Some Problems from the Midterm Examinations

1 Introduction to Analysis (Numbers, Functions, Limits)

Problem 1 The length of a hoop girdling the Earth at the equator is increased by 1 meter, leaving a gap between the Earth and the hoop. Could an ant crawl through this gap? How big would the absolute and relative increases in the radius of the Earth be if the equator were lengthened by this amount? (The radius of the Earth is approximately 6400 km.)

Problem 2 How are the completeness (continuity) of the real numbers, the unboundedness of the series of natural numbers, and Archimedes' principle related? Why is it possible to approximate every real number arbitrarily closely by rational numbers? Explain using the model of rational fractions (rational functions) that Archimedes' principle may fail, and that in such number systems the sequence of natural numbers is bounded and there exist infinitely small numbers.

Problem 3 Four bugs sitting at the corners of the unit square begin to chase one another with unit speed, each maintaining a course in the direction of the one pursued. Describe the trajectories of their motions. What is the length of each trajectory? What is the law of motion (in Cartesian or polar coordinates)?

Problem 4 Draw a flow chart for computing $\sqrt{a}$ $(a > 0)$ by the recursive procedure

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

How is equation solving related to finding fixed points? How do you find $\sqrt[n]{a}$?

Problem 5 Let $g(x) = f(x) + o(f(x))$ as $x \to \infty$. Is it also true that $f(x) = g(x) + o(g(x))$ as $x \to \infty$?

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

Problem 6 By the method of undetermined coefficients (or otherwise) find the first few (or all) coefficients of the power series for $(1+x)^\alpha$ with $\alpha = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}$. (By interpolating the coefficients of like powers of x in such expansions, Newton wrote out the law for forming the coefficients with any $\alpha \in \mathbb{R}$. This result is known as Newton's binomial theorem.)

Problem 7 Knowing the power-series expansion of the function e^x, find by the method of undetermined coefficients (or otherwise) the first few (or all) terms of the power-series expansion of the function $\ln(1+x)$.

Problem 8 Compute $\exp A$ when A is one of the matrices

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Problem 9 How many terms of the series for e^x must one take in order to obtain a polynomial that makes it possible to compute e^x on the interval [–3, 5] within 10^{-2}?

Problem 10 If we know the power expansion of the functions $\sin x$ and $\cos x$, find with the method of the undetermined coefficients (or another one) some few first terms (or all) the power expansion of the function $\tan x$ in a neighborhood of the point $x = 0$.

Problem 11 The length of a band tightening the Earth at the Equator have increased by 1 meter, and after that the band was pulled propping up a vertical column. What is, roughly speaking, the height of the column if the radius of the Earth $\approx$6400 km?

Problem 12 Calculate

$$\lim_{x\to\infty} \left(\mathrm{e}\left(1+\frac{1}{x}\right)^{-x}\right)^x.$$

Problem 13 Sketch the graphs of the following functions:

$$\text{a) } \log_{\cos x} \sin x; \qquad \text{b) } \arctan \frac{x^3}{(1-x)(1+x)^2}.$$

2 One-Variable Differential Calculus

Problem 1 Show that if the acceleration vector $\mathbf{a}(t)$ is orthogonal to the vector $\mathbf{v}(t)$ at each instant of time t, the magnitude $|\mathbf{v}(t)|$ remains constant.

Problem 2 Let (x, t) and $(\tilde{x}, \tilde{t})$ be respectively the coordinate of a moving point and the time in two systems of measurement. Assuming the formulas $\tilde{x} = \alpha x + \beta t$

and $\tilde{t} = \gamma x + \delta t$ for transition from one system to the other are known, find the formula for the transformation of velocities, that is, the connection between $v = \frac{\mathrm{d}x}{\mathrm{d}t}$ and $\tilde{v} = \frac{\mathrm{d}\tilde{x}}{\mathrm{d}\tilde{t}}$.

Problem 3 The function $f(x) = x^2 \sin\frac{1}{x}$ for $x \neq 0$, $f(0) = 0$ is differentiable on $\mathbb{R}$, but f' is discontinuous at $x = 0$ (verify this). We shall "prove", however, that if $f : \mathbb{R} \to \mathbb{R}$ is differentiable on $\mathbb{R}$, then f' is continuous at every point $a \in \mathbb{R}$. By Lagrange's theorem

$$\frac{f(x) - f(a)}{x - a} = f'(\xi),$$

where ξ is a point between a and x. Then if $x \to a$, it follows that $\xi \to a$. By definition,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a),$$

and since this limit exists, the right-hand side of Lagrange's formula has a limit equal to it. That is, $f'(\xi) \to f'(a)$ as $\xi \to a$. The continuity of f' at a is now "proved". Where is the error?

Problem 4 Suppose the function f has $n+1$ derivatives at the point x_0, and let $\xi = x_0 + \theta_x(x - x_0)$ be the intermediate point in Lagrange's formula for the remainder term $\frac{1}{n!} f^{(n)}(\xi)(x - x_0)^n$, so that $0 < \theta_x < 1$. Show that $\theta_x \to \frac{1}{n+1}$ as $x \to x_0$ if $f^{(n+1)}(x_0) \neq 0$.

Problem 5

a) If the function $f \in C^{(n)}([a, b], \mathbb{R})$ vanishes at $n + 1$ points of the interval $[a, b]$, then there exists in this interval at least one zero of the function $f^{(n)}$, the derivative of f of order n.

b) Show that the polynomial $P_n(x) = \frac{\mathrm{d}^n (x^2-1)^n}{\mathrm{d}x^n}$ has n roots on the interval $[-1, 1]$. (Hint: $x^2 - 1 = (x - 1)(x + 1)$ and $P_n^{(k)}(-1) = P_n^{(k)}(1) = 0$, for $k = 0, \dots, n - 1$.)

Problem 6 Recall the geometric meaning of the derivative and show that if the function f is defined and differentiable on an interval I and $[a, b] \subset I$, then the function f' (not even necessarily continuous!) takes all the values between $f'(a)$ and $f'(b)$ on the interval $[a, b]$.

Problem 7 Prove the inequality

$$a_1^{\alpha_1} \cdots a_n^{\alpha_n} \leq \alpha_1 a_1 + \cdots + \alpha_n a_n,$$

where $a_1, \dots, a_n, \alpha_1, \dots, \alpha_n$ are nonnegative and $\alpha_1 + \cdots + \alpha_n = 1$.

Problem 8 Show that

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n = \mathrm{e}^x(\cos y + i\sin y) \quad (z = x+iy),$$

so that it is natural to suppose that $\mathrm{e}^{iy} = \cos y + i\sin y$ (Euler's formula) and

$$\mathrm{e}^z = \mathrm{e}^x\mathrm{e}^{iy} = \mathrm{e}^x(\cos y + i\sin y).$$

Problem 9 Find the shape of the surface of a liquid rotating at uniform angular velocity in a glass.

Problem 10 Show that the tangent to the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ at the point (x_0, y_0) has the equation $\frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1$, and that light rays from a source situated at one of the foci $F_1 = (-\sqrt{a^2-b^2}, 0)$, $F_2 = (\sqrt{a^2-b^2}, 0)$ of an ellipse with semiaxes $a > b > 0$ are reflected by an elliptical mirror to the other focus.

Problem 11 A particle subject to gravity, without any initial boost, begins to slide from the top of an iceberg of elliptic cross-section. The equation of the cross section is $x^2+5y^2=1$, $y \ge 0$. Compute the trajectory of the motion of the particle until it reaches the ground.

Problem 12 The value

$$s_\alpha(x_1, x_2, \dots, x_n) = \left(\frac{x_1^\alpha + x_2^\alpha + \cdots + x_n^\alpha}{n}\right)^{1/\alpha}$$

is called the mean of order α of the numbers $x_1, x_2, \dots, x_n$. In particular, for $\alpha = 1, 2, -1$, we obtain the *arithmetic mean, the mean square, and the harmonic mean*, respectively, of these numbers. We will assume that the numbers $x_1, x_2, \dots, x_n$ are nonnegative and if the exponent α is less than 0, then we will suppose even that they are positive.

a) Show, using Hölder's inequality, that if $\alpha < \beta$, then

$$s_\alpha(x_1, x_2, \dots, x_n) \le s_\beta(x_1, x_2, \dots, x_n),$$

and equality holds only when $x_1 = x_2 = \cdots = x_n$.

b) Show that if α tends to zero, then the value $s_\alpha(x_1, x_2, \dots, x_n)$ tends to $\sqrt[n]{x_1x_2\cdots x_n}$, i.e., to the *harmonic mean* of these numbers. In view of the result of problem a), from here, for example, follows the classical inequality between the geometric and the arithmetic means of nonnegative numbers (write it down).

c) If $\alpha \to \infty$, then $s_\alpha(x_1, x_2, \dots, x_n) \to \max\{x_1, x_2, \dots, x_n\}$, and for $\alpha \to -\infty$, the value $s_\alpha(x_1, x_2, \dots, x_n)$ tends to the lowest of the considered numbers, i.e., to $\min\{x_1, x_2, \dots, x_n\}$. Prove this.

Problem 13 Let $\mathbf{r} = \mathbf{r}(t)$ denote the law of motion of a point (i.e., its radius vectors as a function of time). We suppose that it is a continuously differentiable function on the interval $a \le t \le b$.

a) Is it possible, according to Lagrange's mean value theorem, to claim that there exists a moment ξ on the interval $[a, b]$ such that $\mathbf{r}(b) - \mathbf{r}(a) = \mathbf{r}'(\xi) \cdot (b - a)$? Explain your answer with examples.

b) Let Convex$\{\mathbf{r}'\}$ be the convex hull (of all ends) of the vectors $\mathbf{r}'(t)$, $t \in [a, b]$. Show that there exists a vector $\mathbf{v} \in$ Convex$\{\mathbf{r}'\}$ such that $\mathbf{r}(b) - \mathbf{r}(a) = \mathbf{v} \cdot (b - a)$.

c) The relation $|\mathbf{r}(b) - \mathbf{r}(a)| \le \sup |\mathbf{r}'(t)| \cdot |b - a|$, where the upper bound is taken over $t \in [a, b]$ has an obvious physical sense. What is it? Prove this inequality as a general mathematical fact, developing the classical Lagrange theorem on finite increments.

3 Integration and Introduction to Several Variables

Problem 1 Knowing the inequalities of Hölder, Minkowski, and Jensen for sums, obtaining the corresponding inequalities for integrals.

Problem 2 Compute the integral $\int_0^1 \mathrm{e}^{-x^2}\,\mathrm{d}x$ with a relative error of less than 10 %.

Problem 3 The function $\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} \mathrm{e}^{-t^2}\,\mathrm{d}t$, called the *probability error integral*, has limit 1 as $x \to +\infty$. Draw the graph of this function and find its derivative. Show that as $x \to +\infty$

$$\operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \mathrm{e}^{-x^2} \left(\frac{1}{2x} - \frac{1}{2^2 x^3} + \frac{1 \cdot 3}{2^3 x^5} - \frac{1 \cdot 3 \cdot 5}{2^4 x^7} + o\left(\frac{1}{x^7} \right) \right).$$

How can this asymptotic formula be extended to a series? Are there any values of $x \in \mathbb{R}$ for which this series converges?

Problem 4 Does the length of a path depend on the law of motion (the parametrization)?

Problem 5 You are holding one end of a rubber band of length 1 km. A beetle is crawling toward you from the other end, which is clamped, at a rate of 1 cm/s. Each time it crawls 1 cm you lengthen the band by 1 km. Will the beetle ever reach your hand? If so, approximately how much time will it require? (A problem of L.B. Okun', proposed to A.D. Sakharov.)

Problem 6 Calculate the work done in moving a mass in the gravitational field of the Earth and show that this work depends only on the elevation of the initial and terminal positions. Find the work done in escaping from the Earth's gravitational field and the corresponding escape velocity.

Problem 7 Using the example of a pendulum and a double pendulum explain how it is possible to introduce local coordinate systems and neighborhoods into the set of corresponding configurations and how a natural topology thereby arises making it into the configuration space of a mechanical system. Is this space metrizable under these conditions?

Problem 8 Is the unit sphere in $\mathbb{R}^n$, $\mathbb{R}_0^\infty$ or $C[a, b]$ compact?

Problem 9 A subset of a given set is called an ε-*grid* if any point of the set lies at a distance less than ε from some point of the set. Denote by $N(\varepsilon)$ the smallest possible number of points in an ε-grid for a given set. Estimate the ε-*entropy* $\log_2 N(\varepsilon)$ of a closed line segment, a square, a cube, and a bounded region in $\mathbb{R}^n$. Does the quantity $\frac{\log_2 N(\varepsilon)}{\log_2 (1/\varepsilon)}$ as $\varepsilon \to 0$ give a picture of the dimension of the space under consideration? Can such a dimension be equal, for example, to 0.5?

Problem 10 On the surface of the unit sphere S in $\mathbb{R}^3$ the temperature T varies continuously as a function of a point. Must there be points on the sphere where the temperature reaches a minimum or a maximum? If there are points where the temperature assumes two given values, must there be points where it assumes intermediate values? How much of this is valid when the unit sphere is taken in the space $C[a, b]$ and the temperature at the point $f \in S$ is given as

$$T(f) = \left(\int_a^b |f|(x)\, \mathrm{d}x \right)^{-1}?$$

Problem 11

a) Taking 1.5 as an initial approximation to $\sqrt{2}$, carry out two iterations using Newton's method and observe how many decimal places of accuracy you obtain at each step.

b) By a recursive procedure find a function f satisfying the equation

$$f(x) = x + \int_0^x f(t)\, \mathrm{d}t.$$

4 Differential Calculus of Several Variables

Problem 1 Local linearization. Consider and prove that local linearization is applicable to the following examples: instantaneous velocity and displacement; simplification of the equation of motion when the oscillations of a pendulum are small; computation of linear corrections to the values of $\exp(A)$, A^{-1}, $\det(E)$, $\langle a, b \rangle$ under small changes in the arguments (here A is an invertible matrix, E is the identity matrix, a and b are vectors, and $\langle \cdot, \cdot \rangle$ is the inner product).

Problem 2

a) What is the relative error $\delta = \frac{|\Delta f|}{|f|}$ in computing the value of a function $f(x, y, z)$ at a point (x, y, z) whose coordinates have absolute errors Δx, Δy, and Δz respectively?

b) What is the relative error in computing the volume of a room whose dimensions are as follows: length $x = 5 \pm 0.05$ m, width $y = 4 \pm 0.04$ m, height $z = 3 \pm 0.03$ m?

c) Is it true that the relative error of the value of a linear function coincides with the relative error of the value of its argument?

d) Is it true that the differential of a linear function coincides with the function itself?

e) Is it true that the relation $f' = f$ holds for a linear function f?

Problem 3

a) One of the partial derivatives of a function of two variables defined in a disk equals zero at every point. Does that mean that the function is independent of the corresponding variable in that disk?

b) Does the answer change if the disk is replaced by an arbitrary convex region?

c) Does the answer change if the disk is replaced by an arbitrary region?

d) Let $\mathbf{x} = \mathbf{x}(t)$ be the law of motion of a point in the plane (or in $\mathbb{R}^n$) in the time interval $t \in [a, b]$. Let $\mathbf{v}(t)$ be its velocity as a function of time and $C = \text{conv}\{\mathbf{v}(t) \mid t \in [a, b]\}$ the smallest convex set containing all the vectors $\mathbf{v}(t)$ (usually called the *convex hull* of a set that spans it). Show that there is a vector $\mathbf{v}$ in C such that $\mathbf{x}(b) - \mathbf{x}(a) = \mathbf{v} \cdot (b - a)$.

Problem 4

a) Let $F(x, y, z) = 0$. Is it true that $\frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x} \cdot \frac{\partial x}{\partial z} = -1$? Verify this for the relation $\frac{xy}{z} - 1 = 0$ (corresponding to the Clapeyron equation of state of an ideal gas: $\frac{PV}{T} = R$).

b) Now let $F(x, y) = 0$. Is it true that $\frac{\partial y}{\partial x} \frac{\partial x}{\partial y} = 1$?

c) What can you say in general about the relation $F(x_1, \ldots, x_n) = 0$?

d) How can you find the first few terms of the Taylor expansion of the implicit function $y = f(x)$ defined by an equation $F(x, y) = 0$ in a neighborhood of a point (x_0, y_0), knowing the first few terms of the Taylor expansion of the function $F(x, y)$ in a neighborhood of (x_0, y_0), where $F(x_0, y_0) = 0$ and $F'_y(x_0, y_0)$ is invertible?

Problem 5

a) Verify that the plane tangent to the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ at the point (x_0, y_0, z_0) can be defined by the equation $\frac{xx_0}{a^2} + \frac{yy_0}{b^2} + \frac{zz_0}{c^2} = 1$.

b) The point $P(t) = (\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}, \frac{c}{\sqrt{3}}) \cdot t$ emerged from the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ at time $t = 1$. Let $p(t)$ be the point of the same ellipsoid closest to $P(t)$ at time t. Find the limiting position of $p(t)$ as $t \to +\infty$.

Problem 6

a) In the plane $\mathbb{R}^2$ with Cartesian coordinates (x, y) construct the level curves of the function $f(x, y) = xy$ and the curve $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Using the resulting picture, carry out a complete study of the extremal problem for $f|_S$, the restriction of f to the circle S.

b) What is the physical meaning of the Lagrange multiplier in Lagrange's method of finding extrema with constraints when an equilibrium position is sought for a point mass in a gravitational field if the motion of the point is constrained by ideal relations (for example, relations of the form $F_1(x, y, z) = 0$, $F_2(x, y, z) = 0$)?

Problem 7 If in a vector space V one has a nondegenerate bilinear form $B(x, y)$, then to every linear function $g^* \in V^*$ in this space there corresponds a unique vector g such that $g^*(v) = B(g, v)$, for every vector $v \in V$.

a) Show that if $V = \mathbb{R}^n$, $B(x, y) = b_{ij}x^i x^j$, and $g^*v = g_i v^i$, then the vector g has coordinates $g^j = b^{ij} g_i$, where b^{ij} is the inverse of the matrix (b_{ij}).

A symmetric scalar product $\langle\cdot,\cdot\rangle$ in the Euclidean geometry or a skew-scalar product $\omega(\cdot,\cdot)$ (when the form B is skew-symmetric) in the symplectic geometry appears most often as a bilinear form $B(\cdot,\cdot)$.

b) Let $B(v_1, v_2) = \begin{vmatrix} v_1^1 & v_1^2 \\ v_2^1 & v_2^2 \end{vmatrix}$ be the oriented area of the parallelogram spanned by the vectors $v_1, v_2 \in \mathbb{R}^2$. Find the vector $g = (g^1, g^1)$ corresponding to the linear function $g^* = (g_1, g_2)$ and with respect to the form B if we know the coefficients of g^*.

c) The vector corresponding to the differential of a function $f : \mathbb{R}^n \to \mathbb{R}$ at the point x relative to the scalar product $\langle\cdot,\cdot\rangle$ of the Euclidean space $\mathbb{R}^n$ is called as usual the gradient of the function f at this point and denoted by $\operatorname{grad} f$. Thus, $\mathrm{d}f(x)v := \langle \operatorname{grad} f, v\rangle$ for every vector $v \in T_x\mathbb{R}^n \sim \mathbb{R}^n$, applied at x.

Therefore,

$$f'(x)v = \frac{\partial f}{\partial x^1}(x)v^1 + \cdots + \frac{\partial f}{\partial x^n}(x)v^n = \bigl\langle \operatorname{grad} f(x), v\bigr\rangle = \bigl|\operatorname{grad} f(x)\bigr| \cdot |v| \cos\varphi.$$

c_1) Show that in the standard orthonormal basis, i.e., in Cartesian coordinates, $\operatorname{grad} f(x) = (\frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n})(x)$.

c_2) Show that the rate of growth of the function f under a motion from the point x with unit velocity is maximal when the direction of the movement coincides with the direction of the gradient of the function f at this point and is equal to $|\operatorname{grad} f(x)|$. When the movement has a perpendicular direction to the vector $\operatorname{grad} f(x)$, the function does not change.

c_3) How do the coordinates of the vector $\operatorname{grad} f(x)$ in $\mathbb{R}^2$ change if instead of considering the canonical basis (e_1, e_2), we take an orthogonal basis $(\tilde{e}_1, \tilde{e}_2) = (\lambda_1 e_1, \lambda e_2)$?

c_4) How do we calculate $\operatorname{grad} f$ in polar coordinates? Answer: $(\frac{\partial f}{\partial r}, \frac{1}{r}\frac{\partial f}{\partial \varphi})$.

d) In exercise b) above, we considered a skew-symmetric form $B(v_1, v_2)$ of the oriented area of the parallelogram in $\mathbb{R}^2$.

Since the vector corresponding to $df(x)$ relative to the symmetric form $\langle\cdot,\cdot\rangle$ is called the gradient grad $f(x)$, the vector corresponding to $df(x)$ relative to the skew-symmetric form B is called *skew-gradient* and denoted by sgrad $f(x)$. Write down grad $f(x)$ and sgrad $f(x)$ in Cartesian coordinates.

Problem 8

a) Show that in $\mathbb{R}^3$ (and in general in $\mathbb{R}^{2n+1}$), there are no nondegenerate skew-symmetric bilinear forms.

b) In the oriented space $\mathbb{R}^2$ there is a nondegenerate skew-symmetric bilinear form (the oriented area of the parallelogram), as we have seen. In $\mathbb{R}^{2n}$ with coordinates $(x^1, \dots, x^n, \dots, x^{2n}) = (p^1, \dots, p^n, q^1, \dots, q^n)$, such a form ω also exists: if $v_i = (p_i^1, \dots, p_i^n, q_i^1, \dots, q_i^n)$ $(i = 1, 2)$, then

$$\omega(v_1, v_2) = \begin{vmatrix} p_1^1 & q_1^1 \\ p_2^1 & q_2^1 \end{vmatrix} + \cdots + \begin{vmatrix} p_1^n & q_1^n \\ p_2^n & q_2^n \end{vmatrix}.$$

That means that $\omega(v_1, v_2)$ is the sum of the oriented areas of parallelograms spanned by the projections of the vectors v_1, v_2 in the coordinate plane (p^j, q^j) $(j = 1, \dots, n)$.

b$_1$) Let g^* be a linear function in $\mathbb{R}^{2n}$, given with its coefficients $g^* = (p_1, \dots, p_n, q_1, \dots, q_n)$. Find the coordinates for the vector g mapped by the function g^* through the form ω.

b$_2$) The differential of the function $f : \mathbb{R}^{2n} \to \mathbb{R}$ at the point $x \in \mathbb{R}^{2n}$ through the skew-symmetric form ω is associated with a vector called the skew-gradient of the function f, as we already said, at this point and denoted by sgrad $f(x)$. Find the expression for sgrad $f(x)$ in the canonical Cartesian coordinates in the space $\mathbb{R}^{2n}$.

b$_3$) Find the scalar product $\langle \text{grad}\, f(x), \text{sgrad}\, f(x)\rangle$.

b$_4$) Show that the vector sgrad $f(x)$ is directed along the level surface of the function f.

b$_5$) The law of motion $x = x(t)$ of a point x in the space $\mathbb{R}^{2n}$ is such that $\dot{x}(t) =$ sgrad $f(x(t))$. Show that $f(x(t)) = \text{const}$.

b$_6$) Write down the equation $\dot{x} =$ sgrad $f(x)$ in the canonical Cartesian notation $(p^1, \dots, p^n, q^1, \dots, q^n)$ for the coordinates and $H = H(p, q)$ for the function f. The resulting system, called a system of Hamilton's equations, is one of the central objects of mechanics.

Problem 9 Canonical variables and Hamilton's system of equations.

a) In the calculus of variations and in the fundamental variational principles of classical mechanics, the following *system of Euler–Lagrange equations* plays an important role:

$$\begin{cases} \left(\dfrac{\partial L}{\partial x} - \dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial L}{\partial v}\right)(t, x, v) = 0, \\ v = \dot{x}(t), \end{cases}$$

where $L(t, x, v)$ is a given function in the variables t, x, v, where t usually denotes time, x the coordinate, and v the velocity. This is a system of two equations in three variables. From it, one usually wants to find the dependence relations $x = x(t)$ and $v = v(t)$, which essentially boils down to finding the law of motion $x = x(t)$, since $v = \dot{x}(t)$. Write down the first equation of the system in detail, revealing the derivative $\frac{\mathrm{d}}{\mathrm{d}t}$ given that $x = x(t)$ and $v = v(t)$.

b) Show that in transition from the variables t, x, v, L to the *canonical* variables t, x, p, H by taking the Legendre transform of

$$\begin{cases} p = \dfrac{\partial L}{\partial v}, \\ H = pv - L, \end{cases}$$

with respect to the variables v, L, interchanging them with the variables p, H, the Euler–Lagrange system acquires the symmetric form:

$$\dot{p} = -\frac{\partial H}{\partial x}, \qquad \dot{x} = \frac{\partial H}{\partial p}.$$

c) In mechanics, we often use the notation q and $\dot{q}$, instead of x and v. In many cases when $L(t, q, \dot{q}) = L(t, q^1, \dots, q^m, \dot{q}^1, \dots, \dot{q}^m)$, the Euler–Lagrange system of equations has the form

$$\left(\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i}\right)(t, q, \dot{q}) = 0 \quad (i = 1, \dots, m).$$

Take the Legendre transform with respect to the variables $\dot{q}$ and L, and go from the variables t, q, $\dot{q}$, L to the canonical variables t, q, p, H and show that this Euler–Lagrange system transforms into the following Hamilton equations:

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}, \quad \dot{q}_i = \frac{\partial H}{\partial p^i} \quad (i = 1, \dots, m).$$

Examination Topics

1 First Semester

1.1 Introduction to Analysis and One-Variable Differential Calculus

1. Real numbers. Bounded (from above or below) numerical sets. The axiom of completeness and the existence of a least upper (greatest lower) bound of a set. Unboundedness of the set of natural numbers.
2. Fundamental lemmas connected with the completeness of the set of real numbers $\mathbb{R}$ (nested interval lemma, finite covering, limit point).
3. Limit of a sequence and the Cauchy criterion for its existence. Tests for the existence of a limit of a monotonic sequence.
4. Infinite series and the sum of an infinite series. Geometric progressions. The Cauchy criterion and a necessary condition for the convergence of a series. The harmonic series. Absolute convergence.
5. A test for convergence of a series of nonnegative terms. The comparison theorem. The series $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$.
6. The limit of a function. The most important filter bases. Definition of the limit of a function over an arbitrary base and its decoding in specific cases. Infinitesimal functions and their properties. Comparison of the ultimate behavior of functions, asymptotic formulas, and the basic operations with the symbols $o(\cdot)$ and $O(\cdot)$.
7. The connection of passage to the limit with the algebraic operations and the order relation in $\mathbb{R}$. The limit of $\frac{\sin x}{x}$ as $x \to 0$.
8. The limit of a composite function and a monotonic function. The limit of $(1 + \frac{1}{x})^x$ as $x \to \infty$.
9. The Cauchy criterion for the existence of the limit of a function.
10. Continuity of a function at a point. Local properties of continuous functions (local boundedness, conservation of sign, arithmetic operations, continuity of a composite function). Continuity of polynomials, rational functions, and trigonometric functions.

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

11. Global properties of continuous functions (intermediate-value theorem, maxima, uniform continuity).
12. Discontinuities of monotonic functions. The inverse function theorem. Continuity of the inverse trigonometric functions.
13. The law of motion, displacement over a small interval of time, the instantaneous velocity vector, trajectories and their tangents. Definition of differentiability of a function at a point. The differential, its domain of definition and range of values. Uniqueness of the differential. The derivative of a real-valued function of a real variable and its geometric meaning. Differentiability of $\sin x$, $\cos x$, e^x, $\ln|x|$, and x^α.
14. Differentiability and the arithmetic operations. Differentiation of polynomials, rational functions, the tangent, and the cotangent.
15. The differential of a composite function and an inverse function. Derivatives of the inverse trigonometric functions.
16. Local extrema of a function. A necessary condition for an interior extremum of a differentiable function (Fermat's lemma).
17. Rolle's theorem. The finite-increment theorems of Lagrange and Cauchy (mean-value theorems).
18. Taylor's formula with the Cauchy and Lagrange forms of the remainder.
19. Taylor series. The Taylor expansions of e^x, $\cos x$, $\sin x$, $\ln(1+x)$, and $(1+x)^\alpha$ (Newton's binomial formula).
20. The local Taylor formula (Peano form of the remainder).
21. The connection between the type of monotonicity of a differentiable function and the sign of its derivative. Sufficient conditions for the presence or absence of a local extremum in terms of the first, second, and higher-order derivatives.
22. L'Hôpital rule.
23. Convex functions. Differential conditions for convexity. Location of the graph of a convex function relative to its tangent.
24. The general Jensen inequality for a convex function. Convexity (or concavity) of the logarithm. The classical inequalities of Cauchy, Young, Hölder, and Minkowski.
25. Legendre transform.
26. Complex numbers in algebraic and trigonometric notation. Convergence of a sequence of complex numbers and a series with complex terms. The Cauchy criterion. Absolute convergence and sufficient conditions for absolute convergence of a series with complex terms. The limit $\lim_{n\to\infty}(1+\frac{z}{n})^n$.
27. The disk of convergence and the radius of convergence of a power series. The definition of the functions e^z, $\cos z$, $\sin z$ $(z \in \mathbb{C})$. Euler's formula and the connections among the elementary functions.
28. Differential equations as a mathematical model of reality, examples. The method of undetermined coefficients and Euler's (polygonal) method.
29. Primitives and the basic methods of finding them (termwise integration of sums, integration by parts, change of variable). Primitives of the elementary functions.

2 Second Semester

2.1 Integration. Multivariable Differential Calculus

1. The Riemann integral on a closed interval. Upper and lower sums, their geometric meaning, the behavior under a refinement of the partition, and mutual estimates. Darboux's theorem, upper and lower Darboux's integrals and the criterion for Riemann integrability of real-valued functions on an interval (in terms of sums of oscillations). Examples of classes of integrable functions.
2. The Lebesgue criterion for Riemann integrability of a function (statement only). Sets of measure zero, their general properties, examples. The space of integrable functions and admissible operations on integrable functions.
3. Linearity, additivity and general evaluation of an integral.
4. Evaluating the integral of a real-valued function. The (first) mean-value theorem.
5. Integrals with a variable upper limit of integration, their properties. Existence of a primitive for a continuous function. The generalized primitive and its general form.
6. The Newton–Leibniz formula. Change of variable in an integral.
7. Integration by parts in a definite integral. Taylor's formula with integral remainder. The second mean-value theorem.
8. Additive (oriented) interval functions and integration. The general pattern in which integrals arise in applications, examples: length of a path (and its independence of parametrization), area of a curvilinear trapezoid (area under a curve), volume of a solid of revolution, work, energy.
9. The Riemann–Stieltjes integral. Conditions under which it can be reduced to the Riemann integral. Singularities and the Dirac delta-function. The concept of a generalized function.
10. The concept of an improper integral. Canonical integrals. The Cauchy criterion and the comparison theorem for studying the convergence of an improper integral. The integral test for convergence of a series.
11. Metric spaces, examples. Open and closed subsets. Neighborhoods of a point. The induced metric, subspaces. Topological spaces. Neighborhoods of a point, separation properties (the Hausdorff axiom). The induced topology on subsets. Closure of a set and description of relatively closed subsets.
12. Compact sets, their topological invariance. Closedness of a compact set and compactness of a closed subset of a compact set. Nested compact sets. Compact metric spaces, ε-grids. Criteria for a metric space to be compact and its specific form in $\mathbb{R}^n$.
13. Complete metric spaces. Completeness of $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$, $\mathbb{C}^n$, and the space $C[a, b]$ of continuous functions under uniform convergence.
14. Criteria for continuity of a mapping between topological spaces. Preservation of compactness and connectedness under a continuous mapping. The classical theorems on boundedness, the maximum-value theorem, and the intermediate-value theorem for continuous functions. Uniform continuity on a compact metric space.

15. The norm (length, absolute value, modulus) of a vector in a vector space; the most important examples. The space $L(X, Y)$ of continuous linear transformations and the norm in it. Continuity of a linear transformation and finiteness of its norm.
16. Differentiability of a function at a point. The differential, its domain of definition and range of values. Coordinate expression of the differential of a mapping $f : \mathbb{R}^m \to \mathbb{R}^n$. The relation between differentiability, continuity, and the existence of partial derivatives.
17. Differentiation of a composite function and the inverse function. Coordinate expression of the resulting laws in application to different cases of the mapping $f : \mathbb{R}^m \to \mathbb{R}^n$.
18. Derivative along a vector and the gradient. Geometric and physical examples of the use of the gradient (level surfaces of functions, steepest descent, the tangent plane, the potential of a field, Euler's equation for the dynamics of an ideal fluid, Bernoulli's law, the work of a wing).
19. Homogeneous functions and the Euler relation. The dimension method.
20. The finite-increment theorem. Its geometric and physical meaning. Examples of applications (a sufficient condition for differentiability in terms of the partial derivatives; conditions for a function to be constant in a domain).
21. Higher-order derivatives and their symmetry.
22. Taylor's formula.
23. Extrema of functions (necessary and sufficient conditions for an interior extremum).
24. Contraction mappings. The Picard–Banach fixed-point principle.
25. The implicit function theorem.
26. The inverse function theorem. Curvilinear coordinates and rectification. Smooth k-dimensional surfaces in $\mathbb{R}^n$ and their tangent planes. Methods of defining a surface and the corresponding equations of the tangent space.
27. The rank theorem and functional dependence.
28. Local resolution of a diffeomorphism as the composition of elementary ones (diffeomorphisms changing only one coordinate).
29. Extrema with constraint (necessary condition). Geometric, algebraic, and physical interpretation of the method of Lagrange multipliers.
30. A sufficient condition for a constrained extremum.

Appendix A
Mathematical Analysis (Introductory Lecture)

A.1 Two Words About Mathematics

Mathematics is an abstract science. For example, it teaches how to count and add no matter of whether we count ravens, capital, or something else. Therefore, mathematics is one of the most universal and commonly used applied sciences. Mathematics as a science possesses a large number of features, and therefore it is usually treated with respect. For example, it teaches us to listen to arguments and appreciate the truth.

Lomonosov believed that mathematics leads the mind into order, and Galilei said, "The great book of nature is written in the language of mathematics." The evidence of this is obvious: anyone who wants to read this book, must have studied mathematics. Among them, there are representatives of natural sciences, technical professions, as well as humanities. As examples, there is the chair of mathematics at the Economics Faculty of Moscow State University, and also there is an Institute of Economical Mathematics in the system of the Russian Academy of Sciences. There is even a statement saying that a branch of science contains as much of science as there is mathematics inside of it. Although this point of view is too strong, it is in general a quite sharp observation.

Mathematics has the attributes of a language. However, it is clearly not merely a language (otherwise, it would have been studied by philologists). Mathematics not only can translate a question into mathematical language, but it usually provides the method for solving the formulated mathematical problem.

The capability to pose a question correctly is the great art of researchers in general and mathematicians in particular.

The great mathematician Henri Poincaré, whose influence is clearly seen in most courses in contemporary mathematics, remarked with humor, "Mathematics is the art of calling different things by the same name." For instance, a point is a barely visible particle under a microscope, an airplane on air traffic controls radar, a city on a map, a planet in the sky, and in general all those things, whose dimensions can be neglected on these scales.

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

Therefore, the abstract concepts of mathematics and their relationships, like number, are very useful in a tremendous sphere of specific phenomena and consistent patterns.

A.2 Number, Function, Law

The usual reaction to a miracle, "Is it really possible???" quickly and imperceptibly transforms into, "It could not be otherwise!!!"

We are so used to the fact that $2 + 3 = 5$, that we don't see any miracle there. But here it is not said that two apples and three apples will add up to five apples, it is said that this is the case for apples, elephants, and all other things. We already mentioned this above.

Next, we get used to the fact that $a + b = b + a$, where now the symbols a and b could mean 2, 3, or any other integer.

A function, or functional dependence, is the next mathematical miracle. It is relatively young as a scientific concept: just a little over three hundred years old. Nevertheless, we are confronted with it in nature and even in everyday life no less than with elephants or apples.

Every science or every field of human activity deals with a concrete area of objects and their relations. These relations, laws, or dependences, are described and studied by mathematics in an abstract and general way, relating the terms *function* or *functional dependence* $y = f(x)$ of the state (value) of one variable (y) to some other state (or value) of another variable (x).

It is especially important that now we are not dealing with constants, but with variables x and y and with the rule f relating them. A function is adapted to describe developing processes and phenomena, the nature of change of their states, and in general to describe the dependent variable.

Sometimes the rule f of a relation is known (given), e.g., by a government or by a technological process. We often try then, under the constraints of the acting rule f, to choose a strategy, i.e., some state (value) available to our choice of the independent variable x, in order to obtain the most favorable state (value) for us (in one way or another) of the variable y (given that $y = f(x)$).

In other cases (and this is even more exciting), we search for the law of nature relating certain phenomena. Though it is the job of experts in the corresponding specific branch of science, mathematics might be extremely helpful. Like Sherlock Holmes, it can deduce the new law f from very limited information accessible to the experts in this narrow domain. Following the parallel with Sherlock Holmes, mathematics uses a "deduction method" called *differential equations*, which were unknown to ancient mathematicians, and that emerged with the advent of differential and integral calculus at the seventeenth and eighteenth centuries thanks to the efforts of Newton, Leibniz, and their predecessors and successors.

So, let us begin a primer of modern mathematics.

A.3 Mathematical Model of a Phenomenon (Differential Equations, or We Learn How to Write)

One of the brightest and most long-lived impressions of school mathematics, of course, is that small miracle when you want to find something unknown to you, denoting it by the letter x or by the letters x, y, and then you write something like $a \cdot x = b$ or some system of equations

$$\begin{cases} 2x + y = 1, \\ x - y = 2. \end{cases}$$

After a bit of mathematical hocus-pocus, you discover what was unknown to you: $x = 1$, $y = -1$.

Let us try to learn how to write the equation in a new situation, when we do not have to find a number but instead the unknown rule relating two variables that are important for us, i.e., we look for the requisite function. Let us take a look at some examples.

In order to be more specific, we first shall talk about biology (proliferation of microorganisms, growth of biomass, ecological limitations, etc.). However, it will be clear that all this can be transferred to other areas, as for example the growth of capital, nuclear reactions, atmospheric pressure, and so on.

To warm up, consider the following playful problem:

Consider a primitive organism that replicates itself every second (doubling). Suppose it is put into an empty glass. After one minute, the glass is filled up. How long will it take to fill up the glass if instead of one organism, two of these organisms are put into an empty glass?

Now we get closer to our goal, and the promised examples.

Example 1 It is known that under favorable conditions, the reproduction rate of microorganisms, i.e., the biomass growth rate, is proportional (with a coefficient of proportionality k) to the current amount of biomass. We need to find the rule $x = x(t)$ of the change of biomass over time if the initial condition $x(0) = x_0$ is known.

We expect that if we knew the rule itself $x = x(t)$ of the variable x, we would know its rate of change at any time t. Without going too far into the discussion about how we calculate this rate of change from $x(t)$, we shall denote it by $x'(t)$. Since the function $x' = x'(t)$ is obtained from the function $x(t)$, in mathematics it is called the *derivative* of the function $x = x(t)$ (in order to find out how to calculate the derivative of a function and much more, it is necessary to learn differential calculus. This is coming).

Now we are able to write what is given to us:

$$x'(t) = k \cdot x(t), \tag{A.1}$$

where $x(0) = x_0$, and we want to find the dependence itself.

We just wrote down the first differential equation (A.1). In general, equations containing derivatives are called differential equations (some clarifications and exceptions are not yet relevant). It is worth noting that the independent variable is often omitted, in order to simplify the text while writing the equation. For instance, Eq. (A.1) is written in the form $x' = k \cdot x$. If the desired function is denoted by the letter f or u, then the same equation would have the form $f' = k \cdot f$ or $u' = k \cdot u$ respectively.

It is now clear that if we learn not only how to write, but also to solve or study differential equations, we will be able to predict and know many things. That is why Newton's sacramental phrase referring to the new calculus, was something like this: "It is useful to learn how to solve differential equations."

Problem 1 Write down an equation for the example above of reproduction in a glass. Which coefficient k, initial condition $x(0) = x_0$, and dependence $x = x(t)$ do we have?

Having succeeded with the first equation, let us try to encode by a differential equation several further natural phenomena.

Example 2 Assume now, as is always the case, that there is not infinitely much food and that the environment cannot support more than M individuals, i.e., that the biomass does not exceed the value M. Then the growth rate of the biomass will presumably decrease, for instance proportionally to the remaining conditions of the environment. As a measure of the remaining conditions, you can consider the difference $M - x(t)$, or even better consider the dimensionless value $1 - \frac{x(t)}{M}$. In this situation, instead of Eq. (A.1), we obviously have the equation

$$x' = k \cdot x \cdot \left(1 - \frac{x(t)}{M}\right), \tag{A.2}$$

which reduces to Eq. (A.1) at the stage when $x(t)$ is much less than M. Conversely, when $x(t)$ is very close to M, the growth rate becomes close to zero, i.e., the growth stops, which is natural. After having improved some skills, we shall find out later precisely how the rule $x = x(t)$ looks in this case.

Problem 2 A body having an initial temperature T_0 cools down in an environment that has a constant temperature C. Let $T = T(t)$ be the rule of variation of temperature of the body over time. Write down the equation that this function must satisfy, assuming that the cooling rate is proportional to the temperature difference between the body and the environment.

The velocity $v(t)$, which is the rate of change of the variable $x(t)$, is called the derivative of the function $x(t)$ and denoted by $x'(t)$.

The acceleration $a(t)$, as you may know, is the rate of change of the velocity $v(t)$. This means that $a(t) = v'(t) = (x')'(t)$, i.e., it is the derivative of the derivative of the initial function. It is called the second derivative of the initial function and it is frequently denoted by $x''(t)$ (some other notation will appear later). If we are able to find the first derivative, we may repeat the same process to calculate the derivative $x^{(n)}(t)$ of any order n of the initial function $x = x(t)$.

Example 3 Let $x = x(t)$ be the rule of motion of a point with mass m, i.e., coordinates for the position of the point as function of time. For simplicity, we may assume that the motion is along a straight line (horizontal or vertical); thus there is only one coordinate.

The classical Newton's law $m \cdot a = F$, relating the force acting on a point of mass m with acceleration caused by this action, can be written as

$$m \cdot x''(t) = F(t), \tag{A.3}$$

or in abbreviated form, $m \cdot x'' = F$.

If the acting force $F(t)$ is known, then the relation $m \cdot x'' = F$ can be considered a differential equation (of second order) with respect to the function $x(t)$.

For example, if F is the force of gravity on Earth's surface, then $F = mg$, where g is the acceleration at free fall. In this case, our equation has the form $x''(t) = g$. As you might know, Galileo discovered that in free fall, $x(t) = \frac{1}{2}gt^2 + v_0 t + x_0$, where x_0 is the initial position and v_0 is the initial speed of the point.

In order to check that this function satisfies the equation, it is necessary to be able to differentiate a function, i.e., to find its derivative. In our case, we even need the second derivative.

Just below we shall give a small table of some functions and their derivatives. Its deduction shall be made later, in the systematic presentation of differential calculus. Now try yourself to do the following.

Problem 3 Write down the equation of free fall in the atmosphere. In this case, there will arise a resistance force. Consider it proportional to the first (or second) rate of change of the movement (the speed in free fall does not grow to infinity due to the presence of the force of resistance).

You should be convinced by now that it is worthwhile learning how to calculate derivatives.

A.4 Velocity, Derivative, Differentiation

First, let us consider a familiar situation in which we can trust our intuition (and we change the notation from $x(t)$ to $s(t)$).

Suppose that a point moves on the whole numerical real line, $s(t)$ is its coordinate at the moment t, and $v(t) = s'(t)$ is its velocity at the same moment t. After the time interval h, from its former position at the moment t, the point is located at the position $s(t+h)$. In our picture of velocity the quantity $s(t+h) - s(t)$, the distance traveled in the time interval h after the moment t, and its velocity at the moment t are related by the equation

$$s(t+h) - s(t) \approx v(t) \cdot h; \tag{A.4}$$

in other words, $v(t) \approx \frac{s(t+h)-s(t)}{h}$, and this approximation becomes closer to an equality as the interval h decreases its length.

Thus, we have to assume that

$$v(t) := \lim_{h \to 0} \frac{s(t+h) - s(t)}{h},$$

i.e., we define $v(t)$ as the limit of the quotient between the increment of the function and the increment of its argument as the latter approaches zero.

Now after this example, nothing impedes us from providing the usual definition for the value $f'(x)$ for the derivative f' of the function f at the point x:

$$f'(x) := \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}, \tag{A.5}$$

i.e., $f'(x)$ is the limit of the quotient of increments $\Delta f / \Delta x$ as Δx approaches zero, where $\Delta f = f(x+h) - f(x)$ is the increment of the function with respect to the increment of its argument $\Delta x = (x+h) - x$.

Equation (A.5) can be rewritten in a similar form to (A.4) in the same convenient and useful form

$$f(t+h) - f(t) = f'(t)h + o(h), \tag{A.6}$$

where $o(h)$ is some error (of approximation), small in relation to h, as h approaches zero. (This means that the quotient $o(h)/h$ approaches zero as h goes to zero.)

Now we shall make some concrete calculations.

1. Let f be a constant function, i.e., $f(x) \equiv c$. Then it is clear that $\Delta f = f(x+h) - f(x) \equiv 0$ and $f'(x) \equiv 0$. This is natural, since the velocity of change is equal to zero if there is no change.

2. If $f(x) = x$, then $f(x+h) - f(x) = h$, and therefore $f'(x) \equiv 1$. If $f(x) = kx$, then $f(x+h) - f(x) = kh$ and $f' \equiv k$.

3. By the way, we can make two very common but extremely useful remarks: if the function f has its derivative f', then the function cf, where c is some arbitrary constant, has as derivative cf', i.e., $(cf)' = cf'$; in the same way, $(f+g)' = f' + g'$, i.e, the derivative of a sum is equal to the sum of the derivatives, if these derivatives are defined.

4. Let $f(x) = x^2$. Then $f(x+h) - f(x) = (x+h)^2 - x^2 = 2xh + h^2 = 2xh + o(h)$; hence $f'(x) = 2x$.

5. Analogously, if $f(x) = x^3$, then

$$f(x+h) - f(x) = (x+h)^3 - x^3 = 3x^2h + 3xh^2 + h^3 = 3x^2h + o(h);$$

therefore, $f'(x) = 3x$.

Table A.1

$f(x)$	$f'(x)$	$f''(x)$	$\cdots$	$f^{(n)}(x)$
a^x	$a^x \ln a$	$a^x \ln^2 a$	$\cdots$	$a^x \ln^n a$
e^x	e^x	e^x	$\cdots$	e^x
$\sin x$	$\cos x$	$-\sin x$	$\cdots$	$\sin(x + n\pi/2)$
$\cos x$	$-\sin x$	$-\cos x$	$\cdots$	$\cos(x + n\pi/2)$
$(1+x)^\alpha$	$\alpha(1+x)^{\alpha-1}$	$\alpha(\alpha-1)(1+x)^{\alpha-2}$	$\cdots$	?
x^α	$\alpha x^{\alpha-1}$	$\alpha(\alpha-1)x^{\alpha-2}$	$\cdots$	?

6. Now it is clear that in general, if $f(x) = x^n$, then one has

$$f(x+h) - f(x) = (x+h)^n - x^n = nx^{n-1}h + o(h),$$

and therefore $f'(x) = nx^{n-1}$.
7. This implies that if we have a polynomial

$$P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$

then

$$P'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + \cdots + a_{n-1}.$$

We calculated the above derivatives following the definitions. To develop and master the techniques of differentiation you have to practice. Now as examples and for your illustration we introduce Table A.1 with functions and their derivatives. Later we shall prove the results.

In Table A.1 e is a number (e = 2.7 ...), that appears everywhere in analysis, just like the number π in geometry. The logarithm with base e is frequently denoted by ln, instead of $\log_e$, and you can find it in the first row of the table in the second and third columns. The logarithm with this base is called the *natural logarithm* and it appears in many formulas.

Problem 4 Assuming that the formula of the first derivative f' is correct, verify the expression for $f^{(n)}$ and complete the table at the places where the question marks are. After that, calculate the value $f^{(n)}(0)$ in every case.

Problem 5 Try to find the derivative of the function $f(x) = e^{kx}$ and the solution of Eq. (A.1). Explain at what point in time the initial condition x_0 (capital, biomass, or something else represented by this equation) will double.

A.5 Higher Derivatives, What for?

A wonderful and very useful expansion of the central equation (A.6), which can be written as

$$f(x+h) = f(x) + f'(x)h + o(h), \tag{A.7}$$

turns into the following formula (Taylor's formula):

$$f(x+h) = f(x) + \frac{1}{1!}f'(x)h + \frac{1}{2!}f''(x)h^2 + \cdots + \frac{1}{n!}f^{(n)}(x)h^n + o\big(h^n\big). \quad \text{(A.8)}$$

If we set $x = 0$ and then replace the letter h by the letter x, we obtain

$$f(x) = f(0) + \frac{1}{1!}f'(0)x + \frac{1}{2!}f''(0)x^2 + \cdots + \frac{1}{n!}f^{(n)}(0)x^n + o\big(x^n\big). \quad \text{(A.9)}$$

For instance, if $f(x) = (1+x)^\alpha$, following Newton we find that

$$(1+x)^\alpha = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + o\big(x^n\big). \quad \text{(A.10)}$$

Sometimes in the formula (A.9), it is possible to continue the sum to infinity, removing the reminder term, since it tends to zero as $n \to \infty$.

In particular, we have,

$$\mathrm{e}^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots, \quad \text{(A.11)}$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \cdots + (-1)^k\frac{1}{2k!}x^{2k} + \cdots, \quad \text{(A.12)}$$

$$\sin x = \frac{1}{1!}x - \frac{1}{3!}x^3 + \cdots + (-1)^k\frac{1}{(2k+1)!}x^{2k+1} + \cdots. \quad \text{(A.13)}$$

We have obtained a representation of relatively complex functions as a sum (an infinite sum, a *series*) of simple functions, which can be calculated with the usual arithmetic operations. The finite pieces of these sums are polynomials. They provide good approximations of the functions decomposed in such a series.

A.5.1 Again Toward Numbers

We have always tacitly assumed that we are dealing with functions defined on the set of real numbers. But the right-hand side of Eqs. (A.11), (A.12), and (A.13) make sense when x is replaced by a complex number $z = x + iy$. Then, we are able to say what the expressions e^z, $\cos z$, $\sin z$ mean.

Problem 6 Discover, after Euler, the following impressive formula: $\mathrm{e}^{i\varphi} = \cos\varphi + i\sin\varphi$, linking these elementary functions and the remarkably beautiful equality $\mathrm{e}^{i\pi} + 1 = 0$ resulting from this formula, which combines the basic constants of the mathematical sciences (arithmetic, algebra, analysis, geometry, and even logic).

A.5.2 And What to Do Next?

As we say in Russian, “with the fingers”, without details and justifications, you have been given some idea of differential calculus, which is the core of a first-semester course in mathematical analysis. Step by step, we have become acquainted with the concepts of numbers, functions, limits, derivatives, series, which had been only superficially studied so far.

Now you know why it is necessary to take time to dive into a detailed and careful consideration of all of these concepts and objects. The understanding of them is necessary for a professional mathematician. For the average user, this is not required. Most people drive a car without even opening the hood. But that is reasonable, because someone well versed in engines designed a machine that works reliably.

Appendix B
Numerical Methods for Solving Equations (An Introduction)

B.1 Roots of Equations and Fixed Points of Mappings

Notice that the equation $f(x)=0$ is obviously equivalent to the equation $\alpha(x)f(x) = 0$ if $\alpha(x)\neq 0$. The last equation, in turn, is equivalent to the relation $x = x - \alpha(x)f(x)$, where x can be interpreted as a fixed point of the mapping $\varphi(x) := x - \alpha(x)f(x)$.

Thus, finding roots of equations is equivalent to finding the fixed points of the corresponding mappings.

B.2 Contraction Mappings and Iterative Process

A mapping $\varphi : X \to X$ from the set $X \subset \mathbb{R}$ to itself is called a *contraction mapping* if there exists a number q, $0 \le q < 1$, such that for every pair of points x', x'', their images satisfy the inequality $|\varphi(x') - \varphi(x'')| \le q|x' - x''|$.

It is clear that this definition applies to arbitrary sets, without any changes, on which the distance $d(x', x'')$ between any two points is defined; in our case, $d(x', x'') = |x' - x''|$.

It is also obvious that a contraction mapping is continuous and cannot have more than one fixed point.

Let $\varphi : [a, b] \to [a, b]$ be a contraction mapping from the interval $[a, b]$ into itself. We shall show that the iteration process $x_{n+1} = \varphi(x_n)$, starting at any point x_0 from this interval, leads to the point $x = \lim_{n\to\infty} x_n$, the fixed point of the mapping φ.

We notice first that

$$|x_{n+1} - x_n| \le q|x_n - x_{n-1}| \le \cdots \le q^n|x_1 - x_0|.$$

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

Therefore, for all natural numbers m, n, with $m > n$, inserting intermediate points and using the triangle inequality, we get the estimate

$$\begin{aligned}|x_m - x_n| &\le |x_m - x_{m-1}| + \cdots + |x_{n+1} - x_n| \\ &\le \left(q^{m-1} + \cdots + q^n\right)|x_1 - x_0| < \frac{q^n}{1-q}|x_1 - x_0|,\end{aligned}$$

and from this it follows that the sequence $\{x_n\}$ is fundamental (i.e., is a Cauchy sequence).

Hence, by the Cauchy criterion it converges to a point x in the interval $[a, b]$. This point is a fixed point of the mapping $\varphi : [a, b] \to [a, b]$, since by taking the limit $n \to \infty$ in the relation $x_{n+1} = \varphi(x_n)$, we obtain the equality $x = \varphi(x)$.

(Here we used the obvious fact that a contraction mapping is continuous; actually, it is even uniformly continuous.)

By passing to the limit $m \to \infty$ in the relation $|x_m - x_n| < \frac{q^n}{1-q}|x_1 - x_0|$, one gets the estimate

$$|x - x_n| < \frac{q^n}{1-q}|x_1 - x_0|$$

for the deviation value between the approaching point x_n and the fixed point x of the mapping φ.

B.3 The Method of Tangents (Newton's Method)

We presented a theorem stating that a continuous real-valued function from an interval taking values with different signs at the extremes of the interval has at least one zero in this interval (a point x where $f(x) = 0$). In its proof, we showed the simplest, but universal, algorithm for finding this point (splitting the interval in half). The order of the speed of convergence in this case is 2^{-n}.

In the case of a convex differentiable function, the method proposed by Newton can be much more efficient in terms of speed of convergence.

We draw a tangent to the graph of the given function f at some point $(x_0, f(x_0))$, where $x_0 \in [a, b]$. Next, we find the point x_1 where the tangent intersects the x-axis. By repeating this process, we obtain a sequence $\{x_n\}$ of points that quickly converges to the point x such that $f(x) = 0$. (It is possible to prove that each successive iteration leads to the doubling of the number of correct digits of x found at the previous step.)

It is easy to prove analytically (check it!) that the method of tangents reduces to an iteration process

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

For instance, the solution of the equation $x^m - a = 0$, i.e., the computation of $\sqrt[m]{a}$, according to the formula above, reduces to an iteration process

$$x_{n+1} = \frac{1}{m}\left((m-1)x_n - \frac{a}{x_n^{m-1}}\right).$$

In particular, for calculating $\sqrt{2}$ with the method of tangents, we obtain

$$x_{n+1} = \frac{1}{2}\left(x_n - \frac{a}{x_n}\right).$$

As can be seen from the formulas above, Newton's method looks for fixed points of the mapping $\varphi(x) = x - \frac{f(x)}{f'(x)}$. It is a special case of the mapping $\varphi(x) = x - \alpha f(x)$, discussed in the first section, and we obtain it by setting $\alpha(x) = \frac{1}{f'(x)}$.

Note that in general, the mapping $\varphi(x) = x - \alpha f(x)$ and even the mapping $\varphi(x) = x - \frac{f(x)}{f'(x)}$ involved in the method of tangents do not have to be contraction mappings. Moreover, as can be shown by simple examples, in the case of a general function f, the method of tangents does not always lead to a convergent iteration process.

If in the expression $\varphi(x) = x - \alpha f(x)$, the function $\alpha(x)$ can be chosen in the given interval so that $|\varphi'(x)| \le q < 1$, then the mapping $\varphi : [a,b] \to [a,b]$, of course, will be a contraction.

In particular, if α can be taken as the constant $\frac{1}{f'(x_0)}$, then we obtain $\varphi(x) = x - \frac{f(x)}{f'(x_0)}$ and $\varphi'(x) = 1 - \frac{f'(x)}{f'(x_0)}$. If the derivative of the function f is at least continuous at x_0, then in some neighborhood of x_0, we have $|\varphi'(x)| = |1 - \frac{f'(x)}{f'(x_0)}| \le q < 1$. If the function φ maps this neighborhood into itself (although this is not always the case), then the standard iterative process induced by the contraction mapping φ of this neighborhood will lead to a unique point of the mapping φ in this neighborhood at which the original function f vanishes.

Appendix C
The Legendre Transform (First Discussion)

C.1 Initial Definition of the Legendre Transform and the General Young Inequality

We call the *Legendre transform* of a function f with variable x a new function f^* with new variable x^*, defined by the relation

$$f^*\big(x^*\big) = \sup_x \big(x^* x - f(x)\big), \tag{C.1}$$

where the supremum is taken with respect to the variable x, for a fixed value of the variable x^*.

Problem 1

a) Check that the function f^* is convex in its domain of definition.

b) Draw a graph of the function f, the line $y = kx$ where $k = x^*$, and specify the geometric meaning of the value $f^*(x^*)$.

c) Find $f^*(x^*)$ for $f(x) = |x|$ and $f(x) = x^2$.

d) Prove that from the definition (C.1), it follows clearly that the inequality

$$x^* x \le f^*\big(x^*\big) + f(x) \tag{C.2}$$

is satisfied for all values of the arguments x, x^* in the domain of definition of the functions f and f^*, respectively.

Relation (C.2) is usually called the *general Young's inequality* or the *Fenchel–Young inequality*, and the function f^*, in convex analysis, for instance, is usually called the Young dual of the function f.

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

C.2 Specification of the Definition in the Case of Convex Functions

If the supremum involved in the definition (C.1) is attained at some inner point x from the domain of definition of the function f, and this function is smooth (or at least differentiable), then we find that

$$x^* = f'(x) \tag{C.3}$$

and therefore

$$f^*\bigl(x^*\bigr) = x^*x - f(x) = xf'(x) - f(x). \tag{C.4}$$

Thus in this case, the Legendre transform is specified in the form (C.3), giving the argument x^* and (C.4) providing the value $f^*(x^*)$ of the function f^*, i.e., the Legendre transform of the function f. (Notice that the operator $xf'(x) - f(x)$ was studied already by Euler.)

Moreover, if the function f is convex, then in the first place, the condition (C.3) will not only provide a local extremum, but also will be a local maximum (check it!), which in this case will be the global or absolute maximum.

In the second place, due to the monotonic increase of the derivative of a strictly convex function, Eq. (C.3) is uniquely solvable with respect to x, for such functions.

If Eq. (C.3) admits an explicit solution $x = x(x^*)$, then on substituting it in (C.4), we obtain an explicit expression for $f^*(x^*)$.

Problem 2

a) Find the Legendre transform of the function $\frac{1}{\alpha}x^\alpha$, for $\alpha > 1$, and obtain the classical Young's inequality

$$ab \le \frac{1}{\alpha}a^\alpha + \frac{1}{\beta}b^\beta, \tag{C.5}$$

where $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

b) What is the domain of definition of the Legendre transform of a smooth strictly convex function f having the lines ax and bx as asymptotes, for $x \to +\infty$ and $x \to -\infty$ respectively?

c) Find the Legendre transform of the function e^x and prove the inequality

$$xt \le \mathrm{e}^x + t \ln \frac{t}{\mathrm{e}}. \tag{C.6}$$

C.3 Involutivity of the Legendre Transform of a Function

As we already noted, Eq. (C.2), or equivalently the inequality

$$f(x) \ge xx^* - f^*\bigl(x^*\bigr), \tag{C.7}$$

holds for all values of the arguments x, x^* in the domains of definition of the functions f and f^*, respectively.

At the same time, as shown by formulas (C.3) and (C.4), if x and x^* are linked by the relation (C.3), then the last inequality, (C.7), becomes an equality, at least in the case of a smooth strictly convex function f. Recalling the definition (C.1) of the Legendre transformation, we conclude that in this case,

$$\left(f^*\right)^* = f. \tag{C.8}$$

So the Legendre transform of a smooth strictly convex function is involutive, i.e., the twofold application of this transform leads to the original function.

Problem 3

a) Is it true that $f^{**} = f$ for every smooth function f?

b) Is it true that $f^{***} = f^*$ for every smooth function f?

c) Differentiating Eq. (C.4), using (C.3) and provided that $f''(x) \neq 0$, show that $x = f^*(x^*)$ and therefore $f(x) = xx^* - f^*(x^*)$ (involutivity).

d) Check that at the corresponding points x, x^*, linked by Eq. (C.3), we have $f''(x) = 1/(f^*)''(x^*)$ and $f^{(3)} = -(f^*)^{(3)}(x^*)/((f^*)'')^2(x^*)$.

e) The family of lines $px + p^4$ depending on the parameter p is a family of tangents to a certain curve (the *envelope* of this family). Find the equation of this curve.

C.4 Concluding Remarks and Comments

As part of the discussion about convex functions, we gave the initial presentation of the Legendre transform at the level of functions of one variable. However, here we shall facilitate the perception of the Legendre transform and work with it in a number of important more general cases of applications, such as in theoretical mechanics, thermodynamics, equations of mathematical physics, calculus of variations, convex analysis, contact geometry, ..., and many more yet to be dealt with.

We shall analyze various details and possible developments of the concept of the Legendre transform. Here we add only the following remark. The argument of the Legendre transform is a derivative, or equivalently the differential of the original function, as is shown by Eq. (C.3).

If the argument x is, for instance, a vector of a linear space X equipped with a scalar product $\langle\cdot,\cdot\rangle$, then the generalization of the definition (C.1) is the equation

$$f^*\left(x^*\right) = \sup_x\left(\left\langle x^*, x\right\rangle - f(x)\right). \tag{C.9}$$

If we take x^* as a linear function on the space X, i.e., assume that x^* is an element of the dual space X^* and the action $x^*(x)$ of x^* on a vector x is denoted as before by $\langle x^*, x\rangle$, then Eq. (C.9) will continue to be meaningful, since the function f is defined on the space X, and therefore the Legendre transform f^* is defined in the space X^*, the dual of the space X.

Appendix D
The Euler–MacLaurin Formula

D.1 Bernoulli Numbers

Jacob Bernoulli found that $\sum_{n=1}^{N-1} n^k = \frac{1}{k+1} \sum_{m=0}^{k} C_{k+1}^m B_m N^{k+1-m}$, where $C_n^m = \frac{n!}{m!(n-m)!}$ are the binomial coefficients, and $B_0, B_1, B_2, \ldots$ are rational numbers, now called Bernoulli numbers. These numbers are met in different problems. They have the generating function $\frac{z}{e^z-1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n$, where they occur as coefficients in the Taylor expansion, and in this way they can be calculated.

These numbers can also be calculated with the following recurrence formula:

$$B_0 = 1, \qquad B_n = -\frac{1}{n+1} C_{n+1}^{k+1} B_{n-k}.$$

Problem 1 Find the first Bernoulli numbers and check that all Bernoulli numbers with odd indices, except B_1, are equal to zero, and that the signs alternate for the Bernoulli numbers with even indices. (The function $x/(e^x - 1) + x/2$ is even.)

Euler discovered the connection $B_n = -n\zeta(1-n)$ between the Bernoulli numbers and the Riemann ζ-function.

D.2 Bernoulli Polynomials

Bernoulli polynomials can be defined by various means. For example, Bernoulli polynomials are defined recursively by $B_0(x) \equiv 1$, $B_n'(x) = nB_{n-1}(x)$, with the condition that $\int_0^1 B_n(x)\,dx = 0$; they are also defined through the generating function $\frac{ze^{xz}}{e^z-1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n$, through the formula $B_n(x) = \sum_{k=0}^{n} C_n^k B_{n-k} x^k$, or through $B_n(x) = \sum_{m=0}^{n} \frac{1}{m+1} \sum_{k=0}^{m} (-1)^k C_m^k (x+k)^n$.

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

Problem 2

a) Based on the different definitions, compute the first few Bernoulli polynomials, and check the coincidence between them and the fact that Bernoulli numbers are the values of the Bernoulli polynomials at $x = 0$.

b) By differentiating the generating function, show that the Bernoulli polynomials $B_n(x)$ defined through this function satisfy the recurrence relation above, which in turn means that

$$B_n(x) = B_n + n \int_0^x B_{n-1}(t)\,\mathrm{d}t.$$

D.3 Some Known Operators and Series of Operators

Problem 3

a) If A is an operator, then we employ the notation $\frac{1}{A}$, as is usual for numbers, for referring to the operator A^{-1}, the inverse operator of A.

The integration operator $\int$ is the inverse of the differentiation operator D (with a proper setting of the integration constant). Similarly, the sum operator $\sum$ is the inverse of the difference operator Δ, whose action is defined as $\Delta f(x) = f(x+1) - f(x)$. Specify exactly how to find $\sum f(x)$.

b) Do you agree with the fact that $B_n(x) = D(\mathrm{e}^D - 1)^{-1}x^n$?

c) According to Taylor's formula,

$$\Delta f(x) = f(x+1) - f(x) = \frac{f'(x)}{1!} + \frac{f''(x)}{2!} + \cdots = \left(\frac{D}{1!} + \frac{D^2}{2!} + \cdots\right) f(x);$$

therefore, $\Delta = \mathrm{e}^D - 1$ and $\sum = \Delta^{-1} = (\mathrm{e}^D - 1)^{-1}$, and since $\frac{z}{\mathrm{e}^z - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k$, then

$$\sum = \frac{B_0}{D} + \frac{B_1}{1!} + \frac{B_2}{2!} D + \frac{B_3}{3!} D^2 + \cdots = \int + \sum_{k=1}^{\infty} \frac{B_k}{k!} D^k.$$

D.4 Euler–MacLaurin Series and Formula

By applying this operator relation to the function $f(x)$, we guess the Euler–MacLaurin summation formula, more precisely, not the formula itself but the corresponding series

$$\sum_{a \le n < b} f(n) = \int_a^b f(x)\,\mathrm{d}x + \sum_{k=1}^{\infty} \frac{B_k}{k!} f^{(k-1)}(x)\bigg|_0^1,$$

where a, b, n are integers and k is a natural number.

These series differ in the way that the Taylor series is different from Taylor's formula, which is finite and contains certain information (remainder term) providing the possibility of estimating the remainder (the value of the approximation error).

When the sum is reduced to a single term, the simplest and at the same time basic Euler–MacLaurin formula with a remainder term has the form

$$f(0) = \int_0^1 f(x)\,\mathrm{d}x + \sum_{k=1}^{m} \frac{B_k}{k!} f^{(k-1)}(x)\Big|_0^1 + (-1)^{(m+1)} \int_0^1 \frac{B_m(x)}{m!} f^{(m)}(x)\,\mathrm{d}x.$$

It is assumed here that the original function f is smooth enough, for example, that it possesses continuous derivatives of the required order.

Problem 4 Using the formula of integration by parts, prove by induction the Euler–MacLaurin formula written above. (Recall that Taylor's formula with remainder term of integral type can also be obtained with a simple integration by parts.)

D.5 The General Euler–MacLaurin Formula

The general Euler–MacLaurin formula, providing the value of the sum $\sum_{a\le n<b} f(n)$, has the form

$$\sum_{a\le n<b} f(n) = \int_a^b f(x)\,\mathrm{d}x + \sum_{k=1}^{m} \frac{B_k}{k!} f^{(k-1)}(x)\Big|_a^b + R_m,$$

where a, b, n are integers and k, m are natural numbers,

$$R_m = (-1)^{(m+1)} \int_a^b \frac{B_m(\{x\})}{m!} f^{(m)}(x)\,\mathrm{d}x,$$

and $\{x\}$ is the fractional part of the number x.

Problem 5 Prove this formula, given that every interval $[a, b]$ whose endpoints are integers can be divided into unit intervals (with length 1) and each unit interval can be translated to the interval $[0, 1]$ with a shift.

D.6 Applications

Problem 6

a) Using the Euler–MacLaurin formula and setting $f(x) = x^n$, show that $\sum_{a\le k<b} k^{m-1} = \frac{1}{m}\sum_{k=0}^{m} C_m^k B_k (b^{m-k} - a^{m-k})$, and in particular, obtain Jacob Bernoulli's relation $\sum_{0\le k<b} k^{m-1} = \frac{1}{m}\sum_{k=0}^{m} C_m^k B_k b^{m-k}$.

b) To calculate the asymptotic behavior of a sum or a series, usually the following kind of Euler–MacLaurin formula is used:

$$\sum_{n=a}^{b} f(n) \sim \int_a^b f(x)\,\mathrm{d}x + \frac{f(a)+f(b)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!}\big(f^{(2k-1)}(b) - f^{(2k-1)}(a)\big),$$

where a, b are integers. The formula often remains valid for the extension from the interval $[a,b]$ to the whole line. In many cases, the integral on the right side can be calculated in terms of elementary functions, even if the sum on the left side cannot be expressed. Then all the terms of the asymptotic series can be expressed in terms of elementary functions. For example,

$$\sum_{s=0}^{+\infty} \frac{1}{(z+s)^2} \sim \int_0^{+\infty} \frac{1}{(z+s)^2}\,\mathrm{d}s + \frac{1}{2z^2} + \sum_{k=1}^{+\infty} \frac{B_{2k}}{z^{2k+1}}.$$

Moreover, in this case, the integral can be calculated and is equal to $\frac{1}{z}$.

c) Setting $f(x) = x^{-1}$, prove the asymptotic formula

$$\sum_{k=1}^{n} \frac{1}{k} = \ln n + \gamma + \frac{1}{2n} + \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} - \theta_{m,n} \frac{B_{2m+2}}{(2m+2)n^{2m+2}},$$

where $0 < \theta_{m,n} < 1$, and γ is a constant (Euler's constant).

d) If we take $f(x) = \ln x$, show that

$$\sum_{k=1}^{n} \ln k = n \ln n - n + \sigma - \frac{1}{2}\ln n + \sum_{k=1}^{m} \frac{B_{2k}}{2k(2k-1)n^{2k-1}} - R_{m,n},$$

where

$$R_{m,n} = \phi_{m,n} \frac{B_{2m+2}}{(2m+1)(2m+2)n^{2m+1}},$$

$0 < \phi_{m,n} < 1$, and σ is a constant (in fact, equal to $\ln\sqrt{2\pi}$).

By raising to powers, we can obtain the asymptotic Stirling's formula for the values $n!$ as $n \to \infty$.

D.7 Again to the Actual Euler–MacLaurin Formula

Problem 7

a) If a and n are integers such that $a < n$ and f is a slowly changing function at the interval $[a,b]$, then the sum $S = \frac{1}{2}f(a) + f(a+1) + f(a+2) + \cdots + f(n-1) + \frac{1}{2}f(n)$ is a good approximation of the integral $I = \int_a^b f(x)\,\mathrm{d}x$.

Remember this by drawing a picture showing the geometric meaning of the quantities S and I, and at the same time recalling the numerical methods for calculating the integral.

b) If j is an integer, then integration by parts gives

$$\int_j^{j+1} f(x)\,\mathrm{d}x = \left(x - j - \frac{1}{2}\right) f(x)\Big|_j^{j+1} - \int_j^{j+1} \left(x - j - \frac{1}{2}\right) f'(x)\,\mathrm{d}x,$$

or

$$\frac{1}{2} f(j) + \frac{1}{2} f(j+1) = \int_j^{j+1} f(x)\,\mathrm{d}x + \int_j^{j+1} \omega_1(x) f'(x)\,\mathrm{d}x,$$

where $\omega_1(x) = x - [x] - \frac{1}{2} = \{x\} - \frac{1}{2}$.

(Recall that $[x]$ and $\{x\}$ are the integer and fractional parts of x, respectively.)

Summing these equalities for j from $j = a$ to $j = n - 1$, we obtain

$$S = I + \int_a^n \omega_1(x) f'(x)\,\mathrm{d}x.$$

Draw the graph of the function ω_1.

c) Now integrate the integral $\int_j^{j+1} \omega_1(x) f'(x)\,\mathrm{d}x$ by parts and obtain an expression with a new integral residue $\int_j^{j+1} \omega_2(x) f''(x)\,\mathrm{d}x$, where $\omega_2(x) = \int \omega_1(x)\,\mathrm{d}x$. Show the continuity of the function ω_2, given that $\int_j^{j+1} \omega_1(x)\,\mathrm{d}x = \int_0^1 (x - \frac{1}{2})\,\mathrm{d}x = 0$. Take the sum from $j = a$ to $j = n - 1$, as before, and obtain in one step a more advanced expression for the value of S with a new integral residue.

d) Continuing with this process, obtain the following Euler–MacLaurin formula:

$$S = I + \sum_{s=1}^{m-1} (-1)^{s+1} \omega_{s+1}(0) f^{(s)}(x)\Big|_a^n + (-1)^{m+1} \int_a^n \omega_m(x) f^{(m)}(x)\,\mathrm{d}x.$$

e) Check that $\omega_k(x) = \frac{1}{k!} B_k(\{x\})$ and compare the Euler–MacLaurin formula obtained with the one discussed previously.

Appendix E
Riemann–Stieltjes Integral, Delta Function, and the Concept of Generalized Functions

Basic Background

E.1 The Riemann–Stieltjes Integral

Specific Objective and Some Heuristic Arguments We have considered a number of examples of the effective use of the integral in the calculation of areas, volumes of solids of revolution, lengths of paths, work forces, energy .... We found the potential of the gravitational field and computed the escape velocity for Earth. Using the machinery of the integral calculus, we convinced ourselves of the fact that, for example, the path length does not depend on its parametrization. At the same time, we pointed out that certain calculations (e.g., the length of an ellipse) are associated with nonelementary functions (in this case elliptic functions).

All the quantities mentioned above (length, area, volume, work, ...) are additive like the Riemann integral. We know that every additive function $I[\alpha, \beta]$ of the oriented interval $[\alpha, \beta] \subset [a, b]$ has the form $I[\alpha, \beta] = F(\beta) - F(\alpha)$ if we set $F(x) = I[a, x] + C$. In particular, we can take an arbitrary function F and define an additive function $I[\alpha, \beta] = F(\beta) - F(\alpha)$ from it, considering $F(x) = I[a, x]$. If the function F is discontinuous on the interval $[a, b]$, then the function $I[a, x]$ is also discontinuous there. But then it cannot be represented as a Riemann integral $\int_a^x p(t)\,\mathrm{d}t$ of any Riemann integrable function (a density p), because such an integral, as we know, is continuous at x.

Suppose, for instance, that the interval $[-1, 1]$ is a string having a bead of mass 1 in the middle. If $I[\alpha, \beta]$ is a mass within the interval $[\alpha, \beta] \subset [-1, 1]$, then the function $I[-1, x]$ is equal to zero for $1 \le x < 0$ and equal to 1 for $0 \le x \le 1$. If we tried to describe the distribution of the mass in the segment in terms of the density distribution (i.e., in terms of the limit of the ratio between the mass lying in a neighborhood of the point and the size of the neighborhood when the latter is contracted to a point), then we would have to define the density p as $p(x) = 0$ for $x \ne 0$ and $p(x) = +\infty$ for $x = 0$. Physicists and all scientists after Dirac called p a "function" (this distribution density). Today, we call p a "generalized function" or

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

"distribution". We denote it by δ and it is defined as $\int_\alpha^\beta \delta(x)\,\mathrm{d}x = 1$ if $\alpha < 0 < \beta$ and $\int_\alpha^\beta \delta(x)\,\mathrm{d}x = 0$ if $\alpha < \beta < 0$ or $0 < \alpha < \beta$, no matter what the numbers α and β are.

Of course, according to the traditional definition of the integral, e.g., the Riemann integral, this integral does not make sense (for the simple reason that the integrand is an unbounded "function"). We make liberal use of the integral symbol here. It is used only as a replacement of the additive function $I[\alpha, \beta]$, discussed above, when we considered the bead on a string.

Example 4 (Center of mass) Recall the fundamental equation $m\ddot{r} = F$ describing the motion of a point mass m due to the effect of a force F, where r is the radius vector of the point. If there is a system of n material points, then for each of them we have the equation $m_i\ddot{r}_i = F_i$. By summing all these equalities, we obtain the equation $\sum_{i=1}^n m_i\ddot{r}_i = \sum_{i=1}^n F_i$, which can be rewritten in the form $M\sum_{i=1}^n \frac{m_i}{M}\ddot{r}_i = \sum_{i=1}^n F_i$, or in the form $M\ddot{r}_M = F$, where $M = \sum_{i=1}^n m_i$, $F = \sum_{i=1}^n F_i$, and $\ddot{r}_M = \sum_{i=1}^n \frac{m_i}{M}\ddot{r}_i$. That means that if the total mass of the system is placed at some point in the space, with the radius vector being equal to $r_M = \sum_{i=1}^n \frac{m_i}{M} r_i$, and under the influence of the force $F = \sum_{i=1}^n F_i$, then the mass will move according to Newton's law, no matter how complex the mutual motion of the individual parts of the system is.

The point of the space we found with radius vector $\sum_{i=1}^n \frac{m_i}{M} r_i$ is called the *center of mass of the system of material points*.

Now suppose we have the task of finding the center of mass of a material body, i.e., a region D of space, in which a mass is somehow distributed. Let $\mathrm{d}v$ be the volume element, the concentrated mass $\mathrm{d}m$, and M the total mass of the body D. Then we suppose that $M = \int_D \mathrm{d}m$, and then the center of mass can be found with the formula $\frac{1}{M}\int_D r\,\mathrm{d}m$, where r is the radius vector of the mass element.

We still do not know how to integrate over domains in space, and therefore, we consider the one-dimensional case, which is also quite informative. Thus, instead of the region D, we consider the interval $[a, b]$ of the coordinate axis $\mathbb{R}$.

Then $M = \int_a^b \mathrm{d}m$, and the center of mass is found with the formula $\frac{1}{M}\int_a^b x\,\mathrm{d}m$, where x is the coordinate of the mass element $\mathrm{d}m$, which therefore can be written more precisely as $\mathrm{d}m(x)$.

The meaning of what we wrote evidently must be as follows. Take a partition P of the interval $[a, b]$, with some marked points $\xi_i \in [x_{i-1}, x_i]$. To the interval $[x_{i-1}, x_i]$ corresponds the mass Δm_i. We consider the sums $\sum_i \Delta m_i$, $\sum_i \xi_i \Delta m_i$, and passing to the limit as the parameter $\lambda(P)$ of the partition tends to zero, we obtain respectively what is denoted by $\int_a^b \mathrm{d}m$ and $\int_a^b x\,\mathrm{d}m$.

We arrive at the following generalization of the Riemann integral.

Definition of the Riemann–Stieltjes Integral[1] Let f, g be real-, complex-, or vector-valued functions defined over the interval $[a, b] \subset \mathbb{R}$. Let $(P, \xi) = (a =$

[1]T.J. Stieltjes (1856–1894) – Dutch mathematician.

$x_0 \le \xi_1 \le x_1 \le \cdots \le \xi_n \le x_n = b$) be a partition of this interval with marked points and parameter $\lambda(P)$. We consider the sum $\sum_{i=1}^{n} f(\xi_i)\Delta g_i$, where $\Delta g_i = g(x_i) - g(x_{i-1})$.

The *Riemann–Stieltjes integral of the function f with respect to the function g over the interval* $[a, b]$ is defined as the following value:

$$\int_a^b f(x)\,\mathrm{d}g(x) := \lim_{\lambda(P)\to 0} \sum_{i=1}^{n} f(\xi_i)\Delta g_i \tag{E.1}$$

if the above limit exists.

In particular, when $g(x) = x$, we return to the standard Riemann integral.

E.2 Case in Which the Riemann–Stieltjes Integral Reduces to the Riemann Integral

Note also that if the function g is smooth and f is a Riemann-integrable function over the interval $[a, b]$, then

$$\int_a^b f(x)\,\mathrm{d}g(x) = \int_a^b f(x)g'(x)\,\mathrm{d}x, \tag{E.2}$$

i.e., in this case, the computation of the Riemann–Stieltjes integral reduces to the computation of the Riemann integral of the function fg' on the same interval.

Indeed, using the smoothness of the function g and the mean value theorem, we can rewrite the sum at the right side of Eq. (E.1) in the following form:

$$\sum_{i=1}^{n} f(\xi_i)\Delta g_i = \sum_{i=1}^{n} f(\xi_i)\big(g(x_i) - g(x_{i-1})\big) = \sum_{i=1}^{n} f(\xi_i)g'(\tilde{\xi}_i)(x_i - x_{i-1}) =$$
$$= \sum_{i=1}^{n} f(\xi_i)g'(\xi_i)\Delta x_i + \sum_{i=1}^{n} f(\xi_i)\big(g'(\tilde{\xi}_i) - g'(\xi_i)\big)\Delta x_i .$$

By the uniform continuity of the function g' on the interval $[a, b]$ and the boundedness of the function f, the last sum approaches zero as $\lambda(P) \to 0$. The first sum is the usual integral sum for the integral, which appears at the right side of (E.2). By our assumptions about the functions f and g, the function fg' is Riemann integrable over the interval $[a, b]$. Thus, the sum above approaches the value of this integral for $\lambda(P) \to 0$, which completes the proof of Eq. (E.2).

Problem 1 We achieved the proof of this equality using the mean value theorem, which holds for real-valued functions. Using the general finite increment theorem, complete the proof for vector-valued functions (for instance, complex-valued functions).

E.3 Heaviside Function and an Example of a Riemann–Stieltjes Integral Computation

The *Heaviside step function* is defined through $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $0 \leq x$. Let us calculate the integral $\int_a^b f(x)\,\mathrm{d}H(x)$. Following the definition (E.1), we write the sum $\sum_{i=1}^n f(\xi_i)\Delta H_i = \sum_{i=1}^n f(\xi_i)(H(x_i) - H(x_{i-1}))$. Because of the definition of the Heaviside function, this sum is clearly equal to zero if the point 0 is not contained in the interval $[a, b]$, and equal to $f(x_i)$ if the point 0 falls into some of the intervals $[x_{i-1}, x_i]$ (more precisely, in the interior of it or at its endpoint x_i). In the first case, the integral is of course zero.

In the second case, under the limit $\lambda(P) \to 0$, the point $\xi_i \in [x_{i-1}, x_i]$ approaches 0. Therefore, if the function f is continuous at 0, then the limit of the sum above will be $f(0)$.

If the function f is discontinuous at 0, with small changes of the value of ξ_i it is possible to change essentially the value of $f(\xi_i)$, and thus the sums of integrals will not have a limit for $\lambda(P)$.

It is clear that the last calculation has a general nature, since the occurrence of joint points of discontinuity for the functions f, g involved in the Riemann–Stieltjes integral leads to the nonexistence of the limit if such a joint point occurs in the interior of the integration interval.

Therefore, the calculation above shows that if φ is, for instance, a function of class $C_0(\mathbb{R}, \mathbb{R})$, i.e., a function defined on the whole real line and continuous, that is identically zero outside of some bounded set, then

$$\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}H(x) = \varphi(0). \tag{E.3}$$

E.4 Generalized Functions

E.4.1 Dirac's Delta Function. A Heuristic Description

As we previously remarked, physicists among other scientists use the delta function δ after its introduction by Dirac. This "function" is zero everywhere except at the origin, where its value is infinity. Along with this (and this is really important),

$$\int_\alpha^\beta \delta(x)\,\mathrm{d}x = 1 \quad \text{if } \alpha < 0 < \beta,$$

and

$$\int_\alpha^\beta \delta(x)\,\mathrm{d}x = 0 \quad \text{if } \alpha < \beta < 0 \text{ or if } 0 < \alpha < \beta,$$

for all real values α and β.

It is natural to assume that the multiplication of the integrating function by a number leads to the product of this number with the integral. But then, if a function φ is continuous at the origin, given that it is almost constant in some small neighborhood $U(0)$ of the origin and that $\int_{U(0)} \delta(x)\,\mathrm{d}x = 1$, we conclude that the following relation should hold:

$$\int_{\mathbb{R}} \varphi(x)\delta(x)\,\mathrm{d}x = \varphi(0). \tag{E.4}$$

By comparing Eqs. (E.2), (E.3), and (E.4) and continuing this chain of findings, we conclude that

$$H'(x) = \delta(x). \tag{E.5}$$

Of course, it does not fit into our classical setting. However, these considerations are very constructive, and if one were obliged to write a value $H'(x)$, then one would write what we now write, namely 0 if $x \neq 0$ and $+\infty$ if $x = 0$.

E.5 The Correspondence Between Functions and Functionals

One possible way out of this difficulty consists in the following idea of extension (generalization) of the concept of "function".

We shall look at the function through its interaction with other functions. (As usual, we are not interested in the internal structure of a device, such as a human, and we consider that we know an object if we know how the object responds to certain input actions or incoming questions.)

Take an integrable function f on the interval $[a, b]$, and consider the functional A_f (a function over functions) generated by f:

$$A_f(\varphi) = \int_a^b f(x)\varphi(x)\,\mathrm{d}x. \tag{E.6}$$

In order to simplify the technical difficulties, we shall consider smooth *test functions* even of class $C_0^\infty[a, b]$, i.e., infinitely differentiable functions vanishing in a neighborhood of the endpoints. It is even possible to continue both functions f and φ as zero outside the interval $[a, b]$, and instead of writing the integral over the interval, we can write the integral as

$$A_f(\varphi) = \int_{\mathbb{R}} f(x)\varphi(x)\,\mathrm{d}x. \tag{E.7}$$

Knowing the value of the functional A_f on test functions, we can find easily, if necessary, the value $f(x)$ of the function f at any point where this function is continuous.

Problem 2

a) Prove that the value $\frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon} f(t)\,\mathrm{d}t$ (integral average) tends to $f(x)$ for $\varepsilon \to 0$ at every point of continuity of the integrable function f.

b) Show that the step function $\overline{\delta}_\varepsilon$, equal to zero outside the interval $[\varepsilon, \varepsilon]$ and equal to $\frac{1}{2\varepsilon}$ inside this interval (the function $\overline{\delta}_\varepsilon$ imitates Dirac's δ function), can be approximated with smooth functions δ_ε with these properties: $\delta_\varepsilon(x) \geq 0$ in $\mathbb{R}$, $\delta_\varepsilon(x) = 0$ for $|x| \geq \varepsilon$, and $\int_{\mathbb{R}} \delta_\varepsilon(x)\,\mathrm{d}x = 1$.

c) Show that if $\varepsilon \to 0$, then $\int_{x-\varepsilon}^{x+\varepsilon} f(t)\delta_\varepsilon(x-t)\,\mathrm{d}t \to f(x)$ at every point of continuity of the integrating function f.

E.6 Functionals as Generalized Functions

Thus, an integrable function f provides a linear functional A_f (a linear function in the vector space of functions $C_0^\infty[a,b]$ or $C_0^\infty(\mathbb{R})$) defined through the formula (E.6) or (E.7), and moreover, with the use of the functional A_f, the integrable function f itself can be restored at all its points of continuity (i.e., almost everywhere). Therefore, the functional A_f can be thought of as a different encoding or interpretation of the function f considered in the mirror of functionals.

But in this mirror it is possible to find some other linear functionals that are not given through the integration of any function. As an example, we have the functional that we have studied before, $\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}H(x) = \varphi(0)$, which we denote by A_δ (given that we would like to write $\delta(x)\,\mathrm{d}x$ instead of $\mathrm{d}H(x)$).

The functionals of the first type are called *regular*, while those of the second type are called *singular*.

We shall consider functionals as *generalized functions*. This set of functionals contains our usual functions as a subset, consisting of all regular functionals.

Thus, by relating the Riemann integral and its generalization the Riemann–Stieltjes integral, we gave an overview of the construction of generalized functions. We shall not delve into the details of the theory of generalized functions, which are related, for instance, to different spaces of test functions and the construction of linear functionals (generalized functions) on them. We prefer to prove the rule for differentiating generalized functions. As a final remark showing the usefulness of the Stieltjes integral, we would like to add here that on the space $C[a,b]$ of continuous functions φ on the interval $[a,b]$, every linear continuous functional (either regular or singular) can be represented as a Riemann–Stieltjes integral $\int_a^b \varphi(x)\,\mathrm{d}g(x)$ for some properly selected function g. (For instance, the singular functional A_δ representing the generalized function δ has the form $\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}H(x)$, shown in Eq. (E.3).)

We began with an example in which we found the Stieltjes integral $\int_a^b x\,\mathrm{d}m(x)$ in the determination of the center of mass. The integral $M_n = \int_a^b x^n\,\mathrm{d}m(x)$ is called the *moment of order n* relative to a measure (e.g., a probability measure), mass, or charge distributed over the interval $[a,b]$. The moments M_0, M_1, M_2 are frequently met: M_0 is the total mass (or measure of charge); M_1/M_0 provides the center of mass in mechanics, and M_1 is the mathematical expectation (or expected value) of a random variable in probability theory; M_2 is the moment of inertia in mechanics and the scattering of a random variable with expected value $M_1 = 0$ in probability theory. One of the problems of the theory of moments is the restoration of a distribution through the computation of the moments.

E.7 Differentiation of Generalized Functions

Let A be a generalized function. What generalized function A' should be considered the derivative of A?

Let us consider first the derivative for regular generalized functions, i.e., for functionals A_f generated by a classical function f, in our case a smooth compactly supported function of class $C_0^{(1)}$. Then it is natural to consider $A_{f'}$ the derivative A'_f of A_f, generated by the function f', the derivative of the original function f.

Using integration by parts, we find that

$$A'_f(\varphi) := A_{f'}(\varphi) = \int_{\mathbb{R}} f'(x)\varphi(x)\,\mathrm{d}x = f(x)\varphi(x)\big|_{-\infty}^{\infty} - \int_{\mathbb{R}} f(x)\varphi'(x)\,\mathrm{d}x =$$
$$= -\int_{\mathbb{R}} f(x)\varphi'(x)\,\mathrm{d}x = A_f\big(\varphi'\big).$$

Therefore, we find that in this case,

$$A'_f(\varphi) = -A_f\big(\varphi'\big). \tag{E.8}$$

This provides a reason to adopt the following definition of *derivative*:

$$A'(\varphi) := -A\big(\varphi'\big). \tag{E.9}$$

It is indicated here how the functional A' acts on a function $\varphi \in C_0^{(\infty)}$. Therefore, the functional A' is well defined.

The action of a linear functional on a function φ is frequently written in the form $\langle A, \varphi\rangle$, instead of $A(\varphi)$, recalling the scalar product, to emphasize that this product is linear in both of its variables.

With this notation if f is any generalized function, then according to Eq. (E.9), we have

$$\big\langle f', \varphi\big\rangle = -\big\langle f, \varphi'\big\rangle. \tag{E.10}$$

E.8 Derivatives of the Heaviside Function and the Delta Function

We shall compute the derivative of the Heaviside function, considering it a generalized function acting according to the usual rule of regular generalized functions

$$\langle H, \varphi\rangle = \int_{\mathbb{R}} H(x)\varphi(x)\,\mathrm{d}x.$$

Following the definitions (E.9) and (E.10), we have that

$$\big\langle H', \varphi\big\rangle := -\big\langle H, \varphi'\big\rangle := \int_{\mathbb{R}} H(x)\varphi'(x)\,\mathrm{d}x = -\int_0^{+\infty} \varphi'(x) =$$
$$= -\varphi(x)\big|_0^{+\infty} = \varphi(0).$$

We have shown that $\langle H', \varphi \rangle = \varphi(0)$. However, after the definition of the δ function we have $\langle \delta, \varphi \rangle = \varphi(0)$. Hence, we have proved that in terms of generalized functions, we have the equality

$$H' = \delta.$$

Let us compute, for example, δ' and δ'', i.e., we determine the action of the following functionals:

$$\big\langle \delta', \varphi \big\rangle := -\big\langle \delta, \varphi' \big\rangle := -\varphi'(0);$$
$$\big\langle \delta'', \varphi \big\rangle := -\big\langle \delta', \varphi' \big\rangle := \varphi''(0).$$

It is clear now that in general, we have $\langle \delta^{(n)}, \varphi \rangle = (-1)^n \varphi^{(n)}(0)$.

We realize that generalized functions are infinitely differentiable. This is their remarkable property, which has many consequences. This property allows operations that with usual functions are possible only under very special conditions.

To conclude, we would like to make the following remark of a general nature. Let X be a vector space and X^* its dual space, consisting of linear functions on X, and let X^{**} be the dual space of X^*. We shall write the value $x^*(x)$ of the function $x^* \in X^*$ at the vector $x \in X$ as a scalar product $\langle x^*, x \rangle$, as we did before. By fixing x, we obtain a linear function with respect to x^*. Thus every element of X can be interpreted as an element of X^{**}, i.e., we have an embedding $I : X \to X^{**}$. In the finite-dimensional case, all the spaces X, X^*, and X^{**} are isomorphic, and $I(X) = X^{**}$. In the general case, $I(X) \subsetneq X^{**}$, i.e., $I(X)$ is only a subset of the whole space X^{**}. This is what is observed in the transition from functions (corresponding to regular functionals) to generalized functions, which turned out to be a larger space.

Appendix F
The Implicit Function Theorem (An Alternative Presentation)

F.1 Formulation of the Problem

The formulation of the problem and the heuristic arguments are discussed, of course, in a course lecture; but we shall omit this here, since the relevant material can be read in Sect. 8.5 of Chap. 8.

We shall use a different approach for the proof of the implicit function theorem here, splendid and independent from that we presented in Sect. 8.5 of Chap. 8. This theorem assumes a somehow more advanced audience of readers, despite its conceptual simplicity, beauty, and generality. These readers are in general already more familiar with some general mathematical concepts, presented at the beginning of the second part of the textbook. In any case, all this information allows us to appreciate the real generality of the method, which we can show without loss of generality on simple visual examples in our familiar spaces.

F.2 Some Reminders of Numerical Methods to Solve Equations

By fixing one of the variables in the equation $F(x, y) = 0$, we obtain an equation in terms of the other variable. Therefore, it might be useful to remember how to solve equations $f(x) = 0$.

1) According to the properties of the given function f, one chooses the methods of solution.

For instance, if the function f is real-valued, continuous, and taking values with different signs at the endpoints of the interval $[a, b]$, then we know that in this interval there is at least one root of the equation $f(x) = 0$, and it is possible to find it through successive divisions of the interval. By dividing the interval in half, we get either the root or a half-interval where the function takes values with different signs at the endpoints. Continuing with this dividing process, we obtain a sequence (the endpoints of the intervals) converging to the root of this equation.

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

2) If f is a smooth convex function, then following Newton's method, it is possible to propose in this case a more efficient algorithm, in the sense of speed of convergence, to find a root.

Newton's method or the *method of tangents* works as follows. We build the tangent at the point x_0, we find the intersection point with the x-axis, and by repeating this process, we obtain a sequence of points with a recurrence relation

$$x_{n+1} = x_n - \left(f'(x_n)\right)^{-1} f(x_n) \tag{F.1}$$

converging rapidly to the root. (Estimate the velocity of convergence. Obtain the relation $x_{n+1} = \frac{1}{2}(x_n + a/x_n)$, allowing you to find the positive root of the equation $x^2 - a = 0$. Find $\sqrt{2}$ according to this formula with the desired accuracy and detect how many additional correct digits appear at each step.)

3) Equation (F.1) can be written in the form

$$x_{n+1} = g(x_n), \tag{F.2}$$

where $g(x) = x - (f'(x))^{-1} f(x)$. Thus, finding the roots of the equation for $f(x)$ reduces to finding a fixed point of the mapping g, i.e., a point such that

$$x = g(x). \tag{F.3}$$

This reduction, as we know, applies not only to Newton's method. In fact, the equation $f(x) = 0$ is equivalent to the equation $\lambda f(x) = 0$ (if λ^{-1} exists), and that is equivalent to the equation $x = x + \lambda f(x)$. Setting $g(x) = x + \lambda f(x)$ (λ can be a variable here), we arrive at Eq. (F.3).

The process of solution of (F.3), i.e., finding a fixed point of the mapping g in accordance with the recursive formula (F.2), is called an iterative process or method of iterations, as we already know. This means that the value found in the previous step becomes the argument or input of the function f in the next step. This cyclic process is suitable for implementation in a computer.

If the iteration process (F.2) is done in a region where $|g'(x)| \le q < 1$, then the sequence

$$\begin{aligned}
&x_0,\\
&x_1 = g(x_0),\\
&x_2 = g(x_1) = g^2(x_0),\\
&\vdots\\
x_{n+1} &= g(x_n) = g^n(x_0)
\end{aligned}$$

is always fundamental (or a Cauchy sequence). Indeed, by applying the mean value theorem, we have

$$|x_{n+1} - x_n| \le q|x_n - x_{n-1}| \le \cdots \le q^n|x_1 - x_0|. \tag{F.4}$$

We apply the triangle inequality to it, and we obtain

$$|x_{n+m} - x_n| \le |x_n - x_{n+1}| + \cdots + |x_{n+m-1} - x_{n+m}| \le$$
$$\le \left(q^n + \cdots + q^{n+m-1}\right)|x_1 - x_0| \le \frac{q^n}{1-q}|x_1 - x_0|. \qquad \text{(F.5)}$$

It is useful to remark that if we take the limit $m \to \infty$ in the last inequality, we obtain the estimate

$$|x - x_n| \le \frac{q^n}{1-q}|x_1 - x_0|, \qquad \text{(F.6)}$$

the deviation or evasion of x_n from the fixed point x.

Problem 1 Draw several variants of the curve $y = g(x)$ intersecting the line $y = x$ and a diagram simulating an iterative process $x_{n+1} = g(x_n)$ for finding a fixed point.

F.2.1 The Principle of the Fixed Point

The last arguments (relating formulas (F.3)–(F.6)) can obviously be applied in any metric space in which the Cauchy criterion is valid, i.e., where every fundamental sequence is convergent. Such metric spaces are called *complete metric spaces*. For instance, $\mathbb{R}$ is a complete metric space with respect to the standard distance $d(x', x'') = |x' - x''|$ between points $x', x'' \in \mathbb{R}$. The interval $I = \{x \in \mathbb{R} \mid |x| \le 1\}$ is also a complete metric space with respect to this metric. If we remove a point from $\mathbb{R}$ or I, then clearly, the resulting metric space will not be complete.

Problem 2

a) Prove the completeness of the spaces $\mathbb{R}^n$, $\mathbb{C}$, $\mathbb{C}^n$.

b) Show that the closed ball $B(a, r) = \{x \in X \mid d(a, x) \le r\}$ with radius r and center $a \in X$ in a complete metric space (X, d) is itself a complete metric space with respect to the induced metric d from the embedding $B \subset X$.

We recall now the following definition.

Definition 1 A mapping $g : X \to Y$ from a metric space (X, d_X) into another (Y, d_Y) is called a *contraction* if there exists a number $q \in [0, 1[$ such that for arbitrary points $x', x'' \in X$,

$$d_Y\left(g\left(x'\right), g\left(x''\right)\right) \le q d_X\left(x', x''\right).$$

For example, if $g : \mathbb{R} \to \mathbb{R}$ is a differentiable function with the property that everywhere $|g'(x)| \le q < 1$, then by the mean value theorem, we have $|g(x') - g(x'')| \le q|x' - x''|$, and therefore g is a contraction mapping. The same can be said about a differentiable mapping $g : B \to Y$ from a convex subset B of a normed space X (for instance, the ball $B \subset \mathbb{R}^n$) to a normed space Y if $\|g'(x)\| \leq q$ at every point $x \in B$.

We are able to formulate now the following *fixed-point principle*.

A contraction mapping $g : X \to X$ of a complete metric space into itself has a unique fixed point x.

This point can be found through the iterative process $x_{n+1} = g(x_n)$, starting with any point $x_0 \in X$. The speed of convergence and the error estimate for the approximation are given by the inequality

$$\mathrm{d}(x, x_n) \leq \frac{q^n}{1-q}\,\mathrm{d}(x_1, x_0). \tag{F.6$'$}$$

The proof of this fact was given above by the deduction of formulas (F.4)–(F.6), where instead of $|x' - x''|$ we have to write everywhere $\mathrm{d}(x', x'')$.

In order to appreciate the usefulness and scope for generalizing this principle, consider the following important example.

Example 1 We look for the function $y = y(x)$ satisfying the differential equation $y' = f(x, y)$ and the initial condition $y(x_0) = y_0$.

Using the formula of Newton–Leibniz, we rewrite the problem in the form of the following integral equation for the unknown function $y(x)$:

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t. \tag{F.7}$$

On the right-hand side there is a mapping g, which acts on the function $y(x)$, and we look for the fixed "point" of the mapping (action) g.

For example, let $f(x, y) = y$, $x_0 = 0$, and $y_0 = 1$. Then we deal with the solution of the equation $y' = y$ with the initial condition $y(0) = 1$, and Eq. (F.7) takes the form

$$y(x) = 1 + \int_{0}^{x} y(t)\,\mathrm{d}t. \tag{F.8}$$

We then carry out the iterative process, starting with the function $y_0(x) \equiv 0$, and we successively obtain

$$\begin{aligned}
y_1(x) &= 1,\\
y_2(x) &= 1 + \int_0^x y_1(t)\,\mathrm{d}t = 1 + x,\\
y_3(x) &= 1 + \int_0^x y_2(t)\,\mathrm{d}t = 1 + \int_0^x (1+t)\,\mathrm{d}t = 1 + x + \frac{1}{2}x^2,\\
&\ \ \vdots
\end{aligned}$$

$$y_n(x) = 1 + \frac{1}{1!}x + \cdots + \frac{1}{n!}x^n,$$

$$\vdots$$

It is clear, that we obtain the function $e^x = 1 + \frac{1}{1!}x + \cdots + \frac{1}{n!}x^n + \cdots$.

Problem 3 Show that if $\|f(x, y_1) - f(x, y_2)\| \le M\|y_1 - y_2\|$, then in a neighborhood of the point x_0 the iteration process is applicable in the case of the more general equation (F.7).

In this way, Picard (Émile Picard, 1856–1941) was looking for the solution of the differential equation $y'(x) = f(x, y(x))$, with the initial condition $y(x_0) = y_0$ as a fixed point of the mapping (F.7).

Banach (Stefan Banach, 1882–1945) formulated the fixed-point principle in the abstract form above, and in this form it is often called *Banach's fixed-point principle* or the *Banach–Picard principle*. However, its origins can be traced back to Newton, as we have seen.

F.3 The Implicit Function Theorem

F.3.1 Statement of the Theorem

We return now to the main object of our consideration and prove the implicit function theorem.

Theorem *Let X, Y, Z be normed spaces (for example, $\mathbb{R}^m, \mathbb{R}^n, \mathbb{R}^m$ or even $\mathbb{R}, \mathbb{R}, \mathbb{R}$), and suppose moreover that Y is a complete metric space with respect to the metric induced by the norm. Let $F : W \to Z$ be a mapping defined in a neighborhood W of the point $(x_0, y_0) \in X \times Y$, continuous at (x_0, y_0), together with the partial derivative $F'_y(x, y)$, which is supposed to exist in W. If $F(x_0, y_0) = 0$ and there exist $(F'_y(x_0, y_0))^{-1}$ and $\|(F'_y(x_0, y_0))^{-1}\| < \infty$, then there exist a neighborhood $U = U(x_0)$ of the point x_0 in X, a neighborhood $V = V(y_0)$ of the point y_0 in Y, and a function $f : U \to V$, continuous at x_0, such that $U \times V \subset W$ and*

$$\big(F(x, y) = 0 \text{ within } U \times V\big) \Leftrightarrow \big(y = f(x), x \in U\big). \tag{F.9}$$

In short, under the conditions of the theorem, the set determined by the relation $F(x, y) = 0$ within the neighborhood $U \times V$ is the graph of the function $y = f(x)$.

F.3.2 Proof of the Existence of an Implicit Function

Proof Without loss of generality and for brevity, we may assume that $(x_0, y_0) = (0, 0)$, which can always be achieved by the change of variables $x - x_0$ and $y - y_0$.

For a fixed x we shall solve the equation $F(x, y) = 0$ with respect to y. We look for the solution as the fixed point of the mapping

$$g_x(y) = y - \big(F'_y(0, 0)\big)^{-1} F(x, y). \tag{F.10}$$

This is a simplified version of Newton's formula (F.1), where the coefficient λ is constant (see the paragraph following formula (F.3)). It is immediately clear that $F(x, y) = 0 \Leftrightarrow g_x(y) = y$.

The mapping (F.10) is a contraction if (x, y) is near $(0, 0) \in X \times Y$. Indeed,

$$\frac{\mathrm{d}g_x}{\mathrm{d}y}(y) = E - \big(F'_y(0, 0)\big)^{-1} F'_y(x, y). \tag{F.11}$$

Here E is the identity (unitary) mapping, and since $F'_y(x, y)$ is continuous at the point $(0, 0)$, there exists a number $\Delta \in \mathbb{R}$ such that for $\|x\| < \Delta$ and $\|y\| < \Delta$,

$$\left\| \frac{\mathrm{d}g_x}{\mathrm{d}y} \right\| < \frac{1}{2}. \tag{F.12}$$

Finally, note that for every $\varepsilon \in]0, \Delta[$ there is $\delta \in]0, \Delta[$ such that if $\|x\| < \delta$, then the function g_x maps the interval (ball) $\|y\| \le \varepsilon$ into itself.

Indeed, because of $F(0, 0) = 0$ and from Eq. (F.10), we have $g_0(0) = 0$. In view of the continuity of F at the point $(0, 0)$, it follows from (F.10) that there is $\delta \in]0, \Delta[$ such that $\|g_x(0)\| < \frac{1}{2}\varepsilon$ for $\|x\| < \delta$.

Thus, for $\|x\| < \delta$, the mapping $g_x : B(\varepsilon) \to Y$ displaces the center of the interval $B(\varepsilon) = \{y \in Y \mid \|y\| \le \varepsilon\}$ no more than $\frac{1}{2}\varepsilon$. Therefore, by virtue of (F.12) it decreases $B(\varepsilon)$ by at least a factor of two. Hence, $g_x(B(\varepsilon)) \subset B(\varepsilon)$ for $\|x\| < \delta$.

By assumption, Y is a complete space, and therefore $B(\varepsilon) \subset Y$ is also a complete metric space (with respect to the induced metric).

Then by virtue of the fixed-point principle, there is a unique point $y = f(x) \in B(\varepsilon)$ that is fixed for the mapping $g_x : B(\varepsilon) \to B(\varepsilon)$.

Thus for every x with $\|x\| < \delta$, we have found a unique value $y = f(x)$ ($\|f(x)\| < \varepsilon$) in the neighborhood $B(\varepsilon)$ such that $F(x, f(x)) = 0$.

(The cross section of the domain $P = \{(x, y) \in X \times Y \mid \|x\| < \delta, \|y\| < \varepsilon\}$ passing through the point $(x, 0)$ is the interval (ball) $B(\varepsilon)$ in which lies the corresponding fixed point $y = f(x)$.)

Thus, we have shown that

$$\big(F(x, y) = 0 \text{ for } \|x\| < \delta \text{ and } \|y\| < \varepsilon\big) \Leftrightarrow \big(y = f(x) \text{ for } \|x\| < \delta\big). \tag{F.13}$$

Note that not only have we obtained the relation (F.9), but also, by virtue of the construction, for every $\varepsilon \in]0, \Delta[$ we can choose $\delta > 0$ such that (F.13) holds. Since the function f has been found already and is fixed, we have that $f(0) = 0$ and f is continuous at $x = 0$. □

The theorem just proved can be regarded as the existence theorem of the implicit function $y = f(x)$.

We shall see now what properties of the function F are inherited by the function f.

F.3.3 Continuity of an Implicit Function

If in addition to the conditions of the theorem, we know that the functions F and F'_y are continuous not only at the point (x_0, y_0) but also in some neighborhood of this point, then the implicit function is also continuous in some neighborhood of this point.

Proof Indeed, in this case, the conditions of the theorem will be fulfilled at all the points of the set $F(x, y) = 0$ near (x_0, y_0), and each of them could be considered a starting point (x_0, y_0). The function f has been found already, and therefore is fixed.

Warning! Recall the exercise that if the mapping $A \mapsto A^{-1}$, where A is mapped to its inverse (for example for a matrix A) is defined on A, then it is defined on a neighborhood of A. □

F.3.4 Differentiability of an Implicit Function

If in addition to the conditions of the theorem, we know that the function F is differentiable at the point (x_0, y_0), then the implicit function f is also differentiable at the point x_0, and moreover,

$$f'(x_0) = -\bigl(F'_y(x_0, y_0)\bigr)^{-1} F'_x(x_0, y_0). \tag{F.14}$$

Proof Given the differentiability of F at the point (x_0, y_0), we can write

$$\begin{aligned} &F(x, y) - F(x_0, y_0) = \\ &\quad = F'_x(x_0, y_0)(x - x_0) + F'_y(x_0, y_0)(y - y_0) + o\bigl(|x - x_0| + |y - y_0|\bigr). \end{aligned}$$

Assuming for simplicity $(x_0, y_0) = (0, 0)$ and considering that we are only moving along the curve $y = f(x)$, we obtain

$$0 = F'_x(0, 0)x + F'_y(0, 0)y + o\bigl(|x| + |y|\bigr),$$

or

$$y = -\big(F'_y(0,0)\big)^{-1} F'_x(0,0)x - \big(F'_y(0,0)\big)^{-1} o\big(|x| + |y|\big). \tag{F.15}$$

Since $y = f(x) = f(x) - f(0)$, the formula (F.14) will be justified, if we can show that in the limit $x \to 0$, the second term on the right-hand side of (F.15) is $o(x)$.

But

$$\big|\big(F'_y(0,0)\big)^{-1} o\big(|x|+|y|\big)\big| \le \big\|\big(F'_y(0,0)\big)^{-1}\big\| \cdot \big|o\big(|x|+|y|\big)\big| = o\big(|x|+|y|\big).$$

Further,

$$\big\|\big(F'_y(0,0)\big)^{-1} F'_x(0,0)\big\| \le \big\|\big(F'_y(0,0)\big)^{-1}\big\| \cdot \big\|F'_x(0,0)\big\| = a < \infty.$$

Therefore, from (F.15) we obtain that $|y| \le a|x| + \alpha(|x|+|y|)$, where $y = f(x) \to 0$ and $\alpha = \alpha(x) \to 0$ for $x \to 0$. Hence,

$$|y| \le \frac{a+\alpha}{1-\alpha}|x| < 2a|x|$$

for x sufficiently close to 0. Given this, for $x \to 0$ we obtain from (F.15) that

$$f(x) = -\big(F'_y(0,0)\big)^{-1} F'_x(0,0)x + o(x).$$

In view of $f(0) = 0$, we have (F.14). □

F.3.5 Continuous Differentiability of an Implicit Function

If in addition to the conditions of the theorem, we know that the functions F'_x and F'_y are defined and continuous in some neighborhood of the point (x_0, y_0), then the implicit function f is also continuously differentiable in some neighborhood of the point x_0.

In short, if $F \in C^{(1)}$, then f is also in $C^{(1)}$.

Proof In this case, the conditions of differentiability of f and (F.14) are fulfilled not only at (x_0, y_0) but at all points of the "curve" $F(x, y) = 0$ near (x_0, y_0) (see the above cautionary "Warning!"). Then, according to formula (F.14), in a neighborhood of the point x_0,

$$f'(x) = -\big(F'_y\big(x, f(x)\big)\big)^{-1} F'_x\big(x, f(x)\big), \tag{F.14$'$}$$

from which it is clear that f' is continuous. □

Warning! Recall that the mapping $A \mapsto A^{-1}$ is continuous.

F.3.6 Higher Derivatives of an Implicit Function

If in addition to the conditions of the theorem, we know that the function F is of class $C^{(k)}$ in some neighborhood of the point (x_0, y_0), then the implicit function f is also of class $C^{(k)}$ in a neighborhood of the point x_0.

Proof Suppose, for example, that F is of class $C^{(2)}$. Since f is of class $C^{(1)}$, the right-hand side of the equality (F.14′) can be differentiated according to the differentiation rule for composite functions (chain rule). We obtain then a formula for $f''(x)$, and from it the continuity of $f''(x)$ follows.

Moreover, as on the right-hand side of formula (F.14′), for $f'(x)$ the first partial derivatives of F and the function f itself (but not f') are involved; in the formula for $f''(x)$, the second partial derivatives of F, f, and f' are involved (but not f'').

Thus, if F is of class $C^{(3)}$, then we can differentiate $f''(x)$, and we arrive again at a formula for $f'''(x)$ in which are involved the third partial derivatives of F, and also the derivatives of the functions f (f, f', f'') of order less than three.

By induction, we obtain what we claimed. □

Warning! Recall that the mapping $A \mapsto A^{-1}$ is differentiable and even infinitely differentiable.

Problem 4

a) Find $f''(x)$ (write down the formula for the computation of $f''(x)(h_1, h_2)$ for given displacement vectors h_1, h_2).

b) What does the formula (simplified) for $f''(x)$ look like in the case that x, y, and $z = F(x, y)$ are real or complex variables?

Problem 5 (Method of undetermined coefficients). Suppose that we know the first (or all) coefficients of the Taylor series of the function F. Find the first (or all) coefficients of the Taylor series of the implicit function f.

Problem 6

a) Write in coordinate form the formulation of the implicit function theorem for the cases $F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, when $m = n = 1$ and when $n > 1$.

b) Let $F : \mathbb{R}^m \to \mathbb{R}^n$ $(m > n)$ be a linear mapping with maximal rank $(= n)$. What is the dimension of the subspace $F^{-1}(0) \subset \mathbb{R}^m$ and what is its codimension? Let $F : \mathbb{R}^m \to \mathbb{R}^n$ $(m > n)$ be now an arbitrary smooth mapping, $F(0) = 0$ and $\operatorname{rank} F'(x) = n$. Answer the same questions ($\dim F^{-1}(0) = ?$, $\operatorname{codim} F^{-1}(0) = ?$) with respect to the set $F^{-1}(0)$.

References

1 Classic Works

1.1 Primary Sources

Newton, I.:

– a. (1687): Philosophiæ Naturalis Principia Mathematica. Jussu Societatis Regiæ ac typis Josephi Streati, London. English translation from the 3rd edition (1726): University of California Press, Berkeley, CA (1999).
– b. (1967–1981): The Mathematical Papers of Isaac Newton, D.T. Whiteside, ed., Cambridge University Press.

Leibniz, G.W. (1971): Mathematische Schriften. C.I. Gerhardt, ed., G. Olms, Hildesheim.

1.2 Major Comprehensive Expository Works

Euler, L.

– a. (1748): Introductio in Analysin Infinitorum. M.M. Bousquet, Lausanne. English translation: Springer-Verlag, Berlin – Heidelberg – New York (1988–1990).
– b. (1755): Institutiones Calculi Differentialis. Impensis Academiæ Imperialis Scientiarum, Petropoli. English translation: Springer, Berlin – Heidelberg – New York (2000).
– c. (1768–1770): Institutionum Calculi Integralis. Impensis Academiæ Imperialis Scientiarum, Petropoli.

Cauchy, A.-L.

– a. (1989): Analyse Algébrique. Jacques Gabay, Sceaux.
– b. (1840–1844): Leçons de Calcul Différential et de Calcul Intégral. Bachelier, Paris.

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

1.3 Classical Courses of Analysis from the First Half of the Twentieth Century

Courant, R. (1988): Differential and Integral Calculus. Translated from the German. Vol. 1 reprint of the second edition 1937. Vol. 2 reprint of the 1936 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York.

de la Vallée Poussin, Ch.-J. (1954, 1957): Cours d'Analyse Infinitésimale. (Tome 1 11 éd., Tome 2 9 éd., revue et augmentée avec la collaboration de Fernand Simonart.) Librairie universitaire, Louvain. English translation of an earlier edition: Dover Publications, New York (1946).

Goursat, É. (1992): Cours d'Analyse Mathématiques. (Vol. 1 reprint of the 4th ed. 1924, Vol. 2 reprint of the 4th ed. 1925) Les Grands Classiques Gauthier-Villars. Jacques Gabay, Sceaux. English translation: Dover Publ. Inc., New York (1959).

2 Textbooks[1]

Apostol, T.M. (1974): Mathematical Analysis. 2nd ed. World Student Series Edition. Addison-Wesley Publishing Co., Reading, Mass. – London – Don Mills, Ont.

Courant, R., John F. (1999): Introduction to Calculus and Analysis. Vol. I. Reprint of the 1989 edition. Classics in Mathematics. Springer-Verlag, Berlin.

Courant, R., John F. (1989): Introduction to Calculus and Analysis. Vol. II. With the assistance of Albert A. Blank and Alan Solomon. Reprint of the 1974 edition. Springer-Verlag, New York.

Nikolskii, S.M. (1990): A Course of Mathematical Analysis. Vols. 1, 2. Nauka, Moscow. English translation of an earlier addition: Mir, Moscow (1985).

Rudin, W. (1976): Principals of Mathematical Analysis. McGraw-Hill, New York.

Rudin, W. (1987): Real and Complex Analysis. 3rd ed., McGraw-Hill, New York.

Spivak, M. (1980): Calculus. 2nd ed. Publish Perish, Inc., Berkeley, CA.

Stewart, J. (1991): Calculus. 2nd ed. Thomson Information/Publishing Group. International Student Edition. Brooks/Cole Publishing Company, Pacific Grove, CA.

3 Classroom Materials

Biler, P., Witkowski, A. (1990): Problems in Mathematical Analysis. Monographs and Textbooks in Pure and Applied Mathematics, 132. Marcel Dekker, New York.

[1] For the convenience of the Western reader the bibliography of the English edition has substantially been revised.

Bressoud, D.M. (1994): A Radical Approach to Real Analysis. Classroom Resource Material Series, 2. Mathematical Association of America, Washington, DC.
Demidovich, B.P. (1990): A Collection of Problems and Exercises in Mathematical Analysis. Nauka, Moscow. English translation of an earlier edition: Gordon and Breach, New York – London – Paris (1969).
Gelbaum, B. (1982): Problems in Analysis. Problem Books in Mathematics. Springer-Verlag, New York – Berlin.
Gelbaum, B., Olmsted, J. (1964): Counterexamples in Analysis. Holden-Day, San Francisco.
Hahn, A.J. (1998): Basic Calculus. From Archimedes to Newton to its role in science. Textbooks in Mathematical Sciences. Springer, New York, NY.
Makarov, B.M., Goluzina, M.G., Lodkin, A.A., Podkorytov, A.N. (1992): Selected Problems in Real Analysis. Nauka, Moscow. English translation: Translations of Mathematical Monographs 107, American Mathematical Society, Providence (1992).
Pólya, G., Szegő, G. (1970/1971): Aufgaben und Lehrsätze aus der Analysis. Springer-Verlag, Berlin – Heidelberg – New York. English translation: Springer-Verlag, Berlin – Heidelberg – New York (1972–1976).

4 Further Reading

Arnol'd, V.I.

- a. (1989a): Huygens and Barrow, Newton and Hooke: Pioneers in Mathematical Analysis and Catastrophe Theory, from Evolvents to Quasicrystals. Nauka, Moscow. English translation: Birkhäuser, Boston (1990).
- b. (1989b): Mathematical Methods of Classical Mechanics. Nauka, Moscow. English translation: Springer-Verlag, Berlin – Heidelberg – New York (1997).

Avez, A. (1986): Differential Calculus. Translated from the French. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester.
Bourbaki, N. (1969): Éléments d'Histoire des Mathématiques. 2e édition revue, corrigée, augmentée. Hermann, Paris. English translation: Springer-Verlag, Berlin – Heidelberg – New York (1994).
Cartan, H. (1977): Cours de Calcul Différentiel. Hermann, Paris. English translation of an earlier edition: Differential Calculus. Exercised by C. Buttin, F. Riedeau and J.L. Verley. Houghton Mifflin Co., Boston, Mass. (1971).
Dieudonné, J. (1969): Foundations of Modern Analysis. Enlarged and corrected printing. Academic Press, New York.
Dubrovin, B.A., Novikov, S.P., Fomenko, A.T. (1986): Modern Geometry – Methods and Applications. Nauka, Moscow. English translation: Springer-Verlag, Berlin – Heidelberg – New York (1992).
Einstein, A. (1982): Ideas and Opinions. Three Rivers Press, New York. Contains translations of the papers "Principles of Research" (original German title: "Motive des Forschens"), pp. 224–227, and "Physics and Reality," pp. 290–323.

Feynman, R., Leighton, R., Sands, M. (1963–1965): The Feynman Lectures on Physics, Vol. 1: Modern Natural Science. The Laws of Mechanics. Addison-Wesley, Reading, Mass.

Gel'fand, I.M. (1998): Lectures on Linear Algebra. Dobrosvet, Moscow. English translation of an earlier edition: Dover, New York (1989).

Halmos, P. (1974): Finite-Dimensional Vector Spaces. Springer-Verlag, Berlin – Heidelberg – New York.

Jost, J. (2003): Postmodern Analysis. 2nd ed. Universitext. Springer, Berlin.

Kirillov, A.A. (1993): What Is a Number? Nauka, Moscow (Russian).

Kolmogorov, A.N., Fomin, S.V. (1989): Elements of the Theory of Functions and Functional Analysis. 6th ed., revised, Nauka, Moscow. English translation of an earlier edition: Graylock Press, Rochester, New York (1957).

Kostrikin, A.I., Manin, Yu.I. (1986): Linear Algebra and Geometry. Nauka, Moscow. English translation: Gordon and Breach, New York (1989).

Landau, E. (1946): Grundlagen der Analysis. Chelsea, New York. English translation: Chelsea, New York (1951).

Lax, P.D., Burstein S.Z., Lax A. (1972): Calculus with Applications and Computing. Vol. I. Notes based on a course given at New York University. Courant Institute of Mathematical Sciences, New York University, New York.

Manin, Yu. I. (1979): Mathematics and Physics. Znanie, Moscow. English translation: Birkhäuser, Boston (1979).

Milnor, J. (1963): Morse Theory. Princeton University Press.

Poincaré, H. (1982): The Foundations of Science. Authorized translation by George Bruce Halstead, and an introduction by Josiah Royce, preface to the UPA edition by L. Pearce Williams. University Press of America, Washington, DC.

Pontryagin, L.S. (1974): Ordinary Differential Equations. Nauka, Moscow. English translation of an earlier edition: Addison-Wesley, Reading, Mass. (1962).

Spivak, M. (1965): Calculus on Manifolds: a Modern Approach to the Classical Theorems of Advanced Calculus. W.A. Benjamin, New York.

Uspenskii, V.A. (1987): What is Nonstandard Analysis? Nauka, Moscow (Russian).

Weyl, H. (1926): Die heutige Erkenntnislage in der Mathematik. Weltkreis-Verlag, Erlangen. Russian translations of eighteen of Weyl's essays, with an essay on Weyl: Mathematical Thought. Nauka, Moscow (1989).

Zel'dovich, Ya.B., Myshkis, A.D. (1967): Elements of Applied Mathematics. Nauka, Moscow. English translation: Mir, Moscow (1976).

Zorich, V.A. (2011): Mathematical Analysis of Problems in the Natural Sciences. Springer, Heidelberg.

Subject Index

A
Absolute value, 56
Absorption of radiation, 304
Acceleration, instantaneous, 171–184, 209, 262
Additivity of the integral, 350
Algebra, fundamental theorem, 281, 282
Algebraic closedness of the complex numbers, 280–282
Algebraic form of a complex number, 265
Algorithm, Euclidean, 66, 104
Almost everywhere, 345, 348, 359
Alternant, 170
Angle between curves, 540
Angle between vectors, 432
Area
  of a curvilinear trapezoid, 383
  of an ellipse, 384
Argument
  of a complex number, 265
  of a function, 12
Arithmetic mean, 548
Asymptote
  horizontal, 253
  oblique, 253
  vertical, 253
Asymptotic equivalence of functions, 140
Asymptotics of a function, 137, 141, 225
Average of a function, 371
Axiom
  of choice, 29
  of completeness, 38, 44, 51, 52, 55, 65, 67, 70, 71, 74
  of continuity, 38, 65
  of Dedekind, 65
  of extensionality, 27
  of infinity, 29
  of replacement, 29
  of separation, 27
  pairing, 28
  power set, 28
  union, 28
  Zermelo's, 29
Axiom system
  categorical, 38
  consistent, 38
  for set theory, 6, 27
  for the real numbers, 35, 52, 65, 69
  Zermelo–Fraenkel, 29
Axis
  coordinate, 55
  number, 55

B
Ball
  closed, 411
  open, 411
Base
  in the set of partitions, 333
  of computational system, 65
  of logarithm, 122
Base (filter base), 126, 128
Basis of a vector space, 427
  orthonormal, 431
Bernoulli numbers, 577
Bernoulli polynomials, 577
Bijection, 16
Binomial
  Newton's formula, 66, 207, 222
Binomial differential, 326
Bound of a set of numbers
  greatest lower, 45
  least upper, 44
  lower (minorant), 44

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

upper (majorant), 44
Branch of the argument of a complex number, 265

C
Canonical variables, 496
Cantor set, 78, 348
Cardinal number, 26
Cardinality of a set, 26, 75
Cardinality of the continuum, 76, 77
Cartesian coordinates, 431
Center of curvature, 263, 541
Center of mass, 584
Change of parameter, admissible, 382
Change of variable, 311, 366
Circle, osculating, 263
Closure of a set, 413, 415
Compact set, 160, 413
 in $\mathbb{R}^m$, 424
 in $\mathbb{R}^n$, 413–415
Complement of a set, 9
Complex conjugate, 264
Complex number, 264
 algebraic form, 265
 polar form, 265
Composition
 of mappings, 18, 132, 158, 194, 195, 347, 429, 441
 of relations, 22
Condition
 for a constrained extremum
  necessary, 530
  sufficient, 533
 for a monotonic function, 215, 234
 for an extremum, 235–237
  of a function of several variables, 463
 for convexity, 241–243
 for differentiability, 437, 456
 for integrability
  necessary, 335
  necessary and sufficient, 341, 342, 345
  sufficient, 336–340
 necessary, 2
 sufficient, 2
Constant
 Euler's, 146
 gravitational, 59
 Planck's, 59
Continuation
 of a function, 13
Continuum, 76
Contraction, 593
Contraction mapping, 569
Convergence
 absolute, 268
 necessary condition, 96
 of a sequence, 80
 of a series, 95
  absolute, 97, 98
 of an improper integral, 393
  absolute, 398
  conditional, 402
Convergence test
 Abel–Dirichlet, 403
 Cauchy's, 99
 integral, 399
 Weierstrass', 88, 98
Coordinate of a point, 54, 410
Coordinates
 Cartesian, 431
 curvilinear, 469, 470
  in $\mathbb{R}^m$, 515
 polar, 500, 501
 spherical, 501
Correspondence between functions and functionals, 587
Cosine
 circular, 273
 hyperbolic, 199, 274
Cosine integral, 328
Cotangent
 hyperbolic, 201
Criterion
 Cauchy
  for a function, 130, 418
  for an improper integral, 398
  for sequences, 85, 267, 417
  for series, 95, 268
 for a constant function, 216, 235
 for a constrained extremum
  necessary, 530
  sufficient, 533
 for a monotonic function, 136, 215, 234
 for an extremum, 235
  necessary, 235, 463
  sufficient, 237, 464
 for continuity of a monotonic function, 164
 for integrability
  du Bois-Reymond, 348
  Lebesgue, 344, 348
 for monotonic sequences, 87
 for series of nonnegative terms, 98
 integrability
  Darboux, 348
 necessary, 2
 Sylvester's, 466
Critical point, 464

index, 517
nondegenerate, 512, 516
Curvature, 262
Curve
level, 449, 482
parametrized, 378
simple closed, 378
unicursal, 326
Curvilinear coordinates, 469, 470
Cycloid, 392, 408

D
Decay, radioactive, 291–293
Definition of derivative, 589
Delta function, 584, 586
Dependence, functional, 508, 516
Derivative, 176–178, 277, 433, 561, 563
directional, 446
higher-order, 206
logarithmic, 196
of a function of a complex variable, 277
one-sided, 261
partial, 435
higher-order, 457
with respect to a vector, 444
Diameter of a set, 414, 417
Diffeomorphism, 498, 509
elementary, 509
Difference
finite, 233
of sets, 8
Differential equation, 287–302, 328, 561
of harmonic oscillations, 298–302
with variables separable, 328
Differential of a function, 176–183
of several variables, 433–437
Differential of a mapping, 433, 437
Differentiation
and arithmetic operations, 190–194, 277, 439
of a composite function, 194–197, 441
of a power series, 278
of an implicit function, 202–206
of an inverse function, 197–202, 447
Dimension
of a physical quantity, 452–454
of a surface, 517, 518
Directional derivative, 446
Discontinuity, removable, 155
Disk of convergence, 269
Distance
between sets, 415
in $\mathbb{R}^m$, 410
on the real line, 56
Divisor, 49
greatest common, 66
Domain
in $\mathbb{R}^m$, 423
of a function, 12
of a relation, 20

E
Efficiency, 302
Element
inverse, 37
maximal, 44
minimal, 44
negative, 36
neutral, 36, 37
of a set, 7, 8
zero, 36, 39
Energy
kinetic, 15, 304, 386
potential, 15, 304, 386, 388
total, 15, 304, 387
Equality
of functions, 12
of sets, 7
Equation
differential, 175, 287–302, 328, 561
Euler's (hydrodynamic), 446, 450
heat, 475
Laplace's, 474
Equations
Cauchy–Riemann, 515
Euler–Lagrange, 496
Hamilton, 496
Error
absolute, 59, 60, 80, 196
relative, 59, 60, 196, 450
Error function, 404
Escape velocity, 391
Euclidean structure, 432
Euler–Lagrange equations, 553
Euler–Lagrange system, 554
Euler–MacLaurin formula, 578–580
Euler's identity for homogeneous functions, 451
Euler's substitution, 322
Expansion
partial fraction, 283, 316
Taylor, 208, 220–230, 279
Explosion, 293
Exponent, 118
Exponential function, 187, 294–300
complex, 273, 299
Exponential integral, 328, 407
Extension of a function, 13

Extremum
constrained, 517, 526–539
interior, 212, 235
of a function of several variables, 463–465

F

Factorization of a polynomial, 282
Falling bodies, 293
Fenchel–Young inequality, 573
Fiber, 23
Fibonacci numbers, 105
Field
algebraic, 37
Archimedean, 68
ordered, 68
potential, 446, 449, 543
vector, 446
First mean-value theorem, 354
Fixed-point principle, 594
Force function, 449
Formula
barometric, 289–291, 302
Bonnet's, 358
Cauchy–Hadamard, 269
change of variable
in a definite integral, 366
de Moivre's, 266, 274
Euler's, 272
for change of variable in an indefinite integral, 311
for integration by parts, 310
in an improper integral, 397
Hermite interpolation, 232
integration by parts, 372
Lagrange interpolation, 232
Leibniz', 207
MacLaurin's, 219
Meshcherskii's, 289
Newton–Leibniz, 331, 366, 368, 370
Ostrogradskii's, 325
quadrature, 375
rectangular, 374
Simpson's (parabolic), 374
trapezoidal, 374
Taylor's, 217–230, 461–476
for functions of several variables, 461–476, 513
local, 225, 462
multi-index notation for, 476
with integral form of the remainder, 461, 476, 513
Tsiolkovskii's, 289
Viète's, 147
Fraction
continued, 104
convergents, 104
partial, 283, 316
Fractional part, 54
Function, 12–15, 416, 560
additive interval, 350, 386
analytic at a point, 223, 279
asymptotically equivalent to another, 140
bounded, 111, 416, 424
from above, 111
from below, 111
characteristic, 14
concave, 241
constant, 109
continuous, 421–424
at a point, 149–152
on a set, 152, 422
convex, 241–243
downward, 241
upward, 241
decreasing, 136
differentiable at a point, 175, 176
Dirichlet, 156, 346
exponential, 117–122, 187, 273, 274, 294–300
force, 449
harmonic, 474
homogeneous, 451
hyperbolic, 199
implicit, 206, 209, 480–490
increasing, 135
infinite, 138
of higher order, 138
infinitesimal, 112–114
compared with another, 137
of higher order, 137
integrable, 335
inverse, 16, 17, 163–166, 197, 446, 498
Lagrange, 529, 530, 535
locally homogeneous, 451
logarithmic, 122–125
monotonic, 135
nondecreasing, 135
nonincreasing, 135
of a complex variable, 275
continuous, 277
differentiable, 277
of several variables, 409
differentiable, 432
periodic, 189, 265, 274, 369, 372
power, 125, 126
Riemann, 156, 167, 346
sgn, 108

strictly convex, 241
trigonometric, 379–381
ultimately bounded, 111, 113, 129, 136, 416
ultimately constant, 111
uniformly continuous, 160, 424
Functional, 12, 14, 349
Functional dependence, 508, 516
Functionals as generalized functions, 588
Fundamental theorem of algebra, 281, 282

G
General Euler–MacLaurin formula, 579
General Young's inequality, 573
Generalized functions, 586
Geodesic, 14
Geometric series, 96
Germ of a function, 170
Gradient, 445, 446
Graph of a function, 19, 168, 241–263
of several variables, 469, 470
Group, 36
Abelian, 36, 49
additive, 36
commutative, 36
multiplicative, 37

H
Half-life, 291, 303
Hamilton's system of equations, 553
Harmonic mean, 548
Heaviside function, 586
Hessian, 495, 512, 541
Higher derivatives, 565
Hyperbolic cosine, 199
Hyperbolic cosine integral, 328
Hyperbolic sine, 199
Hyperbolic sine integral, 328

I
Ideal of a ring, 170
maximal, 170
of functions, 170
Identity
in a multiplicative group, 37
in the real numbers, 36, 37
Image, 16, 22
Imaginary part of a complex number, 264
Imaginary unit, 263
Imbedding, 16
Implicit function, 596
Implicit function theorem, 595
Increment
of a function, 176, 177, 433
of an argument, 176, 177, 433
Indefinite integral, 306–314
Index of a critical point, 517
Inequality
Bernoulli's, 66, 89, 238
Cauchy–Bunyakovskii, 360
Hadamard's, 543
Hölder's, 239, 248, 360
Jensen's, 247
Minkowski's, 240, 360, 410
Schwarz, 360
triangle, 57, 240, 410, 411, 430
Young's, 239, 261, 393
Inferior limit, 91
Injection, 16
Inner product, 431
Integer part, 54
Integral
cosine, 328
Darboux, 347
definite, 331–335
elliptic, 323, 324, 326, 327, 383, 401
complete, 383, 389, 407
first kind, 324, 389, 407
second kind, 324, 383
third kind, 324
Euler, 401
Euler–Poisson, 328, 404
exponential, 407
Fresnel, 328, 370, 407
Gaussian, 404
hyperbolic cosine, 328
hyperbolic sine, 328
hyperelliptic, 323
improper, 393–397
absolutely convergent, 398
conditionally convergent, 402
convergent, 393
divergent, 393
with more than one singularity, 404
indefinite, 306–314
logarithmic, 314, 328
of a vector-valued function, 340
Riemann, 334, 335
sine, 328
with variable upper limit, 360
Integration, 307, 331
by parts, 310, 364
in a definite integral, 372
by substitution (change of variable), 311, 366
Intersection of sets, 8, 11

Interval, 56
  closed, 56
  half-open, 56
  multidimensional, 414
  numerical, 55
  open, 56
  unbounded, 56
Isomorphism, 39, 146
Iteration, 32, 168

J
Jacobi matrix, 437, 440, 464
Jacobian, 437
  of the transition to polar coordinates, 501

L
Laplacian, 474, 480
Law, 560
  Bernoulli's, 450
  Clapeyron's (ideal gas), 290, 498
  Kepler's, 454
  Newton's, 171, 211, 294, 446, 449
  of addition of velocities, 202–206
  of refraction, 238, 543
  Ohm's, 25
  Snell's, 237, 543
Least upper bound, 44, 45
Legendre transform, 494, 495, 573
Lemma
  Bolzano–Weierstrass, 72, 90, 94
  Fermat's, 212
  finite covering, 71
  Hadamard's, 476, 512
  least upper bound, 53
  Morse's, 512
  on limit points, 72
  on nested compact sets, 415
Length
  of a curve, 14, 377–379
  of a path, 377–379
  of an ellipse, 383
  of an interval, 52
Level curve, 449, 482
Level of a function, 528, 542
Level set, 449
Level surface, 488
Lifetime, 303
Limit
  of a composite function, 132, 418
  of a function, 106–136, 416
  of a mapping, 416
  of a sequence, 79–82, 85
    inferior, 91
    partial, 93
    superior, 91
  over a base, 126–130
Limits of integration, 334, 350, 360
Linearity of the integral, 349
Logarithm, 122, 196, 286
  natural, 122, 286
Logarithmic integral, 314, 328, 406
Logarithmic scale, 196
Lorentz transformations, 14

M
Mantissa, 69
Mapping, 12, 13, 416
  bijective, 16, 24
  bounded, 111, 416, 424
  constant, 109
  continuous, 149–152, 421–424
  identity, 18
  injective, 16, 24
  inverse, 16, 19, 197, 446, 498, 500
    left, 24
    right, 24
  linear, 176, 180, 428, 429, 439, 441, 447, 503
  one-to-one, 16
  surjective, 16, 24
  tangent, 433, 463, 503, 522
  ultimately bounded, 111, 416
  uniformly continuous, 160, 424
Mass, critical, 293
Maximum, 44, 160, 424, 462
  constrained, 527, 537
  local, 211, 236, 237, 462–471, 543
Mean
  arithmetic, 105, 248, 260
  geometric, 248, 260
  harmonic, 96, 105, 260
  integral, 371
  of order p, 105, 259
  quadratic, 105, 260
Mean square, 548
Mesh of a partition, 333
Method
  dimension, 453
  Euler's, 297
  gradient, 446
  Lagrange multipliers, 529, 543
  of exhaustion, 332
  of least squares, 477
  of undetermined coefficients, 284, 297
  Ostrogradskii's, 324
Method of tangents, 570, 592
Metric, 410, 411, 424
  in $\mathbb{R}^m$, 410, 416

Minimum, 44, 160, 424, 462
  constrained, 527, 537
  local, 211, 236, 237, 462–471, 543
Modulus
  of a complex number, 264
  of a real number, 56
  of a vector, 174, 264
  of continuity, 168
Modulus of a spring, 298, 304, 386
Monotonicity of the integral, 353
Morphism, 12
Multi-index, 475
Multiplicity of a root, 283

N
Natural logarithm, 565
Necessary condition for convergence, 96
Neighborhood
  deleted, 107, 418
  of a point, 57, 72, 107, 411, 416
Newton's method, 570, 592
Node, interpolation, 232, 373
Norm of a vector, 174, 429–432
Normal vector, 471
Nuclear reactor, 293
Number, 560
  algebraic, 52, 67, 76
  complex, 264
  e, 89, 101–103, 122–134, 274, 299
  Fibonacci, 105
  integer, 49
  irrational, 52, 67, 76
  natural, 26, 29, 46
    von Neumann, 29, 32
  negative, 43
  π, 52, 274, 374, 381, 393
  positive, 43
  prime, 49
  rational, 50, 54, 75
  real, 35
  transcendental, 52, 67, 76
Number axis, 55

O
Operation
  addition, 36
  associative, 17, 36, 37
  commutative, 36, 37
  distributive, 37
  multiplication, 36
  of differentiation, 190
  on sets, 8, 11
Operator, 12
  Laplacian, 474, 480
  logical, 7, 30
  shift, 14
  translation, 14
Orbit, planetary, 305
Order
  linear, 38, 55, 68
  partial, 37, 68
Order of contact, 181
Orthogonal vectors, 431
Oscillation, 298–302
  damped, 301
  harmonic, 300, 302
  of a function
    at a point, 151, 422
    on a set, 130, 151, 417
  of a particle in a well, 408
  of a pendulum, 388, 389, 392, 402, 454
  on a set, 336
Oscillator
  linear, 304
  plane, 305
Osculating circle, 263

P
Pair
  ordered, 10, 28
  unordered, 10, 28
Parabolic mirror, 185
Parametrization of a curve, 378, 380
  natural, 387
Partial derivative, 435
Partial fraction, 283, 316
Partial limit, 93
Partition of an interval, 333
  with distinguished points, 333
Path, 377, 421, 423, 448, 470, 517, 540
  closed, 377
  piecewise smooth, 378
  simple, 378
  simple closed, 378
Pendulum, 388, 389, 402, 454
  cycloidal, 392, 408
Period
  of a function, 189, 274, 369
  of oscillation, 389, 392, 402, 408, 454
  of revolution, 305
$\pi(x)$, 136–141
Plane, 466, 469
  complex, 264, 275
  tangent, 470–474, 488, 522, 523
    to a surface, 488, 522, 523
Point
  boundary, 412
  Chebyshev alternant, 170

critical, 464, 521, 542
degenerate, 541
nondegenerate, 512, 516
saddle point, 471
exterior, 412
fixed, 167, 168
in $\mathbb{R}^m$, 410
interior, 412
limit, 72, 413
local maximum, 211, 236, 237, 462–471
local minimum, 211, 236, 237, 462–471
of discontinuity
of a monotonic function, 164
of first kind, 155
of second kind, 156
of inflection, 246, 494
stationary, 464, 543
Polar coordinates, 500, 501
Polar form of a complex number, 265
Polynomial
Chebyshev, 169
Hermite, 232
Lagrange, 232, 373
Legendre, 373
of best approximation, 169, 170
Taylor, 224–228
Potential
Newtonian, 391, 446
of a force, 386–388, 446
of a vector field, 446, 449, 543
Power series, 268–271
absolutely convergent, 269–271
convergent, 269
Pre-image, 16, 17, 23
Primitive, 306–314, 332, 366, 368
generalized, 362
of a rational function, 314–318
Principal value, 406
Principle
Archimedes', 52, 62, 74
Bolzano–Weierstrass, 72, 74
Borel–Lebesgue, 71, 74
Cauchy–Cantor, 71, 74
Fermat's, 238, 543
least upper bound, 67
of induction, 46, 47, 57, 66
Principle of the fixed point, 593
Problem
Buffon needle, 393
Huygens', 468, 478
Kepler's (two-body), 171
Okun's, 549
Procedure
recursive, 18
Product
Cartesian, 10, 28, 32
direct, 10, 28, 32
infinite, 147
inner, 431
of series, 270
Projection, 11, 14, 422
stereographic, 540
Property
global, 158, 424
holding ultimately (over a base), 129–137
local, 157, 170, 422

Q

Quantifier
existence, 7, 30
universal, 7, 30

R

Radioactive decay, 291–293
Radius
critical, 293
of convergence, 269
of curvature, 262
Range
of a function, 12
of a relation, 20
Rank
of a mapping, 503, 516
of a number, 65, 67
of a system of functions, 516
Rational part of an integral, 324, 325
Real part of a complex number, 264
Rearrangement of terms of a series, 97, 270
Rectification, 502
Recursion, 18
Refinement of a partition, 336
Regular, 588
Relation, 5, 20
antisymmetric, 21
equality, 7, 20
equipollence, 26
equivalence, 21, 26
functional, 21, 22
inclusion, 7, 21, 68
order, 21, 55
linear, 21
partial, 21
reflexive, 20
symmetric, 21
transitive, 21, 23
transpose, 23
Remainder in Taylor's formula, 218–226, 375
Cauchy form, 219, 366

integral form, 364, 461, 476
Lagrange form, 219–226, 366, 462, 476
Peano form, 226, 462
Resolution of a diffeomorphism, 509
Restriction of a function, 13
Riemann integral, 585
Riemann–Stieltjes integral, 585
Ring
of continuous functions, 170
of germs of continuous functions, 170
Root
multiplicity, 283
nth, 68, 118
of a complex number, 266
of a polynomial, 169, 279–282
multiple, 232, 283
Rule, l'Hôpital's, 248

S
Saddle point, 471
Secant, 180, 181
Second mean-value theorem, 355, 358
Sequence, 58, 71, 80
bounded, 81, 87
Cauchy, 85, 267, 417
constant, 81
convergent, 80
decreasing, 87
divergent, 80
fundamental, 85, 267, 417
increasing, 87
monotonic, 87
nested, 71, 72
of intervals, 72
nondecreasing, 87
nonincreasing, 87
numerical, 58
of closed intervals, 86
of elements of a set, 58
of nested compact sets, 415
ultimately constant, 81
Series, 95
absolutely convergent, 97
convergent, 95
absolutely, 97
divergent, 95
harmonic, 96
numerical, 95
power, 219, 268–271
Taylor, 220, 279
Set, 5, 26
bounded, 44, 415
from above, 44
from below, 44
Cantor, 348
cardinality, 75
closed, 411–415
connected, 423, 426
pathwise, 423
countable, 74, 75
empty, 8, 27
equipollent to another, 25
finite, 26
inductive, 29, 46, 66
infinite, 26
invariant, 25
level, 449
of integrable functions, 335, 342
of measure zero, 344–346
open, 411–414, 423, 426
stable, 25
unbounded, 418
uncountable, 76
Sine
circular, 107, 273
hyperbolic, 199, 274
Sine integral, 328
Singular, 588
Skew-gradient, 553
Space
configuration, 15
Euclidean, 432
metric, 410
complete, 131, 417
phase, 15
$\mathcal{R}[a, b]$, 335, 342, 344, 349
$\mathbb{R}^m$, 409, 419, 427
tangent, 433, 521–526
vector, 342, 343, 349, 427
Sphere, 412, 424, 425, 539
Spherical coordinates, 501
Stationary point, 464
Streamline, 450
Structure
Euclidean, 432
logical, 30
Subsequence, 90
Subset, 7, 28
empty, 8, 27
proper, 7
Substitution, Euler's, 322
Successor, 28
Sum
Darboux, 340, 347
of a series, 95
partial, 95
Riemann, 332, 341

lower, 340
upper, 340
Superior limit, 91
Support of a path, 378, 381
Surface, 470, 501, 517, 521, 523
level, 488
minimal, 495
Surjection, 16
Symbol
logical, 1
o, 137
O, 139
System of computation, 46, 61, 70
positional, 61, 65
System of functions
dependent, 508, 516
independent, 508, 516

T
Table
of derivatives, 203
of primitives (indefinite integrals), 309
Tangent, 173–182, 213, 244, 480
hyperbolic, 201
Tangent line, 180
Tangent mapping, 433, 463, 503
Tangent plane, 470–474, 488, 522, 523
to a surface, 488, 522, 523
Tangent space, 433, 521–526
Tangent vector, 472
Test
d'Alembert's, 100, 222
for extrema, 212
Gauss', 148
Test functions, 587
Theorem
Abel's, 269
Bolzano–Cauchy, 158
Cantor uniform-continuity, 162
Cantor–Heine, 162
Cantor's, 27
Cauchy's finite-increment, 216
Chebyshev's, 170
comparison
for integrals, 400
for series, 98
Darboux', 231, 347
Dedekind's, 65
finite-increment, 216, 217, 455
fundamental, of arithmetic, 67
Heine–Borel, 71
implicit function, 480–490
Lagrange's finite-increment, 213–217
Liouville's, 67
mean-value, 215, 454
first, 354, 372
second, 355, 358, 372, 403
of dimension theory (Π-theorem), 454
rank, 503
Rolle's, 213, 477
Schröder–Bernstein, 32
Thales', 55
Vallée Poussin's, 169
Weierstrass', 88
Weierstrass maximum-value, 160
Topology, 108
Transform
involutive, 261, 494
Legendre, 261, 494, 495
Transformation, 12
Abel's, 355
Galilean, 13, 25, 203, 204
linear, 428, 429, 439, 503
Lorentz, 14, 25, 205
Trapezoid, curvilinear, 383
Truth table, 4

U
Union of sets, 8, 11
Unit
multiplicative, 40
Unit, imaginary, 263

V
Value
of a function, 12, 22
average, 371
principal, 406
Variable, canonical, 496
Vector
normal, 471
tangent, 472
Vectors, orthogonal, 431
Velocity
escape, 391
instantaneous, 172–175, 192
of light, 13, 59, 205, 238
Volume of a solid of revolution, 384

W
Work, 385
escape, 391

Z
Zero divisor, 69

Name Index

A
Abel, N., 36, 269, 323, 324, 355, 403, 405
Archimedes ('Αρχιμήδης), 52, 53, 332

B
Banach, S., 595
Bernoulli, Daniel, 450
Bernoulli, Jacob, 66, 89, 577, 578
Bernoulli, Johann, 19, 238, 248
Bernstein, F., 26, 32
Bólyai, J., 19
Bolzano, B., 72, 85, 90, 158, 190
Bonnet, P., 358
Borel, É, 71
Bourbaki, N., 5, 128
Buffon, G., 393
Bunyakovskii, B., 360

C
Cantor, G., 5, 26, 71, 76, 85, 162, 348
Carnot, L., 114
Carnot, S., 114
Cartan, H., 128
Cauchy, A., 58, 71, 85–87, 91, 95–97, 99, 100, 106, 130, 131, 133, 158, 216–219, 221–223, 269, 286, 360, 406, 515
Chebyshev, P., 136, 169, 326
Clapeyron, B., 290, 498
Cohen, P., 77

D
d'Alembert, J., 100
Darboux, G., 231, 340, 347
de Moivre, A., 266
de Morgan, A., 9
Dedekind, R., 26, 65, 67
Descartes, R., 10
Dirichlet, P., 156, 403, 405
du Bois-Reymond, P., 348

E
Einstein, A., 14, 428
Euclid (Εὐκλείδης), 66, 77, 136
Euler, L., 89, 146, 271–274, 297, 305, 313, 322, 326, 328, 401, 404, 450, 451, 496, 553, 554, 577–579, 581

F
Fenchel, W., 573
Fermat, P., 10, 212, 235, 237, 543
Fibonacci (Leonardo of Pisa), 105
Fraenkel, A., 29
Fresnel, A., 328, 370, 407

G
Galilei, Galileo, 1, 13
Gauss, C., 19, 148, 156, 264, 282, 404, 477, 478
Gel'fond, A., 52
Germain, M., 478

H
Hadamard, J., 269, 476, 512, 543
Hamilton, W., 496, 553, 554
Heaviside, O., 586
Heine, E., 71, 110, 162
Hermite, Ch., 232
Hilbert, D., 52, 66
Hölder, O., 239, 248, 360
Hooke, R., 304, 305
Huygens, Ch., 392, 468, 478

J
Jacobi, C., 437
Jensen, J., 246, 360

V.A. Zorich, *Mathematical Analysis I*, Universitext,
DOI 10.1007/978-3-662-48792-1

K
Kelvin, Lord (W. Thomson), 290
Kepler, J., 171, 454

L
Lacroix, S., 19
Lagrange, J., 178, 213–217, 219, 220, 222, 225, 226, 243, 462, 496, 529, 543, 553, 554
Laplace, P., 272, 474
Lebesgue, H., 71, 344, 345, 348
Legendre, A., 261, 324, 373, 383, 389, 496, 573
Leibniz, G., 1, 19, 52, 171, 195, 207, 331, 363, 364, 366, 370
L'Hôpital, G., 248
Lindemann, F., 52
Liouville, J., 67, 323
Lobachevskii, N., 19
Lorentz, H., 14, 205, 206

M
MacLaurin, C., 219, 577–579, 581
Maxwell, J., 452
Meshcherskii, I., 289
Minkowski, H., 240, 360
Morse, M., 512, 516

N
Newton, I., 1, 66, 171, 222, 331, 363, 364, 366, 370, 391, 454, 570, 592, 595

O
Ohm, G., 25
Okun', L., 549
Ostrogradskii, M., 324

P
Peano, G., 26, 225, 226, 378, 462
Picard, É., 595
Poincaré, H., 1
Poisson, S., 328, 404

R
Riemann, B., 156, 332–348, 400, 515, 583, 585
Rolle, M., 213, 216, 477
Russell, B., 6, 28

S
Sakharov, A., 549
Schneider, T., 52
Schröder, E., 26, 32
Schwarz, H., 360
Simpson, T., 374, 375
Snell van Royen, W., 237, 543
Stieltjes, T.J., 583, 585
Stokes, G., 364
Sylvester, J., 466

T
Taylor, B., 217, 221, 226, 271, 364, 370, 461, 462, 464, 475
Thomson, W. (Lord Kelvin), 290
Tsiolkovskii, K., 289

V
Vallée Poussin, Ch., 169
van der Waerden, B., 190
Viète, F., 147
von Neumann, J., 29, 32

W
Weierstrass, K., 72, 88, 90, 99, 159, 190

Y
Young, W., 239, 393, 573

Z
Zermelo, E., 29