on extremals having the corner point C. Then

$$\delta\int_{x_0}^{x_1} F(x,\ y,\ y')\,dx=[F+(\varphi'-y')F_{y'}]_{x=x_1-0}\,\delta x_1$$

and

$$\delta\int_{x_1}^{x_2} F(x,\ y,\ y')\,dx=-[F+(\varphi'-y')F_{y'}]_{x=x_1+0}\,\delta x_1$$

(see page 345), where the signs $x=x_1-0$ and $x=x_1+0$ signify that we take the limiting value of the quantity in the parentheses as the point x_1 is approached in the first case from the left (from the side of values of x less than x_1) and in the second case from the right (from the side of values of x greater than x_1). Since only the derivative y' is discontinuous at the point of reflection, it follows that in the first case we should take the left-hand derivative at the corner point and in the second case, the right-hand derivative.

The condition $\delta v=0$ takes the form

$$[F+(\varphi'-y')F_{y'}]_{x=x_1-0}\,\delta x_1-[F+(\varphi'-y')F_{y'}]_{x=x_1+0}\,\delta x_1=0$$

or, since δx_1 varies arbitrarily, then

$$[F+(\varphi'-y')F_{y'}]_{x=x_1-0}=[F+(\varphi'-y')F_{y'}]_{x=x_1+0}$$

or

$$F(x_1, y_1, y'(x_1-0)) + (\varphi'(x_1) - y'(x_1-0)) F_{y'}(x_1, y_1, y'(x_1-0)) = F(x_1, y_1, y'(x_1+0)) + (\varphi'(x_1) - y'(x_1+0)) F_{y'}(x_1, y_1, y'(x_1+0)).$$

This condition of reflection acquires a particularly simple form for functionals of the type

$$v = \int_{x_0}^{x_1} A(x, y)\sqrt{1+y'^2}\,dx,$$

namely:

$$A(x_1, y_1)\left[\sqrt{1+y'^2} + \frac{(\varphi'-y')y'}{\sqrt{1+y'^2}}\right]_{x=x_1-0} = A(x_1, y_1)\left[\sqrt{1+y'^2} + \frac{(\varphi'-y')y'}{\sqrt{1+y'^2}}\right]_{x=x_1+0}$$

or, simplifying and cancelling $A(x_1, y_1)$ on the assumption that $A(x_1, y_1) \neq 0$, we have

$$\left.\frac{1+\varphi'y'}{\sqrt{1+y'^2}}\right|_{x=x_1-0} = \left.\frac{1+\varphi'y'}{\sqrt{1+y'^2}}\right|_{x=x_1+0}.$$

Designating the angle between the tangent to the curve $y = \varphi(x)$ and the axis of abscissas by the letter α, and the slopes, to the

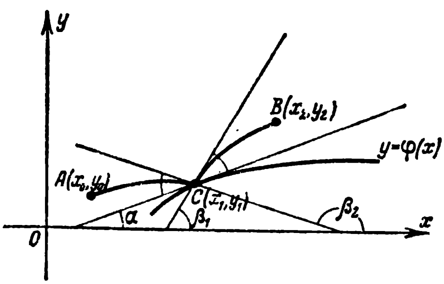


Fig. 7-8

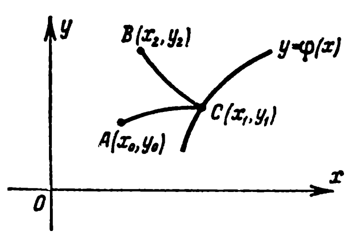


Fig. 7-9

abscissa axis, of the left and right tangents to the extremal at the point of reflection C as β_1 and β_2, respectively (Fig. 7.8), we get

$$y'(x_1-0) = \tan\beta_1, \quad y'(x_1+0) = \tan\beta_2, \quad \varphi'(x_1) = \tan\alpha.$$

The condition at the reflection point takes the form

$$\frac{1+\tan\alpha\cdot\tan\beta_1}{-\sec\beta_1} = \frac{1+\tan\alpha\cdot\tan\beta_2}{\sec\beta_2}$$

or, after simplification and multiplication by $\cos\alpha$:

$$-\cos(\alpha-\beta_1)=\cos(\alpha-\beta_2).$$

Whence follows the equality of the angle of incidence and the angle of reflection.

If a point is in motion in some medium with velocity $v(x, y)$, the time t spent on moving the point from the position $A(x_0, y_0)$

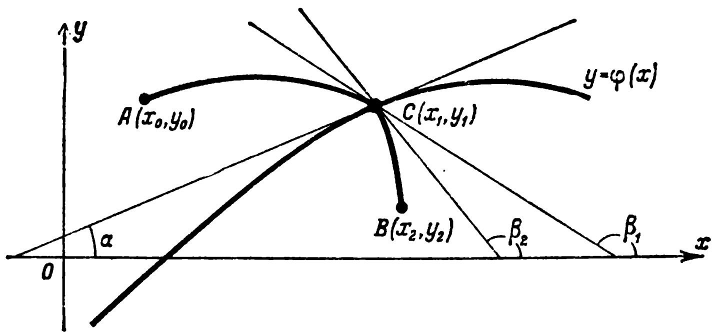


Fig. 7-10

to the position $B(x_1, y_1)$ is equal to the integral $\int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{v(x, y)}\,dx$ which belongs to the type of functional $\int_{x_0}^{x_1} A(x, y)\sqrt{1+y'^2}\,dx$ under consideration and, hence, for any law of variation of velocity $v(x, y)$, at the point of reflection the angle of incidence is equal to the angle of reflection.

If the points A, B and C were arranged differently, for example, as in Fig. 7.9, then in order to obtain the same condition at the point of reflection, it would be more convenient, due to the two-valued nature of the function $y=y(x)$, to carry out the investigation in parametric form.

Refraction of extremals. Let us suppose that in the region under investigation the integrand of the functional $v=\int_{x_0}^{x_2} F(x, y, y')\,dx$ has a discontinuity line $y=\varphi(x)$, and the boundary points A and B are located on different sides of the discontinuity line (Fig. 7.10).

Represent the functional v in the form

$$v=\int_{x_0}^{x_1} F_1(x,\ y,\ y')\,dx+\int_{x_1}^{x_2} F_2(x,\ y,\ y')\,dx,$$

where $F_1(x,\ y,\ y')=F(x,\ y,\ y')$ is on one side of the discontinuity line, and $F_2(x,\ y,\ y')=F(x,\ y,\ y')$ is on the other side.

Suppose that F_1 and F_2 are three times differentiable. It is natural to expect a corner point at the point C of the intersection of the desired curve with the discontinuity line. Arcs AC and CB are obviously extremals (this again follows from the fact that by fixing one of these arcs and varying the other alone we get a problem with fixed boundary points). For this reason, for the comparison curves we can take only polygonal lines consisting of two arcs of extremals, and then the variation, because of the movable nature of the boundary point $C(x_1,\ y_1)$ that is translated along the curve $y=\varphi(x)$, will take the following form (see page 345):

$$\begin{aligned}\delta v=\delta\int_{x_0}^{x_1} F_1(x,\ y,\ y')\,dx+\delta\int_{x_1}^{x_2} F_2(x,\ y,\ y')\,dx=\\ =[F_1+(\varphi'-y')F_{1y'}]_{x=x_1-0}\,\delta x_1-[F_2+(\varphi'-y')F_{2y'}]_{x=x_1+0}\,\delta x_1,\end{aligned}$$

and the basic necessary condition for an extremum, $\delta v=0$, reduces to the equality

$$[F_1+(\varphi'-y')F_{1y'}]_{x=x_1-0}=[F_2+(\varphi'-y')F_{2y'}]_{x=x_1+0}.$$

Since only y' can be discontinuous at the point of refraction, this refraction condition may be written as follows:

$$\begin{aligned}&F_1(x_1,\ y_1,\ y'(x_1-0))+\\ &\quad+(\varphi'(x_1)-y'(x_1-0))F_{1y'}(x_1,\ y_1,\ y'(x_1-0))=\\ &\qquad=F_2(x_1,\ y_1,\ y'(x_1+0))+\\ &\qquad\quad+(\varphi'(x_1)-y'(x_1+0))F_{2y'}(x_1,\ y_1,\ y'(x_1+0)).\end{aligned}$$

The refraction condition, together with the equation $y_1=\varphi(x_1)$, makes it possible to determine the coordinates of the point C.

If, in particular, the functional v is equal to

$$\begin{aligned}&\int_{x_0}^{x_2} A(x,\ y)\sqrt{1+y'^2}\,dx=\\ &=\int_{x_0}^{x_1} A_1(x,\ y)\sqrt{1+y'^2}\,dx+\int_{x_1}^{x_2} A_2(x,\ y)\sqrt{1+y'^2}\,dx,\end{aligned}$$

then the refraction condition takes the form

$$A_1(x, y)\frac{1+\varphi' y'}{\sqrt{1+y'^2}}\bigg|_{x=x_1-0} = A_2(x, y)\frac{1+\varphi' y'}{\sqrt{1+y'^2}}\bigg|_{x=x_1+0},$$

or, retaining the notation of pages 354-355, $y'(x_1-0)=\tan\beta_1$, $y'(x_1+0)=\tan\beta_2$, $\varphi'(x_1)=\tan\alpha$; after simplifying and multiplying by $\cos\alpha$ we get

$$\frac{\cos(\alpha-\beta_1)}{\cos(\alpha-\beta_2)}=\frac{A_2(x_1, y_1)}{A_1(x_1, y_1)} \quad \text{or} \quad \frac{\sin\left[\frac{\pi}{2}-(\alpha-\beta_1)\right]}{\sin\left[\frac{\pi}{2}-(\alpha-\beta_2)\right]}=\frac{A_2(x_1, y_1)}{A_1(x_1, y_1)},$$

which is a generalization of the familiar law of the refraction of light: the ratio of the sine of the angle of incidence to the sine of the angle of refraction is equal to the ratio of the velocities

$$v_1(x, y)=\frac{1}{A_1(x, y)} \quad \text{and} \quad v_2(x, y)=\frac{1}{A_2(x, y)} \quad \text{(compare page 354)}$$

in media, on the boundary between which the refraction occurs.

One should not think that extremals with corners only occur in problems of refraction or reflection of extremals. An extremum may be achieved on extremals with corner points even in extremum problems of the functional $v=\int_{x_0}^{x_1} F(x, y, y')\,dx$, where the function F is three times differentiable, and the permissible curves must pass through the boundary points A and B without any supplementary conditions whatsoever.

Let us investigate, say, the functional

$$v=\int_0^2 y'^2(1-y')^2\,dx, \quad y(0)=0; \quad y(2)=1.$$

Since the integrand is positive, $v \geqslant 0$ and hence if the functional $v=0$ on some curve, it follows that on this curve there will definitely be achieved the absolute minimum of the functional v, that is, the least value of the functional on the permissible curves. It will readily be seen that on the polygonal line $y=x$ for $0 \leqslant x \leqslant 1$ and $y=1$ for $1 < x \leqslant 2$ (Fig. 7.11), the functional $v=0$ since on this polygonal line the integrand is identically zero. Consequently, the absolute minimum of the functional is achieved on this polygonal line.

The absolute minimum of the functional $v=0$ is also reached on the polygonal lines depicted in Fig. 7.13. On the other hand it is clearly seen that on smooth curves the values of the functional are strictly greater than zero, though they may be made arbitrarily close to zero. Indeed, the integrand vanishes only when $y=x+C_1$ or for $y=C_2$, but the lines which are composed of segments of the straight lines of these families and which pass through the points $A(0, 0)$ and $B(2, 1)$, can only be polygonal lines. However, by smoothing the salient points by means of an appropriate variation of the function in an arbitrarily small neighbourhood of these points we can obtain a smooth curve, the value of the functional of which differs by an arbitrarily small value from the values of the functional on the polygonal line. Thus, $v=0$ is the greatest lower bound of the values of the functional v on smooth curves, but this greatest lower bound is not attained on smooth curves, it is attained on piecewise smooth curves.

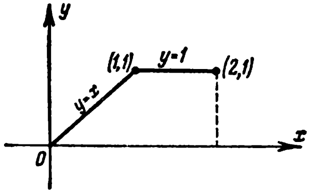


Fig. 7-11

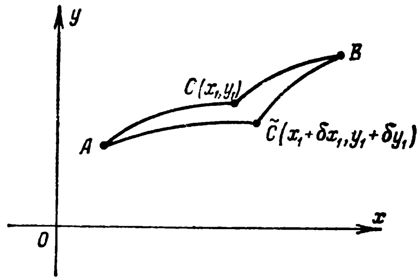


Fig. 7-12

Let us find the conditions that must be satisfied by solutions with corner points of the extremum problems of the functional $v[y(x)]=\int_{x_0}^{x_2} F(x, y, y')\,dx$. It is obvious that the separate smooth arcs which make up the broken-line extremal must be integral curves of the Euler equation. This follows from the fact that if all the segments (except one) of the polygonal line are fixed, then the problem reduces to the most elementary problem with fixed boundaries, and, hence, this segment must be an arc of the extremal.

Assuming that the broken-line extremal has only one corner point (this is to simplify the notation),* we find the conditions that must be satisfied at the corner point:

$$v=\int_{x_0}^{x_2} F(x, y, y')\,dx=\int_{x_0}^{x_1} F(x, y, y')\,dx+\int_{x_1}^{x_2} F(x, y, y')\,dx,$$

where x_1 is the abscissa of the corner point (Fig. 7.12). Taking it that the curves AC and CB are integral curves of the Euler equa-

* If there are several corner points, then the same argument applies to each one.

tion and that the point C can move in arbitrary fashion, we get (according to Sec. 1, page 344):

$$\delta v = (F - y'F_{y'})|_{x=x_1-0}\delta x_1 + \\ + F_{y'}|_{x=x_1-0}\,\delta y_1 - (F - y'F_{y'})|_{x=x_1+0}\,\delta x_1 - F_{y'}|_{x=x_1+0}\,\delta y_1 = 0,$$

whence

$$(F - y'F_{y'})|_{x=x_1-0}\,\delta x_1 + F_{y'}|_{x=x_1-0}\,\delta y_1 = \\ = (F - y'F_{y'})|_{x=x_1+0}\,\delta x_1 + F_{y'}|_{x=x_1+0}\,\delta y_1,$$

or, since δx_1 and δy_1 are independent, we have

$$(F - y'F_{y'})|_{x=x_1-0} = (F - y'F_{y'})|_{x=x_1+0}$$

and

$$F_{y'}|_{x=x_1-0} = F_{y'}|_{x=x_1+0}.$$

These conditions, together with the continuity conditions of the desired extremal, permit determining the coordinates of the corner point.

Example 1. Find the broken-line extremals (if they exist) of the functional $v = \int_0^a (y'^2 - y^2)\,dx$. Write the second of the conditions that must be fulfilled at the point of inflection, $F_{y'}|_{x=x_1-0} = F_{y'}|_{x=x_1+0}$ or, in the given case, $2y'(x_1 - 0) = 2y'(x_1 + 0)$, whence $y'(x_1 - 0) = y'(x_1 + 0)$; that is, the derivative y' at the point x_1 is continuous and there is no point of inflection. Hence, in the problem at hand an extremum may be reached only on smooth curves,

Example 2. Find the broken-line extremals of the functional $v = \int_{x_0}^{x_2} y'^2(1 - y')^2\,dx$. Since the integrand depends only on y', the extremals are the straight lines $y = Cx + \bar{C}$ (see page 314). In this case the conditions at the point of deflection take the form

$$-y'^2(1 - y')(1 - 3y')|_{x=x_1-0} = -y'^2(1 - y')(1 - 3y')|_{x=x_1+0}$$

and

$$2y'(1 - y')(1 - 2y')|_{x=x_1-0} = 2y'(1 - y')(1 - 2y')|_{x=x_1+0}.$$

These conditions, if one disregards the trivial possibility

$$y'(x_1 - 0) = y'(x_1 + 0),$$

are satisfied for

$$y'(x_1 - 0) = 0$$

and

$$y'(x_1 + 0) = 1$$

or

$$y'(x_1-0)=1$$

and

$$y'(x_1+0)=0.$$

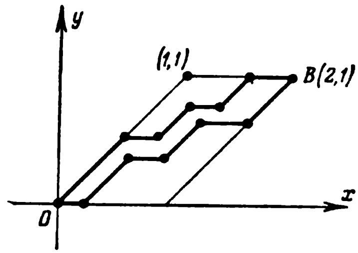


Fig. 7-13

Consequently, broken-line extremals may consist solely of segments of straight lines that belong to the families $y=C_1$ and $y=x+C_2$ (Fig. 7.13).

4. One-Sided Variations

In certain variational problems involving an extremum of a functional $v[y(x)]$, a restriction may be imposed on the class of permissible curves that prohibits them from passing through points of

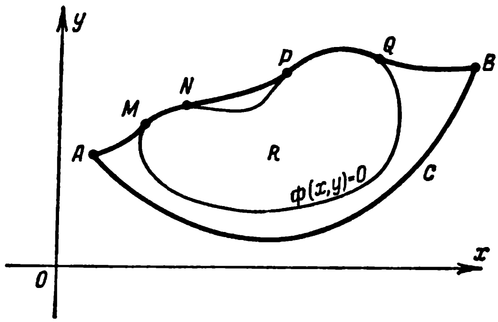


Fig. 7-14

a certain region R bounded by the curve $\Phi(x, y)=0$ (Fig 7.14). In these problems the extremizing curve C either passes completely outside the boundaries of the region R, and then it must be an extremal, since in this case the presence of the prohibited region R does not in the least affect the properties of the functional and its variations in the neighbourhood of the curve C, and the arguments in Chapter 6 hold true, or C consists of arcs lying outside the boundary of R and also consists of parts of the boundary of the region R. In this latter instance, a new situation arises; only one-sided variations of the curve C are possible on parts of the boundary of the region R, since permissible curves cannot enter the region. Parts of the curve C that lie outside the boundary of R must, as before, be extremals, since if we vary the curve C only on such a segment that permits two-sided variations, the presence of the region R will not affect the variations of y, and the conclusions of Chanter 6 continue to hold true.

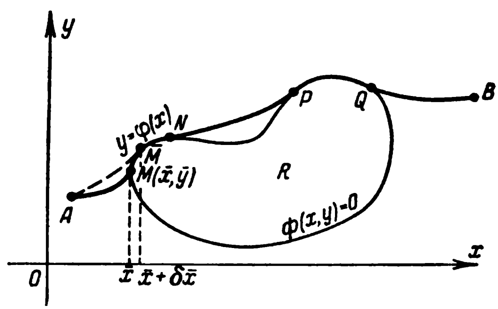


Fig. 7-15

Thus, in the problem under consideration, an extremum can be reached only on curves consisting of arcs of extremals and parts of the boundary of the region R, and hence, in order to construct the desired extremizing curve we have to obtain conditions, at the points of transition of the extremal to the boundary of the region R, which permit determining these points. In the case depicted in Fig. 7.15, it is necessary to obtain conditions at the points M, N, P and Q. Let us, for instance, obtain a condition at the point M. In quite analogous fashion one could obtain conditions also at other points of transition of the extremal to the boundary of the region.

When calculating the variation δv of the functional

$$v=\int_{x_0}^{x_1} F(x, y, y')\,dx=\int_{x_0}^{\bar{x}} F(x, y, y')\,dx+\int_{\bar{x}}^{x_1} F(x, y, y')\,dx$$

we can consider that the variation is caused solely by the displacement of the point $M(\bar{x}, \bar{y})$ on the curve $\Phi(x, y)=0$, i.e. it may be taken that for any position of the point M on the curve

$\Phi(x, y)=0$, the arc AM is already an extremal, and the segment $MNPQB$ does not vary. The functional

$$v_1=\int_{x_0}^{\bar{x}} F(x, y, y')\,dx$$

has a boundary point moving along the boundary of the region R, whose equation is $\Phi(x, y)=0$, or in the form solved for y in the neighbourhood of the point M: $y=\varphi(x)$.

Thus, according to Sec. 1 (page 344)

$$\delta v_1=[F+(\varphi'-y')F_{y'}]_{x=\bar{x}}\,\delta\bar{x}.$$

The functional $v_2=\int_{\bar{x}}^{x_1} F(x, y, y')\,dx$ also has a moving boundary point $(\bar{x}, \bar{y})$. However, in the neighbourhood of this point the curve on which an extremum $y=\varphi(x)$ can be achieved does not vary. Consequently, the variation of the functional v_2 in the translation of point $(\bar{x}, \bar{y})$ to the position $(\bar{x}+\delta\bar{x}, \bar{y}+\delta\bar{y})$ only reduces to a change in the lower limit of integration and

$$\Delta v_2=\int_{\bar{x}+\delta\bar{x}}^{x_1} F(x, y, y')\,dx-\int_{\bar{x}}^{x_1} F(x, y, y')\,dx=$$

$$=-\int_{\bar{x}}^{\bar{x}+\delta\bar{x}} F(x, y, y')\,dx=-\int_{\bar{x}}^{\bar{x}+\delta\bar{x}} F(x, \varphi(x), \varphi'(x))\,dx,$$

since $y=\varphi(x)$ on the interval $(\bar{x}, \bar{x}+\delta\bar{x})$.

Applying the mean-value theorem and taking advantage of the continuity of the function F, we get

$$\Delta v_2=-F(x, \varphi(x), \varphi'(x))\,|_{x=\bar{x}}\,\delta\bar{x}+\beta\cdot\delta\bar{x},$$

where $\delta\to 0$ as $\delta\bar{x}\to 0$.

Consequently, $\delta v_2=-F(x, \varphi(x), \varphi'(x))\,|_{x=\bar{x}}\,\delta\bar{x}$,

$$\delta v=\delta v_1+\delta v_2=[F(x, y, y')+$$
$$+(\varphi'-y')F_{y'}(x, y, y')]_{x=\bar{x}}\,\delta\bar{x}-F(x, y, \varphi')\,|_{x=\bar{x}}\,\delta\bar{x}=$$
$$=[F(x, y, y')-F(x, y, \varphi')-(y'-\varphi')F_{y'}(x, y, y')]_{x=\bar{x}}\,\delta\bar{x},$$

since $y(\bar{x})=\varphi(\bar{x})$.

Due to the arbitrary nature of $\delta\bar{x}$, the necessary condition for an extremum, $\delta v=0$, takes the form

$$[F(x, y, y')-F(x, y, \varphi')-(y'-\varphi')F_{y'}(x, y, y')]_{x=\bar{x}}=0.$$

Applying the mean-value theorem, we obtain

$$(y'-\varphi')[F_{y'}(x, y, q)-F_{y'}(x, y, y')]_{x=\bar{x}}=0,$$

where q is a value intermediate between $\varphi'(\bar{x})$ and $y'(\bar{x})$. Again applying the mean-value theorem, we get

$$(y'-\varphi')(q-y')F_{y'y'}(x, y, \bar{q})|_{x=\bar{x}}=0,$$

where $\bar{q}$ is a value intermediate between q and $y'(\bar{x})$.

Suppose $F_{y'y'}(x, y, \bar{q})\neq 0$. This supposition is natural for many variational problems (see Chapter 8). In this case the condition at the point M is of the form $y'(\bar{x})=\varphi'(\bar{x})$ ($q=y'$ only when $y'(\bar{x})=\varphi'(\bar{x})$, since q is a value intermediate between $y'(\bar{x})$ and $\varphi'(\bar{x})$).

Hence, at the point M the extremal AM and the boundary curve MN have a common tangent (the left tangent for the curve $y=y(x)$ and the right tangent for the curve $y=\varphi(x)$). Thus, *the extremal is tangent to the boundary of the region R at the point M.*

PROBLEMS ON CHAPTER 7

1. Find a solution with one corner point in the minimum problem of the functional

$$v[y(x)]=\int_0^4 (y'-1)^2(y'+1)^2\,dx;\quad y(0)=0;\quad y(4)=2.$$

2. Do solutions with corner points exist in the extremum problem of the functional

$$v[y(x)]=\int_x^{x_1} (y'^2+2xy-y^2)\,dx;\quad y(x_0)=y_0;\quad y(x_1)=y_1?$$

3. Are there any solutions with corner points in the extremum problem of the functional

$$v[y(x)]=\int_0^{x_1} (y'^4-6y'^2)\,dx;\quad y(0)=0;\quad y(x_1)=y_1?$$

4. Find the transversality condition for the functional

$$v[y(x)]=\int_{x_0}^{x_1} A(x, y)\,e^{\arctan y'}\sqrt{1+y'^2}\,dx,\quad A(x, y)\neq 0.$$

5. Using the basic necessary condition for an extremum, $\delta v=0$, find the function on which the functional can be extremized

$$v[y(x)]=\int_0^1 (y''^2-2xy)\,dx;\quad y(0)=y'(0)=0,$$

$$y(1)=\frac{1}{120};\quad y'(1) \text{ is not given.}$$

6. Find curves on which an extremum of the functional

$$v[y(x)]=\int_0^{10} y'^3\,dx;\quad y(0)=0;\quad y(10)=0$$

can be achieved, provided that the permissible curves cannot pass inside a circle bounded by the circumference

$$(x-5)^2+y^2=9$$

7. Find a function on which an extremum of the functional

$$v[y(x)]=\int_0^{\frac{\pi}{4}} (y^2-y'^2)\,dx;\quad y(0)=0$$

can be achieved if the other boundary point is permitted to slide along the straight line $x=\frac{\pi}{4}$.

8. Using only the basic necessary condition $\delta v=0$, find the curve on which an extremum of the functional

$$v[y(x)]=\int_0^{x_1} \frac{\sqrt{1+y'^2}}{y}\,dx;\quad y(0)=0$$

can be achieved if the second boundary point (x_1, y_1) can move along the circumference $(x-9)^2+y^2=9$.

Sufficient conditions for an extremum

1. Field of Extremals

If on an xy-plane, one and only one curve of the family $y=y(x, C)$ passes through every point of a certain region D, then we say that this family of curves forms a *field*, more precisely, a proper field, in the region D. The slope of the tangent line $p(x, y)$ to the curve of the family $y=y(x, C)$ that passes through the point (x, y) is called the *slope of the field* at the point (x, y).

For instance, inside the circle $x^2+y^2\leqslant 1$ the parallel straight lines $y=x+C$ form a field (Fig. 8.1), the slope of which is $p(x, y)=1$. On the contrary, the family of parabolas $y=(x-a)^2-1$ (Fig. 8.2) inside the same circle does not form a field since the parabolas of this family intersect inside the circle.

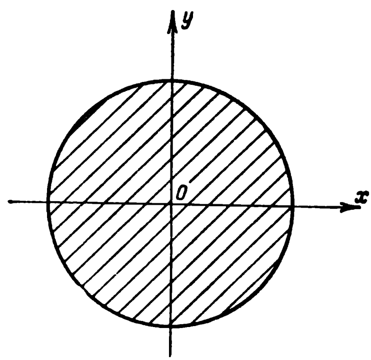


Fig. 8-1

If all the curves of the family $y=y(x, C)$ pass through a certain point (x_0, y_0), i.e. if they form a pencil of curves, then they definitely do not form a proper field in the region D, if the centre of the pencil belongs to D. However, if the curves of the pencil cover the entire region D and do not intersect anywhere in this region, with the exception of the centre of the pencil, i.e. the requirements imposed on the field are fulfilled at all points other than the centre of the pencil, then we say that the family $y=y(x, C)$ also forms a field, but in contrast to the proper field it is called a *central field* (Fig. 8.3).

For example, the pencil of sinusoids $y=C\sin x$ for $0\leqslant x\leqslant a$, $a<\pi$ forms a central field (Fig 8.4). The very same pencil of sinusoids forms a proper field in a sufficiently small neighbourhood of the segment of the axis of abscissas $\delta\leqslant x\leqslant a$, where $\delta>0$, $a<\pi$ (Fig. 8.4). The very same pencil of sinusoids does not form a field in the neighbourhood of the segment of the axis of abscissas $0\leqslant x\leqslant a_1$, $a_1>\pi$ (Fig. 8.4).

If a proper field or a central field is formed by a family of extremals of a certain variational problem, then it is called an *extremal field.*

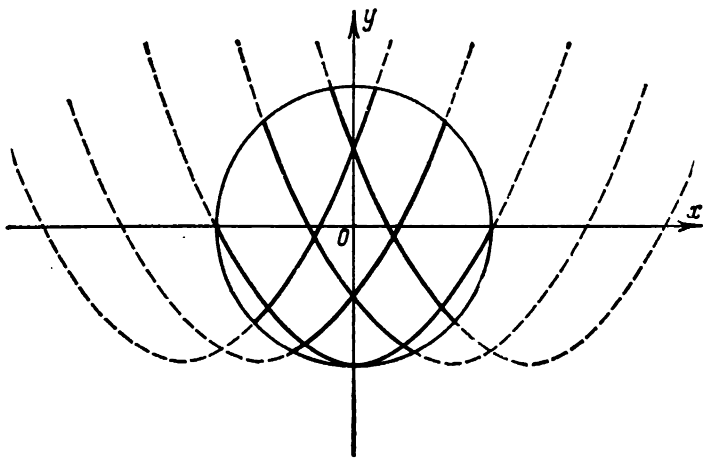


Fig. 8-2

The field concept is extended almost without any changes to the case of a space of any number of dimensions. The family $y_i = y_i(x, C_1, \ldots, C_n)$ $(i = 1, 2, \ldots, n)$ forms a field in the region D of the space $x, y_1, \ldots, y_n$ if through every point of D there passes

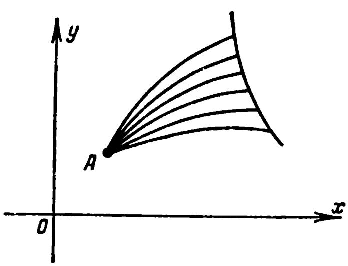


Fig. 8-3

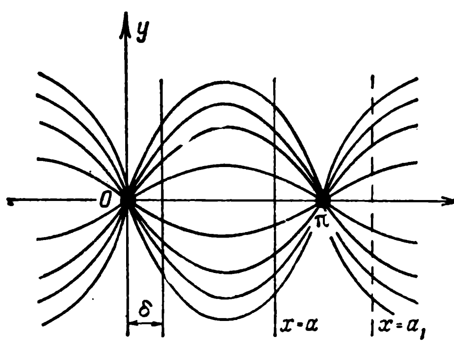


Fig. 8-4

one and only one curve of the family $y_i = y_i(x, C_1, \ldots, C_n)$. The partial derivatives of the functions $y_i(x, C_1, C_2, \ldots, C_n)$ with respect to x calculated at the point $(x, y_1, y_2, \ldots, y_n)$ are called *functions of the slope* of the field $p_i(x, y_1, y_2, \ldots, y_n)$ $(i = 1, 2, \ldots, n)$;

hence, to obtain $p_i(x, y_1, y_2, \ldots, y_n)$ one has to take $\frac{\partial}{\partial x} y_i(x, C_1, C_2, \ldots, C_n)$ and replace $C_1, C_2, \ldots, C_n$ by their expressions given in terms of the coordinates $x, y_1, y_2, \ldots, y_n$. The central field is defined in similar fashion.

Let the curve $y=y(x)$ be an extremal of a variational problem involving the extremum of an elementary functional

$$v[y(x)]=\int_{x_0}^{x_1} F(x, y, y')\,dx,$$

and let the boundary points $A(x_0, y_0)$ and $B(x_1, y_1)$ be fixed. We say that the extremal $y=y(x)$ is included in the extremal field if

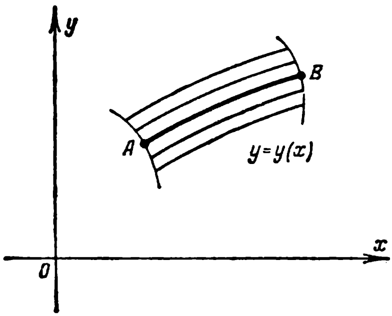


Fig. 8-5

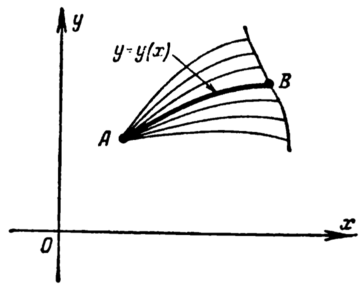


Fig. 8-6

a family of extremals $y=y(x, C)$ has been found that forms a field containing the extremal $y=y(x)$ for some value $C=C_0$, and the extremal $y=y(x)$ does not lie on the boundary of the region D in which the family $y=y(x, C)$ forms the field (Fig. 8.5). If a pencil of extremals centred at the point $A(x_0, y_0)$ forms a field in the neighbourhood of the extremal $y=y(x)$ that passes through this point, then the central field including the given extremal $y=y(x)$ has thus been found. In the given case, for the parameter of the family we can take the slope of the tangent line to the curves of the pencil at the point $A(x_0, y_0)$ (Fig. 8.6).

Example 1. Given the functional

$$\int_0^a (y'^2-y^2)\,dx;$$

it is required to include the arc of the extremal $y=0$ that connects the points (0, 0) and $(a, 0)$, where $0<a<\pi$, in the central field of extremals. The general solution of the Euler equation $y''+y=0$

(see page 310, Example 1) is of the form $y = C_1 \cos x + C_2 \sin x$. From the condition of the passage of extremals through the point (0, 0) we get $C_1 = 0$, $y = C_2 \sin x$, and the curves of this pencil form a central field on the interval $0 \leqslant x \leqslant a$, $a < \pi$ including, for $C_2 = 0$, the extremal $y = 0$. The parameter of the family C_2 is equal to the derivative y'_x at the point (0, 0). But if in the same problem $a \geqslant \pi$, then the family $y = C_2 \sin x$ does not form a field (see page 365).

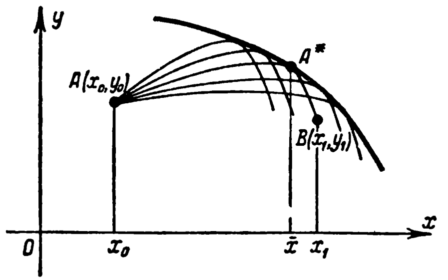


Fig. 8-7

It is known that two infinitely close curves of the family $F(x, y, C) = 0$ intersect at points of the C-discriminant curve defined by the equations

$$F(x, y, C) = 0; \qquad \frac{\partial F}{\partial C} = 0.$$

Recall that the C-discriminant curve includes, among other things, the envelope of the family and the locus of multiple points of the curves of the family. If $F(x, y, C) = 0$ is the equation of the pencil of curves, then the centre of the pencil likewise belongs to the C-discriminant curve. Therefore if we take a pencil of extremals $y = y(x, C)$ passing through the point (x_0, y_0) and determine ts C-discriminant curve $\Phi(x, y) = 0$, then close-lying curves of the family $y = y(x, C)$ will intersect near the curve $\Phi(x, y) = 0$ and, in particular, the curves of this family that are close to the extremal under consideration $y = y(x)$ that passes through the points $A(x_0, y_0)$ and $B(x_1, y_1)$, will intersect at points close to the points of tangency (or intersection) of the curve $y = y(x)$ with the C-discriminant curve (see Fig. 8.7, where the C-discriminant curve is depicted by a heavy line). If the arc AB of the extremal $y = y(x)$ does not have common points (different from A) with the C-discriminant curve of the pencil of extremals that includes the given extremal, then extremals of the pencil sufficiently close to the arc AB do not intersect; that is to say, they form, in the neighbourhood of the arc AB, a central field that includes this arc (Fig. 8.8).

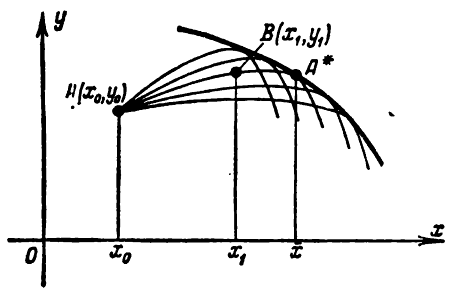


Fig. 8-8

If the arc AB of the extremal $y=y(x)$ has a point A^* (different from A) in common with the C-discriminant curve of the pencil $y=y(x, C)$, then curves of the pencil close to $y=y(x)$ can intersect among themselves and with the curve $y=y(x)$ near the point A^*; generally speaking, they do not form a field (Fig. 8.7). The point A^* is called a point that is *conjugate* to the point A.

The result obtained may be formulated as follows: *to construct a central field of extremals with centre at the point A, which field contains the arc AB of the extremal, it is sufficient for the point A^*, conjugate to A, not to lie on the arc AB.* This condition of the possibility of constructing a field of extremals including a given extremal is called the *Jacobi condition.*

This condition may be readily stated analytically as well. Let $y=y(x, C)$ be the equation of a pencil of extremals with centre at the point A, the parameter C being, for definiteness, considered as coinciding with the slope y' of the extremals of the pencil at the point A. The C-discriminant curve is defined by the equations

$$y=y(x, C); \quad \frac{\partial y(x, C)}{\partial C}=0.$$

Along every fixed curve of the family, the derivative $\frac{\partial y(x, C)}{\partial C}$ is a function of x alone. Denote this function briefly by u: $u=\frac{\partial y(x, C)}{\partial C}$, where C is given; whence $u'_x=\frac{\partial^2 y(x, C)}{\partial C\,\partial x}$. The functions $y=y(x, C)$ are solutions of the Euler equation. Therefore

$$F_y(x, y(x, C), y'_x(x, C))-\frac{d}{dx}F_{y'}(x, y(x, C), y'_x(x, C))\equiv 0.$$

Differentiating this identity with respect to C and putting $\frac{\partial y(x, C)}{\partial C}=u$, we get

$$F_{yy}u+F_{yy'}u'-\frac{d}{dx}(F_{yy'}u+F_{y'y'}u')=0$$

or

$$\left(F_{yy}-\frac{d}{dx}F_{yy'}\right)u-\frac{d}{dx}(F_{y'y'}u')=0.$$

Here, $F_{yy}(x, y, y')$, $F_{yy'}(x, y, y')$, $F_{y'y'}(x, y, y')$ are known functions of x, since the second argument y is equal to a solution of Euler's equation $y=y(x, C)$ taken for the value $C=C_0$ that corresponds to the extremal AB. This homogeneous linear equation of the second order in u is called *Jacobi's equation.*

If the solution of this equation $u=\frac{\partial y(x, C)}{\partial C}$, which vanishes at the centre of the pencil for $x=x_0$ (the centre of the pencil always belongs to the C-discriminant curve) also vanishes at some point of the interval $x_0 < x < x_1$, then the point conjugate to A defined by the equations

$$y = y(x, C_0) \quad \text{and} \quad \frac{\partial y(x, C)}{\partial C} = 0 \text{ or } u = 0$$

lies on the arc AB of the extremal.* But if there exists a solution of the Jacobi equation that vanishes for $x = x_0$ and that does not further vanish at any point of the interval $x_0 \leqslant x \leqslant x_1$, then there are no points conjugate to A on the arc AB: the Jacobi condition is fulfilled, and the arc AB of the extremal may be included in the central field of extremals centred at the point A.

Note. It may be proved that the Jacobi condition is necessary for achieving an extremum; i. e. for the extremizing curve AB the point conjugate to A cannot lie in the interval $x_0 < x < x_1$.

Example 2. Is the Jacobi condition fulfilled for the extremal of the functional $v = \int_0^a (y'^2 - y^2)\,dx$ that passes through the points $A(0, 0)$ and $B(a, 0)$?

The Jacobi equation has the form

$$-2u - \frac{d}{dx}(2u') = 0 \quad \text{or} \quad u'' + u = 0,$$

whence

$$u = C_1 \sin(x - C_2).$$

Since $u(0) = 0$, it follows that $C_2 = 0$; $u = C_1 \sin x$. The function u vanishes at the points $x = k\pi$, where k is an integer, and, hence, if $0 < a < \pi$, then on the interval $0 \leqslant x \leqslant a$ the function u vanishes only at the point $x = 0$ and the Jacobi condition is fulfilled. But if $a \geqslant \pi$, then on the interval $0 \leqslant x \leqslant a$ the function u vanishes in at least one more point $x = \pi$ and the Jacobi condition is not fulfilled (compare with Example 1, page 367).

Example 3. Is the Jacobi condition fulfilled for extremal of the functional

$$v[y(x)] = \int_0^a (y'^2 + y^2 + x^2)\,dx,$$

that passes through the points $A(0, 0)$ and $B(a, 0)$?

The Jacobi equation is of the form $u'' - u = 0$. We take its general solution in the form $u = C_1 \sinh x + C_2 \cosh x$. From the condition

* Note that all nontrivial solutions of the second-order homogeneous linear differential equation that satisfy the condition $u(x_0) = 0$ differ from one another solely in a constant nonzero factor and, hence, vanish simultaneously.

$u(0)=0$ we find $C_2=0$, $u=C_1 \sinh x$. The curves of the pencil $u=C_1 \sinh x$ intersect the x-axis only at the point $x=0$. The Jacobi condition is fulfilled for any a.

2. The Function $E(x, y, p, y')$

Suppose that in the most elementary problem involving an extre mum of the functional

$$v=\int_{x_0}^{x_1} F(x, y, y')\,dx;$$

$$y(x_0)=y_0, \quad y(x_1)=y_1$$

the Jacobi condition is fulfilled and, hence, the extremal C that passes through the points $A(x_0, y_0)$ and $B(x_1, y_1)$ can be included in the central field whose slope is $p(x, y)$ (Fig. 8.9).* To determine

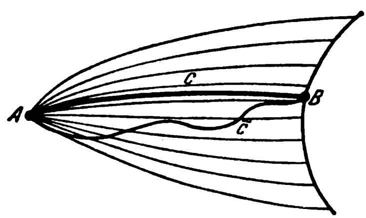


Fig. 8-9

the sign of the increment Δv of the functional v when passing from the extremal C to some close permissible curve $\tilde{C}$, we transform the increment $\Delta v=\int_{\tilde{C}} F(x, y, y')\,dx-\int_{C} F(x, y, y')\,dx$ to a form more convenient to investigate. (The symbols

$$\int_{\tilde{C}} F(x, y, y')\,dx \quad \text{and} \quad \int_{C} F(x, y, y')\,dx$$

represent values of the functional $v=\int_{x_0}^{x_1} F(x, y, y')\,dx$ taken along the arcs of the curves $\tilde{C}$ and C respectively.)

Consider the auxiliary functional

$$\int_{\tilde{C}} \left[F(x, y, p)+\left(\frac{dy}{dx}-p\right) F_p(x, y, p)\right] dx,$$

* It might be assumed that the extremal is included in the proper field and not in the central field.

which turns into $\int_C F(x, y, y')\,dx$ on the extremal C, since $\frac{dy}{dx}=p$ on extremals of the field. On the other hand, the very same auxiliary functional

$$\int_{\tilde{C}} \left[F(x, y, p)+\left(\frac{dy}{dx}-p\right) F_p(x, y, p)\right] dx$$

or

$$\int_{\tilde{C}} [F(x, y, p)-pF_p(x, y, p)]\,dx+F_p(x, y, p)\,dy \tag{8.1}$$

is the integral of an exact differential. Indeed, the differential of the function $\bar{v}(x, y)$, into which the functional $v[y(x)]$ is transformed on the extremals of the field, has the form, according to Sec. 1, Chapter 7 (page 344),

$$d\bar{v}=[F(x, y, y')-y'F_{y'}(x, y, y')]\,dx+F_{y'}(x, y, y')\,dy$$

and differs from the integrand in the auxiliary integral under consideration (8.1) solely in the designation of the slope of the tangent line to the extremals of the field.

Thus, on the extremal C the integral $\int_{\tilde{C}} [F(x, y, p)+(y'-p)F_p]\,dx$ coincides with the integral $\int_C F(x, y, y')\,dx$, and since the functional $\int_{\tilde{C}} [F(x, y, p)+(y'-p)F_p]\,dx$ is the integral of an exact differential and, hence, does not depend on the path of integration, it follows that

$$\int_C F(x, y, y')\,dx=\int_{\tilde{C}} [F(x, y, p)+(y'-p)F_p(x, y, p)]\,dx$$

not only for $\tilde{C}=C$ but for any choice of $\tilde{C}$ as well.

Hence, the increment

$$\Delta v=\int_{\tilde{C}} F(x, y, y')\,dx-\int_C F(x, y, y')\,dx$$

may be transformed to

$$\Delta v=\int_{\tilde{C}} F(x, y, y')\,dx-\int_{\tilde{C}} [F(x, y, p)+(y'-p)F_p(x, y, p)]\,dx=$$

$$=\int_{\tilde{C}} [F(x, y, y')-F(x, y, p)-(y'-p)F_p(x, y, p)]\,dx.$$

The integrand is called the *Weierstrass function* and is denoted by $E(x, y, p, y')$:

$$E(x, y, p, y') = F(x, y, y') - F(x, y, p) - (y' - p) F_p(x, y, p).$$

In this notation,

$$\Delta v = \int_{x_0}^{x_1} E(x, y, p, y')\,dx.$$

Obviously, a sufficient condition for the functional v to achieve a minimum on the curve C is the nonnegativity of the function E, since if $E \geqslant 0$, then also $\Delta v \geqslant 0$, and a sufficient condition for a maximum will be $E \leqslant 0$, since in this case $\Delta v \leqslant 0$ also. Here, it is sufficient for a weak minimum that the inequality $E(x, y, p, y') \geqslant 0$ (or $E \leqslant 0$ in the case of a maximum) be fulfilled for values of x, y close to the value of x, y on the extremal C under study, and for the values of y' close to $p(x, y)$ on the same extremal; for a strong minimum, the same inequality must hold for the same x, y, but now for arbitrary y', since, in the case of a strong extremum, close-lying curves may have arbitrary directions of tangent-lines and in the case of a weak extremum the values of y' on close-lying curves are close to the values of $y' = p$ on the extremal C.

Consequently, the following conditions will be sufficient for a functional v to achieve an extremum on the curve C:

For a weak extremum.

1. The curve C is an extremal that satisfies the boundary conditions.
2. The extremal C may be included in the field of extremals. This condition may be replaced by the Jacobi condition.
3. The function $E(x, y, p, y')$ does not change sign at any point (x, y) close to the curve C and for values of y' close to $p(x, y)$. In the case of a minimum, $E \geqslant 0$, in the case of a maximum, $E \leqslant 0$.

For a strong extremum:

1. The curve C is an extremal satisfying the boundary conditions.
2. The extremal C may be included in the field of extremals. This condition may be replaced by the Jacobi condition.
3. The function $E(x, y, p, y')$ does not change sign at any of the points (x, y) that are close to the curve C and for arbitrary values of y'. In the case of a minimum, $E \geqslant 0$; in the case of a maximum, $E \leqslant 0$.

Note. It may be proved that the Weierstrass condition is necessary. More precisely, if in a central field including the extremal C, the function E has opposite signs at points of the extremal for certain y', then a strong extremum is not achieved. If this property occurs for values of y' arbitrarily close to p, then even a weak extremum is not achieved.

Example 1. Test for an extremum the functional

$$v=\int_0^a y'^3\,dx;\quad y(0)=0,$$

$$y(a)=b,\ a>0,\ b>0.$$

The straight lines $y=C_1x+C_2$ are extremals. An extremum may be achieved only on the straight line $y=\frac{b}{a}x$. The pencil of straight lines $y=C_1x$ centred at the point (0, 0) forms a central field that includes the extremal $y=\frac{b}{a}x$ (Fig. 8.10).

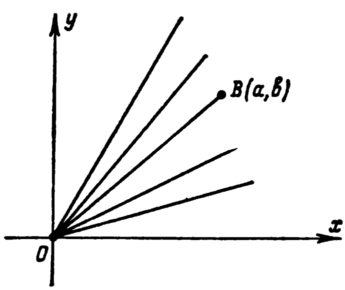


Fig. 8-10

The function

$$E(x,\ y,\ p,\ y')=y'^3-p^3-$$
$$-3p^2(y'-p)=(y'-p)^2\,(y'+2p).$$

On the extremal $y=\frac{b}{a}x$, the slope of the field $p=\frac{b}{a}>0$, and if y' assumes values close to $p=\frac{b}{a}$, then $E\geqslant 0$, and, hence, all the conditions that are sufficient for achieving a weak minimum are fulfilled. Thus, a weak minimum is achieved on the extremal $y=\frac{b}{a}x$. But if y' takes on arbitrary values, then $(y'+2p)$ may have any sign and, hence, the function E changes sign, and the conditions sufficient for achieving a strong minimum are not fulfilled. If we take into account the note on pages 373-374, it will be possible to assert that a strong minimum is not achieved on the straight line $y=\frac{b}{a}x$.

Example 2. Test for an extremum the functional

$$\int_0^a (6y'^2-y'^4+yy')\,dx;\quad y(0)=0;\quad y(a)=b;\quad a>0\ \text{ and }\ b>0$$

in the class of continuous functions with continuous first derivative.

The extremals are the straight lines $y=C_1x+C_2$. The boundary conditions are satisfied by the straight line $y=\frac{b}{a}x$, which is included in the pencil of extremals $y=C_1x$ that form the central field. The function

$$E(x, y, p, y')=6y'^2-y'^4+yy'-6p^2+p^4-yp-(y'-p)\times$$
$$\times(12p-4p^3+y)=-(y'-p)^2[y'^2+2py'-(6-3p^2)].$$

The sign of the function E is opposite that of the last factor

$$y'^2+2py'-(6-3p^2).$$

This factor vanishes and can change sign only when y' passes through the value $y'=-p\pm\sqrt{6-2p^2}$. For $6-2p^2\leqslant 0$, or $p\geqslant\sqrt{3}$ for any y' we have $[y'^2+2py'-(6-3p^2)]\geqslant 0$ but if $6-2p^2>0$ or $p<\sqrt{3}$, then the expression $[y'^2+2py'-(6-3p^2)]$ changes sign. But if, in the process, y' differs by a sufficiently small amount from p, then the latter expression does not change its positive sign for $p>1$ and its negative sign for $p<1$.

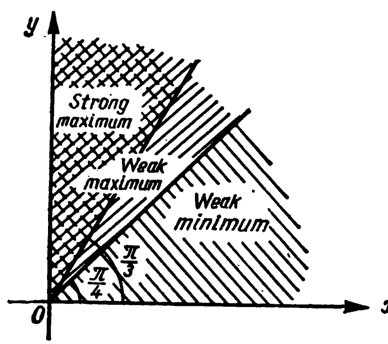


Fig. 8-11

Consequently, for $p=\frac{b}{a}<1$ or $b<a$, we have a weak minimum, since $E\geqslant 0$ for values of y' close to p; for $p=\frac{b}{a}>1$ or $b>a$ we have a weak maximum. For $p=\frac{b}{a}\geqslant\sqrt{3}$ we have a strong maximum since $E\leqslant 0$ for any values of y'. For $p=\frac{b}{a}<\sqrt{3}$, on the basis of the note on pages 373-374, there is neither a strong minimum nor a strong maximum (Fig. 8.11).

Even in the above very simple examples, testing the sign of the function E involved certain difficulties and for this reason it is advisable to replace the condition of retaining the sign by the function E by a more readily verifiable condition. Suppose that the function $F(x, y, y')$ is three times differentiable with respect to the argument y' By Taylor's formula we get

$$F(x, y, y')=$$
$$=F(x, y, p)+(y'-p)F_p(x, y, p)+\frac{(y'-p)^2}{2!}F_{y'y'}(x, y, q),$$

where q lies between p and y'.

The function

$$E(x, y, p, y')=F(x, y, y')-F(x, y, p)-(y'-p)F_p(x, y, p),$$

after replacement of the function $F(x, y, y')$ by its Taylor expansion, takes the form

$$E(x, y, p, y') = \frac{(y'-p)^2}{2!} F_{y'y'}(x, y, q).$$

From this we see that the function E does not change sign if $F_{y'y'}(x, y, q)$ doesn't. When investigating for a weak extremum, the function $F_{y'y'}(x, y, q)$ must retain sign for values of x and y at points close to the points of the extremal under study, and for values of q close to p. If $F_{y'y'}(x, y, y') \neq 0$ at points of the extremal C, then by virtue of continuity this second derivative maintains sign both at points close to the curve C and for values of y' close to the values of y' on the curve C. Thus, when testing for a weak minimum, the condition $E \geqslant 0$ may be replaced by the condition $F_{y'y'} < 0$ on the extremal C, and when testing for a weak maximum the condition $E \leqslant 0$ may be replaced by the condition $F_{y'y'} < 0$ on the curve C. The condition $F_{y'y'} > 0$ (or $F_{y'y'} < 0$) is called the *Legendre condition.* *

When testing for a strong minimum the condition $E \geqslant 0$ may be replaced by the requirement $F_{y'y'}(x, y, q) \geqslant 0$ at points (x, y) close to points of the curve C for arbitrary values of q. Here, of course, it is assumed that the Taylor expansion

$$F(x, y, y') = \\ = F(x, y, p) + (y'-p) F_p(x, y, p) + \frac{(y'-p)^2}{2!} F_{y'y'}(x, y, q)$$

holds true for any y'. When testing for a strong maximum, we get the condition $F_{y'y'}(x, y, q) \leqslant 0$, for the very same assumptions regarding the range of the arguments and the expansibility of the function $F(x, y, y')$ in Taylor's series.

Example 3. Test for an extremum the functional

$$v[y(x)] = \int_0^a (y'^2 - y^2)\, dx, \ a > 0; \ y(0) = 0, \ y(a) = 0.$$

The Euler equation has the form $y'' + y = 0$, its general solution is $y = C_1 \cos x + C_2 \sin x$. Using the boundary conditions, we get $C_1 = 0$ and $C_2 = 0$, if $a \neq k\pi$, where k is an integer.

Thus, for $a \neq k\pi$ an extremum may be achieved only on the straight line $y = 0$. If $a < \pi$, then the pencil of extremals $y = C_1 \sin x$ with centre at the point (0, 0) forms a central field. For $a > \pi$, the Jacobi condition is not fulfilled (see page 365).

---

* The condition $F_{y'y'} > 0$ (or $F_{y'y'} < 0$) is often called the *strong Legendre condition*, while the Legendre condition is the inequality $F_{y'y'} \geqslant 0$ (or $F_{y'y'} \leqslant 0$).

Summary of Sufficient Conditions for a Minimum of the Elementary Functional *

$$v[y(x)]=\int_{x_0}^{x_1} F(x,\ y,\ y')\,dx;\quad y(x_0)=y_0,\quad y(x_1)=y_1$$

Weak minimum	Strong minimum	Weak minimum	Strong minimum	Weak minimum	Strong minimum
1. $F_y-\frac{d}{dx}F_{y'}=0$	1. $F_y-\frac{d}{dx}F_{y'}=0$	1. $F_y-\frac{d}{dx}F_{y'}=0$	1. $F_y-\frac{d}{dx}F_{y'}=0$	1. $F_y-\frac{a}{dx}F_{y'}=0$	1. $F_y-\frac{d}{dx}F_{y'}=0$
2. Jacobi condition	2. Jacobi condition	2. Jacobi condition	2. Jacobi condition	2. An extremal field exists that includes the given extremal	2. An extremal field exists that includes the given extremal
3. $F_{y'y'}>0$ on the extremal under study	3. $F_{y'y'}(x,y,y')\geqslant 0$ for points $(x,\ y)$ close to points on the extremal under study and for arbitrary values of y'. It is here assumed that the function $F(x,\ y,\ y')$ is three times differentiable with respect to y' for any values of y'	3. $E(x,y,p,y')\geqslant 0$ for points (x,y) close to points on the extremal under study and for y' close to $p(x,\ y)$	3. $E(x,y,p,y')\geqslant 0$ for points $(x,\ y)$ close to points on the extremal under study and for arbitra y y'.	3. $E(x,\ y,\ p,\ y')\geqslant 0$ for points $(x,\ y)$ close to the points on the extremal under study and for y' close to $p(x,\ y)$	3. $E(x,y,p,y')\geqslant 0$ for points $(x,\ y)$ close to the points on the extremal under study, and for arbitrary y'

* To obtain the sufficient conditions for a *maximum*, take the inequalities, given here, in the opposite sense.

Since the integrand is three times differentiable with respect to y' for any values of y' and $F_{y'y'}=2>0$ for any values of y', it follows that on the straight line $y=0$ a strong minimum is achieved for $a<\pi$. If we take into account the note on page 370, it may be asserted that for $a>\pi$ a minimum is not achieved on the straight line $y=0$.

Example 4. Test for an extremum the functional

$$v[y(x)]=\int_0^{x_1}\frac{\sqrt{1+y'^2}}{\sqrt{y}}dx, \quad y(0)=0, \quad y(x_1)=y_1$$

(see the problem of the brachistochrone, pages 316-317). The extremals are the cycloids

$$x=C_1(t-\sin t)+C_2,$$
$$y=C_1(1-\cos t).$$

The pencil of cycloids $x=C_1(t-\sin t)$, $y=C_1(1-\cos t)$ with centre

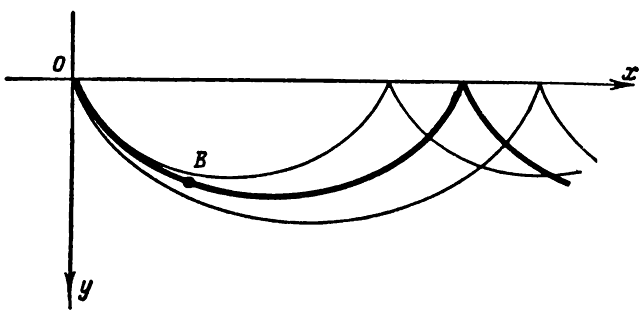


Fig. 8-12

in the point (0, 0) forms a central field including the extremal

$$x=a(t-\sin t),\ y=a(1-\cos t),$$

where a is determined from the condition of the passage of a cycloid through the second boundary point $B(x_1, y_1)$, if $x_1<2\pi a$ (Fig. 8.12).

We have

$$F_{y'}=\frac{y'}{\sqrt{y}\sqrt{1+y'^2}}; \qquad F_{y'y'}=\frac{1}{\sqrt{y}(1+y'^2)^{3/2}}>0$$

for any y'. Hence, for $x_1<2\pi a$, a strong minimum is achieved on the cycloid

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

Example 5. Test for an extremum the functional

$$v[y(x)]=\int_0^a y'^3\,dx;\quad y(0)=0,\ y(a)=b,\ a>0,\ b>0.$$

This example was solved on page 374, but we can now simplify the investigation with respect to a weak extremum.

The extremals are straight lines. The pencil $y=Cx$ forms a central field that includes the extremal $y=\frac{b}{a}x$. On the extremal $y=\frac{b}{a}x$, the second derivative $F_{y'y'}=6y'=6\frac{b}{a}>0$. Hence, the straight line $y=\frac{b}{a}x$ achieves a weak minimum. For arbitrary y', the second derivative $F_{y'y'}=6y'$ changes sign; thus the above-indicated sufficient conditions for achieving a strong minimum are not fulfilled. However, one cannot conclude from this that a strong extremum is not achieved.

Example 6. Test for an extremum the functional

$$v[y(x)]=\int_0^a \frac{y}{y'^2}\,dx;\quad y(0)=1,\ y(a)=b,\ a>0,\ 0<b<1.$$

The first integral of the Euler equation (see case (5), page 315) is of the form

$$\frac{y}{y'^2}+y'\frac{2y}{y'^3}=C\quad\text{or}\quad y'^2=4C_1y;$$

extracting the root, separating the variables and integrating, we get $y=(C_1x+C_2)^2$, which is a family of parabolas. From the condition $y(0)=1$ we find $C_2=1$. The pencil of parabolas $y=(C_1x+1)^2$ with centre in the point $A(0,\ 1)$ has a C_1-discriminant curve $y=0$ (Fig. 8.13). Two parabolas of this pencil pass through the point $B(a,\ b)$. On the arc AB of one of them (L_1) lies the point A^*, which is conjugate to the point A, on the other (L_2) there is no conjugate point and, hence, the Jacobi condition is fulfilled on the arc L_2, and an extremum can be achieved on this arc of the parabola. In the neighbourhood of the extremal under study $F_{y'y'}=\frac{6y}{y'^4}>0$ for arbitrary y'; however, on this basis we cannot assert that a strong minimum is achieved on the arc L_2, since the function $F(x,\ y,\ y')=\frac{y}{y'^2}$ cannot be represented in the form

$$F(x,\ y,\ y')=F(x,\ y,\ p)+(y'-p)F_p(x,\ y,\ p)+\frac{(y'-p)^2}{2!}F_{y'y'}(x,\ y,\ q)$$

for arbitrary values of y' due to the presence of a discontinuity of the function $F(x, y, y')$ when $y'=0$. One can only assert that a weak minimum is achieved on L_2, since for values of y' close to the slope of the field on the curve L_2 we have an expansion of the function $F(x, y, y')$ by Taylor's formula. A full investigation of this function for an extremum involves considering the function $E(x, y, p, y')$:

$$E(x, y, p, y') = \frac{y}{y'^2} - \frac{y}{p^2} + \frac{2y}{p^3}(y'-p) = \frac{y(y'-p)^2(2y'+p)}{y'^2 p^3}.$$

Since the factor $(2y'+p)$ changes sign for arbitrary y', on the basis of the note on pages 373-374 we can assert that a strong minimum is not achieved on the arc L_2.

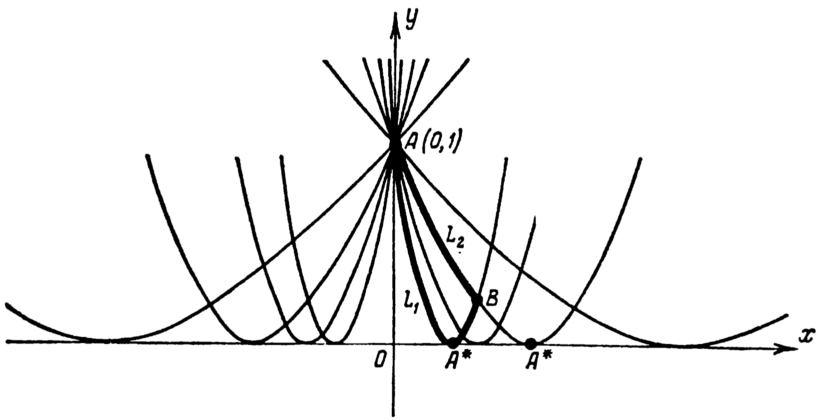


Fig. 8-13

The foregoing theorem can, without substantial modifications, be extended to functionals of the form

$$v[y_1, y_2, \ldots, y_n] = \int_{x}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx;$$

$$y_i(x_0) = y_{i0}, \quad y_i(x_1) = y_{i1} \quad (i = 1, 2, \ldots, n).$$

The function E takes the form

$$E = F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n') - \\ - F(x, y_1, y_2, \ldots, y_n, p_1, p_2, \ldots, p_n) - \\ - \sum_{i=1}^{n} (y_i' - p_i) F_{p_i}(x, y_1, y_2, \ldots, y_n, p_1, p_2, \ldots, p_n),$$

where the p_i are functions of the slope of the field, on which certain restrictions are imposed (under these restrictions it is called a special field).

The Legendre condition $F_{y'y'} \geqslant 0$ is replaced by the following conditions:

$$F_{y_1'y_1'} \geqslant 0, \quad \begin{vmatrix} F_{y_1'y_1'} & F_{y_1'y_2'} \\ F_{y_2'y_1'} & F_{y_2'y_2'} \end{vmatrix} \geqslant 0, \ldots, \begin{vmatrix} F_{y_1'y_1'} & F_{y_1'y_2'} & \ldots & F_{y_1'y_n'} \\ F_{y_2'y_1'} & F_{y_2'y_2'} & \ldots & F_{y_2'y_n'} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ F_{y_n'y_1'} & F_{y_n'y_2'} & \ldots & F_{y_n'y_n'} \end{vmatrix} \geqslant 0.$$

Both in the elementary problem and in more complicated problems, the sufficient conditions for a *weak minimum* may be obtained by a different method based on a study of the sign of the second variation.

Using the Taylor formula, transform the increment of the function in the elementary problem to the following form:

$$\Delta v = \int_{x_0}^{x_1} [F(x, y+\delta y, y'+\delta y') - F(x, y, y')]\,dx =$$

$$= \int_{x_0}^{x_1} (F_y\,\delta y + F_{y'}\,\delta y')\,dx + \frac{1}{2}\int_{x_0}^{x_1} [F_{yy}\delta y^2 + 2F_{yy'}\,\delta y\,\delta y' + F_{y'y'}\,\delta y'^2]\,dx + R,$$

where R is of order higher than second in δy and $\delta y'$. When investigating for a weak extremum, δy and $\delta y'$ are sufficiently small, and in this case the sign of the increment Δv is determined by the sign of the term on the right containing the lowest degrees of δy and $\delta y'$. On the extremal, the first variation

$$\int_{x_0}^{x_1} (F_y\,\delta y + F_{y'}\delta y')\,dx = 0$$

and, hence, the sign of the increment Δv, generally speaking, coincides with the sign of the second variation

$$\delta^2 v = \int_{x_0}^{x_1} (F_{yy}\,\delta y^2 + 2F_{yy'}\,\delta y\,\delta y' + F_{y'y'}\,\delta y'^2)\,dx.$$

The Legendre condition and the Jacobi condition together are the conditions that ensure constancy of sign of the second variation, and thus also constancy of sign of the increment Δv in the problem involving a weak extremum.

Indeed, consider the integral

$$\int_{x_0}^{x_1} [\omega'(x)\,\delta y^2 + 2\omega(x)\,\delta y\,\delta y']\,dx, \tag{8.2}$$

where $\omega(x)$ is an arbitrary differentiable function. This integral is equal to zero:

$$\int_{x_0}^{x_1} [(\omega'(x)\,\delta y^2 + 2\omega(x)\,\delta y\,\delta y']\,dx = \int_{x_0}^{x_1} d(\omega\,\delta y^2)\,dx = [\omega(x)\,\delta y^2]_{x_0}^{x_1} = 0$$

(because $\delta y|_{x_0} = \delta y|_{x_1} = 0$).

Adding the integral (8.2) to the second variation, we get

$$\delta^2 v = \int_{x_0}^{x_1} [(F_{yy} + \omega')\,\delta y^2 + 2(F_{yy'} + \omega)\,\delta y\,\delta y' + F_{y'y'}\,\delta y'^2]\,dx.$$

Choose the function $\omega(x)$ so that the integrand, to within a factor, is transformed into a perfect square, for which purpose the function $\omega(x)$ must satisfy the equation

$$F_{y'y'}(F_{yy} + \omega') - (F_{yy'} + \omega)^2 = 0.$$

For such a choice of the function ω, the second variation takes the form

$$\delta^2 v = \int_{x_0}^{x_1} F_{y'y'}\left(\delta y' + \frac{F_{yy'} + \omega}{F_{y'y'}}\,\delta y\right)^2 dx$$

and, consequently, the sign of the second variation coincides with that of $F_{y'y'}$.

However, such a transformation is possible solely on the assumption that the differential equation

$$F_{y'y'}(\omega' + F_{yy}) - (F_{yy'} + \omega)^2 = 0$$

has a differentiable solution $\omega(x)$ on the interval (x_0, x_1).

Transforming this equation to new variables by the substitution

$$\omega = -F_{yy'} - F_{y'y'}\frac{u'}{u},$$

where u is a new unknown function, we get

$$\left(F_{yy} - \frac{d}{dx}F_{yy'}\right)u - \frac{d}{dx}(F_{y'y'}u') = 0,$$

which is Jacobi's equation (see page 369).

If a solution of this equation exists that does not vanish for $x_0 < x \leqslant x_1$, i.e. if the Jacobi condition is fulfilled, then for the same values of x there exists a continuous and differentiable solution

$$\omega(x) = -F_{yy} - F_{y'y'}\frac{u'}{u}$$

of the equation

$$F_{y'y'}(F_{yy}+\omega') - (F_{yy'}+\omega)^2 = 0.$$

Thus, the Legendre condition and the Jacobi condition guarantee that the sign of the second variation does not change and, hence, they are sufficient conditions for a weak minimum ($F_{y'y'} > 0$) or maximum ($F_{y'y'} < 0$).

3. Transforming the Euler Equations to the Canonical Form

A system of n Euler equations (see page 318)

$$F_{y_i} - \frac{d}{dx}F_{y_i'} = 0 \qquad (i = 1, 2, \ldots, n) \tag{8.3}$$

may be replaced by a system of $2n$ first-order equations. Putting in (8.3)

$$F_{y_k'} = q_k \qquad (k = 1, 2, \ldots, n), \tag{8.4}$$

we obtain

$$\frac{dq_k}{dx} = \frac{\partial F}{\partial y_k} \qquad (k = 1, 2, \ldots, n). \tag{8.5}$$

Solve the system of equations (8.4) for y_k' (to make such a solution possible, assume that

$$\frac{D\left(F_{y_1'}, F_{y_2'}, \ldots, F_{y_n'}\right)}{D(y_1' y_2', \ldots, y_n')} \neq 0), \qquad y_k' = \omega_k(x, y_s, q_s), \tag{8.6}$$

where

$$\omega_k(x, y_s, q_s) = \omega_k(x, y_1, y_2, \ldots, y_n, q_1, q_2, \ldots, q_n),$$

and substitute (8.6) into (8.5). We then get a system of $2n$ first-order equations in the normal form:

$$\left.\begin{aligned} \frac{dy_k}{dx} &= \omega_k(x, y_s, q_s), \\ \frac{dq_k}{dx} &= \left\{\frac{\partial F}{\partial y_k}\right\}. \end{aligned}\right\} \tag{8.7}$$

Here and henceforward the braces signify that in place of y_k' they contain $\omega_k(x, y_s, q_s)$.

With the aid of the function

$$H(x, y_s, q_s) = \sum_1^n \omega_i q_i - \{F\}$$

the system (8.7) may be written in the canonical form:

$$\left.\begin{aligned}\frac{dy_k}{dx}&=\frac{\partial H}{\partial q_k},\\ \frac{dq_k}{dx}&=-\frac{\partial H}{\partial y_k}\end{aligned}\right\}\quad (k=1,\ 2,\ \ldots,\ n). \tag{8.8}$$

Note that if the function $F(y_1,\ y_2,\ \ldots,\ y_n,\ y_1',\ y_2',\ \ldots,\ y_n')$ does not depend explicitly on x, then the system (8.8) has a first integral $H=C$. Indeed, in this case

$$H=\sum_{i=1}^{n}\omega_i q_i-\{F\}$$

does not contain x explicitly either and, hence,

$$\frac{dH}{dx}=\sum_{i=1}^{n}\frac{\partial H}{\partial y_i}\frac{dy_i}{dx}+\sum_{i=1}^{n}\frac{\partial H}{\partial q_i}\frac{dq_i}{dx}.$$

By virtue of the equations (8.8) we get

$$\frac{dH}{dx}=0,\quad H=C$$

along the integral curves of the system (8.8).

This first integral was already obtained on page 316 for the elementary problem.

Example 1. *Law of conservation of energy.* The function

$$H=\sum_{i=1}^{n}\omega_i q_i-\{F\}$$

for the functional

$$\int_{t_0}^{t_1}(T-U)\,dt,\quad T=\sum_{i=1}^{n}m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2),$$

where the notation is that of Example 1, page 333 (T is the kinetic energy of a system of particles and, U is the potential energy), is of the following form:

$$H=\sum_{i=1}^{n}m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2)-(T-U)=T+U$$

is the total energy of the system. Apply the principle of least action. If the potential energy U does not depend explicitly on t, i.e. the system is conservative, then Euler's equations for the functional $\int_{t_0}^{t_1}(T-U)\,dt$ have a first integral $H=C$, $T+U=C$.

Thus, the total energy of a conservative system remains constant when the system is in motion.

Integration of the canonical system (8.8) is equivalent to integration of the partial differential equation

$$\frac{\partial v}{\partial x}+H\left(x,\ y_s,\ \frac{\partial v}{\partial y_s}\right)=0, \tag{8.9}$$

where

$$H\left(x,\ y_s,\ \frac{\partial v}{\partial y_s}\right)=H\left(x,\ y_1,\ y_2,\ \ldots,\ y_n,\ \frac{\partial v}{\partial y_1},\ \frac{\partial v}{\partial y_2},\ \ldots,\frac{\partial v}{\partial y_n}\right).$$

Equation (8.9) is called the *Hamilton-Jacobi equation*.

If a one-parameter family of its solutions $v(x,\ y_s,\ \alpha)$ is known, then the first integral $\frac{\partial v}{\partial \alpha}=\beta$ of the system (8.8) is known; β is an arbitrary constant. Indeed,

$$\frac{d}{dx}\left(\frac{\partial v}{\partial \alpha}\right)=\frac{\partial^2 v}{\partial x\,\partial \alpha}+\sum_{i=1}^{n}\frac{\partial^2 v}{\partial y_i\,\partial \alpha}\frac{\partial y_i}{\partial x}=\frac{\partial^2 v}{\partial x\,\partial \alpha}+\sum_{i=1}^{n}\frac{\partial^2 v}{\partial y_i\,\partial \alpha}\frac{\partial H}{\partial q_i}. \tag{8.10}$$

Differentiating the identity

$$\frac{\partial v(x,\ y_s,\ \alpha)}{\partial \alpha}\equiv -H\left(x,\ y_s,\ \frac{\partial v(x,\ y_s,\ \alpha)}{\partial y_s}\right) \text{ with respect to } \alpha,$$

we get

$$\frac{\partial^2 v}{\partial x\,\partial \alpha}\equiv -\sum_{s=1}^{n}\frac{\partial H}{\partial q_s}\frac{\partial^2 v}{\partial y_s\,\partial \alpha} \tag{8.11}$$

and, substituting (8.11) into (8.10), we get an identical zero on the right of (8.10). Thus

$$\frac{d}{dx}\left(\frac{\partial v}{\partial \alpha}\right)\equiv 0,$$

whence

$$\frac{\partial v}{\partial \alpha}=\beta.$$

Hence, if the complete integral of the Hamilton-Jacobi equation

$$v=v(x,\ y_1,\ y_2,\ \ldots,\ y_n,\ \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)$$

is known, then we also know the n first integrals of the system (8.8):

$$\frac{\partial v}{\partial \alpha_i}=\beta_i \qquad (i=1,\ 2,\ \ldots,\ n). \tag{8.12}$$

If the Jacobian of the system (8.12) is nonzero

$$\left|\frac{\partial^2 v}{\partial y_j\,\partial \alpha_i}\right|=0,$$

then the system (8.12) defines the y_i as functions of the remaining arguments:

$$y_i = y_i(x, \alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n) \quad (i = 1, 2, \ldots, n) \tag{8.13}$$

We have thus obtained a $2n$-parameter family of extremals. It may be proved that (8.13) is the general solution of the system of Euler's equations, and the functions

$$y_i(x\ \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)$$

and

$$q_i = \frac{\partial v(x, y_s, \alpha_s)}{\partial y_i} \qquad (i = 1, 2, \ldots, n)$$

are the general solution of the system (8.8)

Example. Find the equation of geodesics on a surface on which the element of length of the curve is of the form

$$ds^2 = [\varphi_1(x) + \varphi_2(y)](dx^2 + dy^2),$$

that is, find the extremals of the functional

$$S = \int_{x_0} \sqrt{[\varphi_1(x) + \varphi_2(y)](1 + y'^2)}\, dx$$

Since

$$H = \frac{\sqrt{\varphi_1(x) + \varphi_2(y)}}{\sqrt{1 + y'^2}} = \sqrt{\varphi_1(x) + \varphi_2(y)} \cdot \sqrt{1 - q^2},$$

$$q = \frac{y'}{\sqrt{1 + y'^2}}, \quad H^2 + q^2 = \varphi_1(x) + \varphi_2(y),$$

it follows that the Hamilton-Jacobi equation has the form

$$\left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 = \varphi_1(x) + \varphi_2(y)$$

or

$$\left(\frac{\partial v}{\partial x}\right)^2 - \varphi_1(x) = \varphi_2(y) - \left(\frac{\partial v}{\partial y}\right)^2.$$

For equations of this type (equations with separated variables)

$$\Phi_1\left(x, \frac{\partial v}{\partial x}\right) = \Phi_2\left(y, \frac{\partial v}{\partial y}\right)$$

the first integral is easily found. Putting

$$\left(\frac{\partial v}{\partial x}\right)^2 - \varphi_1(x) = \alpha \quad \text{and} \quad \varphi_2(y) - \left(\frac{\partial v}{\partial y}\right) = \alpha$$

or

$$\frac{\partial v}{\partial x}=\sqrt{\varphi_1(x)+\alpha}$$

and

$$\frac{\partial v}{\partial y}=\sqrt{\varphi_2(y)-\alpha},$$

we find

$$v=\int\sqrt{\varphi_1(x)+\alpha}\,dx+\int\sqrt{\varphi_2(y)-\alpha}\,dy;$$

consequently, the equation of geodesic lines $\frac{\partial v}{\partial \alpha}=\beta$ in this case has the form

$$\int\frac{dx}{\sqrt{\varphi_1(x)+\alpha}}-\int\frac{dy}{\sqrt{\varphi_2(y)-\alpha}}=\beta.$$

Note. The Hamilton-Jacobi equation can be approached by different reasoning as well. Consider a central field of extremals with centre in the point $A(x_0, y_0)$ for the functional

$$v[y(x)]=\int_{x_0}^{x_1}F(x, y, y')\,dx.$$

On extremals of the field, the functional $v[y(x)]$ is transformed into the function $\bar{v}(x, y)$ of the coordinates of the second boundary point $B(x, y)$. As was pointed out on page 385,

$$\frac{\partial \bar{v}}{\partial x}=-H(x, y, q), \quad \frac{\partial \bar{v}}{\partial y}=q.$$

Eliminating q, we get

$$\frac{\partial \bar{v}}{\partial x}=-H\left(x, y, \frac{\partial \bar{v}}{\partial y}\right).$$

And so the function $\bar{v}(x, y)$ is a solution of the Hamilton-Jacobi equation. Quite analogous arguments also hold true for the functional

$$\int_{x_0}^{x_1}F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx.$$

PROBLEMS ON CHAPTER 8

Test the following functionals for extrema:

1. $v[y(x)]=\int_0^2(xy'+y'^2)\,dx;\ y(0)=1;\ y(2)=0.$

2. $v[y(x)] = \int_0^a (y'^2 + 2yy' - 16y^2)\,dx;\ a > 0;\ \ y(0) = 0;\ \ y(a) = 0.$

3. $v[y(x)] = \int_{-1}^{2} y'(1 + x^2 y')\,dx;\ y(-1) = 1;\ y(2) = 4.$

4. $v[y(x)] = \int_1^2 y'(1 + x^2 y')\,dx;\ y(1) = 3;\ y(2) = 5.$

5. $v[y(x)] = \int_{-1}^{2} y'(1 + x^2 y')\,dx;\ y(-1) = y(2) = 1.$

6. $v[y(x)] = \int_0^{\frac{\pi}{4}} (4y^2 - y'^2 + 8y)\,dx;\ y(0) = -1;\ y\left(\frac{\pi}{4}\right) = 0.$

7. $v[y(x)] = \int_1^2 (x^2 y'^2 + 12y^2)\,dx;\ y(1) = 1;\ y(2) = 8.$

8. $v[y(x)] = \int_0^1 (y'^2 + y^2 + 2ye^{2x})\,dx;\ y(0) = \frac{1}{3};\ y(1) = \frac{1}{3}e^2.$

9. $v[y(x)] = \int_0^{\frac{\pi}{4}} (y^2 - y'^2 + 6y \sin 2x)\,dx;\ y(0) = 0;\ y\left(\frac{\pi}{4}\right) = 1.$

10. $v[y(x)] = \int_0^{x_1} \frac{dx}{y'};\ y(0) = 0;\ y(x_1) = y_1;\ x_1 > 0;\ y_1 > 0.$

11. $v[y(x)] = \int_0^{x_1} \frac{dx}{y'^2};\ y(0) = 0;\ y(x_1) = y_1;\ \ x_1 > 0;\ y_1 > 0.$

12. $v[y(x)] = \int_1^2 \frac{x^3}{y'^2}\,dx;\ y(1) = 1;\ y(2) = 4.$

13. $v[y(x)] = \int_1^3 (12xy + y'^2)\,dx;\ y(1) = 0;\ y(3) = 26.$

14. $v[y(x)] = \int_0^2 [y^2 + (y')^2 - 2xy]\,dx;\ y(0) = 0;\ y(2) = 3.$

Variational problems involving a conditional extremum

1. Constraints of the Form $\varphi(x, y_1, y_2, \ldots, y_n)=0$

Variational problems involving a conditional extremum are problems in which it is required to find an extremum of a functional v; certain constraints are imposed on the functions on which the functional v is dependent. For example, it is required to investigate for an extremum the functional

$$v[y_1, y_2, \ldots, y_n]=\int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx$$

given the conditions

$$\varphi_i(x, y_1, y_2, \ldots, y_n)=0 \quad (i=1, 2, \ldots, m;\ m<n).$$

Recall the solution of a similar problem dealing with the investigation of the function $z=f(x_1, x_2, \ldots, x_n)$ for an extremum given the constraints

$$\varphi_i(x_1, x_2, \ldots, x_n)=0 \quad (i=1, 2, \ldots, m;\ \ m<n).$$

The most natural way is to solve the system

$$\varphi_i(x_1, x_2, \ldots, x_n)=0 \qquad (i=1, 2, \ldots, m),$$

the equations of which we shall call independent with respect to some kind of m variables, say $x_1, x_2, \ldots, x_m$:

$$\begin{aligned} x_1&=x_1(x_{m+1}, x_{m+2}, \ldots, x_n);\\ x_2&=x_2(x_{m+1}, x_{m+2}, \ldots, x_n);\\ &\ldots\ldots\ldots\ldots\ldots\\ x_m&=x_m(x_{m+1}, x_{m+2}, \ldots, x_n), \end{aligned}$$

and with respect to the substitution of $x_1, x_2, \ldots, x_m$ into $f(x_1, x_2, \ldots, x_n)$. Then the function $f(x_1, x_2, \ldots, x_n)$ becomes a function $\Phi(x_{m+1}, x_{m+2}, \ldots, x_n)$ only of the $n-m$ variables $x_{m+1}, x_{m+2}, \ldots, x_n$, which are already independent, and so the problem has reduced to investigating the function Φ for an unconditional extremum. This approach can also be used, of course, to solve the above variational problem. Solving the system $\varphi_i(x, y_1, y_2, \ldots, y_n)=0$

$(i=1, 2, \ldots, m)$ for $y_1, y_2, \ldots, y_m$ (or any other m functions y_i) and substituting their expressions into $v[y_1, y_2, \ldots, y_n]$, we get the functional $W[y_{m+1}, y_{m+2}, \ldots, y_n)$ which depends only on $n-m$ arguments that are already independent, and, hence, the methods given in Sec. 3, Chapter 6, can now be applied to the functional W. However, both for functions and for functionals a different and more convenient method is commonly employed, that of undetermined coefficients, which retains complete equivalence of all variables. As we know, when investigating a function $z=f(x_1, x_2, \ldots, x_n)$ for an extremum, given the constraints $\varphi_i(x_1, x_2, \ldots, x_n)=0$ $(i=1, 2, \ldots, m)$, this method consists in constructing a new auxiliary function

$$z^* = f + \sum_{i=1}^{m} \lambda_i \varphi_i,$$

where the λ_i are certain constant factors and the function z^* is now investigated for an unconditional extremum; that is, we form a system of equations $\frac{\partial z^*}{\partial x_j}=0$ $(j=1, 2, \ldots, n)$ supplemented by the constraint equations $\varphi_i=0$ $(i=1, 2, \ldots, m)$ from which all the $n+m$ unknowns $x_1, x_2, \ldots, x_n$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ are determined. Also the problem involving a conditional extremum for functional is solved in similar fashion, namely if

$$v=\int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx$$

and

$$\varphi_i(x, y_1, y_2, \ldots, y_n)=0 \quad (i=1, 2, \ldots, m),$$

then the functional

$$v^*=\int_{x_0}^{x_1}\left(F+\sum_{i=1}^{m}\lambda_i(x)\,\varphi_i\right)dx \quad \text{and} \quad v^*=\int_{x_0}^{x_1} F^*\,dx$$

is constructed, where

$$F^*=F+\sum_{i=1}^{m}\lambda_i(x)\,\varphi_i,$$

which is now investigated for an unconditional extremum; that is to say we solve the system of Euler's equations

$$F^*_{y_j}-\frac{d}{dx}F^*_{y_j'}=0 \qquad (j=1, 2, \ldots, n)$$

supplemented by the constraint equations

$$\varphi_i=0 \qquad (i=1, 2, \ldots, m). \qquad (9.1)$$

Generally speaking, the number of equations $m+n$ is sufficient to determine the $m+n$ unknown functions $y_1, y_2, \ldots, y_n$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$, and the boundary conditions $y_j(x_0)=y_{j0}$ and $y_j(x_1)=y_{j1}$ $(j=1, 2, \ldots, n)$, which must not contradict the constraint equations, will, generally speaking, permit determining the $2n$ arbitrary constants in the general solution of the system of Euler's equations.

It is obvious that the curves thus found on which a minimum or maximum of the functional v^* is achived will be solutions of the original variational problem as well. Indeed, for the functions

$$\lambda_i(x) \ (i=1, 2, \ldots, m) \qquad \text{and} \qquad y_j \ (j=1, 2, \ldots, n)$$

found from the system (9.1), all the $\varphi_i=0$ and, hence, $v^*=v$, and if for $y_j=y_j(x)$ $(j=1, 2, \ldots, n)$ determined from the system (9.1) there is achieved an unconditional extremum of the functional v^*, that is, an extremum relative to all close-lying curves (both those that satisfy the constraint equations and those that do not), then, in particular, an extremum is also achieved with respect to a narrower class of curves that satisfy the constraint equations.

However, it does not by any means follow from this argument that all solutions of the original problem involving a conditional extremum will yield an unconditional extremum of the functional v^* and, consequently, it is still not clear whether all solutions can be found by this method. We shall confine ourselves to the proof of a weaker assertion.

Theorem. *Given the conditions*

$$\varphi_i(x, y_1, y_2, \ldots, y_n)=0 \qquad (i=1, 2, \ldots, m;\ m<n)$$

the functions $y_1, y_2, \ldots, y_n$ that extremize the functional

$$v=\int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx$$

satisfy—given an appropriate choice of factors $\lambda_i(x)$ $(i=1, 2, \ldots, m)$—Euler's equations formed for the functional

$$v^*=\int_{x_0}^{x_1}\left(F+\sum_{i=1}^{m}\lambda_i(x)\varphi_i\right)dx=\int_{x_0}^{x_1}F^*\,dx.$$

The functions $\lambda_i(x)$ and $y_i(x)$ are determined from Euler's equations

$$F^*_{y_j}-\frac{d}{dx}F^*_{y_j'}=0 \qquad (j=1, 2, \ldots, n)$$

and

$$\varphi_i=0 \quad (i=1, 2, \ldots, m).$$

The equations $\varphi_i = 0$ can also be considered Euler's equations for the functional v^* if we consider as arguments of the functional not only the functions $y_1, y_2, \ldots, y_n$ but also $\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x)$. The equations $\varphi_i(x, y_1, y_2, \ldots, y_n) = 0$ $(i = 1, 2, \ldots, m)$ are assumed independent, i.e. one of the Jacobians of order m is different from zero, for instance,

$$\frac{D(\varphi_1, \varphi_2, \ldots, \varphi_m)}{D(y_1, y_2, \ldots, y_m)} \neq 0.$$

Proof In the given case, the basic condition of an extremum, $\delta v = 0$, takes the form

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F_{y_j} \delta y_j + F_{y_j'} \delta y_j' \right) dx = 0.$$

Integrating by parts the second terms in each parenthesis and noting that

$$(\delta y_j)' = \delta y_j' \quad \text{and} \quad (\delta y_j)_{x=x_0} = 0; \quad (\delta y_j)_{x=x_1} = 0,$$

we get

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F_{y_j} - \frac{d}{dx} F_{y_j'} \right) \delta y_j dx = 0.$$

Since the functions $y_1, y_2, \ldots, y_n$ are subject to m independent constraints

$$\varphi_i(x, y_1, y_2, \ldots, y_n) = 0 \qquad (i = 1, 2, \ldots, m),$$

it follows that the variations δy_j are not arbitrary and the fundamental lemma cannot as yet be applied. The variations δy_j must satisfy the following conditions obtained by means of varying the constraint equations $\varphi_i = 0$:

$$\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} \delta y_j = 0 \quad (i = 1, 2, \ldots, m)^*$$

* More exactly, applying Taylor's formula to the difference

$$\varphi_i(x, y_1 + \delta y_1, \ldots, y_n + \delta y_n) - \varphi_i(x, y_1, \ldots, y_n)$$

of the left-hand sides of the equations $\varphi_i(x, y_1 + \delta y_1, \ldots, y_n + \delta y_n) = 0$ and $\varphi_i(x, y_1, \ldots, y_n) = 0$, we should write

$$\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} \delta y_j + R_i = 0,$$

where the R_i are of order higher than first in δy_j $(j = 1, 2, \ldots, n)$. However, as may readily be verified, the terms R_i do not exert any appreciable influence on subsequent reasoning, since when calculating the variation of the functional we are only interested in first-order terms in δy $(j = 1, 2, \ldots, n)$.

and, hence, only $n-m$ of the variations δy_j may be considered arbitrary, for example $\delta y_{m+1}, \delta y_{m+2}, \ldots, \delta y_n$, while the rest are determined from the equations that have been obtained.

Multiplying each of these equations term by term by $\lambda_i(x)\,dx$ and integrating from x_0 to x_1, we get

$$\int_{x_0}^{x_1} \lambda_i(x) \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} \delta y_j\, dx = 0 \qquad (i = 1, 2, \ldots, m).$$

Adding termwise all these m equations, which are satisfied by the permissible variations δy_j, with the equation

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F_{y_j} - \frac{d}{dx} F_{y_j'}\right) \delta y_j\, dx = 0.$$

we will have

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left[\frac{\partial F}{\partial y_j} + \sum_{i=1}^{m} \lambda_i(x) \frac{\partial \varphi_i}{\partial y_j} - \frac{d}{dx}\frac{\partial F}{\partial y_j'}\right] \delta y_j\, dx = 0,$$

or, if we introduce the notation

$$F^* = F + \sum_{i=1}^{m} \lambda_i(x) \varphi_i,$$

we get

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F^*_{y_j} - \frac{d}{dx} F^*_{y_j'}\right) \delta y_j\, dx = 0.$$

Here too it is impossible as yet to employ the fundamental lemma due to the fact that the variations δy_j are not arbitrary Choose m factors $\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x)$ so that they should satisfy the m equations

$$F^*_{y_j} - \frac{d}{dx} F^*_{y_j'} = 0 \qquad (j = 1, 2, \ldots, m),$$

or

$$\frac{\partial F}{\partial y_j} + \sum_{i=1}^{m} \lambda_i(x) \frac{\partial \varphi_i}{\partial y_j} - \frac{d}{dx}\frac{\partial F}{\partial y_j'} = 0 \qquad (j = 1, 2, \ldots, m).$$

These equations form a system that is linear in λ_i, with a nonzero determinant

$$\frac{D(\varphi_1, \varphi_2, \ldots, \varphi_m)}{D(y_1, y_2, \ldots, y_m)} \neq 0;$$

hence this system has the solution

$$\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x).$$

Given this choice of $\lambda_1(x), \lambda_2(x), \ldots, \lambda_m(x)$, the basic necessary condition for an extremum

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F^*_{y_j} - \frac{d}{dx} F^*_{y'_j}\right) \delta y_j \, dx = 0$$

takes the form

$$\int_{x_0}^{x_1} \sum_{j=m+1}^{n} \left(F^*_{y_j} - \frac{d}{dx} F^*_{y'_j}\right) \delta y_j \, dx = 0.$$

Since, for the extremizing functions $y_1, y_2, \ldots, y_n$ of the functional v, this functional equation reduces to an identity already for an arbitrary choice of $\delta y_j (j = m+1, m+2, \ldots, n)$, it follows that the fundamental lemma is now applicable. Putting all the δy_j equal to zero in turn, except one, and applying the lemma, we obtain

$$F^*_{y_j} - \frac{d}{dx} F^*_{y'_j} = 0 \qquad (j = m+1, m+2, \ldots, n).$$

Taking into account the above obtained equations

$$F_{y_j} - \frac{d}{dx} F^*_{y'_j} = 0 \qquad (j = 1, 2, \ldots, m),$$

we finally find that the functions which achieve a conditional extremum of the functional v, and the factors $\lambda_i(x)$ must satisfy the system of equations

$$F^*_{y_j} - \frac{d}{dx} F^*_{y'_j} = 0 \qquad (j = 1, 2, \ldots, n),$$

$$\varphi_i(x, y_1, y_2, \ldots, y_n) = 0 \qquad (i = 1, 2, \ldots, m).$$

Example 1. Find the shortest distance between two points $A(x_0, y_0, z_0)$ and $B(x_1, y_1, z_1)$ on the surface $\varphi(x, y, z) = 0$ (see the problem of geodesics, page 295). The distance between two points on a surface is, as we know, given by the formula

$$l = \int_{x_0}^{x_1} \sqrt{1 + y'^2 + z'^2} \, dx.$$

Here we have to find the minimum of l provided $\varphi(x, y, z) = 0$. According to the foregoing, we take the auxiliary functional

$$l^* = \int_{x_0}^{x_1} \left[\sqrt{1 + y'^2 + z'^2} + \lambda(x) \varphi(x, y, z)\right] dx$$

and write the Euler equations for it:

$$\lambda(x) \varphi_y - \frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2 + z'^2}} = 0;$$

$$\lambda(x)\varphi_z - \frac{d}{dx}\frac{z'}{\sqrt{1+y'^2+z'^2}} = 0;$$
$$\varphi(x, y, z) = 0.$$

From these three equations we determine the desired functions

$$y = y(x) \text{ and } z = z(x)$$

on which a conditional minimum of the functional v can be achieved, and the factor $\lambda(x)$.

Example 2. Using the Ostrogradsky-Hamilton principle (see page 333), find the equations of motion of a system of particles of mass m_i $(i = 1, 2, \ldots, n)$ with coordinates (x_i, y_i, z_i) acted upon by forces having the force function $-U$, given the constraints

$$\varphi_j(t, x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n, z_1, z_2, \ldots, z_n) = 0$$
$$(j = 1, 2, \ldots, m).$$

The Ostrogradsky-Hamilton integral

$$v = \int_{t_0}^{t_1} (T - U)\, dt$$

is here of the form

$$v = \int_{t_0}^{t_1} \left[\frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) - U\right] dt,$$

and the auxiliary functional

$$v^* = \int_{t_0}^{t_1} \left[\frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) - U + \sum_{j=1}^{m} \lambda_j(t)\varphi_j\right] dt.$$

The equations of motion will be the Euler equations for the functional v^*. They will have the following form:

$$m_i\ddot{x}_i = -\frac{\partial U}{\partial x_i} + \sum_{j=1}^{m} \lambda_j(t)\frac{\partial \varphi_j}{\partial x_i};$$

$$m_i\ddot{y}_i = -\frac{\partial U}{\partial y_i} + \sum_{j=1}^{m} \lambda_j(t)\frac{\partial \varphi_j}{\partial y_i};$$

$$m_i\ddot{z}_i = -\frac{\partial U}{\partial z_i} + \sum_{j=1}^{m} \lambda_j(t)\frac{\partial \varphi_j}{\partial z_i}$$
$$(i = 1, 2, \ldots, n).$$

2. Constraints of the Form $\varphi(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')=0$

In the preceding section we examined the problem of investigating the functional for an extremum:

$$v=\int_{x_0}^{x_1} F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\,dx;$$

$$y_j(x_0)=y_{j0},\ y_j(x_1)=y_{j1} \qquad (j=1, 2, \ldots, n)$$

given the *finite* constraints

$$\varphi_i(x, y_1, y_2, \ldots, y_n)=0 \quad (i=1, 2, \ldots, m). \tag{9.2}$$

Now suppose that the constraint equations are the *differential equations*

$$\varphi_i(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')=0 \qquad (i=1, 2, \ldots, m).$$

In mechanics, constraints of this type are called *nonholonomic*, while constraints of the type (9.2) are called *holonomic*.

In this case too we can prove the rule of factors, which consists in the fact that a conditional extremum of a functional v is achieved on the same curves on which is achieved an unconditional extremum of the functional

$$v^*=\int_{x_0}^{x_1}\left[F+\sum_{i=1}^{m}\lambda(x)\,\varphi_i\right]dx=\int_{x_0}^{x_1}F^*\,dx,$$

where

$$F^*=F+\sum_{i=1}^{m}\lambda_i(x)\,\varphi_i.$$

However, the proof is considerably more complicated than in the case of finite constraints.

But if we confine the proof to the weaker assertion that the curves on which the conditional extremum of the functional v is achieved, given an appropriate choice of $\lambda_i(x)$, are extremals for the functional v^*, then the proof given in the preceding section may, with slight modifications, be repeated for the given case as well.

Indeed, suppose that one of the functional determinants of order m is different from zero, say,

$$\frac{D(\varphi_1, \varphi_2, \ldots, \varphi_m)}{D(y_1', y_2', \ldots, y_m')}\neq 0.$$

This guarantees independence of the constraints.

Solving the equation $\varphi_i(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n') = 0$ for $y_1', y_2', \ldots, y_m'$, which is possible since

$$\frac{D(\varphi_1, \varphi_2, \ldots, \varphi_m)}{D(y_1', y_2', \ldots, y_m')} \neq 0,$$

we get $y_i' = \psi_i(x, y_1, y_2, \ldots, y_n, y_{m+1}', y_{m+2}', \ldots, y_n')$ $(i = 1, 2, \ldots, m)$. If we consider $y_{m+1}, y_{m+2}, \ldots, y_n$ arbitrarily specified functions, then $y_1, y_2, \ldots, y_m$ are determined from this system of differential equations. Thus, $y_{m+1}, y_{m+2}, \ldots, y_n$ are arbitrary differentiable functions with fixed boundary values and, hence, their variations are arbitrary in the same sense.

Let $y_1, y_2, \ldots, y_n$ be an arbitrary permissible system of functions that satisfies the constraint equations $\varphi_i = 0$ $(i = 1, 2, \ldots, m)$. Vary the constraint equations

$$\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} \delta y_j + \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j'} \delta y_j' = 0 \quad (i = 1, 2, \ldots, m)^*.$$

Multiply term by term each of the equations obtained by the (as yet) undetermined factor $\lambda_i(x)$ and integrate from x_0 to x_1; this yields

$$\int_{x_0}^{x_1} \lambda_i(x) \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} \delta y_j \, dx + \int_{x_0}^{x_1} \lambda_i(x) \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j'} \delta y_j' \, dx = 0;$$

integrating each term of the second integral by parts and taking into consideration that $\delta y_j' = (\delta y_j)'$ and $(\delta y_j)_{x=x_0} = (\delta y_j)_{x=x_1} = 0$, we will have

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left[\lambda_i(x) \frac{\partial \varphi_i}{\partial y_j} - \frac{d}{dx} \left(\lambda_i(x) \frac{\partial \varphi_i}{\partial y_j'} \right) \right] \delta y_j \, dx = 0. \tag{9.3}$$

From the basic necessary condition for an extremum, $\delta v = 0$, we have

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F_{y_j} - \frac{d}{dx} F_{y_j'} \right) \delta y_j \, dx = 0, \tag{9.4}$$

since

$$\delta v = \int_{x_0}^{x_1} \sum_{j=1}^{n} (F_{y_j} \delta y_j + F_{y_j'} \delta y_j') \, dx = \int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F_{y_j} - \frac{d}{dx} F_{y_j'} \right) \delta y_j \, dx.$$

---

* Here too (as on page 392), summands containing terms of order higher than first in δy_j and $\delta y_j'$ $(j = 1, 2, \ldots, n)$ should be included in the left-hand sides of the equations; it is now considerably more difficult to take into account the effect of these nonlinear terms.

Adding termwise all the equations (9.3) and equation (9.4) and introducing the notation $F^* = F + \sum_{i=1}^{m} \lambda_i(x)\varphi_i$, we will have

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \left(F^*_{y_j} - \frac{d}{dx} F^*_{y'_j}\right) \delta y_j \, dx = 0. \tag{9.5}$$

Since the variations δy_j $(j = 1, 2, \ldots, n)$ are not arbitrary, we cannot yet use the fundamental lemma. Choose m factors $\lambda_1(x)$, $\lambda_2(x)$, $\ldots$, $\lambda_m(x)$ so that they satisfy the equations

$$F^*_{y_j} - \frac{d}{dx} F^*_{y'_j} = 0 \qquad (j = 1, 2, \ldots, m),$$

When written in expanded form, these equations form a system of linear differential equations in

$$\lambda_i(x) \quad \text{and} \quad \frac{d\lambda_i}{dx} \qquad (i = 1, 2, \ldots, m).$$

which, given the assumptions we have, has the solution $\lambda_1(x)$, $\lambda_2(x)$, $\ldots$, $\lambda_m(x)$, which depends on m arbitrary constants. With this choice of $\lambda_1(x)$, $\lambda_2(x)$, $\ldots$, $\lambda_m(x)$ the equation (9.5) is reduced to the form

$$\int_{x_0}^{x_1} \sum_{j=m+1}^{n} \left(F^*_{y_j} - \frac{d}{dx} F^*_{y'_j}\right) \delta y_j \, dx = 0,$$

where the variations δy_j $(j = m+1, m+2, \ldots, n)$ are *now arbitrary*, and hence, assuming all variations $\delta y_i = 0$, except some one δy_i, and applying the fundamental lemma, we obtain

$$F^*_{y_i} - \frac{d}{dx} F^*_{y'_j} = 0 \qquad (j = m+1, m+2, \ldots, n).$$

Thus, the functions $y_1(x)$, $y_2(x)$, $\ldots$, $y_n(x)$ that render the functional v a conditional extremum, and the factors $\lambda_1(x)$, $\lambda_2(x)$, $\ldots$, $\lambda_m(x)$ must satisfy the system of $n+m$ equations:

$$F^*_{y_j} - \frac{d}{dx} F^*_{y'_j} = 0 \qquad (j = 1, 2, \ldots, n)$$

and

$$\varphi_i = 0 \qquad (i = 1, 2, \ldots, m),$$

that is they must satisfy the Euler equations of the auxiliary functional v^*, which is regarded as a functional dependent on the $n+m$ functions

$$y_1, y_2, \ldots, y_n, \lambda_1, \lambda_2, \ldots, \lambda_m.$$

3. Isoperimetric Problems

In the strict sense of the word, isoperimetric problems are problems in which one has to find a geometric figure of maximum area for a given perimeter.

Among such extremum problems, which were even studied in ancient Greece, were also variational problems like the one on page 295 (to find a closed curve, without self-intersection, of a given length bounding a maximum area).* Representing the curve in parametric form $x=x(t)$, $y=y(t)$, we can formulate the problem as follows: maximize the functional

$$S=\int_{t_0}^{t_1} x\dot{y}\,dt \quad \text{or} \quad S=\frac{1}{2}\int_{t_0}^{t_1}(x\dot{y}-y\dot{x})\,dt$$

provided that the functional

$$\int_{t_0}^{t_1}\sqrt{\dot{x}^2+\dot{y}^2}\,dt$$

maintains a constant value:

$$\int_{t_0}^{t_1}\sqrt{\dot{x}^2+\dot{y}^2}\,dt=l.$$

We thus have a variational problem involving a conditional extremum with a peculiar condition: the integral $\int_{t_0}^{t_1}\sqrt{\dot{x}^2+\dot{y}^2}\,dt$ maintains a constant value.

At the present time, isoperimetric problems embrace a much more general class of problems, namely: all variational problems in which it is required to determine an extremum of the functional

$$v=\int_{x_0}^{x_1} F(x,\ y_1,\ y_2,\ \ldots,\ y_n,\ y_1',\ y_2',\ \ldots,\ y_n')\,dx$$

given the so-called *isoperimetric conditions*

$$\int_{x_0}^{x_1} F_i(x,\ y_1,\ y_2,\ \ldots,\ y_n,\ y_1',\ y_2',\ \ldots,\ y_n')\,dx=l_i$$
$$(i=1,\ 2,\ \ldots,\ m),$$

where the l_i are constants, m may be greater than, less than or equal to n; and also analogous problems for more complicated functionals.

* Though the solution of this problem was known in ancient Greece, its peculiar variational nature was understood only at the end of the seventeenth century.

Isoperimetric problems can be reduced to conditional-extremum problems considered in the preceding section by introducing new unknown functions. Denote

$$\int_{x_0}^{x} F_i\,dx = z_i(x) \qquad (i = 1, 2, \ldots, m),$$

whence $z_i(x_0) = 0$ and from the condition $\int_{x_0}^{x_1} F_i\,dx = l_i$ we have $z_i(x_1) = l_i$.

Differentiating z_i with respect to x, we get

$$z_i'(x) = F_i(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')$$
$$(i = 1, 2, \ldots, m).$$

In this way, the integral isoperimetric constraints $\int_{x_0}^{x_1} F_i\,dx = l_i$ are replaced by the differential constraints:

$$z_i' = F_i(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')$$
$$(i = 1, 2, \ldots, m)$$

and hence the problem is reduced to the problem considered in the preceding section.

Applying the factor rule, it is possible, given the constraints $F_i - z_i' = 0$ $(i = 1, 2, \ldots, m)$, to replace an investigation of the functional $v = \int_{x_0}^{x_1} F\,dx$ for a conditional extremum by an investigation of the functional

$$v^* = \int_{x_0}^{x_1} \left[F + \sum_{i=1}^{m} \lambda_i(x)(F_i - z_i')\right] dx = \int_{x_0}^{x_1} F^*\,dx$$

for an unconditional extremum; here

$$F^* = F + \sum_{i=1}^{m} \lambda_i(x)(F_i - z_i')$$

The Euler equations for the functional v^* are of the form

$$F^*_{y_j} - \frac{d}{dx} F^*_{y_j'} = 0 \qquad (j = 1, 2, \ldots, n),$$

$$F^*_{z_i} - \frac{d}{dx} F^*_{z_i'} = 0 \qquad (i = 1, 2, \ldots, m),$$

or

$$F_{y_j} + \sum_{i=1}^{m} \lambda_i F_{i y_j} - \frac{d}{dx}\left(F_{y_j'} + \sum_{i=1}^{m} \lambda_i F_{i y_j'}\right) = 0$$

$$(j = 1, 2, \ldots, n),$$

$$\frac{d}{dx}\lambda_i(x) = 0 \qquad (i = 1, 2, \ldots, m).$$

From the last m equations we find that all the λ_i are constant and the first n equations coincide with Euler's equations for the functional

$$v^{**} = \int_{x}^{x_1} \left(F + \sum_{i=1}^{m} \lambda_i F_i\right) dx.$$

We thus get the following rule: to obtain the basic necessary condition in an isoperimetric problem involving finding an extremum of a functional $v = \int_{x_0}^{x_1} F\,dx$, given the constraints $\int_{x_0}^{x_1} F_i\,dx = l_i$ $(i = 1, 2, \ldots, m)$, it is necessary to form the auxiliary functional

$$v^{**} = \int_{x_0}^{x_1} \left(F + \sum_{i=1}^{m} \lambda_i F_i\right) dx,$$

where the λ_i are constants, and write the Euler equations for it.

The arbitrary constants $C_1, C_2, \ldots, C_{2n}$ in the general solution of a system of Euler's equations and the constants $\lambda_1, \lambda_2, \ldots, \lambda_m$ are determined from the boundary conditions

$$y_j(x_0) = y_{j0}, \quad y_j(x_1) = y_{j1} \qquad (i = 1, 2, \ldots, n)$$

and from the isoperimetric conditions

$$\int_{x_0}^{x_1} F_i\,dx = l_i \qquad (i = 1, 2, \ldots, m).$$

The system of Euler's equations for the functional v^{**} does not vary if v^{**} is multiplied by some constant factor μ_0 and, hence, is given in the form

$$\mu_0 v^{**} = \int_{x_0}^{x_1} \sum_{i=0}^{m} \mu_i F_i\,dx,$$

where the notations $F_0 = F$, $\mu_j = \lambda_j \mu_0$, $j = 1, \ldots, m$ have been introduced. Now all the functions F_i enter symmetrically, and therefore the extremals in the original variational problem and in the problem involving finding an extremum of the functional $\int_{x_0}^{x_1} F_s\,dx$, given the isoperimetric conditions

$$\int_{x_0}^{x_1} F_i\,dx = l_i \quad (i=0,\ 1,\ 2,\ \ldots,\ s-1,\ s+1,\ \ldots,\ m)$$

coincide with any choice of s ($s=0,\ 1,\ \ldots,\ n$).

This property is called the *reciprocity principle*. For example, the problem of a maximum area bounded by a closed curve of given length, and the problem of the minimum length of a closed curve bounded by a given area are reciprocal and have common extremals.

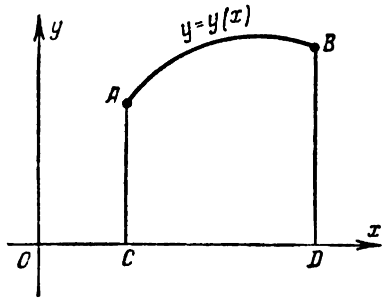


Fig. 9-1

Example 1. Find the curve $y=y(x)$ of given length l, for which the area S of the curvilinear trapezoid $CABD$ depicted in Fig. 9.1 is a maximum.

Investigate for an extremum the functional

$$S=\int_{x_0}^{x_1} y\,dx,\quad y(x_0)=y_0,$$

$y(x_1)=y_1$, given the isoperimetric condition

$$\int_{x_0}^{x_1} \sqrt{1+y'^2}\,dx = l.$$

First form the auxiliary functional

$$S^{**}=\int_{x_0}^{x_1} \left(y+\lambda\sqrt{1+y'^2}\right)dx.$$

Since the integrand does not contain x, Euler's equation for S^{**} has a first integral $F-y'F_{y'}=C_1$ or, in the given case,

$$y+\lambda\sqrt{1+y'^2}-\frac{\lambda y'^2}{\sqrt{1+y'^2}}=C_1,$$

whence

$$y-C_1=\frac{-\lambda}{\sqrt{1+y'^2}}.$$

Introduce the parameter t, putting $y'=\tan t$; this yields

$$y-C_1=-\lambda\cos t;$$

$$\frac{dy}{dx}=\tan t, \text{ whence } dx=\frac{dy}{\tan t}=\frac{\lambda\sin t\,dt}{\tan t}=\lambda\cos t\,dt;$$

$$x=\lambda\sin t+C_2.$$

Thus, the equation of the extremals in parametric form is

$$x-C_2=\lambda\sin t,$$
$$y-C_1=-\lambda\cos t,$$

or, eliminating t, we get $(x-C_2)^2+(y-C_1)^2=\lambda^2$ or a family of circles. The constants C_1, C_2 and λ are determined from the conditions

$$y(x_0)=y_0,$$

$$y(x_1)=y_1 \quad\text{and}\quad \int_{x_0}^{x_1}\sqrt{1+y'^2}\,dx=l.$$

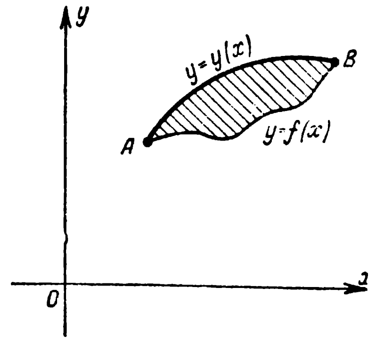


Fig. 9-2

Example 2. Find a curve AB of given length l bounding, together with a given curve $y=f(x)$, the maximum area cross-hatched in Fig. 9.2.

It is required to determine an extremum of the functional

$$S=\int_{x_0}^{x_1}(y-f(x))\,dx;$$

given the condition

$$y(x_0)=y_0,\quad y(x_1)=y_1$$

$$\int_{x_0}^{x_1}\sqrt{1+y'^2}\,dx=l.$$

Form the auxiliary functional

$$S^{**}=\int_{x_0}^{x_1}\left(y-f(x)+\lambda\sqrt{1+y'^2}\right)dx.$$

The Euler equation for this functional does not differ from the Euler equation of the preceding problem and so in the given problem the maximum may be achieved only on arcs of circles.

Example 3. Find the form of an absolutely flexible, nonextensible homogeneous rope of length l suspended at the points A and B (Fig. 9.3).

Since in the equilibrium position, the centre of gravity must occupy the lowest position, the problem reduces to finding the minimum of the static moment P about the x-axis, which is assumed to be horizontal. Investigate for an extremum the functional $P=\int_{x_0}^{x_1} y\sqrt{1+y'^2}\,dx$ provided that $\int_{x_0}^{x_1}\sqrt{1+y'^2}\,dx=l$. Form the auxiliary functional

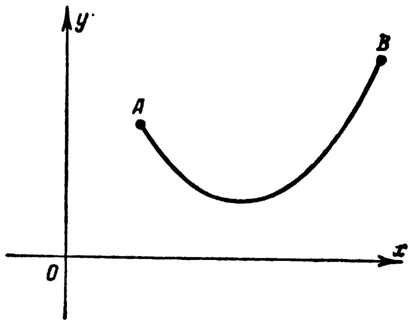


Fig. 9-3

$$P^{**}=\int_{x_0}^{x_1}(y+\lambda)\sqrt{1+y'^2}\,dx,$$

for which Euler's equation has a first integral

$$F-y'F_{y'}=C$$

or, in the given case,

$$(y+\lambda)\sqrt{1+y'^2}-\frac{(y+\lambda)y'^2}{\sqrt{1+y'^2}}=C_1,$$

whence $y+\lambda=C_1\sqrt{1+y'^2}$. Introduce a parameter putting $y'=\sinh t$, whence $\sqrt{1+y'^2}=\cosh t$ and $y+\lambda=C_1\cosh t$; $\frac{dy}{dx}=\sinh t$; $dx=\frac{dy}{\sinh t}=C_1dt$; $x=C_1t+C_2$, or, eliminating t, we get $y+\lambda=C_1\cosh\frac{x-C_2}{C_1}$, which is a family of catenaries.

The foregoing rule for solving isoperimetric problems can also be extended to more complicated functionals.

We shall mention one more problem involving a conditional extremum—the problem of optimal control. Consider the differential equation

$$\frac{dx}{dt}=f(t,\ x(t),\ u(t)) \tag{9.6}$$

with the initial condition $x(t_0)=x_0$.

Besides the unknown function (or vector function) $x(t)$, this equation also contains a so-called *control function* (or vector function) $u(t)$. The control function $u(t)$ has to be chosen so that the given functional

$$v=\int_{t_0}^{t_1} F(x(t),\ u(t))\,dt$$

is extremized.

The function $y(t)$ which yields the solution of this problem is called the *optimal function* or *optimal control.*

This problem may be regarded as a problem involving a conditional extremum of a functional v with differential constraints (9.6). However, in practical problems the optimal functions frequently lie on the boundary of a set of admissible control functions (for example, if the control function is the engine power to be switched on, then obviously this power is bounded by the maximum power output of the engines; and in the solutions of optimum problems it is often necessary to run the engines at peak-power output, at least over certain portions).

Now if the optimal function lies on the boundary of a set of admissible control functions, then the foregoing theory of problems involving a conditional extremum and presuming the possibility of two-sided variations is not applicable.

For this reason, other methods worked out by L. Pontryagin (see [8]) and R. Bellman (see [9]) are ordinarily applied in solving problems of optimal control.

Example. In the system of differential equations

$$\frac{dx}{dt}=v, \quad \frac{dv}{dt}=u \quad (t \text{ is the time}), \tag{9.7}$$

which describe the motion, in a plane, of a particle with coordinates x, v, determine the control function $u(t)$ so that the point $A(x_0, v_0)$ moves to the point $B(0, 0)$ in a least interval of time; $|u| \leqslant 1$ $\left(\text{since } u=\frac{d^2x}{dt^2}, \text{ it follows that } u \text{ may be considered a force acting on a particle of unit mass}\right)$.

The control function $u(t)$ is piecewise continuous. To simplify our reasoning, let us assume that it does not have more than one point of discontinuity, though the final result holds true even without this assumption.

It is almost obvious that on optimal trajectories $u=\pm 1$, since for these values $\left|\frac{dx}{dt}\right|$ and $\left|\frac{dv}{dt}\right|$ attain maximum values and, hence, the particle moves with maximum speed. Putting $u=1$ in (9.7) we get

$$v=t+C_1, \qquad x=\frac{t^2}{2}+C_1t+C_2,$$

or $v^2=2(x-C)$ and similarly for $u=-1$:

$$v=-t+C_1, \quad x=-\frac{t^2}{2}+C_1t+C_2, \quad v^2=-2(x-C).$$

Figures 9.4 and 9.5 depict these families of parabolas, the arrows indicating the direction of motion as t increases. If the point $A(x_0, v_0)$ lies on arcs of the parabolas

$$v = -\sqrt{x} \quad \text{or} \quad v = \sqrt{-x} \tag{9.8}$$

(Fig. 9.6) passing through the coordinate origin, then the optimal trajectory is an arc of one of these parabolas connecting the point

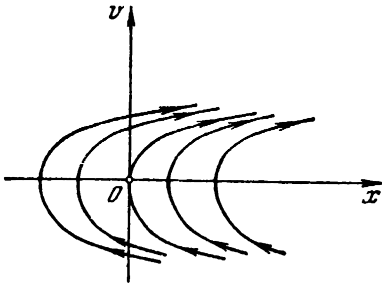


Fig. 9-4

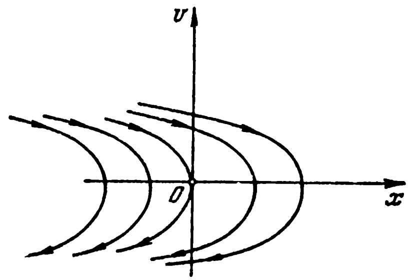


Fig. 9-5

A with the point B. But if A does not lie on these parabolas, then the optimal trajectory is the arc AC of the parabola passing through A, and the arc CB of one of the parabolas (9.8) (see Fig. 9.6 which indicates two possible positions of the points A and C).

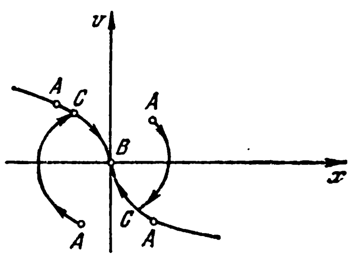


Fig. 9-6

In this problem, the time T of translation of the point from position A to position B is a functional defined by the first of the equations (9.7); the second equation of (9.7) may be regarded as a constraint equation. It would, however, be difficult to apply to this problem the earlier described classical methods of solution, since optimal control lies on the boundary of the region of admissible controls $|u| \leqslant 1$ and two-sided variations are impossible here; moreover, the solution is sought in the class of piecewise continuous controls.

Both of these circumstances are extremely characteristic of most practical problems involving optimal control.

PROBLEMS ON CHAPTER 9

1. Find the extremals of the isoperimetric problem $v[y(x)] = \int_0^1 (y'^2 + x^2)\,dx$ given that $\int_0^1 y^2\,dx = 2$; $y(0)=0$; $y(1)=0$.

2. Find the geodesics of a circular cylinder $r=R$.

Hint. It is convenient to seek the solution in cylindrical coordinates r, φ, z.

3. Find the extremals of the isoperimetric problem

$$v[y(x)] = \int_{x_0}^{x_1} y'^2\,dx \text{ given that } \int_{x_0}^{x_1} y\,dx = a,$$

where a is a constant.

4. Write the differential equation of the extremals of the isoperimetric problem involving extremization of the functional

$$v[y(x)] = \int_0^{x_1} [p(x)\,y'^2 + q(x)\,y^2]\,dx$$

given that $\int_0^{x_1} r(x)\,y^2\,dx = 1$; $y(0)=0$, $y(x_1)=0$.

5. Find the extremal in the isoperimetric problem of the extremization of the functional

$$v[y(x);\ z(x)] = \int_0^1 (y'^2 + z'^2 - 4xz' - 4z)\,dx$$

given that

$$\int_0^1 (y'^2 - xy' - z'^2)\,dx = 2;\quad y(0)=0;\quad z(0)=0;\quad y(1)=1;\quad z(1)=1.$$

CHAPTER 10

Direct methods in variational problems

1. Direct Methods

Differential equations of variational problems can be integrated in closed form only in exceptional cases. This naturally gives rise to the search for other methods of solution. The basic idea of the so-called *direct methods* consists in the following: the variational problem is regarded as a limiting case of a certain extremum problem of a function of a finite number of variables. This extremum problem of a function of a finite number of variables is solved by ordinary methods, then a passage to the limit yields the solution of the appropriate variational problem.

The functional $v[y(x)]$ may be regarded as a function of an infinite set of variables. This assertion becomes quite obvious if we assume that the admissible functions can be expanded in power series:

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

or in Fourier's series:

$$y(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),$$

or in some other kind of series of the form

$$y(x)=\sum_{n=0}^{\infty}a_n\varphi_n(x),$$

where $\varphi_n(x)$ are given functions. To specify a function $y(x)$ that can be represented in the form of a series $y(x)=\sum_{n=0}^{\infty}a_n\varphi_n(x)$, it is sufficient to give the values of all the coefficients a_n, and, hence, the value of the functional $v[y(x)]$ in this case will be given by specifying an infinite sequence of numbers: $a_0, a_1, a_2, \ldots, a_n, \ldots$; i.e. the functional is a function of an infinite set of variables:

$$v[y(x)]=\varphi(a_0, a_1, \ldots, a_n, \ldots).$$

Consequently, the difference between variational probiems and extremum problems of functions of a finite number or variables is that in the variational case one has to investigate, for an extremum, functions of an infinite number of variables. Therefore, the basic idea of the direct methods which consists, as has already been stated above, in regarding the variational problem as a limiting case for the extremum problem of functions of a finite number of variables, is quite natural.

During the first period of investigations into the field of the calculus of variations, Euler employed a method which is now called the direct method of finite differences. For a long time this method was not in use at all and only during the past three decades was revived and successfully used in the works of the Soviet mathematicians L. Lyusternik, I. Petrovsky, and others.

Another direct method, called the Ritz method, in the development of which a very substantial contribution has been made by the Soviet mathematicians N. Krylov, N. Bogolyubov, and others, finds wide application in the solution of various variational problems.

A third direct method, proposed by L. Kantorovich, is applicable to functionals that depend on the functions of several independent variables and is finding ever broader uses in areas in which the Ritz method is employed.

We shall examine only these three basic direct methods (the proofs of many of the assertions will not be given). The reader who wishes to study more closely the direct methods now in use is referred to L. Kantorovich and V. Krylov [10] and S. Mikhlin [11].

2. Euler's Finite-Difference Method

The underlying idea of the method of finite differences is that the values of a functional $v[y(x)]$, for example,

$$\int_{x_0}^{x_1} F(x, y, y')\,dx, \quad y(x_0)=a, \quad y(x_1)=b,$$

are considered not on arbitrary curves that are admissible in the given variational problem, but only on polygonal curves made up of a given number n of straight-line segments with specified abscissas of the vertices:

$$x_0+\Delta x,\ x_0+2\Delta x,\ \ldots,\ x_0+(n-1)\Delta x, \text{ where}$$
$$\Delta x=\frac{x_1-x_0}{n} \text{ (Fig. 10.1).}$$

On such polygonal curves, the functional $v[y(x)]$ is transformed into a function $\varphi(y_1, y_2, \ldots, y_{n-1})$ of the ordinates $y_1, y_2, \ldots, y_{n-1}$ of the vertices of the polygonal curve, since the curve is completely defined by these ordinates.

We choose the ordinates $y_1, y_2, \ldots, y_{n-1}$ so that the function $\varphi(y_1, y_2, \ldots, y_{n-1})$ is extremized, that is we determine $y_1, y_2, \ldots, y_{n-1}$ from the system of equations

$$\frac{\partial\varphi}{\partial y_1}=0, \quad \frac{\partial\varphi}{\partial y_2}=0, \quad \ldots, \quad \frac{\partial\varphi}{\partial y_{n-1}}=0,$$

and then pass to the limit as $n \to \infty$. Given certain restrictions imposed on the function F, we obtain, in the limit, the solution of the variational problem.

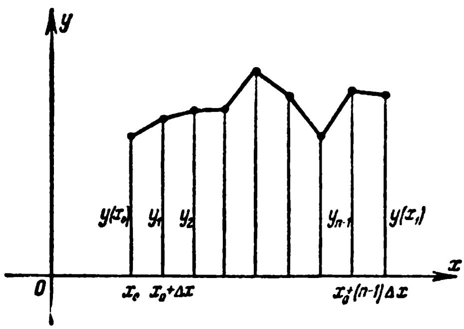


Fig. 10-1

However, it is more convenient to calculate approximately the value of the functional $v[y(x)]$ on the above-indicated polygonal curves; for instance in the most elementary problem it is best to replace the integral

$$\int_{x_0}^{x_1} F(x, y, y')\,dx \approx \sum_{k=0}^{n=1} \int_{x_0+k\Delta x}^{x_0+(k+1)\Delta x} F\left(x, y, \frac{y_{k+1}-y_k}{\Delta x}\right) dx$$

by the integral sum

$$\sum_{i=1}^{n} F\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)\Delta x.$$

By way of an illustration, let us derive Euler's equation for the functional

$$v[y(x)]=\int_{x_0}^{x_1} F(x, y, y')\,dx.$$

In this case, on the polygonal curves under investigation

$$v[y(x)] \approx \varphi(y_1, y_2, \ldots, y_{n-1})=\sum_{i=i}^{n-1} F\left(x_i, y_i, \frac{y_{i+1}-y_i}{\Delta x}\right)\Delta x.$$

Since only two terms of this sum, the ith and the $(i-1)$th, depend on the y_i:

$$F\left(x_i, y_i, \frac{y_{i+1}-y_i}{\Delta x}\right)\Delta x \quad \text{and} \quad F\left(x_{i-1}, y_{i-1}, \frac{y_i-y_{i-1}}{\Delta x}\right)\Delta x,$$

it follows that the equations $\frac{\partial\varphi}{\partial y_i}=0$ $(i=1, 2, \ldots, n-1)$ take the form

$$F_y\left(x_i, y_i, \frac{y_{i+1}-y_i}{\Delta x}\right)\Delta x+F_{y'}\left(x_i, y_i, \frac{y_{i+1}-y_i}{\Delta x}\right)\left(-\frac{1}{\Delta x}\right)\Delta x+$$
$$+F_{y'}\left(x_{i-1}, y_{i-1}, \frac{y_i-y_{i-1}}{\Delta x}\right)\frac{1}{\Delta x}\Delta x=0 \quad (i=1, 2, \ldots, (n-1)),$$

or

$$F_y\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)-\frac{F_{y'}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)-F_{y'}\left(x_{i-1}, y_{i-1}, \frac{\Delta y_{i-1}}{\Delta x}\right)}{\Delta x}=0,$$

or

$$F_y\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)-\frac{\Delta F_{y'}}{\Delta x}=0.$$

Passing to the limit as $n\to\infty$, we get Euler's equation

$$F_y-\frac{d}{dx}F_{y'}=0$$

which must be satisfied by the desired extremizing function $y(x)$. In similar fashion it is possible to obtain the basic necessary extremum condition in other variational problems.

If we do not pass to the limit, then from the system of equations $\frac{\partial\varphi}{\partial y_i}=0$ $(i=1, 2, \ldots, n-1)$ it is possible to determine the desired ordinates $y_1, y_2, \ldots, y_{n-1}$ and thus obtain a polygonal curve which is an approximate solution of the variational problem.

3. The Ritz Method

The underlying idea of the *Ritz method* is that the values of a functional $v[y(x)]$ are considered not on arbitrary admissible curves of a given variational problem but only on all possible linear combinations $y_n=\sum_{i=1}^{n}\alpha_i W_i(x)$ with constant coefficients composed of n first functions of some chosen sequence of functions

$$W_1(x), W_2(x), \ldots, W_n(x), \ldots$$

The functions $y_n=\sum_{i=1}^{n}\alpha_i W_i(x)$ must be admissible in the problem at hand; this imposes certain restrictions on the choice of sequence of functions $W_i(x)$. On such linear combinations, the functional $v[y(x)]$ is transformed into a function $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$. These coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ are chosen so that the function $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is extremized; hence, $\alpha_1, \alpha_2, \ldots, \alpha_n$ must be determined from the system of equations

$$\frac{\partial\varphi}{\partial\alpha_i}=0 \qquad (i=1, 2, \ldots, n).$$

Passing to the limit as $n \to \infty$, if the limit exists, we get the function $y=\sum_{i=1}^{\infty}\alpha_i W_i(x)$, which (for certain restrictions imposed on the functional $v[y(x)]$ and on the sequence $W_1(x), W_2(x), \ldots, W_n(x), \ldots$) is the exact solution of the variational problem at hand. If we do not pass to the limit and confine ourselves only to the first n terms of $y_n=\sum_{i=1}^{n}\alpha_i W_i(x)$, then we obtain an approximate solution of the variational problem.

If this method is used to determine the absolute minimum of the functional, then the approximate value of the minimum of the functional is obtained in excess, since the minimum of the functional on any admissible curves does not exceed the minimum of the same functional on parts of this class of admissible curves, on curves of the form $y_n=\sum_{i=1}^{n}\alpha_i W_i(x)$. When maximizing the function by the same method, we get (for the same reasons) an approximate value of the maximum of the functional in defect

For the functions $y_n=\sum_{i=1}^{n}\alpha_i W_i(x)$ to be admissible, it is first of all necessary to satisfy the boundary conditions (one should not of course forget about other restrictions that may be imposed on admissible functions, say requirements involving their continuity or smoothness). If the boundary conditions are linear and homogeneous, for example, in the elementary problem $y(x_0)=y(x_1)=0$ or

$$\beta_{1j}y(x_j)+\beta_{2j}y'(x_j)=0 \qquad (j=0, 1),$$

where the β_{ij} are constants, then the simplest thing is to choose also coordinate functions such as will satisfy these boundary conditions. Quite obviously, then, $y_n=\sum_{i=1}^{n}\alpha_i W_i(x)$ will also satisfy the same boundary conditions for any α_i. For example, let the boundary conditions have the form $y(x_0)=y(x_1)=0$, then for the coordinate functions we can choose

$$W_i(x)=(x-x_0)(x-x_1)\varphi_i(x),$$

where the $\varphi_i(x)$ are some continuous functions, or

$$W_k(x) = \sin\frac{k\pi(x-x_0)}{x_1-x_0} \qquad (k=1, 2, \ldots),$$

or some other functions that satisfy the conditions

$$W_i(x_0) = W_i(x_1) = 0.$$

If the conditions are nonhomogeneous, for example $y(x_0) = y_0$, $y(x_1) = y_1$, where at least one of the numbers y_0 or y_1 is different from zero, then it is simpler to seek the solution of the variati on problem in the form

$$y_n = \sum_{i=1}^{n} \alpha_i W_i(x) + W_0(x),$$

where $W_0(x)$ satisfies the given boundary conditions $W_0(x_0) = y_0$, $W_0(x_1) = y_1$, and all the remaining $W_i(x)$ satisfy the corresponding homogeneous boundary conditions, i.e. in the case at hand, $W_i(x_0) = W_i(x_1) = 0$. It is obvious that in such a choice, for any α_i, the functions $y_n(x)$ satisfy the given boundary conditions. For the function $W_0(x)$ we can choose, say, the linear function

$$W_0(x) = \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) + y_0.$$

Generally speaking, it is a very complicated problem to solve the system of equations $\frac{\partial\varphi}{\partial\alpha_i} = 0$ $(i = 1, 2, \ldots, n)$. This problem is appreciably simplified if we test for an extremum a functional v that is quadratic in the unknown function and its derivatives, for in this case the equations $\frac{\partial\varphi}{\partial\alpha_i} = 0$ $(i = 1, 2, \ldots, n)$ are linear in α_i.

The choice of the sequence of functions $W_1, W_2, \ldots, W_n, \ldots$, called coordinate functions, affects very appreciably the degree of complexity of subsequent calculations, and for this reason the success of the method depends largely on a proper choice of the coordinate system of functions.

The foregoing fully applies both to the functionals $v[z(x_1, x_2, \ldots, x_n)]$ (in this case of course the functions W_i must already be functions of the variables $x_1, x_2, \ldots, x_n$) and to functionals dependent on several functions.

The Ritz method is frequently employed for exact or approximate solutions of problems in mathematical physics. For example, if it is required, in some domain D, to find a solution of the Poisson equation

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = f(x, y)$$

for specified values of z on the boundary of D, this problem may be replaced by the variational problem on the extremum of a functional for which the given equation is the Ostrogradsky equation (see page 328). In the case at hand, this functional will be

$$\iint_D \left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 2zf(x, y)\right] dx\, dy.$$

The extremizing function z of this functional may be found by any one of the direct methods.

Problems of mathematical physics ordinarily reduce to investigating, for an extremum, functions that are quadratic in the unknown function and its derivatives, and hence, as already indicated, the use of the Ritz method is then simplified.

The question of the convergence of the approximations (obtained by the Ritz method) to the desired solution of the variational problem, and also of evaluating the degree of accuracy of the approximations is extremely complicated. We shall therefore confine ourselves to only a few remarks and refer the interested reader tc texts by Mikhlin [11] and Kantorovich and Krylov [10].

For the sake of definiteness, we will have in view the functiona

$$v[y(x)] = \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx$$

and assume that we are interested in its minimum. We will consider the sequence of coordinate functions $W_1(x), W_2(x), \ldots, W_n(x), \ldots$ complete in the sense that each admissible function can be approximated to any degree of accuracy in the sense of first-order proximity by the linear combination $\sum_{k=1}^{n} \alpha_k W_k(x)$ of coördinate functions, where n is sufficiently large. Then, obviously, the Ritz method may be used to obtain the functions $y_1, y_2, \ldots, y_n, \ldots$, where $y_n = \sum_{k=1}^{n} \alpha_k W_k(x)$, which form a so-called minimizing sequence, i.e., a sequence for which the values of the functional

$$v[y_1], v[y_2], \ldots, v[y_n], \ldots$$

converge to the minimum or to the lower bound of values of the functional $v[y(x)]$. However, from the fact that $\lim_{n\to\infty} v[y_n(x)] = \min v[y(x)]$ it does not in the least follow that $\lim_{n\to\infty} y_n(x) = y(x)$. A minimizing sequence may not tend to the extremizing function in the class of admissible functions.

Indeed, the functional

$$v[y_n(x)] = \int_{x_0}^{x_1} F(x, y_n(x), y'_n(x))\, dx$$

may differ but slightly from

$$v[y(x)] = \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx$$

not only when throughout the interval of integration $y_n(x)$ is close in the sense of first-order proximity to $y(x)$, but also when, over sufficiently small portions of the interval (x_0, x_1), the functions $y_n(x)$ and $y(x)$ or their derivatives differ radically, though remain close on the rest of the interval (x_0, x_1) (Fig. 10.2). For this reason, the minimizing sequence $y_1, y_2, \ldots, y_n$ may not even have a limit in the class of admissible functions, though the functions $y_1, y_2, \ldots, y_n$ will themselves be admissible.

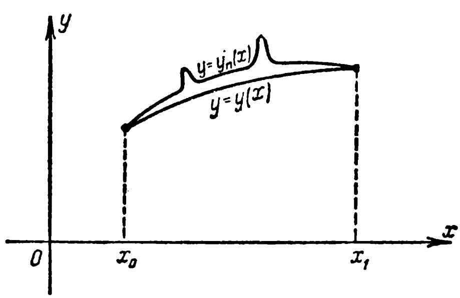


Fig. 10-2

The conditions of convergence of the sequence y_n, obtained by the Ritz method, to a solution of the variational problem and the evaluation of the speed of convergence for concrete, frequently encountered functionals have been worked out by N. Krylov and N. Bogolyubov. For instance, for functionals of the type

$$v = \int_0^1 [p(x)y'^2 + q(x)y^2 + f(x)y]\, dx; \quad y(0) = y(1) = 0,$$

where $p(x) > 0$; $q(x) \geqslant 0$, which are often met with in applications, not only has the convergence been proved of approximations (obtained by the Ritz method) to the function $y(x)$ that minimizes the functional, given the coordinate functions

$$W_k(x) = \sqrt{2} \sin k\pi x \qquad (k = 1, 2, \ldots),$$

but extremely precise error estimations $|y(x) - y_n(x)|$ have been given.

We give one of these estimations of the maximum $|y(x)-y_n(x)|$ on the interval (0, 1):

$$\max|y-y_n|\leqslant\frac{1}{n+1}\left[\max p(x)+\frac{\max q(x)}{(n+1)^2\pi^2}\right]^{\frac{1}{2}}\frac{\sqrt{\int\limits_0^1 f^2(x)\,dx}}{\pi^2\sqrt{2}\,[\min p(x)]^{\frac{5}{2}}}\times$$

$$\times\left[\max|p'(x)|+\frac{1}{\pi}\max q(x)+\pi\min p(x)\right]^*.$$

Even in this comparatively simple case, the estimation of error is very complicated. For this reason, to estimate the accuracy of results obtained by the Ritz method or by other direct methods, one ordinarily uses the following procedure, which is of course theoretically imperfect but sufficiently reliable in a practical way: after calculating $y_n(x)$ and $y_{n+1}(x)$, a comparison is made between them at several points of the interval $[x_0, x_1]$. If their values coincide within the limits of accuracy required, then it is taken that to within the required accuracy the solution of the variational problem at hand is $y_n(x)$. But if the values of $y_n(x)$ and $y_{n+1}(x)$ do not coincide even at some of the chosen points within the limits of the given accuracy, then $y_{n+2}(x)$ is calculated and the values of $y_{n+1}(x)$ and $y_{n+2}(x)$ are compared. This process is continued until the values of $y_{n+k}(x)$ and $y_{n+k+1}(x)$ coincide within the limits of the given accuracy.

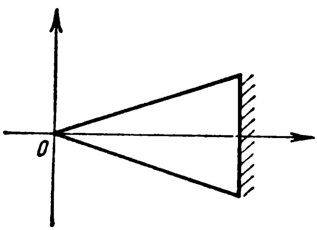


Fig. 10-3

Example 1. In studying the vibrations of a fixed wedge of constant thickness (Fig. 10.3), one has to test for an extremum the functional

$$v=\int\limits_0^1 (ax^3y''^2-bxy^2)\,dx;\qquad y(1)=y'(1)=0,$$

where a and b are positive constants. For the coordinate functions satisfying the boundary conditions we can take

$$(x-1)^2,\ (x-1)^2x,\ (x-1)^2x^2,\ \ldots,\ (x-1)^2x^{k-1},\ \ldots;$$

hence,

$$y_n=\sum_{k=1}^{n}\alpha_k(x-1)^2x^{k-1}.$$

* See Kantorovich and Krylov [10].

Confining ourselves only to the first two terms, we get

$$y_2=(x-1)^2(\alpha_1+\alpha_2 x).$$

then

$$v_2=v[y_2]=\int_0^1 [ax^3(6\alpha_2 x+2\alpha_1-4\alpha_2)^2-bx(x-1)^4(\alpha_1+\alpha_2 x)^2]\,dx=$$

$$=a\left[(\alpha_1-2\alpha_2)^2+\frac{24}{5}\alpha_2(\alpha_1-2\alpha_2)+6\alpha_2^2\right]-b\left(\frac{\alpha_1^2}{30}+\frac{2\alpha_1\alpha_2}{105}+\frac{\alpha_2^2}{280}\right).$$

Here, the necessary conditions for an extremum, $\frac{\partial v_2}{\partial \alpha_1}=0$, $\frac{\partial v_2}{\partial \alpha_2}=0$, take the form

$$\left(a-\frac{b}{30}\right)\alpha_1+\left(\frac{2}{5}a-\frac{b}{105}\right)\alpha_2=0$$

and

$$\left(\frac{2}{5}a-\frac{b}{105}\right)\alpha_1+\left(\frac{2}{5}a-\frac{b}{280}\right)\alpha_2=0.$$

To obtain solutions different from the solution $\alpha_1=\alpha_2=0$, which corresponds to the absence of vibrations of the wedge, it is necessary that the determinant of this homogeneous linear system of equations be zero:

$$\begin{vmatrix} a-\frac{b}{30} & \frac{2}{5}a-\frac{b}{105} \\ \frac{2}{5}a-\frac{b}{105} & \frac{2}{5}a-\frac{b}{280} \end{vmatrix}=0$$

or

$$\left(a-\frac{b}{30}\right)\left(\frac{2}{5}a-\frac{b}{280}\right)-\left(\frac{2}{5}a-\frac{b}{105}\right)^2=0.$$

This equation is called the *frequency equation*. It defines the frequency b of natural vibrations of the wedge, which are described by the function

$$u(x,\ t)=y(x)\cos bt.$$

The smaller of the two roots b_1 and b_2 of the frequency equation yields an approximate value of the frequency of the fundamental tone of vibrations of the wedge.

Example 2. In problems associated with the torsion of a cylinder or prism, one has to investigate the functional

$$v[z(x,\ y)]=\iint_D \left[\left(\frac{\partial z}{\partial x}-y\right)^2+\left(\frac{\partial z}{\partial y}+x\right)^2\right]dx\,dy$$

for an extremum. For a cylinder with an elliptical cross-section the integration domain D will be bounded by the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. In this case, taking only one coordinate function xy, we get

$$z_1=\alpha xy,\quad v[z_1]=v_1=\frac{\pi ab}{4}\left[(\alpha+1)^2a^2+(\alpha-1)^2b^2\right].$$

The necessary condition for an extremum, $\frac{\partial v_1}{\partial \alpha}=0$, in this case takes the form $(\alpha+1)a^2+(\alpha-1)b^2=0$, whence

$$\alpha=\frac{b^2-a^2}{a^2+b^2},\quad z_1=\frac{b^2-a^2}{a^2+b^2}xy.$$

Example 3. If in Example 2 the domain D is a rectangle with sides $2a$ and $2b$, $-a\leqslant x\leqslant a$; $-b\leqslant y\leqslant b$, then, taking for the coordinate functions xy, xy^3, x^3y, that is, putting

$$z_3=\alpha_1xy+\alpha_2xy^3+\alpha_3x^3y,$$

we get

$$v_3=v[z_3]=\int_{-a}^{a}\int_{-b}^{b}\left[\left(\frac{\partial z_3}{\partial x}-y\right)^2+\left(\frac{\partial z_3}{\partial y}+x\right)^2\right]dx\,dy=$$
$$=\frac{4}{3}ab^3(\alpha_1-1)^2+4ab^5\left(\frac{b^2}{7}+\frac{3a^2}{5}\right)\alpha_2^2+4a^5b\left(\frac{a^2}{7}+\frac{3b^2}{5}\right)\alpha_3^2+$$
$$+\frac{4}{3}a^3b(\alpha_1+1)^2+\frac{8}{5}ab^5(\alpha_1-1)\alpha_2+\frac{8}{5}a^3b(\alpha_1+1)\alpha_2-$$
$$-\frac{8}{5}a^5b(\alpha_1+1)\alpha_3-\frac{8}{5}a^3b^3(a^2+b^2)\alpha_2\alpha_3-\frac{8}{3}a^3b^3(\alpha_1-1)\alpha_3.$$

The necessary conditions for an extremum, $\frac{\partial v_3}{\partial \alpha_1}=0$, $\frac{\partial v_3}{\partial \alpha_2}=0$, $\frac{\partial v_3}{\partial \alpha_3}=0$, permit calculating α_1, α_2, α_3:

$$\alpha_1=-\frac{7(a^6-b^6)+135a^2b^2(a^2-b^2)}{7(a^6+b^6)+107a^2b^2(a^2+b^2)},$$
$$\alpha_2=-\frac{7a^2(3a^2+35b^2)}{21(a^6+b^6)+321a^2b^2(a^2+b^2)},$$
$$\alpha_3=-\frac{7b^2(35a^2+3b^2)}{21(a^6+b^6)+321a^2b^2(a^2+b^2)}.$$

Example 4. Find a solution of the equation

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=f(x,\ y)$$

inside the rectangle D, $0\leqslant x\leqslant a$, $0\leqslant y\leqslant b$, that vanishes on the boundary of D. The function $f(x, y)$ is assumed to be expansible (inside this rectangle) in a uniformly convergent double Fourier series:

$$f(x,\ y)=\sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\beta_{pq}\sin p\frac{\pi x}{b}\sin q\frac{\pi y}{b}.$$

This boundary-value problem may be reduced to a variational problem—that of finding a functional for which the given equation is an Ostrogradsky equation; and then, using one of the direct methods, find the function that extremizes this functional, and thus find the solution of the original boundary-value problem. As can easily be verified,

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=f(x, y)$$

is the Ostrogradsky equation for the functional

$$v[z(x, y)]=\iint_D \left[\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2+2zf(x, y)\right] dx\,dy$$

(see page 328). The boundary condition is maintained: $z=0$ on the boundary of the domain D. Let us investigate this functional for an extremum by the Ritz method. For a system of coordinate functions take

$$\sin m\frac{\pi x}{a}\sin n\frac{\pi y}{b} \quad (m, n=1, 2, \ldots).$$

Each of these functions and their linear combinations satisfy the boundary condition $z=0$ on the boundary of D. These functions also possess the property of completeness. Taking

$$z_{n, m}=\sum_{p=1}^{n}\sum_{q=1}^{m}\alpha_{pq}\sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b},$$

we will have

$$v[z_{nm}]=\int_0^a\int_0^b\left[\left(\frac{\partial z_{nm}}{\partial x}\right)^2+\left(\frac{\partial z_{nm}}{\partial y}\right)^2+\right.$$

$$\left.+2z_{nm}\sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\beta_{pq}\sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b}\right]dx\,dy=$$

$$=\frac{\pi^2 ab}{4}\sum_{p=1}^{n}\sum_{q=1}^{m}\left(\frac{p^2}{a^2}+\frac{q^2}{b^2}\right)\alpha_{pq}^2+\frac{ab}{2}\sum_{p=1}^{n}\sum_{q=1}^{m}\alpha_{pq}\beta_{pq}.$$

This result is readily obtained if we take into account that the coordinate functions $\sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b}$ $(p, q=1, 2, \ldots)$ form in D an orthogonal system, i.e.

$$\iint_D \sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b}\sin p_1\frac{\pi x}{a}\sin q_1\frac{\pi x}{b}\,dx\,dy=0$$

for any positive integral p, q, p_1, q_1, with the exception of the case $p=p_1$, $q=q_1$. For $p=p_1$ and $q=q_1$, we get

$$\iint_D \sin^2 p\frac{\pi x}{a}\sin^2 q\frac{\pi y}{b}\,dx\,dy=\frac{ab}{4}.$$

Therefore, of all the terms under the sign of the double integral, equal to $v[z_{nm}]$, only those are taken into account that contain squares of the functions $\sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b}$, $\sin p\frac{\pi x}{a}\cos q\frac{\pi y}{b}$ and $\cos p\frac{\pi x}{a}\sin q\frac{\pi y}{b}$. Obviously, $v[z_{nm}]$ is a function $\varphi(\alpha_{11}, \alpha_{12}, \ldots$ $\ldots, \alpha_{nm})$ of the coefficients $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{nm}$, which are determined from the basic necessary condition for an extremum

$$\frac{\partial\varphi}{\partial\alpha_{pq}}=0 \quad (p=1, 2, \ldots, n;\quad q=1, 2, \ldots, m).$$

In the given case, this system of equations is of the form

$$\alpha_{pq}\left(\frac{p^2}{a^2}+\frac{q^2}{b^2}\right)\pi^2+\beta_{pq}=0 \quad (p=1, 2, \ldots, n;$$
$$q=1, 2, \ldots, m),$$

whence

$$\alpha_{pq}=-\frac{\beta_{pq}}{\pi^2\left(\frac{p^2}{a^2}+\frac{q^2}{b^2}\right)}.$$

Consequently,

$$z_{nm}=-\frac{1}{\pi^2}\sum_{p=1}^{n}\sum_{q=1}^{m}\frac{\beta_{pq}}{\frac{p^2}{a^2}+\frac{q^2}{b^2}}\sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b}.$$

Proceeding to the limit as n and m approach infinity, we obtain an exact solution here:

$$z=-\frac{1}{\pi^2}\sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\frac{\beta_{pq}}{\frac{p^2}{a^2}+\frac{q^2}{b^2}}\sin p\frac{\pi x}{a}\sin q\frac{\pi y}{b}.$$

4. Kantorovich's Method

When applying the Ritz method to functionals $v[z(x_1, x_2, \ldots, x_n)]$ that depend on functions of several independent variables, a coordinate system of functions is chosen

$$W_1(x_1, x_2, \ldots, x_n),\ W_2(x_1, x_2, \ldots, x_n),\ \ldots,$$
$$W_m(x_1, x_2, \ldots, x_n),\ \ldots$$

and an approximate solution of the variational problem is sought in the form $z_m = \sum_{k=1}^{m} \alpha_k W_k(x_1, x_2, \ldots, x_n)$, where the coefficients α_k are constants.

The *Kantorovich method* also requires choosing a coordinate system of functions

$$W_1(x_1, x_2, \ldots, x_n),\ W_2(x_1, x_2, \ldots, x_n),\ \ldots,\ W_m(x_1, x_2, \ldots, x_n),\ \ldots$$

and an approximate solution is also sought in the form

$$z_m = \sum_{k=1}^{m} \alpha_k(x_i) W_k(x_1, x_2, \ldots, x_n);$$

however, the coefficients $\alpha_k(x_i)$ are not constants but are unknown functions of one of the independent variables. On the class of functions of the type

$$z_m = \sum_{k=1}^{m} \alpha_k(x_i) W_k(x_1, x_2, \ldots, x_n)$$

the functional $v[z]$ is transformed to the functional $\tilde{v}[\alpha_1(x_i), \alpha_2(x_i), \ldots, \alpha_m(x_i)]$, which depends on m functions of one independent variable

$$\alpha_1(x_i),\ \alpha_2(x_i),\ \ldots,\ \alpha_m(x_i).$$

The functions $\alpha_1(x_i), \alpha_2(x_i), \ldots, \alpha_m(x_i)$ are chosen so as to extremize the functional $\tilde{v}$

If after that we pass to the limit as $m \to \infty$, then under certain conditions it is possible to obtain an exact solution, but if a passage to the limit is not performed, then this method will yield an approximate solution and, generally speaking, one substantially more exact than when using the Ritz method with the same coordinate functions and the same number of terms m.

The greater precision of this method is due to the fact that the class of functions $z_m = \sum_{k=1}^{m} \alpha_k(x_i) W_k(x_1, x_2, \ldots, x_n)$ with variables $\alpha_k(x_i)$ is considerably broader than the class of functions

$$z_m = \sum_{k=1}^{m} \alpha_k W_k(x_1, x_2, \ldots, x_n)$$

for constant α_k and, hence, among functions of the type

$$z_m = \sum_{k=1}^{m} \alpha_k(x_i) W_k(x_1, x_2, \ldots, x_n)$$

it is possible to find functions that approximate better the solution of the variational problem than among functions of the form $\sum_{k=1}^{m} \alpha_k W_k(x_1, x_2, \ldots, x_n)$ where the α_k are constant.

For example, let it be required to investigate for an extremum the functional

$$v = \int_{x_0}^{x_1} \int_{\varphi_1(x)}^{\varphi_2(x)} F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dx\, dy,$$

extended over the domain D bounded by the curves $y = \varphi_1(x)$, $y = \varphi_2(x)$ and two straight lines $x = x_0$ and $x = x_1$ (Fig. 10.4). Values of the functions $z(x, y)$ are given on the boundary of the domain D.

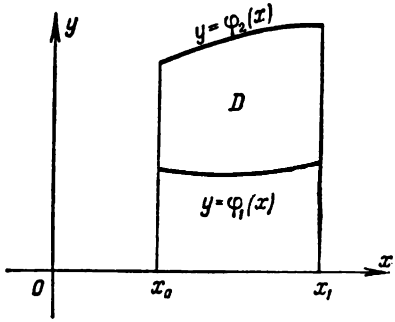


Fig. 10-4

Choose a sequence of coordinate functions:

$$W_1(x, y), \quad W_2(x, y), \ldots, \quad W_n(x, y), \ldots.$$

For the time being we confine ourselves to the first m functions of this sequence and seek the solution of the variational problem in the form of a sum of the functions $z_m = \sum_{k=1}^{m} \alpha_k(x) W_k(x, y)$ or, changing the notation $\alpha_k(x)$ to $u_k(x)$, we get

$$z_m(x, y) = u_1(x) W_1(x, y) + u_2(x) W_2(x, y) + \ldots + u_m(x) W_m(x, y),$$

where the W_k are the functions we chose, and u_k are unknown functions that we difine so that the functional v is extremized. We have

$$v[z_m(x, y)] = \int_{x_0}^{x_1} dx \int_{\varphi_1(x)}^{\varphi_2(x)} F\left(x, y, z_m(x, y), \frac{\partial z_m}{\partial x}, \frac{\partial z_m}{\partial y}\right) dy.$$

Since the integrand is a known function of y, integration with respect to y may be performed and the functional $v[z_m(x, y)]$ will be a functional of the form

$$v[z_m(x, y)] = \int_{x_0}^{x_1} \varphi(x, u_1(x), \ldots, u_m(x), u_1', \ldots, u_m')\, dx$$

The functions $u_1(x), u_2(x), \ldots, u_m(x)$ are chosen so that the functional $v[z_m(x, y)]$ is extremized. Hence, $u_i(x)$ must satisfy the system of Euler's equations

$$\varphi_{u_1}-\frac{d}{dx}\varphi_{u'_1}=0,$$
$$\varphi_{u_2}-\frac{d}{dx}\varphi_{u'_2}=0,$$
$$\cdots\cdots\cdots$$
$$\varphi_{u_m}-\frac{d}{dx}\varphi_{u'_m}=0.$$

The arbitrary constants are chosen so that $z_m(x, y)$ satisfies the given boundary conditions on the straight lines $x=x_0$ and $x=x_1$.

Example 1. Investigate for an extremum the functional

$$v[z(x, y)]=\int_{-a}^{a}\int_{-b}^{b}\left[\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2-2z\right]dx\,dy;$$

on the boundary of the integration domain $z=0$. The integration domain is a rectangle $-a\leqslant x\leqslant a$; $-b\leqslant y\leqslant b$. We seek a solution in the form $z_1=(b^2-y^2)\,u(x)$; then the boundary conditions on the straight lines $y=\pm b$ will be satisfied. The functional

$$v[z_1]=\int_{-a}^{a}\left[\frac{16}{15}b^5u'^2+\frac{8}{3}b^3u^2-\frac{8}{3}b^3u\right]dx.$$

Euler's equation for this functional

$$u''-\frac{5}{2b^2}u=-\frac{5}{4b^2}$$

is a linear equation with constant coefficients, the general solution of which is of the form

$$u=C_1\cosh\sqrt{\frac{5}{2}}\,\frac{x}{b}+C_2\sinh\sqrt{\frac{5}{2}}\,\frac{x}{b}+\frac{1}{2}.$$

The constants C_1 and C_2 are determined from the boundary conditions $z(-a)=z(a)=0$, whence $C_2=0$, $C_1=-\dfrac{1}{2\cosh\sqrt{\frac{5}{2}}\,\frac{a}{b}}$

and we finally get

$$u=\frac{1}{2}\left(1-\frac{\cosh\sqrt{\frac{5}{2}}\,\frac{x}{b}}{\cosh\sqrt{\frac{5}{2}}\,\frac{a}{b}}\right);$$

hence,

$$z_1 = \frac{1}{2}(b^2 - y^2)\left(1 - \frac{\cosh\sqrt{\frac{5}{2}}\,\frac{x}{b}}{\cosh\sqrt{\frac{5}{2}}\,\frac{a}{b}}\right).$$

If a more exact answer is required, the solution may be sought in the form

$$z_2 = (b^2 - y^2)\,u_1(x) + (b^2 - y^2)^2\,u_2(x).$$

Example 2. Find a continuous solution of the equation $\Delta z = -1$ in the domain D, which is an isosceles triangle bounded by the straight lines $y = \pm\frac{\sqrt{3}}{3}x$ and $x = b$ (Fig. 10.5), which solution vanishes on the boundary of the domain.

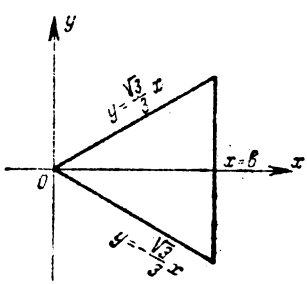


Fig. 10.5

The equation $\Delta z = -1$ is the Ostrogradsky equation for the functional

$$v[z] = \int_0^b \int_{-\frac{\sqrt{3}}{3}x}^{\frac{\sqrt{3}}{3}x} \left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 - 2z\right] dx\,dy,$$

and on the boundary of the integration domain $z = 0$. Proceeding by the Kantorovich method, we shall seek the first approximation in the form

$$z_1 = \left[y^2 - \left(\frac{\sqrt{3}}{3}x\right)^2\right] u(x).$$

For such a choice of z_1 the boundary conditions on the straight lines $y = \pm\frac{\sqrt{3}}{3}x$ are satisfied.

After integration with respect to y, the functional $v[z_1]$ takes the form

$$v[z_1] = \frac{8\sqrt{3}}{405}\int_0^b (2x^5u'^2 + 10x^4uu' + 30x^3u^2 + 15x^3u)\,dx.$$

Euler's equation for this functional will be $x^2u'' + 5xu' - 5u = \frac{15}{4}$. Linear equations of this type are called Euler's equations in the theory of differential equations (page 116).

One particular solution of this nonhomogeneous equation is obvious: $u=-\frac{3}{4}$. We seek a solution of the corresponding homogeneous equation in the form $u=x^k$ and finally get $u=C_1x+C_2x^{-5}-\frac{3}{4}$. Since near the point $x=0$ the solution u must be bounded, it follows that C_2 should be chosen equal to zero, and from the condition $u(b)=0$ we get $C_1=-\frac{3}{4b}$. Thus,

$$z_1=-\frac{3}{4}\left(1-\frac{x}{b}\right)\left(y^2-\frac{1}{3}x^2\right).$$

Note. Boundary-value problems are approximated by yet another direct method (which is not variational). It is called Galerkin's method (B. Galerkin). This method is particularly convenient in the solution of linear boundary-value problems, but can be also applied to many nonlinear problems. For the sake of definiteness we give *Galerkin's method* as applied to frequently encountered linear equations of the second order

$$y''+p(x)y'+q(x)y=f(x) \tag{10.1}$$

with homogeneous boundary conditions $y(x_0)=0$, $y(x_1)=0$ (the nonhomogeneous boundary conditions $y(x_0)=y_0$, $y(x_1)=y_1$ by the change of variables

$$z=y-y_0-\frac{y_1-y_0}{x_1-x_0}(x-x_0)$$

are readily reduced to homogeneous conditions).

Equation (10.1) can be written briefly as

$$L(y)=f(x).$$

On the interval $[x_0, x_1]$ let us choose a complete system of continuous linearly independent functions

$$w_1(x),\ w_2(x),\ \ldots,\ w_n(x),\ \ldots, \tag{10.2}$$

that satisfy the boundary conditions $w_n(x_0)=w_n(x_1)=0$ $(n=1, 2, \ldots)$. We will seek an approximate solution of the boundary-value problem in the form of a linear combination of the first n functions of the system (10.2):

$$y_n=\sum_{i=1}^{n}\alpha_i w_i(x).$$

Substitute y_n into equation (10.1) and choose the coefficients α_i $(i=1, 2, \ldots, n)$ so that the function

$$L\left(\sum_{i=1}^{n} \alpha_i w_i\ (x)\right)-f(x)$$

is orthogonal on the interval $[x_0, x_1]$ of each of the functions $w_i(x)$ $(i=1, 2, \ldots, n)$

$$\int_{x_0}^{x_1}\left[L\left(\sum_{i=1}^{n} \alpha_i w_i(x)\right)-f(x)\right] w_i(x)\,dx=0 \qquad (i=1, 2, \ldots, n). \quad (10.3)$$

It is natural to expect that y_n tends to the exact solution

$$\tilde{y}=\sum_{i=1}^{\infty} \alpha_i w_i(x),$$

as $n \to \infty$, since if the series obtained converges and admits two times termwise differentiation, then the function $L(\tilde{y})-f(x)$ is orthogonal on the interval $[x_0, x_1]$ of each function $w_i(x)$ of the system (10.2), and since the system (10.2) is complete, it follows that $L(\tilde{y})-f(x) \equiv 0$, and this signifies that $\tilde{y}$ is a solution of the equation (10.1). Obviously, $\tilde{y}$ also satisfies the boundary conditions $\tilde{y}(x_0)=\tilde{y}(x_1)=0$ [since all the $w_i(x_0)=w_i(x_1)=0$].

Only very rarely is it possible to determine all the α_i from the system (10.3) that is linear in them and to pass to the limit as $n \to \infty$; for this reason, one ordinarily confines oneself to a finite (very small) number n $(n=2, 3, 4, 5,$ and sometimes even $n=1)$.

Here, of course, one has to choose only n functions $w_i(x)$, and so the condition of completeness is discarded and one has only to choose them linearly independent and satisfying the boundary conditions

$$w_i(x_0)=w_i(x_1)=0.$$

Very often, for these so-called coordinate functions we take the polynomials

$$(x-x_0)(x-x_1),\ (x-x_0)^2(x-x_1),\ (x-x_0)^3(x-x_1),\ \ldots \\ \ldots,\ (x-x_0)^n(x-x_1),\ \ldots \quad (10.4)$$

[it is convenient here to transfer the coordinate origin to the point x_0, and then in (10 4) $x_0=0$] or the trigonometric functions

$$\sin\frac{n\pi(x-x_0)}{x_1-x_0} \qquad (n=1, 2, \ldots).$$

This method is applicable to equations of any order n, to systems of equations, and to partial differential equations.

PROBLEMS ON CHAPTER 10

1. Find an approximate solution of the equation $\Delta z = -1$ inside the square $-a \leqslant x \leqslant a$, $-a \leqslant y \leqslant a$, which vanishes on the boundary of the square.

Hint. The problem reduces to investigating for an extremum the functional

$$\iint_D \left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 - 2z\right] dx\,dy.$$

An approximate solution may be sought in the form

$$z_0 = \alpha (x^2 - a^2)(y^2 - a^2).$$

2. Find an approximate solution of the problem of the extremum of the functional

$$v[y(x)] = \int_0^1 (x^3 y''^2 + 100xy^2 - 20xy)\,dx; \quad y(1) = y'(1) = 0$$

Hint The solution may be sought in the form

$$y_n(x) = (x-1)^2 (\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n);$$

carry out the calculations for $n = 1$.

3. Find an approximate solution of the problem of the minimum of the functional

$$v[y(x)] = \int_0^1 (y'^2 - y^2 - 2xy)\,dx; \quad y(0) = y(1) = 0,$$

and compare it with the exact solution.

Hint. The approximate solution may be sought in the form

$$y_n = x(1-x)(\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n);$$

carry out the calculations for $n = 0$ and $n = 1$.

4. Find an approximate solution of the problem of the extremum of the functional

$$v[y(x)] = \int_1^2 \left(xy'^2 - \frac{x^2-1}{x} y^2 - 2x^2 y\right) dx; \quad y(1) = y(2) = 0,$$

and compare it with the exact solution.

Hint. The solution may be sought in the form

$$y = \alpha (x - 1)(x - 2).$$

5. Using the Ritz method, find an approximate solution of the problem of the minimum of the functional

$$v[y(x)] = \int_0^2 (y'^2 + y^2 + 2xy)\,dx; \quad y(0) = y(2) = 0,$$

and compare it with the exact solution.

Hint. See Problem 3.

6. Using the Ritz method, find an approximate solution of the differential equation $y'' + x^2 y = x$; $y(0) = y(1) = 0$. Determine $y_2(x)$ and $y_3(x)$ and compare their values at the points $x = 0.25$, $x = 0.5$, and $x = 0.75$.

Answers to problems

CHAPTER 1

1. $\sin y \cos x = c$. **2.** $6x^2 + 5xy + y^2 - 9x - 3y = c$. **3.** $x^2 - 2cy = c^2$. **4.** $y = \frac{c}{x} + \frac{x^3}{4}$. **5.** $\frac{y^2}{2} + \frac{y}{x} = c$. **6.** $x = ce^{-3t} + \frac{1}{5}e^{2t}$. **7.** $y = c\cos x + \sin x$. **8.** $e^x - e^y = c$. **9.** $x = ce^t - \frac{1}{2}(\cos t + \sin t)$. **10.** Homogeneous equation: $x = ye^{cy+1}$. **11.** $y = cx$ and $y^2 - x^2 = c$. **12.** $y^2 = \frac{1}{(3x+c)^2}$. **13.** $\ln|t| = c - e^{-\frac{x}{t}}$. **14.** A parameter may be introduced, putting $y' = \cos t$ $\begin{cases} x = \sin t, \\ y = \frac{t}{2} + \frac{\sin 2t}{4} + c. \end{cases}$ **15.** $y = cx + \frac{1}{c}$; singular solution $y^2 = 4x$. **16.** $\begin{cases} x = p^3 - p + 2, \\ y = \frac{3}{4}p^4 - \frac{p^2}{2} + c. \end{cases}$ **17.** Equation is linear in x and $\frac{dx}{dy}$, $x = cy + \frac{y^3}{2}$. **18.** $\begin{cases} x = \frac{4}{3}p^3 - \frac{3}{2}p^2 + c, \\ y = p^4 - p^3 - 2. \end{cases}$ **19** Hyperbolas $x^2 - y^2 = c$. **20.** The differential equation of the required curves is $\frac{y}{2x} = y'$. *Ans.* $y^2 = 2cx$. **21.** The differential equation of the required curves is $y - xy' = x$. *Ans.* $y = cx - x\ln|x|$. **22.** $x^2 + y^2 - 2cy = 0$. The problem is solved very simply in polar coordinates. **23.** The differential equation of the problem is $\frac{dT}{dt} = k(T - 20)$. *Ans.* In one hour. **24.** The differential equation of the problem is $\frac{dv}{dt} = kv$, where v is the velocity. *Ans.* $v \approx 0.466$ km/hr. **25.** If the origin is put in the given point and the x-axis directed parallel to the direction given in the problem, then the differential equation of the curves, the rotation of which forms the desired surface, is of the form $y' = \frac{-x \pm \sqrt{x^2+y^2}}{y}$ (or $dx - d_\rho = 0$, where $\rho = \sqrt{x^2+y^2}$). *Ans.* The axial section of the desired surface is defined by the equation $y^2 = 2cx + c^2$, the surface is a paraboloid of revolution. **26.** $y = 2\sin(x - c)$. **27.** The differential equation of the desired curves is $y' = -\frac{y}{x}$. *Ans.* Hyperbolas $xy = c$. **28.** $(x + y + 1)^3 = c(x - y + 3)$. **29.** $y = \frac{2(1+x)}{c + 2x + x^2}$. **30.** $y(0.5) \approx 0.13$. **31.** $y(0.6) \approx 0.07$. **32.** $y(0.02) \approx 1.984$; $y(0.04) \approx 1.970$; $y(0.06) \approx 1.955$; $y(0.08) \approx 1.942$; $y(0.10) \approx 1.930$; $y(0.12) \approx 1.917$; $y(0.14) \approx 1.907$; $y(0.16) \approx 1.896$; $y(0.18) \approx 1.886$; $y(0.20) \approx 1.877$;

$y(0.22) \approx 1.869$; $y(0.24) \approx 1.861$; $y(0.26) \approx 1.854$; $y(0.28) \approx 1.849$; $y(0.30) \approx$ ≈ 1.841. **33.** $\begin{cases} x = \dfrac{c}{p^2} + \dfrac{2p}{3}, \\ y = 2px - p^2 \end{cases}$ and $y = 0$. **34.** $x + \cot\dfrac{x-y}{2} = c$. **36.** $(x+y+$ $+1)^3 = ce^{2x+y}$. **37.** $y = c$; $y = e^x + c$; $y = -e^x + c$. **38.** $y^2 = 2cx + c^2$. **39.** No. **40.** $y_1 = \dfrac{x^2-1}{2}$; $y_2 = \dfrac{x^2-1}{2} + \dfrac{2}{15} - \dfrac{1}{4}x + \dfrac{x^3}{6} - \dfrac{x^5}{20}$. **41.** $y = 2x^2 - x$. **42.** No. **43.** $x = ce^{\frac{x}{y}}$. **44.** $x^2 + \dfrac{3y^2}{2} = c^2$. **45.** $x = 2t$. **46.** $x = t^2$. **47.** $y = -x + 1$ and $y = -\dfrac{x^2}{4}$. **48.** A real solution does not exist. **49.** $3x - 4y + 1 = ce^{x-y}$. **50.** $x = (4t + c)\sin t$. **51.** $y = cx + \dfrac{c^2 - x^2}{2}$ and the singular solution $y = -x^2$. **52.** $y = \dfrac{7x^3}{x^7 + c}$, $y = 0$. **53.** $x - c = \dfrac{a}{2}(2t - \sin 2t)$, $y = \dfrac{a}{2}(1 - \cos 2t)$ is a family of cycloids. A singular solution is $y = a$. *Hint*: it is convenient to introduce a parameter t, putting $y' = \cot t$. **54.** $3(x^2 + y) + xy^3 = cx$. **55.** $\mu = \dfrac{c}{(y^2 + x)^3}$. **56.** $x = ce^{\frac{x}{y}}$. **57.** $x^2 + 2xy - y^2 - 6x - 2y = c$. **58.** $y = \dfrac{1}{1 + cx + \ln x}$ and $y = 0$. **59.** $(x^2 - 1)y - \sin x = c$. **60.** $8y + 4x + 5 = ce^{4x-8y-4}$. **61.** $y^3 + x^3 - 3xy = c$. **62.** $y = c(x^2 + y^2)$. **63.** $y^3 = x + \dfrac{c}{x}$. **64.** $y = c(x + a) + c^2$ and a singular solution is $y = -\dfrac{(x+a)^2}{4}$. **65.** $x = \dfrac{2}{3}t + \dfrac{c}{t^2}$, $y = 2xt - t^2$ and $y = 0$, $y = \dfrac{3}{4}x^2$ **66.** $y =$ $= \dfrac{c}{1 \pm \cos x}$.

CHAPTER 2

1. $y = 5e^{3x}\sin x + 10$. **2.** $x = c_1 \cos t + c_2 \sin t + \dfrac{1}{3}\cos 2t - \dfrac{t\cos t}{2}$. **3.** $(y - c_3)^2 = c_1 x + c_2$. **4.** $y = c_1 \cos x + c_2 \sin x + \dfrac{\cos^2 x}{\sin x} - \dfrac{1}{2\sin x}$. **5.** $y = c_1 x^2 +$ $+ c_2 x^3 + \dfrac{1}{3}$. **6.** $y = c_1 \sin x + c_2 \cos x + \dfrac{1}{2}\cosh x$. **7.** $y = \dfrac{1}{c_1 x + c_2} + 1$. **8.** $x =$ $= e^{2t}(c_1 + c_2 t) + \dfrac{t^2 e^{2t}}{2} + e^t + \dfrac{1}{4}$. **9.** $= -\dfrac{x}{c_1} + \dfrac{c_1^2 + 1}{c_1^2}\ln|1 + c_1 x| + c_2$. **10.** $c_1 x^2 +$ $+ 1 = c_1^2 (t + c_2)^2$. **11.** $y = c_1 e^{2x} + c_2 e^{-2x} + c_3 \cos 2x + c_4 \sin 2x - \dfrac{x^2}{16} + \dfrac{1}{15}e^x$. **12.** $y = \cos(x - c_1) + c_2 x + c_3$. **13.** $y = c_1 e^x + c_2 e^{-x} + c_3 x^3 + c_4 x^2 + c_5 x + c_6 - \dfrac{x^4}{24}$. **14.** $x = e^t(c_1 + c_2 t) + e^{-t}(c_3 + c_4 t) + 1 + t^2$. **15.** $y = c_0\left(1 - \dfrac{4x^3}{2\cdot 3} + \dfrac{4^2 x^6}{2\cdot 3\cdot 5\cdot 6} - \ldots\right.$ $\left.\ldots + \dfrac{(-1)^k 4^k x^{3k}}{2\cdot 3\cdot 5\cdot 6\ \ldots\ (3k-1)\,3k} + \ldots\right) + c_1\left(x - \dfrac{4x^4}{3\cdot 4} + \dfrac{4^2 x^7}{3\cdot 4\cdot 6\cdot 7} - \ldots +\right.$ $\left.+ \dfrac{(-1)^k 4^k x^{3k+1}}{3\cdot 4\cdot 6\cdot 7\ \ldots\ 3k(3k+1)} + \ldots\right)$. **16.** $y = c_1 J_{\frac{1}{5}}(3x) + c_2 J_{-\frac{1}{5}}(3x)$. **17.** $y = x$.

18. $y=\left(\frac{1}{2}x+1\right)^4$. 19. $y=c_1\cos x+c_2\sin x+1+x\cos x-\sin x\ln|\sin x|$.

20. $u=\frac{c_1}{r}+c_2$. 21. The differential equation of the problem is $\frac{d^2r}{dt^2}=\frac{k}{r^2}$ or $v\frac{dv}{dr}=\frac{k}{r^2}$, where r is the distance from the centre of the earth to the body, v is the velocity, and $k=-6400^2\,g$. *Ans.* $v\approx 11$ km/s. 22. The differential equation of motion is $\frac{d^2x}{dt^2}=-g+k\left(\frac{dx}{dt}\right)^2$. *Ans.* $x=\frac{75^2}{g}\ln\cosh\frac{g}{75}t$. 23. The differential equation of motion is $\frac{d^2s}{dt^2}=k(s+1)$ or $\frac{d^2s}{dt^2}=\frac{g}{6}(s+1)$. *Ans.* $t=\sqrt{\frac{6}{g}}\ln(6+\sqrt{35})$. 24. $t=\frac{3}{\sqrt{g}}\ln(9+\sqrt{80})$. 25. $s=\frac{F-a}{b}t-$ $-\frac{(F-a)p}{b^2g}\left(1-e^{-\frac{bg}{p}t}\right)$. 26. $x=A\cos\sqrt{\frac{g}{a}}t$. 27. $x=a\cos\sqrt{\frac{g}{a}}t$.

28. The differential equation of motion is $\ddot{x}+k_1\dot{x}-k_2x=0$, $k_2>0$. *Ans.* $x=c_1e^{\left(-\frac{k_1}{2}+\sqrt{\frac{k_1^2}{4}+k_2}\right)t}+c_2e^{\left(-\frac{k_1}{2}-\sqrt{\frac{k_1^2}{4}+k_2}\right)t}$. 29. $x=$ $=12\sum_{n=1}^{\infty}\frac{(-1)^n\sin nt}{(n^2-2)n^3}$. 30. $y^2=c_1(x^2+x\sqrt{1+x^2}+\ln|x+\sqrt{1+x^2}|)+c_2$.

31. $y=c_2e^{c_1x}+c_1$, $y=\frac{4}{c-x}$. 32. $x=c_1\cos 3t+c_2\sin 3t-\frac{1}{12}t^2\cos 3t+\frac{1}{36}\sin 3t$.

33. $y=e^{-x}\left(c_1+c_2x-\frac{1}{4}x^2\right)+\frac{1}{8}e^x$. 34. $y=c_1e^x+e^{-\frac{1}{2}x}\left(c_2\cos\frac{\sqrt{3}}{2}x+\right.$ $\left.+c_3\sin\frac{\sqrt{3}}{2}x\right)+\frac{1}{3}xe^x$. 35. $y=e^x(c_1\cos x+c_2\sin x)+\frac{xe^x\cos x}{4}+\frac{x^2e^x\sin x}{4}$.

36 $y=c_1(x-x^3)+c_2\left[4-6x^2+3(x^3-x)\right]\ln\left|\frac{x+1}{x-1}\right|\Big]-\frac{1}{6}$. 37. $u=$ $=c_1\ln(x^2+y^2)+c_2$. 38. $u=\frac{c_1}{\sqrt{x^2+y^2+z^2}}+c_2$. 39. The differential equation of motion is $m\ddot{x}=mg-k\dot{x}$. *Ans.* $x=\frac{mg}{k}t-\frac{m^2g}{k^2}\left(1-e^{-\frac{k}{m}t}\right)$.

40 (a) $t-t_0=\int_{x_0}^{x}\frac{dx}{\sqrt{v^2+\frac{2}{m}\int_{x_0}^{x}f(x)\,dx}}$. (b) $x-x_0=m\int_{v_0}^{v}\frac{v\,dv}{f(v)}$; $t-t_0=$ $=m\int_{v_0}^{v}\frac{dv}{f(v)}$, where $v=\dot{x}$. 41. $y=c_1+c_2x+c_3x^2+e^x(c_4+c_5x+c_6x^2)-\frac{x^3}{2}-\frac{x^4}{24}$.

42. $x=(c_1+c_2t)\cos t+(c_3+c_4t)\sin t-\frac{1}{8}t^2\cos t$. **43.** $y=c_1\cos\ln(1+x)+$ $+c_2\sin\ln(1+x)+\ln(1+x)\sin\ln(1+x)$. **44.** $x=\sum_{n=1}^{\infty}\frac{(2-n^2)\sin nt-2n\cos nt}{[(2-n^2)^2+4n^2]\,n^4}$.

45. $x=\frac{\alpha_0}{2a_2}+\sum_{n=1}^{\infty}\left[\frac{-(n^2-a_2)\,\alpha_n-a_1n\beta_n}{(n^2-a_2)^2+a_1^2n^2}\cos nt+\frac{a_1n\alpha_n-(n^2-a_2)\,\beta_n}{(n^2-a_2)^2+a_1^2n^2}\sin nt\right]$,

where α_0, α_n, β_n are Fourier's coefficients of the function $f(t)$. **46.** $x=\frac{\cos t}{2}+$ $+\frac{\mu}{24}(1+3\cos 2t)$. **47.** $y=c_1x+c_2xe^{-\frac{1}{x}}$. **48.** $x^2y''+xy'-y=0$.

49. $x=e^{\frac{\sqrt{2}}{2}t}\left(c_1\cos\frac{\sqrt{2}}{2}t+c_2\sin\frac{\sqrt{2}}{2}t\right)+e^{-\frac{\sqrt{2}}{2}t}\left(c_3\cos\frac{\sqrt{2}}{2}t+\right.$ $\left.+c_4\sin\frac{\sqrt{2}}{2}t\right)+t^3$. **50.** $x=t^3+t+1$, $y=\frac{3}{8}t^6+\frac{3}{10}t^5+\frac{3}{16}t^4+\left(c_1+\frac{1}{6}\right)t^3+$ $+c_1t+c_2$. **51** $x=(c_1+c_2t)\,e^{-5t}+\frac{2^t}{(5+\ln 2)^2}+\frac{t^3e^{-5t}}{6}$. **52.** $y=c_2e^{c_1x^2}$.

53. $y=c_1e^x+c_2e^{-x}+e^{\frac{x}{2}}\left(c_3\cos\frac{\sqrt{3}}{2}x+c_4\sin\frac{\sqrt{3}}{2}x\right)+e^{-\frac{x}{2}}\left(c_5\cos\frac{\sqrt{3}}{2}x+\right.$ $\left.+c_6\sin\frac{\sqrt{3}}{2}x\right)+\frac{e^{2x}}{63}$. **54.** $y=(c_1x+c_2)\cos x+(c_3x+c_4)\sin x+c_5+c_6x+$ $+\frac{x^3}{6}+\frac{1}{4}e^x$. **55.** $y=(c_1x+c_2)^8+c_3x+c_4$. **56.** $y=e^{1+c_1x}\left(\frac{x}{c_1}-\frac{1}{c_1^2}\right)+c_2$.
57. $y=c_1\cos x+c_2\sin x-\frac{\sin 2x}{6}-\frac{\sin 4x}{30}$. **58.** $y=-\frac{1}{x-2}$. **59.** $y=c_2e^{c_1x}+\frac{1}{c_1}$.

CHAPTER 3

1. $x=\sin t$, $y=\cos t$. **2.** $x_1=2e^t$, $x_2=2e^t$. **3** $x=c_1e^{(-1+\sqrt{15})\,t}+$ $+c_2e^{(-1-\sqrt{15})\,t}+\frac{2}{11}e^t+\frac{1}{6}e^{2t}$; we find y from the first equation: $y=e^t-\frac{dx}{dt}-5x$. **4** $x=c_1e^t+e^{-\frac{1}{2}t}\left(c_2\cos\frac{\sqrt{3}}{2}t+c_3\sin\frac{\sqrt{3}}{2}t\right)$; y and z are found from the equations $y=\frac{dx}{dt}$, $z=\frac{d^2x}{dt^2}$. **5.** $x=c_1e^{c_2t}$; $y=c_1c_2e^{c_2t}$.
6. $x=c_1\cos t+c_2\sin t+3$; $y=-c_1\sin t+c_2\cos t$. **7.** $y=c_1J_0(x)+c_2Y_0(x)$; $z=x\left[c_1J_0'(x)+c_2Y_0'(x)\right]$. **8.** $x+y+z=c_1$, $x^2+y^2+z^2=c_2^2$. **9.** $x=c_1e^t+c_2e^{-2t}$, $y=c_1e^t+c_3e^{-2t}$; $z=c_1e^t-(c_2+c_3)\,e^{-2t}$. **10.** $x=c_1t+\frac{c_2}{t}$; $y=-c_1t+\frac{c_2}{t}$.
11 $x=c_1\cos t+c_2\sin t-t\cos t+\sin t\ln|\sin t|$; y is determined from the equation $y=\frac{dx}{dt}-1$. **12.** $x^2-y^2=c_1$, $y-x-t=c_2$. **13.** $x=c_1e^t+c_2e^{-t}+\sin t$;

$y = -c_1e^t + c_2e^{-t}$. **14.** $x = e^t$; $y = 4e^t$. **15.** $0(1) \approx 0.047$. **16.** $x = e^{at}(c_1 \cos t + c_2 \sin t)$, $y = e^{at}(c_1 \sin t - c_2 \cos t)$. **17.** $x = 2c_1e^{-t} + c_2e^{-7t}$, $y = -c_1e^{-t} + c_2e^{-7t}$. **18.** $x = e^{-6t}(2c_1 \cos t + 2c_2 \sin t)$, $y = e^{-6t}[(c_1 - c_2)\cos t + (c_1 + c_2)\sin t]$. **19.** $x = c_1e^t + c_2$, $y = (c_1t + c_3)e^t - t - 1 - c_2$, $z = y - c_1e^t$. **20.** $x + y + z = c_1$, $xyz = c_2$. **21.** $x^2 + y^2 + z^2 = c_1^2$, $xyz = c_2$. **22.** $X = \left\| \begin{matrix} c_1e^t + c_2e^{-t} \\ c_1e^t + 3c_2e^{-t} \end{matrix} \right\|$.

CHAPTER 4

1. The rest point is asymptotically stable. **2.** The rest point is unstable. **3.** For $\alpha < -\frac{1}{2}$, the rest point is asymptotically stable, for $\alpha = -\frac{1}{2}$ stable, and for $\alpha > -\frac{1}{2}$ unstable. **4.** For $\alpha \leqslant 0$, the rest point is asymptotically stable, for $\alpha > 0$ unstable. **5.** For $1 < t < 2$ $x(t, \mu) \to \sqrt{4 - t^2}$; for $2 < t < 3$ $x(t, \mu) \to -\sqrt{9 - t^2}$; for $t > 3$ $x(t, \mu) \to \infty$. **6.** $x(t, \mu) \to \infty$. **7.** The rest point is unstable. **8.** The rest point is stable. **9.** The rest point is unstable. **10.** The rest point is stable. **11.** Saddle point. **12.** The periodic solution $x = \frac{1}{5}\sin t - \frac{2}{5}\cos t$ is asymptotically stable. **13.** All solutions, including periodic solutions, are asymptotically stable. **14.** The rest point is unstable. The function $v = x^4 - y^4$ satisfies the conditions of the Chetayev theorem. **15.** All solutions are unstable. **16.** The solution $x \equiv 0$ is unstable. **17.** For $1 < \alpha < 2$ the solution $x \equiv 0$ is asymptotically stable. For $\alpha = 1$ and for $\alpha = 2$ the solution $x \equiv 0$ is stable. For $\alpha > 2$ and for $\alpha < 1$ the solution $x \equiv 0$ is unstable. **18.** The solution $x \equiv 0$, $y \equiv 0$ is stable for constantly acting perturbations. The function $v = 4x^2 + 3y^2$ satisfies the conditions of Malkin's theorem. **19.** The solution $x(t) \equiv 0$ is unstable. **20.** All solutions are stable, but there is no asymptotic stability. **21.** All solutions are stable, but there is no asymptotic stability. **22.** The periodic solution $x = \frac{\cos t - \sin t}{2}$ is unstable. **23.** The region of stability is $0 \leqslant \alpha \leqslant 1$, the region of asymptotic stability is $0 < \alpha < 1$. **24.** The region of stability is $\alpha \geqslant 5$, the region of asymptotic stability is $\alpha > 5$.

CHAPTER 5

1. $z = \Phi(x + y)$. **2.** $z = e^{2x}\Phi(x - y)$. **3.** $z = e^{\frac{y}{x}}\Phi(x)$. **4.** $\Phi\left(z,\ ye^{\frac{x}{z}}\right) = 0$. **5.** $z = 5 + \frac{\Phi(x^3y^5)}{y^5}$. **6.** $u = \Phi(x - y,\ y - z)$. **7.** $u = x^4\Phi\left(\frac{y}{x^2}, \frac{z}{x^3}\right)$. **8.** $z = x\Phi_1(y) + \Phi_2(y)$. **9.** $z = (x^2 + y - 1)^2$. **10.** $z = ye^{\frac{x-2}{y}}$. **11.** $z = 3x$. **12.** $z = \left(y^2 - \frac{2x}{z}\right)^{3/2}$. **13.** $\Phi(z^2 + x^2,\ x^2 - y^2) = 0$. **14.** $\Phi(z^2 - x^2,\ x^2 - y^2) = 0$. **15.** No. **16.** $2xy + y^2 + 6xz^2 = c$. **17** $z = ax^3 + \frac{y^3}{9a} + b$ (other answers are also possible). **18.** $z = ax + by + a^3b^3$ (other answers are also possible). **19.** $z = be^{\frac{3}{a}(a^2x + y)}$ (other answers are also possible). **20.** $z = x\sin a + ay + b$ (other answers are also possible). **21.** $x^2y - 3xyz = c$. **22.** There is no such family of surfaces, since the condition $(\mathbf{F} \cdot \operatorname{rot} \mathbf{F}) = 0$ is not fulfilled. **23.** The equations of the vector lines are $\frac{y}{x} = c_1$, $xz = c_2$. The equation of the vector surfaces is

$z=\frac{1}{x}\Phi\left(\frac{y}{x}\right)$. The equation of surfaces orthogonal to vector lines is $x^2+y^2-z^2=c$. **24.** $z=xy+1$. **25.** $z=3xy$. **26.** $z=x^2+y^2$.

CHAPTER 6

1. The extremals are the circles $(x-C_1)^2+y^2=C_2^2$. **2.** The integral is independent of the path of integration. The variation problem is meaningless. **3.** An extremum is not achieved in the class of continuous functions. **4.** The extremals are the hyperbolas $y=\frac{C_1}{x}+C_2$. **5.** $y=C_1\sin(4x-C_2)$. **6.** $y=-\frac{x^2}{4}+C_1x+C_2$. **7.** $y=\sinh(C_1x+C_2)$. **8.** $y=C_1e^x+C_2e^{-x}+\frac{1}{2}\sin x$. **9.** $y=C_1e^{2x}+C_2e^{-2x}+C_3\cos 2x+C_4\sin 2x$. **10.** $y=\frac{x^7}{7!}+C_1x^5+C_2x^4+C_3x^3+C_4x^2+C_5x+C_6$. **11.** $y=(C_1x+C_2)\cos x+(C_3x+C_4)\sin x$, $z=2y+y''$, whence z is readily determined. **12.** $\frac{\partial^2 z}{\partial x^2}-\frac{\partial^2 z}{\partial y^2}=0$ **13.** $\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=f(x, y, z)$. **14.** $y=C_1x^4+C_2$. **15.** $y=\frac{1}{2}xe^x+C_1e^x+C_2e^{-x}$. **16** $y=-\frac{x\cos x}{2}+C_1\cos x+C_2\sin x$. **17.** $y=C_1\cosh x+C_2\sinh x+x\sinh x-\cosh x\ln\cosh x$. **18.** $y=C_1x+\frac{C_2}{x^2}+\frac{1}{3}x\ln|x|$. **19.** $y=(C_1+C_2x)\cos x+(C_3+C_4x)\sin x-\frac{x^2\sin x}{4}$. **20.** $y=C_1e^x+C_2e^{-x}+e^{\frac{x}{2}}\left(C_3\cos\frac{\sqrt{3}}{2}x+C_4\sin\frac{\sqrt{3}}{2}x\right)+e^{-\frac{x}{2}}\left(C_5\cos\frac{\sqrt{3}}{2}x+C_6\sin\frac{\sqrt{3}}{2}x\right)+x^3$.

CHAPTER 7

1. $y=-x$ for $0\leqslant x\leqslant 1$; $y=x-2$ for $1<x\leqslant 4$ and $y=x$ for $0\leqslant x\leqslant 3$, $y=-x+6$ for $3<x\leqslant 4$. The functional achieves an absolute minimum on both polygonal lines. **2.** No. **3.** The polygonal lines passing through the given boundary points are composed of rectilinear segments with slopes $\sqrt{3}$ and $-\sqrt{3}$. **4.** $\frac{\varphi'-y'}{1+y'\varphi'}=1$, i.e., the extremals must intersect the curve $y_1=\varphi(x_1)$, along which the boundary point slides, at an angle $\frac{\pi}{4}$. **5.** $y=\frac{x^5}{120}+\frac{1}{24}(x^2-x^3)$. **6.** $y=\pm\frac{3}{4}x$ for $0\leqslant x\leqslant\frac{16}{5}$; $y=\pm\sqrt{9-(x-5)^2}$ for $\frac{16}{5}<x\leqslant\frac{34}{5}$; $y=\mp\frac{3}{4}(x-10)$ for $\frac{34}{5}<x\leqslant 10$; that is, the curve consists of a segment of a straight line that is tangent to the circle, an arc of the circle, and again a segment tangent to the circle. **7.** $y\equiv 0$. **8.** Arcs of the circle $y=\pm\sqrt{8x-x^2}$.

CHAPTER 8

1 A strong minimum is achieved for $y=-\frac{x^2}{4}+1$. **2.** A strong minimum is achieved for $y=0$ if $0<a<\frac{\pi}{4}$, but if $a>\frac{\pi}{4}$, then there is no minimum.

3. An extremum is not achieved on continuous curves. 4. There is a strong minimum for $y=7-\frac{4}{x}$. 5. There is a strong minimum for $y=1$. 6. A strong maximum is achieved for $y=\sin 2x-1$. 7. A strong minimum is achieved for $y=x^3$. 8. A strong minimum is achieved for $y=\frac{1}{3}e^{3x}$. 9. A strong maximum is achieved for $y=\sin 2x$. 10. A weak minimum is achieved on the straight line $y=\frac{y_1}{x_1}x$. 11. A weak minimum is achieved on the straight line $y=\frac{y_1}{x_1}x$. 12. A weak minimum is achieved for $y=x^2$. 13. A strong maximum is achieved for $y=x^3-1$. 14. A strong minimum is achieved for $y=\frac{\sinh x}{\sinh 2}+x$.

CHAPTER 9

1. $y=\pm 2\sin n\pi x$, where n is an integer. 2. $\varphi=C_1+C_2z$; $r=R$. 3. $y=\lambda x^2+C_1x+C_2$, where C_1, C_2 and λ are determined from the boundary conditions and from the isoperimetric condition. 4. $\frac{d}{dx}(p(x)y')+[\lambda r(x)-q(x)]y=0$; $y(0)=0$; $y(x_1)=0$. The trivial solution $y\equiv 0$ does not satisfy the isoperimetric condition, and nontrivial solutions, as is known, exist only for certain values of λ called eigenvalues. Hence, λ must be an eigenvalue. One arbitrary constant of the general solution of Euler's equation is determined from the condition $y(0)=0$, the other, from the isoperimetric condition. 5. $y=-\frac{5}{2}x^2+\frac{7}{2}x$; $z=x$.

CHAPTER 10

1. $z_1=\frac{5}{16a^2}(x^2-a^2)(y^2-b^2)$. If greater accuracy is needed, then the solution may be sought in the form $z_2=(x^2-a^2)(y^2-b^2)[\alpha_0+\alpha_1(x^2+y^2)]$. 2. $y_1=(x-1)^2(0.124+0.218x)$. 3. The exact solution is $y=\frac{\sin x}{\sin 1}-x$. 4. The solution of Euler's equation is $y=3.6072J_1(x)+0.75195Y_1(x)-x$, where J_1 and Y_1 are Bessel functions. 5. The exact solution is $y=\frac{2\sinh x}{\sinh 2}-x$. 6. If the solution is sought in the form: $y_2=x(x-1)(\alpha_1+\alpha_2x)$, $y_3=x(x-1)(\alpha_1+\alpha_2x+\alpha_3x^2)$, then $y_2=x(x-1)(0.1708+0.17436x)$, $y_3=x(x-1)(0.1705+0.1760x-0.0018x^2)$. The values of y_2 and y_3 coincide at the specified points to within 0.0001.

Recommended Literature

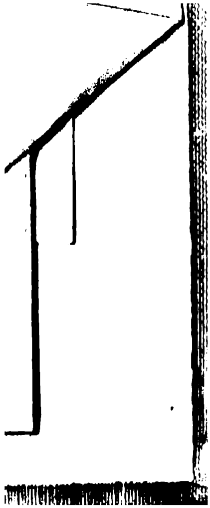
PART I

1. Petrovskii, I. *Lektsii po teorii obyknovennykh differentsial'nykh uravnenii* (Lectures on Ordinary Differential Equations). 5th ed., Nauka, 1964
2. Malkin, I. *Teoriya ustoychivosti dvizheniya* (Theory of Stability of Motion). Gostekhizdat, 1952 (to Ch. 4)
3. Malkin, I. *Nekotoryie zadachi teorii nelineinykh kolebanii* (Certain Problems in the Theory of Non-Linear Vibrations). Gostekhizdat, 1956 (to Sec. 8, Ch. 2)
4. Tikhonov, A. *O zavisimosti reshenii differentsial'nykh uravnenii ot malogo parametra* (On the Dependence of Solutions of Differential Equations on a Small Parameter). Matematicheskii sbornik, Vol. 22 (64):2 (1948) and Vol. 31 (72):3 (1952) (to Sec. 6, Ch. 4)
5. Stepanov, V. *Kurs differentsial'nykh uravnenii* (A Course of Differential Equations), 8th ed., Fizmatgiz, 1959
6. Krylov, A. *Lektsii o priblizhennykh vychisleniyakh* (Lectures on Approximate Calculations). 5th ed., Gostekhizdat, 1950 (to Sec. 7, Ch. 1 and Sec. 6, Ch. 3)
7. Berezin, I., Zhidkov, N. *Metody vychislenii* (Methods of Calculations). Vol. II. Fizmatgiz, 1960 (to Sec. 7, Ch. 1 and Sec. 6, Ch. 3)

PART II

1. Gelfand, I., Fomin, S. *Variatsionnoye ischisleniye* (The Calculus of Variations). Fizmatgiz, 1961
2. Lavrentiev, M., Lyusternik, L. *Kurs variatsionnogo ischisleniya* (A Course of the Calculus of Variations). 2nd ed. Gostekhizdat, 1950
3. Smirnov, V., Krylov, V., Kantorovich, L. *Variatsionnoye ischisleniye* (The Calculus of Variations). KUBUCH, 1933
4. Smirnov, V. *Kurs vysshey matematiki* (A Course of Higher Mathematics). Vol. 4, 4th ed. Fizmatgiz, 1958
5. Günther, N. *Kurs variatsionnogo ischisleniya* (A Course of the Calculus of Variations). Gostekhizdat, 1941
6. Akhiezer, N. *Lektsii po variatsionnomu ischisleniyu* (Lectures on the Calculus of Variations). Gostekhizdat, 1955
7. Lavrentiev, M., Lyusternik, L. *Osnovy variatsionnogo ischisleniya* (Fundamentals of the Calculus of Variations). Parts 1 and 2. Gostekhizdat, 1935
8. Pontryagin, L., Boltyanskii, V., Gamkrelidze, R., Mishchenko, E. *Matematicheskaya teoriya optimal'nykh protsessov* (Mathematical Theory of Optimal Processes). Fizmatgiz, 1961
9. Bellman, R. Dynamic Programming, Princeton University Press, 1957
10. Kantorovich, L., Krylov, V. *Priblizhennyie metody vysshego analiza* (Approximate Methods of Higher Analysis). 5th ed. Fizmatgiz, 1962
11. Mikhlin, S. *Pryamyie metody v matematicheskoy fizike* (Direct Methods in Mathematical Physics). Gostekhizdat, 1950

Index



L. ELSGOLTS

Differential Equations and the Calculus of Variations