**EMS Series of Lectures in Mathematics**
Edited by Andrew Ranicki (University of Edinburgh, U.K.)

*EMS Series of Lectures in Mathematics* is a book series aimed at students, professional mathematicians and scientists. It publishes polished notes arising from seminars or lecture series in all fields of pure and applied mathematics, including the reissue of classic texts of continuing interest. The individual volumes are intended to give a rapid and accessible introduction into their particular subject, guiding the audience to topics of current research and the more advanced and specialized literature.

Previously published in this series:
Katrin Wehrheim, *Uhlenbeck Compactness*
Torsten Ekedahl, *One Semester of Elliptic Curves*
Sergey V. Matveev, *Lectures on Algebraic Topology*
Joseph C. Várilly, *An Introduction to Noncommutative Geometry*
Reto Müller, *Differential Harnack Inequalities and the Ricci Flow*
Eustasio del Barrio, Paul Deheuvels and Sara van de Geer, *Lectures on Empirical Processes*

Iskander A. Taimanov

# Lectures on Differential Geometry

Translated from the Russian by Gleb V. Dyatlov

European Mathematical Society

Author:

Iskander A. Taimanov
Mechanics and Mathematics Department
Chair of Geometry and Topology
Novosibirsk State University
2 Pirogova street
630090 Novosibirsk, Russia

and

Sobolev Institute of Mathematics
Siberian Branch of the Russian Academy of Sciences
4 Acad. Koptyug avenue
630090 Novosibirsk, Russia

E-mail: taimanov@math.nsc.ru

Translator:

Gleb V. Dyatlov
Sobolev Institute of Mathematics
Siberian Branch of the Russian Academy of Sciences
4 Acad. Koptyug avenue
630090 Novosibirsk, Russia

Originally published by Institute of Computer Studies, Moscow-Izhevsk 2002, 2006, under the title (Russian)
Лекции по дифференциальной геометрии

2000 Mathematical Subject Classification (primary; secondary): 53-01; 20C05, 20G20, 37J05, 37J35, 53A10, 53D05, 53D17, 57S17, 58A05

Key words: Differential geometry, Riemannian geometry, smooth manifolds, minimal surfaces, Lie groups, representation theory, symplectic and Poisson geometry, finite-dimensional integrable systems

ISBN 978-3-03719-050-0

The Swiss National Library lists this publication in The Swiss Book, the Swiss national bibliography, and the detailed bibliographic data are available on the Internet at http://www.helveticat.ch.



Contact address:
European Mathematical Society Publishing House
Seminar for Applied Mathematics
ETH-Zentrum FLI C4
CH-8092 Zürich
Switzerland
Phone: +41 (0)1 632 34 36
Email: info@ems-ph.org
Homepage: www.ems-ph.org

Typeset using the author's TEX files: I. Zimmermann, Freiburg
Printed in Germany

9 8 7 6 5 4 3 2 1

# Preface

Differential geometry studies geometrical objects using analytical methods. Like modern analysis itself, differential geometry originates in classical mechanics. For instance, geodesics and minimal surfaces are defined via variational principles and the curvature of a curve is easily interpreted as the acceleration with respect to the path length parameter.

Modern differential geometry in its turn strongly contributed to modern physics when, for instance, at the beginning of the 20th century it was discovered by Einstein that a gravitational field is just a pseudo-Riemannian metric on space time. The basic equations of gravity theory were written in terms of the curvature of a metric, which is a geometric quantity. More recently the modern theory of elementary particles was based on gauge fields, which mathematically are connections on fiber bundles.

In this book we attempt to give an introduction to the basics of differential geometry, keeping in mind the natural origin of many geometrical quantities, as well as the applications of differential geometry and its methods to other sciences.

The book is divided into three parts. The first part covers the basics of curves and surfaces, while the second part is designed as an introduction to smooth manifolds and Riemannian geometry. In particular, in Chapter 5 we give short introductions to hyperbolic geometry and geometrical principles of special relativity theory. Here, only a basic knowledge of algebra, calculus and ordinary differential equations is required.

The third part is more advanced and assumes that the reader is familiar with the first two parts of the book. It introduces the reader to Lie groups and Lie algebras, the representation theory of groups, symplectic and Poisson geometry, and the applications of complex analysis in surface theory. We use Lie groups as important examples of smooth manifolds and we expose symplectic and Poisson geometry in their close relation with mechanics and the theory of integrable systems.

This book is a translation (with minor revisions and corrections) of Лекции по дифференциальной геометрии,[6] which is based on lecture notes[7] that arose from a one-semester course given in Novosibirsk State University, and on some advanced minicourses.

The contents and the concise style of exposition are due to the expectation that the book will be suitable for a full semester course at the upper undergraduate or beginning graduate level. During the course at Novosibirsk State University we included a complete treatment of Chapters 1, 2, and 5. Chapters 3 and 4 give a much more detailed exposition of some further material which was touched upon in the course.

[6] Lectures on Differential Geometry, 2nd ed., Moscow-Izhevsk 2006.

[7] Lectures on differential geometry. I. Curves and surfaces. Novosibirsk: NSU, 1998; II. Riemannian geometry. Novosibirsk: NSU, 1998; III. Supplement chapters. Novosibirsk: NSU, 1999.

Note that, unless specified to the contrary, repeated upper and lower indices imply summation. For example

$$a^i_k v^k := \sum_{k=1}^{n} a^i_k v^k \quad \text{where } v \in V, \ n = \dim V.$$

Although this notation rarely appears in textbooks, it is widely used in scientific publications, especially in physical applications of differential geometry. It is one of the more practical ways in which the book will be useful for students with wider interests in such applications of differential geometry.

The Bibliography is shorter than the list that appeared in the Russian edition, which contains many publications unavailable in English translation. Hence we have added supplementary reading sources in English. For further bibliography we refer to [4].

Novosibirsk, February 2008 I. A. Taimanov

# Contents

# Part I

# Curves and surfaces

# 1

# Theory of curves

## 1.1 Basic notions of the theory of curves

Let $\mathbb{R}^n$ be the $n$-dimensional *Euclidean space* with coordinates $x^1, \ldots, x^n$. The *distance* $\rho(x, y)$ between points $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^n)$ is defined by the formula

$$\rho(x, y) = \sqrt{(x^1 - y^1)^2 + \cdots + (x^n - y^n)^2}.$$

For brevity, we denote an $n$-dimensional vector space over the field $\mathbb{R}$ also by $\mathbb{R}^n$, assuming that it is always clear from the context which object is meant. We use the notation $(v, w)$ for the standard inner product in the vector space $\mathbb{R}^n$:

$$(v, w) = v^1 w^1 + \cdots + v^n w^n. \tag{1.1}$$

By "smooth" we mean "differentiable as many times as required". In all assertions, the reader can easily reconstruct what smoothness we need or understand the word "smooth" as "infinitely differentiable".

By a parameterized *curve* in the Euclidean space $\mathbb{R}^n$ we mean a path given by functions of one-dimensional parameter:

$$\gamma(t) = (x^1(t), \ldots, x^n(t)).$$

Also, for brevity, we confine ourselves to the case when this parameter varies in an interval $[a, b]$ or on the whole real axis $\mathbb{R}$. Accepting this definition, by a *point* $P$ of the curve $\gamma$ we mean the image of the point together with the value of the parameter $t \in [a, b]$:

$$P = (\gamma(t) \in \mathbb{R}^n, \quad t \in [a, b]).$$

We restrict our exposition to studying regular curves:

- a curve $\gamma$ is *smooth* if such is the map $\gamma$.
- a smooth curve $\gamma$ is *regular* if the derivative of $\gamma$ with respect to the parameter $t$ is nonzero at all interior points of interval $[a, b]$:

$$\frac{d\gamma}{dt}(s) \neq 0 \quad \text{for } a < s < b$$

  and the derivative has nonzero unilateral limits at the endpoints $a$ and $b$.

For simplicity, in the definition of regular curves we exclude the possible existence of singular points at which $\frac{d\gamma}{dt} = 0$.

It is natural to identify curves obtained by traversing the same path with different velocities:

- regular curves

  $$\gamma_1 \colon [a_1, b_1] \to \mathbb{R}^n \quad \text{and} \quad \gamma_2 \colon [a_2, b_2] \to \mathbb{R}^n$$

  are *equivalent* if there is a map

  $$\varphi \colon [a_1, b_1] \to [a_2, b_2]$$

  such that $\varphi$ is invertible, the maps $\varphi$ and $\varphi^{-1}$ are smooth, and

  $$\gamma_1(t) = \gamma_2(\varphi(t)) \quad \text{for } t \in [a_1, b_1].$$

We will identify equivalent curves and consider $t$ and $s = \varphi(t)$ as different parameters on the same curve.

- The *length* of a (parameterized) curve $\gamma \colon [a, b] \to \mathbb{R}^n$, where $-\infty < a < b < +\infty$, is the quantity

  $$\mathrm{L}(\gamma) = \int_a^b \left| \frac{d\gamma(t)}{dt} \right| dt = \int_a^b \sqrt{\left(\frac{dx^1(t)}{dt}\right)^2 + \cdots + \left(\frac{dx^n(t)}{dt}\right)^2}\, dt, \quad (1.2)$$

  where $\gamma(t) = (x^1(t), \ldots, x^n(t))$.

We now have the following lemma.

**Lemma 1.1.** *The length of a curve is independent of the choice of the parameter on this curve.*

*Proof.* Suppose that $\gamma_1 \colon [a_1, b_1] \to \mathbb{R}^n$ and $\gamma_2 \colon [a_2, b_2] \to \mathbb{R}^n$ define the same curve and the parameters $t \in [a_1, b_1]$ and $s \in [a_2, b_2]$ are connected by a monotone map $s = s(t)$ such that the derivative $\frac{ds}{dt}$ is defined and positive everywhere. Then it follows from the chain and substitution rules that

$$\int_{a_1}^{b_1} \left| \frac{d\gamma_1}{dt} \right| dt = \int_{a_1}^{b_1} \left| \frac{d\gamma_2}{ds} \frac{ds}{dt} \right| dt = \int_{a_1}^{b_1} \left| \frac{d\gamma_2}{ds} \right| \frac{ds}{dt} dt = \int_{a_2}^{b_2} \left| \frac{d\gamma_2}{ds} \right| ds.$$

Lemma 1.1 is proven. □

We see that this definition of length satisfies the following natural requirements: the length of two successively traversed curves equals the sum of their lengths and the length of a line segment coincides with the distance between the endpoints.

The notion of length is connected with the notion of an *arc length parameter*, i.e., a parameter $l$ such that the length of the part of the curve corresponding to the change of $l$ from $a_1$ to $b_1 > a_1$ is equal to $(b_1 - a_1)$.

We now have the following lemma.

**Lemma 1.2.** (1) *If* $l \in [a,b]$ *is an arc length parameter on the curve* $\gamma\colon [a,b] \to \mathbb{R}^n$, *then*

$$\left|\frac{d\gamma}{dl}\right| = 1$$

*at smooth points.*

(2) *Each regular curve has an arc length parameter.*

*Proof.* Assertion (1) is immediate from the definition of the length of a curve.

Let $\gamma\colon [a,b] \to \mathbb{R}^n$ be a curve with a parameter $t$. Consider the differential equation

$$\frac{dl}{dt} = \left|\frac{d\gamma}{dt}\right|.$$

Since the right-hand side is a smooth function, this equation has a unique solution with the initial data $l(a) = 0$. Clearly, $l(b) = \mathrm{L}(\gamma)$. Take the inverse function $t = t(l)\colon [0, \mathrm{L}(\gamma)] \to [a,b]$, and define $l$ as the parameter on $\gamma$ by the formula

$$\gamma_0(l) = \gamma(t(l)).$$

It is obvious that

$$\left|\frac{d\gamma_0}{dl}\right| = \left|\frac{d\gamma}{dt}\right| \frac{dt}{dl} = 1.$$

Consequently, $l$ is an arc length parameter on $\gamma$. Lemma 1.2 is proven. □

## 1.2 Plane curves

Let $\gamma\colon [a,b] \to \mathbb{R}^2$ be a regular curve on the two-dimensional plane with an arc length parameter $l$. Assuming the plane to be oriented, at each point of the curve choose a basis $v, n$ for $\mathbb{R}^2$ such that

(1) $v = \frac{d\gamma}{dl}$, in particular, $|v| = 1$ since the parameter $l$ is arc length;

(2) the vector $n$ is a unit vector orthogonal to $v$ and the basis $(v, n)$ has positive orientation.

These conditions determine the frames $(v, n)$ uniquely. They are called the *Frenet frames*.

**Theorem 1.1.** *As the arc length parameter* $l$ *varies along a plane curve* $\gamma$, *the Frenet frame changes according to the equations*

$$\frac{d}{dl}\begin{pmatrix} v \\ n \end{pmatrix} = \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix} \begin{pmatrix} cv \\ n \end{pmatrix}. \tag{1.3}$$

*Proof.* Since $(v,v) \equiv (n,n) \equiv 1$, we have

$$\frac{d(v,v)}{dl} = 2\left(\frac{dv}{dl}, v\right) = 0, \quad \frac{d(n,n)}{dl} = 2\left(\frac{dn}{dl}, n\right) = 0.$$

Consequently, $v \perp \frac{dv}{dl}$ and $n \perp \frac{dn}{dl}$. Since $(v,n)$ is an orthonormal basis for $\mathbb{R}^2$, there are functions $\alpha(l)$ and $\beta(l)$ such that

$$\frac{dv}{dl} = \alpha n, \quad \frac{dn}{dl} = \beta v.$$

But $(v,n) \equiv 0$ and therefore

$$\frac{d(v,n)}{dl} = \left(\frac{dv}{dl}, n\right) + \left(v, \frac{dn}{dl}\right) = \alpha + \beta = 0.$$

Put $k = \alpha = -\beta$. Theorem 1.1 is proven. □

We arrive at some important notions:

- equations (1.3) are referred to as the *Frenet equations* (or *Frenet formulas*) for a plane curve;
- the coefficient $k$ in (1.3) is called the *curvature of a* (*plane*) *curve*;
- the *radius of curvature* of a curve is the value $R = |k|^{-1}$.

**Problem 1.1.** If the curvature $k$ of a plane curve $\gamma$ is constant and nonzero, then $\gamma$ is an arc of a circle of radius $R = |k|^{-1}$. If the curvature $k$ of a plane curve $\gamma$ is zero everywhere, then $\gamma$ is a line segment.

To solve this problem, it suffices to find solutions to equations (1.3) with constant coefficients.

The curvature of a plane curve determines the curve up to motion of the plane $\mathbb{R}^2$.

**Theorem 1.2.** (1) *Let $k\colon [0,L] \to \mathbb{R}$ be a smooth function. Then there is a smooth curve $\gamma\colon [0,L] \to \mathbb{R}^2$ whose curvature equals $k(l)$.*

(2) *Suppose that $\gamma_1\colon [0,L] \to \mathbb{R}^2$ and $\gamma_2\colon [0,L] \to \mathbb{R}^2$ are arc length parameterized regular curves and their curvatures coincide: $k_1(l) = k_2(l)$ for all $l \in [0,L]$. Then there is an orientation-preserving motion $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $\gamma_2(l) = \varphi(\gamma_1(l))$ for all $l \in [0,L]$.*

*Proof.* (1) Choose a positively oriented orthonormal basis $(v_0, n_0)$ for $\mathbb{R}^2$ and consider a solution to equations (1.3) with the initial conditions $v(0) = v_0$ and $n(0) = n_0$. This is an ordinary differential equation in $\mathbb{R}^4$ with a smooth right-hand side. Therefore, it has a unique solution. The inner products of the vectors $v$ and $n$ satisfy the system of equations

$$\frac{d(v,v)}{dl} = -2k(v,n), \quad \frac{d(n,n)}{dl} = 2k(v,n), \quad \frac{d(v,n)}{dl} = k(n,n) - k(v,v),$$

which, being supplemented with the initial data $(v_0, v_0) = (n_0, n_0) = 1$ and $(v_0, n_0) = 0$, has a unique solution; note that this solution is constant. Therefore, for every value of $l$ the vectors $v(l)$ and $n(l)$ constitute an orthonormal basis for $\mathbb{R}^2$. Now, define the curve $\gamma$ by the formula

$$\gamma(l) = \int_0^l v(s)\, ds.$$

It is easy to note that $l$ is an arc length parameter on the curve and $v$ is the velocity vector with respect to this parameter; moreover, since $(v, n)$ is an orthonormal basis and $\frac{dv}{dl} = kn$, $k$ is the curvature of $\gamma$.

(2) Recall that the group of orientation-preserving motions of the Euclidean space $\mathbb{R}^2$ is generated by translations $T_a \colon x \to r + a$ and rotations $\Omega \colon \mathbb{R}^2 \to \mathbb{R}^2$ around a fixed point.

Denote the Frenet frames of the curves $\gamma_1$ and $\gamma_2$ by $(v_1, n_1)$ and $(v_2, n_2)$. Define the motion $\varphi$ as the composition of the translation $T_a$ and a rotation $\Omega \colon \mathbb{R}^2 \to \mathbb{R}^2$ around the point $\gamma_2(0)$:

$$\varphi = \Omega \cdot T_a \colon \mathbb{R}^2 \to \mathbb{R}^2,$$

where

$$a = \gamma_2(0) - \gamma_1(0), \quad \Omega v_1(0) = v_2(0).$$

The Frenet frame of $\varphi(\gamma_1(l))$ has the form $(\Omega v_1, \Omega n_1)$. Since $(v_1, n_1)$ and $(v_2, n_2)$ both satisfy equations (1.3) and

$$\begin{aligned} \frac{d}{dl}\begin{pmatrix} \Omega v_1 \\ \Omega n_1 \end{pmatrix} &= \begin{pmatrix} \Omega & 0 \\ 0 & \Omega \end{pmatrix}\begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ n_1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix}\begin{pmatrix} \Omega & 0 \\ 0 & \Omega \end{pmatrix}\begin{pmatrix} v_1 \\ n_1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix}\begin{pmatrix} \Omega v_1 \\ \Omega n_1 \end{pmatrix}, \end{aligned}$$

$(\Omega v_1, \Omega n_1)$ satisfies the same equation. Further, $(\Omega v_1(0), \Omega n_1(0)) = (v_2(0), n_2(0))$ by the choice of $\Omega$. Now, uniqueness of the solution to equations (1.3) with the given initial data gives the equality

$$(\Omega v_1(l), \Omega n_1(l)) \equiv (v_2(l), n_2(l)).$$

It follows from the choice of the translation $T_a$ that

$$\gamma_2(l) = \gamma_2(0) + \int_0^l v_2(t)\, dt = \varphi(\gamma_1(l)).$$

Theorem 1.2 is proven. □

The equation $k = k(l)$ is called the *natural equation* of a plane curve. By the previous theorem, it determines the curve up to motion of the plane.

## 1.3 Curves in three-dimensional space

The *curvature* of a regular space curve is the value

$$k = \left|\frac{d^2\gamma}{dl^2}\right|, \tag{1.4}$$

where $l$ is the arc length parameter.

If we consider a curve $\gamma(l)$ as the trajectory of a point mass which moves along a path with unit velocity (i.e., $\left|\frac{d\gamma}{dl}\right| \equiv 1$) then the curvature of the curve is the modulus of the acceleration vector.

Note that, with this definition, the curvature is always nonnegative unlike the case of plane curves.

In the case of space curves there are infinitely many ways to define the normal to a curve. Therefore, henceforth we consider only biregular curves:

- a regular arc length parameterized curve $\gamma\colon [a,b] \to \mathbb{R}^3$ is *biregular* if $\frac{d^2\gamma}{dl^2} \neq 0$ everywhere;
- the *normal* to a biregular curve is the vector

$$n = \frac{\frac{d^2\gamma}{dl^2}}{\left|\frac{d^2\gamma}{dl^2}\right|}.$$

To obtain the *Frenet frame* of a curve in $\mathbb{R}^3$, we complement the vectors $v = \frac{d\gamma}{dl}$ and $n$ to an orthonormal basis for $\mathbb{R}^3$ with the third vector, the *binormal*

$$b = [v \times n].$$

**Theorem 1.3.** *Let $\gamma$ be a biregular curve in $\mathbb{R}^3$. As the arc length parameter $l$ varies, the Frenet frame changes according to the Frenet equations for a space curve*

$$\frac{d}{dl}\begin{pmatrix} v \\ n \\ b \end{pmatrix} = \begin{pmatrix} 0 & k & 0 \\ -k & 0 & \varkappa \\ 0 & -\varkappa & 0 \end{pmatrix}\begin{pmatrix} v \\ n \\ b \end{pmatrix}. \tag{1.5}$$

The *proof* is similar to that of Theorem 1.1. Since $(v(l), n(l), b(l))$ is an orthonormal basis for each value of $l$, the lengths of these vectors are preserved and consequently the derivative of each of them is orthogonal to the vector itself; therefore,

$$\frac{d}{dl}\begin{pmatrix} v \\ n \\ b \end{pmatrix} = \begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{pmatrix}\begin{pmatrix} v \\ n \\ b \end{pmatrix}.$$

We are left with noting that

$$\frac{d(v,n)}{dl} = a_{12} + a_{21} = 0, \quad \frac{d(n,b)}{dl} = a_{23} + a_{32} = 0, \quad \frac{d(v,b)}{dl} = a_{13} + a_{31} = 0.$$

From the definition of $n$ we obtain $a_{12} = k$ and $a_{13} = 0$. Putting $\varkappa = a_{23}$, we arrive at equations (1.5). Theorem 1.3 is proven.

The value $\varkappa$ in (1.5) is called the *torsion* of the curve $\gamma$. The origin of this term is explained in the following problem.

**Problem 1.2.** Let $k > 0$ and $\varkappa$ be constants. Then the helix $\gamma(l) = (R\cos(\lambda l), R\sin(\lambda l), \mu l)$ with $\lambda = \sqrt{k^2 + \varkappa^2}$, $a = \varkappa/\lambda$, and $R = k/\lambda^2$ is parameterized by the arc length, its curvature is everywhere equal to $k$, and its torsion is identically equal to $\varkappa$.

**Problem 1.3.** Suppose that $\gamma(t)$ is a smooth curve with some (not necessarily arc length) parameter $t$. Prove the following:

(1) If the curve $\gamma(t)$ is regular, i.e., $\gamma' \neq 0$, then its curvature is equal to

$$k = \frac{|\gamma' \times \gamma''|}{|\gamma'|^3};$$

in particular, the curvature of a plane curve given by the equations $\gamma(t) = (x(t), y(t))$ is equal to

$$k = \frac{|x''y' - x'y''|}{(x'^2 + y'^2)^{3/2}}$$

(here $x$, $y$, and $z$ are the Euclidean coordinates in $\mathbb{R}^3$).

(2) The curve $\gamma(t)$ is biregular if and only if $k \neq 0$.

(3) If the curve $\gamma(t)$ is biregular then its torsion is equal to

$$\varkappa = \frac{([\gamma' \times \gamma''], \gamma''')}{|[\gamma' \times \gamma'']|^2}.$$

There is also an analog of Theorem 1.2:

**Theorem 1.4.** (1) *Suppose that* $k\colon [0, L] \to \mathbb{R}$ *and* $\varkappa\colon [0, L] \to \mathbb{R}$ *are smooth functions and the function* $k$ *is positive. Then there is a smooth curve* $\gamma\colon [0, L] \to \mathbb{R}^3$ *with curvature* $k(l)$ *and torsion* $\varkappa(l)$.

(2) *Suppose that* $\gamma_1\colon [0, L] \to \mathbb{R}^3$ *and* $\gamma_2\colon [0, L] \to \mathbb{R}^3$ *are arc length parameterized biregular curves and their curvatures and torsions coincide:* $k_1(l) = k_2(l)$ *and* $\varkappa_1(l) = \varkappa_2(l)$ *for all* $l \in [0, L]$. *Then there is an orientation-preserving motion* $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^3$ *such that* $\gamma_2(l) = \varphi(\gamma_1(l))$ *for all* $l \in [0, L]$.

The proof of Theorem 1.4 is similar to that of Theorem 1.2.

The equations $k = k(l)$ and $\varkappa = \varkappa(l)$ are called the *natural equations* for a space curve. It follows from Theorem 1.4 that they determine the curve up to motion of the three-dimensional space.

## 1.4 The orthogonal group

The Frenet equations (1.3) and (1.5) were derived from the fact that the inner products of the vectors which constitute the Frenet frames are preserved. We now make a brief digression devoted to the orthogonal group and derivation of these equations from the general viewpoint.

A linear transformation $A\colon \mathbb{R}^n \to \mathbb{R}^n$ of the vector space $\mathbb{R}^n$ is called *orthogonal* if it preserves the inner product (1.1). The following lemma is obvious.

**Lemma 1.3.** *Orthogonal transformations constitute a group with respect to the composition operation:* $A \cdot B(v) = A(B(v))$.

The group formed by all orthogonal transformations is called the *orthogonal group* and is denoted by $\mathrm{O}(n)$.

We identify the set of linear transformations $A\colon \mathbb{R}^n \to \mathbb{R}^n$ with the set of $(n \times n)$-matrices as follows: let $e_1, \dots, e_n$ be an orthonormal basis for $\mathbb{R}^n$, then every matrix $A = (a_{jk})_{1 \le j,k \le n}$ is uniquely associated with the transformation acting on the basis vectors by the formula

$$A\, e_k = \sum_{j=1}^{n} a_{jk}\, e_j .$$

**Lemma 1.4.** *A linear transformation is orthogonal if and only if its matrix A satisfies the equation*

$$A^\top A = 1_n, \tag{1.6}$$

*where* $A^\top$ *is the transpose of the matrix* $A$ ($a^\top_{jk} = a_{kj}$) *and* $1_n$ *is the identity* $(n \times n)$-*matrix.*

*Proof.* Write down the inner product (1.1) as the product of matrices (here we consider vectors as $(n \times 1)$-matrices):

$$(v, w) = v^\top 1_n w.$$

Then

$$(Av, Aw) = (Av)^\top 1_n (Aw) = v^\top A^\top 1_n A w = v^\top (A^\top A) w$$

and $A$ is orthogonal if and only if

$$v^\top (A^\top A) w = v^\top 1_n w \quad \text{for all } v, w \in \mathbb{R}^n,$$

which is equivalent to (1.6). Lemma 1.4 is proven. □

Considering the entries $a_{jk}$ as coordinates in the $n^2$-dimensional space, we identify the space of $(n \times n)$-matrices with the $n^2$-dimensional Euclidean space $\mathbb{R}^{n^2}$. The orthogonal group is distinguished in this space by $n(n+1)/2$ polynomial equations (1.6):

$$F_{jk}(a_{11},\dots,a_{nn}) - \delta_{jk} = \sum_{m=1}^{n} a_{mj}a_{mk} - \delta_{jk} = 0, \quad 1 \le j \le k \le n. \tag{1.7}$$

Now, we recall the implicit function theorem:

**The Implicit Function Theorem.** *Let $F\colon U \times V \to \mathbb{R}^n$ be a smooth map from the direct product of domains $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^k$ with coordinates $x$ and $y$ to $\mathbb{R}^n$. Suppose that $F(x_0, y_0) = 0$ and the matrix*

$$J = \left(\frac{\partial F^j}{\partial x^m}\right)_{1\le j,m\le n} \tag{1.8}$$

*is invertible at $(x_0, y_0) \in U \times V$. Then there is a neighborhood $W \subset U \times V$ of the point $(x_0, y_0)$ and a neighborhood $V_0 \subset V$ of $y_0 \in \mathbb{R}^k$ such that*

(1) *smooth functions $\psi^1,\dots,\psi^n$ are defined in $V_0$;*

(2) *$F(x,y) = 0$ for $(x,y) \in W$ if and only if $x^1 = \psi^1(y),\dots,x^n = \psi^n(y)$.*

The matrix

$$\left(\frac{\partial F^j(x_0)}{\partial x^m}\right)$$

is called the *Jacobian matrix* (or the *Jacobian*) of $F$ at $x_0$.

It follows from the implicit function theorem that in a neighborhood of the point $(x_0, y_0)$ the zero set of $F$ looks like the graph of a map $V_0 \subset \mathbb{R}^k \to \mathbb{R}^n$; therefore, the points of this set can be parameterized in a smooth fashion by points of a domain in $\mathbb{R}^k$.

**Theorem 1.5.** *Let $M$ be a set of points in $\mathbb{R}^{n+k}$. Then the following conditions are equivalent:*

(1) *In a sufficiently small neighborhood of each point, $M$ is the graph of a smooth map*

$$\begin{aligned} x^1 &= \psi^1(x^{n+1},\dots,x^{n+k}),\\ &\vdots\\ x^n &= \psi^n(x^{n+1},\dots,x^{n+k}) \end{aligned} \tag{1.9}$$

(*with an appropriate numbering of the coordinates $x^1,\dots,x^{n+k}$*).

(2) *In a sufficiently small neighborhood of each point, $M$ is the zero set of a smooth map $F\colon W \subset \mathbb{R}^{n+k} \to \mathbb{R}^n$ such that the matrix* (1.8) *is invertible in this neighborhood* (*with an appropriate numbering of the coordinates $x^1,\dots,x^{n+k}$*).

*Proof.* It follows from the implicit function theorem that condition (2) implies condition (1). The converse is also valid: define $F$ by the formulas $F^1(x) = x^1 - \psi^1(x^{n+1}, \dots, x^{n+k}), \dots, F^n(x) = x^n - \psi^n(x^{n+1}, \dots, x^{n+k})$. Theorem 1.5 is proven. □

A subset $M$ which satisfies either of two equivalent conditions from Theorem 1.5 is called a $k$-dimensional *smooth submanifold* (without boundary) in $\mathbb{R}^{n+k}$.

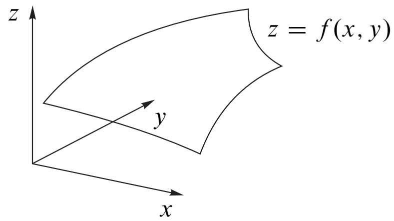


Figure 1.1. The graph of a function of two variables.

**Example.** A *sphere* of dimension $n$ is given in $\mathbb{R}^{n+1}$ by the equation

$$(x^1)^2 + \cdots + (x^{n+1})^2 = 1.$$

It cannot be defined globally as the graph of a function of $n$ variables.

Suppose that, in a neighborhood of a point $x \in M$, the submanifold $M$ is defined as the graph of a map (1.9). Then $y^1 = x^{n+1}, \dots, y^k = x^{n+k}$ determine *local coordinates* in a neighborhood of $x$ and the function $f$ on $M$ is called *smooth at* $x$ if it is smooth as a function of the local coordinates: $f = f(y^1, \dots, y^k)$. Similarly, we introduce the notion of smoothness for other objects on submanifolds, in particular, for vector fields.

Let $M$ be a $k$-dimensional submanifold in $\mathbb{R}^{n+k}$ and take $x \in M$. Consider the set of all smooth curves on $M$ passing through $x$ and lying in its neighborhood given by (1.9). With each curve $\gamma$ we associate its velocity vector at $x$: $v = \frac{d\gamma(t_0)}{dt}$, where $\gamma(t_0) = x$. Such vectors constitute the *tangent space* $T_x M$ of $M$ at $x$.

**Lemma 1.5.** *Let $M$ be a $k$-dimensional submanifold in $\mathbb{R}^{n+k}$. Then the tangent space at every point of $M$ is a $k$-dimensional vector space.*

*Proof.* Every curve $\gamma$ on $M$ passing through $x$, is the lifting of a curve $\gamma_0 \colon [a, b] \to \mathbb{R}^k$ by means of (1.9). Therefore, the velocity vector $\frac{d\gamma}{dt}$ has the form

$$\frac{d\gamma}{dt} = B\left(\frac{d\gamma_0}{dt}\right) = \left(\Psi_*\left(\frac{d\gamma_0}{dt}\right), \left(\frac{d\gamma_0}{dt}\right)\right),$$

where $\Psi_*$ is the differential of the map $\Psi = (\psi^1, \dots, \psi^n)$ of the form (1.9). The linear map $B$ has full rank $k$ and therefore is an isomorphism onto its image: it establishes an isomorphism between the vector spaces $\mathbb{R}^k$ and $T_x M$. Lemma 1.5 is proven. □

There is also another equivalent definition of the tangent space:

**Problem 1.4.** If a submanifold $M \subset \mathbb{R}^{n+k}$ is defined in a neighborhood $W$ of a point $x \in M$ as the zero set of a map $F\colon W \to \mathbb{R}^n$, then the tangent space of $M$ at $x$ coincides with the kernel of the linear map

$$F_*\colon \mathbb{R}^{n+k} \to \mathbb{R}^n, \quad F_*(v) = \lim_{\varepsilon \to 0} \frac{F(x + \varepsilon v) - F(x)}{\varepsilon}.$$

Important examples of smooth submanifolds are the orthogonal groups $\mathrm{O}(n)$:

**Theorem 1.6.** *The orthogonal group* $\mathrm{O}(n)$ *is an* $\frac{n(n-1)}{2}$*-dimensional smooth submanifold of the* $n^2$*-dimensional Euclidean space constituted by* $(n \times n)$*-matrices. Moreover, the tangent space* $T_1\mathrm{O}(n)$ *at the identity of the group consists of all antisymmetric matrices.*

*Proof.* The identity transformation given by the identity $(n \times n)$-matrix $1_n$ is the identity of the group $\mathrm{O}(n)$. In a neighborhood of $1_n \in \mathrm{O}(n)$ divide the variables $a_{jk}$ into two groups: $x$ corresponds to $j \le k$ and $y$ corresponds to $j > k$. Consider the map

$$F\colon \mathbb{R}^{n^2} \to \mathbb{R}^{n(n+1)/2}$$

given by (1.7). We can easily verify that the following equalities hold at $1_n \in \mathbb{R}^{n^2}$ for $j \le k, r \le s$:

$$\frac{\partial F_{jk}}{\partial a_{rs}} = \begin{cases} 2 & \text{for } j = r \text{ and } k = s, \\ 1 & \text{for } j = r \text{ and } k = s \text{ with } j \ne k, \\ 0 & \text{otherwise.} \end{cases}$$

Consequently, the matrix $\left(\frac{\partial F^j}{\partial x^m}\right)$ is invertible in a neighborhood of $1_n \in \mathbb{R}^{n^2}$. Since $\mathrm{O}(n)$ is defined by the equations $F = 0$, by the implicit function theorem, $\mathrm{O}(n)$ is a smooth submanifold in a neighborhood of $1_n$.

If $A \in \mathrm{O}(n)$ then define the map $F_A$ by the formula $F_A(X) = F(X \cdot A^{-1})$. Since $\mathrm{O}(n)$ is a group, $\mathrm{O}(n)$ is also defined by the equations $F_A = 0$. It follows from the definition of $F_A$ that the rank of the Jacobian matrix of $F_A$ at $A$ coincides with the rank of the Jacobian matrix of $F$ at $1_n$ and therefore is equal to $\frac{n(n+1)}{2}$. Hence, we conclude that $\mathrm{O}(n)$ is a smooth submanifold in a neighborhood of every point.

The dimension of the submanifold $\mathrm{O}(n)$ equals $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

Let $\gamma$ be a smooth curve in $\mathrm{O}(n)$ passing through $1_n$. We can assume that

$$\gamma(\varepsilon) = 1_n + X\varepsilon + O(\varepsilon^2),$$

where $X \in T_1\mathrm{O}$ is a tangent matrix. It follows from (1.6) that

$$\gamma(\varepsilon)^\top \gamma(\varepsilon) = (1_n + (X^\top + X)\varepsilon + O(\varepsilon^2)) \equiv 1_n.$$

Hence, $X$ is an antisymmetric matrix:

$$X^\top + X = 0.$$

Since the dimension of the space of antisymmetric matrices coincides with the dimension of $\mathrm{O}(n)$, the tangent space of $\mathrm{O}(n)$ at the identity coincides with the space of antisymmetric matrices. Theorem 1.6 is proven. □

Now, we return to the Frenet equations. Denote by $R(l)$ the Frenet frame corresponding to the parameter $l$ on a curve $\gamma$. The passage from $R(0)$ to $R(l)$ is given by an orthogonal transformation $A(l)$, since all these frames are orthonormal. It means that we have a smooth curve $A(l)$ in $\mathrm{O}(n)$ and the columns of $A(l)$ determine the decompositions of the vectors of $R(l)$ in the basis $R(0)$. The Frenet equation has the form

$$\frac{dA(l)}{dl} = \lim_{\varepsilon \to 0} \frac{A(l+\varepsilon) \cdot A^{-1}(l) - 1_n}{\varepsilon} \cdot A(l) = B(l) \cdot A(l),$$

where the matrix $B(l)$ is antisymmetric, since it is a tangent vector at the identity of the group $\mathrm{O}(n)$.

**Problem 1.5.** Prove that for every $(n \times n)$-matrix $A$ the series

$$\exp A = 1_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n + \cdots$$

converges. (Hint: the series for the exponential function $\exp \|A\|$ of the norm converges.)

**Problem 1.6.** Let $A$ be an antisymmetric $(n \times n)$-matrix. Then $\gamma(t) = \exp(At)$ is a smooth curve in $\mathrm{O}(n)$.

# 2

# Theory of surfaces

## 2.1 Metrics on regular surfaces

A *regular surface* in $\mathbb{R}^3$ is a two-dimensional smooth submanifold in $\mathbb{R}^3$.

Using the implicit function theorem, we can make this definition more specific. A set of points $\Sigma \subset \mathbb{R}^3$ is called a regular surface if one of the following conditions is satisfied:

(1) In a sufficiently small neighborhood of each point, it represents the zero set of a smooth function $F(x^1, x^2, x^3)$ such that

$$\frac{\partial F}{\partial x^3} \neq 0$$

in this neighborhood with an appropriate numbering of the coordinates.

(2) In a sufficiently small neighborhood of each point, it is the graph of a smooth map

$$x^3 = \psi(x^1, x^2)$$

with an appropriate numbering of the coordinates.

(3) In a sufficiently small neighborhood of each point, it is the range of a smooth map

$$\boldsymbol{r} : U \subset \mathbb{R}^2 \to \mathbb{R}^3,$$

where $U$ is a domain in the $(u^1, u^2)$-plane; moreover, the vectors $\boldsymbol{r}_1 = \partial \boldsymbol{r}/\partial u^1$ and $\boldsymbol{r}_2 = \partial \boldsymbol{r}/\partial u^2$ are linearly independent at all points of this domain.

**Lemma 2.1.** *Definitions* (1)–(3) *of a regular surface are equivalent.*

*Proof.* The equivalence of definitions (1) and (2) was already established in §1.4 in the most general form, i.e., for smooth submanifolds in $\mathbb{R}^n$. To prove the lemma, it suffices to establish equivalence of definitions (2) and (3).

It is obvious that (2) implies (3): it suffices to consider the map $\boldsymbol{r}(u^1, u^2) = (u^1, u^2, \psi(u^1, u^2))$. To prove that (3) implies (2), apply the following consequence of the implicit function theorem:

**The Inverse Function Theorem.** *Suppose that $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map of a domain in $\mathbb{R}^n$ to $\mathbb{R}^n$ and the Jacobian matrix of this map is invertible at a point $x_0 \in U$:*

$$\det \left( \frac{\partial F^j}{\partial x^k} \right)_{1 \leq j,k \leq N} \neq 0.$$

*Then, in some neighborhood $V$ of $F(x_0)$, a smooth map $G\colon V \to \mathbb{R}^n$ is defined which is the inverse of $F$: $G \cdot F(x) \equiv x$ in the domain $G(V) \subset U$.*

To prove this theorem, it suffices to apply the inverse function theorem to the map $\Psi\colon U \times \mathbb{R}^n \to \mathbb{R}^n$ of the form $\Psi(x, y) = y - F(x)$.

Now, show that (3) implies (2). Take $x \in \Sigma$. Without loss of generality we may assume that the matrix

$$\begin{pmatrix} \partial x^1/\partial u^1 & \partial x^1/\partial u^2 \\ \partial x^2/\partial u^1 & \partial x^2/\partial u^2 \end{pmatrix}$$

at $x$ is invertible and therefore, by the inverse function theorem, in a neighborhood of $x$ the parameters $u^1$ and $u^2$ are expressed uniquely as functions of $x^1$ and $x^2$: the map $(u^1, u^2) \to (x^1(u^1, u^2), x^2(u^1, u^2))$ is invertible. Now, note that, in a neighborhood of $x$, the surface is the graph of the smooth map

$$x^3 = x^3(u^1(x^1, x^2), u^2(x^1, x^2)).$$

Lemma 2.1 is proven. □

Although all three definitions are equivalent, henceforth we use the third one, since it introduces the notion of *local coordinates* $(u^1, u^2) \in U \subset \mathbb{R}^2$, i.e., quantities which determine coordinates on the surface near each point.

This definition is convenient in consideration of curves on a surface. In this case there is no need to define a curve in $\mathbb{R}^3$ and then impose the analytic conditions which would guarantee that it lies on the surface; on the contrary it suffices to define a curve

$$\gamma\colon [a, b] \to U \subset \mathbb{R}^2$$

as a smooth map into a domain $U$. If a curve does not lie in a neighborhood with the same local coordinates, then it is given as a family of maps to such domains. These maps are glued in the overlapping domains described by different local coordinates so that the curve and its velocity vectors (which are different in different coordinate systems) are defined correctly. Definition of this gluing requires introduction of tensors; therefore, we postpone it until §3.3, restricting our consideration until then to parts of surfaces covered by one coordinate system.

**Theorem 2.1.** *Let a smooth map $\boldsymbol{r}: U \to \mathbb{R}^3$ define a regular surface $\Sigma$ with local coordinates $(u^1, u^2) \in U$. Then*

(1) *at each point $\boldsymbol{r}(u^1, u^2)$, the vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ determine a basis for the tangent plane of $\Sigma$;*

(2) *the length of a regular curve $\gamma\colon [a, b] \to U$ on the surface is equal to*

$$\mathrm{L}(\gamma) = \int_a^b \sqrt{\boldsymbol{I}(\dot\gamma, \dot\gamma)}\, dt,$$

*where*

$$I(v, w) = \begin{pmatrix} v^1 & v^2 \end{pmatrix} \cdot \begin{pmatrix} E & F \\ F & G \end{pmatrix} \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix}$$

*is the symmetric bilinear form on the vector space* $\mathbb{R}^2$ *depending on the point of the surface and defined by the formulas*

$$E = (\boldsymbol{r}_1, \boldsymbol{r}_1), \quad F = (\boldsymbol{r}_1, \boldsymbol{r}_2), \quad G = (\boldsymbol{r}_2, \boldsymbol{r}_2),$$

*and* $\dot{\gamma} = \frac{d\gamma}{dt} = (\dot{u}^1, \dot{u}^2)$ *is the velocity vector in local coordinates.*

*Proof.* Assertion (1) follows from the definition of the tangent space. Assertion (2) is derived from (1.2). Indeed, by definition,

$$\mathrm{L}(\gamma) = \int_a^b \sqrt{\left(\frac{d(\boldsymbol{r} \cdot \gamma)}{dt}, \frac{d(\boldsymbol{r} \cdot \gamma)}{dt}\right)} dt$$

and, by straightforward calculations, we obtain

$$\left(\frac{d(\boldsymbol{r} \cdot \gamma)}{dt}, \frac{d(\boldsymbol{r} \cdot \gamma)}{dt}\right) = (\boldsymbol{r}_1\dot{u}^1 + \boldsymbol{r}_2\dot{u}^2, \boldsymbol{r}_1\dot{u}^1 + \boldsymbol{r}_2\dot{u}^2) = E(\dot{u}^1)^2 + 2F\dot{u}^1\dot{u}^2 + G(\dot{u}^2)^2.$$

Theorem 2.1 is proven. □

By Theorem 2.1, the tangent space at each point possesses the inner product defined by the *first fundamental form* $\boldsymbol{I}(v, w)$. For each pair of vectors $v$ and $w$ it determines the angle $\varphi_{v,w}$ between them, and for every vector $v$, its length $|v|$:

$$|v| = \sqrt{\boldsymbol{I}(v, v)}, \quad \cos\varphi_{v,w} = \frac{\boldsymbol{I}(v, w)}{|v||w|} \quad \text{for } v \neq 0 \text{ and } w \neq 0.$$

The following definition of the area of a part of a surface, as well as the definition of the length, is quite natural, additive, and coincides with the usual definition for plane surfaces:

- if $V \subset U$, then the *area* of $r(V) \subset \Sigma$ is equal to

$$\iint_V \sqrt{EG - F^2}\, du^1\, du^2.$$

## 2.2 Curvature of a curve on a surface

Denote by $\boldsymbol{n}(u^1, u^2)$ the normal to a regular surface at a point $\boldsymbol{r}(u^1, u^2)$ such that $(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{n})$ is a positively-oriented frame for $\mathbb{R}^3$. It can be found explicitly by the formula

$$\boldsymbol{n} = \frac{[\boldsymbol{r}_1 \times \boldsymbol{r}_2]}{|[\boldsymbol{r}_1 \times \boldsymbol{r}_2]|}.$$

Denote by $\boldsymbol{r}_{jk}$ the vector $\frac{\partial^2 \boldsymbol{r}}{\partial u^j \partial u^k}$.

Let

$$\gamma\colon [a,b] \to U \subset \mathbb{R}^2 \xrightarrow{\boldsymbol{r}} \mathbb{R}^3$$

be a regular curve with parameter $t$ and let $l$ be the arc length parameter on $\gamma$. By (1.4),

$$\frac{d^2(\boldsymbol{r}\cdot\gamma)}{dl^2} = k\, n_\gamma,$$

where $k$ is the curvature of the curve and $n_\gamma$ is the normal to $\gamma$. Also, from the proof of Lemma 1.2 we know that $\frac{dl}{dt} = \left|\frac{d(\boldsymbol{r}\cdot\gamma)(t)}{dt}\right| = \sqrt{\boldsymbol{I}(\dot\gamma,\dot\gamma)}$.

**Theorem 2.2.** *The curvature of a curve $\gamma$ on a surface satisfies the equation*

$$k\, \cos\theta = k(\boldsymbol{n}, n_\gamma) = \frac{\boldsymbol{II}(\dot\gamma,\dot\gamma)}{\boldsymbol{I}(\dot\gamma,\dot\gamma)},$$

*where $\theta$ is the angle between the normal $n_\gamma$ to $\gamma$ and the normal $\boldsymbol{n}$ to the surface and the form $\boldsymbol{II}$ is defined as*

$$\boldsymbol{II}(v,w) = \begin{pmatrix} v^1 & v^2 \end{pmatrix} \cdot \begin{pmatrix} L & M \\ M & N \end{pmatrix} \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix},$$

*where*

$$L = (\boldsymbol{r}_{11}, \boldsymbol{n}), \quad M = (\boldsymbol{r}_{12}, \boldsymbol{n}), \quad N = (\boldsymbol{r}_{22}, \boldsymbol{n}).$$

*Proof.* By straightforward computations, we obtain

$$\begin{aligned}
\frac{d^2(\boldsymbol{r}\cdot\gamma)}{dl^2} &= k\, n_\gamma \\
&= \frac{d}{dl}\left(\boldsymbol{r}_1 \dot u^1 \frac{dt}{dl} + \boldsymbol{r}_2 \dot u^2 \frac{dt}{dl}\right) \\
&= (\boldsymbol{r}_{11}(\dot u^1)^2 + 2\boldsymbol{r}_{12}\dot u^1 \dot u^2 + \boldsymbol{r}_{22}(\dot u^2)^2 + \boldsymbol{r}_1 \ddot u^1 + \boldsymbol{r}_2 \ddot u^2)\left(\frac{dt}{dl}\right)^2 \\
&\quad + (\boldsymbol{r}_1 \dot u^1 + \boldsymbol{r}_2 \dot u^2)\frac{d^2 t}{dl^2}
\end{aligned}$$

and, since $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are orthogonal to $\boldsymbol{n}$, we derive

$$k(\boldsymbol{n}, n_\gamma) = \big((\boldsymbol{r}_{11}, \boldsymbol{n})(\dot u^1)^2 + 2(\boldsymbol{r}_{12}, \boldsymbol{n})\dot u^1 \dot u^2 + (\boldsymbol{r}_{22}, \boldsymbol{n})(\dot u^2)^2\big)\left(\frac{dt}{dl}\right)^2.$$

We are left with noting that $\left(\frac{dt}{dl}\right)^2 = \left(\frac{dl}{dt}\right)^{-2} = \boldsymbol{I}(\dot\gamma,\dot\gamma)^{-1}$. Theorem 2.2 is proven. □

The form $\boldsymbol{II}(v, w)$ is called the *second fundamental form* of the surface.

Suppose that $v$ is a tangent vector to a surface at a point $\boldsymbol{r}(u^1, u^2)$ and $\boldsymbol{n}$ is the normal vector to the surface at the same point. Consider the two-dimensional plane passing through the point $\boldsymbol{r}(u^1, u^2)$ and spanned by the vectors $\boldsymbol{n}$ and $v$. The intersection of this plane and the surface is the curve $\gamma_{u^1,u^2,v}$ called the *normal section* corresponding to the point $\boldsymbol{r}(u^1, u^2)$ and the tangent vector $v$. From Theorem 2.2 we obtain

**Corollary 2.1.** *The curvature of the normal section $\gamma_{u^1,u^2,v}$ is equal to*

$$k = \pm \frac{\boldsymbol{II}(v, v)}{\boldsymbol{I}(v, v)},$$

*where the plus sign is taken if the normals to the surface and to the normal section coincide, and the minus sign, otherwise.*

**Problem 2.1.** Suppose that a surface is given as the graph of a function:

$$x^1 = u^1, \quad x^2 = u^2, \quad x^3 = f(u^1, u^2).$$

Then the first fundamental form is calculated as follows:

$$E = 1 + f_1^2, \quad F = f_1 f_2, \quad G = 1 + f_2^2,$$

the normal to the surface is given by the formula

$$\boldsymbol{n} = \frac{1}{\sqrt{1 + f_1^2 + f_2^2}}(-f_1, -f_2, 1),$$

and the second fundamental form is calculated as follows:

$$L = \frac{f_{11}}{\sqrt{1 + f_1^2 + f_2^2}}, \quad M = \frac{f_{12}}{\sqrt{1 + f_1^2 + f_2^2}}, \quad N = \frac{f_{22}}{\sqrt{1 + f_1^2 + f_2^2}},$$

where $f_{j_1 \dots j_k}$ denote the derivatives $\frac{\partial^k f}{\partial u^{j_1} \dots \partial u^{j_k}}$.

**Problem 2.2.** Given a function $f(x)$, construct the surface of revolution (obtained by revolution of the graph of this function around the $Ox$-axis):

$$r(u, v) = (u, f(u) \cos v, f(u) \sin v).$$

Show that the first fundamental form of this surface is calculated as follows:

$$E = 1 + f'^2, \quad F = 0, \quad G = f^2,$$

the normal to the surface is given by the formula

$$\boldsymbol{n} = \frac{1}{\sqrt{1 + f'^2}}(f', \cos v, \sin v),$$

and the second fundamental form is calculated as follows:

$$L = -\frac{f''}{\sqrt{1 + f'^2}}, \quad M = 0, \quad N = \frac{f}{\sqrt{1 + f'^2}}.$$

## 2.3 Gaussian curvature

The ratio $\frac{\boldsymbol{II}(v,v)}{\boldsymbol{I}(v,v)}$ is called the *normal curvature* of the surface in the direction $v$.

**Lemma 2.2.** *Suppose that two symmetric bilinear forms $\boldsymbol{I}$ and $\boldsymbol{II}$ are given on the vector space $\mathbb{R}^n$ and the form $\boldsymbol{I}$ is positive definite. Then there is a basis $e_1, \dots, e_n$ for $\mathbb{R}^n$ in which these forms become*

$$\boldsymbol{I}(v,w) = v^1 w^1 + \cdots + v^n w^n, \quad \boldsymbol{II}(v,w) = \lambda_1 v^1 w^1 + \cdots + \lambda_n v^n w^n$$

(*i.e., the form $\boldsymbol{I}$ is given by the identity matrix and the form $\boldsymbol{II}$ is diagonal*).

The proof of this lemma, known from a linear algebra course, is simple. Consider the form $\boldsymbol{I}$ as an inner product on $\mathbb{R}^n$ and, using the Gram–Schmidt orthogonalization, find an orthonormal basis for $\mathbb{R}^n$. Now, consider $\boldsymbol{II}$ as a symmetric bilinear form on $\mathbb{R}^n$ with the inner product $\boldsymbol{I}$. It is well known that every symmetric form becomes diagonal in some orthonormal basis. Lemma 2.2 is proven.

Now, denote by $T_x\Sigma$ the tangent plane of a surface $\Sigma$ at a point $x$. Suppose that the fundamental forms $\boldsymbol{I}$ and $\boldsymbol{II}$ are given on $T_x\Sigma$. Apply Lemma 2.2 in this situation to obtain the following result:

**Lemma 2.3.** *There is a basis $e_1, e_2$ for $T_x\Sigma$ in which both forms $\boldsymbol{I}$ and $\boldsymbol{II}$ are diagonal:*

$$\boldsymbol{I}(v,w) = v^1 w^1 + v^2 w^2,$$
$$\boldsymbol{II}(v,w) = k_1 v^1 w^1 + k_2 v^2 w^2.$$

The directions of the vectors $e_1$ and $e_2$ are called the *principal directions* and they are determined uniquely if $k_1 \neq k_2$. The values $k_1$ and $k_2$ of the normal curvatures along the principal directions are called the *principal curvatures*. They are the extremal values for the normal curvatures at the given point, which is a consequence of the following (now obvious) assertion.

**Theorem 2.3** (Euler's formula).

$$\frac{\boldsymbol{II}(v,v)}{\boldsymbol{I}(v,v)} = k_1 \cos^2\varphi + k_2 \sin^2\varphi,$$

*where $\varphi$ is the angle between the vectors $e_1$ and $v$.*

The principal curvatures are independent of the choice of the basis for the tangent space and are invariants of the pair of fundamental forms $\boldsymbol{I}$ and $\boldsymbol{II}$:

**Lemma 2.4.** *Suppose that the fundamental forms of a surface are determined by symmetric $(2\times 2)$-matrices $\boldsymbol{I}$ and $\boldsymbol{II}$. Then the principal curvatures $k_1$ and $k_2$ are the roots of the equation*

$$P(\lambda) = \det(\boldsymbol{II} - \lambda \boldsymbol{I}) = 0.$$

*Proof.* Assume that in a basis $e_1, e_2$ a quadratic form $J$ is given by a symmetric matrix denoted by the same letter $J$,

$$J(v,w) = \begin{pmatrix} v^1 & v^2 \end{pmatrix} \cdot J \cdot \begin{pmatrix} w^1 \\ w^2 \end{pmatrix},$$

and the coordinates $\hat{v}$ in a new basis $\hat{e}_1, \hat{e}_2$ are connected with the coordinates in the former basis by the relation $v = A\hat{v}$. Then

$$J(v,w) = \begin{pmatrix} \hat{v}^1 & \hat{v}^2 \end{pmatrix} \cdot A^\top \cdot J \cdot A \cdot \begin{pmatrix} c\widehat{w}^1 \\ \widehat{w}^2 \end{pmatrix} = \hat{J}(\hat{v}, \widehat{w}).$$

Hence, in the latter basis the fundamental form $J$ is given by the matrix $A^\top JA$.

We know that, in some basis, the principal curvatures $k_1$ and $k_2$ are the roots of the equation

$$\det(\boldsymbol{II} - \lambda \boldsymbol{I}) = 0.$$

But since

$$\det(A^\top \boldsymbol{II} A - \lambda A^\top \boldsymbol{I} A) = \det A^\top \cdot \det A \cdot \det(\boldsymbol{II} - \lambda \boldsymbol{I})$$

and $\det A^\top = \det A \neq 0$, in every other basis the corresponding equation has the same roots. Lemma 2.4 is proven. □

It follows from Lemma 2.4 that the product and the sum of the principal curvatures at a point of a surface are geometric invariants:

- the product of the principal curvatures at a point is called the *Gaussian curvature* of the surface at this point:

$$K = k_1 k_2 = \frac{\det \boldsymbol{II}}{\det \boldsymbol{I}} = \frac{LN - M}{EG - F^2};$$

- the half-sum of the principal curvatures at a point is called the *mean curvature* of the surface at this point:

$$H = \frac{k_1 + k_2}{2}.$$

**Problem 2.3.** Under the conditions of Problem 2.1, the Gaussian curvature has the form

$$K = \frac{f_{11} f_{22} - f_{12}^2}{(1 + f_1^2 + f_2^2)^2}.$$

**Problem 2.4.** For a surface of revolution (see Problem 2.2) the principal curvatures have the form

$$k_1 = -\frac{f''}{(1 + f'^2)^{3/2}}, \quad k_2 = \frac{1}{f\sqrt{1 + f'^2}};$$

the Gaussian curvature is equal to

$$K = -\frac{f''}{f(1+f'^2)^2};$$

and the mean curvature is equal to

$$H = \frac{1+f'^2 - ff''}{2f(1+f'^2)^{3/2}}.$$

**Problem 2.5.** The Gaussian curvature of a hyperbolic paraboloid $z = xy$ is negative everywhere and is equal to

$$K = -\frac{1}{(1+x^2+y^2)^2}.$$

The Gaussian curvature of the round sphere $x^2 + y^2 + z^2 = R^2$ of radius $R$ is everywhere positive and is equal to $K = R^{-2}$.

Problem 2.5 explains the geometric meaning of the sign of the Gaussian curvature:

(a) Let $K > 0$ at $x \in \Sigma$. Then a small neighborhood of $x$ lies to one side from the tangent plane $T_x M$. Indeed, the plane $T_x M$ separates the space into two half-spaces and the normals of all normal sections are directed to the same half-space; the surface looks like a "cap" near the point $x$.

(b) Let $K < 0$ at $x \in \Sigma$. Then the curvatures of two normal sections are zero (the tangent directions of these sections are called *asymptotic*), and the normals of the sections corresponding to the principal curvatures are directed to different half-spaces with respect to the plane $T_x M$; the surface looks like a "saddle" near $x$.

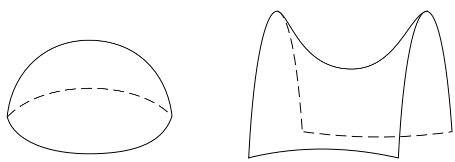

Figure 2.1. A cap and a saddle.

## 2.4 Derivational equations and Bonnet's theorem

An analog of the Frenet formulas for surfaces is the equations which describe the deformation of the basis $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{n}$ along the surface. The deformation of the vectors $\boldsymbol{r}_1$

and $\boldsymbol{r}_2$ is described by the *Gauss equations* which express the vectors $\boldsymbol{r}_{jk} = \frac{\partial^2 \boldsymbol{r}}{\partial u^j \partial u^k}$ in terms of the basis vectors:

$$\begin{aligned} \boldsymbol{r}_{11} &= \Gamma^1_{11}\boldsymbol{r}_1 + \Gamma^2_{11}\boldsymbol{r}_2 + b_{11}\boldsymbol{n}, \\ \boldsymbol{r}_{12} &= \Gamma^1_{12}\boldsymbol{r}_1 + \Gamma^2_{12}\boldsymbol{r}_2 + b_{12}\boldsymbol{n}, \\ \boldsymbol{r}_{22} &= \Gamma^1_{22}\boldsymbol{r}_1 + \Gamma^2_{22}\boldsymbol{r}_2 + b_{22}\boldsymbol{n}. \end{aligned} \tag{2.1}$$

The symbols $\Gamma^i_{jk}$ are called the *Christoffel symbols*; moreover, it is easy to note that $b_{jk}$ are the coefficients of the second fundamental form:

$$L = b_{11}, \quad M = b_{12} = b_{21}, \quad N = b_{22}.$$

Before starting to derive the equations for the deformation of the vector $\boldsymbol{n}$ we make two agreements.

(1) If some index appears twice in a formula, once as a superscript and once as a subscript, then summation is assumed over this index:

$$a^j b_j := \sum_j a^j b_j.$$

(2) We denote the inverse of the matrix to $a_{ij}$ by $a^{kl}$, assuming by definition that

$$a^{ij} a_{jk} = \delta^i_k = \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

Since $(\boldsymbol{n}, \boldsymbol{n}) \equiv 1$, we have $(\boldsymbol{n}_j, \boldsymbol{n}) = 0$, and consequently the vector $\boldsymbol{n}_j = \frac{\partial \boldsymbol{n}}{\partial u^j}$ is expressed as a linear combination of the vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Now, we derive

$$\frac{\partial(\boldsymbol{n}, \boldsymbol{r}_j)}{\partial u^k} = (\boldsymbol{n}_k, \boldsymbol{r}_j) + (\boldsymbol{n}, \boldsymbol{r}_{jk}) = 0. \tag{2.2}$$

Introduce the coefficients $a^j_k$ so that

$$\boldsymbol{n}_k = a^1_k \boldsymbol{r}_1 + a^2_k \boldsymbol{r}_2 = a^j_k \boldsymbol{r}_j$$

and denote by $g_{jk} = (\boldsymbol{r}_j, \boldsymbol{r}_k)$ the coefficients of the first fundamental form:

$$E = g_{11}, \quad F = g_{12} = g_{21}, \quad G = g_{22}.$$

Rewrite (2.2) as

$$a^1_1 g_{11} + a^2_1 g_{12} = -b_{11}, \quad a^1_1 g_{12} + a^2_1 g_{22} = -b_{12},$$

$$a^1_2 g_{11} + a^2_2 g_{12} = -b_{21}, \quad a^1_2 g_{12} + a^2_2 g_{22} = -b_{22}.$$

The resulting equations split into two systems of equations in $a_1^1, a_1^2$ and $a_2^1, a_2^2$ which can be written as the system

$$a_i^m g_{jm} = -b_{ij}.$$

Since the matrix $g_{jk}$ is invertible and symmetric, it has the symmetric inverse matrix $g^{jk} = g^{kj}$. Multiply both sides of this system by $g^{jk}$ and sum the result over the repeated index $j$. Eventually, we obtain the solution to the system:

$$-b_{ij} g^{jk} = a_i^m g_{jm} g^{jk} = a_i^m \delta_m^k = a_i^k.$$

We have thus proven the following theorem:

**Theorem 2.4** (The Weingarten equations).

$$\boldsymbol{n}_i = -b_{ij}\, g^{jk}\, \boldsymbol{r}_k.$$

These equations, together with the Gauss equations, constitute a complete set of *derivational equations*. As with the Frenet equations, we can write them in the form

$$\frac{\partial}{\partial u^j} \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \boldsymbol{n} \end{pmatrix} = A_j \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \boldsymbol{n} \end{pmatrix},$$

where

$$A_1 = \begin{pmatrix} \Gamma_{11}^1 & \Gamma_{11}^2 & b_{11} \\ \Gamma_{12}^1 & \Gamma_{12}^2 & b_{12} \\ -b_{1j} g^{j1} & -b_{1j} g^{j2} & 0 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} \Gamma_{21}^1 & \Gamma_{21}^2 & b_{21} \\ \Gamma_{22}^1 & \Gamma_{22}^2 & b_{22} \\ -b_{2j} g^{j1} & -b_{2j} g^{j2} & 0, \end{pmatrix}$$

and the Christoffel symbols are symmetric in the subscripts: $\Gamma_{jk}^i = \Gamma_{kj}^i$. Unlike the Frenet equations, in this case nontrivial compatibility conditions arise:

$$\frac{\partial}{\partial u^1} \frac{\partial}{\partial u^2} \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \boldsymbol{n} \end{pmatrix} = \frac{\partial}{\partial u^2} \frac{\partial}{\partial u^1} \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \boldsymbol{n} \end{pmatrix},$$

which are equivalent to the system of equations

$$\frac{\partial}{\partial u^1} A_2 - \frac{\partial}{\partial u^2} A_1 = A_1 A_2 - A_2 A_1$$

called the *Gauss–Codazzi equations*.

Another dependence between the values in (2.1) and the first fundamental form is expressed by the following formulas and is a consequence of the Gauss equations.

**Theorem 2.5.**
$$\Gamma_{ij}^k = \frac{1}{2} g^{lk} \left( \frac{\partial g_{il}}{\partial u^j} + \frac{\partial g_{jl}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^l} \right). \tag{2.3}$$

*Proof.* By (2.1), we have
$$\frac{\partial g_{ij}}{\partial u^k} = \frac{\partial(\boldsymbol{r}_i, \boldsymbol{r}_j)}{\partial u^k} = \Gamma_{ik}^l\, g_{lj} + \Gamma_{jk}^l\, g_{li}.$$

Hence,
$$\frac{\partial g_{jl}}{\partial u^i} + \frac{\partial g_{il}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^l} = 2\,\Gamma_{ij}^m\, g_{ml},$$

which implies that
$$g^{lk} g_{ml} \Gamma_{ij}^m = \delta_m^k\, \Gamma_{ij}^m = \Gamma_{ij}^k = \frac{1}{2} g^{lk} \left( \frac{\partial g_{il}}{\partial u^j} + \frac{\partial g_{jl}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^l} \right).$$

Theorem 2.5 is proven. □

Now, we can state analogs of Theorems 1.2 and 1.4 for surfaces.

**Theorem 2.6** (Bonnet's theorem). *Suppose that*
$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21}(= g_{12}) & g_{22} \end{pmatrix} \quad \textit{and} \quad \begin{pmatrix} ccb_{11} & b_{12} \\ b_{21}(= b_{12}) & b_{22} \end{pmatrix}$$

*are smooth fundamental forms in a domain $U$ homeomorphic to the interior of a disk*[1]*; moreover, the first form is positive definite and the coefficients of these forms satisfy the Gauss–Codazzi equations.*

*Then there is a unique (up to motion) surface in $\mathbb{R}^3$ for which these forms are the first and the second fundamental forms.*

*Proof.* We will only sketch the proof, indicating the principal difference from the proofs of Theorems 1.2 and 1.4, namely the necessity of validity of the compatibility conditions (the Gauss–Codazzi equations).

Take $u_0 = (u_0^1, u_0^2) \in U$. Choose vectors $\boldsymbol{a}_0$, $\boldsymbol{b}_0$, and $\boldsymbol{c}_0$ such that
$$(\boldsymbol{a}_0, \boldsymbol{a}_0) = g_{11}(u_0), \quad (\boldsymbol{a}_0, \boldsymbol{b}_0) = g_{12}(u_0), \quad (\boldsymbol{b}_0, \boldsymbol{b}_0) = g_{22}(u_0),$$
$$(\boldsymbol{a}_0, \boldsymbol{c}_0) = (\boldsymbol{b}_0, \boldsymbol{c}_0) = 0, \quad (\boldsymbol{c}_0, \boldsymbol{c}_0) = 1$$

as the initial data for the Gauss–Weingarten equations:
$$\boldsymbol{r}_1(u_0) = \boldsymbol{a}_0, \quad \boldsymbol{r}_2(u_0) = \boldsymbol{b}_0, \quad \boldsymbol{n}(u_0) = \boldsymbol{c}_0.$$

[1]As we will see from the proof, the only essential condition is the property that the domain is simply connected (the definition of a simply connected domain is given in $3.4): to construct the surface by integration of the form $d\boldsymbol{r} = \boldsymbol{r}_x dx + \boldsymbol{r}_y dy$, it is required that the integral of a closed 1-form along a curve depend only on the endpoints of the curve.

Since the coefficients of the forms satisfy the Gauss–Codazzi equations, the Gauss–Weingarten equations constructed for these forms are compatible and have a unique solution $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ with the given initial data at $u_0$.

Since $\frac{\partial \boldsymbol{a}}{\partial u^2} = \frac{\partial \boldsymbol{b}}{\partial u^1}$ and the domain $U$ is homeomorphic to the interior of a disk, the following equality holds for every closed smooth curve $\gamma$ in the domain $U$ without self-intersections and every $j = 1, 2, 3$:

$$\int_\gamma (a^j \, du^1 + b^j \, du^2) = 0, \tag{2.4}$$

where $\boldsymbol{a} = (a^1, a^2, a^3)$ and $\boldsymbol{b} = (b^1, b^2, b^3)$. Indeed, the contour $\gamma$ bounds some domain $V \in U$ and, by the Stokes formula,

$$\int_\gamma (a^j \, du^1 + b^j \, du^2) = \int_V \left( \frac{\partial b^j}{\partial u^1} - \frac{\partial a^j}{\partial u^2} \right) du^1 \wedge du^2.$$

But the right-hand side of this equality vanishes, since $\frac{\partial \boldsymbol{a}}{\partial u^2} - \frac{\partial \boldsymbol{b}}{\partial u^1} = 0$.

Now, it is easy to construct a map $\boldsymbol{r} : U \to \mathbb{R}^3$ such that

$$\boldsymbol{a} = \boldsymbol{r}_1, \quad \boldsymbol{b} = \boldsymbol{r}_2.$$

To this end, map the point $u_0 \in U$ to the point $(0, 0, 0) \in \mathbb{R}^3$. Join every point $u \in U$ with the point $u_0$ by some smooth curve $\gamma$ in the domain $U$ and define the point $\boldsymbol{r}(u)$ by the following formula:

$$\boldsymbol{r}(u) = \int_\gamma (\boldsymbol{a} \, du^1 + \boldsymbol{b} \, du^2).$$

It follows from (2.4) that this map is independent of the choice of $\gamma$.

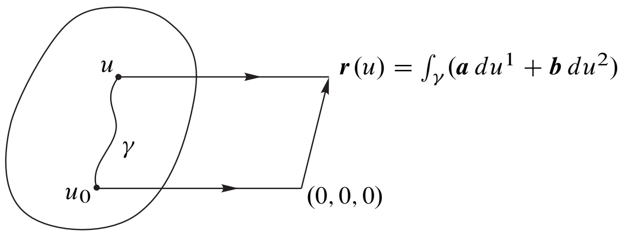


Figure 2.2. The map $\boldsymbol{r}$.

As in the proof of Theorem 1.4, by simple calculations we prove that the first and second fundamental forms of this surface coincide with the given forms.

Uniqueness of the surface up to motion of $\mathbb{R}^3$ is proven by analogy with the case of a plane curve as in Theorem 1.2. These arguments are simple and therefore omitted.

Theorem 2.6 is proven. □

**Problem 2.6.** Suppose that the first fundamental form is diagonal:

$$g_{11} = \lambda(u^1, u^2), \quad g_{12} = 0, \quad g_{22} = \mu(u^1, u^2).$$

Then the Christoffel symbols have the form

$$\Gamma^1_{11} = \frac{1}{2\lambda}\frac{\partial\lambda}{\partial u^1}, \quad \Gamma^1_{12} = \frac{1}{2\lambda}\frac{\partial\lambda}{\partial u^2}, \quad \Gamma^1_{22} = -\frac{1}{2\lambda}\frac{\partial\mu}{\partial u^1},$$
$$\Gamma^2_{11} = -\frac{1}{2\mu}\frac{\partial\lambda}{\partial u^2}, \quad \Gamma^2_{12} = \frac{1}{2\mu}\frac{\partial\mu}{\partial u^1}, \quad \Gamma^2_{22} = \frac{1}{2\mu}\frac{\partial\mu}{\partial u^2}.$$

**Problem 2.7.** Prove that for a surface of revolution (see Problem 2.2) the Christoffel symbols have the form

$$\Gamma^1_{11} = \frac{f'f''}{1+f'^2}, \quad \Gamma^2_{12} = \frac{f'}{f}, \quad \Gamma^1_{22} = -\frac{ff'}{1+f'^2}, \quad \Gamma^1_{12} = \Gamma^2_{11} = \Gamma^2_{22} = 0,$$

where $(u^1, u^2) := (u, v)$.

## 2.5 The Gauss theorem

Gauss proved that the Gaussian curvature is determined only by the first fundamental form. This follows from a formula derived by him and which is included in the system of the Gauss–Codazzi equations.

**Theorem 2.7** (The Gauss theorem).

$$K = \frac{b_{11}b_{22} - b_{12}^2}{g_{11}g_{22} - g_{12}^2}$$
$$= \frac{1}{g_{11}g_{22} - g_{12}^2}\Big((\Gamma^k_{12}\Gamma^l_{12} - \Gamma^k_{11}\Gamma^l_{22})g_{kl} + \frac{\partial^2 g_{12}}{\partial u^1 \partial u^2} - \frac{1}{2}\frac{\partial^2 g_{11}}{\partial u^2 \partial u^2} - \frac{1}{2}\frac{\partial^2 g_{22}}{\partial u^1 \partial u^1}\Big).$$

*Proof.* It follows from (2.1) that

$$(\boldsymbol{r}_{11}, \boldsymbol{r}_{22}) - (\boldsymbol{r}_{12}, \boldsymbol{r}_{12}) = b_{11}b_{22} - b_{12}^2 + \Gamma^k_{11}\Gamma^l_{22}\, g_{kl} - \Gamma^k_{12}\Gamma^l_{12}\, g_{kl}.$$

But

$$(\boldsymbol{r}_{11}, \boldsymbol{r}_{22}) - (\boldsymbol{r}_{12}, \boldsymbol{r}_{12}) = \frac{\partial^2 g_{12}}{\partial u^1 \partial u^2} - \frac{1}{2}\frac{\partial^2 g_{11}}{\partial u^2 \partial u^2} - \frac{1}{2}\frac{\partial^2 g_{22}}{\partial u^1 \partial u^1}$$

which is a consequence of the relation

$$\frac{\partial^2 g_{ij}}{\partial u^k \partial u^l} = \frac{\partial^2 (\boldsymbol{r}_i, \boldsymbol{r}_j)}{\partial u^k \partial u^l} = (\boldsymbol{r}_{ik}, \boldsymbol{r}_{jl}) + (\boldsymbol{r}_{il}, \boldsymbol{r}_{jk}) + (\boldsymbol{r}_i, \boldsymbol{r}_{jkl}) + (\boldsymbol{r}_{ikl}, \boldsymbol{r}_j).$$

To complete the proof, we only need to compare the expressions for $(\boldsymbol{r}_{11}, \boldsymbol{r}_{22}) - (\boldsymbol{r}_{12}, \boldsymbol{r}_{12})$. Theorem 2.7 is proven. □

Theorems 2.5 and 2.7 yield

**Corollary 2.2.** *The Gaussian curvature is expressed in terms of the coefficients of the first fundamental form and their first and second derivatives.*

Regular surfaces $\boldsymbol{r}: U \to \mathbb{R}^3$ and $\tilde{\boldsymbol{r}}: U \to \mathbb{R}^3$ are *isometric* if for every regular curve $\gamma: [a,b] \to U$ the lengths of the curves $\boldsymbol{r}(\gamma)$ and $\tilde{\boldsymbol{r}}(\gamma)$ coincide. This is equivalent to the fact that the values of the first fundamental forms at all points coincide:

$$(\boldsymbol{r}_i, \boldsymbol{r}_j) = (\tilde{\boldsymbol{r}}_i, \tilde{\boldsymbol{r}}_j).$$

Now, we can restate Corollary 2.2 as follows:

**Corollary 2.3.** *If two surfaces $\boldsymbol{r}: U \to \mathbb{R}^3$ and $\tilde{\boldsymbol{r}}: U \to \mathbb{R}^3$ are isometric, then their Gaussian curvatures at the respective points $\boldsymbol{r}(u^1, u^2)$ and $\tilde{\boldsymbol{r}}(u^1, u^2)$ coincide.*

**Problem 2.8.** Prove that the Gaussian curvature of the cone $x^2 + y^2 = z^2$ at a point $(x, y, z) \neq (0,0,0)$ is zero.

## 2.6 Covariant derivative and geodesics

Suppose that $\boldsymbol{r}: U \to \mathbb{R}^3$ determines a regular surface, $\gamma: [a,b] \to U$ is a regular curve on the surface, and $v$ is a smooth *vector field* along the curve $\gamma$ such that

(1) for each $t \in [a,b]$ the vector $v(t)$ lies in the tangent space of the surface at the point $\gamma(t)$;

(2) the vectors $v(t)$ depend smoothly on $t$.

Decompose the vector field $v(t)$ in the bases for the tangent planes:

$$v(t) = v^i(t)\boldsymbol{r}_i(t).$$

The derivative of the vector field along the curve has the form

$$\dot{v} = \frac{dv^i}{dt}\boldsymbol{r}_i + v^i \dot{\boldsymbol{r}}_i = \frac{dv^i}{dt}\boldsymbol{r}_i + v^i \dot{u}^j \boldsymbol{r}_{ij},$$

where $\dot{\gamma} = \left(\dot{u}^1, \dot{u}^2\right)$. Decompose the right-hand side of the last equality into two summands one of which is tangent to the surface and the other is orthogonal:

$$\dot{v} = \left(\frac{dv^i}{dt}\boldsymbol{r}_i + \Gamma^i_{jk} v^j \dot{u}^k \boldsymbol{r}_i\right) + v^j \dot{u}^k b_{jk} \boldsymbol{n}.$$

The *covariant derivative* $\nabla_{\dot{\gamma}} v$ of the vector field $v$ along the curve $\gamma$ is the orthogonal projection of the derivative of $v$ along the curve to the tangent plane of the surface:

$$\frac{Dv}{\partial t} = \left(\frac{dv^i}{dt} + \Gamma^i_{jk} v^j \dot{u}^k\right)\boldsymbol{r}_i.$$

Now, suppose that the vector field $v$ is given in a domain on the surface rather than on a curve only. Then we can consider the covariant derivatives of the field along various curves and the formula for the *covariant derivative of the field $v$ in the direction of a vector $w$* takes the form

$$\nabla_w v = \left(\frac{\partial v^i}{\partial u^k} + \Gamma^i_{jk} v^j\right) w^k \boldsymbol{r}_i,$$

where $w$ is a tangent vector to the surface. This operation possesses a series of remarkable properties.

**Lemma 2.5.** (1) *The map $(w, v) \to \nabla_w v$ is linear in $v$ and $w$:*

$$\nabla_{\alpha_1 w_1 + \alpha_2 w_2} v = \alpha_1 \nabla_{w_1} v + \alpha_2 \nabla_{w_2} v,$$
$$\nabla_w (\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 \nabla_w v_1 + \alpha_2 \nabla_w v_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

(2) *If $f : U \to \mathbb{R}$ is a smooth function, then*

$$\nabla_{fw} v = f \nabla_w v, \quad \nabla_w f v = \partial_w f\, v + f \nabla_w v,$$

*where $\partial_w f$ is the derivative of $f$ in the direction of $w$:*

$$\partial_w f = \sum_j \frac{\partial f}{\partial u^j} w^j;$$

(3) $\nabla_{\boldsymbol{r}_j} \boldsymbol{r}_i = \Gamma^k_{ij} \boldsymbol{r}_k$.

(4) $\Gamma^k_{ij} = \Gamma^k_{ji}$.

(5) *The derivative of the inner product of vector fields is calculated by the formula*

$$\partial_w (v_1, v_2) = (\nabla_w v_1, v_2) + (v_1, \nabla_w v_2).$$

*Proof.* Assertions (1)–(3) are immediate from the definition of the covariant derivative. Assertion (4) follows from the definition of the Christoffel symbols given above and Theorem 2.5.

Assertion (5) is proven by straight calculations:

$$\begin{aligned}
&D_w((\boldsymbol{r}_i, \boldsymbol{r}_j) v_1^i v_2^j) \\
&= w^k \frac{\partial}{\partial u^k}((\boldsymbol{r}_i, \boldsymbol{r}_j) v_1^i v_2^j) \\
&= w^k ((\boldsymbol{r}_{ik}, \boldsymbol{r}_j) + (\boldsymbol{r}_i, \boldsymbol{r}_{jk})) v_1^i v_2^j + w^k (\boldsymbol{r}_i, \boldsymbol{r}_j) \left(\frac{\partial v_1^i}{\partial u^k} v_2^j + v_1^i \frac{\partial v_2^j}{\partial u^k}\right) \\
&= w^k (\Gamma^m_{ik} (\boldsymbol{r}_m, \boldsymbol{r}_j) + \Gamma^m_{jk} (\boldsymbol{r}_i, \boldsymbol{r}_m)) v_1^i v_2^j + w^k (\boldsymbol{r}_i, \boldsymbol{r}_j) \left(\frac{\partial v_1^i}{\partial u^k} v_2^j + v_1^i \frac{\partial v_2^j}{\partial u^k}\right) \\
&= w^k \left(\frac{\partial v_1^i}{\partial u^k} + \Gamma^i_{mk} v_1^m\right) (\boldsymbol{r}_i, \boldsymbol{r}_j) v_2^j + w^k \left(\frac{\partial v_2^j}{\partial u^k} + \Gamma^j_{mk} v_2^m\right) (\boldsymbol{r}_i, \boldsymbol{r}_j) v_1^i \\
&= (\nabla_w v_1, v_2) + (v_1, \nabla_w v_2).
\end{aligned}$$

□

Taking the projection to the tangent plane in the definition of the covariant derivative leads to consideration of the intrinsic geometry of the surface; i.e., we forget about the ambient space and construct the whole geometry from the first fundamental form only. The Gauss theorem is fundamental in the following sense: the Gaussian curvature (unlike the second fundamental form) is an object of the intrinsic geometry.

If a bilinear operation on vector fields on a submanifold of the Euclidean space is such that the assertions (1) and (2) of Lemma 2.5 hold, then we say that an *affine connection* is given on the submanifold. The connection is determined by the values $\Gamma_{ij}^k$. The connection constructed above satisfies two additional conditions: it is *symmetric*; i.e., the assertion (4) of Lemma 2.5 holds and *compatible with the metric*; i.e., the assertion (5) of Lemma 2.5 holds. We give the following fact:[2]

- *a symmetric affine connection compatible with the metric is unique and is determined by the first fundamental form $g_{ij}$ according to* (2.3).

The notion of covariant derivative leads to the notion of *parallel translation*: a vector field $v(t)$ along a curve $\gamma$ is *parallel* if the covariant derivative of $v$ along $\gamma$ is everywhere zero:

$$\frac{Dv}{\partial t} = \left(\frac{dv^i}{dt} + \Gamma_{jk}^i v^j \dot{u}^k\right)\boldsymbol{r}_i = 0. \tag{2.5}$$

Now, by analogy with the Euclidean plane, define an analog of a straight line: a curve $\gamma$ is called a *geodesic* if its velocity vector is parallel along the curve:

$$\frac{D\dot{\gamma}}{\partial t} = 0.$$

**Lemma 2.6.** (1) *The geodesics are the curves satisfying the equation*

$$\ddot{u}^i + \Gamma_{jk}^i \dot{u}^j \dot{u}^k = 0. \tag{2.6}$$

(2) *If $\gamma(t)$ is a geodesic and $C \in \mathbb{R}$, then the curve $\gamma_C(t) = \gamma(Ct)$ is a geodesic, too.*

(3) *A smooth curve $\gamma$ on a surface $\Sigma \subset \mathbb{R}^3$ is a geodesic if and only if, at each point of the curve, the normal to the curve is orthogonal to the surface* (*i.e., is collinear with the normal of the surface*).

The proof of Lemma 2.6 is carried out by elementary calculations. Now we point out one important consequence of the theorem on existence and uniqueness of a solution of an ordinary differential equation.

**Lemma 2.7.** *Suppose that $x_0$ is a point on a regular surface $\Sigma$, $v_0$ is a tangent vector to $\Sigma$ at $x_0$, and $\gamma$ is a curve passing through $x_0$: $\gamma(0) = x_0$. Then*

(1) *there is a unique parallel vector field $v(t)$ along $\gamma$ such that $v(0) = v_0$;*

[2] Its generalization for all dimensions is given by Theorem 4.1.

(2) *for every sufficiently small* $\varepsilon > 0$, *there is a unique geodesic* $\gamma\colon [-\varepsilon, \varepsilon] \to \Sigma$ *such that*

$$\gamma(0) = x_0, \quad \dot{\gamma}(0) = v_0.$$

*Proof.* (1) Since (2.5) is a first-order ordinary differential equation with smooth coefficients, for arbitrary initial data, i.e., for every value of $v$ at zero: $v(0) = v_0$, this equation has a unique solution.

(2) Equation (2.6) is a first-order ordinary differential equation on the set of pairs of the form $(x, v)$, where $x \in M$ and $v \in T_x\Sigma$. In a neighborhood of $x_0$ such a pair is parameterized by points $(u^1, u^2, v^1, v^2) \in \mathbb{R}^4$: $x = \boldsymbol{r}(u^1, u^2)$ and $v = v^1\boldsymbol{r}_1 + v^2\boldsymbol{r}_2$. Equation (2.6) takes the form

$$\dot{u}^i = v^i, \quad \dot{v}^i = -\Gamma^i_{jk}(u^1, u^2)\, v^j v^k, \quad i = 1, 2.$$

Now, assertion (2), like assertion (1), follows from the theorem on existence and uniqueness of a solution to an ordinary differential equation.

Lemma 2.7 is proven. □

The space constituted by the pairs $(x, v)$, where $x \in \Sigma$ and $v \in T_x\Sigma$, is called the *tangent bundle* and is denoted by $T\Sigma$.

**Lemma 2.8.** *Let* $\Sigma$ *be a two-dimensional submanifold in* $\mathbb{R}^3$. *Then* $T\Sigma$ *is a four-dimensional submanifold in* $\mathbb{R}^6$.

*Proof.* Take $(x_0, v_0) \in T\Sigma$. We can assume that in a neighborhood of $x_0$ the surface $\Sigma$ is given by the equation $F(x) = 0$ and $\left(\frac{\partial F}{\partial x^1}, \frac{\partial F}{\partial x^2}, \frac{\partial F}{\partial x^3}\right) \neq 0$. Without loss of generality we may assume that $\frac{\partial F}{\partial x^1} \neq 0$ in this neighborhood. Then, in a neighborhood of $(x_0, v_0)$, the tangent space, as a submanifold in $\mathbb{R}^6$, is determined by the equations

$$F(x, v) = 0, \quad G(x, v) = \sum_{j=1}^{3} \frac{\partial F}{\partial x^j} v^j = 0$$

(see §1.4, Problem 1.4). Now, the lemma follows from the implicit function theorem, since the determinant of the matrix

$$\begin{pmatrix} \partial F/\partial x^1 & \partial F/\partial v^1 \\ \partial G/\partial x^1 & \partial G/\partial v^1 \end{pmatrix} = \begin{pmatrix} \partial F/\partial x^1 & 0 \\ * & \partial F/\partial x^1 \end{pmatrix}$$

is zero in a neighborhood of $(x_0, v_0)$. Lemma 2.8 is proven. □

In view of Lemma 2.7, a *flow* or a *dynamical system* is given on $T\Sigma$; i.e., a system of ordinary differential equations

$$\dot{x} = v(x),$$

where $v(x)$ is a vector field (it is always supposed to be smooth) on the space which is called the *phase space*. This space foliates into the trajectories of this system or the *integral curves* of the vector field $v$.

A *first integral* of a flow is a function on the phase space which is constant along the trajectories. Let us indicate an almost obvious first integral of the geodesic flow.

**Lemma 2.9.** *If $\gamma$ is a geodesic then*

$$\frac{d}{dt}(\dot\gamma, \dot\gamma) = 0.$$

*Proof.* It follows from the assertion (5) of Lemma 2.5 that

$$\frac{d}{dt}(\dot\gamma, \dot\gamma) = \left(\frac{D}{\partial t}\dot\gamma, \dot\gamma\right) + \left(\dot\gamma, \frac{D}{\partial t}\dot\gamma\right) = 2\left(\frac{D}{\partial t}\dot\gamma, \dot\gamma\right).$$

But from the definition of geodesics we obtain $\frac{D\dot\gamma}{\partial t} = 0$. Lemma 2.9 is proven. □

**Problem 2.9.** Suppose that a function $f(x)$ determines a surface of revolution (see Problems 2.2 and 2.7). The parallels are intersections of the surface with planes orthogonal to the axis of revolution (they are given by the equations $u = \text{const}$). Denote by $\alpha$ the angle between the velocity vector of the geodesic and the parallel. Denote by $R(x)$ the distance from $x \in \Sigma$ to the axis of revolution. Prove that the function $I = R\cos\alpha$ is a first integral of the geodesic flow on the surface of revolution.[3]

## 2.7 The Euler–Lagrange equations

Suppose that the surface $\Sigma$ is given by a map

$$\boldsymbol{r} : U \to \mathbb{R}^3$$

and

$$L \colon T\Sigma \to \mathbb{R}$$

is a smooth function on the tangent bundle. Denote by $u^1$ and $u^2$ coordinates in the domain $U \subset \mathbb{R}^2$ and consider them as coordinates on the surface.

Choose two points $x, y \in \Sigma$ and consider the set $\Lambda = \Lambda_{x,y}$ of all parameterized smooth curves

$$\gamma \colon [a, b] \to \Sigma$$

with beginning $x$ and end $y$:

$$\gamma(a) = x, \quad \gamma(b) = y.$$

[3]The function $I$ is called the *Clairaut integral*.

The *action functional*

$$S\colon \Lambda \to \mathbb{R}, \quad S(\gamma) = \int_a^b L(\gamma(t), \dot{\gamma}(t))\, dt$$

is defined on $\Lambda$.

Many problems of mechanics and physics reduce to finding a point in a certain infinite-dimensional functional space $\Lambda$ which minimizes the action functional $S$ corresponding to the problem. A more general problem is the problem of description of extremal points of the functional $S$. Let us explain what it means in our case.

A smooth *variation* of a curve $\gamma$ is a one-parameter family of curves $\gamma_\varepsilon$ ($\gamma_\varepsilon(t) = \gamma(t, \varepsilon)$) such that

(1) $\gamma\colon [a, b] \times (-\varepsilon_0, \varepsilon_0) \to \Sigma$ is a smooth function of both $t \in [a, b]$ and $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$;

(2) $\gamma(a, \varepsilon) = x$ and $\gamma(b, \varepsilon) = y$ for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$;

(3) $\gamma(t, 0) = \gamma(t)$ for all $t \in [a, b]$.

A *variation field* $W(t)$ is a vector field along $\gamma$ of the form

$$W(t) = \frac{\partial \gamma(t, \varepsilon)}{\partial \varepsilon}\Big|_{\varepsilon=0}.$$

A variation field is a natural analog of a tangent vector to $\Lambda$ at $\gamma \in \Lambda$.

A curve $\gamma \in \Lambda$ is called an *extremal function* (*extremal*) of the functional $S$ if

$$\frac{dS(\gamma_\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} = 0$$

for every smooth variation $\gamma_\varepsilon$ of the curve $\gamma$. Obviously, if $\gamma$ minimizes the functional $S$, then the curve $\gamma$ is an extremal of the functional $S$. The converse is not true: as in the case of functions on a finite-dimensional space not each extremum is a minimum.

**Theorem 2.8.** *If a curve $\gamma$ is an extremal of $S$, then the following equations hold along this curve:*

$$\frac{\partial L}{\partial u^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{u}^i}. \tag{2.7}$$

*Proof.* Suppose that $\gamma$ is an extremal, $\gamma_\varepsilon(t) = (u^1(t, \varepsilon), u^2(t, \varepsilon))$ is its variation, and $W(t)$ is the corresponding variation field. Then

$$\begin{aligned}\frac{dS(\gamma_\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} &= \int_a^b \frac{dL(\gamma(t, \varepsilon), \dot{\gamma}(t, \varepsilon))}{d\varepsilon}\Big|_{\varepsilon=0} dt \\ &= \int_a^b \left(\frac{\partial L(\gamma(t), \dot{\gamma}(t))}{\partial u^i}\frac{\partial u^i(t, 0)}{\partial \varepsilon} + \frac{\partial L(\gamma(t), \dot{\gamma}(t))}{\partial \dot{u}^i}\frac{\partial^2 u^i(t, 0)}{\partial \varepsilon\, \partial t}\right) dt.\end{aligned}$$

Integrate the last expression by parts, using the fact that

$$W(a) = \frac{\partial u^i(a,0)}{\partial \varepsilon} = 0, \quad W(b) = \frac{\partial u^i(b,0)}{\partial \varepsilon} = 0. \tag{2.8}$$

We obtain

$$\frac{dS(\gamma_\varepsilon)}{d\varepsilon}\bigg|_{\varepsilon=0} = \int_a^b \left(\frac{\partial L}{\partial u^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{u}^i}\right) W^i(t)\,dt = 0. \tag{2.9}$$

Since every smooth vector field $W(t)$ along $\gamma$, satisfying (2.8), is obviously a variation field for some smooth variation, equation (2.9) implies equations (2.7). Theorem 2.8 is proven. □

Equations (2.7) are called the *Euler–Lagrange equations* for the variational problem with the *Lagrangian function* $L(x, \dot{x})$.

An example of the Euler–Lagrange equations is the geodesic equations (2.6).

**Theorem 2.9.** *If* $L(u, \dot{u}) = |\dot{u}|^2 = g_{ij}\dot{u}^i\dot{u}^j$, *then the Euler–Lagrange equations for this Lagrangian are the geodesic equations* (2.6).

*Proof.* Write down the Euler–Lagrange equations:

$$\frac{\partial L}{\partial u^i} = \frac{\partial g_{jk}}{\partial u^i}\dot{u}^j\dot{u}^k, \quad \frac{\partial L}{\partial \dot{u}^i} = 2g_{ij}\dot{u}^j,$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{u}^i}\right) = 2\frac{\partial g_{ij}}{\partial u^k}\dot{u}^j\dot{u}^k + 2g_{ij}\ddot{u}^j = \left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{ik}}{\partial u^j}\right)\dot{u}^j\dot{u}^k + 2g_{ij}\ddot{u}^j.$$

The Euler–Lagrange equations take the form

$$g_{ij}\ddot{u}^j + \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{ik}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^i}\right)\dot{u}^j\dot{u}^k = 0,$$

or

$$\ddot{u}^m + \frac{1}{2}g^{mi}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{ik}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^i}\right)\dot{u}^j\dot{u}^k = 0.$$

Inserting the expression (2.3) for the Christoffel symbols in the last formula, we obtain the geodesic equation

$$\ddot{u}^m + \Gamma^m_{jk}\dot{u}^j\dot{u}^k = 0.$$

Theorem 2.9 is proven. □

Consider the variational problem for the Lagrangian

$$L_0(u, \dot{u}) = |\dot{u}| = \sqrt{g_{ij}\dot{u}^i\dot{u}^j}.$$

The corresponding action functional is the length of the curve. The Lagrangian is not differentiable at the points where $\dot{u} = 0$. But the value $S$ is independent of the choice of parameterization on the curve: it is the same for equivalent curves (see Lemma 1.1: we restrict our consideration to regular curves, but it is easy to show that this constraint is unessential in this case). Therefore, if a curve is an extremal of the length functional, then the equivalent curve with the parameter proportional to the arc length parameter ($|\dot{u}| = \text{const}$) is an extremal of the length functional, too.

**Theorem 2.10.** *The extremals of the length functional coincide with the geodesics up to equivalence.*

*Proof.* The Euler–Lagrange equations for the length functional take the form

$$\frac{1}{2\sqrt{g_{lm}\dot{u}^l\dot{u}^m}}\frac{\partial g_{jk}}{\partial u^i}\dot{u}^j\dot{u}^k = \frac{d}{dt}\left(\frac{1}{\sqrt{g_{lm}\dot{u}^l\dot{u}^m}}g_{ij}\dot{u}^j\right). \tag{2.10}$$

The parameter on the geodesic is proportional to the arc length parameter (see Lemma 2.9). Putting $g_{lm}\dot{u}^l\dot{u}^m = \text{const} \neq 0$ in (2.10), we obtain the geodesic equations (2.6) and conclude that the geodesics are extremals of the length functional.

We can assume that on each extremal of the length functional the parameter is proportional to the arc length parameter. An extremal with this parameter is described by equations (2.6) and hence is a geodesic.

Theorem 2.10 is proven. □

We introduce the *distance* between points of a surface as follows:

$$d(x, y) = \inf_{\gamma \in \Lambda_{x,y}} \mathrm{L}(\gamma).$$

It follows from Theorem 2.10 that if there is a curve in $\Lambda_{x,y}$ with length $d(x, y)$, then it is a geodesic (up to parametrization). We restrict our consideration to the following local fact:

**Theorem 2.11.** *Let $\gamma$ be a geodesic passing through a point $x_0$. There is a neighborhood of $x_0$ on $\gamma$ such that, for an arbitrary pair $x_1$ and $x_2$ in this neighborhood, the shortest curve connecting the points $x_1$ and $x_2$ is a segment of the geodesic $\gamma$.*

*Proof.* First prove the following technical lemma:

**Lemma 2.10.** *Let $x_0$ be a point on the surface and let $\gamma$ be a geodesic passing through $x_0$. Then, in a neighborhood of $x_0$, we can choose coordinates $u^1$ and $u^2$ such that the first fundamental form takes the shape*

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & G \end{pmatrix}$$

*and the equation for $\gamma$ becomes $u^2 = 0$. With these coordinates the curves $u^2 = const$ are geodesics.*

*Proof.* Choose coordinates $y^1$ and $y^2$ in a neighborhood of $x_0$ so that $x_0 = (0,0)$ and the tangent vector to $\gamma$ at $x_0$ is equal to $(1,0)$.

Through each point with coordinates $(0,s)$ draw a geodesic with the initial velocity vector $(1,0)$. Then there is a function $\varphi$ in a neighborhood of $x_0$ such that these geodesics are given by the equations $y^2 = \varphi(y^1,s)$. Since the initial data for the geodesics depend smoothly on $s$, $\varphi$ is a smooth function. Consider the map $(y^1,s) \to (y^1, y^2 = \varphi(y^1,s))$. Its Jacobian matrix at $x_0$ is equal to

$$\begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}$$

and, by the inverse function theorem, it is invertible near the point $x_0$. Consequently, $y^1$ and $s$ determine local coordinates in a neighborhood of $x_0$ and the curves $s = \text{const}$ are geodesics.

At each point of this neighborhood of $x_0$, take a tangent vector $w(y^1,s)$ orthogonal to the curve $s = \text{const}$ and such that $(w,w) = 1$ and $(w, \boldsymbol{r}_s) > 0$. A small neighborhood of the point $x_0$ foliates into the trajectories of the smooth vector field $w$, i.e., the solutions to the equation $\dot{x} = w(x)$. Introduce the coordinate $t$ instead of $y^1$, by putting $x = (t,s)$ if the point $x$ lies on the trajectory of the flow $\dot{x} = w(x)$ which intersects the curve $s = 0$ at $(t,0)$. The first fundamental form in the coordinates $(t,s)$ is equal to $\tilde{g}_{11} = \lambda$, $\tilde{g}_{22} = \mu$, $\tilde{g}_{12} = 0$.

Each curve $s = \text{const}$ is a geodesic and consequently $\Gamma^2_{11} \equiv 0$. Hence, $\partial\lambda/\partial s = 0$ (see Problem 2.6) and the coordinates

$$u^1 = \int_0^t \sqrt{\lambda(\tau)}\, d\tau, \quad u^2 = s$$

are those we seek. Lemma 2.10 is proven. □

We turn to proving Theorem 2.11. In a neighborhood of $x_0$ introduce the coordinate system $(u^1,u^2)$ connected with $x_0$ and $\gamma$ as in Lemma 2.10. We can assume that $x_0 = (0,0)$. Choose $\varepsilon > 0$ such that the disk $B = \{|u| \leq \varepsilon\}$ lies completely in this neighborhood. Let $C = \min_{x \in B} G(x)$ and $D = \min\{1, C\}$.

If $x_1$ and $x_2$ lie in the disk $B_0 = \{|u| \leq \rho\varepsilon\}$, where $\rho = D/(D+2)$, then

(1) $|u^1(x_1) - u^1(x_2)| \leq 2\rho\varepsilon$;

(2) every smooth curve connecting $x_1$ and $x_2$ and leaving the disk $B$ at some moment has a length greater than $4\rho\varepsilon$;

(3) the length of a smooth curve connecting $x_1$ and $x_2$ and lying completely in $B$ is equal to

$$\int \sqrt{(\dot{u}^1)^2 + G(\dot{u}^2)^2}\, dt \geq \int \sqrt{(\dot{u}^1)^2} \geq |u^1(x_1) - u^1(x_2)|. \tag{2.11}$$

But the equality in (2.11) is attained precisely at the segment $\gamma_0$.

Theorem 2.11 is proven. □

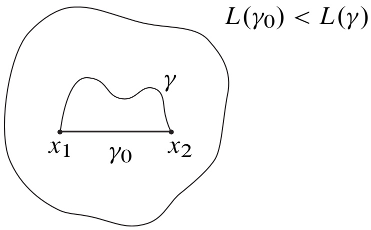


Figure 2.3. The shortest path in the semigeodesic coordinates.

Lemma 2.10 claims that in a neighborhood of every point there are special coordinates convenient for calculations. They are called the semigeodesic coordinates. We now give a rigorous definition.

Coordinates $(u^1, u^2) = (x, y)$ are called *semigeodesic* if the first fundamental form has the shape

$$g_{ij}du^i du^j = dx^2 + Gdy^2.$$

**Problem 2.10.** Prove that, in semigeodesic coordinates $(x, y)$, every curve of the form $y = \text{const}$ is a geodesic.

**Example** (Spherical coordinates). Define the spherical coordinates $r$, $\theta$, $\varphi$ in $\mathbb{R}^3$ by the formulas

$$x^1 = r\cos\varphi\sin\theta, \quad x^2 = r\sin\varphi\sin\theta, \quad x^3 = r\cos\theta,$$

where $r \geq 0$, $0 \leq \theta \leq \pi$, and $0 \leq \varphi \leq 2\pi$. For each fixed value $r = \text{const} \neq 0$ the values $\varphi$ and $\theta$ are coordinates on the sphere of radius $r$. The metric tensor is equal to

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\varphi^2.$$

Therefore, for every value of $r$, the coordinates $(x, y) = (r\theta, \varphi)$ are semigeodesic coordinates and the great circles $\varphi = \text{const}$ are geodesics on the sphere of radius $r$.

**Problem 2.11.** Prove that the Gaussian curvature of a surface with semigeodesic coordinates $(x, y)$ is equal to

$$K = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial x^2}.$$

**Problem 2.12.** Prove that if the Gaussian curvature $K$ of a surface is constant and $K \neq 0$, then there exist semigeodesic coordinates $(x, y)$ in which the first fundamental form has the shape

$$\begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\sqrt{K}x) \end{pmatrix} \quad \text{for } K > 0,$$

$$\begin{pmatrix} 1 & 0 \\ 0 & \sinh^2(\sqrt{-K}x) \end{pmatrix} \quad \text{for } K < 0.$$

## 2.8 The Gauss–Bonnet formula

Geodesics represent a natural generalization of straight lines to the case of arbitrary surfaces (straight lines are geodesics on the plane: in this case $\Gamma^i_{jk} = 0$ in the linear coordinates and the geodesic equations become linear: $\ddot{u}^i = 0, i = 1, 2$). The deviation of an arbitrary curve from a geodesic is described by the analog of the curvature of a plane curve, namely the geodesic curvature. We now give its definition.

Let $\gamma\colon [a,b] \to U \subset \mathbb{R}^2$ be a naturally parameterized curve on a surface $\boldsymbol{r} : U \to \mathbb{R}^3$. Choose an orientation in the tangent planes to the surface, assuming that the basis $(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is positively oriented. At each point $\gamma(l)$ of the curve choose an orthonormal, positively oriented basis $(\dot{\gamma}, n)$ for the tangent space. The *geodesic curvature* is the value

$$k_g = (\nabla_{\dot{\gamma}}\dot{\gamma}, n),$$

i.e., the immediate analog of the curvature of a plane curve (see (1.3)). It follows from Lemma 2.6 that a smooth curve is a geodesic if and only if its geodesic curvature is everywhere zero.

Let $W$ be a small neighborhood of a point of a surface with semigeodesic coordinates $(x, y) := (u^1, u^2)$ and let $\gamma = \gamma_1 \cup \cdots \cup \gamma_n$ be a piecewise smooth closed contour (i.e., a family of successively traversed regular curves: the endpoint of $\gamma_j$ coincides with the starting point of $\gamma_{j+1}$; moreover, put $\gamma_{n+1} = \gamma_1$). We suppose that the contour $\gamma$ has no self-intersections and bounds a domain $V \subset W$.

Denote by $\alpha_j$ the internal (with respect to $V$) angle between $\gamma_j$ and $\gamma_{j+1}$ at a common endpoint and by $d\sigma$ the area form on the surface

$$d\sigma = \sqrt{g_{11}g_{22} - g_{12}^2}\, dx \wedge dy.$$

**Theorem 2.12** (The Gauss–Bonnet formula).

$$\int_\gamma k_g dl = 2\pi - \sum_{j=1}^{n} (\pi - \alpha_j) - \int_V K\, d\sigma. \tag{2.12}$$

*Proof.* In semigeodesic coordinates $(x, y) := (u^1, u^2)$ the first fundamental form has the shape

$$dx^2 + G dy^2,$$

and the Christoffel symbols are equal to (see Problem 2.6)

$$\Gamma^1_{22} = -\frac{1}{2}G_x, \quad \Gamma^2_{12} = \frac{1}{2G}G_x, \quad \Gamma^2_{22} = \frac{1}{2G}G_y,$$

with the other Christoffel symbols equal to zero. The normal to the curve is taken with respect to the first fundamental form and is calculated explicitly

$$n = \frac{1}{\sqrt{G}}(-G\dot{y}\boldsymbol{r}_1 + \dot{x}\boldsymbol{r}_2)$$

(recall that the dot denotes differentiation with respect to the arc length parameter $l$: $|\dot\gamma| = 1$). Now, insert these expressions in the formula for the geodesic curvature to obtain

$$k_g = \sqrt{G}\Big(-\ddot x\dot y + \dot x\ddot y + \frac{1}{2}G_x\dot y^3 + \frac{1}{G}G_x\dot x^2\dot y + \frac{1}{2G}G_y\dot x\dot y^2\Big).$$

Since $|\dot\gamma|^2 = \dot x^2 + G\dot y^2 = 1$, we have

$$\frac{d}{dl}\arctan\left(\frac{\sqrt{G}\dot y}{\dot x}\right) = \sqrt{G}\Big(-\ddot x\dot y + \dot x\ddot y + \frac{1}{2G}G_x\dot x^2\dot y + \frac{1}{2G}G_y\dot x\dot y^2\Big),$$

$$(\sqrt{G})_x\dot y = (\dot x^2 + G\dot y^2)\frac{G_x\dot y}{2\sqrt{G}} = \sqrt{G}\Big(\frac{1}{2}G_x\dot y^3 + \frac{1}{2G}G_x\dot x^2\dot y\Big).$$

Hence,

$$k_g\,dl = d\arctan\left(\frac{\sqrt{G}\dot y}{\dot x}\right) + (\sqrt{G})_x\dot y\,dl.$$

From the Stokes formula we obtain

$$\int_\gamma(\sqrt{G})_x\dot y\,dl = \int_\gamma(\sqrt{G})_x\,dy = \int_V(\sqrt{G})_{xx}\,dx\wedge dy$$

and, using the formula for the Gaussian curvature in semigeodesic coordinates (see Problem 2.11), we derive

$$\int_\gamma(\sqrt{G})_x\dot y\,dl = -\int_V K\,d\sigma.$$

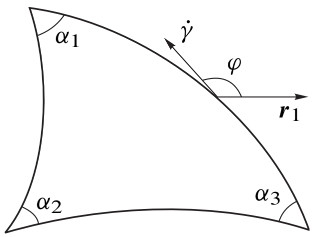


Figure 2.4. A polygon, the angles $\alpha_j$ and $\varphi$ in the Gauss–Bonnet formula.

Now, note that the angle

$$\arctan\left(\frac{\sqrt{G}\dot y}{\dot x}\right)$$

is equal to the angle $\varphi$ between $\dot\gamma$ and $\boldsymbol{r}_1$ (up to $\pi$). If the contour $\gamma$ is smooth, then

$$\int_\gamma d\varphi = 2\pi,$$

while if the successive parts $\gamma_j$ meet at nonzero angles, then it is easy to note that

$$\int_\gamma d\varphi = 2\pi - \sum_{j=1}^{n}(\pi - \alpha_j).$$

Theorem 2.12 is proven. □

Show that the Gauss–Bonnet formula is also valid for large domains $V$ homeomorphic to a disk. First of all define the notion of simplicial partition.

Let $V$ be either a closed domain (the closure of an open set) on the plane with piecewise smooth boundary or a compact surface in $\mathbb{R}^3$. A *simplicial partition*[4] of the surface $V$ is its representation in the form of a finite union of triangles

$$V = \bigcup_j \delta_j$$

such that

(1) the interior of each triangle $\delta_j$ is a domain in $V$ and the closure of this domain is homeomorphic to a triangle;

(2) three vertices are selected on the boundary of each triangle and the parts of the boundary between them are called *edges*;

(3) two different triangles can intersect only along one common side or at one common vertex, and two different sides can intersect only at one common vertex.

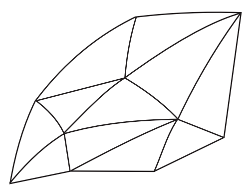

Figure 2.5. A simplicial partition.

If the boundaries of the triangles are piecewise smooth contours, then the partition is called *piecewise smooth*.

Let $\Delta$ be a simplicial partition of a closed domain $V$. Denote by $a_0$ the number of vertices, by $a_1$ the number of edges, and by $a_2$ the number of faces (triangles). The quantity

$$\chi(\Delta) = a_0 - a_1 + a_2$$

is called the *Euler characteristic* of the partition.

[4]For an exposition of the topological theory of simplicial complexes of an arbitrary dimension see [2], III, and [6].

**Theorem 2.13.** *If a closed domain $V$ is homeomorphic to a disk on the plane, then the Euler characteristic of every simplicial partition of $V$ equals* 1.

*Proof.* We proceed by induction on $a_2$. For $a_2 = 1$ the assertion is obvious. Assume that it is valid for $a_2 \leq k$.

Take an arbitrary partition $\Delta$ of the closed domain $V$ with $a_2 = k + 1$. Choose a triangle $\delta_j$ whose side $\gamma^*$ lies on the boundary. Delete the side $\gamma^*$ and the interior of $\delta_j$ from $V$, thereby obtaining a new closed domain $V'$ with the partition $\Delta \setminus \delta_j$. One of the following two situations is possible:

(1) the closed domain $V'$ is homeomorphic to a disk;

(2) the closed domain $V'$ is homeomorphic to the union of two closed domains $V_1$ and $V_2$ with partitions $\Delta_1$ and $\Delta_2$ and these domains intersect at a common vertex.

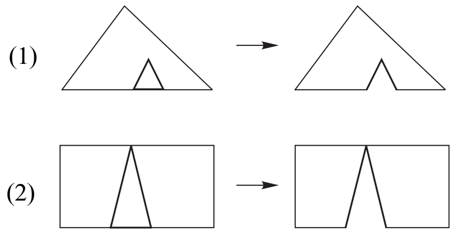


Figure 2.6. "Two different situations".

In the first case we obviously have $\chi(\Delta) = \chi(\Delta') = 1$. In the second case $\chi(\Delta) = \chi(\Delta') = \chi(\Delta_1) + \chi(\Delta_2) - 1 = 1$.

Theorem 2.13 is proven. □

We now prove the Gauss–Bonnet formula for large domains.

**Theorem 2.14.** *If $V$ is a closed domain on a surface which is homeomorphic to a disk and has a piecewise smooth boundary, then the Gauss–Bonnet formula* (2.12) *is valid for it.*

*Proof.* Choose a piecewise smooth simplicial partition of the domain $V$ into small triangles $\delta_k$ each of which lies in a domain with semigeodesic coordinates. Denote by $c_0$ the number of vertices lying on the boundary $\partial V$ of the domain $V$ and by $c_1$ the number of sides lying on the boundary $\partial V$. Since the boundary $\partial V$ is homeomorphic to a circle, $c_0 = c_1$.

Write down (2.12) for each triangle $\delta_k$ and sum them. Since the integrals of the geodesic curvature $k_g$ along the interior edges are taken twice with different signs, the sum of the left-hand sides is equal to $\int_\gamma k_g \, dl$, where $\gamma = \partial V$ is the boundary of $V$. On the right-hand side we obtain

$$2\pi a_2 - 3\pi a_2 + 2\pi(a_0 - c_0) + \sum_j \alpha_j - \int_V K \, d\sigma,$$

where $\alpha_j$ are the angles between the smooth parts of the boundary $V$. It is obvious that $3a_2 = 2a_1 - c_1$ and we find that

$$\int_\gamma k_g \, dl = 2\pi a_2 - 2\pi a_1 + \pi c_1 + 2\pi(a_0 - c_0) + \sum_j \alpha_j - \int_V K \, d\sigma.$$

Since $c_0 = c_1$, from Theorem 2.13 we obtain

$$\begin{aligned}\int_\gamma k_g \, dl &= 2\pi(a_2 - a_1 + a_0) - \sum_j (\pi - \alpha_j) - \int_V K \, d\sigma \\ &= 2\pi - \sum_j (\pi - \alpha_j) - \int_V K \, d\sigma.\end{aligned}$$

Theorem 2.14 is proven. □

The Gauss–Bonnet formula has a series of beautiful consequences.

First, we can apply it to closed surfaces in $\mathbb{R}^3$, i.e., to compact surfaces without boundary. A surface is called orientable if we can choose an orientation in the tangent space at each point so that it changes continuously as the point moves over the surface. The simplest examples of such surfaces are a torus and a sphere. Now, delete $g$ disks from a sphere to obtain a sphere with $g$ holes. Take $g$ tori with the interior of a disk deleted and glue each of these tori to the sphere with holes, identifying the boundaries of the contours. We obtain a *sphere with $g$ handles*. It is well known that

- *each closed orientable surface has the structure of a sphere with handles*.[5]

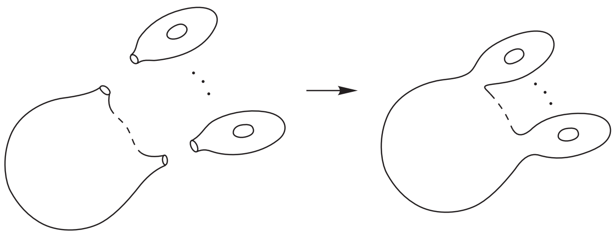

Figure 2.7. A construction of a sphere with handles.

Once an orientation on the surface is chosen, the surface is called oriented and we can take surface integrals over it, in particular, the integrals of $K \, d\sigma$.

[5]This means that it is homeomorphic to a sphere with handles (a strict definition of a homeomorphism is given below in §6). A proof of this fact can be found, for example, in [6], [4].

**Theorem 2.15.** *Let $\Sigma$ be a closed oriented surface in $\mathbb{R}^3$. Then the following equality holds for every simplicial partition $\Delta$ of $\Sigma$:*

$$\int_\Sigma K\,d\sigma = 2\pi\chi(\Delta).$$

*Proof.* Every simplicial partition can be slightly deformed so that it becomes piecewise smooth (this follows from the Weierstrass theorem on approximation of a continuous function by polynomials). Moreover, the numbers of vertices, sides, and (faces) triangles remain the same.

Suppose that $\Delta$ is a piecewise smooth simplicial partition and apply the Gauss–Bonnet formula (Theorem 2.14) to each triangle in $\Delta$. Sum these formulas. Since the integrals of $k_g$ along the sides are taken twice with different signs, the left-hand side is zero. The right-hand side is equal to

$$2\pi a_2 - 3\pi a_2 + 2\pi a_0 - \int_\Sigma K\,d\sigma,$$

but $3a_2 = 2a_1$, since all edges are interior. Eventually, we conclude that

$$2\pi\chi(\Delta) = \int_\Sigma K\,d\sigma.$$

Theorem 2.15 is proven. □

**Corollary 2.4.** *For a closed orientable surface $\Sigma$ the Euler characteristic of a simplicial partition is independent of the partition and is determined only by the surface. It is called the Euler characteristic $\chi(\Sigma)$ of the surface $\Sigma$.*

**Problem 2.13.** Prove that the Euler characteristic of a sphere with $g$ handles is equal to $2 - 2g$ and, in particular, the Euler characteristics of a sphere and a torus are equal to 2 and 0, respectively.

Another remarkable application of the Gauss–Bonnet formula is the formula for the sum of the angles of a triangle. A domain which is homeomorphic to the interior of a triangle and is bounded by three segments of geodesics is called a *geodesic triangle*.

**Theorem 2.16.** *The sum of the angles $\alpha_1$, $\alpha_2$, and $\alpha_3$ of a geodesic triangle $\triangle$ on a surface is equal to*

$$\alpha_1 + \alpha_2 + \alpha_3 = \pi + \int_\triangle K\,d\sigma.$$

The proof of the theorem is immediate from the Gauss–Bonnet formula, since the geodesic curvature of the sides of a geodesic triangle is everywhere zero. It follows from the theorem that if the Gaussian curvature $K$ is positive (negative) then the sums of the angles of triangles are greater (less) than $\pi$.

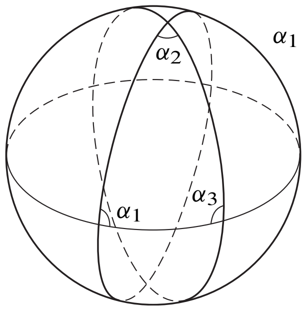


Figure 2.8. A geodesic triangle and the sum of its angles (on a sphere).

## 2.9 Minimal surfaces

Minimal surfaces generalize geodesics to two dimensions.

On an oriented surface we can define the area form

$$d\sigma = \sqrt{\det(g_{ij})}\, du^1 \wedge du^2 = \sqrt{g_{11}g_{22} - g_{12}^2}\, du^1 \wedge du^2,$$

where the frame $e_1$, $e_2$ corresponding to the coordinates $u^1$ and $u^2$ in each tangent space is positively oriented.

We say that a one-parameter family of surfaces $\Sigma_\varepsilon$ is a smooth deformation of a surface $\Sigma$ if

(1) $\Sigma_0 = \Sigma$;

(2) the surfaces $\Sigma_\varepsilon$ are given by functions which depend smoothly on the deformation parameter $\varepsilon$.

We say that a closed domain $V \subset \Sigma$ is the support of the deformation if the part of the surface lying outside $V$ does not deform.

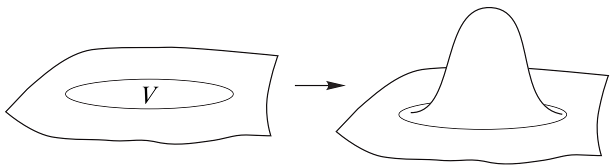


Figure 2.9. A deformation of a surface and its support.

If the support $V$ of the deformation is compact, then the area $S(\varepsilon)$ of the deformed part $V_\varepsilon$ is finite and is a smooth function of the parameter $\varepsilon$. A surface is *minimal* if

$$\frac{d}{d\varepsilon} S(\varepsilon)\Big|_{\varepsilon=0} = 0$$

for every deformation with compact support $V$.

This notion has a rather clear origin: if $\gamma$ is a closed contour in $\mathbb{R}^3$ and there is a surface $\Sigma$ bounded by $\gamma$ which has the least area among all surfaces bounded by the contour $\gamma$, then this surface $\Sigma$ is minimal.[6]

In the case of geodesics we consider one-dimensional objects, i.e., curves, which minimize the one-dimensional volume, the length, among all curves bounded by a pair of points. As in the case of geodesics, a minimal surface bounded by a contour $\gamma$ may fail to provide a minimum of the area functional: it only satisfies the Euler–Lagrange equations for this functional. We now derive them.

**Theorem 2.17.** *A regular surface $\Sigma$ given by a map $\boldsymbol{r} : U \to \mathbb{R}^3$ is minimal if and only if its mean curvature is everywhere zero:*

$$H = 0.$$

*Proof.* Let $V$ be a closed subdomain in $U$ and let $\gamma$ be the boundary of $V$. A deformation of the surface concentrated on $V$ has the form

$$\boldsymbol{r}^{\varepsilon}(u^1, u^2) = \boldsymbol{r}(u^1, u^2) + \varepsilon\varphi\boldsymbol{n} + O(\varepsilon^2),$$

where the function $\varphi$ is zero outside $V$. The area of the deformed part of the surface $\boldsymbol{r}^{\varepsilon}(V)$ is equal to

$$S(\varepsilon) = \int_V \sqrt{(\boldsymbol{r}_1^{\varepsilon}, \boldsymbol{r}_1^{\varepsilon})(\boldsymbol{r}_2^{\varepsilon}, \boldsymbol{r}_2^{\varepsilon}) - (\boldsymbol{r}_1^{\varepsilon}, \boldsymbol{r}_2^{\varepsilon})(\boldsymbol{r}_1^{\varepsilon}, \boldsymbol{r}_2^{\varepsilon})}\, du^1 du^2.$$

Since

$$\boldsymbol{r}_k^{\varepsilon} = \boldsymbol{r}_k + \varepsilon\varphi\boldsymbol{n}_k + \varepsilon\varphi_k\boldsymbol{n} + O(\varepsilon^2)$$

and $(\boldsymbol{r}_1, \boldsymbol{n}) = (\boldsymbol{r}_2, \boldsymbol{n}) = 0$, we derive

$$(\boldsymbol{r}_i^{\varepsilon}, \boldsymbol{r}_j^{\varepsilon}) = (\boldsymbol{r}_i, \boldsymbol{r}_j) + \varepsilon\varphi((\boldsymbol{r}_i, \boldsymbol{n}_j) + (\boldsymbol{r}_j, \boldsymbol{n}_i)) + O(\varepsilon^2).$$

It follows from (2.2) that $(\boldsymbol{r}_i, \boldsymbol{n}_j) = -b_{ij}$ and hence

$$\begin{aligned} S(\varepsilon) &= \int_V \sqrt{1 - 2\varepsilon\varphi\frac{b_{11}g_{22} + b_{22}g_{11} - b_{12}g_{21} - b_{21}g_{12}}{g_{11}g_{22} - g_{12}^2} + O(\varepsilon^2)}\, d\sigma \\ &= S(0) - \varepsilon\int_V \frac{b_{11}g_{22} + b_{22}g_{11} - b_{12}g_{21} - b_{21}g_{12}}{g_{11}g_{22} - g_{12}^2}\varphi\, d\sigma + O(\varepsilon^2). \end{aligned}$$

[6] Plateau's famous problem was to prove existence of a minimal surface with an arbitrary piecewise smooth boundary in $\mathbb{R}^3$; its study had a great influence on the development of the theory of partial differential equations. Its final solution was obtained by Douglas and Rado in the 1930s by means of variational methods: a minimal surface was obtained as an extremal of another functional, the Dirichlet functional.

It is easy to see that the sum of the roots $k_1$ and $k_2$ of the equation $P(\lambda) = \det(b_{ij} - \lambda g_{ij}) = 0$ is equal to

$$k_1 + k_2 = \frac{b_{11}g_{22} + b_{22}g_{11} - b_{12}g_{21} - b_{21}g_{12}}{g_{11}g_{22} - g_{12}^2}.$$

By Lemma 2.4, this is exactly the doubled mean curvature of the surface: $2H = k_1 + k_2$.

Eventually, we obtain

$$\frac{d}{d\varepsilon}S(\varepsilon)\Big|_{\varepsilon=0} = -2\int_V H\varphi\, d\sigma.$$

This value vanishes for all deformations supported in $V$, i.e., for arbitrary smooth functions $\varphi$ equal to zero outside the closed domain $V$, if and only if $H = 0$ at each interior point of $V$.

Theorem 2.17 is proven. □

**Problem 2.14.** Prove that

(1) the surfaces of revolution (see Problem 6) obtained by revolution of the graphs of the functions $f(x) = a\cosh(x/a + b)$, where $a \neq 0$, are minimal surfaces (they are called *catenoids*);

(2) if a surface of revolution is minimal, then it is either a catenoid or a plane obtained by revolution of a straight line orthogonal to the axis of revolution.

# Part II

# Riemannian geometry

# 3

# Smooth manifolds

## 3.1 Topological spaces

A *topological space* is a set $X$ of points with a collection of so-called *open* subsets. This collection satisfies the following conditions:

(1) the union of arbitrarily many open sets is open;

(2) the intersection of finitely many open sets is open;

(3) the whole set $X$ and its empty subset are open.

The complement to an open set is called a *closed* set.

The family of open sets is called the *base* (*of topology*) if every open set is representable as the union of sets of this family. Sometimes, it is simpler to define a *topology* (the structure of a topological space) on a set $X$ by taking only the additive generators, i.e., the topology base, rather than prescribing all open sets. For example, we can define the topology on metric spaces.

A *metric* on a set $X$ is a function

$$\rho \colon X \times X \to \mathbb{R}$$

such that the following conditions are satisfied:

(1) $\rho(x, y) = \rho(y, x)$;

(2) $\rho(x, x) = 0$ and $\rho(x, y) > 0$ for $x \neq y$;

(3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ (the triangle inequality).

A *metric space* is a set $X$ with a metric $\rho$ and the topology defined by the base constituted by all open balls $B_{x,\varepsilon} = \{y \in X \mid \rho(x, y) < \varepsilon\}$, where $x \in X$ and $\varepsilon > 0$.

Another way to introduce the topology is to induce it. Namely, each subset $Y \subset X$ of a topological space $X$ is endowed with the *subspace* (or *induced*) topology in which a set $V \subset Y$ is open if and only if it is representable as the intersection $V = U \cap Y$, where $U$ is an open set in $X$. Later, unless the contrary is specified, by the topology on subsets of topological spaces we always mean the induced topology.

Once a topology is given, we can define the proximity relation between points:

- a *neighborhood* of a point $x \in X$ is an open set $U$ containing $x$: $x \in U$;

and continuity of maps from $X$ to $Y$:

- a map $f : X \to Y$ of topological spaces is *continuous at a point* $x \in X$ if for every neighborhood $V$ of the point $f(x)$ there is a neighborhood $U$ of $x$ such that $f(U) \subset V$;

- a map $f : X \to Y$ of topological spaces is *continuous* if one of the following equivalent conditions holds:
  (a) the map $f$ is continuous at each point of $X$;
  (b) for every open set $V \subset Y$ its inverse image $f^{-1}(V) \subset X$ is open.

**Problem 3.1.** Prove equivalence of conditions (a) and (b) in the definition of a continuous map.

A continuous map $f : X \to \mathbb{R}$ is called a *continuous function*.

For maps of metric spaces, the notion of continuity generalizes the notion of a continuous function on an interval.

**Problem 3.2.** A map $f : X \to Y$ of metric spaces $X$ and $Y$ is continuous at a point $x \in X$ if and only if, for every $\varepsilon > 0$, there is $\delta > 0$ such that if $\rho_X(x, x') < \delta$, then $\rho_Y(f(x), f(x')) < \varepsilon$.

Let

$$\rho(x, y) = \sqrt{\sum_{i=1}^{n} (x^i - y^i)^2}$$

be the distance between points $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^n)$ in the Euclidean space $\mathbb{R}^n$. Then the metric topology on $\mathbb{R}^n$ is the usual topology known from a calculus course.

We introduce some classes of topological spaces:

- a space $X$ is a *Hausdorff space* if, for every pair of different points $x, y \in X$, there are disjoint neighborhoods $U$ and $V$ of $x$ and $y$: $x \in U$, $y \in V$ and $U \cap V = \emptyset$;
- a space $X$ is *connected* if it cannot be represented as the union of two nonempty subsets each of which is open and closed simultaneously;
- a space $X$ is *path-connected* if every pair of different points $x, y \in X$ can be joined by a continuous curve; i.e., there is a continuous map $f : [0, 1] \to X$ such that $f(0) = x$ and $f(1) = y$.

Let $X$ be a topological space and let $\{U_\alpha\}_{\alpha \in A}$ be a family of its subsets indexed by elements of a set $A$ such that the union of $U_\alpha$ coincides with $X$:

$$X = \bigcup_{\alpha \in A} U_\alpha.$$

In this case the family $\{U_\alpha\}$ is called a *covering* of $X$. If $\{U_\beta\}_{\beta \in B}$ is a subfamily (i.e., $B \subset A$) which is a covering of $X$, then it is called a *subcovering* of the covering $\{U_\alpha\}_{\alpha \in A}$. If all sets $U_\alpha$ are open, then the covering is called *open*. The following notion is very important in topology and analysis:

- a space $X$ is *compact* if from every open covering of $X$ we can extract a finite subcovering.

Examples of compact spaces are intervals $[a, b] \subset \mathbb{R}$, $-\infty < a, b < \infty$ (the Heine–Borel lemma).

**Problem 3.3.** A subset $X$ of the Euclidean space $\mathbb{R}^n$ is compact if and only if it is closed in $\mathbb{R}^n$ and bounded (is contained completely in a finite cube $\{|x| \leq N\}$ for some $N < \infty$).

**Problem 3.4.** If $f: X \to Y$ is a continuous map and $X$ is compact, then its image $f(X)$ is compact.

From the assertions of Problems 3.3 and 3.4 we obtain the following property:

**Problem 3.5.** If $f: X \to \mathbb{R}$ is a continuous function and $X$ is a compact space, then $f$ attains its maximum and minimum values: there exist points $x_{\min}, x_{\max} \in X$ such that $f(x_{\min}) \leq f(x) \leq f(x_{\max})$ for all $x \in X$.

The above-listed properties (the Hausdorff property, connectedness, path connectedness, compactness) are topological invariants; i.e., they remain the same for spaces which cannot be distinguished as topological spaces without any additional structure. Formally, this notion of identity can be worded as follows.

- The spaces $X$ and $Y$ are *homeomorphic* if there exist continuous maps $f: X \to Y$ and $g: Y \to X$ which are inverse to each other: $gf: X \to X$ and $fg: Y \to Y$ are the identity maps of $X$ and $Y$. Such maps $f$ and $g$ are called *homeomorphisms*.
- A property is *topologically invariant* if the fact that it holds for a space $X$ implies that it also holds for every space $Y$ which is homeomorphic to $X$.

## 3.2 Smooth manifolds and maps

A *topological $n$-dimensional manifold* is a Hausdorff space $M$ such that each point of $M$ has a neighborhood homeomorphic to a domain in $\mathbb{R}^n$.

An open covering $\{U_\alpha\}$ of a manifold $M$ such that, for each element $U_\alpha$, there is a homeomorphism $\varphi_\alpha$ between $U_\alpha$ and a domain $W_\alpha$ in $\mathbb{R}^n$:

$$\varphi_\alpha: W_\alpha \to U_\alpha,$$

is called an *atlas*. Each homeomorphism $\varphi_\alpha$ determines *local coordinates* in the domain $U_\alpha$ called the *chart*. Namely, if $x = (x^1, \ldots, x^n) \in \mathbb{R}^n$ then the local coordinates $x^1_\alpha, \ldots, x^n_\alpha$ of a point $\varphi_\alpha(x)$ are $x^1, \ldots, x^n$. In the intersection of coordinate domains $U_\alpha$ and $U_\beta$ the local coordinates are connected by the *transition maps*:

$$x^i_\alpha = f^i_{\alpha\beta}(x^1_\beta, \ldots, x^n_\beta),$$

where $f_{\alpha\beta} = \varphi_\alpha \varphi_\beta^{-1}$.

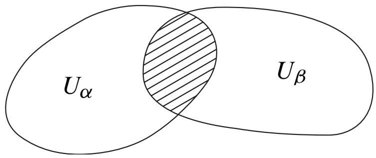


Figure 3.1. Overlapping charts.

We say that a topological manifold is endowed with a *smooth structure* of class $C^k$ if it has an atlas in which all transition maps are $k$ times continuously differentiable. A manifold with such covering is called $C^k$ *smooth*[1] (or *differentiable*), and the corresponding local coordinates are called *smooth*. Henceforth for simplicity by smoothness we mean $C^\infty$ smoothness.

The maps $f_{\alpha\beta}$ are invertible, since their composition $f_{\alpha\beta} f_{\beta\alpha}$ is the identity map. Consequently, their Jacobian determinants are nonzero everywhere:

$$\det\left(\frac{\partial f^i_{\alpha\beta}}{\partial x^j_\beta}\right) = \det\left(\frac{\partial x^i_\alpha}{\partial x^j_\beta}\right) \neq 0.$$

**Example.** Recall that the *sphere* of dimension $n$ is the smooth submanifold in $\mathbb{R}^{n+1}$ given by the equation

$$(x^1)^2 + \cdots + (x^{n+1})^2 = 1.$$

Construct its atlas with two charts. Let $P_{\mathrm{N}}$ be the "north" pole; i.e., the point with coordinates $(0, \ldots, 0, 1)$, and let $P_{\mathrm{S}}$ be the "south pole" $(0, \ldots, 0, -1)$. On the complement to each pole define the projection from the given pole to the plane $x^{n+1} = 0$: we join the pole and the point on the sphere by a straight line. The projection of the given point is the intersection of this straight line with the plane $x^{n+1} = 0$. These projections are the sought maps $\varphi_{\mathrm{N}}$ and $\varphi_{\mathrm{S}}$:

$$\varphi_{\mathrm{N}} \colon U_{\mathrm{N}} = S^2 \setminus P_{\mathrm{N}}, \quad \varphi_{\mathrm{N}}(x^1, \ldots, x^{n+1}) = \left(\tfrac{x^1}{1-x^{n+1}}, \ldots, \tfrac{x^n}{1-x^{n+1}}\right)$$

(the projection from the north pole),

$$\varphi_{\mathrm{S}} \colon U_{\mathrm{S}} = S^2 \setminus P_{\mathrm{S}}, \quad \varphi_{\mathrm{S}}(x^1, \ldots, x^{n+1}) = \left(\tfrac{x^1}{1+x^{n+1}}, \ldots, \tfrac{x^n}{1+x^{n+1}}\right)$$

(the projection from the south pole). The charts $U_{\mathrm{N}}$ and $U_{\mathrm{S}}$ overlap over the domain $S^2 \setminus \{P_{\mathrm{N}}, P_{\mathrm{S}}\}$ and the formulas for the transition map from $U_{\mathrm{N}}$ to $U_{\mathrm{S}}$ has the form

$$f_{\mathrm{SN}}(x^1, \ldots, x^n) = \left(\tfrac{x^1}{|x|^2}, \ldots, \tfrac{x^n}{|x|^2}\right), \quad |x|^2 = (x^1)^2 + \cdots + (x^n)^2$$

(we leave the derivation of this formula as an exercise).

[1]It is known that even in dimension $n = 4$ there are compact topological manifolds which admit no smooth structure.

By the chain rule, if a function $f$ given on the intersection $U_\alpha \cap U_\beta$ is of class $C^k$ in the variable $x^i_\alpha$, then it has the same smoothness in the variables $x^j_\beta$. This allows us to define the notion of a *smooth* or *differentiable map*:

- a map $f\colon M \to N$ between smooth manifolds is $C^k$ smooth if, for smooth local coordinates $\{x^i\}$ on $M$ and $\{y^j\}$ on $N$, it is given by a $C^k$ smooth vector function $(y^1, \dots, y^m) = f(x^1, \dots, x^n)$.

Smooth maps $f\colon M \to \mathbb{R}$ are called *smooth functions*, and smooth maps $\gamma\colon [a,b] \to M$ are called *smooth curves*.

Note that the definition of smoothness of a map $f\colon M \to N$ is based on smooth structures on both $M$ and $N$. If some topological manifold $X$ carries two smooth structures, then we obtain two smooth manifolds $M_1$ and $M_2$. The smooth structures are assumed to coincide if the identity maps $M_1 \to M_2$ and $M_2 \to M_1$ are smooth.

Manifolds $M$ and $N$ are *diffeomorphic* if there exist smooth maps $f\colon M \to N$ and $g\colon N \to M$ which are mutually inverse; i.e., $fg\colon N \to N$ and $gf\colon M \to M$ are the identity maps. Such smooth homeomorphisms $f$ and $g$ are called *diffeomorphisms*.[2]

Let $\gamma(t)$ be a smooth curve in $M$. In local coordinates $\{x^i_\alpha\}$ the curve $\gamma(t)$ is written in the form

$$t \to (x^1_\alpha(t), \dots, x^n_\alpha(t))$$

and its velocity vector at a point $\gamma(t)$ is equal to

$$\dot\gamma(t) = (\dot x^1_\alpha(t), \dots, \dot x^n_\alpha(t)).$$

In different local coordinates $\{x^i_\beta\}$ the curve $\gamma(t)$ and its velocity vector $\dot\gamma(t)$ have the form

$$t \to (x^1_\beta(t), \dots, x^n_\beta(t)), \quad \dot\gamma(t) = (\dot x^1_\beta(t), \dots, \dot x^n_\beta(t)).$$

Hence we derive the following formula which connects the velocity vectors in different local coordinates:

$$\dot\gamma(t) = (\dot x^1_\beta(t), \dots, \dot x^n_\beta(t)) = \left(\frac{\partial x^1_\beta}{\partial x^j_\alpha}\dot x^j_\alpha(t), \dots, \frac{\partial x^n_\beta}{\partial x^j_\alpha}\dot x^j_\alpha(t)\right),$$

which implies

$$\dot x^i_\beta = \frac{\partial x^i_\beta}{\partial x^j_\alpha}\dot x^j_\alpha.$$

[2]It is known that there are infinitely many pairwise nondiffeomorphic smooth manifolds homeomorphic to $\mathbb{R}^4$ (for $n \neq 4$ such manifolds are unique up to diffeomorphism), while on the topological space $S^7$ homeomorphic to the unit seven-dimensional sphere in $\mathbb{R}^8$ ($S^7 = \{v \in \mathbb{R}^8 \mid |v| = 1\}$) there are exactly twenty-eight different (up to diffeomorphism) smooth structures. All these structures on $S^7$ are realized by the submanifolds of $\mathbb{R}^{10}$ (see Lemma 3.1) given by the equations

$$z_1^{6k-1} + z_2^3 + z_3^2 + z_4^2 + z_5^2 = 0,$$
$$|z_1|^2 + |z_2|^2 + |z_3|^2 + |z_4|^2 + |z_5|^2 = 1,$$

for $k = 1, \dots, 28$ (here $z_1, \dots, z_5$ are the complex-valued coordinates in $\mathbb{C}^5 \approx \mathbb{R}^{10}$); see, e.g., [3].

The velocity vectors of curves in $M$ are tangent vectors to $M$ and the vector $\dot\gamma(t)$ is attached to the point $\gamma(t)$. We arrive at the following definition:

- a *tangent vector* to an $n$-dimensional manifold $M$ at a point $x$ is an object which is given in local coordinates $\{x^i_\alpha\}$ by an ordered collection of numbers $(v^1_\alpha, \dots, v^n_\alpha)$ and its expression $(w^1_\beta, \dots, w^n_\beta)$ in other local coordinates $\{x^j_\beta\}$ satisfies the equation
$$w^j_\beta = \frac{\partial x^j_\beta}{\partial x^i_\alpha} v^i_\alpha. \tag{3.1}$$

The tangent vectors at a point $x$ constitute an $n$-dimensional vector space which is called the *tangent space* $T_xM$ at $x$. Each coordinate system determines the basis for the tangent space denoted by

$$\partial_1 = \frac{\partial}{\partial x^1}, \dots, \partial_n = \frac{\partial}{\partial x^n}. \tag{3.2}$$

The velocity vector $\dot x$ is decomposed in this basis by the formula

$$\dot x = \dot x^1 \partial_1 + \cdots + \dot x^n \partial_n.$$

A manifold is *orientable* if there exist coordinate domains $U_\alpha$ covering the whole manifold and such that on the intersection $U_\alpha \cap U_\beta$ of every pair of domains the following inequality holds:

$$\det\left(\frac{\partial x^i_\alpha}{\partial x^j_\beta}\right) > 0.$$

Suppose that a manifold is connected and orientable. Then, assuming that the basis $\frac{\partial}{\partial x^1_\alpha}$, $\dots$, $\frac{\partial}{\partial x^n_\alpha}$ is positively or negatively oriented, we define the *orientation* in all tangent spaces, assuming that the bases $\frac{\partial}{\partial x^1_\beta}, \dots, \frac{\partial}{\partial x^n_\beta}$ are positively or negatively oriented, too. Once some orientation is chosen, the manifold is called *oriented.*

A smooth map $f : M \to N$ induces the linear map

$$f_* \colon T_xM \to T_{f(x)}N$$

of the tangent spaces under which the velocity vector of the curve $\gamma(t)$ goes into the velocity vector of the curve $f(\gamma(t))$. Choose local coordinates $\{x^i\}$ in a neighborhood of $x$ and local coordinates $\{y^j\}$ in a neighborhood of $f(x)$. Then the map $f$ takes the form

$$y^j = f^j(x^1, \dots, x^n)$$

and we obtain

$$\dot f^j(\gamma(t)) = \frac{\partial f^j}{\partial x^i} \dot x^i(t).$$

It means that in the given coordinates the map $f_*$ has the form

$$v^i \to w^j = \frac{\partial f^j}{\partial x^i} v^i.$$

Hence, in particular, we see that it is linear.

A subspace $N$ is a $k$-dimensional *submanifold* in $M$ if each point $x \in N$ has a neighborhood $U$ with local coordinates $x^1, \dots, x^n$ such that the intersection $U \cap N$ is determined by the equations $x^1 = \dots = x^{n-k} = 0$. Moreover, considering $y^1 = x^{n-k+1}, \dots, y^k = x^n$ as local coordinates on $N$, we define the structure of a smooth manifold on $N$.

We introduce the following important notion:

- a smooth map $F\colon M \to N$ is *regular* at $x \in M$ if the rank of the Jacobian matrix of $F$ at $x$ written in some (hence arbitrary) local coordinates $y^j = F^j(x^1, \dots, x^n)$, $j = 1, \dots, k$, is equal to the dimension of $N$:

$$\operatorname{rank}\left(\frac{\partial F^j}{\partial x^i}\right) = \dim N.$$

  If this condition is satisfied, then $x$ is called a *regular* point of the map $F$.

Examples of submanifolds are regular zero sets of smooth maps.

**Lemma 3.1.** *Suppose that $F\colon M \to \mathbb{R}^{n-k}$ is a smooth map on an $n$-dimensional manifold $M$ and its zero set $M_0 = F^{-1}(0)$ consists of regular points. Then $M_0$ is a smooth submanifold in $M$.*

*Proof.* Take $x \in M_0$. In a neighborhood of $x$ we have

$$\operatorname{rank}\left(\frac{\partial F^j}{\partial x^i}\right) = n - k,$$

where $\{x^i\}$ are the coordinates on $M$. Without loss of generality we may assume that the minor of the matrix corresponding to the coordinates $x^1, \dots, x^{n-k}$ at $x$ is nonzero. By the implicit function theorem, in a small neighborhood $U \subset M$ of $x$ we can define functions $\psi_1, \dots, \psi_{n-k}$ such that the equality $F(x) = 0$ holds if and only if $x^1 = \psi_1(x^{n-k+1}, \dots, x^n), \dots, x^{n-k} = \psi_{n-k}(x^{n-k+1}, \dots, x^n)$. Now, let

$$\hat{x}^1 = x^1 - \psi_1(x^{n-k+1}, \dots, x^n), \dots, \hat{x}^{n-k} = x^{n-k} - \psi_1(x^{n-k+1}, \dots, x^n),$$
$$\hat{x}^{n-k+1} = x^{n-k+1}, \dots, \hat{x}^n = x^n$$

be local coordinates in a neighborhood $V$ of $x$. With these coordinates, the set $V \cap M_0$ is determined by the equations $\hat{x}^1 = \dots = \hat{x}^{n-k} = 0$. Lemma 3.1 is proven. □

This lemma yields a great source of examples of smooth manifolds which can be constructed as submanifolds of other, already known, manifolds. For $M = \mathbb{R}^n$ we obtain smooth submanifolds of the Euclidean space (see §1.4).

**Example** (Spheres). The zero set of the smooth map

$$f(x^1, \ldots, x^{n+1}) = (x^1)^2 + \cdots + (x^{n+1})^2 - 1$$

of the Euclidean space $\mathbb{R}^{n+1}$ to $\mathbb{R}$ represents a smooth hypersurface, namely the $n$-dimensional sphere $S^n$. Every smooth manifold diffeomorphic to this sphere is called an $n$-dimensional sphere, too.

Let $f: N \to M$ be a smooth map such that

(1) $f$ is a homeomorphism between $N$ and $f(N) \subset M$;

(2) at each point $x \in N$ the map $f_*$ is an embedding of tangent spaces.

Such a map $f$ is called an *embedding* of $N$ into $M$. If only condition (2) is satisfied, then the map $f$ is called an *immersion*.

**Lemma 3.2.** *If $f: N \to M$ is an embedding, then $f(N)$ is a submanifold in $M$ and $f: N \to f(N)$ is a diffeomorphism.*

*Proof.* In local coordinates $\{y^j\}$ on $N$ and $\{x^j\}$ on $M$ the embedding $f$ takes the form $x^j = x^j(y^1, \ldots, y^n)$ and, since the point $x$ is regular for $f$,

$$\operatorname{rank}\left(\frac{\partial x^j}{\partial y^i}\right) = k.$$

Without loss of generality we may assume that

$$\det\left(\frac{\partial x^j}{\partial y^i}\right)_{1 \le i,j \le k} \neq 0$$

in a neighborhood of $x$. By the inverse function theorem, in a small neighborhood $U$ of $f(x)$ we can define smooth functions $\varphi_1, \ldots, \varphi_n$ such that $y^j = \varphi_j(x^1, \ldots, x^k)$ for $1 \le j \le k$ if and only if $x^j = x^j(y^1, \ldots, y^k)$. Now, take

$$F: U \to \mathbb{R}^{n-k}$$

in the form

$$F^j(x^1, \ldots, x^n) = x^{k+j} - x^{k+j}(\varphi_1(x^1, \ldots, x^k), \ldots, \varphi_k(x^1, \ldots, x^k)).$$

By construction, the zero set of the map $F$ coincides with $f(N) \cap U$ and $F$ is regular on the latter set. Applying Lemma 3.1, we complete the proof of Lemma 3.2. □

Suppose that $f: M \to \mathbb{R}$ is a smooth function on a manifold $M$ and its zero set $M_0$ consists of regular points. Then $M_0$ is a submanifold which divides $M$ into two parts, where $f < 0$ and $f > 0$. In this case the closed domain $N$ selected by the inequality $f(x) \ge 0$ is called a *manifold with boundary* $\partial N = M_0$. Removing from

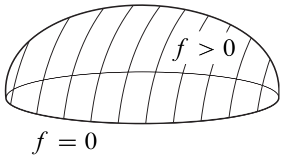


Figure 3.2. Boundary of a manifold.

the manifold $N$ with boundary its boundary $\partial N$, we obtain a submanifold $N \setminus \partial N \subset M$ whose dimension coincides with that of $M$ and is called the dimension of a manifold with boundary. If a manifold with boundary is $n$-dimensional, then its boundary is an $(n-1)$-dimensional manifold without boundary.

A map $f: N_1 \to N_2$ between manifolds $N_1 \subset M_1$ and $N_2 \subset M_2$ with boundary determined in $M_1$ and $M_2$ by inequalities is called *smooth* if there is an open domain $U \subset M_1$ such that $N_1 \subset U$ and $f$ extends to a smooth map $g: U \to M_2$.

By analogy with the case of manifolds without boundary, we define a diffeomorphism of manifolds with boundary. Note that if $N$ is an $n$-dimensional manifold with boundary, then each point of its boundary has a neighborhood diffeomorphic to the intersection of the $n$-dimensional ball $\{|x| < 1\} \subset \mathbb{R}^n$ and the half-space $x^n \geq 0$.

If a manifold has no boundary and is compact, then it is called *closed*.

Define a smooth structure on the direct product of smooth manifolds. Let $\{U_\alpha\}$ be a covering of $M$ by domains with coordinates $\{x^i_\alpha\}$ and let $\{V_\beta\}$ be a covering of $N$ by domains with coordinates $\{y^j_\beta\}$. Then $\{U_\alpha \times V_\beta\}$ is a covering of $M \times N$ by domains with smooth coordinates $\{x^1_\alpha, \dots, x^k_\alpha, y^1_\beta, \dots, y^l_\beta\}$ (here $k = \dim M$ and $l = \dim N$). Henceforth by a smooth manifold $M \times N$ we mean the manifold described above.

**Examples.** (1) *Tori*. The direct product of $n$ copies of the circle $S^1$ is called the $n$-dimensional torus:

$$T^n = \underbrace{S^1 \times \cdots \times S^1}_{n}.$$

A two-dimensional torus is obtained if we take the square $\{0 \leq x, y \leq 1\}$ on the plane and glue the opposite sides by the rule:

$$(0, y) \sim (1, y), \quad (x, 0) \sim (x, 1).$$

The values $x$ and $y$ are coordinates on the so-obtained torus defined modulo 1.

(2) *The Klein bottle*. Take the square $\{0 \leq x, y \leq 1\}$ on the plane and glue the opposite sides by the rule

$$(0, y) \sim (1, y), \quad (x, 0) \sim (1 - x, 1).$$

We assume that in a sufficiently small neighborhood of every point of the so-obtained space $M^2$ the values $x$ and $y$ determine smooth local coordinates. The so-obtained closed manifold is called the Klein bottle.

**Problem 3.6.** Prove that the Klein bottle is not orientable, since under a continuous change of the orientation in the tangent spaces along the loop $r(t) = (1/2, t)$, where $t = [0, 1]$, the orientation changes.

**Problem 3.7.** Prove that the inequality $\frac{1}{4} \leq x \leq \frac{3}{4}$ distinguishes a submanifold in the Klein bottle which is not orientable and whose boundary consists of one component homeomorphic to the circle.[3]

**Problem 3.8.** Let $\mathbb{R}_1$ and $\mathbb{R}_2$ be two straight lines with respective coordinates $x$ and $y$. Let $X$ be the set obtained from these straight lines by identification of the points $x \in \mathbb{R}_1$ and $y \in \mathbb{R}_2$ for $x = y$ and $x \neq 0$. The straight line $\mathbb{R}_j$ is naturally embedded into $X$: $p_j : \mathbb{R}_j \to X$. Introduce the topology on $X$, by assuming that $U \subset X$ is open if and only if its inverse images $p_1^{-1}(U) \subset \mathbb{R}_1$ and $p_2^{-1}(U) \subset \mathbb{R}_2$ are open. Show that $X$ possesses all properties of a smooth manifold, except the Hausdorff property.

## 3.3 Tensors

Smooth manifolds have the structure of Euclidean space near each point, and functions on manifolds are smooth if they are smooth functions of local coordinates.

The expressions $v_\alpha$ and $w_\beta$ of the same tangent vector in different local coordinates $\{x_\alpha\}$ and $\{x_\beta\}$ are connected by (3.1):

$$w_\beta^j = \frac{\partial x_\beta^j}{\partial x_\alpha^i} v_\alpha^i.$$

Let $f : M \to \mathbb{R}$ be a smooth function on $M$. In the coordinates $\{x_\alpha^i\}$ its gradient grad $f$ has the form

$$v^\alpha = \left(\frac{\partial f}{\partial x_\alpha^1}, \dots, \frac{\partial f}{\partial x_\alpha^n}\right),$$

and in the coordinates $\{x_\beta^i\}$, the form

$$w^\beta = \left(\frac{\partial f}{\partial x_\beta^1}, \dots, \frac{\partial f}{\partial x_\beta^n}\right) = \left(\frac{\partial f}{\partial x_\alpha^i}\frac{\partial x_\alpha^i}{\partial x_\beta^1}, \dots, \frac{\partial f}{\partial x_\alpha^i}\frac{\partial x_\alpha^i}{\partial x_\beta^n}\right).$$

The formula which connects the expressions of the gradient in different local coordinates differs from the formula (3.1) for vectors.

A *covector* at a point $x$ is an object given in local coordinates $\{x_\alpha^i\}$ by an ordered collection of numbers $(v_1^\alpha, \dots, v_\alpha^n)$ and the corresponding collection $(w_1^\beta, \dots . w_\beta^n)$ for another coordinate system $\{x_\beta^j\}$ satisfies the equation

$$w_j^\beta = \frac{\partial x_\alpha^i}{\partial x_\beta^j} v_i^\alpha. \tag{3.3}$$

[3] This manifold with boundary is called the *Möbius strip*.

To stress the difference between vectors and covectors we use in notation superscripts (upper indices) for vectors and subscripts (lower indices) for covectors.

Now we can make the following conclusion:

**Lemma 3.3.** *The gradient of a function is a covector.*

The derivative of a function in the direction of a tangent vector $v$ is equal to $D_v f = (\text{grad } f)_i v^i$, is independent of the local coordinates, and is a linear function on the tangent space. The last assertion is valid for all covectors.

**Lemma 3.4.** *A covector $w$ at $x \in M$ is a linear function on the tangent space $T_x M$ which is written in local coordinates by the formula $\langle w, v \rangle = w_i v^i$.*

*Proof.* It suffices to show that the value $\langle w, v \rangle$ is independent of the choice of local coordinates. It follows from (3.1) and (3.3) that

$$w_i^\beta v_\beta^i = \frac{\partial x_\alpha^j}{\partial x_\beta^i} w_j^\alpha \frac{\partial x_\beta^i}{\partial x_\alpha^k} v_\alpha^k,$$

and, since

$$\frac{\partial x_\alpha^j}{\partial x_\beta^i} \frac{\partial x_\beta^i}{\partial x_\alpha^k} = \frac{\partial x_\alpha^j}{\partial x_\alpha^k} = \delta_k^j,$$

we obtain

$$\langle w, v \rangle = w_i^\beta v_\beta^i = w_k^\alpha v_\alpha^k.$$

Lemma 3.4 is proven. □

Lemma 3.4 yields the following assertion:

**Lemma 3.5.** *Covectors at a point $x \in M$ constitute a vector space $T_x^* M$ of dimension $n = \dim M$ dual to the tangent space $T_x M$.*

The space $T_x^* M$ is called the *cotangent* space at $x \in M$.

A generalization of vectors and covectors is the notion of tensor:

- a *tensor* of type $(k, l)$ at $x \in M$ is an object

  $$T_{j_1 \dots j_l}^{i_1 \dots i_k}, \quad i_1, \dots, i_k, \ j_1, \dots, j_l = 1, \dots, \ n = \dim M,$$

  given for all values of the superscripts $i$ and the subscripts $j$ such that its expressions in different local coordinates are connected by the formula

  $${}^{(\beta)}T_{j_1 \dots j_l}^{i_1 \dots i_k} = \frac{\partial x_\beta^{i_1}}{\partial x_\alpha^{r_1}} \cdots \frac{\partial x_\beta^{i_k}}{\partial x_\alpha^{r_k}} \frac{\partial x_\alpha^{s_1}}{\partial x_\beta^{j_1}} \cdots \frac{\partial x_\alpha^{s_l}}{\partial x_\beta^{j_l}} \, {}^{(\alpha)}T_{s_1 \dots s_l}^{r_1 \dots r_k}, \tag{3.4}$$

- a family of tensors of the same type depending continuously on $x \in M$ is called a *tensor field* on $M$.

A particular case of a tensor field is a *vector field.* If a tensor (in particular, vector) field depends smoothly on the point, i.e., all functions $T^{i_1 \dots i_k}_{j_1 \dots j_k}$ are smooth, then we say that the field is *smooth*.

We see that

(1) a tangent vector is a tensor of type $(1, 0)$;

(2) the gradient of a function is a tensor of type $(0, 1)$.

Since vectors $v$ are linear functions on covectors $w$ of the form $v(w) = w(v)$, tensors $T$ of type $(k, l)$ are linear functions of $k$ covectors $u^{(m)}$ and $l$ vectors $v_{(n)}$:

$$T(u^{(1)}, \dots, u^{(k)}, v_{(1)}, \dots, v_{(l)}) = T^{i_1 \dots i_k}_{j_1 \dots j_l} u^{(1)}_{i_1} \dots u^{(k)}_{i_k} v^{j_1}_{(1)} \dots v^{j_l}_{(l)}.$$

Now, formula (3.4) is derived componentwise from formulas (3.1) and (3.3) which follow for tangent vectors and gradients of functions from the chain rule.

Formula (3.4) is obvious for tensor products of tensors of types $(1, 0)$ and $(0, 1)$: the *tensor product* of tensors $A$ and $B$ of the respective types $(k, l)$ and $(p, q)$ is the tensor $A \otimes B$ of type $(k + p, l + q)$ such that

$$(A \otimes B)^{i_1 \dots i_k r_1 \dots r_p}_{j_1 \dots j_l s_1 \dots s_q} = A^{i_1 \dots i_k}_{j_1 \dots j_l} \cdot B^{r_1 \dots r_p}_{s_1 \dots s_q}.$$

We have already indicated the basis

$$e_1 = \partial_1, \ \dots, \ e_n = \partial_n$$

of the form (3.2) for the space of vectors. Its associated dual basis for the space of covectors is

$$e^1 = dx^1, \ \dots, \ e^n = dx^n.$$

It is uniquely determined by the values of its elements at the basis vectors:

$$e^i(e_j) = \delta^i_j.$$

The tensor products of the form

$$e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_l}$$

determine a basis for the space of tensors of type $(k, l)$ at every point. For such tensors we obviously have the decomposition

$$T = T^{i_1 \dots i_k}_{j_1 \dots j_l} e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_l}.$$

If an inner product $(v_1, v_2)$ is given in each tangent space $T_x M$ (for surfaces in $\mathbb{R}^3$ this is the first fundamental form) then we obtain the *metric*

$$g_{ij} e^i \otimes e^j = g_{ij} dx^i dx^j,$$

which is a tensor of type $(0, 2)$:

$$(v_1, v_2) = g_{ij} v_1^i v_2^j, \quad g_{ij} = (e_i, e_j).$$

The tangent vectors, i.e., pairs $(x, v)$ with $x \in M$ and $v \in T_x M$, constitute the *tangent bundle* $TM$.

**Theorem 3.1.** *The tangent bundle $TM$ of an $n$-dimensional manifold $M$ carries the structure of a smooth manifold such that*

(a) *the projection $\pi : TM \to M$ is a smooth map;*

(b) *every point $x \in M$ has a neighborhood $U$ such that the inverse image $\pi^{-1}(U)$ of this neighborhood is diffeomorphic to the direct product $U \times \mathbb{R}^n$: $f : \pi^{-1}(U) \to U \times \mathbb{R}^n$; moreover, $\pi(f^{-1}(x, v)) = x$, where $x \in U$ and $v \in \mathbb{R}^n$* (*the diffeomorphism $f$ agrees with the projection*).

*Proof.* For each point $x \in M$ take some coordinate neighborhood $U_\alpha$ with coordinates $(x_\alpha^1, \dots, x_\alpha^n)$. Let $W_\alpha$ be the set of tangent vectors $v$ attached to points of $U$. Introduce the coordinates $(x_\alpha^1, \dots, x_\alpha^n, v_\alpha^1, \dots, v_\alpha^n)$ on $W_\alpha$, where $v \in T_x M$ has the decomposition $v = v_\alpha^i \partial_i$ in the basis (3.2).

The family $W_\alpha$ constitutes a covering of $TM$ and, assuming that the coordinates $(x_\alpha^1, \dots, x_\alpha^n, v_\alpha^1, \dots, v_\alpha^n)$ are smooth, we introduce the sought structure of a smooth manifold on $TM$. Theorem 3.1 is proven. □

Henceforth by the smooth structure on $TM$ we mean the structure constructed in the proof of Theorem 3.1.

Similarly, we can prove that the set of covectors attached to points of $M$, i.e., the *cotangent bundle* $T^*M$ of $M$, carries the structure of a smooth manifold such that the projection $\pi : T^*M \to M$ is a smooth map and every point $x \in M$ has a neighborhood $U$ such that the inverse image $\pi^{-1}(U)$ of this neighborhood is diffeomorphic to the direct product $U \times \mathbb{R}^n$, where $n = \dim M$.

A tensor $T_{i_1 \dots i_k}$ of type $(0, k)$ is *antisymmetric* if its value changes sign as arbitrary two indices interchange:

$$T_{i_1 \dots i_p \dots i_q \dots i_k} = -T_{i_1 \dots i_q \dots i_p \dots i_k}.$$

There is a one-to-one correspondence between the antisymmetric tensors of type $(0, k)$ and differential $k$-forms:

$$T \leftrightarrow \omega_T = \sum_{i_1 < \dots < i_k} T_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

Indeed, we have

$$dx^1 \wedge \dots \wedge dx^k = e^1 \wedge \dots \wedge e^k = \sum_{\iota \in S_k} (-1)^{\sigma(\iota)} e^{i_1} \otimes \dots \otimes e^{i_k},$$

where $S_k$ is the permutation group of a $k$-element set and $\sigma(\iota)$ equals 0 or 1 depending on whether the permutation $\iota = \left(\begin{smallmatrix} 1 & \dots & k \\ i_1 & \dots & i_k \end{smallmatrix}\right)$ is even or odd. Hence,

$$\omega_T = T_{i_1 \dots i_k} dx^{i_1} \otimes \dots \otimes dx^{i_k},$$

where the repeated indices $i_1, \dots, i_k$ imply summation.

Moreover, the *exterior* (or *wedge*) product of forms $\omega_1 = \alpha dx^{i_1} \wedge \dots \wedge dx^{i_k}$ and $\omega_2 = \beta dx^{j_1} \wedge \dots \wedge dx^{j_l}$ is defined as

$$\omega_1 \wedge \omega_2 = \alpha\beta dx^{i_1} \wedge \dots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \dots \wedge dx^{j_l};$$

moreover, $dx^{p_1} \wedge \dots \wedge dx^{p_m} = 0$ (by definition) if some of the indices $p_1, \dots, p_m$ coincide. This product extends to all differential forms by linearity.

Once a Riemannian metric $g_{ij} dx^i dx^j$ is given on an oriented manifold $M^n$, we can define the *volume form*

$$d\mu = \sqrt{g}\, dx^1 \wedge \dots \wedge dx^n, \quad g = \det(g_{ij}),$$

which behaves like an antisymmetric tensor of type $(0, n)$ under orientation-preserving changes of coordinates. In general relativity, the volume form for pseudo-Riemannian metrics of type $(1, 3)$ is defined as

$$d\mu = \sqrt{-g}\, dx^1 \wedge \dots \wedge dx^n.$$

This form is needed for integration: namely, by definition, the *integral* of a function $f(x)$ over the manifold is defined to be the quantity

$$\int_{M^n} f(x)\, d\mu.$$

**Problem 3.9.** Let $\omega_1$ be a $k$-form and let $\omega_2$ be an $l$-form. Prove that

$$\omega_1 \wedge \omega_2 = (-1)^{kl} \omega_2 \wedge \omega_1.$$

**Problem 3.10.** With each vector field $v$ we can associate the operator $i_v$ which takes $n$-forms into $(n-1)$-forms by the following rule:

$$i_v \omega(\eta_1, \dots, \eta_{n-1}) = \omega(v, \eta_1, \dots, \eta_{n-1})$$

(or $[i_v \omega]_{j_1 \dots j_{n-1}} = v^j \omega_{j j_1 \dots j_{n-1}}$). Prove that this operator is anti-derivation:

$$i_v(\omega_1 \wedge \omega_2) = (i_v \omega_1) \wedge \omega_2 + (-1)^k \omega_1 \wedge (i_v \omega_2),$$

where $\omega_1$ is a $k$-form.

**Problem 3.11.** Suppose that a linear operator $T \colon V \to V$ is defined in some coordinates in $V$ by values $T^i_j$:

$$T\xi = \eta, \quad \eta^i = T^i_j \xi^j.$$

Prove that these values constitute a tensor of type $(1, 1)$.

## 3.4 Action of maps on tensors

In §3.2 we indicated how a smooth map

$$f\colon M^n \to N^k, \quad y^1 = (x^1, \dots, x^n), \ \dots, \ y^k = y^k(x^1, \dots, x^n),$$

acts on tangent vectors. The map

$$f_*\colon T_x M^n \to T_{f(x)} N^k$$

of tangent spaces takes the velocity vector $v$ of a curve $\gamma(t)$ at a point $x = \gamma(0)$ into the velocity vector of the curve $\tilde{\gamma}(t) = f(\gamma(t))$ at $t = 0$ and has the form

$$v \to w\colon \ w^i = \frac{\partial y^i}{\partial x^j} v^j. \tag{3.5}$$

Since covectors are linear functions of vectors, the map $f$ acts on covectors in the reverse direction by the rule

$$f^*\colon T^*_{f(x)} N^k \to T^*_x M^n\colon \ \langle f^*(\eta), v\rangle = \langle \eta, f_*(v)\rangle,$$

or in the coordinate notation

$$\eta_i(y) dy^i \to \eta_i(y(x)) dy^i(x) = \eta_i(y(x)) \frac{\partial y^i}{\partial x^j} dx^j,$$

i.e.,

$$\eta \to \xi\colon \ \eta_i \to \xi_j = \frac{\partial y^i}{\partial x^j} \eta_i. \tag{3.6}$$

Although formulas (3.5) and (3.6) look similar, they are essentially different in one respect: the left-hand side in (3.5) and the right-hand side in (3.6) are a vector $v$ and a covector $\xi$ attached to $x \in M^n$, while the right-hand side in (3.5) and the left-hand side in (3.6) are a vector and a covector attached to $f(x)$.

The linear map $f^*$ extends naturally to tensors of type $(0, m)$:

$$T(y) dy^{i_1} \otimes \dots \otimes dy^{i_m} \to T(y(x)) \frac{\partial y^{i_1}}{\partial x^{j_1}} \dots \frac{\partial y^{i_m}}{\partial x^{j_m}} dx^{j_1} \otimes \dots \otimes dx^{j_m}$$

and, in particular, to antisymmetric tensors, i.e., differential forms:

$$T(y) dy^{i_1} \wedge \dots \wedge dy^{i_m} \to T(y(x)) \frac{\partial y^{i_1}}{\partial x^{j_1}} \dots \frac{\partial y^{i_m}}{\partial x^{j_m}} dx^{j_1} \wedge \dots \wedge dx^{j_m}, \tag{3.7}$$

and relates to an $l$-form $\omega$ on $N^k$ the form $f^*\omega$ of the same type on $M^n$.

**Problem 3.12.** Prove the formula

$$f^*(\omega_1 \wedge \omega_2) = f^*\omega_1 \wedge f^*\omega_2.$$

**Example** (Induced metric). Let $g_{ij}dy^i dy^j$ be a Riemannian metric on $N^k$ and let $f: M^n \to N^k$ be an embedding. The action of $f^*$ on the metric tensor $g_{ij}$ determines the induced metric $g_{ij}(f(x))\frac{\partial y^i}{\partial x^p}\frac{\partial y^j}{\partial x^q}dx^p dx^q$. In the case $n = 2$ and $N^k = \mathbb{R}^3$ we obtain the first fundamental form on a surface.

Recall some facts from calculus: the *exterior derivative* of a $k$-form

$$\omega = \sum_{i_1<\dots<i_k} T_{i_1\dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k}$$

is the $(k+1)$-form

$$d\omega = \sum_{i_1<\dots<i_k} \frac{\partial T_{i_1\dots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

A form $\omega$ is *closed* if $d\omega = 0$. If $\omega = d\eta$ for some form $\eta$, then the form $\omega$ is *exact*.

**Problem 3.13.** Prove that all closed forms are exact:

$$d^2 = 0.$$

**Problem 3.14.** Prove the formula

$$f^* d\omega = df^*\omega.$$

We introduce one more important definition: a topological space $X$ is *simply connected* if every continuous map

$$f: S^1 \to X$$

of the circle $S^1 = \{x^2 + y^2 = 1\}$ (in the plane $\mathbb{R}^2$) extends to a continuous map

$$f: D^2 \to X$$

of the disk $D^2 = \{x^2 + y^2 \le 1\}$. This property is especially important for the following reason:

- *on a simply connected smooth manifold $M^n$, the value of the integral of a closed 1-form $\omega$ along a curve $\gamma: [0, T] \to M^n$ depends only on the endpoints of the curve:*

$$\int_0^T \gamma^*\omega = f(P, Q), \quad P = \gamma(0),\ Q = \gamma(T).$$

Indeed, if two smooth curves $\gamma_1$ and $\gamma_2$ connect two points $P$ and $Q$, then they together constitute a map $f: S^1 \to X$ which determines the curve $\gamma_1$ traversed clockwise for $y \ge 0$ and the curve $\gamma_2$ traversed counter-clockwise for $y \le 0$, and we obtain

$$\int_{S^1} f^*\omega = \int_{\gamma_1} f^*\omega - \int_{\gamma_2} f^*\omega.$$

If the map of the circle extends to a map $g\colon D^2 \to M^n$ of the disk then, as easy to show, this extension can be made smooth, and then the Stokes theorem gives

$$\int_{S^1} f^*\omega = \int_{D^2} df^*\omega = \int_{D^2} f^* d\omega = 0,$$

i.e., $\int_{\gamma_1} f^*\omega = \int_{\gamma_2} f^*\omega$ (the integral depends only on the endpoints).

We turn to the definition of the Lie derivative of a tensor field.

Consider a smooth vector field $v(x)$ on $M^n$. For a small neighborhood $U$ of each point $x_0 \in M^n$ and for small $t < T$, there is a unique solution to the equation

$$\dot{y} = v(y)$$

with the initial data $y(t) = x$, where $x \in U$ (if $M^n$ is compact then, certainly, this solution exists globally for all times). Denote by $\varphi_t$ the point $\varphi_t(x) = y(t)$, where $y(t)$ is the solution to the equation with $y(0) = x$. From the uniqueness theorem we obtain:

$$\varphi_{s+t}(x) = \varphi_s(\varphi_t(x)), \quad 0 \le s, t,\ s+t < T.$$

The trajectories of the points $x$, i.e., the curves $\gamma(t) = \varphi_t(x)$, are called the *integral curves* of the vector field and the existence and uniqueness theorem claims that, for each point, there is a unique integral curve passing through this point. If $v(x_0) = 0$, then the curve degenerates into a fixed point $x_0$.

Consider how a smooth tensor field $T^{j_1\dots j_k}_{i_1\dots i_l}$ of type $(k,l)$ evolves along the integral curves of a field $v$. Denote by $e_1, \dots, e_n$ and $e^1 = dx^1, \dots, e^n = dx^n$ the basis vector and covector fields (we consider only a small coordinate neighborhood of a point $x \in M^n$). Now, define a function of the parameter $t$ whose values are tensors of type $(k,l)$ at $x$:

$$\begin{aligned} &S^{j_1\dots j_k}_{i_1\dots i_l}(t) e^{i_1} \otimes \dots \otimes e^{i_l} \otimes e_{j_1} \otimes \dots \otimes e_{j_k} \\ &\quad = T^{j_1\dots j_k}_{i_1\dots i_l}(\varphi_t(x)) \varphi_t^*(e^{i_1}) \otimes \dots \otimes \varphi_t^*(e^{i_l}) \otimes (\varphi_t)_*^{-1}(e_{j_1}) \otimes \dots (\varphi_t)_*^{-1}(e_{j_k}). \end{aligned}$$

The *Lie derivative* of a tensor field $T$ along a vector field $v$ (at $x$) is the tensor

$$(L_v T)^{j_1\dots j_k}_{i_1\dots i_l} = \frac{d}{dt} S^{j_1\dots j_k}_{i_1\dots i_l}(t)\Big|_{t=0}.$$

Calculate this derivative. For small values of $t$ we have (in local coordinates)

$$\varphi_t(x)^i = x^i + t v^i(x) + o(t), \quad i = 1, \dots, n.$$

Therefore,

$$((\varphi_t)_* T)^i = \Big(\delta^i_j + t \frac{\partial v^i(x)}{\partial x^j} + o(t)\Big) T^j$$

and

$$(\varphi_t^* T)_i = \Big(\delta_i^j + t\frac{\partial v^j(x)}{\partial x^i} + o(t)\Big)T_j.$$

Eventually, we obtain

$$S^{j_1\dots j_k}_{i_1\dots i_l}(t) = T^{q_1\dots q_k}_{p_1\dots p_l}(\varphi_t(x))\Big(\delta^{j_1}_{q_1} - t\frac{\partial v^{j_1}}{\partial x^{q_1}}\Big)\dots\Big(\delta^{j_k}_{q_k} - t\frac{\partial v^{j_k}}{\partial x^{q_k}}\Big)$$
$$\cdot\Big(\delta^{p_1}_{i_1} + t\frac{\partial v^{p_1}}{\partial x^{i_1}}\Big)\dots\Big(\delta^{p_l}_{i_l} + t\frac{\partial v^{p_l}}{\partial x^{i_l}}\Big) + o(t).$$

Hence we conclude that

$$(L_v T)^{j_1\dots j_k}_{i_1\dots i_l}(x) = v^m\frac{\partial T^{j_1\dots j_k}_{i_1\dots i_l}}{\partial x^m} + T^{j_1\dots j_k}_{m\dots i_l}\frac{\partial v^m}{\partial x^{i_1}} + \cdots$$
$$\cdots + T^{j_1\dots j_k}_{i_1\dots m}\frac{\partial v^m}{\partial x^{i_l}} - T^{m\dots j_k}_{i_1\dots i_l}\frac{\partial v^{j_1}}{\partial x^m} - \cdots - T^{j_1\dots m}_{i_1\dots i_l}\frac{\partial v^{j_k}}{\partial x^m}.$$

If the tensor field $T$ depends also on time, $T = T(x,t)$, then the *total derivative* along the vector field $v$ is the value

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + L_v T.$$

**Problem 3.15.** Prove the formula

$$L_v(R\otimes S) = (L_v R)\otimes S + R\otimes(L_v S).$$

**Problem 3.16.** Let $\omega_1$ and $\omega_2$ be differential forms. Prove that

$$L_v(\omega_1\wedge\omega_2) = (L_v\omega_1)\wedge\omega_2 + \omega_1\wedge(L_v\omega_2).$$

**Problem 3.17.** Prove that, for a scalar function, the Lie derivative along a field $v$ coincides with the derivative along $v$:

$$L_v f = v^i\frac{\partial f}{\partial x^i} = \partial_v f,$$

and for a vector field $w$ the Lie derivative equals the commutator of the fields:

$$L_v w = [v,w], \quad [v,w]^i = v^k\frac{\partial w^i}{\partial x^k} - w^k\frac{\partial v^i}{\partial x^k}.$$

**Problem 3.18.** Let $i_v$ be the operator of Problem 3.10. Prove the identity

$$i_v d + d i_v = L_v.$$

## 3.5 Embedding of smooth manifolds into the Euclidean space

In the beginning of the 20th century a smooth manifold was understood as the regular zero set of a map $f : \mathbb{R}^n \to \mathbb{R}^k$, which is now called a submanifold of the Euclidean space. In fact, the intrinsic definition given in §3.2 is not more general. Namely, the following Whitney theorem holds:

- *each smooth n-dimensional manifold with a countable base* (*i.e., having a countable base of open sets*) *is embedded into* $\mathbb{R}^{2n}$.

We restrict our consideration to proving the following simpler assertion:

**Theorem 3.2.** *Let $M$ be a closed smooth manifold. Then there is an embedding $\varphi\colon M \to \mathbb{R}^N$ into the Euclidean space of a sufficiently large dimension.*

*Proof.* We can restrict our consideration to the case of a connected manifold. Let $n$ be the dimension of the manifold.

We say that a *partition of unity* is given on a manifold $M$ if we are given a covering $\{U_\alpha\}$ of the manifold $M$ by open sets and a family of functions $\varphi_\alpha\colon M \to \mathbb{R}$ such that

(1) $\varphi_\alpha(x) = 0$ if $x$ does not lie in $U_i$;

(2) all functions $\varphi_\alpha$ are nonnegative: $\varphi_\alpha \geq 0$;

(3) at each point $x$ only finitely many functions $\varphi_\alpha$ are nonzero and

$$\sum_\alpha \varphi_\alpha(x) \equiv 1.$$

For a compact manifold $M$, take an open covering $\{U_\alpha\}$ possessing some additional properties. Namely,

(1) local coordinates $x_\alpha$ are given in each domain $U_\alpha$ which vary in the interior of the ball $\sum_i \left(x_\alpha^k\right)^2 < 1$;

(2) for some $\varepsilon > 0$, the subdomains $V_\alpha \subset U_\alpha$ defined by the condition $\sum_i \left(x_\alpha^i\right)^2 < 1 - \varepsilon$ constitute a covering of the manifold $M$, too.

Obviously, we can easily find such finite covering in view of compactness.

For this covering we construct a smooth partition of unity as follows: take some smooth function $f(r)$ defined on the half-axis $r \geq 0$ and possessing the following properties:

(1) $f(r) \equiv 1$ for $r \leq 1 - \varepsilon$;

(2) $f(r) \equiv 0$ for $r \geq 1$;

(3) $0 < f < 1$ for $1 - \varepsilon < r < 1$.

Define the function $\psi_\alpha$ as

$$\psi_\alpha(x) = \begin{cases} f((x_\alpha^1)^2 + \cdots + (x_\alpha^n)^2) & \text{for } x \in U_\alpha, \\ 0 & \text{otherwise,} \end{cases}$$

and put

$$\varphi_\alpha = \frac{\psi_\alpha}{\sum_\beta \psi_\beta}, \quad \alpha = 1, \dots, k,$$

constructing thereby a smooth partition of unity subordinate to the covering $\{U_\alpha\}$.

Given $\alpha = 1, \dots, k$, define the smooth map $f_\alpha : M \to \mathbb{R}^{n+1}$ by the formula

$$f_\alpha(x) = \begin{cases} (\varphi_\alpha(x)x_\alpha^1, \dots, \varphi_\alpha(x)x_\alpha^n, \varphi_\alpha(x)) & \text{for } x \in U_\alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Since each map $f_\alpha$ determines an embedding of the ball $V_\alpha$ into $\mathbb{R}^{n+1}$, the rank of the map $f_\alpha$ is maximal at the points of $V_\alpha$ (is equal to $n$). Therefore, the map

$$F : M \to \mathbb{R}^{(n+1)k}, \quad F(x) = (f_1(x), \dots, f_k(x)),$$

is an immersion of $M$ into $\mathbb{R}^{k(n+1)}$.

Note that

(1) if $x_1, x_2 \in V_\alpha$ and $x_1 \neq x_2$, then $f_\alpha(x_1) \neq f_\alpha(x_2)$;

(2) if $x_1 \in V_\alpha$ and $x_2$ does not lie in $V_\alpha$, then the $(n+1)$st coordinates of the map $f_\alpha$ at these points are different: it equals 1 for $x_1$ and is less than 1 for $x_2$.

Consequently, the map $F$ is an embedding of $M$ into $\mathbb{R}^{(n+1)k}$. The theorem is proven. □

**Problem 3.19.** (1) Prove that every product of spheres of the form $S^{k_1} \times \cdots \times S^{k_n}$ is embedded into the Euclidean space $\mathbb{R}^{k_1+\cdots+k_n+1}$ whose dimension is greater by one than the sum of the dimensions.

(2) Suppose that a closed $n$-dimensional manifold $M^n$ is embedded into $\mathbb{R}^{n+1}$. Prove that $M^n \times S^k$ is embedded into $\mathbb{R}^{n+k+1}$.

**Problem 3.20.** Suppose that a manifold $M^n$ is embedded into $\mathbb{R}^{n+1}$ as the zero set of a smooth function $f : \mathbb{R}^{n+1} \to \mathbb{R}$ and the gradient of $f$ is nonzero everywhere on the set $\{f = 0\}$. Prove that the manifold $M^n$ is orientable.

# 4

# Riemannian manifolds

## 4.1 Metric tensor

Let $M$ be a smooth manifold and the tangent space at each point of $M$ be endowed with an inner product defined by a symmetric tensor $g_{ij}$:

$$v, w \in T_x M \to (v, w) = g_{ij} v^i w^j \in \mathbb{R}, \quad g_{ij} = g_{ji}.$$

Suppose that $g_{ij}(x)$ depends continuously on $x \in M$. Such a manifold is called a *Riemannian manifold* and the tensor $g_{ij}$ is called the *metric tensor* (or the *Riemannian metric*).[1] If the metric depends smoothly on $x$, then we have a smooth Riemannian manifold.

The simplest examples of smooth Riemannian manifolds are submanifolds in $\mathbb{R}^n$.

Let $f : M \to \mathbb{R}^n$ be an embedding of a manifold $M$ into $\mathbb{R}^n$. Define the metric $(\cdot, \cdot)$ on $M$ *induced by the embedding*. Given $v, w \in T_x M$ and $f_*(v), f_*(w) \in T_{f(x)}\mathbb{R}^n$, we put

$$(v, w) := (f_*(v), f_*(w)),$$

where in the right-hand side we take the standard inner product in $\mathbb{R}^n$. In the case of surfaces in $\mathbb{R}^3$ the metric tensor induced by the embedding is called the first fundamental form.

We can give a more general definition: if $f : M \to N$ is a smooth embedding of $M$ into a Riemannian manifold $N$, then the metric on the manifold $M$ given by the formula

$$(v, w) = (f_*(v), f_*(w))_N,$$

where $(\cdot, \cdot)$ is the metric on $N$, is said to be *induced* by the embedding $f$.

From Theorem 3.2 we obtain

**Corollary 4.1.** *Each closed smooth manifold carries a smooth Riemannian metric.*[2]

[1] Do not confuse this notion with a metric as the distance between two points.

[2] Moreover: each Riemannian metric is induced by some embedding into the Euclidean space of sufficiently large dimension. This embedding is called isometric. Namely,

- each Riemannian $n$-dimensional manifold $M^n$ admits a $C^1$-smooth isometric embedding $M^n \to \mathbb{R}^{2n}$ (moreover, it is possible that the metric tensor is continuous but not differentiable) (Nash and Kuiper);
- each smooth Riemannian $n$-dimensional manifold $M^n$ admits a smooth isometric embedding into $\mathbb{R}^{\frac{1}{2}(3n^2+11n)(n+1)}$. For closed manifolds, there is an isometric embedding into $\mathbb{R}^{\frac{1}{2}(3n^2+11n)}$; if the Riemannian manifold is real-analytic, then the embedding can be chosen real-analytic, too (Nash).

We have stated these theorems in their original form as obtained by Nash and Kuiper. Later these estimates in the Nash theorem were improved.

We introduce $g^{ij}$ which is uniquely determined by the following relations:

$$g^{ij} g_{jk} = \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

A generalization of a Riemannian metric is the notion of a pseudo-Riemannian metric: suppose that the tensor $g_{ij}$ is symmetric but not necessarily positive definite and the corresponding matrix has $k$ positive and $(n-k)$ negative eigenvalues, where $\dim M = n$. Then it determines a *pseudo-Riemannian metric* of signature $(k, n-k)$.

The simplest example of a manifold with such a metric is the *pseudo-Euclidean space* $\mathbb{R}^{k,n}$ diffeomorphic to $\mathbb{R}^{k+n}$ and such that each tangent plane is endowed with the indefinite inner product

$$(v, w)_{k,n} = v^1 w^1 + \cdots + v^k w^k - v^{k+1} w^{k+1} - \cdots - v^{k+n} w^{k+n}.$$

The space $\mathbb{R}^{1,3}$ is called the *Minkowski space* and known as "spacetime" in special relativity (see Chapter 5).

## 4.2 Affine connection and covariant derivative

As a rule, formulas (3.4) which connect the expressions of tensors in different coordinate systems are nonlinear with respect to the coordinates on the manifold. Therefore, definition of the derivative of tensors which would be independent of coordinates requires introduction of affine connections.

Let $M$ be a smooth manifold.

Suppose that, at each point $x \in M$, we have a function which relates to each vector field $v$ on $M$ and each vector $w \in T_x M$ the new tangent vector

$$v, w \to \nabla_w v \in T_x M \tag{4.1}$$

and the following conditions are satisfied:

(1) the map (4.1) is linear in both arguments:

$$\nabla_{\alpha_1 w_1 + \alpha_2 w_2} v = \alpha_1 \nabla_{w_1} v + \alpha_2 \nabla_{w_2} v,$$
$$\nabla_w (\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 \nabla_w v_1 + \alpha_2 \nabla_w v_2, \quad \alpha_1, \alpha_2 \in \mathbb{R};$$

(2) if $f : M \to \mathbb{R}$ is a smooth function, then

$$\nabla_{fw} v = f \nabla_w v, \quad \nabla_w f v = (D_w f)\, v + f \nabla_w v,$$

where $D_w f$ is the derivative of $f$ the direction of $w$.

This function (4.1) is called an *affine connection* and its value $\nabla_w v$ is called the *covariant derivative* of the vector field $v$ in the direction of $w$.

If the vector field $\nabla_w v$ is smooth for smooth vector fields $v$ and $w$, then the connection is called smooth. Henceforth we assume that all connections are smooth.

Note that, defining the connection, we do not require $M$ be a Riemannian manifold.

In coordinates, the connection is written in terms of the *Christoffel symbols* $\Gamma^k_{ij}$. Namely, let $x^1, \dots, x^n$ be coordinates in a domain of $M$ and let $\partial_1, \dots, \partial_n$ be the corresponding vector fields (3.2) which generate bases for the tangent spaces. Define the Christoffel symbols by the formula

$$\nabla_{\partial_j} \partial_i = \Gamma^k_{ij} \partial_k. \tag{4.2}$$

It follows from the definition of connection that for arbitrary vector fields $v = v^i \partial_i$ and $w = w^j \partial_j$ the covariant derivative has the form

$$\begin{aligned}\nabla_w v = \nabla_{w^j \partial_j}(v^i \partial_i) &= w^j \nabla_{\partial_j}(v^i \partial_i) \\ &= w^j \left( \frac{\partial v^k}{\partial x^j} \partial_k + v^i \nabla_{\partial_j} \partial_i \right) = w^j \left( \frac{\partial v^k}{\partial x^j} + v^i \Gamma^k_{ij} \right) \partial_k;\end{aligned}$$

eventually we obtain

$$(\nabla_w v)^k = w^j \left( \frac{\partial v^k}{\partial x^j} + v^i \Gamma^k_{ij} \right). \tag{4.3}$$

From these calculations we conclude that

- *the Christoffel symbols determine the connection uniquely.*

Moreover, it follows from them that the Christoffel symbols do not constitute a tensor. Show this.

Suppose that $y^1, \dots, y^n$ are other coordinates in the same domain and $\tilde{\partial}_1, \dots, \tilde{\partial}_n$ are the corresponding coordinate vector fields. It follows from the chain rule that

$$\partial_i = \frac{\partial}{\partial x^i} = \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} = \frac{\partial y^j}{\partial x^i} \tilde{\partial}_j.$$

Insert this in (4.2):

$$\begin{aligned}\Gamma^k_{ij} \partial_k = \nabla_{\partial_j} \partial_i &= \nabla_{\partial_j} \left( \frac{\partial y^l}{\partial x^i} \tilde{\partial}_l \right) \\ &= \frac{\partial^2 y^l}{\partial x^i \partial x^j} \tilde{\partial}_l + \frac{\partial y^l}{\partial x^i} \frac{\partial y^m}{\partial x^j} \nabla_{\tilde{\partial}_m} \tilde{\partial}_l \\ &= \frac{\partial^2 y^l}{\partial x^i \partial x^j} \tilde{\partial}_l + \frac{\partial y^l}{\partial x^i} \frac{\partial y^m}{\partial x^j} \tilde{\Gamma}^n_{lm} \tilde{\partial}_n \\ &= \frac{\partial^2 y^m}{\partial x^i \partial x^j} \frac{\partial x^k}{\partial y^l} \partial_k + \frac{\partial y^l}{\partial x^i} \frac{\partial y^m}{\partial x^j} \tilde{\Gamma}^n_{kl} \frac{\partial x^k}{\partial y^m} \partial_k.\end{aligned}$$

Comparing the coefficients of $\partial_k$, we finally obtain

$$\Gamma^k_{ij} = \frac{\partial^2 y^p}{\partial x^i \partial x^j}\frac{\partial x^k}{\partial y^p} + \frac{\partial y^l}{\partial x^i}\frac{\partial y^m}{\partial x^j}\frac{\partial x^k}{\partial y^n}\widetilde{\Gamma}^n_{lm}$$

(here for clarity we replaced the summation index $l$ in the first summand with $p$). Hence,

- *the Christoffel symbols transform like tensors only under linear changes of coordinates* (in this case $\frac{\partial^2 y^m}{\partial x^i \partial x^j} \equiv 0$),
- *the values* $T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}$ *constitute a tensor* (*which is called the torsion tensor*).

The covariant derivative extends to arbitrary tensors as follows.

If $f$ is a smooth function on a manifold $M$, then we naturally assume that

$$\nabla_w f = w^j \frac{\partial f}{\partial x^j},$$

i.e., the covariant derivative coincides with the derivative of a function along the field $w$.

Let $u$ be a covector field on $M$ and let $v$ be a vector field on $M$. Then $f = u_i v^i$, i.e., the value of $u$ at $v$ is a smooth function and

$$\nabla_w f = w^j \frac{\partial (u_i v^i)}{\partial x^j}.$$

Under the natural assumption that the covariant derivative satisfies the *Leibniz rule*

$$\nabla_w (u_i v^i) = (\nabla_w u)_i v^i + u_i (\nabla_w v)^i,$$

from (4.3) we derive

$$(\nabla_w u)_i = w^j \frac{\partial u_i}{\partial x^j} - \Gamma^k_{ij} w^j u_k.$$

Every tensor of type $(k, l)$ is representable as a linear combination of elementary tensors of the form

$$T^{i_1 \dots i_k}_{j_1 \dots j_l} = A^{i_1}_{(1)} \dots A^{i_k}_{(k)} A^{(k+1)}_{j_1} \dots A^{(k+l)}_{j_l},$$

i.e., tensor products of tensors of type $(0, 1)$ or $(1, 0)$. Now, supposing that the covariant derivative satisfies the Leibniz rule

$$\nabla_w (A \otimes B) = (\nabla_w A) \otimes B + A \otimes (\nabla_w B),$$

and knowing its action on vectors and covectors, we can write its action on every tensor field. Moreover,

- *the covariant derivative of a tensor is a tensor.*

**Problem 4.1.** Show that the so-defined covariant derivative of a tensor has the form

$$\nabla_w T^{(i)}_{(k)} = \Big( w^j \frac{\partial T^{(i)}_{(k)}}{\partial x^j} - T^{(i)}_{k k_2 \dots k_n} \Gamma^k_{k_1 j} - \dots - T^{(i)}_{k_1 \dots k} \Gamma^k_{k_n j} + T^{i i_2 \dots i_m}_{(k)} \Gamma^{i_1}_{ij} + \dots + T^{i_1 \dots i}_{(k)} \Gamma^{i_m}_{ij} \Big).$$

In particular, the following formula is valid for tensors of type $(0, 2)$:

$$\nabla_w g_{ij} = w^k \frac{\partial g_{ij}}{\partial x^k} - w^k (\Gamma^l_{ik} g_{lj} + \Gamma^l_{jk} g_{il}).$$

Note that we started with the definition of the covariant derivative of vector fields and never supposed that the manifold is Riemannian. Such a definition of connection can be introduced on every smooth *vector bundle*, i.e., on a manifold $E$ with a smooth map $\pi : E \to B$ such that

(1) there is a covering of the manifold $B$ by domains $U_\alpha$ such that, for each domain $U_\alpha$, there is a diffeomorphism

$$p_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F,$$

where $F$ is a vector space which is the same for all $\alpha$ and the diffeomorphism $p_\alpha$ agrees with the projection:

$$\pi(p_\alpha^{-1}(x, v)) = x, \quad x \in U_\alpha, \ v \in F;$$

(2) if the intersection $U_\alpha \cap U_\beta$ of such domains is nonempty, then the map

$$p_\alpha p_\beta^{-1} : (U_\alpha \cap U_\beta) \times F \to (U_\alpha \cap U_\beta) \times F$$

has the form $(x, v) \to (x, A_x(v))$, where, for each point $x$, the linear transformation $A_x$ belongs to some subgroup $G$ of the group $\mathrm{GL}(F)$ of all invertible linear maps of $F$ into itself.

In this case $E$ is called the *space of the bundle*, $B$ is called the *base space*, $G$ is called the *structure group* of the bundle, and $F$ is called the *fiber* of the bundle.

The simplest example of a vector bundle after the direct product $M \times F$ is the tangent bundle of an $n$-dimensional manifold $M$: its fiber is isomorphic to $\mathbb{R}^n$, the structure group is $\mathrm{GL}(n, \mathbb{R})$, and the map $\pi$ relates to a tangent vector the reference point, i.e., the point to which the vector is attached (Theorem 3.1).[3]

In mathematics, there are more general *bundles*, when the fiber $F$ is not necessarily a vector space and the maps $p_\alpha p_\beta^{-1}$ have the form $(x, y) \to (x, f_x(y))$, where $f_x : F \to F$ are diffeomorphisms depending continuously on $x$. Henceforth we consider only vector bundles.

A *section* $v : U \to E$ *of a bundle*, given in some domain $U \subset B$, is a function which relates to each point $x \in U$ some vector lying in the fiber $F_x = \pi^{-1}(x)$ over this point.

A *vector field* is merely a section of the tangent bundle.

The connection in a bundle is defined as the rule for calculation of the derivatives of smooth sections:

$$\nabla_j v = \Big( \frac{\partial}{\partial x^j} + A_j \Big) v, \tag{4.4}$$

[3] For instance, it is known that the tangent bundle to the $n$-dimensional sphere $S^n$ is homeomorphic to the product $S^n \times \mathbb{R}^n$ if and only if $n = 1, 3$, or 7.

where $(x^1, \dots, x^n)$ are the coordinates in $U$ and $A_j(x), i = j, \dots, n$ are matrices in the Lie algebra of the group $G$ (see Chapter 7). The result is again a section of the same bundle. The connections in the tangent bundle are given by the matrices

$$A_j = (\Gamma^k_{ij}) \tag{4.5}$$

(here $i$ and $k$ are matrix indices and $j$ is the index of the matrix itself; obviously, in the general case when the dimensions of the base space and the fiber are different these indices vary over different values).

Connections of this general form play a fundamental role in physics of elementary particles. For example, an electromagnetic field is given as a connection in the bundle with the structure group $G = \mathrm{U}(1)$ (see Chapter 7).

The connection determines the *parallel translation* in the bundle.

Let $\gamma\colon [0, T] \to B$ be a smooth curve on the base of a bundle. Then a vector field $v(t)$ (i.e., a function of $t$ with values in the fibers of the bundle: $v(t) \in \pi^{-1}(\gamma(t))$) is called *parallel* along $\gamma$ if $\pi(v(t)) = \gamma(t)$ for every $t$ and the following equation holds:

$$\frac{Dv}{\partial t} = \nabla_{\dot\gamma} v = 0. \tag{4.6}$$

Expanding the left-hand side, we find that the field $v$ is parallel if and only if

$$\left(\frac{Dv}{\partial t}\right)^i = \dot\gamma^k \left(\frac{\partial v^i}{\partial x^k} + \Gamma^i_{jk} v^j\right) = 0, \quad k = 1, \dots, n.$$

Equation (4.6), being a first-order equation, has a unique solution for an arbitrary initial value $v(0)$; and we say that the vector $v(1)$ is obtained from the vector $v(0)$ by parallel translation along the curve $\gamma$.

## 4.3 Riemannian connections

Let $M$ be a smooth manifold with a metric $g_{ij}$ which can be either Riemannian or pseudo-Riemannian and let its tangent bundle carry an affine connection.

**Lemma 4.1.** *The following assertions are equivalent.*

(1) *For every vector field $w$ on $M$,*

$$\nabla_w g_{ij} = 0.$$

(2) *If $\gamma(t)$ is a smooth curve on $M$ and $v$ and $w$ are parallel vector fields along $\gamma$, then their inner product is constant along $\gamma$:*

$$\frac{d}{dt}(v(t), w(t)) = 0.$$

(3) *If $\gamma(t)$ is a smooth curve on $M$ and $v$ and $w$ are vector fields along $\gamma$, then*

$$\frac{d}{dt}(v(t), w(t)) = (\nabla_{\dot\gamma} v, w) + (v, \nabla_{\dot\gamma} w).$$

*Proof.* First of all write down $\frac{d}{dt}(v(t), w(t))$ in local coordinates and, using Problem 4.1, obtain its invariant form:

$$\begin{aligned}
\frac{d}{dt}(v(t), w(t)) &= \frac{d}{dt}(g_{ij} v^i w^j) \\
&= \frac{dg_{ij}}{dt} v^i w^j + g_{ij}\frac{dv^i}{dt} w^j + g_{ij} v^i \frac{dw^j}{dt} \\
&= (\nabla_{\dot\gamma} g_{ij} + \Gamma^k_{il}\dot\gamma^l g_{kj} + \Gamma^k_{jl}\dot\gamma^l g_{ik}) v^i w^j + g_{ij}\frac{dv^i}{dt} w^j + g_{ij} v^i \frac{dw^j}{dt} \\
&= \nabla_{\dot\gamma} g_{ij} v^i w^j + g_{ij}\left(\frac{dv^i}{dt} + \Gamma^i_{kl}\dot\gamma^l v^k\right) w^j \\
&\quad + g_{ij} v^i \left(\frac{dw^j}{dt} + \Gamma^j_{kl}\dot\gamma^l w^k\right) \\
&= \nabla_{\dot\gamma} g_{ij} v^i w^j + (\nabla_{\dot\gamma} v, w) + (v, \nabla_{\dot\gamma} w).
\end{aligned}$$

It is clear from the last formula that (1) implies (3). The fact that (2) follows from (3) is obvious. We are left with proving that (2) implies (1). Take an arbitrary point $x \in M$ and arbitrary vectors $u, v, w \in T_x M$. Draw a curve from $x$ in the direction of $u$ and extend (uniquely) $v$ and $w$ to parallel vector fields on the curve. It follows from (2) that $\nabla_u g_{ij} v^i w^j = 0$. Since all data are arbitrary, (1) follows. Lemma 4.1 is proven. □

If a connection satisfies one of the conditions (1)–(3) of Lemma 4.1, then it is *compatible with the metric*.

Another important class of connections are symmetric connections: a connection is *symmetric* if its torsion tensor is identically zero,

$$T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji} = 0.$$

Obviously, this identity is equivalent to the following assertion: the relation

$$\nabla_u v - \nabla_v u = [u, v] \tag{4.7}$$

holds for every pair of vector fields $u$ and $v$, where

$$[u, v]^i = u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j}$$

is the *commutator of the vector fields* $u$ and $v$. The equivalence is proven by straightforward computations.

The following theorem holds:

**Theorem 4.1.** *On each Riemannian manifold, there is a unique symmetric connection compatible with the metric. Its Christoffel symbols have the form*

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left( \frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{lj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l} \right). \tag{4.8}$$

*Proof.* Let $\partial_1, \dots, \partial_n$ be the basis for $T_x M$ corresponding to coordinates $x^1, \dots, x^n$ in a domain $U \subset M$. The tensor $g_{ij}$ has the form $g_{ij} = (\partial_i, \partial_j)$.

Suppose that $\nabla$ is a symmetric connection on the manifold compatible with the metric. It follows from Lemma 4.1 that the equality

$$\frac{\partial g_{ij}}{\partial x^l} = (\nabla_{\partial_l} \partial_i, \partial_j) + (\partial_i, \nabla_{\partial_l} \partial_j) = \Gamma^m_{il} g_{mj} + \Gamma^m_{jl} g_{im} \tag{4.9}$$

holds at each point of $U$. Inserting this equality in

$$\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{lj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}$$

and recalling that the connection and the metric tensor are both symmetric; i.e., $\Gamma^k_{ij} = \Gamma^k_{ji}$ and $g_{ij} = g_{ji}$ for all $i$, $j$, and $k$, we obtain

$$2\Gamma^m_{ij} g_{ml} = \frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{lj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}.$$

Multiplying both sides of the last formula by $2g^{kl}$ and summing over $l$, we arrive at (4.8).

Conversely, connection (4.8) is obviously symmetric. By Lemma 4.1, it is compatible with the metric, since it satisfies equation (4.9). Theorem 4.1 is proven. □

This theorem is a generalization of Theorem 2.5 for surfaces in $\mathbb{R}^3$ to arbitrary Riemannian and pseudo-Riemannian manifolds.

The case of surfaces in $\mathbb{R}^3$ suggests the following construction of such connections on submanifolds.

Let $M$ be a Riemannian manifold with a symmetric connection $\nabla$ compatible with the metric. Let $N$ be a submanifold in $M$ with the induced metric. Consider a vector field $v$ on $N$ and let $w \in T_x N$. The covariant derivative $\nabla_w v$ need not be tangent to $N$; therefore, decompose it into components

$$\nabla_w v = \widetilde{\nabla}_w v + B(w, v),$$

where $\widetilde{\nabla}_w v \in T_x N$ and the vector $B(w, v)$ is orthogonal to $T_x N$. It is easy to check that

- $\widetilde{\nabla}$ *determines a symmetric connection on* $N$ *compatible with the induced metric*;
- $B$ *is a symmetric bilinear operator* (*the second fundamental form*) *on the tangent bundle of the submanifold* $N$ *with values in the orthogonal complements of the tangent spaces* $T_x N$.

In the case of surfaces in $\mathbb{R}^3$ the Christoffel symbols are determined this way from the trivial connection on $\mathbb{R}^3$ ($\Gamma_{ij}^k = 0$ in the Euclidean coordinates), and the second fundamental form multiplied by the normal to the surface coincides with $B$.

## 4.4 Curvature

Let $M$ be a manifold with a connection in the tangent bundle and let $U$ be a domain on $M$ with local coordinates $x^1, \dots, x^n$. Take $x \in M$. Suppose for simplicity that the coordinates of $x$ are $(0, \dots, 0)$.

Consider a small square in $U$ whose sides are segments of the coordinate lines and join successively the points $x = (0, 0, 0, \dots)$, $y = (\varepsilon, 0, 0, \dots)$, $z = (0, \varepsilon, 0, \dots)$, and $s = (\varepsilon, \varepsilon, 0, \dots)$, where $\dots$ stands for zeros. Define two operators $T_{1\varepsilon}$ and $T_{2\varepsilon}$ from $T_x M$ to $T_s M$. Each of them is a composition of successive parallel translations along the sides of the square: $T_{1\varepsilon}$ is the composition of the translations from $x$ to $y$ and then from $y$ to $s$, and $T_{2\varepsilon}$ is the composition of the translations from $x$ to $z$ and then from $z$ to $s$.

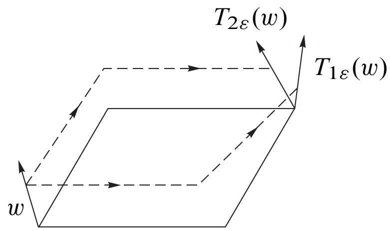


Figure 4.1. Parallel translations along the sides of a square.

In the general case $T_{1\varepsilon}(w) \neq T_{2\varepsilon}(w)$ and the difference between these vectors is described by the *curvature of the connection*. The notion of curvature is defined for every connection on every vector bundle. However, the general definition is based on the theory of Lie groups and algebras. Therefore, we restrict ourselves to the case of connections on the tangent bundle.

Find $T_{1\varepsilon}(w) - T_{2\varepsilon}(w)$ up to lower order terms in $\varepsilon$. Denote the operator of parallel translation from $x$ to $y$ by $\tau_1$ and the operator of parallel translation from $y$ to $s$ by $\tau_2$. We decompose all tangent vectors in the bases $\partial_1, \dots, \partial_n$ for the tangent spaces.

It follows from (4.6) that

$$\tau_1(w)^k = w^k - \varepsilon\Gamma^k_{1j}(x)w^j + O(\varepsilon^2),$$

$$\tau_2(\widetilde{w})^k = \widetilde{w}^k - \varepsilon\Gamma^k_{2j}(y_1)\widetilde{w}^j + O(\varepsilon^2) = \widetilde{w}^k - \varepsilon\Big(\Gamma^k_{2j}(x) + \varepsilon\frac{\partial\Gamma^k_{2j}(x)}{\partial x^1}\Big)\widetilde{w}^j + O(\varepsilon^2).$$

Also, we can easily estimate the terms of order $\varepsilon^2$: $w$ is a solution to the equation (4.6) of the form $\dot{w} = Aw$ and substituting for $A$ and $w$ their Taylor series, we obtain

$$w(\varepsilon) = w(0) - A(0)w(0)\varepsilon - (\dot{A}(0)w(0) + A(0)\dot{w}(0))\frac{\varepsilon^2}{2} + O(\varepsilon^3).$$

Now, when all computations become clear, we can give the final answer:

$$T_{1\varepsilon}(w)^k - T_{2\varepsilon}(w)^k = \Big(-\frac{\partial\Gamma^k_{2j}}{\partial x^1} + \frac{\partial\Gamma^k_{1j}}{\partial x^2} + \Gamma^k_{2l}\Gamma^l_{1j} - \Gamma^k_{1l}\Gamma^l_{2j}\Big)w^j\cdot\varepsilon^2 + O(\varepsilon^3).$$

The last expression has the form

$$T_{1\varepsilon}(w) - T_{2\varepsilon}(w) = (\nabla_{\partial_2}\nabla_{\partial_1} - \nabla_{\partial_1}\nabla_{\partial_2})w\cdot\varepsilon^2 + O(\varepsilon^3). \tag{4.10}$$

These calculations lead to introduction of the *curvature tensor* (or *Riemann* (*curvature*) *tensor*), which represents a linear function of three vector fields $u$, $v$, and $w$:

$$R(u,v)w = (\nabla_v\nabla_u - \nabla_u\nabla_v + \nabla_{[u,v]})w.$$

(obviously, $[\partial_i, \partial_j] = 0$).

**Lemma 4.2.** *The value $R(u,v)w$ at a point $x$ depends only on the values of $u$, $v$, and $w$ at $x$ and the map*

$$R\colon T_xM\times T_xM\times T_xM\to T_xM$$

*is linear in each of the arguments; i.e., $R^l_{ijk}$ is a tensor, where*

$$(R(u,v)w)^l = R^l_{ijk}u^kv^jw^i,$$

$$R^l_{ijk} = \frac{\partial\Gamma^l_{ik}}{\partial x^j} - \frac{\partial\Gamma^l_{ij}}{\partial x^k} + \Gamma^l_{mj}\Gamma^m_{ik} - \Gamma^l_{mk}\Gamma^m_{ij}. \tag{4.11}$$

*Proof.* We can demonstrate straightforwardly that, multiplying some of three fields, say, the field $u$ by a smooth function $f$, we obtain $R(fu,v)w = fR(u,v)w$. Indeed, we have

$$\begin{aligned} R(fu,v)w &= (\nabla_v\nabla_{fu} - \nabla_{fu}\nabla_v + \nabla_{[fu,v]})w \\ &= (\nabla_v f\nabla_u - f\nabla_u\nabla_v + \nabla_{(f[u,v]-D_vfu)})w \\ &= (D_vf\nabla_u + f\nabla_v\nabla_u - f\nabla_u\nabla_v + f\nabla_{[u,v]} - D_vf\nabla_u)w \\ &= fR(u,v)w. \end{aligned}$$

For $v$ and $w$ this identity is proven by similar calculations. Decomposing the fields in the basis $\{\partial_i\}$ and applying this property, we obtain

$$R(u^k\partial_k, v^j\partial_j)(w^i\partial_i) = u^k v^j w^i \cdot R(\partial_k, \partial_j)\partial_i,$$

which completes the proof of Lemma 4.2. □

Note that in terms of the matrices $A_j = (\Gamma^l_{ij})$ (see (4.5)) formula (4.11) takes a very simple form:

$$\Omega_{jk} = \frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k} + A_j A_k - A_k A_j = \frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k} + [A_j, A_k],$$

where $\Omega_{jk} = R^l_{ijk}$. For a general connection (4.4), $\Omega$ is called the *curvature of the connection*; in small domains in the base over which the bundle is trivial, $\Omega$ is the matrix-valued 2-form $\Omega = \sum_{j<k} \Omega_{jk} dx^j \wedge dx^k$ (globally on a manifold it is a 2-form with values in fibers of some bundle which is isomorphic to the Lie algebra of the group $G$).

A connection has zero curvature, if it satisfies the *zero curvature equations*

$$\frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k} + [A_j, A_k] = 0, \quad j,k = 1,\ldots,n.$$

In one important particular case, these equations appeared already in Chapter 2 as the Gauss–Codazzi equations.

Consider differential forms with values in fibers of a bundle and define the differential $d^\nabla$ by the formula

$$d^\nabla\Big(\sum_{i_1<\cdots<i_k} T_{i_1\ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}\Big) = \sum_{i,i_1<\cdots<i_k} \nabla_i T_{i_1\ldots i_k} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \quad (4.12)$$

where $\nabla_i = \nabla_{\partial_i}$.

**Problem 4.2.** Prove that the square of the differential $d^\nabla$ is the operator of multiplication by the curvature form:

$$(d^\nabla)^2\omega = \Omega \wedge \omega.$$

The following lemma is the infinitesimal version of (4.10).

**Lemma 4.3.** *Let $\boldsymbol{r} : U \to M$ be an immersion of a surface $M$ and let $x$ and $y$ be the coordinates on the surface. Then*

$$\frac{D}{\partial y}\frac{D}{\partial x} - \frac{D}{\partial x}\frac{D}{\partial y} = R(\boldsymbol{r}_x, \boldsymbol{r}_y).$$

Here $D/\partial x$ and $D/\partial y$ stand for the operators (4.6) of taking the total derivative along the coordinate lines on the surface.

*Proof.* The local coordinates $x$ and $y$ on the surface extend to coordinates $x, y, z^1, \ldots, z^{n-2}$ on the $n$-dimensional manifold and the total derivatives coincide with the covariant derivatives along the coordinate fields. From the definition of the curvature tensor we obtain

$$R(\boldsymbol{r}_x, \boldsymbol{r}_y) = -\nabla_{\boldsymbol{r}_x}\nabla_{\boldsymbol{r}_y} + \nabla_{\boldsymbol{r}_y}\nabla_{\boldsymbol{r}_x} + \nabla_{[\boldsymbol{r}_x,\boldsymbol{r}_y]} = -\nabla_{\boldsymbol{r}_x}\nabla_{\boldsymbol{r}_y} + \nabla_{\boldsymbol{r}_y}\nabla_{\boldsymbol{r}_x},$$

since the coordinate vector fields commute: $[\boldsymbol{r}_x, \boldsymbol{r}_y] = 0$. The lemma is proven. □

The curvature tensor satisfies some additional relations.

**Lemma 4.4.** *Let $R^l_{ijk}$ be the curvature tensor of a connection (in the tangent bundle). Then the following relations hold for arbitrary vector fields $u$, $v$, $w$, and $z$:*

(1) $(u, v)w = -R(v, u)w$ *or* $R^l_{ijk} = -R^l_{ikj}$, $i, j, k, l = 1, \dots, n$;

(2) *if the connection is symmetric, then*

$$R(u, v)w + R(v, w)u + R(w, u)v = 0$$

*or*

$$R^l_{ijk} + R^l_{kij} + R^l_{jki} = 0, \quad i, j, k, l = 1, \dots, n;$$

(3) *if a Riemannian metric is given on $M$ and the connection is compatible with the metric, then*

$$(R(u, v)w, z) = -(R(u, v)z, w);$$

(4) *if a Riemannian metric is given on $M$ and the connection is compatible with the metric and is symmetric, then*

$$(R(u, v)w, z) = (R(w, z)u, v).$$

*Proof.* Assertion (1) is obvious.

Since $R$ is a tensor, it suffices to prove assertion (2) for pairwise commuting (for example, coordinate) fields. Then

$$\begin{aligned} &R(u, v)w + R(v, w)u + R(w, u)v \\ &\quad = (-\nabla_u\nabla_v w + \nabla_v\nabla_u w) + (-\nabla_v\nabla_w u + \nabla_w\nabla_v u) + (-\nabla_w\nabla_u v + \nabla_u\nabla_w v) \end{aligned}$$

and, applying (4.7), we demonstrate that the right-hand side of this formula is identically zero.

Assertion (3) means that the expression $(R(u, v)w, z)$ is antisymmetric with respect to $w$ and $z$ which is obviously equivalent to the identity $(R(u, v)w, w) = 0$. Since $R$ is a tensor, we can restrict again our exposition to the case of $[u, v] = 0$. In this case, denoting by $D_u$ and $D_v$ the derivatives of functions along the fields $u$ and $v$ and using the fact that the connection is compatible with the metric, we obtain

$$\begin{aligned} D_u D_v(w, w) &= D_u(\nabla_v(w, w)) \\ &= 2D_u(\nabla_v w, w) \\ &= 2\nabla_u(\nabla_v w, w) \\ &= 2(\nabla_u\nabla_v w, w) + 2(\nabla_u w, \nabla_v w) \end{aligned}$$

and similarly

$$D_v D_u(w, w) = 2(\nabla_v \nabla_u w, w) + 2(\nabla_u w, \nabla_v w).$$

Since $[u, v] = 0$, we have $D_u D_v = D_v D_u$, whence we derive

$$(D_u D_v - D_v D_u)(w, w) = -(R(u, v)w, w) = 0,$$

as required.

Define the tensor

$$R_{ijkl} = g_{im} R^m_{jkl} \tag{4.13}$$

or

$$R_{ijkl} u^l v^k w^j z^i = \langle R(u, v)w, z \rangle.$$

Assertion (2) takes the form

$$R_{ijkl} + R_{iklj} + R_{iljk} = R_{i[jkl]} = 0.$$

Assertions (1) and (3) mean that the tensor $R_{ijkl}$ is antisymmetric in the second and the first pairs of indices:

$$R_{ijkl} = -R_{ijlk}, \quad R_{ijkl} = -R_{jikl}.$$

Using these identities, we can write

$$R_{i[jlk]} + R_{l[ijk]} - R_{j[ikl]} - R_{k[ilj]} = 2(R_{lijk} - R_{jkli}) = 0,$$

which proves assertion (4). Lemma 4.4 is proven. □

From the definition of the curvature tensor of an arbitrary connection we see that it obviously satisfies the assertion (1) of the above lemma: $F_{jk} = -F_{kj}$. Suppose that an inner product is given in the fibers of a vector bundle (in this case the structure group is a subgroup of the orthogonal group $O(n)$, $n = \dim F$) which is compatible with the connection: $\nabla_w(u, v) = (\nabla_w u, v) + (u, \nabla_w v)$. Then this connection satisfies the assertion (3) of the above lemma and its proof repeats verbatim the proof above).

**Problem 4.3.** Suppose that a connection on a Riemannian manifold $M$ is symmetric and compatible with the metric. If $\dim M = 2$, then tensor (4.13) is determined completely by one component, namely $R_{1212}$, and if $\dim M = 3$, then by six components.

Also it follows from Lemma 4.4 that if $u, v \in T_x M$ and $\sigma$ is a two-dimensional subspace in $T_x M$ generated by $u$ and $v$, then the value

$$K(\sigma) = \frac{(R(u, v)u, v)}{(u, u)(v, v) - (u, v)^2}$$

depends only on $\sigma$. It is called the *sectional curvature* in the two-dimensional direction $\sigma$.

**Lemma 4.5.** *If $M$ is a two-dimensional surface in $\mathbb{R}^3$ with the induced metric and the connection is compatible with the metric and is symmetric, then its sectional curvature $K$ coincides with the Gaussian curvature.*

*Proof.* Take $p \in M$. Since the Gaussian and sectional curvatures are independent of the choice of local coordinates, it suffices to show that these values, calculated in some special coordinate system, coincide.

Choose orthonormal coordinates $(x^1, x^2, x^3)$ in $\mathbb{R}^3$ so that the surface in a neighborhood of the point $p$ is the graph of a function $x^3 = f(x^1, x^2)$. Then $g_{ii} = 1 + f_i^2$, $g_{12} = f_1 f_2$, and

$$K = \frac{f_{11} f_{22} - f_{12}^2}{1 + f_1^2 + f_2^2},$$

where the subscripts of $f$ denote derivatives with respect to $x^i$ (Problems 2.1 and 2.3). This expression becomes much simpler if we direct the $x^3$-axis along the normal to the surface at the point $p$: in this case $\operatorname{grad} f(p) = 0$. In particular, we find that

$$K = f_{11} f_{22} - f_{12}^2$$

at $p$ and all Christoffel symbols are zero. It follows from (4.8) and (4.11) that

$$\begin{aligned} R_{1212} &= \frac{\partial^2 g_{12}}{\partial x^1 \partial x^2} - \frac{1}{2}\left(\frac{\partial^2 g_{11}}{\partial x^2 \partial x^2} + \frac{\partial^2 g_{22}}{\partial x^1 \partial x^1}\right) \\ &= f_{11} f_{22} + f_{12}^2 - \frac{1}{2} \cdot 4 f_{12}^2 \\ &= f_{11} f_{22} - f_{12}^2. \end{aligned}$$

Hence in the given coordinates the sectional and Gaussian curvatures are both equal to $f_{11} f_{22} - f_{12}^2$ and thereby coincide. Lemma 4.5 is proven. □

In applications, of particular interest is the trace of the curvature tensor, the *Ricci tensor*,

$$R_{ik} = R^l_{ilk} = \frac{\partial \Gamma^l_{ik}}{\partial x^l} - \frac{\Gamma^l_{il}}{\partial x^k} + \Gamma^l_{ml} \Gamma^m_{ik} - \Gamma^l_{mk} \Gamma^m_{il}$$

rather than the curvature tensor itself.

The trace of the Ricci tensor is called the *scalar curvature*

$$R = g^{ik} R_{ik}.$$

Riemannian manifolds of constant sectional curvature have sufficiently simple structure. Every $n$-dimensional manifold of constant sectional curvature $K > 0$ in a neighborhood of each point looks like an $n$-dimensional sphere of radius $\frac{1}{\sqrt{K}}$. It means that a small neighborhood of each point can be mapped homeomorphically onto a domain on the sphere with the lengths of all curves preserved. Manifolds of constant negative curvature will be considered in §5.1.

**Problem 4.4.** Prove that for two-dimensional Riemannian manifolds the Ricci tensor takes the form

$$R_{ik} = 2Kg_{ik},$$

where $K$ is the Gaussian curvature; moreover, $R = 2K$.

**Problem 4.5.** Prove that for three-dimensional Riemannian or pseudo-Riemannian manifolds the curvature tensor is determined completely by the Ricci tensor as follows:

$$R_{ijkl} = R_{ik}g_{jl} + R_{jl}g_{ik} - R_{il}g_{jk} - R_{jk}g_{il} + \frac{R}{2}(g_{il}g_{jk} - g_{ik}g_{jl}).$$

**Problem 4.6.** Prove that for manifolds of dimension $n \geq 4$ the curvature tensor is not determined uniquely by the Ricci tensor.

**Problem 4.7.** (1) Prove the *Bianchi identity* for a connection on the tangent bundle:

$$\nabla_m R^i_{jkl} + \nabla_k R^i_{jlm} + \nabla_l R^i_{jmk} = 0$$

for arbitrary $i$, $j$, $k$, $l$, and $m$.

(2) Prove the Bianchi identity for a connection in an arbitrary vector bundle:

$$d^{\nabla}\Omega = 0,$$

where $\Omega$ is the curvature form and the operator $d^{\nabla}$ is defined by (4.12).

## 4.5 Geodesics

Suppose that an affine connection is defined on the tangent bundle of a smooth manifold $M$. A *geodesic* is a curve $x(t)$ satisfying the equation

$$\frac{D\dot{x}}{\partial t} = \nabla_{\dot{x}}\dot{x} = 0 \tag{4.14}$$

which is a second-order equation in local coordinates:

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0. \tag{4.15}$$

Introduce coordinates $x^1, \dots, x^n, v^1, \dots, v^n$ in the tangent bundle as in the proof of Theorem 3.1. In these coordinates, equation (4.15) becomes the following system of first-order ordinary differential equations:

$$\dot{x}^i = v^i, \quad \dot{v}^i = -\Gamma^i_{jk}(x)v^j v^k.$$

Applying the existence and uniqueness theorem for ordinary differential equations, we obtain the following assertion:

**Lemma 4.6.** *For each point $x \in M$ of a Riemannian manifold $M$, there exist a neighborhood $U$ of $x$ and a constant $\varepsilon > 0$ such that, for every point $y \in U$ and every vector $v \in T_y M$ of length $< \varepsilon$, there is a unique geodesic $\gamma\colon (-1,1) \to M$ satisfying the initial conditions*

$$\gamma(0) = y, \quad \dot{\gamma}(0) = v. \tag{4.16}$$

*Proof.* The theorem on existence and uniqueness of a solution to an ordinary differential equation formally implies that, for some neighborhood $W$ of a point $(x,0) \in TM$, there exists $\delta > 0$ such that for every point $(y,v) \in W$, there is a unique geodesic $\gamma_1\colon (-\delta,\delta) \to M$ with the initial data (4.16).

Now choose a constant $\varepsilon_1 > 0$ and a neighborhood $U \subset M$ of $x$ such that $(y,v) \in W$ for $y \in U$ and $|v| < \varepsilon_1$. It follows from (4.15) that if the curve $\tilde{\gamma}(t)$ is a geodesic, then for every constant $C$ the curve $\hat{\gamma}(t) = \tilde{\gamma}(Ct)$ is a geodesic, too. Now take $\varepsilon = \delta\varepsilon_1$ and $\gamma(t) = \gamma_1(\delta t)$. Lemma 4.6 is proven. □

It follows from Lemma 4.6 that, for every point $x \in M$ in a small ball $B_{x,\varepsilon} = \{v \in T_x M \mid |v| < \varepsilon\}$, we can define the *exponential map*

$$\exp_x\colon B_{x,\varepsilon} \to M,$$

which relates to a point $(x,v)$ the endpoint $\gamma(1)$ of the geodesic.

Below, we restrict our exposition to geodesics of symmetric connections compatible with the metrics. By analogy with the case of surfaces, we can prove that

- *the geodesics coincide precisely with solutions to the Euler–Lagrange equations with the Lagrangian $L(x,\dot{x}) = g_{ij}\dot{x}^i\dot{x}^j$*, where $g_{ij}$ is the Riemannian metric;
- *the geodesics are naturally parameterized ($|\dot{x}| = const$) extremal functions of the length functional.*

We apply the exponential map to prove the following fact:

**Lemma 4.7.** *For each point $x \in M$, there is a neighborhood $V$ and a constant $\eta > 0$ such that*

(1) *arbitrary pairs of points of $V$ are joined by a unique geodesic of length $\leq \eta$ and this geodesic depends smoothly on the endpoints;*

(2) *for each point $y \in V$, the map* $\exp_y$ *takes the ball $B_{y,\eta} \subset T_y M$ diffeomorphically onto a neighborhood of $y$ in $M$.*

*Proof.* Consider the function

$$F\colon \widetilde{U} \to M \times M,$$

where $\widetilde{U} \subset TM$ is the domain of points $(x,v)$ with $x \in U$ and $|v| < \varepsilon$ and $F(x,v) = (x, \exp_x(v))$. Its derivatives at $(x,0)$ have the form

$$\frac{\partial F^i}{\partial x^j} = \delta^i_j, \quad \frac{\partial F^{i+n}}{\partial x^j} = \delta^i_j, \quad \frac{\partial F^i}{\partial v^j} = 0, \quad \frac{\partial F^{i+n}}{\partial v^j} = \delta^i_j,$$

where $F = (F^1, \dots, F^{2n})$ and $i = 1, \dots, n$. Since the Jacobian matrix of this map at $(x, 0)$ is nondegenerate, this map is invertible in a small neighborhood of $(x, x)$ by the inverse function theorem.

Choose a domain $\widetilde{U}' \subset \widetilde{U}$ of points $(x, v)$, where $x \in U'$ and $|v| < \eta$, on which the map $F$ acts diffeomorphically onto the image. Now find a neighborhood $V \subset M$ of the point $x$ such that $V \times V \subset F(\widetilde{U}')$. It is easy to note that this neighborhood is the sought one, since the length of the segment of the geodesic $\gamma$ from $x$ to $\exp_x(v)$ is equal to $|v|$ (this follows from the arc length parameterization of the geodesic).

Lemma 4.7 is proven. □

In a neighborhood $V$ of a point $x$, introduce the coordinates which relate to $y \in M$ the coordinates $(v^1, \dots, v^n)$, where $\exp_x(v) = y$. These coordinates are called the *geodesic* coordinates.

**Lemma 4.8.** *In the geodesic coordinates, all Christoffel symbols at $x$ are zero:*

$$\Gamma^i_{jk}(x) = 0.$$

*Proof.* In the geodesic coordinates the equation for the geodesic $\gamma(t)$ with the initial data $\gamma(0) = x$ and $\dot{\gamma}(0) = v$ has the form $\gamma(t) = tv$. Writing down equation (4.15) for this geodesic, we obtain

$$\Gamma^i_{jk} v^j v^k = 0$$

along the geodesic. Since the direction $v$ at $x$ can be chosen arbitrary, the Christoffel symbols are identically zero at this point. Lemma 4.8 is proven. □

**Lemma 4.9.** *Let $\sigma$ be a two-dimensional plane in $T_xM$ and let $\Sigma$ be an embedded two-dimensional surface formed by the geodesics drawn in the directions tangent to $\sigma$ (i.e., $\Sigma = \exp_x(\sigma \cap B_{x,\eta})$) and endowed with the induced metric. Then the Gaussian curvature of $\Sigma$ at $x$ coincides with the sectional curvature of $M$ at $x$ in the two-dimensional direction $\sigma$.*

*Proof.* Choose a basis $\partial_1, \dots, \partial_n$ for $T_xM$ such that $(\partial_i, \partial_j) = g_{ij}(x) = \delta_{ij}$ and construct the geodesic coordinates in $V$ for this basis. Without loss of generality we may assume that $\sigma$ is spanned by $\partial_1$ and $\partial_2$.

Since all Christoffel symbols are zero and $g_{11}g_{22} - g_{12}^2 = 1$ at the point $x$, the Gaussian curvature $\Sigma$ at $x$ (calculated for the metric according to the Gauss theorem (see §2.5)) is equal to

$$K = \frac{\partial^2 g_{12}}{\partial x^1 \partial x^2} - \frac{1}{2}\frac{\partial^2 g_{11}}{\partial x^2 \partial x^2} - \frac{1}{2}\frac{\partial^2 g_{22}}{\partial x^1 \partial x^1}.$$

The expression for the sectional curvature $K(\sigma)$ becomes essentially simpler, too:

$$K(\sigma) = -\frac{\partial \Gamma^2_{12}}{\partial x^1} + \frac{\partial \Gamma^2_{11}}{\partial x^2}.$$

Expanding the last expression, using (4.8), we obtain

$$R_{1212} = \frac{\partial^2 g_{12}}{\partial x^1 \partial x^2} - \frac{1}{2}\frac{\partial^2 g_{11}}{\partial x^2 \partial x^2} - \frac{1}{2}\frac{\partial^2 g_{22}}{\partial x^1 \partial x^1}.$$

Lemma 4.9 is proven. □

For small $t$ the geodesics $\exp_x(tv)$ remain in the neighborhood $V$ (see Lemma 4.7); moreover, the spheres $\exp_x S(x,\tau)$, where $S_{x,\tau} = \{v \in T_xM \mid |v| = \tau < \eta\} = \partial B_{x,\tau}$, are submanifolds in $M$.

**Lemma 4.10.** *The geodesics* $\exp_x(tv)$ *are orthogonal to the spheres* $\exp_x(S_{x,\tau})$.

*Proof.* Let $v(s)$ be an arbitrary smooth curve on the sphere $S_{x,\eta/2}$. It suffices to prove that for the surface $f(s,t) = \exp_x(tv(s))$ the vector fields $f_t = \frac{\partial f}{\partial t}$ and $f_s = \frac{\partial f}{\partial s}$ are orthogonal everywhere.

First of all show that

$$\frac{D}{\partial t}(f_s, f_t) = 0.$$

Since the connection is compatible with the metric, we have

$$\frac{D}{\partial t}(f_s, f_t) = \left(\frac{D}{\partial t} f_s, f_t\right) + \left(f_s, \frac{D}{\partial t} f_t\right).$$

It follows from the symmetry of the connection that the following identity holds for every immersed surface $f(s,t)$, which fact is verified by decomposition in the basis:

$$\frac{D}{\partial t} f_s = \frac{D}{\partial s} f_t.$$

Since the curves $f(s,t)$ for fixed $s$ are geodesics, we obtain

$$\frac{Df_t}{\partial t} = 0,$$

and finally

$$\frac{D}{\partial t}(f_s, f_t) = \left(\frac{D}{\partial s} f_t, f_t\right) = \frac{1}{2}\frac{D}{\partial s}(f_t, f_t) = \frac{1}{2}\frac{d}{ds}|v(s)|^2 = 0.$$

But $f_s = 0$ at $t = 0$ and consequently $(f_s, f_t) = 0$ at $t = 0$ and hence everywhere. Lemma 4.10 is proven. □

This lemma is a particular case of the following more general assertion: suppose that $\Sigma$ is a hypersurface in $M$ and $f(s,t)$ is a family of geodesics parameterized by the points $s \in \Sigma$ and $t$; moreover, $f(s,0) \in \Sigma$ and the vector $\frac{\partial f(s,0)}{\partial t}$ is orthogonal to $\Sigma$

(i.e., orthogonal to the tangent space of $\Sigma$). Then these geodesics are orthogonal to the surfaces $\Sigma_t = \{f(s,t) \mid s \in \Sigma\}$ for all $t$.

Coordinates $x^1, \dots, x^n$ in a domain $W$ are called *semigeodesic* if the metric tensor satisfies the conditions $g_{in} = 0$ for $i = 1, \dots, (n-1)$ and $g_{nn} = 1$:

$$g_{ij} dx^i dx^j = \sum_{1 \le i,j \le n-1} g_{ij} dx^i dx^j + (dx^n)^2 . \tag{4.17}$$

**Lemma 4.11.** *If $x^1, \dots, x^n$ are semigeodesic coordinates, then for every vector $\xi \in \mathbb{R}^{n-1}$ the curve $(x^1, \dots, x^{n-1}) = \xi$ is a geodesic with the arc length parameter $t = x^n$.*

*Proof.* From the formulas for the Christoffel symbols we obtain

$$\Gamma^i_{nn} = \frac{1}{2} g^{ij} \left( 2 \frac{\partial g_{jn}}{\partial x^n} - \frac{\partial g_{nn}}{\partial x^j} \right) = 0$$

since all values $g_{in} = 0$ are constant. Similarly, we find that

$$\Gamma^n_{ij} = -\frac{1}{2} \frac{\partial g_{ij}}{\partial x^n} .$$

The geodesic equations take the form

$$\begin{aligned} \ddot{x}^i + \sum_{1 \le j,k \le n-1} \Gamma^i_{jk} \dot{x}^j \dot{x}^k + 2 \sum_{1 \le j \le n-1} \Gamma^i_{jn} \dot{x}^i \dot{x}^n + \Gamma^i_{nn} \dot{x}^n \dot{x}^n \\ = \ddot{x}^i + \sum_{1 \le j,k \le n-1} \Gamma^i_{jk} \dot{x}^j \dot{x}^k + 2 \sum_{1 \le j \le n-1} \Gamma^i_{jn} \dot{x}^i \dot{x}^n = 0 \end{aligned}$$

for $i = 1, \dots, n$. Obviously, every curve with $\dot{x}^1 = \dots = \dot{x}^{n-1} = 0$ satisfies these equations for $i = 1, \dots, n-1$. In this case the last equation for $i = n$ takes the form

$$\ddot{x}^n = 0$$

and determines the arc length parameter $t = x^n$ on the curve $(x^1, \dots, x^{n-1}) = \xi$. The lemma is proven. □

**Lemma 4.12.** *Let $x \in M$ and let $V$ be the neighborhood of Lemma* 4.7. *Then the following holds.*

(1) *In the deleted ball* $\exp_x(B_{x,\eta}) \setminus x$ *there are semigeodesic coordinates.*

(2) *Let $\gamma_y$ be the unique geodesic of length $\le \eta$ lying in $V$ and joining the points $x, y \in V$. For every piecewise smooth curve $\omega \colon [0,T] \to M$ joining $x$ and $y$ and lying in $V$ the length of $\omega$ is not less than the length of $\gamma_y$ and is equal to the former only if the curves $\omega$ and $\gamma_y$ coincide.*

*Proof.* Take $x^1, \dots, x^{n-1}$ to be the coordinates on the sphere $S_{x,1}$ and let $x^n$ be an arc length parameter on the geodesic starting at $x$. By construction, these coordinates have the form (4.17) and therefore are semigeodesic.[4]

The proof of assertion (2) is similar to that of Theorem 2.11. Write down $\dot{\omega}$ in the semigeodesic coordinates and put $w_1 = (\dot{\omega}^1, \dots, \dot{\omega}^{n-1}, 0)$ and $w_2 = (0, \dots, 0, \dot{\omega}^n)$. Let $y^n$ be the $n$th coordinate of the point $y$. Then

$$\mathrm{L}(\gamma_y) = y^n, \quad \int_0^T w_2 \, dt = y^n,$$

$$\mathrm{L}(\omega) = \int \sqrt{(w_1, w_1) + (w_2, w_2)} \, dt \geq \int \sqrt{(w_2, w_2)} \, dt \geq y^n$$

and, obviously, the lengths of $\omega$ attain their minimum in the case $w_1 = 0$ and $w_2 =$ const. In this case it equals $y^n$ and the curve coincides with $\gamma_y$.

Lemma 4.12 is proven. □

**Theorem 4.2.** *If a curve $\omega\colon [0, T] \to M$ parameterized by the arc length is not longer than any other curve from $\omega(0)$ to $\omega(T)$, then the curve is a geodesic.*

The proof of Theorem 4.2 is simple: for every point $x = \omega(t)$ a small neighborhood of the curve lies in a domain $V$ (see Lemmas 4.7 and 4.12) and has a length less than $\eta$. By Lemma 4.12, this segment is a geodesic. Consequently the whole curve $\omega$ is a geodesic.

In the case of the Euclidean space the geodesics are line segments. Theorem 4.2 claims that for a general Riemannian manifold the geodesics are the natural analog of straight lines as the shortest curves.

[4]For surfaces we gave another derivation of existence of semigeodesic coordinates (Lemma 2.10).

# 5

# The Lobachevskii plane and the Minkowski space

## 5.1 The Lobachevskii plane

Let $\mathbb{R}^2$ be the two-dimensional Euclidean plane with coordinates $x$ and $y$ (in tensor notation we assume that $x$ is the first coordinate and $y$ is the second).

Consider the upper half-plane $\mathcal{H} = \{(x, y) \mid y > 0\}$ and define on it another metric

$$g_{11} = g_{22} = \frac{1}{y^2}, \quad g_{12} = 0,$$

or, which is more convenient in the theory of surfaces,

$$\frac{dx^2 + dy^2}{y^2}. \tag{5.1}$$

Using (4.8), we can easily find that

$$\Gamma^2_{11} = \frac{1}{y}, \quad \Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -\frac{1}{y}$$

and the other Christoffel symbols are zero.

The geodesic equations take the form

$$\ddot{x} - \frac{2}{y}\dot{x}\dot{y} = 0, \quad \ddot{y} + \frac{1}{y}(\dot{x}^2 - \dot{y}^2) = 0.$$

We know that the length of a velocity vector is preserved (it follows from the definition of a geodesic that $\frac{D(\dot{\gamma},\dot{\gamma})}{\partial t} = 2\big(\frac{D\dot{\gamma}}{\partial t}, \dot{\gamma}\big) = 0$); therefore, the value

$$I_1 = \frac{\dot{x}^2 + \dot{y}^2}{y^2}$$

(the square of this length) is a first integral. Another first integral, as we can verify, is

$$I_2 = x + \frac{\dot{y}}{\dot{x}}y.$$

**Problem 5.1.** Suppose that $\gamma$ is a geodesic and its velocity vector $v(t) = \dot{\gamma}(t)$ at the point $\gamma(t)$ is not vertical. Draw the straight line $l(t)$ through the point $\gamma(t)$ (in the Euclidean metric on $\mathbb{R}^2$) which is orthogonal to the vector $v(t)$. Then the $x$-coordinate of the intersection point of $l(t)$ and the $Ox$-axis is equal to $I_2$.

**Theorem 5.1.** *The geodesics of the Lobachevskii plane* (*in terms of the Euclidean metric on* $\mathbb{R}^2$) *are*

(1) *rays orthogonal to the* $Ox$*-axis;*

(2) *half-circles lying in the upper half-plane and crossing the* $Ox$*-axis at the angle* $\pi/2$.

**Problem 5.2.** Prove Theorem 5.1.

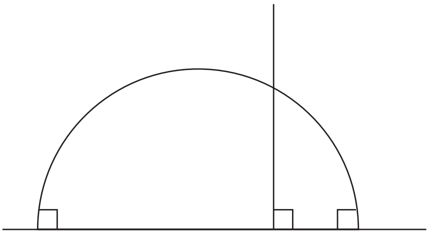

Figure 5.1. Geodesics on the Lobachevskii plane.

From Theorem 5.1 we obtain

**Lemma 5.1.** *Arbitrary pairs of points of* $\mathcal{H}$ *are joined by a unique geodesic.*

The Lobachevskii plane possesses the following fundamental property:

**Theorem 5.2.** *The sectional curvature of the space* $\mathcal{H}$ *is constant and equal to* $K = -1$.

**Problem 5.3.** Prove Theorem 5.2.

It is essential that we introduce an abstract Riemannian metric on the upper half-plane, since, by the Hilbert theorem, there is no immersion of the upper half-plane into $\mathbb{R}^3$ such that the induced metric would be (5.1).

The group $\mathrm{PSL}(2,\mathbb{R}) = \mathrm{SL}(2,\mathbb{R})/\mathbb{Z}_2$ acts on $\mathcal{H}$. The group $\mathrm{SL}(2,\mathbb{R})$ consists of all real $(2 \times 2)$-matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with determinant 1:

$$ad - bc = 1. \tag{5.2}$$

It contains the subgroup $\mathbb{Z}_2$ constituted by the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

and this subgroup is normal. The quotient group $SL(2, \mathbb{R})/\mathbb{Z}_2$ is denoted by $PSL(2, \mathbb{R})$. Its action on $\mathcal{H}$ has the form

$$z = x + iy \to A(z) = \frac{az + b}{cz + d}, \quad A \in PSL(2, \mathbb{R}). \tag{5.3}$$

Indeed, the following assertions can be verified straightforwardly:

**Problem 5.4.** Show that

(1) if $\operatorname{Im} z > 0$ then $\operatorname{Im} \frac{az+b}{cz+d} > 0$;

(2) $A_2(A_1(z)) = (A_2 \cdot A_1)(z)$;

(3) the matrix

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

determines the transformation inverse to *(*5.3*)*;

(4) the subgroup $\Gamma$ generated by the matrices

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

consists precisely of those elements which leave the point $(0, 1)$ fixed.

Choose the following parameterization of the group $SL(2, \mathbb{R})$. Since $a^2 + c^2 \neq 0$, we can assume that

$$a = r\cos\varphi, \quad c = r\sin\varphi.$$

Then condition (5.2) takes the form $r\cos\varphi \cdot d - r\sin\varphi \cdot b = 1$ and its general solution depending on one parameter $s \in \mathbb{R}$ has the form

$$b = -r^{-1}\sin\varphi + s\cos\varphi, \quad d = r^{-1}\cos\varphi + s\sin\varphi.$$

The parameters $r$, $s$, and $\varphi$ determine uniquely the elements of $SL(2, \mathbb{R})$:

$$\begin{pmatrix} r\cos\varphi & -\frac{1}{r}\sin\varphi + s\cos\varphi \\ r\sin\varphi & \frac{1}{r}\cos\varphi + s\sin\varphi \end{pmatrix},$$

where $r > 0$, $s \in \mathbb{R}$, and the parameter $\varphi$ is defined modulo $2\pi$.

**Problem 5.5.** If we consider $a$, $b$, $c$, and $d$ as real coordinates in the space $\mathbb{R}^4$ identified with the space of real $(2 \times 2)$-matrices, then equation (5.2) determines a smooth submanifold $SL(2, \mathbb{R}) \subset \mathbb{R}^4$ and the parameters $r$, $s$, and $\varphi$ constructed above are smooth coordinates.

Note that in §1.4 we proved a similar assertion for the groups $O(n)$.

Find the transformations in $PSL(2, \mathbb{R})$ which take a point $z = \lambda + i\mu$ to the point $i$. The corresponding matrices in $SL(2, \mathbb{R})$ are selected by the conditions

$$\lambda = -\frac{ab + cd}{a^2 + c^2}, \quad \mu = \frac{1}{a^2 + c^2}$$

and, using the above parameterization, we find that these are precisely the matrices with

$$r = \frac{1}{\sqrt{\mu}}, \quad s = -\frac{\lambda}{\sqrt{\mu}}. \tag{5.4}$$

**Lemma 5.2.** *Every point of $\mathcal{H}$ can be transformed to the point* $(0, 1)$ *and hence to any other point of $\mathcal{H}$ by a transformation of* $PSL(2, \mathbb{R})$. *The space $\mathcal{H}$ is naturally identified with the space* $PSL(2, \mathbb{R})/\Gamma$ *of left conjugacy classes* $PSL(2, \mathbb{R})$ *by the subgroup* $\Gamma$.

*Proof.* By (5.4), for every point $z = \lambda + i\mu \in \mathcal{H}$, we construct a transformation $A \in PSL(2, \mathbb{R})$ such that $A(z) = i$. Let $z_1, z_2 \in \mathcal{H}$ and $A_1(z_1) = A_2(z_2) = i$. Then $A_2^{-1}A_1(z_1) = z_2$.

Suppose that $B_1(i) = B_2(i)$. Then $B_2^{-1}B_1(i) = i$ and hence $B_2^{-1}B_1 \in \Gamma$. Obviously, if $B_2^{-1}B_1 \in \Gamma$ then $B_1(i) = B_2(i)$. The condition $B_2^{-1}B_1 \in \Gamma$ is equivalent to the fact that $B_2\Gamma = B_1\Gamma$; i.e., these left conjugacy classes coincide. Since $PSL(2, \mathbb{R})(i) = \mathcal{H}$, it follows that $\mathcal{H} = PSL(2, \mathbb{R})/\Gamma$.

Lemma 5.2 is proven. □

Assume that an *action of a group G* on a manifold $M$ is given; i.e., a map

$$G \times M \to M \tag{5.5}$$

of the form $(g, x) \to g(x)$ such that $g_1(g_2(x)) = (g_1 \cdot g_2)(x)$ for arbitrary $g_1, g_2 \in G$ and $e(x) = x$, where $e$ is the identity of the group $G$.

Also suppose that

(1) the group $G$ is a smooth manifold (see Chapter 7);
(2) the map (5.5) is smooth;
(3) for every pair $x_1, x_2 \in M$, there is at least one element $g \in G$ such that $g(x_1) = x_2$.

If the action satisfies condition (3), then it is called *transitive*.

If these conditions are satisfied, then $M$ is called the *homogeneous space* of the group $G$ and is identified with the space $G/\Gamma_x$ of left conjugacy classes of the group $G$ by the stationary subgroup of an arbitrary point $x$ ($\Gamma_x = \{g \in G \mid g(x) = x\}$).

Under some natural additional conditions,[1] the homogeneous space carries the canonical structure of a smooth manifold (see, e.g., [4]).

Now, we can restate Lemma 5.2 as follows:

[1]For example, if $G$ is a compact Lie group and $H = \Gamma_x$ is a closed subgroup in $G$ (see Chapter 7).

- *$\mathcal{H}$ is the homogeneous space of the group* PSL(2, $\mathbb{R}$).

In Riemannian geometry we are interested in the case when the group $G$ acts by isometries. A map $f: M \to M$ is called an *isometry* if it preserves the length of every smooth curve $\gamma$: $\mathrm{L}(f(\gamma)) = \mathrm{L}(\gamma)$. Obviously, this is equivalent to the fact that the map $f_*: TM \to TM$ preserves the lengths of all vectors.

**Lemma 5.3.** *The group* PSL(2, $\mathbb{R}$) *acts on $\mathcal{H}$ by isometries.*

*Proof.* Write the metric (5.1) in the form

$$\frac{dz\,d\bar{z}}{(\operatorname{Im} z)^2}.$$

A transformation $A$ preserves the lengths of vectors if and only if

$$\frac{dz\,d\bar{z}}{(\operatorname{Im} z)^2} = \frac{dA(z)\,d\overline{A(z)}}{(\operatorname{Im} A(z))^2}. \tag{5.6}$$

Indeed, every vector $v \in T_z M$ is written in the form $w\partial + \bar{w}\bar{\partial}$, where $\partial = \frac{1}{2}(\partial_x - i\,\partial_y)$. The map $A_*$ acts on tangent vectors as follows:

$$A_*(\partial) = \frac{\partial A(z)}{\partial z}\partial, \quad A_*(\bar{\partial}) = \frac{\partial \overline{A(z)}}{\partial \bar{z}}\bar{\partial},$$

$$|A_*(v)|^2 = \frac{\partial A(z)}{\partial z}\frac{\partial \overline{A(z)}}{\partial \bar{z}}\frac{w\bar{w}}{(\operatorname{Im} A(z))^2} \quad \text{and} \quad |v|^2 = \frac{w\bar{w}}{(\operatorname{Im} z)^2};$$

i.e., condition (5.6) is proven.

It is immediately verified that

$$\operatorname{Im} A(z) = \frac{\operatorname{Im} z}{|cz+d|^2}, \quad \frac{\partial A(z)}{\partial z} = \frac{1}{(cz+d)^2}.$$

Inserting these expressions into the right-hand side of (5.6), we establish that every $A \in \mathrm{PSL}(2, \mathbb{R})$ preserves the lengths of vectors.

Lemma 5.3 is proven. □

**Theorem 5.3.** *The group* PSL(2, $\mathbb{R}$) *consists of all isometric transformations of $\mathcal{H}$ preserving the orientation.*

*Proof.* Every isometry takes geodesics to geodesics. Let $T$ be an isometry of the space $\mathcal{H}$ which preserves the orientation and is such that $B(i) = z$.

Take a map $A \in \mathrm{PSL}(2, \mathbb{R})$ such that $A(i) = z$. Then $BA^{-1}$ is an isometry which leaves the point $i$ fixed. By Lemma 5.1, each point of $\mathcal{H}$ is joined with the point $i$ by a unique geodesic. Therefore, since an isometry takes geodesics to geodesics, the map $BA^{-1}$ is determined completely by the generated rotation $T_i\mathcal{H} \to T_i\mathcal{H}$ of the tangent space at the point $i$ and each such rotation determines uniquely the isometry. But all these isometries belong to the group $\Gamma$ and thereby $BA^{-1} \in \Gamma$. Hence, $B \in \mathrm{PSL}(2, \mathbb{R})$. Theorem 5.3 is proven. □

The Lobachevskii plane appears in many fields of mathematics. We will only point out one of the most important properties of the Lobachevskii plane, namely it is a model for non-Euclidean geometry.

The "Euclid's fifth postulate" claims:

- *Let $l$ be a straight line on the plane and let $x$ be a point on the plane not belonging to $l$. Then there is a unique straight line parallel to the straight line $l$ passing through $x$.*

The attempts to derive this axiom from the other postulates of the Euclidean geometry failed and finally led to development of non-Euclidean geometries. Namely, in 1825 Lobachevskii introduced the geometry with Lobachevskii's postulate:

- *Let $l$ be a straight line on the plane and let $x$ be a point on the plane not belonging to $l$. Then there are infinitely many straight lines parallel to $l$ passing through $x$.*

Starting with this assertion, Lobachevskii proved that there should be infinitely many such straight lines and that the sum of the angles of a triangle in such geometry must be less than $\pi$. Later, Bolyai independently arrived at the same conclusions.

The Lobachevskii plane is a model for such geometry.

Indeed, considering the geodesics as straight lines and assuming that straight lines are parallel if they do not intersect, we conclude that on the space $\mathcal{H}$ homeomorphic to the two-dimensional plane we implement the geometry which satisfies all Euclid's postulates except the fifth one. It follows from Theorem 5.1 that Lobachevskii's postulate holds in this geometry. Moreover, the $Ox$-axis is a "straight line at infinity": the length of every curve approaching this axis is infinite: this follows from the fact the integral $\int \frac{dy}{y}$ diverges as the lower limit tends to zero.

It is easy to verify that the sums of the angles of geodesic triangles are less than $\pi$. However, this follows from the Gauss–Bonnet formula whose proof given in §2.8 is translated without changes to arbitrary two-dimensional Riemannian manifolds.

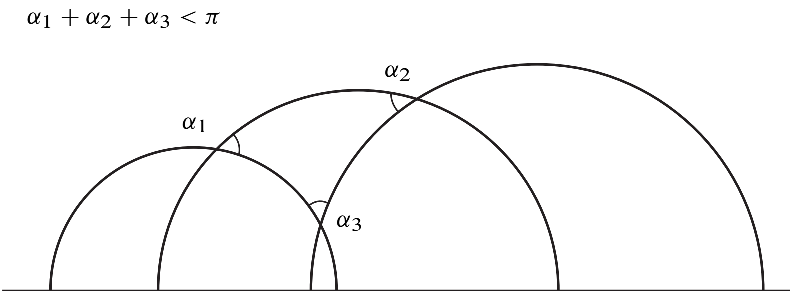


Figure 5.2. A geodesic triangle on the Lobachevskii plane.

The Lobachevskii space (or the *hyperbolic space*) $\mathcal{H}^n$ of dimension $n \geq 2$ is the upper half-plane $x_n \geq 0$ in $\mathbb{R}^n$ with the Riemannian metric

$$ds^2 = \frac{dx_1^2 + \cdots + dx_n^2}{x_n^2}$$

and coordinates $x_1, \ldots, x_n$. For $n = 2$ we obtain the Lobachevskii plane. It is easy to figure out that the sectional curvature of the space $\mathcal{H}^n$ is constant and equals $-1$. Every space of constant sectional curvature $K = -1$ is locally isometric to the Lobachevskii space of the same dimension and is the homogeneous space of the Lobachevskii space of the same dimension under the discrete action of some group.

## 5.2 Pseudo-Euclidean spaces and their applications in physics

Let $\mathbb{R}^{1,n}$ be the pseudo-Euclidean space with coordinates $x^0, x^1, \ldots, x^n$ and the metric

$$(dx^0)^2 - (dx^1)^2 - \cdots - (dx^n)^2. \tag{5.7}$$

In the tangent space at each point we have the indefinite inner product

$$(v, w)_{1,n} = v^0 w^0 - v^1 w^1 - \cdots - v^n w^n.$$

The linear transformations which preserve this inner product constitute the group $\mathrm{O}(1, n)$. The groups $\mathrm{O}(1, n)$ generalize the groups $\mathrm{O}(n)$ and their elements are given by $((n+1) \times (n+1))$-matrices $A$ satisfying the conditions

$$A^\top \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1_n \end{pmatrix} \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & -1_n \end{pmatrix}, \tag{5.8}$$

where $1_n$ is the identity $(n \times n)$-matrix (the proof of this fact is similar to that of Lemma 1.4).

**Problem 5.6.** Prove that equations (5.8) determine a smooth manifold of dimension $\frac{n(n+1)}{2}$ in the space of $(n+1) \times (n+1)$-matrices (Hint: see the proof of the similar fact for the groups $\mathrm{O}(n)$ in §1.4).

The space $\mathbb{R}^{1,3}$ arises in physics as the "space of events" in *special relativity*, the *spacetime* (also called the *Minkowski space*). Therewith the metric

$$(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$$

is called the *Minkowski metric*. The coordinates $x^1$, $x^2$, and $x^3$ are the spatial coordinates and $x^0 = ct$ is the time coordinate (here $c$ is the speed of light and $t$ is time). Therefore, if $(v, v)_{1,3} < 0$ then a vector $v$ is called *space-like*; a vector $v$ is called *time-like* if $(v, v)_{1,3} > 0$; and if $(v, v)_{1,3} = 0$, then $v$ is called *light-like*. The last definition is grounded on the following physical base:

- *the speed of light is constant and equal to c in every inertial coordinate system.*

Therefore, if $x(t)$ is a curve in $\mathbb{R}^{1,3}$ along which a light ray propagates, then

$$c^2 - \left(\frac{dx^1}{dt}\right)^2 - \left(\frac{dx^2}{dt}\right)^2 - \left(\frac{dx^3}{dt}\right)^2 = 0.$$

The group $\mathrm{O}(n)$ has two connected components constituted by orientation-preserving transformations and transformations which change the orientation (i.e., those with $\det A = 1$ and $\det A = -1$). The group $\mathrm{O}(1, n)$ has four components: each family of transformations preserving or changing the orientation splits also into two components depending on whether the transformation changes the "time direction" or not: $(e_0, A(e_0))_{1,3} < 0$ or $(e_0, A(e_0))_{1,3} > 0$, where $e_0$ is the basis vector corresponding to the coordinate $x^0$.

**Problem 5.7.** Show that the group $\mathrm{O}(1, 1)$ consists of the transformations given by the matrices

$$\begin{pmatrix} \varepsilon & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} \cosh\varphi & \sinh\varphi \\ \sinh\varphi & \cosh\varphi \end{pmatrix}, \quad \varphi \in \mathbb{R},$$

where $\varepsilon, \delta = \pm 1$, $\cosh\varphi = (e^{\varphi} + e^{-\varphi})/2$ and $\sinh\varphi = (e^{\varphi} - e^{-\varphi})/2$. For $\varepsilon > 0$ the transformation preserves the time direction and for $\varepsilon < 0$, changes. For $\varepsilon\delta = 1$ the transformation preserves the orientation, and for $\varepsilon\delta = -1$, changes.

Consider the following problem. Assume that we have two frames: $K$ with coordinates $ct$, $x^1$, $x^2$, and $x^3$ and $\tilde{K}$ with coordinates $c\tilde{t}$, $\tilde{x}^1$, $\tilde{x}^2$, and $\tilde{x}^3$. Suppose that for $t = \tilde{t} = 0$ they coincide and the frame $\tilde{K}$ moves along the $Ox^1$-axis with a constant velocity $v$ with respect to the frame $K$ (the frame $\tilde{K}$ is inertial). Find formulas for transition from one frame to the other.

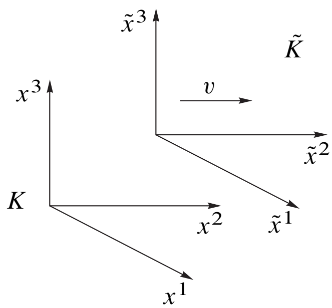


Figure 5.3. A frame $\tilde{K}$ moving uniformly along the $Ox$-axis.

In Galilean mechanics, time is universal and the transition formulas have the following simple form (*Galilean transformations*)

$$t = \tilde{t}, \quad x^1 = \tilde{x}^1 + vt, \quad x^2 = \tilde{x}^2, \quad x^3 = \tilde{x}^3.$$

In special relativity, we postulate that the distance (in metric (5.7)) between events, i.e., points of the spacetime, is preserved under the passage to another inertial frame. It follows that such a passage is given by an element of the *Poincaré group*, the group of all motions of $\mathbb{R}^{1,3}$. This group is generated by translations and elements of $\mathrm{O}(1,3)$.

Since the origins coincide for $t = \tilde{t}$, the passage from $K$ to $\widetilde{K}$ is given by a transformation of $\mathrm{O}(1,3)$. Decreasing continuously the velocity $v$ to zero, we obtain the identity transformation; hence the sought transformation preserves also the orientation and the time direction (belongs to the same component as the identity transformation). Since this transformation relates only to the coordinates $t$, $x^1$ and $\tilde{t}$, $\tilde{x}^1$, it has the form

$$ct = c\tilde{t}\cosh\varphi + \tilde{x}^1\sinh\varphi, \quad x^1 = c\tilde{t}\sinh\varphi + \tilde{x}^1\cosh\varphi, \quad x^2 = \tilde{x}^2, \quad x^3 = \tilde{x}^3.$$

For $\tilde{x}^1 = 0$,

$$ct = c\tilde{t}\cosh\varphi, \quad x^1 = c\tilde{t}\sinh\varphi$$

and, eventually, we obtain

$$\frac{x^1}{ct} = \tanh\varphi.$$

But the point $\tilde{x}^1 = \tilde{x}^2 = \tilde{x}^3 = 0$ moves along the $Ox^1$-axis with the constant velocity $v$ and hence

$$\frac{x^1}{ct} = \frac{v}{c} = \tanh\varphi.$$

Since

$$\cosh\varphi = \frac{1}{\sqrt{1-\tanh^2\varphi}}, \quad \sinh\varphi = \frac{\tanh\varphi}{\sqrt{1-\tanh^2\varphi}},$$

we obtain the final form of the *Lorentz transformations*:

$$t = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(\tilde{t} + \tilde{x}^1\frac{v}{c^2}\right), \quad x = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(\tilde{x}^1 + v\tilde{t}).$$

If $v$ is very small as compared with the speed of light,

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \approx 1,$$

then the Galilean transformations give a rather nice approximation to the Lorentz transformations.

Indicate some obvious consequence of the obtained formulas known from popular exposition of special relativity.

The rest length of an object is the length in a frame in which it is at rest. Suppose that a rod has a rest length $\Delta l$ being at rest in a frame $K$ and its ends have coordinates $x^1 = a, b$, where $b - a = \Delta l$. In has a different length in the frame $\widetilde{K}$: it follows that, according to the Lorentz transformations, at each fixed time $\tilde{t}$ the difference of the $\tilde{x}^1$-coordinates of its ends is equal to

$$\Delta\tilde{l} = \sqrt{1 - \frac{v^2}{c^2}} \cdot \Delta l;$$

i.e.,

- *the rod contracts in the direction of motion of the frame* $\widetilde{K}$.

Another “paradox” is the time dilation.

Suppose that a clock rests in a frame $\widetilde{K}$. The time difference between two events at the same points of the space ($\tilde{x}^1$, $\tilde{x}^2$, and $\tilde{x}^3$ are fixed) in this frame is equal to $\Delta\tilde{t}$, while the time difference between these events in the frame $K$ is equal to

$$\Delta t = \frac{\Delta\tilde{t}}{\sqrt{1 - \frac{v^2}{c^2}}};$$

i.e.,

- *time in a moving frame delays.*

General relativity, created in Einstein’s works, required already the whole machinery of Riemannian geometry and essentially stimulated its development. According to this theory, a gravitational field is a pseudo-Riemannian metric $g_{ij}$ on the four-dimensional space $M^4$. The metric has signature $(1,3)$, i.e., at each point (but not in the whole domain in general) there is a change of coordinates which reduces it to the Minkowski metric

$$g_{ij}dx^i dx^j = (dx^0)^2 - \sum_{i=1}^{3}(dx^i)^2.$$

The space $M^4$ is the space of events and can differ topologically from $\mathbb{R}^4$. The metric itself must satisfy some nonlinear equations called the *Einstein equations*. They claim that in vacuum (in the absence of other fields) the Ricci tensor of the gravitational field is zero,

$$R_{ik} = 0.$$

If other fields are given (for example, electromagnetic) then the Einstein equations take the form

$$R_{ik} - \frac{1}{2}Rg_{ik} = T_{ik},$$

where $R = g^{ik}R_{ik}$ is the *scalar curvature* of the metric $g_{ik}$ and $T_{ik}$ is the so-called *energy-stress tensor* of these fields. This equation describes the interaction of the gravitational field with the other fields, while the behavior of the other fields themselves is described by special equations (for example, by the *Maxwell equations* in the case of an electromagnetic field).

**Problem 5.8.** Suppose that $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$, $t$ is a coordinate in $\mathbb{R}$, and $r$, $\theta$, and $\varphi$ are the spherical coordinates in $\mathbb{R}^3$ (see §2.7). Show that the metric[2]

$$dl^2 = \left(1 - \frac{r_g}{r}\right) c^2 dt^2 - \frac{dr^2}{1 - \frac{r_g}{r}} - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \quad r_g = \text{const} > 0,$$

satisfies the Einstein vacuum field equations $R_{ik} = 0$.

[2]This metric is called the *Schwarzschild metric* and describes stationary black holes.

# Part III

# Supplement chapters

# 6

# Minimal surfaces and complex analysis

## 6.1 Conformal parameterization of surfaces

Let $M$ and $\widetilde{M}$ be Riemannian manifolds with respective metrics $g_{jk}$ and $\tilde{g}_{jk}$. A smooth map $f: M \to \widetilde{M}$ is *conformal* if it preserves the angles between tangent vectors. Formally, this is written as follows:

$$\tilde{g}_{lm}(f(x))\frac{\partial f^l(x)}{\partial x^j}\frac{\partial f^m(x)}{\partial x^k}dx^j dx^k = \lambda(x)g_{jk}(x)dx^j dx^k,$$

where $\lambda(x)$ is a function on the manifold $M$, $f(x) = (f^1(x), \dots, f^q(x))$, $x^1, \dots, x^p$ are local coordinates on the manifold $M$, $p = \dim M$, and $q = \dim \widetilde{M}$. It is obvious that $p \leq q$.

Riemannian manifolds are *conformally equivalent* if the map $f$ is a diffeomorphism.

A Riemannian manifold of dimension $n$ is *conformally Euclidean* if a neighborhood of each point is conformally equivalent to a domain of the $n$-dimensional Euclidean space. In this case the linear coordinates $x^1, \dots, x^n$ in the Euclidean space determine coordinates on the manifold in which the metric takes the form $g_{jk} = \lambda(x)\delta_{jk}$ (such coordinates are called *conformally Euclidean coordinates*).

**Theorem 6.1.** *Each two-dimensional Riemannian manifold is conformally Euclidean.*

This theorem cannot be generalized to manifolds of greater dimension. We begin the proof of the theorem with the following lemma:

**Lemma 6.1.** *Let $u$ and $v$ be local coordinates in a neighborhood of a point $p$ on a surface $\Sigma$ and let $Edu^2 + 2Fdudv + Gdv^2$ be a metric on $\Sigma$. Suppose that, in a neighborhood of $p$, there exist conformally Euclidean coordinates $x$ and $y$. Then they satisfy the following two equivalent systems of equations:*

$$\begin{aligned}\frac{\partial x}{\partial u} &= \frac{1}{\sqrt{EG - F^2}}\left(\frac{\partial y}{\partial v}E - \frac{\partial y}{\partial u}F\right),\\ \frac{\partial x}{\partial v} &= \frac{1}{\sqrt{EG - F^2}}\left(\frac{\partial y}{\partial v}F - \frac{\partial y}{\partial u}G\right),\end{aligned} \tag{6.1}$$

$$\begin{aligned}\frac{\partial y}{\partial u} &= \frac{1}{\sqrt{EG - F^2}}\left(-\frac{\partial x}{\partial v}E + \frac{\partial x}{\partial u}F\right),\\ \frac{\partial y}{\partial v} &= \frac{1}{\sqrt{EG - F^2}}\left(-\frac{\partial x}{\partial v}F + \frac{\partial x}{\partial u}G\right).\end{aligned} \tag{6.2}$$

*Proof of Lemma* 6.1. Rewrite the equality

$$\lambda(dx^2 + dy^2) = Edu^2 + 2Fdudv + Gdv^2$$

as

$$\lambda\left(\left(\frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv\right)^2 + \left(\frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv\right)^2\right) = Edu^2 + 2Fdudv + Gdv^2.$$

It is equivalent to the system of equations

$$\begin{aligned} \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 &= \lambda^{-1}E, \\ \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} &= \lambda^{-1}F, \\ \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 &= \lambda^{-1}G, \end{aligned}$$

which is written in the form of two systems of linear equations

$$A\begin{pmatrix} \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial u} \end{pmatrix} = \lambda^{-1}\begin{pmatrix} E \\ F \end{pmatrix}, \quad A\begin{pmatrix} \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial v} \end{pmatrix} = \lambda^{-1}\begin{pmatrix} F \\ G \end{pmatrix}, \tag{6.3}$$

where

$$A = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

The area element is the 2-form

$$\sqrt{EG - F^2}du \wedge dv = \lambda dx \wedge dy = \lambda\left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right)du \wedge dv.$$

Hence,

$$\lambda \det A = \sqrt{EG - F^2}.$$

Therefore,

$$A^{-1} = \frac{\lambda}{\sqrt{EG - F^2}}\begin{pmatrix} \frac{\partial y}{\partial v} & -\frac{\partial y}{\partial u} \\ -\frac{\partial x}{\partial v} & \frac{\partial x}{\partial u} \end{pmatrix}$$

and, solving formally systems (6.3), we obtain (6.1) and (6.2).

Solving (6.1) for $\frac{\partial y}{\partial u}$ and $\frac{\partial y}{\partial v}$, we obtain (6.2). Similarly, solving (6.2) for $\frac{\partial x}{\partial u}$ and $\frac{\partial x}{\partial v}$, we obtain (6.1). Consequently, these systems are equivalent. Lemma 6.1 is proven. □

To draw the analogy to the multidimensional case, denote $u = w^1$, $v = w^2$, $g_{11} = E, g_{12} = F$, and $g_{22} = G$. Introduce the following differential operator acting on functions on a Riemannian manifold:

$$\Delta\varphi = \frac{1}{\sqrt{g}} \frac{\partial}{\partial w^j}\left(\sqrt{g} g^{jk} \frac{\partial \varphi}{\partial w^k}\right),$$

where $g = \det g_{jk}$. It is called the *Laplace–Beltrami operator*. For the Euclidean metric it coincides with the usual *Laplace operator*

$$\Delta = \sum_j \frac{\partial^2}{(\partial w^j)^2}. \tag{6.4}$$

This operator depends on the Riemannian metric on the manifold and is independent of the choice of coordinates. The latter is established as follows. Let $p$ be an arbitrary point on the manifold and let $U$ be its neighborhood with coordinates $x^1, \ldots, x^n$. Consider all possible smooth functions on the manifold vanishing outside $U$. Define the inner product of their gradients by

$$(\operatorname{grad}\varphi, \operatorname{grad}\psi) = \int_U g^{jk} \frac{\partial \varphi}{\partial x^j} \frac{\partial \psi}{\partial x^k}\, d\sigma,$$

where $d\sigma = \sqrt{g} dx^1 \wedge \cdots \wedge dx^n$. It is obvious that the value of the integral is independent of the choice of coordinates, since such is the integrand. Integrating by parts, we obtain

$$(\operatorname{grad}\varphi, \operatorname{grad}\psi) = -\int_U \Delta\varphi \cdot \psi\, d\sigma.$$

Since the function $\psi$ is arbitrary, the value $\Delta\varphi$ is independent of the choice of coordinates everywhere.

**Lemma 6.2.** *If $x$ and $y$ are twice continuously differentiable functions satisfying system* (6.1) (*or the equivalent system* (6.2)) *then*

$$\Delta x = \Delta y = 0. \tag{6.5}$$

The equation $\Delta\varphi = 0$ is called the *Beltrami equation* and its solutions are called *harmonic functions* on the Riemannian manifold. In the case of the Euclidean space we obtain the usual Laplace equation and harmonic functions.

*Proof of Lemma* 6.2. Since the function $x(u, v)$ is twice continuously differentiable, we have

$$\frac{\partial^2 x}{\partial u\, \partial v} = \frac{\partial^2 x}{\partial v\, \partial u}$$

or, by (6.1),

$$\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{EG - F^2}}\left(\frac{\partial y}{\partial v} F - \frac{\partial y}{\partial u} G\right)\right) - \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{EG - F^2}}\left(\frac{\partial y}{\partial v} E - \frac{\partial y}{\partial u} F\right)\right) = 0.$$

It is easy to verify by straightforward calculations that the latter equation has the form $\Delta y = 0$. Similarly, we prove that $\Delta x = 0$. Lemma 6.2 is proven. □

We give the following lemma without a proof, which can be obtained in the framework of the theory of elliptic partial differential equations.

**Lemma 6.3.** *If, in a neighborhood of a point $p$, the coefficients $E$, $F$, and $G$ of the metric tensor are $n$ times continuously differentiable with respect to $u$ and $v$ (i.e., belong to the class $C^n$), where $n \geq 2$, then in some sufficiently small neighborhood of $p$, there exist*

(1) *a solution $x(u, v)$ of class $C^n$ to equation* (6.5) *such that*

$$\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial x}{\partial v}\right)^2 \neq 0$$

*in this neighborhood;*

(2) *a solution $y(u, v)$ of class $C^n$ to equation* (6.2).

For the Euclidean metric this assertion is obvious. For the proof it suffices to construct $x(u, v)$ and then $y(u, v)$ is found from first-order linear differential equations.

**Lemma 6.4.** *If $x(u, v)$ and $y(u, v)$ satisfy equations* (6.1) *and* (6.2)*, then the Jacobian of the map $(u, v) \to (x, y)$ is equal to*

$$\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} = \frac{1}{\sqrt{EG - F^2}}(V, V),$$

*where $V = \left(\frac{\partial x}{\partial v}, -\frac{\partial x}{\partial u}\right)$ and the inner product is taken with respect to the metric on the surface.*

The proof of Lemma 6.4 is obtained by straightforward substitution:

$$\begin{aligned}
&\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} \\
&\quad = \frac{\partial x}{\partial u} \cdot \frac{1}{\sqrt{EG - F^2}}\left(-\frac{\partial x}{\partial v}F + \frac{\partial x}{\partial u}G\right) - \frac{\partial x}{\partial v} \cdot \frac{1}{\sqrt{EG - F^2}}\left(-\frac{\partial x}{\partial v}E + \frac{\partial x}{\partial u}F\right) \\
&\quad = \frac{1}{\sqrt{EG - F^2}}\left(G\left(\frac{\partial x}{\partial u}\right)^2 - 2F\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + E\left(\frac{\partial x}{\partial v}\right)^2\right) \\
&\quad = \frac{1}{\sqrt{EG - F^2}}(V, V).
\end{aligned}$$

Lemma 6.4 is proven.

Now, the proof of Theorem 6.1 is plain: using Lemma 6.3, find $C^n$-smooth solutions $x(u, v)$ and $y(u, v)$ to equations (6.1) and (6.2). By the choice of the solution $x(u, v)$, the vector $V$ is nowhere zero in a neighborhood $U$ of the point $p \in \Sigma$. Therefore, by

Lemma 6.4, the Jacobian of the map $(u, v) \to (x, y)$ is nowhere zero in $U$. Moreover, by the inverse function theorem, we can choose the neighborhood $U$ so small that the map is invertible in it: the functions $x$ and $y$ determine new local coordinates. It follows from (6.1) and (6.2) that in these coordinates the metric takes the form $\lambda(dx^2 + dy^2)$; i.e., the map $(u, v) \in U \to (x, y)$ determines a conformal equivalence.

Theorem 6.1 is proven.

Having conformally Euclidean coordinates $x$ and $y$ on a surface, we construct the complex-valued *conformal parameter*

$$z = x + iy$$

whose usage connects the theory of surfaces and complex analysis.

## 6.2 The theory of surfaces in terms of the conformal parameter

Let $\boldsymbol{r}: U \to \mathbb{R}^3$ be a regular surface in the three-dimensional Euclidean space. By Theorem 6.1, in a neighborhood of every point of the surface we can choose the conformal parameter $z = x + iy$. For simplicity, assume that the parameter $z$ is given in a domain $U \subset \mathbb{C}$.

The conformality condition means that the first fundamental form has the shape

$$\boldsymbol{I} = \lambda(z, \bar{z})(dx^2 + dy^2) = \lambda(z, \bar{z})dz d\bar{z}.$$

Since

$$\boldsymbol{I} = (\boldsymbol{r}_z, \boldsymbol{r}_z)dz^2 + (\boldsymbol{r}_{\bar{z}}, \boldsymbol{r}_{\bar{z}})d\bar{z}^2 + 2(\boldsymbol{r}_z, \boldsymbol{r}_{\bar{z}})dz d\bar{z},$$

the condition for conformality of the parameter $z$ is equivalent to the equations

$$(\boldsymbol{r}_z, \boldsymbol{r}_z) = (\boldsymbol{r}_{\bar{z}}, \boldsymbol{r}_{\bar{z}}) = 0, \quad (\boldsymbol{r}_z, \boldsymbol{r}_{\bar{z}}) = \frac{\lambda(z, \bar{z})}{2}, \tag{6.6}$$

where the subscripts $z$ and $\bar{z}$ denote the derivatives with respect to the corresponding variables and $(\cdot, \cdot)$ is the usual inner product in $\mathbb{R}^3$ extended by linearity to $\mathbb{C}^3$:

$$(V, W) = \sum_{j=1}^{3} V^j W^j, \quad V, W \in \mathbb{C}^3.$$

Since the surface is regular, at each point, the vectors $\boldsymbol{r}_z$, $\boldsymbol{r}_{\bar{z}}$, and $\boldsymbol{n}$, where $\boldsymbol{n}$ is the normal to the surface, constitute a basis for $\mathbb{C}^3$. Indeed, it follows from regularity that $\boldsymbol{r}_x$, $\boldsymbol{r}_y$, and $\boldsymbol{n}$ constitute a basis for $\mathbb{R}^3$; moreover, $\boldsymbol{r}_z = (\boldsymbol{r}_x - i\boldsymbol{r}_y)/2$ and $\boldsymbol{r}_{\bar{z}} = (\boldsymbol{r}_x + i\boldsymbol{r}_y)/2$; i.e., $\boldsymbol{r}_z$ and $\boldsymbol{r}_{\bar{z}}$ are expressed linearly in terms of $\boldsymbol{r}_x$ and $\boldsymbol{r}_y$ over the field of complex numbers.

Write down the derivational equations (see §2.4) in terms of the conformal parameter. These equations define the deformation of the frame $\boldsymbol{r}_z, \boldsymbol{r}_{\bar z}, \boldsymbol{n}$ and have the form

$$\frac{\partial}{\partial z}\begin{pmatrix}\boldsymbol{r}_z\\ \boldsymbol{r}_{\bar z}\\ \boldsymbol{n}\end{pmatrix} = \boldsymbol{U}\begin{pmatrix}\boldsymbol{r}_z\\ \boldsymbol{r}_{\bar z}\\ \boldsymbol{n}\end{pmatrix}, \quad \frac{\partial}{\partial \bar z}\begin{pmatrix}\boldsymbol{r}_z\\ \boldsymbol{r}_{\bar z}\\ \boldsymbol{n}\end{pmatrix} = \boldsymbol{V}\begin{pmatrix}\boldsymbol{r}_z\\ \boldsymbol{r}_{\bar z}\\ \boldsymbol{n}\end{pmatrix}.$$

Find the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$. First of all introduce two additional functions which play the role of the Christoffel symbols:

$$A = (\boldsymbol{r}_{zz}, \boldsymbol{n}), \quad B = (\boldsymbol{r}_{z\bar z}, \boldsymbol{n}).$$

Note that the function $A$ is complex and the function $B$ is real.

**Lemma 6.5.** *The matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ have the form*

$$\boldsymbol{U} = \begin{pmatrix}(\ln\lambda)_z & 0 & A\\ 0 & 0 & B\\ -2B/\lambda & -2A/\lambda & 0\end{pmatrix},$$

$$\boldsymbol{V} = \begin{pmatrix}0 & 0 & B\\ 0 & (\ln\lambda)_{\bar z} & \bar A\\ -2\bar A/\lambda & -2B/\lambda & 0\end{pmatrix}.$$

*Proof.* We will only find the form of the matrix $\boldsymbol{U}$, since the derivation for the other matrix is similar.

By definition, $\boldsymbol{r}_{zz} = \boldsymbol{U}_{11}\boldsymbol{r}_z + \boldsymbol{U}_{12}\boldsymbol{r}_{\bar z} + \boldsymbol{U}_{13}\boldsymbol{n}$. Note that

$$\boldsymbol{U}_{13} = (\boldsymbol{r}_{zz}, \boldsymbol{n}) = A$$

and from (6.6) we obtain the following equalities:

$$\boldsymbol{U}_{12}\frac{\lambda}{2} = (\boldsymbol{r}_{zz}, \boldsymbol{r}_z) = \frac{1}{2}\frac{\partial}{\partial z}\langle \boldsymbol{r}_z, \boldsymbol{r}_z) = 0,$$
$$\boldsymbol{U}_{11}\frac{\lambda}{2} = (\boldsymbol{r}_{zz}, \boldsymbol{r}_{\bar z}) = \frac{\partial}{\partial z}(\boldsymbol{r}_z, \boldsymbol{r}_{\bar z}) - (\boldsymbol{r}_z, \boldsymbol{r}_{z\bar z}) = \frac{1}{2}\frac{\partial\lambda}{\partial z} - \frac{1}{2}\frac{\partial}{\partial \bar z}(\boldsymbol{r}_z, \boldsymbol{r}_z) = \frac{1}{2}\frac{\partial\lambda}{\partial z},$$

which imply $\boldsymbol{U}_{11} = (\ln\lambda)_z$ and $\boldsymbol{U}_{12} = 0$.

Similarly, $\boldsymbol{r}_{z\bar z} = \boldsymbol{U}_{21}\boldsymbol{r}_z + \boldsymbol{U}_{22}\boldsymbol{r}_{\bar z} + \boldsymbol{U}_{23}\boldsymbol{n}$,

$$\boldsymbol{U}_{23} = (\boldsymbol{r}_{z\bar z}, \boldsymbol{n}) = B,$$

and it follows from (6.6) that

$$\boldsymbol{U}_{21}\frac{\lambda}{2} = (\boldsymbol{r}_{z\bar z}, \boldsymbol{r}_{\bar z}) = \frac{1}{2}\frac{\partial}{\partial z}(\boldsymbol{r}_{\bar z}, \boldsymbol{r}_{\bar z}) = 0,$$
$$\boldsymbol{U}_{22}\frac{\lambda}{2} = (\boldsymbol{r}_{z\bar z}, \boldsymbol{r}_z) = \frac{1}{2}\frac{\partial}{\partial \bar z}(\boldsymbol{r}_z, \boldsymbol{r}_z) = 0.$$

Consequently, $\boldsymbol{U}_{21} = \boldsymbol{U}_{22} = 0$.

By definition, we have $\boldsymbol{n}_z = \boldsymbol{U}_{31}\boldsymbol{r}_z + \boldsymbol{U}_{32}\boldsymbol{r}_{\bar{z}} + \boldsymbol{U}_{33}\boldsymbol{n}$. Since $(\boldsymbol{n},\boldsymbol{n}) \equiv 1$, we have

$$\boldsymbol{U}_{33} = (\boldsymbol{n}_z, \boldsymbol{n}) = \frac{1}{2}\frac{\partial}{\partial z}(\boldsymbol{n},\boldsymbol{n}) = 0.$$

Since, by definition, $(\boldsymbol{n}, \boldsymbol{r}_z) = (\boldsymbol{n}, \boldsymbol{r}_{\bar{z}}) = 0$; therefore,

$$\boldsymbol{U}_{31}\frac{\lambda}{2} = (\boldsymbol{n}_z, \boldsymbol{r}_{\bar{z}}) = \frac{\partial}{\partial z}(\boldsymbol{n}, \boldsymbol{r}_{\bar{z}}) - (\boldsymbol{n}, \boldsymbol{r}_{z\bar{z}}) = -B,$$
$$\boldsymbol{U}_{32}\frac{\lambda}{2} = (\boldsymbol{n}_z, \boldsymbol{r}_z) = \frac{\partial}{\partial z}(\boldsymbol{n}, \boldsymbol{r}_z) - (\boldsymbol{n}, \boldsymbol{r}_{zz}) = -A,$$

which implies $\boldsymbol{U}_{31} = -2B/\lambda$ and $\boldsymbol{U}_{32} = -2A/\lambda$.

Lemma 6.5 is proven. □

Recall that the second fundamental form is equal to

$$\boldsymbol{II} = (\boldsymbol{r}_{xx}, \boldsymbol{n})dx^2 + 2(\boldsymbol{r}_{xy}, \boldsymbol{n})dxdy + (\boldsymbol{r}_{yy}, \boldsymbol{n})dy^2;$$

the roots of the equation

$$F(k) = \det\begin{pmatrix} (\boldsymbol{r}_{xx}, \boldsymbol{n}) - \lambda k & (\boldsymbol{r}_{xy}, \boldsymbol{n}) \\ (\boldsymbol{r}_{xy}, \boldsymbol{n}) & (\boldsymbol{r}_{yy}, \boldsymbol{n}) - \lambda k \end{pmatrix} = 0$$

are the principal curvatures $k_1$ and $k_2$; their product is called the Gaussian curvature $K = k_1 k_2$, and the half-sum is called the mean curvature $H = (k_1 + k_2)/2$.

Making straightforward calculations, we obtain

**Lemma 6.6.**

$$\boldsymbol{II} = (2B + (A + \bar{A}))dx^2 + 2i(A - \bar{A})dxdy + (2B - (A + \bar{A}))dy^2,$$
$$H = \frac{2B}{\lambda}, \quad K = \frac{4(B^2 - A\bar{A})}{\lambda^2}.$$

The Gauss–Codazzi equations represent the compatibility condition for the system of derivational equations:

$$\boldsymbol{U}_{\bar{z}} - \boldsymbol{V}_z + [\boldsymbol{U}, \boldsymbol{V}] = 0.$$

Making straightforward calculations, we obtain

**Lemma 6.7.** *The Gauss–Codazzi equations take the form*

$$(\ln\lambda)_{z\bar{z}} + \frac{2}{\lambda}(B^2 - A\bar{A}) = 0, \tag{6.7}$$

$$A_{\bar{z}} - B_z + (\ln\lambda)_z B = 0. \tag{6.8}$$

By the formula for the Gaussian curvature of Lemma 6.6, the first of these equations is

$$K = -2(\ln \lambda)_{z\bar{z}}$$

(the Gauss theorem). The second equation splits into two equations, since it claims that the real and imaginary parts of its left-hand side, which is complex-valued in general, are zero. Using Lemma 6.6, we can rewrite them in the form

$$A_{\bar{z}} = \frac{\lambda}{2} H_z.$$

The form $\Omega = A dz^2$ (which is not antisymmetric) is called the *Hopf differential*. It is a *quadratic differential*. This means that if another conformal parameter $w$ is given on the surface, then the Hopf differential in terms of this new parameter has the form $\Omega = \widetilde{A} dw^2$, where

$$\tilde{A} = (\boldsymbol{r}_{ww}, \boldsymbol{n}) = \left(\frac{\partial z}{\partial w}\right)^2 (\boldsymbol{r}_{zz}, \boldsymbol{n}) = \left(\frac{\partial z}{\partial w}\right)^2 A.$$

It possesses a series of remarkable properties whose proofs are left as an exercise (they follow from Lemma 6.7):

**Lemma 6.8.** (1) *The Hopf differential at $p \in \Sigma$ is zero if and only if $p$ is an umbilic point* (*i.e., the principal curvatures at this point coincide: $k_1 = k_2$*).

(2) *The Hopf differential is holomorphic* (*i.e., $A_{\bar{z}} = 0$*) *on a surface if and only if the surface has constant mean curvature* (*$H \equiv$ const*).

Now, the Bonnet theorem which claims that every surface is determined by the first and the second fundamental forms (Theorem 2.6) takes the following form:

- *a surface is determined uniquely* (*up to motions of* $\mathbb{R}^3$) *by the metric $\lambda\, dz d\bar{z}$, the mean curvature $H = 2B/\lambda$, and the Hopf differential $A dz^2$, provided that they satisfy equations* (6.7) *and* (6.8);
- *every collection $\{\lambda\, dz d\bar{z}, H, A dz^2\}$ satisfying equations* (6.7) *and* (6.8) *determines a surface*.

Note that there exist surfaces which can be smoothly deformed with the metric and the mean curvature preserved, and these deformations do not reduce to motions of surfaces in the Euclidean space.

Observe (without proof) that the spaces of holomorphic quadratic differentials on closed surfaces are finite-dimensional.

**Lemma 6.9.** *If $\Omega$ is a holomorphic quadratic differential on the two-dimensional sphere $S^2$, then $\Omega = 0$ everywhere.*

*Proof.* The two-dimensional sphere is covered by two charts each of which is identified with the complex plane. The complex coordinates $z$ and $w$ in these charts are connected by the transition formula

$$w = \frac{1}{z}.$$

We see that the sphere is obtained from the $z$-plane by addition of a point at infinity with the coordinate $w = 0$. If $\alpha(z)dz^2$ and $\beta(w)dw^2$ are the expressions of the differential $\Omega$ in these charts, then in the domain $z \neq 0$ in which both coordinate systems are defined, we obtain

$$\beta(w)dw^2 = \beta(w(z))\left(\frac{dw}{dz}\right)^2 dz^2 = \beta(w(z))z^{-4}dz^2 = \alpha(z)dz^2.$$

Since the value of the function $\beta(w)$ at $w = 0$ is finite, it follows that the holomorphic function $\alpha$ on the $z$-plane has the asymptotic expansion

$$\alpha(z) = O\left(\frac{1}{z^4}\right) \quad \text{as } z \to \infty.$$

But a holomorphic function decreasing at infinity is equal to zero everywhere.

Lemma 6.9 is proven. □

From this lemma we obtain

**Theorem 6.2** (The Hopf theorem). *If a sphere immersed in* $\mathbb{R}^3$ *has a constant mean curvature* $H$, *then it coincides* (*up to parallel translation*) *with the standard sphere of radius* $1/H$.

*Proof.* Assume that the mean curvature is constant: $H = \text{const}$. It follows from Lemmas 6.8 and 6.9 that $A \equiv 0$. Lemma 6.5 implies that $\boldsymbol{n}_z = H\boldsymbol{r}_z$ which yields

$$\boldsymbol{n} = H\boldsymbol{r} + \boldsymbol{c},$$

where $\boldsymbol{c}$ is a constant vector. If $H = 0$, then the normal vector is constant and the surface cannot be a sphere (it lies completely in the plane orthogonal to this vector). If $H \neq 0$, then we translate the surface by $\boldsymbol{c}/H$: $\boldsymbol{r} \to \tilde{\boldsymbol{r}} = \boldsymbol{r} + \boldsymbol{c}/H$. For the translated surface we obtain: $(H\tilde{\boldsymbol{r}}, H\tilde{\boldsymbol{r}}) = (\boldsymbol{n}, \boldsymbol{n}) = 1$ and eventually $(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{r}}) = 1/H^2$.

Theorem 6.2 is proven. □

## 6.3 The Weierstrass representation

From the derivational equations (Lemma 6.5) we find that the following equation holds for an arbitrary surface:

$$\boldsymbol{r}_{z\bar{z}} = B\boldsymbol{n},$$

which together with Lemma 6.6 gives

$$\boldsymbol{r}_{z\bar{z}} = \frac{1}{2} H\lambda \boldsymbol{n}. \tag{6.9}$$

In terms of the conformal parameter the Laplace–Beltrami operator becomes

$$\Delta = \frac{4}{\lambda} \frac{\partial^2}{\partial z \partial \bar{z}}$$

and therefore (6.9) takes the form

$$\Delta \boldsymbol{r} = 2H\boldsymbol{n}. \tag{6.10}$$

From (6.10) we obtain

**Lemma 6.10.** *A surface is minimal* ($H = 0$) *if and only if its coordinate functions which describe its immersion into* $\mathbb{R}^3$ *are harmonic.*

It follows from (6.9) that the complex-valued vector function $\boldsymbol{r}_z$ for minimal surfaces is holomorphic. By condition (6.6), it also satisfies the equation $(\boldsymbol{r}_z, \boldsymbol{r}_z) = 0$. It turns out that the formal validity of these conditions for a vector function $\varphi$ is sufficient for the equality $\varphi = \boldsymbol{r}_z$ to hold for some conformally parameterized minimal surface.

**Lemma 6.11.** *Let* $U \subset \mathbb{C}$ *be a simply connected domain and let* $\varphi \colon U \to \mathbb{C}^3$ *be a vector function such that*

(1) $\varphi$ *is holomorphic* (*i.e., each component* $\varphi_j$ *is holomorphic*);

(2) $\varphi_1^2 + \varphi_2^2 + \varphi_3^2 = 0$;

(3) $\varphi$ *has no zeros:* $\varphi(U) \subset \mathbb{C}^3 \setminus \{0\}$.

*Then there is a regular minimal surface* $\boldsymbol{r} : U \to \mathbb{R}^3$ *such that* $\varphi = \boldsymbol{r}_z$ *and* $z$ *is the conformal parameter on it.*

*Proof.* Suppose that $\varphi_j = a_j + ib_j$, where $a_j$ and $b_j$ are real functions. The fact that $\varphi_j$ is holomorphic means that the Cauchy–Riemann conditions are satisfied:

$$\frac{\partial}{\partial \bar{z}}(a_j + ib_j) = \frac{1}{2}\left(\left(\frac{\partial a_j}{\partial x} - \frac{\partial b_j}{\partial y}\right) + i\left(\frac{\partial a_j}{\partial y} + \frac{\partial b_j}{\partial x}\right)\right) = 0.$$

In particular, the fact that the imaginary part $\frac{\partial \varphi_j}{\partial \bar{z}}$ is zero means that the 1-form $\omega_j = a_j\,dx - b_j\,dy$ is closed: $d\omega_j = 0$. Take an arbitrary point $z_0 \in U$ and define a smooth real-valued function

$$f_j(z) = 2\int_{z_0}^{z} \omega_j$$

on $U$. Since $U$ is a simply connected domain, by the Stokes theorem, the value of the integral is independent of the choice of the curve in $U$ joining $z_0$ with $z$. Straightforward calculations demonstrate that

$$\frac{\partial f_j}{\partial z} = \varphi_j.$$

Consequently, the functions $f_1$, $f_2$, and $f_3$ determine a map

$$\boldsymbol{r}\colon U \to \mathbb{R}^3, \quad \boldsymbol{r} = (f_1, f_2, f_3),$$

such that $\boldsymbol{r}_z = \varphi$. Since $\varphi(U) \subset \mathbb{C}^3 \setminus \{0\}$, this map has full rank at each point, i.e., is an immersion.

It follows from $(\varphi, \varphi) = 0$ that $z$ is a conformal parameter on the immersed surface. Since $\boldsymbol{r}_{z\bar{z}} = 0$, this surface is minimal by Lemma 6.10.

Lemma 6.11 is proven. □

From the above lemma we obtain the following theorem:

**Theorem 6.3** (The Weierstrass theorem). *Let $U \subset \mathbb{C}$ be a simply connected domain and let $\psi_1, \psi_2\colon U \to \mathbb{C}$ be holomorphic functions such that $|\psi_1|^2 + |\psi_2|^2$ is nonzero everywhere in $U$. Let $(x_0^1, x_0^2, x_0^3) \in \mathbb{R}^3$ and $p \in U$ be given points.*

*Then the formulas*[1]

$$\begin{aligned} x^1(z) &= x_0^1 + \frac{1}{2}\int_{z_0}^{z} ((\psi_1^2 - \psi_2^2)dz + (\bar{\psi}_1^2 - \bar{\psi}_2^2)d\bar{z}), \\ x^2(z) &= x_0^2 - \frac{i}{2}\int_{z_0}^{z} ((\psi_1^2 + \psi_2^2)dz - (\bar{\psi}_1^2 + \bar{\psi}_2^2)d\bar{z}), \\ x^3(z) &= x_0^3 + \int_{z_0}^{z} (\psi_1\psi_2 dz + \bar{\psi}_1\bar{\psi}_2 d\bar{z}) \end{aligned} \tag{6.11}$$

*determine an immersion of the minimal surface into $\mathbb{R}^3$. The induced metric has the form*

$$(|\psi_1|^2 + |\psi_2|^2)^2 dz d\bar{z}.$$

*Every minimal surface in $\mathbb{R}^3$ in a sufficiently small neighborhood of every point admits such representation.*

*Proof.* The general solution to the equation

$$\varphi_1^2 + \varphi_2^2 + \varphi_3^2 = 0$$

is representable as

$$\varphi_1 = \frac{1}{2}(\psi_1^2 - \psi_2^2), \quad \varphi_2 = -\frac{i}{2}(\psi_1^2 + \psi_2^2), \quad \varphi_3 = \psi_1\psi_2; \tag{6.12}$$

moreover, to every nonzero solution $\varphi$ there correspond two pairs $(\psi_1, \psi_2)$ and $(-\psi_1, -\psi_2)$. The vector function $\varphi$ is holomorphic, since such are $\psi_1$ and $\psi_2$. It follows from here and Lemma 6.11 that formulas (6.11) determine a conformal immersion of $U$ into $\mathbb{R}^3$ and the constructed surface is minimal. The form of the induced metric $\lambda dz d\bar{z}$ is established by elementary calculations: $2(\boldsymbol{r}_z, \boldsymbol{r}_{\bar{z}}) = \lambda$.

[1] Usually, the Weierstrass–Enneper formulas are formulated in terms of $f = \psi_1^2$ and $g = \psi_2/\psi_1$.

It follows from (6.12) that, in a neighborhood of every point of a minimal surface, $\boldsymbol{r}_z$ admits such parameterization and hence the surface is representable in the form (6.11).

Theorem 6.3 is proven. □

Before giving examples of minimal surfaces, we observe the following fact.

**Theorem 6.4.** *In $\mathbb{R}^3$ there are no minimal compact surfaces without boundary.*

*Proof.* By the maximum principle, a nonconstant harmonic function $f$ (i.e., such that $f_{z\bar{z}} = 0$) defined in some domain of $\mathbb{C}$ cannot take its maximal value at an interior point of the domain. Suppose that $\Sigma$ is a minimal compact surface without boundary immersed into $\mathbb{R}^3$. Since it is compact, each coordinate function $x^j$ attains its maximum on this surface. Suppose, for definiteness, that $x^1$ attains a maximum at a point $p$. In a neighborhood of $p$, there is a conformal parameter $z$ such that $x^1_{z\bar{z}} = 0$ (by Lemma 6.10); i.e., a harmonic function attains its maximum at an interior point ($\Sigma$ has no boundary). This implies that $x^1$ is constant on the surface. Repeating this argument for $x^2$ and $x^3$, we prove that all coordinate functions must be constant. The so-obtained contradiction proves Theorem 6.4. □

**Examples of minimal surfaces.** (1) *Plane*: $\psi_1 = 1, \psi_2 = 0$.

(2) *Catenoid* (Figure 6.1):

$$\psi_1 = \frac{\sqrt{a}}{\sqrt{2}}, \quad \psi_2 = \frac{\sqrt{a}}{\sqrt{2}z}, \qquad a \in \mathbb{R},\ a > 0.$$

This is a minimal surface of revolution. Here $z \in \mathbb{C} \setminus \{0\}$ and the domain of variation of $z$ (the plane with a deleted point) is not homeomorphic to a disk. However,

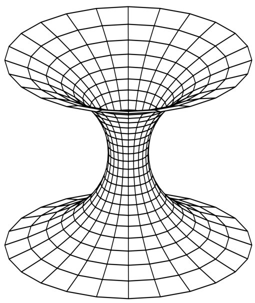

Figure 6.1. Catenoid.

in Theorem 3 we imposed this requirement on the domain to guarantee that the integrals in (6.11) are determined uniquely. In the case of the catenoid these integrals are determined uniquely and the catenoid is an embedded (i.e., without self-intersections) minimal cylinder. We already pointed out this surface in Problem 2.14.

The explicit formulas for the catenoid are

$$x^1 = a \cosh u \cos v, \quad x^2 = a \cosh u \sin v, \quad x^3 = au,$$

where $z = e^{u+iv}$. From these formulas we see that the catenoid is described by the equation

$$(x^1)^2 + (x^2)^2 = \cosh^2\left(\frac{x^3}{a}\right).$$

(3) *Helicoid* (Figure 6.2):

$$\psi_1 = \frac{\sqrt{a}\sqrt{i}}{\sqrt{2}}, \quad \psi_2 = \frac{\sqrt{a}\sqrt{i}}{\sqrt{2}z}, \quad a \in \mathbb{R},\ a > 0.$$

This is a line surface obtained by uniform revolution of a straight line $l$ intersecting the axis of revolution and orthogonal to it and simultaneous uniform translation of $l$ parallel to the axis of revolution; moreover, the ratio of the speed of translation and the speed of revolution equals $a$.

Here again $z \in \mathbb{C} \setminus \{0\}$; however, in this case the integrals in (6.11) are not determined uniquely. To retain univalence, delete the negative real half-axis $\mathbb{R}_- = \{z = x + iy \mid x \leq 0, y = 0\}$ from $\mathbb{C}$. Eventually, we obtain one turn of the helicoid corresponding to one revolution of $l$ around the axis of revolution. Now, we have to extend analytically the embedding, by gluing the upper coast of one copy of $\mathbb{C} \setminus \mathbb{R}_-$ with the lower coast of another such copy, and so on.

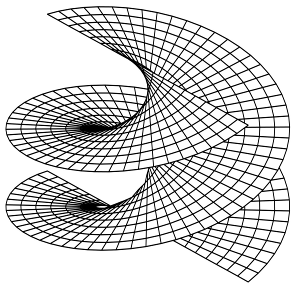

Figure 6.2. Helicoid.

The explicit formulas for the helicoid are

$$x^1 = a \sinh v \sin u, \quad x^2 = a \sinh v \cos u, \quad x^3 = au,$$

where $z = e^{-i(u+iv)} = e^{v-iu}$. Hence, the helicoid is given by the equation

$$\tan \frac{x^3}{a} = \frac{x^1}{x^2}.$$

The catenoid and helicoid corresponding to the same value of the parameter $a$ are locally isometric: the functions $\psi$ which determine them differ by the factor $\sqrt{i}$; however, globally they are different: topologically, the helicoid is an embedded plane, while the catenoid is a cylinder.

(4) *The Enneper surface*: $\psi_1 = 1, \psi_2 = z$. This surface is not embedded but, like the previous surfaces, is complete in the sense that every geodesic on it extends indefinitely.

**Problem 6.1.** Prove that the graph of the function $z = f(x, y)$ (in the three-dimensional Euclidean space with coordinates $x$, $y$, and $z$) is a minimal surface if and only if

$$(1 + f_y^2) f_{xx} - 2 f_x f_y f_{xy} + (1 + f_x^2) f_{yy} = 0$$

(*minimal surface equation*).

# 7

# Elements of Lie group theory

## 7.1 Linear Lie groups

A smooth manifold $G$ whose points constitute a group and all group operations are smooth is called a *Lie group*. This means that:

(1) a smooth map $G \times G \to G$, called *multiplication*, is given which relates to a pair of points $x, y \in G$ their product $xy$;

(2) $(xy)z = x(yz)$ for all $x, y, z \in G$;

(3) there is an *identity element* $e \in G$ such that $xe = ex = x$ for all $x \in G$;

(4) a smooth map $G \to G$, $x \mapsto x^{-1}$, called *inversion*, is given which relates to each element $x \in G$ the inverse element $x^{-1} \in G$ such that their product is equal to the identity $e$:

$$xx^{-1} = x^{-1}x = e.$$

It is easy to show that the element $x^{-1}$ meeting the latter equalities is unique.

A subgroup $H$ of a Lie group $G$ is a *Lie subgroup* if it is simultaneously a smooth submanifold of $G$.

We have already indicated one example of a Lie group in §1.4, the real orthogonal group $\mathrm{O}(n)$ which is an $\frac{(n-1)n}{2}$-dimensional submanifold in the space $\mathbb{R}^{n^2}$ realized as the space of real $(n \times n)$-matrices. For the beginning, we will focus on other examples of Lie groups realized by matrices.

(1) *Complete linear group* $\mathrm{GL}(n)$. Let $\mathbb{R}^n$ be the $n$-dimensional vector space over the field of reals. Choosing a basis $e_1, \dots, e_n$ for $\mathbb{R}^n$, we relate each linear map $A\colon \mathbb{R}^n \to \mathbb{R}^n$ with a unique $(n \times n)$-matrix $A = (a_{jk})$ which determines the map in this basis:

$$A\colon \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \dots & \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}.$$

Although we know that a linear map is determined by a tensor of type $(1,1)$ (see §6.3), in this chapter we enumerate the entries of matrices by two subscripts like in most linear algebra courses. Thereby the space $M(n, \mathbb{R})$ constituted by $(n \times n)$-matrices is identified with the Euclidean space $\mathbb{R}^{n^2}$ in which the coordinates are the entries $a_{jk}$ of the matrices. Linear maps $\mathbb{R}^n \to \mathbb{R}^n$ are also called *transformations* of the (vector) space $\mathbb{R}^n$.

We denote by $1_n$ the identity transformation of the space $\mathbb{R}^n$ and the identity $(n \times n)$-matrix which determines this transformation.

The *complete linear group* $\mathrm{GL}(n) = \mathrm{GL}(n, \mathbb{R})$ is the group constituted by all invertible linear transformations of the space $\mathbb{R}^n$.

**Theorem 7.1.** *The group* $\mathrm{GL}(n)$ *is a noncompact Lie group of dimension* $n^2$.

*Proof.* The smooth function (determinant)

$$\det\colon M(n, \mathbb{R}) \to \mathbb{R},$$

is defined on the space $M(n, \mathbb{R})$; this function vanishes only on the matrices which determine noninvertible maps. Consequently, each element $A \in \mathrm{GL}(n)$ is contained in $\mathrm{GL}(n)$ together with its neighborhood; therefore, $\mathrm{GL}(n)$ is an open set and consequently an $n^2$-dimensional submanifold in $M(n, \mathbb{R}) = \mathbb{R}^{n^2}$. Since the function det is unbounded, the manifold $\mathrm{GL}(n)$ is noncompact (otherwise the function would attain its maximum and minimum). Obviously, the operations of multiplication and inversion of matrices are smooth as functions of the entries. Consequently, $\mathrm{GL}(n)$ is a Lie group. Theorem 7.1 is proven. □

From Theorem 7.1 we immediately derive the following fact about the orthogonal group $\mathrm{O}(n)$ introduced in §1.4.

**Corollary 7.1.** $\mathrm{O}(n)$ *is a compact Lie subgroup of the group* $\mathrm{GL}(n)$.

Here and in the sequel we identify elements of $\mathrm{O}(n)$ with the matrices which determine them in some chosen orthonormal basis.

Corollary 7.1 is plain. The transformations in $\mathrm{O}(n)$ are those which preserve the inner product

$$(u, v) = \sum_{j=1}^{n} u^j v^j.$$

This is equivalent to the fact that

$$A^\top A = 1_n, \quad A \in O(n), \tag{7.1}$$

where $A^\top$ is the transpose of $A$ (see §1.4). Consequently, all transformations in $\mathrm{O}(n)$ are invertible and the embedding $\mathrm{O}(n) \subset \mathrm{GL}(n)$ is a smooth map. Hence, $\mathrm{O}(n)$ is a Lie subgroup of the group $\mathrm{GL}(n)$.

The assertion about compactness is also plain. It follows from (7.1) that

$$\sum_{k=1}^{n} a_{jk}^2 = 1, \quad j = 1, \dots, n,$$

for each $j$th row of the matrix $A \in \mathrm{O}(n)$. Therefore, all entries $a_{jk}$ of the matrix $A$ are bounded in magnitude by 1: $|a_{jk}| \le 1$. Since the set of solutions to equation (7.1) is closed in $\mathbb{R}^{n^2} = M(n, \mathbb{R})$, the submanifold $\mathrm{O}(n)$ is a closed and bounded subset in $\mathbb{R}^{n^2}$; consequently, it is compact. Corollary 7.1 is proven.

Subgroups of the complete linear groups $\mathrm{GL}(n)$ are called *linear* (or *matrix*) *groups*.

The following lemma has been already applied in §1.4 in the case $H = \mathrm{O}(n)$:

**Lemma 7.1.** *Let* $H \subset \mathrm{GL}(n)$ *be a subgroup which is also a smooth submanifold in a neighborhood of the identity* $1_n \in \mathrm{GL}(n)$. *Then* $H$ *is a smooth submanifold and consequently a Lie subgroup of* $\mathrm{GL}(n)$.

*Proof.* The fact that the subgroup $H$ is a submanifold in a neighborhood $U$ of the identity matrix $1_n$ means that in this neighborhood $U \cap H$ is given as the zero set of a smooth map $F\colon U \to \mathbb{R}^k$ whose Jacobian matrix has full rank (see Theorem 1.5). Take $A \in H$. This element has a neighborhood $AU$ constituted by the points of the form $AB$, where $B \in U$. Since $H$ is a subgroup of $\mathrm{GL}(n)$, the elements of $H$ in $AU$ are determined by the equation $F_A(X) = 0$, where $F_A(X) = F(A^{-1}X)$. The Jacobian matrix of the map $F_A\colon AU \to \mathbb{R}^k$ at $A$ has the same rank as the Jacobian matrix of the map $F$ at $1_n$. Consequently this rank is maximal, and in a neighborhood of every point $A \in H$ the subgroup $H$ is a smooth submanifold. Lemma 7.1 is proven. □

(2) *Unimodular* (*or special linear*) *group* $\mathrm{SL}(n)$. Denote by $\mathrm{SL}(n)$ the level surface $\det = 1$ of the determinant, i.e., the set of points at which $\det = 1$.

**Theorem 7.2.** $\mathrm{SL}(n)$ *is a Lie subgroup of* $\mathrm{GL}(n)$. *It is noncompact and has dimension* $n^2 - 1$.

*Proof.* It is easy to find that

$$\frac{\partial \det A}{\partial a_{11}}\bigg|_{A=1_n} = 1.$$

Consequently, in a neighborhood of $1_n$ the group $\mathrm{SL}(n)$ is a smooth submanifold. It follows from Lemma 7.1 that $\mathrm{SL}(n)$ is a Lie subgroup of the group $\mathrm{GL}(n)$. By the implicit function theorem, $\dim \mathrm{SL}(n) = \dim \mathrm{GL}(n) - 1$. Since the entries are unbounded on $\mathrm{SL}(n)$, this group is noncompact. Theorem 7.2 is proven. □

(3) *Special orthogonal group* $\mathrm{SO}(n)$. Take $A \in \mathrm{O}(n)$. It follows from the equalities $\det A \det A^{-1} = 1$ and $\det A = \det A^\top$ that $\det A = \pm 1$ for $A \in \mathrm{O}(n)$.

The group $\mathrm{SO}(n)$ is constituted by the matrices in $\mathrm{O}(n)$ with $\det A = 1$.

**Theorem 7.3.** $\mathrm{SO}(n)$ *is a connected component of the Lie group* $\mathrm{O}(n)$.

*Proof.* First of all prove two lemmas.

**Lemma 7.2.** *The group* $\mathrm{O}(2)$ *has two components constituted by the matrices of the form*

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \quad \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}, \tag{7.2}$$

*where* $\varphi \in \mathbb{R}$. *Each component is diffeomorphic to the circle* $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ *and the first of them coincides with the group* SO(2).

*Proof of Lemma* 7.2. Take $A \in O(2)$. Denote

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Equation (7.1) takes the form

$$a^2 + b^2 = c^2 + d^2 = 1, \quad ac + bd = 0.$$

The general solution to the equation $a^2+b^2 = 1$ is representable in the form $a = \cos\varphi$, $b = \sin\varphi$. Hence, $c = \pm\sin\varphi$ and $d = \mp\cos\varphi$. Consequently, all entries of O(2) have the form (7.2). The determinant of the matrices of the first form equals $\det A = 1$, while the determinant of the matrices of the second form equals $\det A = -1$. Lemma 7.2 is proven. □

**Lemma 7.3.** *Let* $A \in O(n)$. *Then the vector space* $\mathbb{R}^n$ *is representable as the direct sum of one- and two-dimensional subspaces* $V_j$ *such that they are pairwise orthogonal; i.e.,* $V_i \perp V_j$ *for* $i \neq j$, *and each of them is invariant under* $A$; *i.e.,* $A(V_j) = V_j$ *for all* $j$.

*Proof of Lemma* 7.3. We proceed by induction on $n$. For $n = 1$ the assertion is trivial. Assume that the lemma is proven for $n < k$. Take $A \in O(k)$. Write the characteristic equation

$$\det(A - \lambda 1_k) = 0.$$

If it has a real root $\lambda$, then there is a vector $v \in \mathbb{R}^k$ such that $Av = \lambda v$. In this case the space $\mathbb{R}^k$ splits into the direct sum $\mathbb{R}^k = \mathbb{R}\cdot v + (\mathbb{R}\cdot v)^\perp$ of the one-dimensional subspace $\mathbb{R}\cdot v$ spanned by $v$ and its orthogonal complement. Both subspaces are invariant under $A$. Consequently, the restriction of $A$ to $(\mathbb{R}\cdot v)^\perp$ is given by a matrix $A_1 \in O(k-1)$. By the induction assumption, the assertion of the lemma is valid for $A_1$. Therefore, it is also valid for $A$.

If the equation $\det(A - \lambda 1_k) = 0$ has no real roots, then take a complex root $\mu$ and a complex vector $w = w_1 + iw_2 \in \mathbb{C}^k$, with $w_1, w_2 \in \mathbb{R}^k$, such that $Aw = \mu w$. It is obvious that $A(w_1 - iw_2) = \overline{\mu}(w_1 - iw_2)$ and the vectors $w_1$ and $w_2$ generate a two-dimensional invariant subspace $\mathbb{R}\cdot w_1 + \mathbb{R}\cdot w_2$. Since the transformation $A$ is orthogonal, the orthogonal complement to $\mathbb{R}\cdot w_1 + \mathbb{R}\cdot w_2$ is also invariant. Now, arguing as in the previous case, we derive the assertion of the lemma for $A$ from the induction assumption. Lemma 7.3 is proven. □

We turn to proving Theorem 7.3. To show that SO($n$) is connected, it suffices to deform every transformation $A \in SO(n)$ within SO($n$) into the identity transformation $1_n$. It follows from Lemma 7.3 that in an appropriate basis every element $A \in O(n)$

is representable as a block diagonal matrix

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix},$$

where the matrices $A_j$ are either one-dimensional (in this case they are equal to the eigenvalues of the orthogonal transformation; i.e., $\pm 1$) or two-dimensional (in this case they have the form (7.2)). In the latter case, changing continuously the angle $\varphi$, we can deform every matrix of the form (7.2) into the diagonal matrix with diagonal entries equal to $\pm 1$. Hence, $A$ is deformed into the element $B$ which is given, in an appropriate basis, by a diagonal matrix with $\pm 1$ on the diagonal. Since $\det A = \det B = 1$, the diagonal contains an even number of entries equal to $-1$. Split them into pairs and for each pair choose the corresponding eigenvectors $v_1$ and $v_2$. The restriction of $B$ to $\mathbb{R} \cdot v_1 + \mathbb{R} \cdot v_2$ is equal to $-1_2$. This matrix has the first form in (7.2) for $\varphi = \pi$ and, deforming $\varphi$ into 0, we take the restriction of $B$ to this subspace into $1_2$. Repeating this procedure for all pairs, we bring the matrix $B$ to the form $1_n$, i.e., deform the linear transformation $A$ into the identity transformation. We have thus shown that $\mathrm{SO}(n)$ is a connected manifold.

Let $A \in \mathrm{O}(n)$ and $\det A = -1$. The image of a smooth and invertible (i.e., diffeomorphism) map $\mathrm{SO}(n) \to \mathrm{O}(n)$ of the form $X \to AX$ is the set constituted by all matrices in $\mathrm{O}(n)$ with $\det = -1$. Hence, the manifold $\mathrm{O}(n)$ splits into two connected components on which the determinant has different signs. Theorem 7.3 is proven. □

We outlined the proof of this theorem in view of the importance of the groups $\mathrm{SO}(n)$ and the fact that its analogs for other groups are simpler but based on the same idea.

**Problem 7.1.** Using the real Jordan form of matrices, show that:

($a$) for every $n$, the group $\mathrm{GL}(n)$ has two connected components selected by the sign of $\det A$;

($b$) for every $n$, the group $\mathrm{SL}(n)$ is connected.

(4) *General linear group* $\mathrm{GL}(n, \mathbb{C})$. Invertible linear transformations of the vector space $\mathbb{C}^n$ over the field $\mathbb{C}$ constitute the group $\mathrm{GL}(n, \mathbb{C})$. Choose a basis $e_1, \dots, e_n$ for $\mathbb{C}^n$ and relate to each linear map $A \colon \mathbb{C}^n \to \mathbb{C}^n$ the $(n \times n)$-matrix $A$ with complex entries which determines the map in this basis. Denote the space of such matrices by $M(n, \mathbb{C})$. The transformations $A \in \mathrm{GL}(n, \mathbb{C})$ are determined by the condition $\det A \neq 0$. Hence, as in the case of the group $\mathrm{GL}(n)$, we conclude that $\mathrm{GL}(n, \mathbb{C})$ is a Lie group of (real) dimension $2n^2 = \dim M(n, \mathbb{C})$.

(5) *Unitary group* $\mathrm{U}(n)$. Let $\mathbb{C}^n$ be an $n$-dimensional vector space over the field $\mathbb{C}$. Denote by $(u, v)$ the *Hermitian product* on $\mathbb{C}^n$ which has the following form in an

orthonormal basis $e_1, \dots, e_n$:

$$(u, v) = \sum_{j=1}^{n} u^j \bar{v}^j .$$

The transformation $A \in M(n, \mathbb{C})$ is *unitary* if it preserves this Hermitian product. Obviously, it has the zero kernel and is invertible.

The space $M(n, \mathbb{C})$ is naturally identified with $\mathbb{C}^{n^2}$ and hence with $\mathbb{R}^{2n^2}$ in which the coordinates are the real and imaginary parts of the entries: $\operatorname{Re} a_{jk}$ and $\operatorname{Im} a_{jk}$. Since each linear transformation in $M(n, \mathbb{C})$ determines a linear transformation $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$, where $\mathbb{R}^{2n}$ is generated over $\mathbb{R}$ by the vectors $e_1, ie_1, \dots, e_n, ie_n$, the group $M(n, \mathbb{C})$ is naturally embedded into $M(2n, \mathbb{R})$.

**Theorem 7.4.** *The group* $\mathrm{U}(n)$*, constituted by all unitary transformations of* $\mathbb{C}^n$*, is determined in* $\mathrm{GL}(n, \mathbb{C})$ *by the equation*

$$A^\top \bar{A} = 1_n,$$

*and is a compact Lie subgroup of* $\mathrm{GL}(2n)$ *of dimension* $n^2$.

*Proof.* The product of two unitary transformations and the inverse of a unitary transformation are unitary. Therefore, unitary transformations constitute a group which we denote by $\mathrm{U}(n)$. The Hermitian product can be written in the form

$$(u, v) = u^\top 1_n \bar{v},$$

where $u = (u_1, \dots, u_n)^\top$ and $v = (v_1, \dots, v_n)^\top$ are vectors in $\mathbb{C}^n$ written as $(n \times 1)$-matrices. A matrix $A \in M(n, \mathbb{C})$ is unitary if and only if

$$(Au, Av) = u^\top A^\top 1_n \bar{A} \bar{v} = (u, v)$$

for all $u, v \in \mathbb{C}^n$. This means that the following complex analog of equation (7.1) holds:

$$A^\top \bar{A} = 1_n, \quad A \in U(n). \tag{7.3}$$

This equation is rewritten as the following system of $n^2$ polynomial equations:

$$F_{jk} = \sum_{m=1}^{n} (\operatorname{Re} a_{mj} \operatorname{Re} a_{mk} + \operatorname{Im} a_{mj} \operatorname{Im} a_{mk}) - \delta_{jk} = 0$$

with $1 \le j \le k \le n$, and

$$G_{jk} = \sum_{m=1}^{n} (\operatorname{Re} a_{mj} \operatorname{Im} a_{mk} - \operatorname{Im} a_{mj} \operatorname{Re} a_{mk}) = 0,$$

with $1 \le j < k \le n$.

The functions $F_{jk}$ and $G_{jk}$ determine the map

$$F\colon M(n,\mathbb{C}) \to \mathbb{R}^{n^2}.$$

Prove that the Jacobian matrix of this map at $1_n \in M(n,\mathbb{C})$ has full rank equal to $n^2$. Choose $n^2$ variables $\operatorname{Re} a_{jk}$, $1 \le j \le k \le n$, and $\operatorname{Im} a_{jk}$, $1 \le j < k \le n$ and calculate the minor of the Jacobian matrix corresponding to these variables for $A = 1_n$:

$$\frac{\partial F_{jk}}{\partial \operatorname{Re} a_{rs}} = \begin{cases} 2 & \text{for } j = r, k = s, j = k, \\ 1 & \text{for } j = r, k = s, j < k, \\ 0 & \text{otherwise,} \end{cases}$$

$$\frac{\partial G_{jk}}{\partial \operatorname{Im} a_{rs}} = \begin{cases} 1 & \text{for } j = r, k = s, \\ 0 & \text{otherwise,} \end{cases}$$

$$\frac{\partial F_{jk}}{\partial \operatorname{Im} a_{rs}} = \frac{\partial G_{jk}}{\partial \operatorname{Re} a_{rs}} = 0 \quad \text{for all } j, k, r, \text{ and } s.$$

Consequently, by Lemma 7.1, $\mathrm{U}(n)$ is a smooth submanifold in $\mathbb{R}^{2n^2}$ of dimension $\dim \mathrm{U}(n) = \dim \mathbb{R}^{2n^2} - n^2 = n^2$.

It follows from (7.3) that $|a_{jk}| \le 1$. Therefore, $\mathrm{U}(n)$ is a bounded subset of $\mathbb{R}^{2n^2}$; since it is closed, it is compact. Theorem 7.4 is proven. □

The chain of embeddings

$$\mathrm{U}(n) \subset M(n,\mathbb{C}) \subset M(2n,\mathbb{R})$$

demonstrates that the group $\mathrm{U}(n)$ is linear.

(6) *Special unitary group* $\mathrm{SU}(n)$. It follows from equation (7.3) that $|\det A| = 1$ for $A \in \mathrm{U}(n)$. The matrices in $\mathrm{U}(n)$ with $\det A = 1$ constitute the subgroup which we denote by $\mathrm{SU}(n)$.

**Theorem 7.5.** $\mathrm{SU}(n)$ *is an* $(n^2-1)$*-dimensional Lie subgroup of the group* $\mathrm{U}(n)$.

*Proof.* In $M(n,\mathbb{C})$ the group $\mathrm{SU}(n)$ is selected by the same equations $F_{jk} = G_{jk} = 0$ as $\mathrm{U}(n)$ and one additional equation

$$\det A = 1,$$

which we replace with

$$H(A) = \operatorname{Im} \det A = 0.$$

This equation determines two components $\det A = \pm 1$ one of which is precisely $\mathrm{SU}(n)$. The Jacobian matrix of the map $\widetilde{F}\colon M(n,\mathbb{C}) \to \mathbb{R}^{n^2+1}$ of the form

$$\widetilde{F}(A) = (F(A), H(A))$$

at the point $A = 1_n$ has full rank. To show this, complement the set of the variables $\mathrm{Re}\, a_{jk}$, $1 \le j \le k \le n$, and $\mathrm{Im}\, a_{jk}$, $1 \le j < k \le n$, with the variable $\mathrm{Im}\, a_{11}$ and note that $\frac{\partial H}{\partial \mathrm{Im}\, a_{11}} = 1$ at the point $A = 1_n$. Using the calculations carried out in the proof of Theorem 7.4, we conclude that the minor of the Jacobian matrix of the map $\widetilde{F}$ corresponding to these $n^2 + 1$ variables is invertible at the point $A = 1_n \in \mathrm{SU}(n)$. Consequently, $\mathrm{SU}(n)$ is a smooth submanifold $\mathrm{U}(n)$ and $\dim \mathrm{SU}(n) = n^2 - 1$. Theorem 7.5 is proven. □

**Problem 7.2.** Using the complex Jordan form of a matrix, show that $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are connected for all values of $n$.

## 7.2 Lie algebras

The tangent space of a linear group $G \subset \mathrm{GL}(n)$ at the identity $1_n$ is a vector subspace in $M(n, \mathbb{R})$. Since a sufficiently small neighborhood of each invertible matrix in $M(n, \mathbb{R})$ consists of invertible matrices (this follows from continuity of the determinant), the tangent space of $\mathrm{GL}(n)$ at each point coincides with $M(n, \mathbb{R})$. Similarly, we can prove that the tangent spaces of the group $\mathrm{GL}(n, \mathbb{C})$ coincide with $M(n, \mathbb{C})$.

**Theorem 7.6.** *The tangent space $T_e G$ of the group $G$ at the identity $e$ is constituted by:*

(1) *the traceless matrices*

$$\mathrm{Tr}\, X = 0$$

*for* $G = \mathrm{SL}(n)$;

(2) *the anti-symmetric matrices*

$$X^\top = -X, \quad x_{jk} = -x_{kj}$$

*for* $G = \mathrm{O}(n)$;

(3) *the matrices* $X \in M(n, \mathbb{C})$ *such that*

$$\bar{X}^\top = -X, \quad \bar{x}_{jk} = -x_{kj}$$

*for* $G = \mathrm{U}(n)$;

(4) *the matrices* $X \in M(n, \mathbb{C})$ *determined by the equations*

$$\bar{x}_{jk} = -x_{kj} \quad \textit{and} \quad \mathrm{Tr}\, x = 0$$

*for* $G = \mathrm{SU}(n)$.

*Proof.* Let $\gamma(t) = 1 + Xt + O(t^2)$ be a smooth curve in $G$ passing through the identity $e = 1$. If $G = \mathrm{SL}(n)$, then $\det \gamma(t) \equiv 1$ and $\det \gamma(t) = 1 + \operatorname{Tr} X \cdot t + O(t^2) = 1$. Hence, we find that $\operatorname{Tr} X = 0$. Since the dimension of the space of matrices with the zero trace coincides with $\dim \mathrm{SL}(n)$, every such a matrix is realized as a tangent vector.

We have already given the proof for $G = \mathrm{O}(n)$ in §1.4 which can be generalized to the case of $G = \mathrm{U}(n)$ as follows: $\overline{\gamma}(t)\gamma^\top(t) = 1_n$ and from

$$(1_n + \bar{X}t + O(t^2))(1_n + X^\top t + O(t^2)) = 1_n + (X^\top + \bar{X})t + O(t^2) = 1_n$$

we find that $X^\top = -\bar{X}$. The dimension of the space of such matrices coincides with $\dim \mathrm{U}(n)$.

In the case $G = \mathrm{SU}(n)$ it suffices to combine the arguments for $\mathrm{SL}(n)$ and $\mathrm{U}(n)$. Theorem 7.6 is proven. □

The deviation of a group from a commutative group is described by its *commutators* $[x, y] = xyx^{-1}y^{-1}$. The vanishing of all commutators means that the group is commutative. Given a Lie group $G$, we define on its tangent space at the identity $e \in G$ the operation

$$[u, v] = \frac{1}{2}\frac{\partial^2 [x(t), y(t)]}{\partial t^2} \quad \text{at } t = 0,$$

where

$$u = \frac{\partial x(t)}{\partial t}, \quad v = \frac{\partial y(s)}{\partial t} \quad \text{at } t = 0, \quad x(0) = y(0) = e.$$

This operation in a Lie algebra is called the *commutator* as well and can be defined in a different way.

Suppose that $\{x^j\}$ are local coordinates in a neighborhood of the identity $e \in G$ such that $x^j(e) = 0$ for all $j$. We can identify $T_eG$ with $\mathbb{R}^n$, where $n = \dim G$, and write down multiplication in terms of local coordinates in the form of a Taylor series

$$(xy)^j = x^j + y^j + b^j_{kl}x^k y^l + (\text{terms of order} \geq 3), \tag{7.4}$$

where $x^j$ and $y^k$ are assumed to be the first-order terms. Then

$$[u, v]^j = (b^j_{kl} - b^j_{lk})u^k v^l = c^j_{kl}u^k v^l.$$

**Theorem 7.7.** *The operation* $[\,,\,]$ *on the tangent space* $T_eG$ *at the identity of the Lie group is linear in each variable and satisfies the relations*

$$[u, v] = -[v, u], \tag{7.5}$$

$$[[u, v], w] + [[v, w], u] + [[w, u], v] = 0 \quad (\textit{Jacobi identity}). \tag{7.6}$$

Equivalence of different definitions of the commutator and identities (7.5) and (7.6) are proven by straightforward and simple calculations by means of Taylor series. We leave them as exercises, observing that their detailed exposition can be found in [5].

A vector space (over some field) endowed with a commutator which is linear in both variables and satisfies identities (7.5) and (7.6) is called a *Lie algebra*.

If $e_1, \dots, e_n$ is a basis for a Lie algebra, then the commutator is determined uniquely by its values on the basis elements:

$$[e_i, e_j] = c_{ij}^k e_k$$

(here the repeated index $k$ implies summation). The constants $c_{ij}^k$ are called the *structure constants* of the Lie algebra.

Speaking of the Lie algebra of a Lie group, we mean the tangent space at the identity with the commutator operation introduced above.

It is known that every finite-dimensional Lie algebra over $\mathbb{R}$ or $\mathbb{C}$ is the Lie algebra of some matrix group.

A *homomorphism of Lie groups* $G$ and $H$ is a smooth map $f : G \to H$ which is a homomorphism of the groups. Each such a map determines a map of the tangent spaces at the identities which preserves the commutators, namely the *homomorphism of the Lie algebras*

$$f_* : \boldsymbol{g} \to \boldsymbol{h},$$

i.e., the linear map of the corresponding Lie algebras such that

$$f_*([u, v]) = [f_*(u), f_*(v)] \tag{7.7}$$

for all $u, v \in \boldsymbol{g} = T_e G$. Indeed, equality (7.7) follows from the definition of the commutator of elements of the Lie algebra in terms of the commutators of curves in Lie groups:

$$\begin{aligned} f_*([\dot{x}(0), \dot{y}(0)]) &= \frac{1}{2} \frac{\partial^2 f([x(t), y(t)])}{\partial t^2}\Big|_{t=0} \\ &= \frac{1}{2} \frac{\partial^2 ([f(x(t)), f(y(t))])}{\partial t^2}\Big|_{t=0} = [f_*(\dot{x}(0)), f_*(\dot{y}(0))]. \end{aligned}$$

Obviously, if Lie groups are smoothly isomorphic, then their Lie algebras are isomorphic. This simple assertion makes it possible to reduce a great part of the classification problem for Lie groups to problems of linear algebra. The Lie algebra is constructed by the quadratic part of the multiplication rule (7.4) which determines the multiplication in the group. The last condition is very strong and, roughly speaking, allows us to reconstruct the whole Taylor series for multiplication: the structure constants $c_{kl}^j$ determine completely multiplication in a sufficiently small neighborhood of the identity (the Lie theorem). Detailed proofs and classification results can be found, for example, in [5].

As we will show now, the commutator of matrix groups has a simple form.

A *one-parameter subgroup* $F(t)$ of a Lie group $G$ is a subgroup $F(t)$ which is the image of a homomorphism of Lie groups $F\colon \mathbb{R} \to G$, where $\mathbb{R}$ is the additive group of reals.

**Lemma 7.4.** *For linear groups $G \subset \mathrm{GL}(n)$ the one-parameter subgroups are the subgroups $F(t) = \exp(Xt)$, where $X \in T_eG$.*

*Proof.* Let $X = \partial F(0)/\partial t$ and denote by $\dot{F}$ the derivative of $F$ with respect to $t$. Then

$$\dot{F}(t) = \lim_{s\to 0} \frac{F(s+t) - F(t)}{s} = \lim_{s\to 0} \frac{F(s) - 1_n}{s} F(t) = X\,F(t)$$

and, solving the matrix equation

$$\dot{F} = XF$$

with the constant matrix $A$ and the initial condition $F(0) = 1_n$, we obtain

$$F(t) = \exp(Xt) = 1_n + Xt + \frac{1}{2}X^2t^2 + \cdots + \frac{1}{n!}X^nt^n + \cdots .$$

Lemma 7.4 is proven. □

For a subgroup $e^{Xt} \in G$ the element $X$ is called the *generator* of $e^{Xt}$ (for example, one can often see the expression "the generator of rotations around an axis").

**Theorem 7.8.** *For linear groups $G \subset \mathrm{GL}(n)$ the commutator on the Lie algebras is the usual commutator of matrices*

$$[X, Y] = XY - YX \in \boldsymbol{g} = T_{E_n}G \subset M(n, \mathbb{R}).$$

*Proof.* We obtain

$$\begin{aligned}[\exp(Xt), \exp(Yt)] &= (1_n + Xt + \tfrac{1}{2}X^2t^2 + O(t^3))(1_n + Yt + \tfrac{1}{2}Y^2t^2 + O(t^3)) \\ &\quad \cdot (1_n - Xt + \tfrac{1}{2}X^2t^2 + O(t^3))(1_n - Yt + \tfrac{1}{2}Y^2t^2 + O(t^3)) \\ &= 1_n + [X, Y]t^2 + O(t^3)\end{aligned}$$

and

$$\frac{1}{2}\frac{\partial^2[\exp(Xt), \exp(Yt)]}{\partial t^2}\Big|_{t=0} = [X, Y].$$

Theorem 7.8 is proven. □

The Lie algebras of matrix groups are denoted by the same but lower case letters as the group. For example, $gl(n)$ is the Lie algebra of the group $\mathrm{GL}(n)$.

**Examples of Lie algebras.** (1) $so(2)$ is one-dimensional and commutative:

$$so(2) = \mathbb{R} \cdot e_1, \quad [e_1, e_1] = 0.$$

(2) $so(3)$ is generated by the matrices

$$e_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

satisfying the relations

$$[e_j, e_k] = \varepsilon^l_{jk} e_l, \tag{7.8}$$

where $\varepsilon^l_{jk} = 1$ if the permutation

$$\begin{pmatrix} 1 & 2 & 3 \\ j & k & l \end{pmatrix}$$

is even, $\varepsilon^l_{jk} = -1$ if this permutation is odd, and $\varepsilon^l_{jk} = 0$ if at least two indices among $j$, $k$, and $l$ coincide.

(3) $su(2)$ is generated (over the field $\mathbb{R}$) by the matrices

$$e_1 = \begin{pmatrix} 0 & -\frac{i}{2} \\ -\frac{i}{2} & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} -\frac{i}{2} & 0 \\ 0 & \frac{i}{2} \end{pmatrix} \tag{7.9}$$

satisfying relations (7.8).

Let $e_1$, $e_2$, $e_3$ be a positively-oriented orthonormal basis for $\mathbb{R}^3$. Then the vector products of these basis vectors satisfy relations (7.8), too.

**Corollary 7.2.** *The Lie algebras $so(3)$ and $su(2)$ are isomorphic to the algebra of vectors in $\mathbb{R}^3$ with the cross product operation.*

(4) $sl(2)$ is generated by the matrices

$$e_1 = \begin{pmatrix} 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

satisfying the relations

$$[e_1, e_3] = e_1, \quad [e_2, e_3] = -e_2, \quad [e_1, e_2] = e_3.$$

**Corollary 7.3.** *$sl(2)$ is nonisomorphic to the algebras $so(3)$ and $su(2)$.*

Indeed, the vector product $[u, v]$ in $\mathbb{R}^3$ is orthogonal to the vectors $u$ and $v$. Therefore, the equation $[u, v] = u$ has no solution in the algebra $so(3)$, and the generators $e_1$ and $e_3$ of the algebra $sl(2)$ satisfy the relation $[e_1, e_3] = e_1$.

## 7.3 Geometry of the simplest linear groups

(1) SU(2). It follows from equations (7.3) that a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

belongs to the group $\mathrm{U}(2) \subset M(2,\mathbb{C})$ if and only if

$$\bar{a}b + \bar{c}d = 0, \quad |a|^2 + |c|^2 = |b|^2 + |d|^2 = 1.$$

The additional condition that $A \in \mathrm{SU}(2) \subset \mathrm{U}(2)$ takes the form

$$ad - bc = 1.$$

Rewrite the first and last equations as the system of equations

$$\begin{pmatrix} a & -c \\ \bar{c} & \bar{a} \end{pmatrix} \begin{pmatrix} d \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Solving this system using $|a|^2 + |c|^2 = 1$, we find that $b = -\bar{c}$ and $d = \bar{a}$. We have thus proven the following theorem:

**Theorem 7.9.** *The group* $\mathrm{SU}(2) \subset M(2,\mathbb{C})$ *is constituted by the matrices of the form*

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

*where* $|a|^2 + |b|^2 = 1$.

**Corollary 7.4.** *The Lie group* $\mathrm{SU}(2)$ *is diffeomorphic to the three-dimensional sphere.*

Indeed, the $n$-dimensional sphere is the submanifold $\mathbb{R}^{n+1}$ determined by the equation

$$\sum_{j=1}^{n+1} (x^j)^2 = 1,$$

and, putting $a = x^1 + ix^2$ and $b = x^3 + ix^4$, we bring the equation $|a|^2 + |b|^2 = 1$ to the form of the equation for the three-dimensional sphere.

(2) *Quaternions*. The algebra $\mathbb{H}$ of *quaternions* is the four-dimensional algebra over the field $\mathbb{R}$ with the summation and multiplication operations such that

(i) the summation and multiplication operations are associative; i.e.,

$$u(vw) = (uv)w, \quad u + (v + w) = (u + v) + w$$

for $u, v, w \in \mathbb{H}$ and satisfy also the conditions

$$v + w = w + v, \qquad \lambda(v + w) = \lambda v + \lambda w,$$
$$(\lambda + \mu)v = \lambda v + \mu v, \quad (\lambda v)(\mu w) = (\lambda\mu)(vw)$$

for $v, w \in \mathbb{H}$ and $\lambda, \mu \in \mathbb{R}$;

(ii) there exist linear generators $\mathbf{1}$, $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ of the algebra $\mathbb{H}$ such that $\mathbf{1}v = v\mathbf{1}$ for every $v \in \mathbb{H}$, and multiplication of the other generators is given by the rule

$$\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}, \quad \boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}, \quad \boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}, \quad \boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -\mathbf{1}. \tag{7.10}$$

These rules determine uniquely the multiplication operation which is noncommutative and possesses one remarkable property: each nonzero element has an inverse. Such an algebra is called a *division algebra.*

Adams' theorem proven by topological methods claims that

- *the dimension of a finite-dimensional division algebra over the field* $\mathbb{R}$ *can take only the following values:* 1*,* 2*,* 4*, and* 8.

All these cases are realized by the algebra of reals (dim $= 1$), the algebra of complex numbers (dim $= 2$), the algebra of quaternions (dim $= 4$), and the Cayley algebra of octonions (dim $= 8$).[1]

The four algebras indicated above exhaust all alternative division algebras. This assertion generalizes the Frobenius theorem which claims that *every associative division algebra over the field* $\mathbb{R}$ *is isomorphic to either* $\mathbb{R}$*,* $\mathbb{C}$*, or* $\mathbb{H}$.

Put

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{i} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \quad \boldsymbol{j} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{k} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

The following lemma is proven by simple verification:

**Lemma 7.5.** *The vector subspace in* $M(2, \mathbb{C})$ *generated by these matrices with the matrix summation and multiplication operation is isomorphic to the algebra of quaternions.*

The matrices

$$\sigma_x = i\boldsymbol{i}, \quad \sigma_y = i\boldsymbol{j}, \quad \sigma_z = i\boldsymbol{k}$$

are applied in theoretical physics and are called the *Pauli matrices*.

As in the case of the algebra of complex numbers, on the algebra of quaternions we have the conjugation operation

$$q = u\mathbf{1} + x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} \mapsto \bar{q} = u\mathbf{1} - x\boldsymbol{i} - y\boldsymbol{j} - z\boldsymbol{k}.$$

[1]The multiplication in the Cayley algebra is neither commutative nor even associative. This algebra is an example of an *alternative* algebra, i.e., an algebra in which every subalgebra generated by two elements is associative.

We can immediately derive the formulas

$$\overline{v+w} = \bar{v} + \bar{w}, \quad \overline{vw} = \bar{w}\bar{v}.$$

A quaternion $q$ is called *imaginary* if $\bar{q} = -q$. It is obvious that the imaginary quaternions are exactly those which have the form

$$q = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}.$$

The algebra $\mathbb{H}$ is *normed*; i.e., for each element $q$ we can define its norm $|q| \in \mathbb{R}$; moreover, the following conditions are satisfied:

$$|v+w| \leq |v| + |w|, \quad |vw| = |v||w|,$$
$$|\lambda v| = |\lambda||v| \quad \text{for } v, w \in \mathbb{H},\ \lambda \in \mathbb{R},$$
$$|v| \geq 0 \quad \text{and} \quad |v| = 0 \text{ if and only if } v = 0.$$

This norm is given by the formula

$$|q| = \sqrt{q\bar{q}} = \sqrt{u^2 + x^2 + y^2 + z^2}$$

and a quaternion $q$ is called a *unit quaternion* if $|q| = 1$.

From Theorem 7.9 and Lemma 7.5 we immediately obtain the following lemma:

**Lemma 7.6.** *The group* SU(2) *is isomorphic to the group of unit quaternions with the multiplication operation.*

(3) *The groups* SO(3) *and* SO(4). Henceforth we will identify the group SU(2) with the group of unit quaternions, the linear space $\mathbb{R}^3$ with the linear space of imaginary quaternions and the linear space $\mathbb{R}^4$ with the space of all quaternions.

If $v$ is an imaginary quaternion and $q \in \mathbb{H}$, then

$$\overline{qv\bar{q}} = q\bar{v}\bar{q} = -qv\bar{q}$$

and consequently $qv\bar{q}$ is also an imaginary quaternion. Therefore, for each $q \in \mathrm{SU}(2) \subset \mathbb{H}$ we can define the linear map

$$\mathbb{R}^3 \xrightarrow{Q} \mathbb{R}^3 \colon Q(v) = qv\bar{q}. \tag{7.11}$$

This linear map preserves the norms of elements,

$$|Q(v)| = |q|\,|v|\,|\bar{q}| = |v|$$

which are written in terms of the inner product $(v_1, v_2) = x_1x_2 + y_1y_2 + z_1z_2$ as follows:

$$|v| = \sqrt{(v, v)}.$$

Since

$$(v + w, v + w) = (v, v) + 2(v, w) + (w, w)$$

and the norms of the elements $v$, $w$, and $(v + w)$ are preserved under the action of $Q$, $Q$ preserves the inner product on $\mathbb{R}^3 = \mathrm{Im}\,\mathbb{H}$:

$$(Q(v), Q(w)) = (v, w).$$

Two different elements $q_1$ and $q_2$ determine the same linear transformation if and only if the following equality holds for each imaginary quaternion $v$:

$$(\bar{q}_2 q_1) v (\bar{q}_1 q_2) = v,$$

which is equivalent to the commutativity condition

$$(\bar{q}_2 q_1) v = v (\bar{q}_2 q_1).$$

From relations (7.10) we can easily find that the quaternion $\bar{q}_2 q_1$ commutes with each imaginary quaternion if and only if it is real: $\bar{q}_2 q_1 = \lambda \mathbf{1}$, where $\lambda \in \mathbb{R}$. But, since $\bar{q}_2 q_1$ is a unit quaternion, it must be equal to $\pm \mathbf{1}$.

We have proven the following lemma:

**Lemma 7.7.** *Formula* (7.11) *determines a homomorphism of groups*

$$\rho\colon \mathrm{SU}(2) \to \mathrm{O}(3),$$

*whose kernel is equal to* $\{\pm 1_2\}$ *and consequently isomorphic to* $\mathbb{Z}_2$.

In terms of the entries $a$ and $b$ which are smooth functions on SU(2), the homomorphism $\rho$ is written by simple formulas which imply that the map $\rho$ is smooth, i.e., a homomorphism of Lie groups.

Since $1_2 \in \mathrm{SU}(2)$ goes into $1_3$ and SU(2) is connected, the image of $\rho$ lies in the connected component of O(3) containing the identity, i.e., in SO(3).

Now, observe one general fact.

**Lemma 7.8.** *Let* $f: G \to H$ *be a homomorphism of connected linear Lie groups of the same dimension whose kernel is a discrete subgroup* $\Gamma \subset G$ (*in the topology induced by the embedding, each point of* $\Gamma$ *has a neighborhood containing no other points of* $\Gamma$). *Then*

(1) *the differential of the map*

$$f_*: T_g G \to T_{f(g)} H$$

*is an isomorphism for each* $g \in G$;

(2) $f(G) = H$;

(3) *the Lie algebras of the groups* $G$ *and* $H$ *are isomorphic.*

*Proof.* Since the homomorphism of Lie groups takes the one-parameter subgroups $\exp(Xt)$ into subgroups $\exp(f_*(X)t)$ and the kernel $f$ is discrete, the map $f_*$ is an embedding. Since $\dim G = \dim H$, the map $f_*$ is an even isomorphism of the tangent spaces at the identities. These tangent spaces at the identities are endowed with the structures of Lie algebras and $f_*$ determines an isomorphism of Lie algebras.

Take $g \in G$ and let $\boldsymbol{g}$ be the Lie algebra of the group $G$. Each element in a neighborhood of $g$ has the form $g\exp(X)$, where $X \in \boldsymbol{g}$, and goes into $f(g)\exp(f_*X)$ under the action of $f$. Consequently, the differential $f_*$ at $g$ has the same rank as at the identity of the group. Hence, this rank is maximal everywhere.

Since $f_*$ has full rank at each point and the dimensions of $G$ and $H$ coincide, the image of $f$ is a smooth submanifold $f(G) \subset H$. But $f(G)$ and $H$ are connected and therefore coincide. Lemma 7.8 is proven. □

Applying this lemma to the map $\rho\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$, we obtain

**Theorem 7.10.** *The Lie group* $\mathrm{SO}(3) = \mathrm{SU}(2)/\{\pm 1\}$ *is diffeomorphic to the manifold which is obtained from the unit three-dimensional sphere in* $\mathbb{R}^4$ *by identification of the opposite points.*

Indeed, $\rho$ has rank 3 everywhere and the inverse image of each element $p \in \mathrm{SO}(3)$ consists of a pair of points $\pm q \in \mathrm{SU}(2)$.

The manifold SO(3) is an example of a *real projective space* $\mathbb{R}P^n$ for $n = 3$, which is defined as the manifold obtained from the sphere

$$S^n = \left\{(x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \mid \textstyle\sum_{j=1}^{n+1}(x^j)^2 = 1\right\}$$

by identification of the opposite points. The following two simpler examples have already appeared in the analytic geometry course: the projective straight line $\mathbb{R}P^1$ and the projective plane $\mathbb{R}P^2$.

Now, let $q_1, q_2$ be a pair of unit quaternions. With this pair we associate the linear map $\mathcal{H} = \mathbb{R}^4 \to \mathbb{R}^4$ of the form

$$v \mapsto q_1 v \overline{q_2}. \tag{7.12}$$

By analogy with Lemma 7.7, we can prove the following lemma:

**Lemma 7.9.** *Formula* (7.12) *defines a homomorphism of Lie groups*

$$\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4),$$

*whose kernel is constituted by the elements* $(1_2, 1_2)$ *and* $(-1_2, -1_2)$.

From the above lemma and Lemma 7.8 we derive

**Theorem 7.11.** *The Lie group* SO(4) *is diffeomorphic to the manifold which is obtained from the direct product of two unit three-dimensional spheres* $S^3 \times S^3$ *by identification of pairs* $(v, w)$ *and* $(-v, -w)$.

*The Lie algebra* $so(4)$ *of the group* SO(4) *is isomorphic to the direct sum of the Lie algebras*

$$so(4) = so(3) \oplus so(3).$$

Here by the direct sum $\boldsymbol{g} \oplus \boldsymbol{h}$ we mean the Lie algebra constituted by the pairs $(X, Y)$, $X \in \boldsymbol{g}$, $Y \in \boldsymbol{h}$, with the commutator operation

$$[(X_1, Y_1), (X_2, Y_2)] = ([X_1, X_2], [Y_1, Y_2]).$$

If $\boldsymbol{g}$ and $\boldsymbol{h}$ are the Lie algebras of the groups $G$ and $H$, then $\boldsymbol{g} \oplus \boldsymbol{h}$ is obviously the Lie algebra of the direct product $G \times H$.

Although $so(4) = so(3) \oplus so(3)$, it is easy to show that SO(4) is not homeomorphic to the direct product SO(3) × SO(3).

Decomposition of the Lie algebra $so(4)$ into the nontrivial direct sum of Lie algebras is a paradoxical property of the dimension 4: for all other values of $n$ the algebra $so(n)$ is not representable as the direct sums of nonzero Lie algebras. There is one more peculiarity of the dimension 4 connected with this phenomenon: only for $n = 4$ there are pairwise nondiffeomorphic manifolds homeomorphic to $\mathbb{R}^n$ (we have already observed this in §3.2).

# 8

# Elements of representation theory

## 8.1 The basic notions of representation theory

Linear Lie groups have numerous applications and serve for studying abstract groups by means of *representation theory*.

A *representation* of a group $G$ is a homomorphism

$$\rho\colon G \to \mathrm{GL}(V),$$

where $V$ is a vector space called the *representation space* and $\mathrm{GL}(V)$ is the group of its invertible linear transformations. If $\dim V < \infty$, then the representation is called *finite-dimensional* and $\dim V$ is called the *dimension of the representation*.

A representation is *real* if $\mathrm{GL}(V) = \mathrm{GL}(n, \mathbb{R})$ and *complex* if $\mathrm{GL}(V) = \mathrm{GL}(n, \mathbb{C})$. If a real finite-dimensional representation $\rho$ on $V = \mathbb{R}^n$ is such that there is a nondegenerate inner product which is preserved by all transformations $\rho(g)$ with $g \in G$ or, which is the same, if

$$\rho(G) \subset \mathrm{O}(n) \subset \mathrm{GL}(n),$$

where the subgroup $\mathrm{O}(n)$ is defined by this inner product, then the representation $\rho$ is called *orthogonal*. A finite-dimensional complex representation is *unitary* if it preserves a nondegenerate Hermitian product on $V = \mathbb{C}^n$. For brevity orthogonal representations are also called unitary.

The notion of a unitary representation is naturally generalized to infinite-dimensional representations in the case when $V$ is a Hilbert space: in this case preservation of the Hilbert product is required. Henceforth we consider only finite-dimensional real or complex representations.

A representation $\rho$ is called *faithful* if its kernel contains only the identity $e$ of the group $G$: $\mathrm{Ker}\,\rho = e$. Now, we can easily reformulate the definition of a linear group:

- *a Lie group is called linear if it has a faithful finite-dimensional representation.*

Here we do not specify that the map $\rho$ must be smooth; we can show that this holds automatically.

The general definition of a linear group is the following:

- *a group $G$ is linear if it is isomorphic to a subgroup of the group* $\mathrm{GL}(n)$ *for some $n$.*

Two representations $\rho_1\colon G \to \mathrm{GL}(V)$ and $\rho_2\colon G \to \mathrm{GL}(W)$ are *equivalent* if there is an invertible linear transformation

$$A\colon V \to W$$

such that the equality

$$A\rho_1(g) = \rho_2(g)A\colon V \to W$$

holds for all $g \in G$. Below, for enumerating or listing various representations, we do not distinguish between equivalent representations.

A subspace $V' \subset V$ is called *invariant* (under a representation $\rho$) if

$$\rho(g)V' \subset V'$$

for all $g \in G$. Simple examples of nontrivial invariant subspaces are obtained by summation of representations. Namely, a representation $\rho_1 \oplus \rho_2$ is called the *direct sum of representations* $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ if it has the form

$$(\rho_1 \oplus \rho_2)(g) = (\rho_1(g), \rho_2(g)) \in \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \subset \mathrm{GL}(V_1 \oplus V_2),$$

where the embedding $\mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \subset \mathrm{GL}(V_1 \oplus V_2)$, roughly speaking, is implemented by block matrices. In this case $V_1$ and $V_2$ are invariant subspaces of the representation $\rho_1 \oplus \rho_2$.

A representation is *irreducible* if it has no nontrivial invariant subspaces (i.e., different from the null space and the whole space $V$). The importance of this definition, for example, for unitary representations is demonstrated by the following assertion:

**Theorem 8.1.** *Each unitary representation $\rho$ decomposes into the direct sum of irreducible representations:*

$$\rho = \rho_1 \oplus \cdots \oplus \rho_k.$$

*Proof.* We proceed by induction on $\dim V$, taking the dimension over $\mathbb{R}$ for real representations and over $\mathbb{C}$ for complex representations. For $\dim V = 1$ the assertion is obvious. Suppose that it is proven for $\dim V < k$.

Let $\rho\colon G \to \mathrm{GL}(V)$ be a representation with $\dim V = k$. If $V$ contains no invariant subspaces, then $\rho$ is irreducible and the assertion is proven. Let $V_1$ be a nontrivial invariant subspace and let $V_1^{\perp}$ be its orthogonal complement. Since $\rho$ is unitary, we have $\rho(g)V_1^{\perp} \subset V_1^{\perp}$. The restrictions of $\rho$ to $V_1$ and $V_2 = V_1^{\perp}$ determine the representations $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ whose direct sum is equal to $\rho$:

$$\rho = \rho_1 \oplus \rho_2.$$

Since the dimensions of $V_1$ and $V_2$ are less than $k$, by the induction assumption, $\rho_1$ and $\rho_2$ decompose into the direct sum of irreducible representations. Theorem 8.1 is proven. □

The theorem implies that the classification of unitary representations of a group $G$ reduces to the classification of unitary irreducible representations of $G$.

The following theorem is simple but has fundamental consequence:

**Theorem 8.2** (Schur's lemma). *Let $\rho_1\colon G \to \mathrm{GL}(V)$ and $\rho_2\colon G \to \mathrm{GL}(W)$ be irreducible representations connected by the relation*

$$A\rho_1 = \rho_2 A, \tag{8.1}$$

*where $A\colon V \to W$ is a linear transformation. Then one of the following two possibilities holds:*

(1) $A = 0$;

(2) *$A$ is an isomorphism.*

*If the representations are complex and $\rho_1 = \rho_2$, then $A$ is the operator of multiplication by a constant $\lambda \in \mathbb{C}$.*

*Proof.* It is immediate from (8.1) that the kernel $\operatorname{Ker} A$ of the transformation $A$ is invariant under $\rho_1$ and its image $\operatorname{Im} A$ is invariant under $\rho_2$. Since these representations are irreducible, we have either $\operatorname{Ker} A = V$ and then $A = 0$, or $\operatorname{Ker} A = 0$ and then $\operatorname{Im} A = W$ and $A$ is an isomorphism.

If the representation is complex, then $A$ has an eigenvector $v \in V$: $Av = \lambda v$. If $\rho_1 = \rho_2$, then the operator $A - \lambda = A - \lambda \cdot 1_V$, where $1_V\colon V \to V$ is the identity transformation, satisfies the following relation of type (8.1):

$$(A-\lambda)\rho_1 = \rho_1(A-\lambda).$$

Since $\operatorname{Ker}(A-\lambda) \neq 0$, we have $A - \lambda = 0$. Theorem 8.2 is proven. □

**Corollary 8.1.** *If a group is commutative, then all irreducible complex representations of the group are one-dimensional.*

*Proof.* Let $\rho\colon G \to \mathrm{GL}(V)$ be an irreducible representation and take $h \in G$. Then the relation

$$\rho(h)\rho(g) = \rho(g)\rho(h)$$

holds for all $g \in G$. We conclude that the operator $\rho(h)$ is the multiplication by a constant $\lambda_h$ for each $h \in G$; since this representation is irreducible, $\dim_{\mathbb{C}} V = 1$. Corollary 8.1 is proven. □

For real representations the assertion of the corollary is false. As an example we can take the following two-dimensional representation of the group $\mathrm{U}(1) = \mathbb{R}/2\pi\mathbb{Z}$ of rotations of the plane:

$$\varphi \mapsto \rho(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

The transformations $\rho(\varphi)$ have no eigenvectors in $\mathbb{R}^2$ for $\rho(\varphi) \neq \pm 1_2$.

The *character of a representation* $\rho\colon G \to \mathrm{GL}(V)$, where $V$ is a vector space over the field $F = \mathbb{R}$ or $\mathbb{C}$, is a function

$$\chi_\rho\colon G \to F\colon \chi_\rho(g) = \operatorname{Tr} \rho(g).$$

Corollary 8.1 implies that an irreducible complex representation of a commutative group “coincides” with its character.

The following properties of characters are well known from linear algebra.

**Lemma 8.1.** (1) $\chi_\rho(e) = \dim V$;

(2) $\chi_\rho(hgh^{-1}) = \chi_\rho(g)$, *i.e., a character is defined on the set of conjugacy classes;*

(3) *if representations* $\rho_1$ *and* $\rho_2$ *are equivalent, then their characters coincide:* $\chi_{\rho_1} = \chi_{\rho_2}$;

(4) *if a representation* $\rho$ *decomposes into the direct sum of representations* $\rho_1, \dots, \rho_k$, *then*

$$\chi_\rho = \chi_{\rho_1} + \dots + \chi_{\rho_k}.$$

*Proof.* The proof of assertions (1) and (4) consists in computation of the traces of block matrices, while the assertions (2) and (3) follow from the obvious identity

$$\operatorname{Tr}(ABA^{-1}) = \operatorname{Tr} B$$

where $A$ and $B$ are square matrices. □

**Examples.** (1) $G = \mathrm{U}(1) = \mathbb{R}/2\pi\mathbb{Z}$. Let $\varphi$ be a linear parameter on $\mathbb{R}$ which determines the parameter on U(1) defined modulo $2\pi$. The one-dimensional complex representations have the form

$$\rho_n(\varphi) = e^{in\varphi},$$

where $n \in \mathbb{Z}$. The characters of these representations “coincide” with themselves:

$$\chi_n(\varphi) = \chi_{\rho_n}(\varphi) = e^{in\varphi}.$$

All these representations are irreducible and, by Corollary 8.1, each irreducible complex representation of the group U(1) has the indicated form. Since the characters $\chi_n$ and $\chi_m$ do not coincide for $n \neq m$, the representations $\rho_n$ and $\rho_m$ are not equivalent. Hence,

$$\{\rho_n\}_{n\in\mathbb{Z}}$$

is the set of all equivalence classes of irreducible complex representations of U(1). Note that the representations $\rho_{\pm 1}$ are faithful.

(2) $G$ is the $n$-dimensional torus $T^n = \mathrm{U}(1)^n = \underbrace{\mathrm{U}(1) \times \dots \times \mathrm{U}(1)}_{n}$. Let $\varphi_1, \dots, \varphi_n$ be linear parameters on the factors of the form U(1). By analogy with the case

$G = \mathrm{U}(1)$ we prove that the set of irreducible complex representations of the group $T^n$ has the form

$$\rho_{\boldsymbol{m}}(\varphi_1, \dots, \varphi_n) = \exp(i(m_1\varphi_1 + \dots + m_n\varphi_n)),$$

where $\boldsymbol{m} = (m_1, \dots, m_n) \in \mathbb{Z}^n$. All these representations are pairwise nonequivalent and "coincide" with their characters $\chi_{\boldsymbol{m}}$. The characters $\chi_{\boldsymbol{m}}$ constitute a *Fourier basis* for the space of periodic functions. This is not a contingency, since character theory provides a basis for harmonic analysis, the theory of the Fourier transform. Expansion of a periodic function $f(x)$ (or a function on a torus, which is the same) in a series in the characters $\rho_{\boldsymbol{m}}$:

$$f(x) = \sum_{\boldsymbol{m}} a_{\boldsymbol{m}} \rho_{\boldsymbol{m}}(x)$$

is called the *Fourier series* of a periodic function and the coefficients of the series are called the *Fourier coefficients*.

Obviously, the representation $\rho_{\boldsymbol{m}}$ is unitary; moreover, the following general fact is valid:

- *every finite-dimensional representation of a finite group or a compact Lie group is unitary.*

We will prove this assertion for finite groups in §8 and discuss the idea of the proof in the general case; meanwhile, below we restrict our consideration to unitary representations.

Each real representation $\rho$ can be complexified as follows. Let $V = \mathbb{R}^n$ be the representation space and let $e_1, \dots, e_n$ be a basis for $V$. Put $V^{\mathbb{C}} = \mathbb{C}e_1 + \dots + \mathbb{C}e_n$ and to each transformation $A\colon V \to V$ relate the transformation $A^{\mathbb{C}}\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ which is given in the basis $\{e_j\}$ by the same matrix as $A$. The so-obtained representation $\rho^{\mathbb{C}}\colon G \to \mathrm{GL}(V^{\mathbb{C}})$ is called the *complexification of the representation* $\rho$; it is obvious that the characters of $\rho$ and $\rho^{\mathbb{C}}$ coincide:

$$\chi_\rho = \chi_{\rho^{\mathbb{C}}}.$$

Therefore, henceforth we restrict the exposition to complex representations.

In line with the tradition of monographs on representation theory, in the next two sections we will write entries of the matrices $A$ defining linear maps with two subscripts

$$A(e_j) = \sum_{k=1}^{n} a_{jk} f_k,$$

where $A\colon \mathbb{C}^m \to \mathbb{C}^n$, $\{e_j\}$ is a basis for $\mathbb{C}^m$, and $\{f_k\}$ is a basis for $\mathbb{C}^n$.

## 8.2 Representations of finite groups

Recall that from now on we consider only complex representations. Let $G$ be a finite group and denote its order by $|G|$.

**Theorem 8.3.** *Let $\rho\colon G \to \mathrm{GL}(n,\mathbb{C})$ be a finite-dimensional representation. Then on $\mathbb{C}^n$ there is a Hermitian product preserved by the operators $\rho(g)$ and, hence, $\rho$ is unitary.*

*Proof.* The proof is based on the averaging method. Let $(\cdot,\cdot)_0$ be some Hermitian product on $\mathbb{C}^n$. Define the new Hermitian product to be

$$(v,w) = \sum_{g\in G} (\rho(g)v, \rho(g)w)_0. \tag{8.2}$$

It is obvious that $(\rho(g)v, \rho(g)w) = (v,w)$ for all $g \in G$. Theorem 8.3 is proven. □

Applying Theorem 8.1, we obtain

**Corollary 8.2.** *Each complex representation of a finite group decomposes into the direct sum of irreducible representations.*

Unitary representations possess the following important property:

**Lemma 8.2.** *If a representation $\rho\colon G \to \mathrm{GL}(n,\mathbb{C})$ is unitary, then*

$$\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)} \quad \textit{for all } g \in G.$$

This assertion is valid for all (not only finite) groups.

*Proof.* Since the representation $\rho$ is unitary, the matrices which define the representation in an orthonormal basis for $\mathbb{C}^n$ satisfy the equation

$$\rho(g^{-1})_{jk} = \rho(g)^{-1}_{jk} = \overline{\rho(g)_{kj}};$$

consequently, $\mathrm{Tr}\,\rho(g^{-1}) = \overline{\mathrm{Tr}\,\rho(g)}$. Lemma 8.2 is proven. □

Denote by $L(G)$ the space of complex-valued functions on the group $G$. It is isomorphic to the linear space $\mathbb{C}^{|G|}$. The isomorphism is plain: identify the generators of $\mathbb{C}^{|G|}$ with the elements of the group and define the one-to-one correspondence:

$$f \in L(G) \longleftrightarrow \sum_{g\in G} f(g)\cdot g \in \mathbb{C}^{|G|}.$$

Therefore, the elements of the group determine the natural basis $\{\hat{g}\}_{g\in G}$ for $L(G)$:

$$\hat{g}(h) = \begin{cases} 1 & \text{for } h = g, \\ 0 & \text{for } h \neq g. \end{cases}$$

The group $G$ acts on $L(G)$ by left and right translations denoted by $l_h$ and $r_h$:

$$l_h\colon f \mapsto l_h(f) = \sum_{g\in G} f(g)\widehat{hg}, \quad r_h\colon f \mapsto r_h(f) = \sum_{g\in G} f(g)\widehat{gh^{-1}}$$

or, which is the same,

$$l_h(f)(g) = f(h^{-1}g), \quad r_h(f)(g) = f(gh).$$

These formulas determine representations, since

$$l_{h_1}l_{h_2}(f)(g) = l_{h_2}(f)(h_1^{-1}g) = f(h_2^{-1}h_1^{-1}g) = l_{h_1h_2}(f)(g)$$

and similarly $r_{h_1h_2} = r_{h_1}r_{h_2}$. The representations $l_h$ and $r_h$ are called the *left regular* and *right regular* representations and the representation space $L(G)$ is called the *group algebra* of the (finite) group $G$.

Regular representations act on the basis $\{\hat{g}\}$ by permutations; moreover, the permutation corresponding to $h \in G$ is the identity permutation if and only if $h$ is the identity of the group: $h = e \in G$. Hence, we obtain the following theorem:

**Theorem 8.4.** *Regular representations are faithful, and therefore every finite group is linear.*

Define the following Hermitian product on $L(G)$:

$$(f_1, f_2) = \frac{1}{|G|} \sum_{g\in G} f_1(g)\overline{f_2(g)}. \tag{8.3}$$

The basis $\{\hat{g}\}$, $g \in G$, is orthonormal with respect to this product; moreover, since the left and right regular representations act on it by permutations, the following assertion is valid:

**Lemma 8.3.** *The left and right regular representations of G preserve the Hermitian product* (8.3).

Now, prove the following technical lemma:

**Lemma 8.4.** *Let $\rho\colon G \to \mathrm{GL}(n,\mathbb{C})$ and $\varphi\colon G \to \mathrm{GL}(m,\mathbb{C})$ be representations of a group $G$ and let $A\colon \mathbb{C}^m \to \mathbb{C}^n$ be a linear map. Then the map $\hat{A}\colon \mathbb{C}^m \to \mathbb{C}^n$ given by the formula*

$$\hat{A} = \sum_{g\in G} \rho(g)^{-1}A\varphi(g), \tag{8.4}$$

*connects the representations $\rho$ and $\varphi$ by the following relation of type* (*8.1*):

$$\rho\hat{A} = \hat{A}\varphi.$$

*Proof.* Write down $\rho(h)\hat{A}$ in detail:

$$\begin{aligned}\rho(h)\hat{A} &= \rho(h)\sum_{g\in G}\rho(g)^{-1}A\varphi(g)\\ &= \sum_{g\in G}\rho(h)\rho(g)^{-1}A\varphi(g)\\ &= \sum_{g\in G}\rho(hg^{-1})A\varphi(g)\\ &= \sum_{u=gh^{-1}\in G}\rho(u)^{-1}A\varphi(uh)\\ &= \Big(\sum_{u\in G}\rho(u)^{-1}A\varphi(u)\Big)\varphi(h)\\ &= \hat{A}\varphi(h).\end{aligned}$$

Lemma 8.4 is proven. □

We will need Lemma 8.3 for proving the following theorem:

**Theorem 8.5.** *Let $\rho\colon G \to \mathrm{GL}(n,\mathbb{C})$ and $\varphi\colon G \to \mathrm{GL}(m,\mathbb{C})$ be irreducible representations given by unitary matrices: $\rho(g)^{-1}_{jk} = \overline{\rho(g)_{kj}}$ and $\varphi(g)^{-1}_{jk} = \overline{\varphi(g)_{kj}}$. Then their entries satisfy the following orthogonality relations:*

$$(\rho_{rj}, \varphi_{sk}) = 0 \text{ if } \rho \text{ and } \varphi \text{ are not equivalent}, \tag{8.5}$$

$$(\rho_{rj}, \rho_{sk}) = \frac{1}{n}\delta_{jk}\delta_{rs}. \tag{8.6}$$

*Proof.* If the representations $\rho$ and $\varphi$ are not equivalent then, applying the Schur lemma to the operator $\hat{A}$ corresponding to any linear map $A\colon \mathbb{C}^m \to C^n$, we obtain $\hat{A} = 0$. Now, write down the right-hand side of (8.4) in the matrix form:

$$\hat{a}_{jk} = \sum_{g\in G}\sum_{r,s}\rho(g)^{-1}_{jr}a_{rs}\varphi(g)_{sk}.$$

The right-hand side is a linear form of $a_{lm}$; since it is identically zero, all coefficients of this form are zero:

$$\sum_{g\in G}\rho(g)^{-1}_{jr}\varphi(g)_{sk} = 0.$$

If the representations are given by unitary matrices, then this equality takes the form (8.5).

Now, prove (8.6). For the operators $A = A^{(r,s)}$ with

$$a^{(r,s)}_{jk} = \delta_{jr}\delta_{ks},$$

find their averages $\hat{A}$ given by (8.4) for $\rho = \varphi$:

$$\hat{a}_{jk}^{(r,s)} = \frac{1}{|G|} \sum_{g \in G} \sum_{u,v} \rho(g)_{ju}^{-1} \delta_{ur} \delta_{sv} \rho(g)_{vk} = \frac{1}{|G|} \sum_{g \in G} \rho(g)_{jr}^{-1} \rho(g)_{sk}.$$

By Lemma 8.4, each of the operators $\hat{A}^{(r,s)}$ satisfies identity (8.1):

$$\hat{A}^{(r,s)} \rho = \rho \hat{A}^{(r,s)};$$

therefore, the Schur lemma implies that $\hat{A}^{(r,s)}$ is the multiplication by a constant which is obviously equal to

$$\lambda = \frac{1}{n} \operatorname{Tr} \hat{A}^{(r,s)}.$$

But for $\rho = \varphi$ the trace of the averaged operator coincides with the trace of the initial operator:

$$\operatorname{Tr} \hat{A} = \operatorname{Tr} \Big( \frac{1}{|G|} \sum_{g \in G} \rho(g)^{-1} A \rho(g) \Big) = \operatorname{Tr} A.$$

Hence, we find that

$$\frac{1}{|G|} \sum_{g \in G} \rho(g^{-1})_{jr} \rho(g)_{sk} = \delta_{jk} \delta_{rs} \frac{1}{n} \operatorname{Tr} \hat{A}^{(r,s)} = \frac{1}{n} \delta_{jk} \delta_{rs}.$$

In the case when the matrix defining the representation is unitary this equality coincides with (8.6). Theorem 8.5 is proven. □

Now, we turn to the corollaries of this theorem:

**Corollary 8.3.** *Let* $\rho \colon G \to \mathrm{GL}(n, \mathbb{C})$ *and* $\varphi \colon G \to \mathrm{GL}(m, \mathbb{C})$ *be irreducible representations. Then*

$$(\chi_\rho, \chi_\varphi) = \begin{cases} 1 & \text{if } \rho \text{ and } \varphi \text{ are equivalent,} \\ 0 & \text{otherwise.} \end{cases}$$

To derive this corollary of Theorem 8.5, it suffices to note that $\operatorname{Tr} \rho(g) = \sum_j \rho(g)_{jj}$ and that the characters of equivalent representations coincide, afterwards we are left with applying relations (8.5) and (8.6).

**Corollary 8.4.** *Let*

$$\rho = \rho_1 \oplus \cdots \oplus \rho_k \tag{8.7}$$

*be a decomposition of a representation* $\rho$ *into the direct sum of irreducible representations. Then the number of factors equivalent to an irreducible representation* $\varphi$ *is equal to the Hermitian product of characters*

$$(\chi_\rho, \chi_\varphi)$$

*and hence decomposition* (8.7) *is unique* (*up to equivalence*).

*Proof.* Note that

$$(\chi_\rho, \chi_\varphi) = \sum_j (\chi_{\rho_j}, \chi_\varphi)$$

and the right-hand side is equal to the number of the components $\rho_j$ equivalent to $\varphi$. This number is called the *multiplicity* of $\varphi$ in $\rho$. Corollary 8.4 is proven. □

**Corollary 8.5.** *Representations are equivalent if and only if their characters coincide.*

*Proof.* Indeed, the previous Corollary 8.4 implies that the character determines completely the decomposition of the representation into irreducible representations and each representation of a finite group has such a decomposition (Corollary 8.2). Corollary 8.5 is proven. □

**Corollary 8.6.** *If $\chi$ is the character of $\rho$, then $(\chi, \chi)$ is a positive integer equal to 1 if and only if $\rho$ is irreducible.*

*Proof.* We only need to calculate

$$(\chi, \chi) = \sum_j m_j^2,$$

where $m_j$ is the multiplicity of the irreducible representation $\rho_j$ into the decomposition of $\rho$ in irreducible representations:

$$\rho = m_1\rho_1 \oplus \cdots \oplus m_k\rho_k.$$

Corollary 8.6 is proven. □

**Corollary 8.7.** *Each irreducible representation of a group $G$ is contained in a regular representation with multiplicity equal to its dimension.*

*Proof.* The left and right regular representations are equivalent which follows, in view of Corollary 8.5, from the coincidence of their characters that are equal to

$$\chi_{\mathrm{reg}}(g) = \begin{cases} |G| & \text{if } g(=e) \text{ is the identity of the group } G, \\ 0 & \text{otherwise.} \end{cases}$$

Now, let $\rho\colon G \to \mathrm{GL}(\mathbb{C}, n)$ be an irreducible representation of dimension $n$. Then

$$(\chi_{\mathrm{reg}}, \chi_\rho) = \frac{1}{|G|}\chi_{\mathrm{reg}}(e)\overline{\chi_\rho(e)} = \frac{1}{|G|}|G|n = n.$$

Corollary 8.7 is proven. □

**Corollary 8.8.** *The order of the group $G$ is equal to the sum of the squares of the dimensions of its irreducible representations.*

*Proof.* Let $\rho_1, \dots, \rho_k$ be all irreducible representations of $G$ and let $m_1, \dots, m_k$ be their dimensions. Then, by Corollary 8.7,

$$\chi_{\mathrm{reg}} = \sum_j m_j \chi_{\rho_j},$$

and we conclude that

$$(\chi_{\mathrm{reg}}, \chi_{\mathrm{reg}}) = |G| = \sum_j m_j^2.$$

Corollary 8.8 is proven. □

Denote by $H(G)$ the subspace $L(G)$ constituted by the *central* functions, i.e., functions satisfying the relation

$$f(x) = f(yxy^{-1}) \quad \text{for all } x, y \in G$$

or, equivalently,

$$f(xy) = f(yx) \quad \text{for all } x, y \in G.$$

The simplest examples of central functions are the characters (Lemma 8.1).

**Lemma 8.5.** *Let $f \in H(G)$ and let $\rho$ be a representation of $G$. Denote by $A_f$ the operator*

$$A_f = \sum_{g \in G} \overline{f(g)} \rho(g).$$

*Then*

$$A_f \rho = \rho A_f.$$

The proof of Lemma 8.5 is carried out by simple verification:

$$\begin{aligned}
A_f \rho(h) &= \Big( \sum_{g \in G} \overline{f(g)} \rho(g) \Big) \rho(h) \\
&= \sum_{g \in G} \overline{f(g)} \rho(gh) \\
&= \sum_{u = gh \in G} \overline{f(uh^{-1})} \rho(u) \\
&= \sum_{u \in G} \overline{f(h^{-1}u)} \rho(u) \\
&= \rho(h) \sum_{u \in G} \overline{f(h^{-1}u)} \rho(h^{-1}u) \\
&= \rho(h) \sum_{v = h^{-1}u} \overline{f(v)} \rho(v) \\
&= \rho(h) A_f.
\end{aligned}$$

Lemma 8.5 is proven and we now use it for proving the following fact:

**Theorem 8.6.** *The characters of irreducible representations constitute an orthonormal basis for the space $H(G)$.*

*Proof.* Let $f$ be a central function orthogonal to all characters. For $\overline{f}$ and the left regular representation $\rho_{\text{reg}}$ we construct the operator $A_f = \sum_{g \in G} \overline{f(g)} l_g$ which must by Schur's lemma be the operator of multiplication by the constant $\lambda = \operatorname{Tr} A_f / |G|$.

Calculate the trace of the operator $A_f$:

$$\operatorname{Tr} A_f = \sum_{g \in G} \overline{f(g)} \operatorname{Tr} \rho_{\text{reg}}(g) = \sum_{g \in G} \overline{f(g)} \chi_{\text{reg}}(g) = (f, \chi_{\text{reg}}).$$

From the fact that the character of a regular representation is a linear combination of the characters of irreducible representations, we obtain the equality

$$(f, \chi_{\text{reg}}) = 0,$$

which implies

$$A_f = 0.$$

Since $l_g(\hat{e}) = \hat{g}$, the action of $A_f$ on $\hat{e}$ yields

$$A_f(\hat{e}) = \sum_{g \in G} \overline{f(g)} \hat{g} = 0.$$

The functions $\hat{g}$ are linearly independent in $L(G)$ and therefore

$$f(g) = 0 \quad \text{for every } g \in G.$$

Theorem 8.6 is proven. □

**Corollary 8.9.** *The number of irreducible representations of a group G coincides with the number of conjugacy classes.*

*Proof.* Each central function is given by its values on the conjugacy classes ($x$ and $y$ belong to one class if $x = gyg^{-1}$ for some $g \in G$) and the collection of these values can be arbitrary. Hence, $\dim H(G) = k$, where $k$ is the number of conjugacy classes. It follows from Theorems 8.5 and 8.6 that $\dim H(G)$ is equal to the number of irreducible representations. Corollary 8.9 is proven. □

Now, Corollary 8.1 can be improved as follows:

**Corollary 8.10.** *A finite group is commutative if and only if all irreducible representations of the group are one-dimensional.*

*Proof.* We have

$$|G| = \sum m_j^2;$$

moreover, if all $m_j$ are equal to 1, then $G$ has $|G|$ different irreducible representations. Hence, each conjugacy class contains exactly one element which is equivalent to commutativity of the group. Corollary 8.10 is proven. □

## 8.3 On representations of Lie groups

**Compact groups.** Many assertions proven in §8.2 for finite groups remain valid for compact Lie groups. Moreover, the proofs are similar and are obtained by changing summation over the elements of $G$ with integration, i.e., "summation over a continuous set of parameters".

To define integration on an $n$-dimensional group $G$, we have to prescribe a measure: a smooth $n$-form

$$dg = f(g)dx^1 \wedge \cdots \wedge dx^n,$$

where $\{x^j\}$ are local coordinates on $G$ and $f(g) \geq 0$. The *measure* of a set $U \subset G$ is the value of the integral

$$\mu(U) = \int_U dg.$$

A measure is called *right-invariant* (*left-invariant*) if the form $dg$ is invariant under the right (left) translations:

$$r_h : G \mapsto G,\ g \mapsto gh \quad (l_h : G \to G,\ g \to hg).$$

**Lemma 8.6.** (1) *On every Lie group there is a right-invariant measure given by a smooth $n$-form.*

(2) *On every Lie group there is a unique* (*up to multiplication by a constant*) *right-invariant measure.*

(3) *A right-invariant measure on a compact Lie group is left-invariant.*

*Proof.* Let $c \in \mathbb{R}$ and $C > 0$. Take the form $dg$ at the identity of the group $G$ by the formula

$$dg = c dx^1 \wedge \cdots \wedge dx^n$$

and at $g \in G$ by the formula

$$dg = c r_g^*(dx^1 \wedge \cdots \wedge dx^n).$$

It is obvious that every right-invariant measure has the indicated form and is determined completely by the constant $c$. Measures corresponding to different constants are proportional.

Suppose that the group $G$ is compact and the measure $dg$ is right-invariant. Its left translations $l_h^* dg$ are also right-invariant, since the left and right translations commute. Hence, $l_h^* dg = c_h dg$, where $c_h$ is a constant for each $h \in G$. But

$$\int_G dg = \int_G l_h^* dg = c_h \int_G dg$$

and hence $c_h \equiv 1$. Lemma 8.6 is proven. □

A two-sided invariant measure on the group $G$ is called a *Haar measure*.

**Theorem 8.7.** *Let $G$ be a compact Lie group.*

(1) *Each finite-dimensional representation $\rho\colon G \to \mathrm{GL}(n, \mathbb{C})$ is unitary and decomposes into the direct sum of irreducible representations.*

(2) *Let $\rho\colon G \to \mathrm{GL}(m, \mathbb{C})$ and $\varphi\colon G \to \mathrm{GL}(n, \mathbb{C})$ be nonequivalent irreducible representations given by unitary matrices. Then their entries satisfy the relations*

$$(\rho_{rj}, \varphi_{sk}) = 0, \quad (\rho_{rj}, \rho_{sk}) = \frac{1}{n}\delta_{jk}\delta_{rs}, \tag{8.8}$$

*where*

$$(f_1, f_2) = \frac{1}{\int_G dg} \int_G f_1(g)\overline{f_2(g)}\, dg. \tag{8.9}$$

Assertion (1) is proven by analogy with Theorem 8.3 with the Hermitian product (8.2) replaced by

$$(v, w) = \int_G (\rho(g)v, \rho(g)w)\, dg.$$

Assertion (2) is proven by analogy with its analogs for finite groups (Theorem 8.5) with (8.3) replaced with (8.9). Here the volume $\int_G dg$ of the group is the analog of the order $|G|$ of the group $G$.

The proof of the analog of Theorem 8.4 is essentially more complicated; we give the corresponding assertion without proof.

**Theorem 8.8.** *Each compact Lie group has a faithful finite-dimensional representation and consequently is linear.*

Note that

- *there are noncompact Lie groups which have no faithful finite-dimensional representations.*[1]

Denote by $L_2(G)$ the space of complex-valued square integrable functions on $G$ with the inner product

$$(\psi, \varphi) = \frac{1}{\int_G dg} \int_G \psi(g)\overline{\varphi(g)}\, dg.$$

This space contains all smooth functions and, in particular, the entries $\rho_{jk}$ of the operators defining representations. Relations (8.8) demonstrate that these matrix entries constitute an orthogonal system.

In the case of a finite group this system contains exactly $|G|$ different functions (Corollary 8.8) and, since $\dim L(G) = |G|$, we obtain

**Theorem 8.9.** *The entries irreducible unitary representations of a finite group constitute an orthogonal basis for $L(G)$.*

[1]Such an example is given by the group $\widetilde{\mathrm{SL}(2, \mathbb{R})}$, the universal covering group of $\mathrm{SL}(2, \mathbb{R})$ (see, for instance, [2], II, or [4]).

If a group $G$ is not finite, then $L_2(G)$ is an infinite-dimensional Hilbert space (with the Hilbert product (8.9)). By analogy with the case of finite groups, we might expect that the entries generate $L_2(G)$ in a sense. This is really true. To improve this assertion, below we give Theorem 8.11 (without proof).

But first we need to introduce necessary notions and indicate the following fact (again without proof).

**Theorem 8.10.** *Irreducible representations of a compact Lie group are finite-dimensional and their number is countable.*

**Examples.** (1) SU(2). For each $n \geq 0$ define $V_n$ as the vector space constituted by all complex homogeneous polynomials of degree $n$ in two variables $z_1, z_2 \in \mathbb{C}$. It is obvious that $\dim V_n = n + 1$ and the polynomials of the form $z_1^k z_2^{n-k}$, where $k = 0, \ldots, n$, constitute a basis for $V_n$. Define the representation $\rho_n \colon \mathrm{SU}(2) \to \mathrm{GL}(V_n) = \mathrm{GL}(n+1, \mathbb{C})$ by the formula

$$\rho_n(g)(f)(z_1, z_2) = f(a_{11}z_1 + a_{21}z_2, a_{12}z_1 + a_{22}z_2),$$

which can be written as

$$\rho_n(g)(f)(z) = f(g^\top z),$$

where $g = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathrm{SU}(2)$, $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$. It is easy to verify that this formula determines a representation, since

$$\rho(g)\rho(h)(f)(z) = \rho(h)(f)(g^\top z) = f(h^\top g^\top z) = f((gh)^\top z) = \rho(gh)(f)(z).$$

For $n = 0$ we obtain the trivial representation: $\rho(g) \equiv 1$; and for $n = 1$, the standard embedding $\mathrm{SU}(2) \to \mathrm{GL}(2, \mathbb{C})$.

- *The representations $\rho_n$, $n \geq 0$, are precisely all irreducible complex representations of* SU(2).

Since the dimensions of the representation spaces are different, these representations are pairwise nonequivalent.

(2) SO(3). We know that $\mathrm{SU}(2)/\{\pm 1\} = \mathrm{SO}(3)$ (see §7.3). For even values $n$ the representations $\rho_n$ possess the property $\rho(1) = \rho(-1)$. It means that they descend through SO(3):

$$\rho_{2k} \colon \mathrm{SU}(2) \to \mathrm{SU}(2)/\{\pm 1\} = \mathrm{SO}(3) \to \mathrm{GL}(2k+1, \mathbb{C}).$$

- *The representations $\rho_{2k}$, $k \geq 0$, are precisely all irreducible complex representations of* SO(3).
- *The representations $\rho_{2k+1}$, $k \geq 0$, give rise to two-valued* (*spinor*) *representations of* SO(3).

Let $\{\rho^l\}_{l=1}^{\infty}$ be a complete system of irreducible representations of a group $G$, and let $\rho^l_{jk}$ be the matrix entries of the unitary operators which determine these representations. The *Fourier series* of a function $f \in L_2(G)$ is the series

$$f(g) \sim \sum_{l,j,k} c^l_{jk} \rho^l(g)_{jk},$$

where the *Fourier coefficients* $c^l_{jk}$ are determined by the formula

$$c^l_{jk} = \frac{1}{\sqrt{n_l}}(f, \rho^l_{jk})$$

and $n_l$ is the dimension of the representation $\rho^l$.

**Theorem 8.11.** (1) (Peter–Weyl theorem) *A continuous function on a compact group $G$ can be approximated by linear combinations of the functions $\rho^l_{jk}$ with arbitrary accuracy: for every $\varepsilon > 0$, there is a finite linear combination $\sum a^l_{jk}\rho^l_{jk}$ such that*

$$\left| f(g) - \sum a^l_{jk} \rho^l(g)_{jk} \right| < \varepsilon$$

*for all $g \in G$;*

(2) *The Fourier series of a function $f \in L_2(G)$ converges to $f$ in the norm of $L_2(G)$:*

$$\lim_{N\to\infty} \Big( f - \sum_{l \le N} c^l_{jk} \rho^l(g)_{jk}, f - \sum_{l \le N} c^l_{jk} \rho^l(g)_{jk} \Big) = 0;$$

*consequently, the system of functions*

$$\frac{1}{\sqrt{n_l}} \rho^l_{jk}$$

*is a complete orthonormal system in $L_2(G)$.*

In the case $G = \mathrm{U}(1) = \mathbb{R}/2\pi\mathbb{Z}$ we obtain the usual Fourier series of periodic functions: the entries of the irreducible representations are the characters $e^{inx}$.

For noncommutative Lie groups the functions $\rho^l_{jk}$ turn out to be special functions and appear in mathematical physics as well as representation theory itself.

Note that the entries $\rho^l_{jk}$ satisfy many special relations resulting from the simple identity

$$\rho^l(gh) = \rho^l(g)\rho^l(h), \quad g, h \in G.$$

In the case when $\rho\colon \mathrm{SO}(2) \to \mathrm{GL}(2)$ is the standard embedding, we obtain the trigonometric summation theorems:

$$\rho(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

and the equality $\rho(\alpha + \beta) = \rho(\alpha)\rho(\beta)$ is equivalent to the identities

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta,$$
$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta.$$

For multidimensional groups the summation theorems for special functions (matrix entries) look essentially more complicated and follow from representation theory.

**Noncompact groups.** As a rule, unitary representations of noncompact Lie groups are infinite-dimensional. Their study was connected with development of analysis on these groups and construction of the analog of the Fourier transform.

Suppose that a group is unimodular; i.e., it has a two-sided invariant measure $d\mu$. Denote by $L_1(G)$ and $L_2(G)$ the spaces of integrable and square integrable (with respect to the measure $d\mu$) complex-valued functions on $G$.

The Fourier transform is defined as follows. We are given a space $\widehat{G}$ whose points parameterize all irreducible unitary representations of the group $G$ and a measure $d\hat{\mu}$ on it.

The *Fourier transform* of a function $f \in L_1(G)$ is defined by the equality

$$\hat{f}(\lambda) = T_\lambda(f) = \int f(g)T_\lambda(g)\,d\mu,$$

where $\lambda \in \widehat{G}$ and $T_\lambda$ is the unitary representation corresponding to $\lambda$. As we see, in the case when the unitary representation is infinite-dimensional, the Fourier transform of a function on a group is the integral of unitary operators in the Hilbert space $H_\lambda$ (the space of the representation $T_\lambda$).

If $f \in L_1(G) \cap L_2(G)$, then for almost all $\lambda$ (with respect to the measure $d\hat{\mu}$) the operators $T_\lambda(f)\colon H_\lambda \to H_\lambda$ are Hilbert–Schmidt operators; moreover, the *Plancherel formula* holds:

$$\int_G |f(g)|^2 d\mu = \int_{\widehat{G}} \operatorname{Tr}(T_\lambda(f)T_\lambda^*(f))\,d\hat{\mu}.$$

This correspondence extends to an isometry $L_2(G) \approx L_2(\widehat{G})$.

For functions in an everywhere dense subspace $\mathcal{A} \subset L_1(G) \cap L_2(G)$ we can define the inverse Fourier transform:

$$f(g) = \int_{\widehat{G}} \operatorname{Tr}(T_\lambda(f)T_\lambda^*(g))\,d\hat{\mu}.$$

At present, the Fourier transform (harmonic analysis) is constructed for a wide class of noncompact Lie groups, including, actually, all cases interesting for applications.

To illustrate the basic concepts on harmonic analysis on noncompact Lie groups we expose a couple of examples.

**Examples.** (1) $\mathbb{R}^n$ (the classical Fourier transform). Since this group is commutative, all irreducible representations are one-dimensional and coincide with the characters. The unitary irreducible representations have the form

$$T_\lambda(x) = e^{i(\lambda,x)},$$

where $\lambda = \widehat{\mathbb{R}}^n = \mathbb{R}^n$ and $(\lambda, x) = \sum_{i=1}^n \lambda_i x_i$. The Fourier transform takes the form

$$\hat{f}(\lambda) = \int_{\mathbb{R}^n} f(x) e^{i(\lambda,x)} dx_1 \dots dx_n,$$

where $d\mu = dx_1 \dots dx_n$, and the inverse transform takes the form

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\lambda) e^{-i(\lambda,x)} d\lambda_1 \dots d\lambda_n,$$

where $d\hat{\mu} = (2\pi)^{-n} d\lambda_1 \dots d\lambda_n$.

(2) $\mathrm{SL}(2,\mathbb{R})$. We denote the elements of this group by

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad ad - bc = 1,$$

or

$$g = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

The formula

$$d\mu = dt \wedge dx \wedge d\theta$$

determines a two-sided invariant measure on $G = \mathrm{SL}(2,\mathbb{R})$.

Consider the irreducible representations $T_\lambda^+$, $T_\lambda^-$, $\lambda \in \mathbb{R}, \lambda \geq 0$, and $D_n^+$, $D_n^-$, $n = 1, 2, 3, \dots$. All these representations are irreducible, unitary, and pairwise nonequivalent. They do not exhaust the whole set $\widehat{G}$ but are sufficient for construction of the Fourier transform.

First of all we describe these representations:

(a) $T_\lambda^+$ and $T_\lambda^-$ act on $L_2(\mathbb{R})$ by the following formulas:

$$[T_\lambda^+(g)f](x) = |bx + d|^{i\lambda - 1} f\left(\frac{ax + c}{bx + d}\right),$$

$$[T_\lambda^-(g)f](x) = \mathrm{sgn}(bx + d)|bx + d|^{i\lambda - 1} f\left(\frac{ax + c}{bx + d}\right),$$

where $f \in L_2(\mathbb{R})$.

(b) Denote by $H_n^\pm$ the Hilbert space of analytic functions in the half-plane $\pm \operatorname{Im} z > 0$ for which the integrals $(\psi, \psi)$, $\psi \in H_n^\pm$, are finite, where

$$(\psi, \varphi) = \int_{\pm \operatorname{Im} z > 0} y^{n-1} \psi(x, y) \overline{\varphi(x, y)} dx dy, \quad z = x + iy.$$

The last formula defines the inner product in $H_m^{\pm}$. The representations $D_n^{\pm}$ are given by the formulas

$$[D_n^+(g)\psi](z) = (bz+d)^{-n-1}\psi\left(\frac{az+c}{bz+d}\right),$$
$$[D_n^-(g)\psi](z) = (bz+d)^{-n-1}\psi\left(\frac{az+c}{bz+d}\right).$$

The Fourier transform has the form

$$\hat{f}(\lambda,\pm) = \int f(g)T_\lambda^{\pm}(g)\,d\mu, \quad \hat{f}(n,\pm) = \int f(g)D_n^{\pm}(g)\,d\mu,$$

and the Plancherel formula (for the group $\mathrm{SL}(2,\mathbb{R})$) becomes

$$\begin{aligned}\int |f(g)|^2 d\mu = {} & \frac{1}{8\pi^2}\int_0^\infty \mathrm{Tr}(T_\lambda^+(f)T_\lambda^+(f)^*)\lambda\tanh\frac{\pi\lambda}{2}\,d\lambda \\ & + \frac{1}{8\pi^2}\int_0^\infty \mathrm{Tr}(T_\lambda^-(f)T_\lambda^-(f)^*)\lambda\coth\frac{\pi\lambda}{2}\,d\lambda \\ & + \sum_{n=1}^{\infty} 2n[\mathrm{Tr}(D_n^+(g)D_n^+(g)^*) + \mathrm{Tr}(D_n^-(g)D_n^-(g)^*)].\end{aligned}$$

For the subspace $\mathcal{A} \subset L_1(G)\cap L_2(G)$ constituted by all smooth compactly-supported functions, the inverse Fourier transform is given by the formula

$$\begin{aligned}f(g) = {} & \frac{1}{4\pi^2}\int_0^\infty \mathrm{Tr}(T_\lambda^+(f)T_\lambda^+(g)^*)\lambda\tanh\frac{\pi\lambda}{2}\,d\lambda \\ & + \frac{1}{4\pi^2}\int_0^\infty \mathrm{Tr}(T_\lambda^-(f)T_\lambda^-(g)^*)\lambda\coth\frac{\pi\lambda}{2}\,d\lambda \\ & + \sum_{n=1}^{\infty} 2n[\mathrm{Tr}(D_n^+(f)D_n^+(g)^*) + \mathrm{Tr}(D_n^-(f)D_n^-(g)^*)],\end{aligned}$$

where $f \in \mathcal{A}$.

# 9

# Elements of Poisson and symplectic geometry

## 9.1 The Poisson bracket and Hamilton's equations

In §2.7, considering the equation for geodesics on a surface we introduced the Euler–Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i}, \quad i = 1, \dots, n, \tag{9.1}$$

which describe the extremals of the functional

$$S(\gamma) = \int_\gamma L(x, \dot{x})dt$$

on the space of curves on the manifold $M^n$.

Consider the function

$$L(x, \dot{x}) = \frac{m\dot{x}^2}{2} - U(x),$$

where $m$ is the mass of a point body which moves in the $n$-dimensional Euclidean space $\mathbb{R}^n$. The point $x \in \mathbb{R}^n$ determines the position of the body. The function $U(q)$ is called the potential (or the *potential energy*). The Euler–Lagrange equations take the form

$$m\ddot{x}^i = -\frac{\partial U}{\partial x^i}, \quad i = 1, \dots, n.$$

Introducing the notation for the momentum

$$p_i = m\dot{x}^i, \tag{9.2}$$

we can write the above equations in the form of *Newton's equations*

$$\dot{p}_i = -\frac{\partial U}{\partial x^i}, \quad i = 1, \dots, n. \tag{9.3}$$

In terms of the variables $x$ and $p$, equations (9.2) and (9.3) take the form of *Hamilton's equations*

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \quad i = 1, \dots, n, \tag{9.4}$$

where the function

$$H(x, p) = \frac{p^2}{2m} + U(x)$$

is called the *Hamiltonian function* or the *Hamiltonian*.

To give the most general definition of Hamilton's equations, rewrite (9.4) in the form

$$\dot{x}^i = \{x^i, H\}, \quad \dot{p}_i = \{p_i, H\}, \quad i = 1, \dots, n, \tag{9.5}$$

where

$$\{f, g\} = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial x^i} \right). \tag{9.6}$$

A *Poisson bracket* on a smooth manifold $M$ is a bilinear product

$$f, g \to \{f, g\}$$

on the space $C^\infty(M)$ of all smooth functions on $M$ such that the space $C^\infty(M)$ is a Lie algebra with respect to this form; i.e., this product is antisymmetric:

$$\{f, g\} = -\{g, f\},$$

and satisfies the Jacobi identity

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0;$$

moreover, it agrees with multiplication of functions in the sense that the *Leibniz identity* holds:

$$\{fg, h\} = f\{g, h\} + g\{f, h\}. \tag{9.7}$$

In this case we say that a Poisson structure is given on $M$ and $M$ itself is called a *Poisson manifold*.

The following lemma is proven by simple verification:

**Lemma 9.1.** *Formula* (*9.6*) *defines a Poisson bracket on* $C^\infty(M)$, *where* $M$ *is the Euclidean space* $\mathbb{R}^{2n}$ *with coordinates* $x^1, \dots, x^n, p_1, \dots, p_n$.

Every smooth function $H$ on a Poisson manifold $M$ determines uniquely the dynamical system by formula

$$\frac{df}{dt} = \{f, H\} \tag{9.8}$$

which describes the change of the value of an arbitrary smooth function $f$ along the trajectories of this system. The so-obtained system is called a *Hamiltonian* system, equations (9.8) are called *Hamilton's equations*, and the function $H$ is called the *Hamiltonian function* or the *Hamiltonian* of the system.[1]

The advantage of representation of the equations of mechanics in the form of Hamilton's equations remained unclear until the discovery of quantum mechanics, when it

[1] Here and in the sequel we suppose that the Hamiltonian does not depend explicitly on time; i.e., it is a function of the point of the Poisson manifold only.

turned out that Hamilton's equations provide a convenient method for quantization of classical systems.[2]

In quantum mechanics, observables (in particular, coordinates and momenta) are associated with operators. Consider one simple example. Suppose that we have a system with $n$ degrees of freedom for which there are given the coordinates $x \in \mathbb{R}^n$ of position of the system and momenta $p \in \mathbb{R}^n$. The state space of a quantum system (the space of *wave functions*) is a Hilbert space $\mathcal{H}$. Take $\mathcal{H}$ to be the space $L_2(\mathbb{R}^n)$ constituted by all square summable complex-valued functions of the variables $x^1, \dots, x^n$ with the inner product

$$(\psi, \varphi) = \int_{\mathbb{R}^n} \psi \bar{\varphi}\, dx,$$

where $dx = dx^1 \dots dx^n$.

To the coordinate $x^i$ there corresponds the operator of multiplication by the function $x^i$:

$$X_i(f)(x) = x^i f(x),$$

and to the momentum $p_i$, the differential operator $P_i$:

$$P_i(f)(x) = -i\hbar \frac{\partial f(x)}{\partial x^i},$$

where $\hbar = \frac{h}{2\pi}$ and $h$ is the Planck constant. The domains of these operators are everywhere dense in $H$. Physically important values are the mean values of the observables, namely the mean value of an observable $A$ in a state $\psi$ is equal to

$$\langle A|\psi\rangle = (A\psi, \psi) = \int_{\mathbb{R}^n} (A\psi)\overline{\psi}\, dx.$$

Let $H(x, p)$ be the Hamiltonian of a classical system. Suppose, for simplicity, that it is the sum of analytic functions of $x$ and $q$:

$$H(x, p) = F(x) + G(q),$$

for example, it has the form

$$H = \frac{p^2}{2m} + U(q),$$

where the function $\mathrm{U}(x)$ expands in a converging series in the powers of $x$. In quantum mechanics this Hamiltonian is associated with the *Hamiltonian operator*

$$H(X, P) = F(P) + G(X)$$

which is obtained from $H(x, p)$ by the formal change of the variables $x^i$ and $p_i$ in the series for $F$ and $G$ with the operators $X_i$ and $P_i$. Note that the coordinate operators as well as the momentum operators commute; therefore, the value of substitution is independent of the order of these operators. In the general case, inserting $X_j$ and $P_k$ in the analytic expression for $H(x, p)$, we have to prescribe additionally the order of the operators $X_j$ and $P_j$, which do not commute.

[2]For a detailed exposition of quantization of classical systems we refer the reader to P. A. M. Dirac, *The Principles of Quantum Mechanics*. Fourth edition, Clarendon Press, Oxford 1958.

**Example** (Harmonic oscillator.) Let $n = 1$ and consider the Hamiltonian function

$$H(x, p) = \frac{p^2}{2m} + \frac{\omega^2 x^2}{2}.$$

The Hamiltonian operator is equal to

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{\omega^2 x^2}{2}.$$

The time evolution of the quantum system is described by the (*nonstationary*) *Schrödinger equation*

$$i\hbar\frac{d\psi}{dt} = H\psi.$$

A translation along the trajectories of this system by a time $t$ is a transformation of the Hilbert state space:

$$T_t : \mathcal{H} \to \mathcal{H}.$$

For some physical reasons, the operator $H$ must be self-adjoint which implies that the transformation $T_t$ is unitary (preserves the inner product) for all $t$:

$$(T_t\psi, T_t\varphi) = (\psi, \varphi).$$

Another description of the evolution is given by the *Heisenberg equation*

$$\frac{dA}{dt} = \{A, H\}_q$$

where

$$A(t) = T_t^{-1} A_0 T_t = T_t^* A_0 T_t,$$

$A_0$ is the operator of the observable (at $t = 0$), $\{\cdot, \cdot\}_q$ is the quantum Poisson bracket which is equal to

$$\{A, B\}_q = \frac{1}{i\hbar}(AB - BA)$$

for every pair of operators $A$ and $B$. Moreover, note that

$$\{X_j, P_k\}_q = \{x_j, p_k\} = \delta_{jk}.$$

The first approach (the Schrödinger equation) assumes that the states change with time, and the observables are preserved; while the second approach (the Heisenberg equation) assumes that the observables change, and the states are preserved. However, the mean values of the observables in both cases coincide:

$$\langle A(0)\psi(t)|\psi(t)\rangle = (A(0)T_t\psi(0), T_t\psi_t) = (T_t^* A(0)T_t\psi(0), \psi(0)) = \langle A(t)\psi(0)|\psi(0)\rangle.$$

Consequently, the physical pictures given by two approaches are equivalent.

We now return to Poisson manifolds.

**Lemma 9.2.** *Let $M$ be a Poisson manifold and let $x^1, \dots, x^n$ be local coordinates in a domain $U$ of $M$. Suppose that $f$ and $g$ are polynomials in the variables $x^1, \dots, x^n$. Then*

$$\{f, g\} = \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^k} \{x^j, x^k\} = \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^k} h^{jk}(x). \tag{9.9}$$

*Proof.* Let $f = (x^1)^{k_1}(x^2)^{k_2} \dots (x^n)^{k_n} = (x^1)^{k_1} h(x^2, \dots, x^n)$. From the Leibniz identity (9.7) we obtain

$$\{f, g\} = k_1 (x^1)^{k_1 - 1} h\{x^1, g\} + (x^1)^{k_1}\{h, g\} = \frac{\partial f}{\partial x^1}\{x^1, g\} + (x^1)^{k_1}\{h, g\}.$$

Repeating these arguments for $h = (x^2)^{k_2} \dots (x^n)^{k_n}$, eventually, we obtain the identity

$$\{f, g\} = \frac{\partial f}{\partial x^j}\{x^j, g\},$$

which extends by linearity to all polynomials $f$. If $g$ is another polynomial, then the right-hand side of the last formula takes the form

$$\frac{\partial f}{\partial x^j}\{x^j, g\} = \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^k}\{x^j, x^k\}$$

and we obtain identity (9.9). The lemma is proven. □

It is obvious that identity (9.9) extends by linearity to functions $f$ and $g$ representable by converging series in the powers of $x^1, \dots, x^n$. Henceforth we consider only the Poisson bracket of the form (9.9).

**Theorem 9.1.** *Let $M$ be an $n$-dimensional manifold and let $x^1, \dots, x^n$ be local coordinates in a domain of $M$.*

(1) *If a Poisson bracket is given on $M$, then the values $h^{jk}(x) = \{x^j, x^k\}$ constitute an antisymmetric tensor of type* $(2, 0)$ *and the Poisson bracket in arbitrary coordinates $y^1, \dots, y^n$ has the form*

$$\{f, g\} = \frac{\partial f}{\partial y^j} \frac{\partial g}{\partial y^k}\{y^j, y^k\}.$$

(2) *Given an antisymmetric tensor $h^{jk}(x)$, formula* (9.9) *defines an antisymmetric bilinear form on the space of functions which satisfies the Leibniz identity.*

*This form satisfies the Jacobi identity* (*and hence is a Poisson bracket*) *if and only if*

$$h^{il} \frac{\partial h^{jk}}{\partial x^l} + h^{jl} \frac{\partial h^{ki}}{\partial x^l} + h^{kl} \frac{\partial h^{ij}}{\partial x^l} = 0, \quad j, k, l = 1, \dots, n. \tag{9.10}$$

*Proof.* (1) Apply formula (9.9) to the coordinate functions $y^k = y^k(x^1, \dots, x^n)$, $k = 1, \dots, n$:

$$\tilde{h}^{lm} = \{y^l, y^m\} = \frac{\partial y^l}{\partial x^j} \frac{\partial y^m}{\partial x^k}\{x^j, x^k\} = \frac{\partial y^l}{\partial x^j} \frac{\partial y^m}{\partial x^k} h^{jk}.$$

Consequently, the values $h^{jk}$ determine a tensor of type $(2,0)$ which is obviously antisymmetric: $h^{jk} = -h^{kj}$. Inserting the last formula in (9.9), from the chain rule we derive

$$\begin{aligned}\{f,g\} &= \frac{\partial f}{\partial x^j}\frac{\partial g}{\partial x^k}\{x^j,x^k\}\\ &= \left(\frac{\partial f}{\partial y^l}\frac{\partial y^l}{\partial x^j}\right)\left(\frac{\partial g}{\partial y^m}\frac{\partial y^m}{\partial x^k}\right)\{x^j,x^k\}\\ &= \frac{\partial f}{\partial y^l}\frac{\partial g}{\partial y^m}\left(\frac{\partial y^l}{\partial x^j}\frac{\partial y^m}{\partial x^k}\{x^j,x^k\}\right)\\ &= \frac{\partial f}{\partial y^l}\frac{\partial g}{\partial y^m}\{y^l,y^m\}.\end{aligned}$$

(2) If the tensor $h^{jk}$ is antisymmetric then, obviously, formula (9.9) determines an antisymmetric bilinear form which satisfies the Leibniz identity. Straightforward calculations show that the following identity holds:

$$\begin{aligned}&\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\}\\ &\qquad = \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j}\frac{\partial h}{\partial x^k}(\{x^i,\{x^j,x^k\}\} + \{x^j,\{x^k,x^i\}\} + \{x^k,\{x^i,x^j\}\}).\end{aligned}$$

Therefore, it suffices to verify the Jacobi identity for the coordinate functions $x^i$, $i = 1,\dots,n$. Making simple calculations again, we show that

$$\{x^i,\{x^j,x^k\}\} + \{x^j,\{x^k,x^i\}\} + \{x^k,\{x^i,x^j\}\} = h^{il}\frac{\partial h^{jk}}{\partial x^l} + h^{jl}\frac{\partial h^{ki}}{\partial x^l} + h^{kl}\frac{\partial h^{ij}}{\partial x^l},$$

where $h^{ij} = \{x^i,x^j\}, i,j = 1,\dots,n$. The lemma is proven. □

**Corollary 9.1.** *For every antisymmetric tensor $h^{ij}$ with constant coefficients* (*in coordinates $x^1,\dots,x^n$ in a domain $U \subset \mathbb{R}^n$*) *formula* (9.9) *defines a Poisson bracket on $C^\infty(U)$.*

The Hamiltonian system (9.8) generated by the Hamiltonian function $H$ is a dynamical system on a Poisson manifold $M$ and has a very simple coordinate form:

$$\dot{x} = v_H(x), \quad v_H^i = h^{ik}\frac{\partial H}{\partial x^k}, \quad i = 1,\dots,n.$$

Since $h^{ik}$ is a tensor of type $(2,0)$ and the gradient grad $f$ is a covector field, the formula

$$f \to v_f, \quad v_f^i = h^{ik}\frac{\partial f}{\partial x^k}, \quad i = 1,\dots,n, \tag{9.11}$$

relates to each smooth function $f$ on $M$ the vector field $v_f$. A vector field of the form $v_f$ is called a *Hamiltonian field.*

**Corollary 9.2.** *Given two functions $f$ and $g$ on a Poisson manifold $M$, the commutator $[v_f, v_g]$ of vector fields $v_f$ and $v_g$ is equal to*

$$[v_f, v_g] = -v_{\{f,g\}}.$$

*Proof.* Write down explicitly the commutator of two vector fields:

$$\begin{aligned}
[v_f, v_g]^i &= h^{jk}\frac{\partial f}{\partial x^k}\frac{\partial}{\partial x^j}\left(h^{il}\frac{\partial g}{\partial x^l}\right) - h^{jl}\frac{\partial g}{\partial x^l}\frac{\partial}{\partial x^j}\left(h^{ik}\frac{\partial f}{\partial x^k}\right) \\
&= \frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^l}\left(h^{jk}\frac{\partial h^{il}}{\partial x^j} - h^{jl}\frac{\partial h^{ik}}{\partial x^j}\right) + h^{jk}\frac{\partial f}{\partial x^k}h^{il}\frac{\partial^2 g}{\partial x^j \partial x^l} \\
&\quad - h^{jl}\frac{\partial g}{\partial x^l}h^{ik}\frac{\partial^2 f}{\partial x^j \partial x^k}.
\end{aligned}$$

Since all indices in the third term except $i$ imply summation, we can make the following permutation of indices without changing the result: $j \to k$, $k \to l$, and $l \to j$. From antisymmetry of the tensor $h^{jk}$ and the Jacobi identity (9.10) we obtain

$$h^{jk}\frac{\partial h^{il}}{\partial x^j} - h^{jl}\frac{\partial h^{ik}}{\partial x^j} = h^{jk}\frac{\partial h^{il}}{\partial x^j} + h^{jl}\frac{\partial h^{ki}}{\partial x^j} = -h^{ji}\frac{\partial h^{lk}}{\partial x^j}.$$

Making the substitution, from the last identity we derive

$$\begin{aligned}
[v_f, v_g]^i &= -h^{ji}\frac{\partial h^{lk}}{\partial x^j}\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^l} + h^{jk}\frac{\partial f}{\partial x^k}h^{il}\frac{\partial^2 g}{\partial x^j \partial x^l} - h^{kj}\frac{\partial g}{\partial x^j}h^{il}\frac{\partial^2 f}{\partial x^k \partial x^l} \\
&= -h^{ji}\frac{\partial h^{lk}}{\partial x^j}\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^l} + h^{jk}h^{il}\frac{\partial}{\partial x^l}\left(\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^j}\right) \\
&= -h^{li}\frac{\partial h^{jk}}{\partial x^l}\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^l} + h^{jk}h^{il}\frac{\partial}{\partial x^l}\left(\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^j}\right)
\end{aligned}$$

(here we interchanged the indices $j$ and $l$ in the first term)

$$\begin{aligned}
&= h^{il}\frac{\partial}{\partial x^l}\left(h^{jk}\frac{\partial f}{\partial x^k}\frac{\partial g}{\partial x^k}\right) \\
&= h^{il}\frac{\partial\{g, f\}}{\partial x^l} = -h^{il}\frac{\partial\{f, g\}}{\partial x^l}.
\end{aligned}$$

The corollary is proven. □

At each point $x \in M$ the Poisson bracket is determined by the matrix $h^{jk}$ and its rank is called the *rank of the Poisson bracket* at $x$.

Suppose that the Poisson bracket has full rank at each point. Then the matrix $h^{jk}$ at each point is invertible: there is an antisymmetric tensor $h_{jk}$ such that

$$h_{jk}h^{kl} = \delta_j^l$$

everywhere. In this case we say that $M$ is a *symplectic manifold*, and the 2-form

$$\omega = \sum_{j<k} h_{jk} dx^j \wedge dx^k$$

is called the *symplectic form*.

Note that the form $\omega$ is nondegenerate by definition: the matrix $h_{jk}$ has full rank at each point or, which is the same, for each tangent vector $\xi \neq 0$ attached to an arbitrary point, there is another tangent vector $\eta$ attached to the same point such that

$$\omega(\xi, \eta) = h_{jk}\xi^j \eta^k \neq 0.$$

It is well known from linear algebra that an antisymmetric nondegenerate 2-form in a vector space always reduces, in appropriate coordinates $p$ and $q$, to the form

$$\sum_{i=1}^{k} dp_i \wedge dq^i, \qquad n = 2k.$$

Thus, we have the following lemma:

**Lemma 9.3.** (1) *The Poisson bracket at each point has an even rank.*
(2) *A symplectic manifold is even-dimensional.*

Consider the question when a nondegenerate 2-form determines a Poisson bracket.

**Theorem 9.2.** *If* $\omega = \sum_{j<k} h_{jk} dx^j \wedge dx^k$ *is a nondegenerate* 2*-form, then the formula*

$$\{f, g\} = h^{jk} \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial x^k}, \quad h_{jk}h^{kl} = h^{lk}h_{kj} = \delta^l_j,$$

*determines a Poisson bracket if and only if this form is closed:*

$$d\omega = 0.$$

*Proof.* An antisymmetric tensor $h^{jk}$ determines a Poisson bracket if and only if

$$h^{il}\frac{\partial h^{jk}}{\partial x^l} + h^{jl}\frac{\partial h^{ki}}{\partial x^l} + h^{kl}\frac{\partial h^{ij}}{\partial x^l} = 0$$

for all $i$, $j$, and $k$ (Theorem 9.1). Multiply the above relation by $h_{ri}h_{js}h_{kt}$ and contract terms with coinciding upper and lower indices, i.e.,sum over $i$, $j$, and $k$:

$$h_{ri}h_{js}h_{kt}\left(h^{il}\frac{\partial h^{jk}}{\partial x^l} + h^{jl}\frac{\partial h^{ki}}{\partial x^l} + h^{kl}\frac{\partial h^{ij}}{\partial x^l}\right) = 0. \tag{9.12}$$

Consider, for example, the first summand on the left-hand side:

$$\begin{aligned} h_{ri}h_{js}h_{kt}h^{il}\frac{\partial h^{jk}}{\partial x^l} &= \delta^l_r h_{js}h_{kt}\frac{\partial h^{jk}}{\partial x^l} \\ &= h_{js}h_{kt}\frac{\partial h^{jk}}{\partial x^r} \\ &= h_{js}\left(\frac{\partial h^{jk}h_{kt}}{\partial x^r} - h^{jk}\frac{\partial h_{kt}}{\partial x^r}\right) \\ &= h_{js}\left(\frac{\partial \delta^j_t}{\partial x^r} - h^{jk}\frac{\partial h_{kt}}{\partial x^r}\right) \\ &= -h_{js}h^{jk}\frac{\partial h_{kt}}{\partial x^r} = \delta^k_s\frac{\partial h_{kt}}{\partial x^r} = \frac{\partial h_{st}}{\partial x^r}. \end{aligned}$$

Transforming similarly the other summands in (9.12), we find that (9.12) is equivalent to the fact that the relations

$$a_{rst} = \frac{\partial h_{st}}{\partial x^r} + \frac{\partial h_{tr}}{\partial x^s} + \frac{\partial h_{rs}}{\partial x^t} = 0$$

hold for all $r$, $s$, and $t$; but the last is equivalent to the fact that the form $\omega$ is closed, since

$$d\omega = \sum_{r<s<t} a_{rst}dx^r \wedge dx^s \wedge dx^t.$$

The theorem is proven. □

Thereby we arrive at the classical definition of a symplectic manifold:

- a manifold $M$ is called *symplectic* if a nondegenerate closed 2-form $\omega$ is given on $M$; this form is called the *symplectic form* and determines the symplectic structure on $M$.

The Poisson bracket on a symplectic manifold can also be written in the form

$$\{f, g\} = -\omega(v_f, v_g) = \omega(v_g, v_f).$$

Indeed, expand the right-hand side of the last formula:

$$\omega(v_g, v_f) = h_{jl}\Big(h^{jk}\frac{\partial g}{\partial x^k}\Big)\Big(h^{li}\frac{\partial f}{\partial x^i}\Big) = \delta^i_j h^{jk}\frac{\partial g}{\partial x^k}\frac{\partial f}{\partial x^i} = h^{ik}\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^k} = \{f, g\}.$$

For completeness of exposition we give another important definition:

- an $n$-dimensional submanifold $L$ of a $2n$-dimensional symplectic manifold $M$ is called *Lagrangian* if the restriction of the symplectic form $\omega$ onto $L$ vanishes: $\omega|_L = 0$.

We will not discuss Lagrangian submanifolds. However we remark that they play an important role, in particular, in the theory of partial differential equations (see, for instance, [4]).

## 9.2 Lagrangian formalism

**1 Lagrangian formalism.** Suppose that a smooth function $L(x, \dot{x})$ is given on the tangent bundle $TM^n$ of an $n$-dimensional manifold $M^n$, where $x \in M^n$ and $\dot{x}$ is a tangent vector to $M^n$ at $x$. This function determines a Lagrangian system as follows.

The points of the space $M^n$ determine different configurations (positions) of this system; therefore, the space $M^n$ is called the *configuration space*. The trajectories of the system are found from the Maupertuis–Fermat variational principle (*the principle of least action*):

- *each part $\gamma_0(t)$, $t \in [a, b]$, of a trajectory $\gamma_0$ is an extremal of the functional*

$$S(\gamma) = \int_\gamma L(x, \dot{x})\, dt$$

*defined on the space all piecewise smooth curves $\gamma(t)$, $t \in [a, b]$, with the same endpoints as $\gamma_0(t)$: $\gamma(a) = x(a)$, $\gamma(b) = x(b)$.*

In §2.7, for two-dimensional manifolds, we proved Theorem 2.8 which claims that a curve $x(t)$ is an extremal of $S$ if and only if the *Euler–Lagrange equations* are satisfied along this curve:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i}, \quad i = 1, \dots, n. \tag{9.13}$$

The proof given in §2.7 can be translated without changes to the general case of manifolds $M^n$ of arbitrary dimension. Recall that the function $L(x, \dot{x})$ is called the *Lagrangian* or the *Lagrangian function* of the system.

We see that a trajectory of the system is determined uniquely not only by its position at every given time $t = t_0$ but the pair of values: its position $x \in M^n$ and the velocity vector $\dot{x} = \frac{dx}{dt}$ at $t = t_0$. We obtain a dynamical system on the space $TM^n$ called the *phase space* of the dynamical system.

The *energy* (of a Lagrangian system) is the following (scalar) value:

$$E = \dot{x}^i \frac{\partial L}{\partial \dot{x}^i} - L,$$

and the *momentum* is the covector $p$ with coordinates

$$p_i = \frac{\partial L}{\partial \dot{x}^i}, \quad i = 1, \dots, n.$$

The *force* is the covector $f$ with coordinates

$$f_i = \frac{\partial L}{\partial x^i}, \quad i = 1, \dots, n.$$

In terms of these covectors the Euler–Lagrange equations take the form of Newton's law:

$$\dot{p} = f.$$

The *variational derivative* of a functional $S$ is defined as

$$\frac{\delta S}{\delta x^i(t)} = \frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right);$$

by definition, it equals zero for all values of $t$ and $i$ precisely for the extremals of $S$ (as in the finite-dimensional case when we speak of a function and its derivatives). Its meaning is rather simple: if we consider a small variation

$$x(t) \to x(t) + \varepsilon\eta(t)$$

of a smooth curve $x(t), t \in [a,b]$, where $\eta(t)$ is a tangent vector field along the curve ($\eta(t)$ is a tangent vector at $x(t)$ for every $t$) with $\eta(a) = \eta(b) = 0$, then we obtain the formal decomposition

$$S(x + \varepsilon\eta) = S(x) + \varepsilon \int \frac{\delta S}{\delta x^i(t)} \eta^i \, dt + O(\varepsilon^2)$$

(this formula was established in the proof of Theorem 2.8).

**Lemma 9.4.** (1) (Energy conservation law). *The energy of a system is conserved along the trajectories of the flow:*

$$\frac{dE}{dt} = 0.$$

(2) ( Momentum conservation law). *If the Lagrangian does not depend explicitly on a variable $x^i$ then the corresponding coordinate of the momentum is conserved along the flow: $\dot{p}_i = 0$.*

*Proof.* (1) Straightforward calculations show that

$$\begin{aligned} \frac{dE}{dt} &= \frac{d}{dt}\left(\dot{x}^i \frac{\partial L}{\partial \dot{x}^i} - L\right) \\ &= \ddot{x}^i \frac{\partial L}{\partial \dot{x}^i} + \dot{x}^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i}\dot{x}^i - \frac{\partial L}{\partial \dot{x}^i}\ddot{x}^i \\ &= \dot{x}^i \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i}\right) = 0. \end{aligned}$$

(2) We have

$$\dot{p}_i = \frac{\partial L}{\partial x^i} = 0.$$

The lemma is proven. □

**2 Examples of Lagrangian systems.** (a) Consider the classical problem of motion of a point mass in the $n$-dimensional Euclidean space in a potential field $U(x)$:

$$L(x, \dot{x}) = \frac{m\dot{x}^2}{2} - U(x),$$

where $x \in \mathbb{R}^n$ and $U(x)$ is the potential energy. The total energy of the body is the sum

$$E = \frac{m\dot{x}^2}{2} + U(x)$$

of the *kinetic energy* $T = \frac{m\dot{x}^2}{2}$ and the potential energy $U(x)$.[3] The momentum equals $p = m\dot{x}$ and, if the potential energy is zero, i.e., $U(x) = 0$, then the momentum is conserved.

(b) Now, consider the motion of a point mass in a nonhomogeneous medium, where the kinetic energy is the scalar square of the velocity vector (times $\frac{m}{2}$) calculated by means of a more general Riemannian inner product. We have

$$L(x, \dot{x}) = \frac{m}{2} g_{ik}(x)\dot{x}^i \dot{x}^k - U(x), \tag{9.14}$$

where $x \in M^n$ and $g_{ik}$ is a Riemannian metric on the manifold $M^n$. The momentum equals

$$p_i = g_{ik}x^k$$

and in the absence of a potential field (U$(x) = 0$) we obtain (as the Lagrangian system) the geodesic flow on $M^n$.

(c) Extend the previous Lagrangian system by adding the magnetic field and assuming that the point mass has charge $e$. By the Maxwell equations, the magnetic field is given by the closed 2-form

$$F = \sum_{i<k} F_{ik} dx^i \wedge dx^k$$

which determines the contribution of the magnetic field in the Euler–Lagrange equations:

$$\frac{\partial L_0}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L_0}{\partial \dot{x}^i}\right) + eF_{ik}\dot{x}^k = 0,$$

where $L_0$ is the Lagrangian of the form (9.14) or simply the Lagrangian of some system without the magnetic field. If the form $F$ is exact; i.e., there is a 1-form $A_i dx^i$ such that $F = dA$, then the presence of the magnetic field results in the following change of the Lagrangian:

$$L_0 \to L_0 + eA_i\dot{x}^i.$$

[3] In mechanics, Lagrangian systems with the Lagrangian $L = T - U$ (the difference of the kinetic and potential energies) are called *natural mechanical systems*.

In the general case we can make it only locally: take a domain $U$ in which all 2-forms are exact (this is true, for example, in every domain homeomorphic to the interior of the $n$-dimensional ball) and consider the form $A^U = A_i^U dx^i$ in this domain such that $dA^U = F$ in $U$. Then the trajectories of the system in $U$ satisfy the Euler–Lagrange equations for the modified Lagrangian

$$L^U = L_0 + eA_i^U \dot{x}^i.$$

If a curve $\gamma$ lies simultaneously in two such domains $U$ and $V$, then we have the values of two such functionals $S^U = \int L^U dt$ and $S^V = \int L^V dt$ (or, more exactly, two branches of the multivalued functional) whose variational derivatives coincide:[4]

$$\frac{\delta S^U}{\delta x^i(t)} \equiv \frac{\delta S^V}{\delta x^i(t)}, \quad x(t) \subset U \cap V.$$

The trajectories of the Lagrangian system with the Lagrangian (for curves lying in domains $U$)

$$L^U(x, \dot{x}) = g_{ik}(x)\dot{x}^i \dot{x}^k + A_i^U(x)\dot{x}^i,$$

where $dA^U = F$ is a globally defined 2-form, are now called *magnetic geodesics*.

## 9.3 Examples of Poisson manifolds

**1 The Legendre transform: passage from the Lagrangian formalism to the Hamiltonian formalism. The canonical Poisson bracket on the cotangent bundle.** We are not going to give the general definition of the *Legendre transform*. In the case in question it determines the passage from the function $L(x, \dot{x})$ of the variables $x$ and $\dot{x}$ to the function $H(x, p)$ of the variables $x$ and $p$ by the following formulas:

$$L(x, \dot{x}) \to H(x, p) = p_i \dot{x}^i - L(x, \dot{x}), \quad p_i = \frac{\partial L}{\partial \dot{x}^i}, \quad i = 1, \dots, n.$$

Naturally, we assume that these formulas determine a one-to-one change of variables $(x, \dot{x}) \to (x, p(x, \dot{x}))$ and the variable $\dot{x}$ in the formula for $H(x, p)$ is considered as a function of $x$ and $p$.

In the new variables $x$ and $p$ the Euler–Lagrange equations take the Hamiltonian form. Indeed, write down the differential of $H(x, p)$:

$$dH = \dot{x}^i dp_i + p_i d\dot{x}^i - \frac{\partial L}{\partial x^i} dx^i - \frac{\partial L}{\partial \dot{x}^i} d\dot{x}^i,$$

[4] This is a general procedure for obtaining multivalued functionals with correctly defined Euler–Lagrange equations which appear not only in problems of mechanics but also in quantum field theory (the Wess–Zumino–Novikov–Witten functional). For detailed exposition see Supplement 1 (by S. P. Novikov) to [2], III.

which is equal to

$$dH = \dot{x}^i dp_i - \frac{\partial L}{\partial x^i} dx^i$$

by the Euler–Lagrange equations. This equality means that

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = \frac{\partial L}{\partial x^i} = -\frac{\partial H}{\partial x^i};$$

i.e., we obtain Hamilton's equations

$$\dot{x}^i = \{x^i, H\}, \quad \dot{p}_i = \{p_i, H\}$$

for the Poisson structure

$$\{x^i, p_k\} = \delta^i_k, \quad \{x^i, x^k\} = 0, \quad \{p_i, p_k\} = 0, \quad i, k = 1, \dots, n. \tag{9.15}$$

We conclude that if $x(t)$ is a trajectory of the Lagrangian system, then $(x(t), p(t))$ is a trajectory of the Hamiltonian system with $H = p_i \dot{x}^i - L$. It is easy to verify that the converse is also true; therefore, the Lagrangian and Hamiltonian systems are equivalent.

Note that $p = \frac{\partial L}{\partial \dot{x}}$ is a covector and therefore the phase space of the Hamiltonian system is the cotangent bundle $T^*M^n$ of the configuration space $M^n$ (see §3.3).

On the cotangent bundle $T^*M^n$ we have the Poisson structure (9.15) which already appeared above. It is easy to verify that this structure is symplectic. The corresponding symplectic form on the cotangent bundle is equal to

$$\omega = dp_i \wedge dx^i \tag{9.16}$$

(here and in the sequel the repeated index $i$ implies summation). The coordinates $x$ and $p$ in which the symplectic form takes the shape (9.16) are called *canonical*.

The Hamiltonian coincides with the energy of the Lagrangian system:

$$H = E = \dot{x}^i p_i - L.$$

In this event, the energy conservation law goes into the identity which follows obviously from the anticommutativity of the Poisson bracket:

$$\frac{dH}{dt} = \{H, H\} = 0.$$

In general, a function $F$ whose values do not change along the trajectories of the Hamiltonian system:

$$\dot{F} = \{F, H\} = 0,$$

is called a *first integral* (or *integral of motion*) of the system. The Hamiltonian is always a first integral of the system generated by it.[5]

[5] Recall that we suppose that the Hamiltonian does not depend explicitly on time. The situation when this is not true is discussed below in §9.5.

Consider the example of a natural mechanical system with the Lagrangian $L(x, \dot{x}) = \frac{1}{2} g_{ik} \dot{x}^i \dot{x}^k - U(x)$. We have

$$p_i = g_{ik} \dot{x}^k, \quad H(x, p) = \frac{1}{2} g^{ik} p_i p_k + U(x), \tag{9.17}$$

where $g^{ij} g_{jk} = \delta^i_k$. In particular,

- the geodesic flow on a Riemannian manifold $M^n$ with a metric $g_{ik}$ is the Hamiltonian system on the cotangent bundle $T^* M^n$ with the Hamiltonian (9.17) with $U = 0$ and the Poisson structure (9.15).

**2 The magnetic Poisson bracket on the cotangent bundle.** If the magnetic field is given by an exact form $F = dA$, then construction of the Hamilton formalism for the magnetic geodesics is plain: instead of (9.17) we have

$$p_i = g_{ik} \dot{x}^k + eA_i, \quad H = \frac{1}{2} g^{ik} (p_i - eA_i)(p_k - eA_k),$$

where the Poisson structure has the form (9.15). Note that this method does not work if the covector field $A$ is not given globally. However, there is a quite general approach based on changing the Poisson structure. If $F = \sum_{i<k} F_{ik} dx^i \wedge dx^k$ is the magnetic field, then the 2-form

$$\omega = dp_i \wedge dx^i + F$$

on $T^* M^n$ is still nondegenerate and closed (recall that $dF = 0$). This form is called the *twisted symplectic form* and to this form there corresponds the *magnetic* (or *twisted*) *Poisson bracket* on the cotangent bundle:

$$\{x^i, p_k\} = \delta^i_k, \quad \{x^i, x^k\} = 0, \quad \{p_i, p_k\} = F_{ik}(x), \quad i, k = 1, \dots, n.$$

It is easy to check that the Hamiltonian system with this Poisson bracket and the Hamiltonian

$$H = \frac{1}{2} g^{ik} p_i p_k$$

is the magnetic geodesic flow (to this end, it suffices to insert $p_i = g_{ik} \dot{x}^k$ into Hamilton's equations).

**3 The Lie–Poisson bracket.** Let $x_1, \dots, x_n$ be coordinates in $\mathbb{R}^n$.[6] Consider the following simple question:

When does the formula

$$\{x_i, x_j\} = c^k_{ij} x_k \tag{9.18}$$

determine a Poisson bracket?

[6] Here we deliberately use subscripts, since $\mathbb{R}^n$ will be identified with the dual space of some Lie algebra $\boldsymbol{g}$.

Here $c_{ij}^k$ are constant and independent of $x$. Obviously, we have to require the symbols $c_{ij}^k$ be antisymmetric in the subscripts:

$$c_{ij}^k = -c_{ji}^k, \quad i, j, k = 1, \dots, n.$$

By Theorem 9.1, this is necessary and sufficient for the expression (9.18) to be antisymmetric in $f$ and $g$ and satisfy the Leibniz identity. Moreover, by the same theorem, for validity of the Jacobi identity, it is necessary and sufficient that $h^{ij} = c_{ij}^k x_k$ satisfy relations (9.10). Expand these relations and obtain

$$c_{il}^m x_m c_{jk}^l + c_{jl}^m x_m c_{ki}^l + c_{kl}^m x_m c_{ij}^l = (c_{il}^m c_{jk}^l + c_{jl}^m c_{ki}^l + c_{kl}^m c_{ij}^l) x_m = 0.$$

Since these identities must hold for all $x$, they are equivalent to the identities

$$c_{il}^m c_{jk}^l + c_{jl}^m c_{ki}^l + c_{kl}^m c_{ij}^l = 0,$$

which mean precisely that $c_{ij}^k$ are the structure constants of some Lie algebra (see §7.2). Thus, the relations

$$[e_i, e_j] = c_{ij}^k e_k$$

determine a Lie algebra $\boldsymbol{g}$.

**Theorem 9.3.** *Formula* (9.18) *determines a Poisson bracket on* $\mathbb{R}^n$ (*linear in the coordinates*) *if and only if* $c_{ij}^k$, $i, j, k = 1, \dots, n$, *are the structure constants of some* $n$*-dimensional Lie algebra.*

The bracket (9.18) is called the *Lie–Poisson bracket*. It is natural to consider it as a bracket on the dual space $\boldsymbol{g}^*$ of the Lie algebra $\boldsymbol{g}$, i.e., the space constituted by all linear functionals on the algebra $\boldsymbol{g}$. In this case the elements of the Lie algebra determine the functions $x_1, \dots, x_n$ on the space $\boldsymbol{g}^*$ by the rule $x(f) = f(x)$, $x \in \boldsymbol{g}$, $f \in \boldsymbol{g}^*$.

Consider one important example, namely the Poisson bracket connected with the algebra $so(3)$ constituted by all anti-symmetric matrices of order 3. We have

$$c_{ij}^k = \varepsilon_{ijk} = \begin{cases} 1 & \text{if the permutation } \begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix} \text{ is even;} \\ -1 & \text{if the permutation } \begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix} \text{ is odd;} \\ 0 & \text{if at least two of the indices } i, j, \text{ and } k \text{ coincide.} \end{cases}$$

Since the space $\mathbb{R}^3$ is odd-dimensional, this Poisson bracket is not symplectic. It is easy to verify that

$$\{f, x_1^2 + x_2^2 + x_3^2\} = 0 \quad \text{for all } f.$$

**Problem 9.1.** Prove that we can correctly restrict the Poisson bracket corresponding to the algebra $so(3)$ to every sphere of the form

$$x_1^2 + x_2^2 + x_3^2 = c \neq 0,$$

obtaining thereby a symplectic manifold.

Consider the Hamiltonian system determined by the quadratic Hamiltonian

$$H = \frac{1}{2}\left(\frac{x_1^2}{A} + \frac{x_2^2}{B} + \frac{x_3^2}{C}\right)$$

with respect to this bracket (see Problem 9.1). After the changes

$$p = -\frac{x_1}{A}, \quad q = -\frac{x_2}{B}, \quad r = -\frac{x_3}{C},$$

Hamilton's equations

$$\dot{x}_i = \{x_i, H\} = \frac{\partial H}{\partial x_j} c_{ij}^k x_k, \quad i = 1, 2, 3,$$

take the form

$$A\dot{p} = (B - C)qr, \quad B\dot{q} = (C - A)pr, \quad C\dot{r} = (A - B)pq.$$

For positive values of $A$, $B$, and $C$ we obtain the equations of motion of a (three-dimensional) rigid body around a fixed point (*Euler's equations*). Moreover, the constants $A$, $B$, and $C$ are the moments of inertia of the body.

## 9.4 Darboux's theorem and Liouville's theorem

**Theorem 9.4** (Darboux's theorem). *Let $M$ be a Poisson manifold. If, in a neighborhood $U$ of $y_0 \in M$, the Poisson bracket has constant rank equal to $2m$, then there are a neighborhood $V \subset U$ of the point $y_0$ and local coordinates $x^1, \dots, x^m, p_1, \dots, p_m$, and $z^1, \dots, z^{n-2m}$ in this neighborhood such that the Poisson bracket in the whole domain $V$ takes the form*

$$\{x^i, p_j\} = \delta_j^i, \quad \{x^i, z^k\} = \{p_j, z^k\} = 0, \tag{9.19}$$

*where $i, j = 1, \dots, m$ and $k = 1, \dots, n - 2m$.*

Such coordinates on a Poisson manifold are called canonical coordinates. In the symplectic case when $n = 2m$, we obtain canonical coordinates on a symplectic manifold.

*Proof.* Let $y^1, \dots, y^n$ be local coordinates in $U$. If the Poisson bracket is identically zero in a neighborhood of $y_0$ ($m = 0$):

$$\{y^i, y^j\} = h^{ij} = 0, \quad i, j = 1, \dots, n,$$

then we can take $z^1, \dots, z^n$ to be arbitrary coordinates in this neighborhood.

If $m \neq 0$, then take $x^1$ to be a function $f$ for which the vector $v_f$ is nonzero at $y_0$ and hence in its small neighborhood (recall that its coordinates are equal to $v_f^i = h^{ik} \frac{\partial f}{\partial y^k}$). The function $f$ considered as a Hamiltonian generates some Hamiltonian system; take $p_1$ to be the function $p_1 = t$, where $t$ is the time coordinate for this Hamiltonian system.

We now give a rigorous definition. It is well known from the theory of ordinary differential equations that if a smooth vector field $v$ is nonzero at $y_0$, i.e., $v(y_0) \neq 0$, it can be linearized by a change of coordinates. This means that in a neighborhood $U'$ of the point $y_0$ we can introduce coordinates $u^1, \dots, u^n$ in which the field $v$ takes the form

$$v = (1, 0, \dots, 0).$$

For $v = v_{x^1}$ we put $p_1 = -u^1$. By construction,

$$\dot{p}_1 = -\dot{t} = -1 = \{p_1, x^1\} = -\{x^1, p_1\}.$$

Since $x^1$ is conserved along the vector field $v_{x^1}$, we have $x^1 = x^1(u^2, \dots, u^n)$.

The function $p_1$ also generates a Hamiltonian system with the vector field $v_{p_1}$. The vector fields $v_{x^1}$ and $v_{p_1}$ commute, since

$$[v_{x^1}, v_{p_1}] = -v_{\{x^1, p_1\}} = 0$$

(the function $\{x^1, p_1\}$ is constant), and the function $p_1$ is conserved along the trajectories of the field $v_{p_1}$ (as the Hamiltonian of the system). Therefore,

$$v_{p_1} = (0, *, \dots, *), \quad v_{p_1} = v_{p_1}(u^2, \dots, u^n)$$

(the first equality follows from the conservation of $p_1$ and the second, from the fact that the fields commute).

Now, we linearize the field $v_{p_1}$, replacing $u^2, \dots, u^n$ with coordinates $w^0, \dots, w^{n-2}$ such that

$$v_{p_1} = (1, 0, \dots, 0).$$

Moreover, take $w^0$ to be $x^1$.

Eventually, in a neighborhood $\widetilde{U}$ of $y_0$ (perhaps smaller than $U$) we obtain coordinates $x^1, p_1, w^1, \dots, w^{n-2}$ in which the Poisson bracket has the form

$$\{x^1, p_1\} = 1, \quad \{x^1, w^i\} = \{p_1, w^i\} = 0, \quad i = 1, \dots, n-2.$$

If $n - 2 = 0$, then these are the sought coordinates. If $n - 2 > 0$, then we repeat the above procedure for the coordinates $w^1, \dots, w^{n-2}$ until we obtain the sought coordinates. The theorem is proven. □

**Corollary 9.3.** *In a neighborhood of each point of a symplectic manifold we can introduce canonical coordinates, i.e., coordinates $x^1, \dots, x^n, p_1, \dots, p_n$ in which the symplectic form has the shape*

$$\omega = dp_i \wedge dx^i. \tag{9.20}$$

**Corollary 9.4.** *On a symplectic manifold there is a nowhere zero volume form*

$$d\mu = \omega^n;$$

*in particular, a symplectic manifold is always orientable.*

Indeed, in the canonical coordinates we have

$$\frac{1}{n!}\omega^n = dp_1 \wedge dx^1 \wedge \cdots \wedge dp_n \wedge dx^n \neq 0,$$

and as a positively oriented basis for the tangent space at the point we can take a basis $e_1, \dots, e_{2n}$ such that

$$\omega^n(e_1, \dots, e_{2n}) > 0.$$

This determines an orientation on $M$.

Note that, if a symplectic manifold $M$ is closed (compact and without boundary), then the form $\omega^n$ cannot be exact. Otherwise, if $\omega^n = d\tau$, then by the Stokes formula we would have

$$\int_M \omega^n = \int_M d\tau = \int_{\partial M} \tau = 0,$$

since the boundary $\partial M$ is empty. But this contradicts the fact that $\int_M d\mu > 0$. Hence, all forms $\omega, \omega^2, \dots, \omega^{n-1}$ are not exact: otherwise, if $\omega^k = d\sigma$, then we would have $\omega^n = d(\sigma \wedge \omega^{n-k})$. This is a very strong topological constraint on symplectic manifolds. It means that the real cohomology classes $[\omega^k]$, $k = 1, \dots, n$, are nonzero.[7] For example, among all spheres only the two-dimensional sphere $S^2$ carries a symplectic form.

Canonical coordinates enable us to simplify essentially the proof of the following theorem:

**Theorem 9.5.** *Let $M$ be a symplectic manifold.*

(1) *The symplectic form $\omega$ is preserved along the trajectories of the Hamiltonian system on $M$:*

$$\dot{\omega} = 0.$$

(2) *If the symplectic form is preserved along the trajectories of a vector field $v$ on $M$, then $v$ is locally Hamiltonian: in some neighborhood $U$ of each point $x \in M$, there is a smooth function $f$ such that*

$$v = v_f \text{ in } U.$$

*Proof.* In canonical coordinates we have

$$\dot{\omega} = d\dot{p}_i \wedge dx^i + dp_i \wedge d\dot{x}^i.$$

[7] An explanation of what it means can be found, for example, in [4].

If the system is Hamiltonian, then we can rewrite the right-hand side of the last relation as

$$-d\left(\frac{\partial H}{\partial x^i}\right) \wedge dx^i + dp_i \wedge d\left(\frac{\partial H}{\partial p_i}\right) = -d\left(\frac{\partial H}{\partial x^i}dx^i + \frac{\partial H}{\partial p_i}dp_i\right)$$
$$= -d(dH) = -d^2H = 0,$$

since $d^2 \equiv 0$.

If $\dot{\omega} = 0$ along the trajectories of the field $v = (\dot{x}, \dot{p})$, then

$$0 = d\,\dot{p}_i \wedge dx^i + dp_i \wedge d\dot{x}^i = -d(\dot{x}^i\,dp_i - \dot{p}_i\,dx^i).$$

It is well known from calculus that if, in a simply connected domain $U \subset \mathbb{R}^n$, the differential of a 1-form $\alpha$ is zero, then this form is the differential of some smooth function $f$ in $U$:

$$\alpha = df.$$

For $\alpha = \dot{x}^i\,dp_i - \dot{p}_i\,dx^i$ (here the repeated index $i$ implies summation) we find that in a small neighborhood of each point we have a function $f = H$ such that

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \quad i = 1, \dots, n;$$

i.e., the field $v$ is locally Hamiltonian. The theorem is proven. □

A transformation of the space

$$\varphi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$$

with coordinates $x^1, \dots, x^n, p_1, \dots, p_n$ is called *canonical* if it preserves the form (9.20). In general, this term is now used for transformations of an arbitrary symplectic manifold $M$ which preserves the symplectic form $\omega$.

The assertion (1) of Theorem 9.5 means that translations along the trajectories of Hamiltonian vector fields are canonical transformations.

**Corollary 9.5** (Liouville's theorem on the volume preservation). *A Hamiltonian system preserves the volume form*

$$d\mu = \omega^n = \underbrace{\omega \wedge \cdots \wedge \omega}_{n}$$

*on the phase space $M$.*

Indeed,

$$\frac{d\omega^n}{dt} = n\dot{\omega} \wedge \omega^{n-1} = 0,$$

since $\dot{\omega} = 0$.

This corollary makes it possible to study Hamiltonian systems by the methods of ergodic theory which deals with measure-preserving transformations of measure spaces.

As we see, the Poisson bracket can be degenerate. A function $f$ such that

$$\{f, g\} = 0 \quad \text{for all } g$$

is called a *Casimir function* (or simply a *casimir*) of the bracket.

In general, every algebra $A$ with commutative multiplication and a bilinear operation

$$a, b \to \{a, b\}$$

which is antisymmetric and satisfies the Jacobi identity and the Leibniz identity

$$\{ab, c\} = a\{b, c\} + b\{a, c\}$$

is called a *Poisson algebra.* The elements $z$ of the center $Z_A$:

$$\{z, a\} = 0 \quad \text{for all } a \in A,$$

of the algebra $A$ are called *Casimirs*.

A *symplectic leaf* is a submanifold $L$ of a Poisson manifold $M$ such that

(1) the rank of the Poisson bracket at the points of $L$ is equal to the dimension of $L$ and the manifold $L$ carries a symplectic form $\omega_L$ which generates the Poisson bracket $\{\cdot, \cdot\}_L$ on $L$;

(2) if $f$ and $g$ are functions on $M$, then their Poisson bracket $\{f, g\}$ at the points of $L$ coincides with $\{f|_L, g|_L\}_L$ (here $f|_L$ and $g|_L$ are the restrictions of $f$ and $g$ to $L$):

$$\{f, g\}|_L = \{f|_L, g|_L\}_L.$$

**Examples.** (1) Suppose that the rank of the bracket is constant in a domain $U$ and we have canonical coordinates:

$$\{x^i, p_j\} = \delta^i_j, \quad \{x^i, z^k\} = \{p_j, z^k\} = 0.$$

Casimirs in the Poisson algebra $C^\infty(U)$ are precisely all functions of the $z$-variables and the symplectic leaves are the submanifolds $z^1 = c_1, \dots, z^{n-2m} = c_{n-2m}$, where $c_1, \dots, c_{n-2m}$ are constants. The symplectic forms on these leaves are given by the formula $\omega_L = dp_i \wedge dx^i$.

(2) For the Lie–Poisson bracket on $\mathbb{R}^3$ connected with the algebra $so(3)$, the Casimirs constitute the ring of functions of the form $f(x_1, x_2, x_3) = h(x_1^2 + x_2^2 + x_3^2)$. The symplectic leaves are the spheres

$$x_1^2 + x_2^2 + x_3^2 = c$$

of dimension 2 for $c \neq 0$ and 0 for $c = 0$.

These two examples lead to the following important constructions:

**(a) Foliation into symplectic leaves.** Consider all Hamiltonian vector fields $v_f$. At each point $x \in M$ the values $v_f(x)$ of the Hamiltonian vector fields generate a subspace $P_x \subset T_x M$ of the tangent space $T_x M$ at the point $x$. Every nonzero vector in $P_x$ has the form $v_f$; moreover, if $t$ is the time coordinate for the Hamiltonian system with $H = f$, then $\{t, f\} = 1$ (see the proof of Darboux's theorem), and therefore the rank of the Poisson bracket at $x$ is equal to the dimension of $P_x$; in particular, it is even.

Since $[v_f, v_g] = -v_{\{f,g\}}$, the Hamiltonian fields constitute a subalgebra of the Lie algebra of vector fields. By the Frobenius theorem, for a small neighborhood $U$ of each point $x$ there is a submanifold $L$ embedded into $U$ such that at every point $y \in L$ the tangent space of $L$ is $P_y$. The whole neighborhood $U$ foliates into such submanifolds which are symplectic leaves. Note that, for each point, there is exactly one symplectic leaf passing through this point. The leaf $L$ extends naturally to some subset also denoted by $L \subset M$ which is the image of (possibly improper)[8] embedding of the symplectic manifold into $M$. Such sets $L$ are also called symplectic leaves.

We arrive at the following conclusion:

- *A Poisson manifold foliates into a disjoint union of symplectic leaves and this representation is unique.*

**(b) Symplectic leaves of the Lie–Poisson bracket.** Let $\boldsymbol{g}^* = \mathbb{R}^n$ be the space of linear functions on the Lie algebra $\boldsymbol{g}$ of a group $G$. A Lie–Poisson bracket is given on $\boldsymbol{g}^*$.

Take $x \in \boldsymbol{g}^*$. Consider the space $P_x$ constituted by the values of all Hamiltonian fields $v_f$ at $x$. The value of a field $v_f$ at a given point depends only on the linear part of the Taylor series expansion of the function $f$ and therefore, to obtain the whole space $P_x$, it suffices to consider only linear functions $a(x) = a^i x_i$. The space of linear functions on $\boldsymbol{g}^*$ coincides with the Lie algebra $\boldsymbol{g}$:

$$a(x) = x(a), \quad x \in \boldsymbol{g}^*, \quad a \in (\boldsymbol{g}^*)^* = \boldsymbol{g}.$$

The Hamiltonian vector fields at $x = (x_1, \dots, x_n)$ (recall that we use subscripts for coordinates in $\boldsymbol{g}^*$) take the values

$$(v_a(x))_j = c_{ij}^k x_k a^i, \quad j = 1, \dots, n.$$

Considering them as elements in $\boldsymbol{g}^*$, we see that their values at $b \in \boldsymbol{g}$ are equal to

$$v_a(x)(b) = c_{ij}^k x_k a^i b^j = x_k([a,b])^k = x([a,b]). \tag{9.21}$$

[8] Recall that a map is *proper* if the inverse image of each compact set is compact. A typical example of an improper embedding is an irrational winding of the torus $T^2$ with angular coordinates $\varphi_1$ and $\varphi_2$ defined modulo integer numbers. This winding has the form: $t \in \mathbb{R} \to \psi(t) = (t, \alpha t) \in T^2$, where $\alpha$ is an irrational number. We have $\psi(t_1) \neq \psi(t_2)$ for $t_1 \neq t_2$, $d\psi \neq 0$ everywhere, and the closure of $\psi(\mathbb{R})$ coincides with the whole torus. Similar situations appear also for symplectic leaves $L$ in Poisson manifolds.

Recall that to each element $a \in \boldsymbol{g}$ of the Lie algebra there corresponds the one-parameter subgroup $F(t) = e^{at} \in G$ (for matrix groups it has the form of the exponential function of matrices $at$). On the Lie group we have families of diffeomorphisms: the *left* and *right translations*:

$$L_g : G \to G,\ L_g(x) = gx; \quad R_g : G \to G,\ R_g(x) = xg,$$

where $g, x \in G$ (for matrix groups $G$ these are the left and right multiplications by a matrix $g \in G$). The differentials of the right and left translations determine the following commuting maps of the tangent spaces:

$$L_g^* : T_x G \to T_{gx} G, \quad R_g^* : T_x G \to T_{xg} G,$$

and generate, by the obvious rules

$$p(L_g^* v) = L_{g^{-1}}^*(p)(v), \quad p(R_g^* v) = R_{g^{-1}}^*(p)(v),$$

the maps of the cotangent spaces: $L_g^* : T_{gx}^* G \to T_x^* G$ and similarly for $R_g^*$. The composition of commuting diffeomorphisms $L_g$ and $R_{g^{-1}}$ is an inner automorphism of the group:

$$x \to gxg^{-1},$$

and the differential of this map at the identity determines the action of $G$ on the Lie algebra $\boldsymbol{g}$:

$$\mathrm{Ad}_g = L_g^* R_{g^{-1}}^* : \boldsymbol{g} \to \boldsymbol{g}.$$

We obtain the *adjoint representation* of $\mathrm{Ad}_g$; for matrix groups it has a very simple form:

$$\mathrm{Ad}_g(a) = gag^{-1}, \quad g \in G,\ a \in \boldsymbol{g}.$$

This representation generates the *coadjoint representation* of $\mathrm{Ad}_g^*$ on $\boldsymbol{g}^*$:

$$(\mathrm{Ad}_g^*(x))(a) = x(\mathrm{Ad}_g(a)), \quad x \in \boldsymbol{g}^*,\ a \in \boldsymbol{g},\ g \in G.$$

Consider the differential of the map $\mathrm{Ad}_g$ at the identity:

$$\frac{d}{dt} \mathrm{Ad}_{e^{at}}(b) = \mathrm{ad}_a(b).$$

For linear groups we obviously have

$$\mathrm{ad}_a(b) = \frac{d}{dt}(e^{at} b e^{-at}) = ab - ba = [a, b];$$

moreover, this formula is valid in the general case (to show this, it suffices to use the Taylor series expansion of the multiplication at the identity of the group as in §7.2). Note that for the differential $\mathrm{ad}^*$ of the coadjoint representation we have

$$\mathrm{ad}_a^*(x)(b) = x(\mathrm{ad}_a(b)) = x([a, b])$$

and compare this formula with (9.21). We see that the values of the Hamiltonian vector fields at $x$ coincide with the tangent fields to the orbit of the coadjoint representation at $x$.

Hence, we conclude that

- *the orbits of the coadjoint representation of the group $G$ are symplectic leaves of the Lie–Poisson bracket on $\mathbf{g}^*$.*

## 9.5 Hamilton's variational principle

Since the Hamiltonian $H$ is conserved by a Hamiltonian system, every trajectory of this system lies completely on some level surface $M_c = \{H = c\}$.

Consider the case when the hypersurface $M_c$ is a smooth submanifold in the initial symplectic manifold $M$ (with the symplectic form $\omega$). In this case the manifold $M_c$ is odd-dimensional and the restriction of the form $\omega$ to $M_c$ cannot be nondegenerate. At each point $x \in M_c$ the tangent space of $M_c$ contains a nonzero subspace $A_x$ called the *annihilator* of the form $\omega$ such that

$$\omega(\xi, \eta) = 0 \quad \text{for all } \xi \in A_x,\ \eta \in T_x M_c;$$

i.e., the linear form $l(\eta) = \omega(\xi, \eta)$ is identically zero.[9]

If $A_x$ contained two linearly independent vectors $\xi_1$ and $\xi_2$, then for every tangent vector $\eta$ at $x$ such that $\eta$ and $T_x M_c$ generate the whole tangent space of $M$, we could choose numbers $a_1$ and $a_2$ such that

$$\omega(a_1\xi_1 + a_2\xi_2, \eta) = 0$$

and the vector $a_1\xi_1 + a_2\xi_2$ would lie in the annihilator of the form $\omega$ on the whole tangent space of $M$. But this is impossible, since the form $\omega$ is nondegenerate by definition. Consequently, at each point of $M_c$ the annihilator $A_x$ is one-dimensional.

We say that a curve $r(t) \subset M_c$ is a *characteristic* of the form $\omega$ if at each point $r(t)$ of the curve the tangent vector to the curve lies in the annihilator of the form $A_x$.

In a neighborhood of every point $x_0$, we can choose a vector field $v(x)$ such that $v(x) \neq 0$ and $v(x)$ generates $A_x$ (this may fail to hold globally on the whole manifold $M_c$ in view of some topological obstacles). Moreover, the characteristics are precisely the integral curves of the field and we conclude that

- for each point of the level surface $M_c = \{H = c\}$ there is only one characteristic of the form $\omega$ passing through this point; moreover, a sufficiently small neighborhood of the point foliates into these characteristics.

[9]This definition of annihilator is common for all linear spaces with a bilinear form $B(\xi, \eta)$ (symmetric or antisymmetric): a vector $\xi$ belongs to the annihilator if and only if $B(\xi, \eta) = 0$ for all vectors $\eta$. For example, the Casimirs constitute the annihilator of the Poisson bracket on the space of functions.

**Theorem 9.6.** *The trajectories of the Hamiltonian system on the level surface $M_c$ are characteristics of the form $\omega$ on $M_c$.*

*Proof.* We introduce the canonical coordinates $x^1, \dots, x^n, p_1, \dots, p_n$ in a neighborhood of the point $x$. The condition that $\eta$ is a tangent vector to $M_c$ at $x$ takes the form

$$dH(\eta) = \eta^i \frac{\partial H(x)}{\partial x^i} + \eta^{i+n} \frac{\partial H(x)}{\partial p_i} = 0$$

(here we mean summation over $i$), and a vector $\xi$ lies in the annihilator $A_x$ if and only if

$$\omega(\xi, \eta) = [dp_i \wedge dx^i](\xi, \eta) = \eta^1 \xi^{i+1} + \dots + \eta^n \xi^{2n} - \eta^{i+1} \xi^1 - \dots - \eta^n \xi^{2n} = 0$$

for all $\eta \in T_x M_c$. Comparing these two formulas, we see that the vector

$$\xi = v_H = \left( \frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x} \right)$$

belongs to $A_x$. The theorem is proven. □

Now, consider the 1-form $\alpha$ given by the condition

$$d\alpha = \omega \quad \text{on } M_c,$$

and the variational problem for the functional

$$S(\gamma) = \int_\gamma \alpha$$

on the spaces of curves satisfying some additional conditions.

Perhaps, the form $\alpha$ is not defined globally. This case is typical, for example, when $M$ is a closed symplectic manifold. We arrive again at the multivalued functional $S$ (see §9.2) for which, as we will see below, the equations for extremals are written down globally.

We consider the following two spaces of curves on $M_c$:

(1) the space of closed curves;

(2) the space of curves joining two submanifolds $V_1$ and $V_2$ such that the restrictions of $\alpha$ to $V_1$ and $V_2$ are zero.

**Theorem 9.7.** *The extremal functions of the functional $S = \int d^{-1}\omega$ on each of the above spaces $\Gamma$ of curves are precisely the characteristics of the form $\omega$ lying in $\Gamma$.*

*Proof.* Let $\gamma$ be an extremal of $S$. If the tangent vector $\dot{\gamma}$ at $x \in \gamma$ does not belong to the annihilator $A_x$, then there is a vector $\eta \in T_x M_c$ such that $\omega(\dot{\gamma}, \eta) \neq 0$. Changing, if necessary, the direction of $\eta$, we can assume that $\omega(\dot{\gamma}, \eta) > 0$ and extend $\eta$ to a small part of the curve near $x$ so that the inequality holds on this part and $\eta = 0$ outside it. By the Stokes theorem, local variation of the curve $\gamma \to \gamma + \varepsilon\eta$ leads to variation of the functional

$$S(\gamma) \to S(\gamma) + \varepsilon \int \omega(\dot{\gamma}, \eta) + O(\varepsilon^2) \tag{9.22}$$

which is nonzero for small $\varepsilon$, which contradicts the condition that the curve $\gamma$ is extremal. Consequently, $\gamma$ is a characteristic.

It follows from (9.22) that the converse is also true: if $\gamma$ is a characteristic, then $\delta S = 0$. The theorem is proven. □

**Corollary 9.6.** *The closed trajectories of a Hamiltonian system on a level surface $M_c = \{H = c\}$ are precisely the closed characteristics of the form $\omega$ on $M_c$.*

We now give another corollary. Let $M$ be $\mathbb{R}^{2n}$ or, in general, the cotangent bundle $T^*X^n$ to an $n$-dimensional manifold $X$ with canonical coordinates $x^i$ and $p_k$. The form $\alpha = d^{-1}\omega$ is defined globally:

$$d^{-1}\omega = p_i dx^i.$$

**Corollary 9.7.** *The trajectories of a Hamiltonian system on $T^*X^n$ lying on the level surface $M_c = \{H = c\}$ are extremals of the functional*

$$S = \int p_i dx^i$$

(*here we naturally assume summation over $i$*) *on the space of curves lying in $M_c$ and joining the spaces $x = x_0$ and $x = x_1$.*

On the whole symplectic manifold the variational principle has the more habitual form and we leave its proof as an exercise (all necessary ideas and constructions are already exposed above):

**Theorem 9.8** (Hamilton's variational principle). *The trajectories of a Hamiltonian system on a symplectic manifold $M$ are extremal functions of the functional*

$$S = \int (d^{-1}\omega - H dt),$$

*where $H$ is the Hamiltonian.*

Moreover, formally, this is the form of Corollary 9.7 for systems with the Hamiltonian depending on time. In this case we obtain the usual Hamilton equations

$$\frac{df}{dt} = \{f, H\},$$

where $f$ is an arbitrary function on the manifold $M$, but now the Hamiltonian depends explicitly on time:

$$H = H(x, p, t), \quad \frac{dE}{dt} = \frac{\partial H}{\partial t}.$$

To reduce this nonautonomous system to the autonomous form, we use the standard method, making the *time* one of the variables. We have to consider the new symplectic manifold, the *extended phase space* $M \times \mathbb{R}^2$, with the symplectic form

$$\omega_0 = \omega + dt \wedge dE,$$

where "time" $t$ and "energy" $E$ are coordinates on $\mathbb{R}^2$. Now, replacing the Hamiltonian $H(x, p, t)$ with

$$\widehat{H}(x, p, t, E) = H(x, p, t) - E,$$

we obtain the Hamiltonian system

$$\dot{f} = \{f, H\}, \quad \dot{E} = \{E, H\} = \frac{\partial H}{\partial t}, \quad \dot{t} = \{t, H\} = 1,$$

whose restriction to the level surface $\{\widehat{H} = 0\}$ coincides with the initial system:

We make one small remark: the form $\alpha = d^{-1}\omega$ on the $(2n-1)$-dimensional level surface $M_c$ (if defined globally) satisfies the inequality

$$\alpha \wedge (d\alpha)^{n-1} \neq 0$$

everywhere. Manifolds of dimension $2n - 1$ with such 1-form are called *contact manifolds* and the form itself is called the *contact form*. We have the following analog of Darboux's theorem which is proven by means of the indicated theorem:

- *in a neighborhood of every point of a contact manifold, there exist local coordinates* $x^1, \dots, x^{n-1}, p_1, \dots, p_{n-1}, t$ *such that the contact form takes the shape*

$$\alpha = p_i dx^i - dt$$

*in this neighborhood.*

## 9.6 Reduction of the order of the system

**1 First integrals and reduction of the order of dynamical systems.** Let

$$\dot{x} = v(x)$$

be a dynamical system (flow) on an $n$-dimensional manifold $M$.

Recall that in local coordinates $x^1, \dots, x^n$ this is simply a system of ordinary differential equations:

$$\dot{x}^i = v^i(x^1, \dots, x^n), \quad i = 1, \dots, n.$$

These expressions agree on the intersections of charts with different coordinates. This means that $v(x)$ is a vector field on $M$.

If we know a nontrivial function $F$ which is conserved along the trajectories of the system:

$$\frac{dF}{dt} = v^i \frac{\partial F}{\partial x^i} = 0$$

(i.e., a first integral of the flow), then we can *reduce the order* of this system by 1 in domains where grad $F \neq 0$. To this end, consider the level surfaces $M_c = \{F = c\}$. By the implicit function theorem, if grad $F(x) \neq 0$ at a point $x \in M_c$, then the level surface $M_c$ is an $(n-1)$-dimensional submanifold in a neighborhood of $x$ with some local coordinates $y^1, \dots, y^{n-1}$. Since the trajectories of the system starting at the points of $M_c$ remain in $M_c$, the field $v$ is tangent to $M_c$ everywhere and the system restricted to $M_c$ takes the form

$$\dot{y}^i = w^i(y^1, \dots, y^{n-1}), \quad i = 1, \dots, n-1.$$

Here $w$ is the restriction of the field $v$ to $M_c$. As a result of this procedure we simplify the system, reducing the number of variables, the order of the system, by 1 (see Figure 9.1).

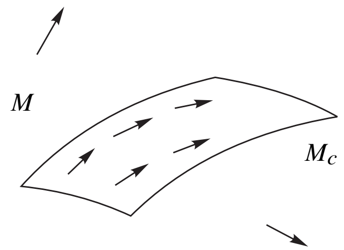


Figure 9.1. The invariant submanifold $M_c$.

Assume that we know a greater number of first integrals $F_1, \dots, F_k$ and these integrals are (*functionally*) *independent* at some point $x_0$; i.e., the system of the covectors grad $F_1, \dots,$ grad $F_k$ is linearly independent. Then, taking the restriction of the flow to the common level surface

$$\Gamma_{c_1,\dots,c_k} = \{F_1 = c_1 = F_1(x_0), \dots, F_k = c_k = F_k(x_0)\},$$

we reduce the order of the system by $k$. Moreover, if $k = n-1$ and $v(x_0) \neq 0$, then this level surface is one-dimensional and therefore is a trajectory of the flow. Thus, we obtain a (local) solution to the equation $\dot{x} = v(x)$ with the initial data $x(0) = x_0$.

**2 Reduction of the order of a Hamiltonian system.** For Hamiltonian systems on a symplectic manifold, the presence of one additional integral allows one to reduce the order of the system by 2. We describe this procedure which was implicitly used already in the proof of Darboux's theorem (§9.4).

Let $H$ be the Hamiltonian of the system and let $f$ be an extra first integral. In view of Hamilton's equations, $\dot{f} = \{f, H\} = 0$. Assume that grad $f \neq 0$ at a point $x_0$ and $f(x_0) = c$. The Hamiltonian system can be restricted to the level surface

$$\Gamma_c = \{f = c\};$$

moreover, on this level surface the initial flow and the flow $\dot{x} = v_f$ (the Hamiltonian system generated by the Hamiltonian $f$) commute. This follows from the identity $[v_f, v_h] = -v_{\{f,h\}}$ (see Lemma 9.2). Now, take one more integral of motion to be the function $g$, the time variable on the trajectories of the flow $\dot{x} = v_f$ (its correct construction was given in the proof of Darboux's theorem).

Suppose for simplicity that $\{g, H\} = 0$. Then we restrict the Hamiltonian system to the common level surface

$$\Gamma_{c,0} = \{f = c, g = 0\},$$

reducing its order by 2.

The above procedure is local. It can be made globally under two assumptions:

(1) the level surface $\Gamma_c = \{f = c\}$ is a submanifold;

(2) all trajectories of the flow $\dot{x} = v_f$ are closed and have the same period $T$. In this case we can suppose that the translations $\varphi_t : \Gamma_c \to \Gamma_c$ along the trajectory for time $t$ determine the action of the group $S^1 = \mathbb{R}/T\mathbb{Z}$.

In this case the initial Hamiltonian system can be restricted to the symplectic manifold

$$\Gamma_{c,0} = \Gamma_c / S^1.$$

The construction of the symplectic structure on this manifold is similar to that in the proof of Darboux's theorem: at each point, it is obtained by taking the restriction of the initial form to the orthogonal (with respect to this form) complement to the space generated by the coordinates $f$ and $g$, where $\{f, g\} = 1$ (in the proof of Darboux's theorem we had the coordinates $x^1$ and $p_1$, where $\{x^1, p_1\} = 1$).

In the general case when $\{g, H\} \neq 0$ we still can reduce the order of the system by 2, passing to the quotient space $\Gamma_c$ by the action of translations along the integral curves of the field $v_f$. Before giving an explanation of this procedure, we define the notion of quotient space and introduce a smooth structure on it.

**3 Smooth structure on the quotient space.** We should explain what is meant by the expression

$$M = X/G$$

(for example, $\Gamma_{c,0} = \Gamma_c/S^1$). This is the quotient space of the space $X$ by the action of the group $G$. Its points are sets of the form $Gx = \{gx \mid g \in G\}$, i.e., the *orbits* of the action. We can always introduce the canonical topology on this quotient space: a set $U \in X/G$ is open if and only if its inverse image $\pi^{-1}(U) \subset X$ is open. Here $\pi\colon X \to X/G$ is the projection which relates to a point $x$ its orbit $Gx$.

If the action, i.e., the map $X \times G \to X$ of the form $(x, g) \to gx$, is smooth, the group $G$ is compact, and the action of the group is *free* (it means that the equality $gx = x$ holds only for $g = 1$), then we can introduce the structure of a smooth manifold on the quotient space. This is done as follows.

A small neighborhood $U$ of an arbitrary point $x_0 \in X$ splits into orbits of the action of the group $G$. Consider a submanifold $Z$ in $U$ such that, at each point $z$, the tangent space of $X$ splits into the direct sum

$$T_zX = T_z(Gz) + T_zZ$$

of the tangent space of the orbit $Gz$ of the group $G$ and the tangent space of $Z$. Then we can introduce the local coordinates $y^1, \dots, y^m, z^1, \dots, z^{k-m}$ in $U$, where $y$ are local coordinates in a neighborhood $V$ of the identity of the group $G$ and $z$ are coordinates on $Z$ (see Figure 9.2).

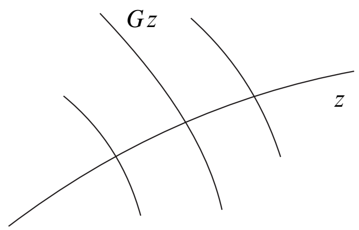


Figure 9.2. The quotient space $X/G$.

Every point $x \in U$ can be uniquely represented in the form $x = gz$, where $g \in V$ and $z \in Z$. Moreover, suppose that the neighborhood $U$ is so small that the equality $g_1z_1 = g_2z_2$, where $g_1, g_2 \in G$ and $z_1, z_2 \in Z$, implies that $g_1 = g_2$ and $z_1 = z_2$. (For existence of such unique decomposition, it is necessary that $G$ be compact and consequently its orbit be submanifolds.) Now, relate to the point $x$ the coordinates of the points $g = (y^1, \dots, y^m)$ and $z = (z^1, \dots, z^{k-m})$. The coordinates $(z^1, \dots, z^{k-m})$ determine coordinates on the quotient space $X/G$. In this event, the tangent vectors to $Z$ are tangent vectors to $X/G$. If the submanifold $Z$ is chosen in a different way, then we obtain other coordinates in a neighborhood of $Gx \in X/G$ and other representatives of the tangent vectors to $X/G$. The details can be found in [4].

**4 The Poisson action of a group and symplectic reduction.** We say that functions $F$ and $G$ on a Poisson manifold are *in involution* if their Poisson bracket is zero:

$$\{F, G\} = 0.$$

Suppose that we have $m$ functions $F_1, \dots, F_m$ on a symplectic manifold which are not necessarily in pairwise involution. This means that they generate a noncommutative Lie algebra $\boldsymbol{g}$ with respect to the Poisson bracket. We consider the case when they determine a linearly independent basis for this algebra $\boldsymbol{g}$:

$$\{F_i, F_j\} = -c_{ij}^k F_k.$$

These functions considered as the Hamiltonian determine $m$ Hamiltonian flows with the vector fields $v_1 = v_{f_1}, \dots, v_m = v_{f_m}$ which do not commute but satisfy the relation

$$[v_i, v_j] = c_{ij}^k v_k$$

(see Lemma 9.2). Translations along these flows generate the action of some connected Lie group $G$ on $M$ with the following properties:

(1) the algebra $\boldsymbol{g}$ is the Lie algebra of the group $G$ and its elements $a \in \boldsymbol{g}$ are identified with the vector fields $a_1 v_1 + \dots + a_m v_m$.

Moreover, since translation along each flow preserves the symplectic form and each element of $G$ is the composition of such translations, we have

(2) $G$ acts on $M$ by the canonical transformations (or *simplectomorphisms*, i.e., diffeomorphisms $\varphi\colon M \to M$ preserving the form $\omega$: $\varphi^*\omega = \omega$):

$$M \xrightarrow{g} M, \quad g^*\omega = \omega, \quad g \in G.$$

We arrive at the following important notion: the action of the group $G$ on a symplectic manifold is called a *Poisson action* if it satisfies condition (2).

This definition is more general than the case under consideration: it does not assume that, for each element $a \in \boldsymbol{g}$, the one-parameter subgroup $e^{at}$ is a group of translations along the trajectories of the Hamiltonian flow; it is sufficient that the vector field $v(x) = \frac{d}{dt} e^{at}(x)$ is locally Hamiltonian.

In our situation with the Poisson action we associate the *moment map*:

$$\mathcal{M}\colon M \to \boldsymbol{g}^*,$$

which relates to each point $x \in M$ the following linear function on the Lie algebra $\boldsymbol{g}$:

$$[\mathcal{M}(x)](a) = a_1 F_1(x) + \dots + a_m F_m(x), \quad a = a_1 v_1 + \dots + a_m v_m \in \boldsymbol{g}.$$

**Lemma 9.5.** *The equality*

$$\mathcal{M}(gx) = \mathrm{Ad}_g^* (\mathcal{M}(x)) \tag{9.23}$$

*holds for all $g \in G, x \in M$; i.e., the moment map is equivariant under the coadjoint representation of the group $G$.*

*Proof.* It suffices to consider the case of a one-parameter subgroup $g = e^{bt}$, $b \in \boldsymbol{g}$. By definition,

$$[\mathcal{M}(e^{bt}x)](a) = a_1 F_1(e^{bt}x) + \cdots + a_m F(e^{bt}x).$$

Differentiate both sides of the equality with respect to $t$:

$$\begin{aligned}\frac{d}{dt}[\mathcal{M}(e^{bt}x)](a) &= a_1\{F_1, F_b\}(e^{bt}x) + \cdots + a_m\{F_m, F_b\}(e^{bt}x)\\ &= \{F_a, F_b\}(e^{bt}x) = -F_{[a,b]}(e^{bt}x) = F_{\mathrm{ad}_b^*(a)}(e^{bt}x),\end{aligned}$$

where $F_a = a_1F_1 + \cdots + a_mF_m$ and $F_b = b_1F_1 + \cdots + b_mF_m$. We have thus shown that the derivatives of the left- and right-hand sides of (9.23) coincide for all $t$. Hence, observing that the identity is obviously valid at $t = 0$, we immediately obtain the assertion of the lemma. □

Denote the inverse image of the moment map at $\tau \in \boldsymbol{g}^*$ by

$$\mathcal{M}^{-1}(\tau) = M_\tau.$$

**Corollary 9.8.** (1) *$M_0$ is invariant under the action of the group $G$.*

(2) *If the group $G$ is commutative then, for every $\tau \in \boldsymbol{g}^*$, the set $M_\tau$ is invariant under the action of the group $G$.*

**Theorem 9.9.** *Suppose that $M_0$ is a smooth submanifold in $M$ and a group $G$ is compact and acts freely on $M_0$. Then the quotient space $M_{\mathrm{red}} = M_0/G$ carries the natural symplectic structure.*

*Proof.* Introducing a smooth structure on the quotient space, we introduce coordinates $z^1, \dots, z^{k-m}$, $k = \dim M - m = \dim M_0$, as coordinates on some patch, a submanifold $Z$. Define the antisymmetric product on the tangent space of $Z$ by the usual formula

$$\omega'(\xi, \eta) = \omega(\xi, \eta), \tag{9.24}$$

where $\omega$ is the symplectic form on the initial manifold $M$ and

$$\xi, \eta \in T_z Z \subset T_z M_0 \subset T_z M.$$

It would be natural to define the symplectic structure on $M_{\mathrm{red}}$ this way, especially as the form $\omega'$ is closed as the restriction of the closed form $\omega$ to the submanifold $Z$. But we should justify this definition, by proving that it is independent of

(1) the choice of coordinates on $M_{\mathrm{red}}$;

(2) the choice of the point $x$ on the orbit.

Also, we should demonstrate that

(3) the form $\omega'$ is nondegenerate.

We now prove these three assertions. First of all we give some auxiliary assertion:

**Lemma 9.6.** *The spaces $T_x M_0$ and $T_x(Gx)$ are orthogonal complements of each other in $T_x M$ with respect to the form $\omega$.*

*Proof.* Since the group $G$ acts by Hamiltonian transformations, the tangent vector $\eta$ to the orbit has the form $v_f$ for some function of the form $f = a_1 F_1 + \cdots + a_m F_m$:

$$\eta^j = v_f^j = h^{jk} \frac{\partial f}{\partial x^k}.$$

We have

$$\omega(\xi, \eta) = h_{ij} \xi^i \eta^j = h_{ij} \xi^i h^{jk} \frac{\partial f}{\partial x^k} = \delta_i^k \xi^i \frac{\partial f}{\partial x^k} = \xi^i \frac{\partial f}{\partial x^i} = \partial_\xi f,$$

i.e., $\omega(\xi, \eta)$ is the derivative of the function $f$ in the direction of $\xi$. Therefore, $\xi$ lies in the orthogonal complement to $T_x(Gx)$ if and only if $\partial_\xi \mathcal{M} = 0$, but this condition determines the tangent space of $M_0$. We have thus demonstrated that

$$(T_x(Gx))^\perp = T_x M_0.$$

It is easy to check the converse assertion

$$(T_x M_0)^\perp = T_x(Gx).$$

The lemma is proven. □

We turn to proving correctness of the definition of the symplectic structure on $M_{\text{red}}$:

(1) Actually, the change of coordinates consists in either changing coordinates on $Z$ or changing of $Z$ itself. We only need to consider the second case, since in this case the tangent vectors in $TM_0$ representing tangent vectors to $M_{\text{red}}$ vary:

$$\xi \to \xi + \xi', \quad \eta \to \eta + \eta',$$

where the vectors $\xi'$ and $\eta'$ touch the orbits of the group $G$. By the previous lemma, the value of the symplectic product (9.24) does not vary upon this change:

$$\omega(\xi + \xi', \eta + \eta') = \omega(\xi, \eta).$$

(2) Since the action of the group $G$ is Hamiltonian, it preserves the symplectic product:

$$\omega(\xi, \eta) = \omega(g^* \xi, g^* \eta);$$

therefore, the definition of form (9.24) is independent of the choice of the point $x$ on the orbit.

(3) If $\omega(\xi, \eta) = 0$ for all vectors $\eta \in T_x Z$, then $\xi$ is orthogonal to all vectors in $T_x M_0$ and, by Lemma 9.6, it touches the orbits of the group $G$, i.e., represents the zero vector in $T_x M_{\text{red}}$.

The theorem is proven. □

Let $\tau \in \boldsymbol{g}^*$ and $\tau \neq 0$. Denote by $G_\tau$ the subgroup in $G$ constituted by all elements $g$ which do not change $\tau$:

$$\mathrm{Ad}_g^*(\tau) = \tau.$$

**Corollary 9.9.** *Suppose that $M_\tau$ is a submanifold and a group $G_\tau$ is compact and acts freely on $M_\tau$. Then the quotient space*

$$M_\tau / G_\tau$$

*carries the natural symplectic structure.*

*Proof.* Denote by $O_\tau$ the orbit of $\tau \in \boldsymbol{g}^*$ under the adjoint representation $\mathrm{Ad}^*$ of the group $G$. Consider the symplectic form $\omega_\tau$ on $O_\tau$ as on a symplectic leaf of the Lie–Poisson bracket (see §9.4). The adjoint representation determines the Poisson action of the group $G$ on $O_\tau$ under which the moment map coincides with the embedding $O_\tau \subset \boldsymbol{g}^*$. Consider the symplectic form $\omega_M - \omega_\tau$ (i.e., the sum of two forms on the factors) on the direct product $M \times O_\tau$. Then $M_\tau / G_\tau = (M \times O_\tau)_{\mathrm{red}}$. The corollary is proven. □

Note that all arguments in the proof of the theorem are actually local and therefore in the general case we can consider a local reduction which simplifies the system in some domain. Moreover, the situation when the action of $G$ is free and we obtain the manifold $M_{\mathrm{red}}$ is rather rare, meanwhile the converse holds: the group $G$ contains finite subgroups which leave some points fixed. Eventually, under factorization, to these points there correspond conic singularities in $M_{\mathrm{red}}$. However, away from these singularities we obtain a nice symplectic manifold.

If the Hamiltonian $H$ is in involution with the functions $F_1, \dots, F_m$, then the corresponding Hamiltonian system reduces to a Hamiltonian system on the $2(n-m)$-dimensional symplectic manifold $M_{\mathrm{red}}$, the *reduced phase space* with the symplectic form $\omega_{\mathrm{red}}$. Since $\{H, F_i\} = 0$, the Hamiltonian $H$ is conserved upon translations along the Hamiltonian flows given by the Hamiltonians $F_1, \dots, F_m$; consequently, it is conserved under the action of $G$ and the projection $\pi\colon M_0 \to M_{\mathrm{red}} = M_0/G$. Therefore, the initial Hamiltonian system reduces to the system on $M_{\mathrm{red}}$ with the *reduced Hamiltonian* $H_{\mathrm{red}}(x) = H(\pi^{-1}(x))$.

Moreover, if a function $f$ is $G$-invariant, then it descends to the function $f(x) = f_{\mathrm{red}}(\pi(x))$ to the quotient space, and the variation of $f$ along the trajectories of the flow is described by the equations

$$\dot{f} = \{f, H\} = \{f_{\mathrm{red}}, H_{\mathrm{red}}\}_{\mathrm{red}},$$

where the Poisson bracket $\{\cdot, \cdot\}_{\mathrm{red}}$ corresponds to the form $\omega_{\mathrm{red}}$.

This method of a reduction of a Hamiltonian system $(M, H)$ to the Hamiltonian system $(M_{\mathrm{red}}, H_{\mathrm{red}})$ is called a *symplectic reduction*. For concrete examples it was known for a long time and it has been used in various problems of mechanics. It applies when the Hamiltonian system admits a symmetry group and, as in many concrete examples, the action of this group is Poisson.

Before turning to examples, we make one remark: there exist interesting problems in which, on the level surface of a first integral $F$, the gradients of the functions $F$ and $H$ are not linearly independent everywhere. To the point, where grad $F$ and grad $H$ are linearly dependent, there corresponds a fixed point (equilibrium) of the reduced Hamiltonian system, however this point is not an equilibrium state of the initial system. A trajectory projected into an equilibrium state on $M_{\text{red}}$ is called a *relative equilibrium* of the initial system. For example,

- *if the Hamiltonian is invariant under the Poisson action of a group $G$, then the relative equilibria are exactly the orbits of the one-parameter subgroups of $G$.*

## 5 Examples

**(a) Motion in a central field.** Suppose that $M = T^*\mathbb{R}^3 = \mathbb{R}^3_x \times \mathbb{R}^3_p$ is the cotangent bundle to $\mathbb{R}^3$, $x^1$, $x^2$, and $x^3$ are the Euclidean coordinates in $\mathbb{R}^3$, and $p_1$, $p_2$, and $p_3$ are coordinates in the momentum space. Consider the Hamiltonian

$$H(x, p) = \frac{p^2}{2m} + U(r),$$

which describes the motion of a particle of mass $m$ in a central field $U(r)$, $r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$. For example, for $U(r) = \frac{\alpha}{r}$, $\alpha = \text{const} < 0$, we obtain *Kepler's problem* on the motion of a planet around the Sun (with the other forces neglected). The Hamiltonian is invariant under the Poisson action of the rotation group $G = \mathrm{SO}(3)$. In this case the Lie algebra $\boldsymbol{g}$ is generated by the components of the kinetic momentum vector:

$$F = (F_1, F_2, F_3), \quad F = \mu = [x \times p].$$

The level surface $M_\tau = \{\mu = \tau\}, \tau \neq 0$, has dimension 3 and the group $G_\tau$ is one-dimensional. Therefore, the dimension of the reduced phase space equals 2. We now give explicit formulas.

Executing a change of variables in $\mathbb{R}^3$, we can reduce everything to the case $\tau = (0, 0, c)$. The level surface $M_\tau$ is given by the equations

$$x^3 = p_3 = 0, \quad x^1 p_2 - x^2 p_1 = c.$$

The group $G_\tau$ is the group SO(2) constituted by all rotations around the $Ox^3$-axis. We introduce the polar coordinates $r$ and $\varphi$ in the $(x^1, x^2)$-plane and the conjugate variables $p_{\boldsymbol{r}}$ and $p_\varphi$ in the momentum space:

$$x^1 = r\cos\varphi, \quad x^2 = r\sin\varphi,$$

$$p_1 = p_r \cos\varphi - \frac{p_\varphi}{r}\sin\varphi, \quad p_2 = p_r \sin\varphi + \frac{p_\varphi}{r}\cos\varphi,$$
$$\{r, p_r\} = \{\varphi, p_\varphi\} = 1,$$

with the zero Poisson brackets for the other collections of coordinates. In the new coordinates, the surface $M_\tau$ is given by the equations

$$x^3 = p_3 = 0, \quad p_\varphi = c.$$

The group $G_\tau$ acts by rotations: $\varphi \to \varphi + \theta$, $0 \le \theta \le 2\pi$, and the factorization $M_\tau \to M_\tau / G_\tau$ reduces to elimination of the angular variable $\theta$. Eventually, we obtain the reduced phase space $M_{\text{red}}$ with the coordinates $r$ and $p_r$, the symplectic form $dp_r \wedge dr$, and the reduced Hamiltonian

$$H_{\text{red}}(r, p_r) = \frac{1}{2}\left(p_r^2 + \frac{c^2}{r^2}\right) + U(r).$$

**(b) Kepler's problem.** ($U = \frac{\alpha}{r}, \alpha < 0$) has extra first integrals, namely the components of the Laplace–Runge–Lenz vector

$$\nu = \left[\frac{p}{m} \times \mu\right] + \frac{\alpha x}{|x|} = \left[\frac{p}{m} \times [x \times p]\right] + \frac{\alpha x}{|x|}.$$

**Problem 9.2.** (1) Prove that the inner product of the kinetic momentum vector and the Laplace–Runge–Lenz vector equals zero:

$$(\mu, \nu) = 0,$$

but the components of these vectors at a generic point give five functionally independent first integrals of the Kepler problem.

(2) Prove that the Poisson brackets of the components of the vectors $\mu$ and $\nu$ are equal to

$$\{\mu_i, \mu_j\} = \varepsilon_{ijk}\mu_k, \quad \{\mu_i, \nu_j\} = \varepsilon_{ijk}\nu_k, \quad \{\nu_i, \nu_j\} = -\frac{2E}{m}\varepsilon_{ijk}\mu_k,$$

where $E = H = \frac{p^2}{2m} + \frac{\alpha x}{|x|}$ is the energy of the system. Derive that, at different energy level surfaces, the algebra of the integrals $\mu_i$ and $\nu_j$ closes and is isomorphic to

(a) the Lie algebra of the group SO(4) for $E < 0$;

(b) the Lie algebra of the group of motions of the three-dimensional space for $E = 0$;

(c) the Lie algebra of the group SO(1, 3) for $E > 0$.

In particular, the common level surface $\mu = \mu_0 = \text{const}, \nu = \nu_0 = \text{const}$ is one-dimensional and is exactly the trajectory of Kepler's problem.

**(c) Orbits of the coadjoint representation of Lie groups.** Let $M = T^*G$, where $G$ is a Lie group which acts on itself (and thereby on $T^*G$) by left translations. This action is Poisson and the corresponding Hamiltonians are linear in the momentum and have the form

$$F_a(x, p) = p(ag) = p\big(\tfrac{de^{at}}{dt}\big|_{t=0}g\big), \quad a \in \boldsymbol{g},$$

where $p(v)$ is the value of a covector (linear function) $p$ at a vector $v$. The expression $ag$ stands for the right translation $R_g^* a$ of a vector $a \in \boldsymbol{g} = T_1 G$ to a point $g = x \in G$, and we have

$$p(R_g^* a) = R_{g-1}^* p(a)$$

(see §9.4). Therefore, the moment map for this action has a very simple form:

$$\mathcal{M}(g) = R_g^*(p) \in \boldsymbol{g}^* = T_1^* G.$$

The inverse image $\mathcal{M}^{-1}(\tau)$ of the moment map at a point $\tau \in \boldsymbol{g}^*$ is diffeomorphic to the group $G$ and is constituted by points of the form $(g, R_{g-1}^* \tau) \in T^*G$. Consider the map

$$M_\tau \xrightarrow{L_g^*} \boldsymbol{g}^* \colon (g, p) \to L_g^* p = L_g^* R_{g-1}^* \tau.$$

The image of this map coincides with the orbit of the point $\tau$ in the coadjoint representation and the fiber coincides with the orbit of the action of the group $G_\tau$ constituted by all elements $g \in G$ such that

$$L_g^* R_{g-1}^* \tau = \operatorname{Ad}_g^* \tau = \tau.$$

Thus, we obtain the reduced phase space $M_{\text{red}}$ diffeomorphic to the orbit of the point $\tau$ in the coadjoint representation. The constructed symplectic structure is the same as the one on this orbit as a symplectic leaf of the Lie–Poisson bracket on $\boldsymbol{g}^*$ (see §9.4).

## 9.7 Euler's equations

On the cotangent bundle to a Lie group $G$, consider a Hamiltonian system invariant under the left translations.

For example, assume that a left-invariant metric is given on a Lie group. This means that the tangent bundle is endowed with a metric such that

$$(\xi, \eta) = (L_g^* \xi, L_g^* \eta) \quad \text{for all } g \in G$$

and for all vectors $\xi$ and $\eta$ tangent to $G$ at the same point. It is obvious that every such metric is determined by the inner product at the identity $1 \in G$. In this case the inner product of tangent vectors at an arbitrary point $g \in G$ is defined as

$$(\xi, \eta) = (L_{g-1}^* \xi, L_{g-1}^* \eta).$$

The geodesic flow is invariant under the left translations on the Lie group, since such is its Hamiltonian.

The left-invariant Hamiltonian flow reduces to the orbits of the coadjoint action of $G$ on $\boldsymbol{g}^*$ (see Example (c) in §9.6). Since these orbits are symplectic leaves (see §9.4), Hamilton's equations take the form

$$\dot{x} = \{x, H\},$$

where the right-hand side is the Lie–Poisson bracket of $x \in \boldsymbol{g}^*$ and the reduced Hamiltonian $H = H_{\text{red}}$ of the geodesic flow. For the geodesic flow of the left-invariant metric generated by the metric $g_{ik}$ on $\boldsymbol{g}$, we have

$$H(x) = \frac{1}{2} g^{ik} x_i x_k.$$

In coordinates, Hamilton's equations take the form

$$\dot{x}_i = \frac{\partial H}{\partial x^j}\{x_i, x_j\} = c_{ij}^k x_k \frac{\partial H}{\partial x^j}.$$

Consider $\dot{x}$ as a linear functional on $\boldsymbol{g}$:

$$\begin{aligned}\dot{x}(\xi) &= c_{ij}^k x_k \frac{\partial H}{\partial x^j} \xi^i \\ &= x_k \left( c_{ij}^k \xi^i \frac{\partial H}{\partial x^j} \right) = x([\xi, \operatorname{grad} H]) = -x(\operatorname{ad}_{\operatorname{grad} H} \xi) = -\operatorname{ad}^*_{\operatorname{grad} H}(x)(\xi),\end{aligned}$$

where $\operatorname{grad} H \in \boldsymbol{g}$ is the gradient of the function $H$ on $\boldsymbol{g}^*$. We have obtained another form of Hamilton's equations:

$$\dot{x} = -\operatorname{ad}^*_{\operatorname{grad} H}(x). \tag{9.25}$$

Since in the case of the left-invariant metric on the group $G = \mathrm{SO}(3)$ we obtain Euler's equations of motion of a rigid body around a fixed point, the general equation (9.25) with the quadratic Hamiltonian $H$ is called *Euler's equation* on $\boldsymbol{g}^*$.[10]

Suppose that a nondegenerate inner product is given on a Lie algebra and this product is invariant under the adjoint representation. As usual, the nondegeneracy of the product allows us to identify $\boldsymbol{g}$ and its dual space $\boldsymbol{g}^*$ by the rule

$$x(\xi) = (x, \xi).$$

The invariance under $\mathrm{Ad}_g$ implies the identity

$$([z, x], y) + (x, [z, y]) = 0 \quad \text{for all } x, y, z \in \boldsymbol{g}.$$

[10]For a detailed derivation and analysis of Euler's equations, in particular, for infinite-dimensional Lie algebras see Supplement 2 to [1].

Indeed, we have

$$(\mathrm{Ad}_{e^{zt}}\, x, \mathrm{Ad}_{e^{zt}}\, y) = (x, y), \quad t \in \mathbb{R},$$

and therefore

$$\frac{d}{dt}(\mathrm{Ad}_{e^{zt}}\, x, \mathrm{Ad}_{e^{zt}}\, y)\Big|_{t=0} = (\mathrm{ad}_z\, x, y) + (x, \mathrm{ad}_z\, y) = ([z, x], y) + (x, [z, y]) = 0.$$

Now, rewrite Euler's equation as an equation on the Lie algebra $\boldsymbol{g}$:

$$(\dot{x}, y) = (x, [y, \operatorname{grad} H]) = -([\operatorname{grad} H, x], y),$$

and finally

$$\dot{x} = [x, \operatorname{grad} H], \quad x \in \boldsymbol{g}.$$

Consider important examples when $G$ is the Lie group $E(3)$ constituted by all (orientation-preserving) motions of the three-dimensional Euclidean space:

$$x \to Ax + b, \quad A \in \mathrm{SO}(3),\ b \in \mathbb{R}^3.$$

**Problem 9.3.** Prove that

(1) the Lie algebra $\boldsymbol{g} = e(3)$ of the group $E(3)$ is generated by the elements $M_i$ and $p_k$, $i, k = 1, 2, 3$, where $M_i$ is the generator of rotations around the $Ox^i$-axis and $p_k$ is the generator of translations along the $Ox^k$-axis; moreover, the following commutation relations hold:

$$[M_i, M_j] = \varepsilon_{ijk} M_k, \quad [M_i, p_j] = \varepsilon_{ijk} p_k, \quad [p_i, p_j] = 0 \tag{9.26}$$

(the repeated index $k$ implies summation).

(2) the Lie–Poisson bracket on $e(3)^*$ has two Casimirs:

$$I_1 = p^2 = p_1^2 + p_2^2 + p_3^2, \quad M_2 = (M, p) = ps = M_1 p_1 + M_2 p_2 + M_3 + p_3.$$

The level surface

$$M_{p,s} = \{I_1 = p^2,\ I_2 = ps\} \subset e(3)^*$$

is diffeomorphic to the cotangent bundle of the two-dimensional sphere $S^2$. The symplectic structure is given by the form

$$\omega = d\xi_i \wedge dq^i + sf(q^1, q^2)dq^1 \wedge dq^2 = d\xi_i \wedge dq^i + sF,$$

where $q^1$ and $q^2$ are the coordinates on the sphere, $\xi_1$ and $\xi_2$ are the momentum coordinates, and $F$ is a closed 2-form on $S^2$. Moreover,

$$\int_{S^2} F = 4\pi.$$

Consequently, for $s \neq 0$ we obtain the magnetic Poisson bracket on $T^*S^2$.[11]

The *equations of motion of a heavy rigid body around a fixed point* correspond to the Hamiltonian

$$H(M, p) = \sum_{i,k} a_{ik} M_i M_k + \sum_k l_k p_k = (AM, M) + l(p)$$

which is quadratic in the variable $M$ and linear in the variable $p$.

The *Kirchhoff equations* which describe the motion of a rigid body in an ideal fluid are Euler's equations corresponding to the Hamiltonian quadratic in the variables $M$ and $p$:

$$H(M, p) = \sum_{ik,k} (a_{ik} M_i M_k + b_{ik} M_i p_k + c_{ik} p_i p_k),$$

where $a_{ik} = a_{ki}$ and $c_{ik} = c_{ki}$. Since the Hamiltonian in the Kirchhoff problem has the meaning of the kinetic energy, it is assumed to be positive definite.

Executing the changes of coordinates in $\boldsymbol{g}$,

$$M \to \mathrm{Ad}_g M, \quad p \to g(p), \quad g \in \mathrm{SO}(3),$$

preserving the form of bracket (9.26), we can reduce the Hamiltonian of the equations of motion of a rigid body around a fixed point to the form

$$H(M, p) = \sum_i a_i M_i^2 + \sum l_k p_k,$$

and the Hamiltonian of the Kirchhoff equations to the form

$$H(M, p) = \sum_i a_i M_i^2 + \sum_{i,k} \left( \frac{b_{ik}}{2} (M_i p_k + M_k p_i) + c_{ik} p_i p_k \right), \quad c_{ik} = c_{ki}.$$

Obviously, it follows from the fact that every positive definite quadratic form in $\mathbb{R}^n$ can be reduced to the diagonal form by an orthogonal transformation (in our situation this is the Gram–Schmidt orthogonalization of the form $A(M, M)$).

Note that the appearance of the group $E(3)$ in these two problems is of absolutely different nature.

(a) The Kirchhoff equations are the equations for geodesics of a left-invariant metric on $E(3)$.

(b) In the case of the equations of motion of a heavy rigid body around a fixed point, choose an orthonormal coordinate system connected with the body with the origin at a fixed point. In this coordinate system, $v = -M$ is the kinetic momentum vector, while $p$ is the unit vector in the direction of the gravitational force.[12] In this problem one of the Casimirs equals 1 by definition, $I_1 = p^2 = 1$, and the other equals $I_2 = ps$.

[11] For the proofs of these assertions see Supplement 1 (by S. P. Novikov) to [2], III.

[12] If the coordinate axes are directed along the momentum axes of the body and $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ are the

## 9.8 Integrable Hamiltonian systems

The modern definition of the integrability of finite-dimensional systems is based on the following result known as the Liouville theorem on integrable systems.

**Theorem 9.10.** *Let $M$ be a symplectic manifold of dimension $2n$ on which a Hamiltonian system with a Hamiltonian $H$ is given. Suppose that we have $n$ first integrals $I_1, \dots, I_{n-1}, I_n = H$ of this system which are in involution, $\{I_j, I_k\} = 0$, $j, k = 1, \dots, n$, and functionally independent almost everywhere. Denote by $\mathcal{M}\colon M \to \mathbb{R}^n$ the moment map: $\mathcal{M}(x) = (I_1(x), \dots, I_n(x))$.*

*Let $X$ be a component of the set $\mathcal{M}^{-1}(\tau)$ such that $I_1, \dots, I_n$ are functionally independent on $X$ and $X$ is compact.*

*Then $X$ is diffeomorphic to the $n$-dimensional torus $T^n$ and has a neighborhood $U$ such that*

(1) *$U$ is diffeomorphic to the direct product of the $n$-dimensional open ball $D^n$ and $T^n$;*

(2) *the action–angle coordinates $s_1, \dots, s_n$ and $\varphi_1, \dots, \varphi_n$ are given in $U$; moreover, the level surfaces $s = (s_1, \dots, s_n) = \mathrm{const}$ are invariant tori of the Hamiltonian system and the angular coordinates $\varphi_1, \dots, \varphi_n$ are coordinates on these tori defined modulo $2\pi$; furthermore, the symplectic form takes the form*

$$\omega = \sum_k ds_k \wedge d\varphi_k;$$

(3) *in the action–angle coordinates, the Hamiltonian system takes the form*

$$\dot{s}_j = 0, \quad \dot{\varphi}_k = \psi_k(s_1, \dots, s_n), \quad j, k = 1, \dots, n.$$

We now make some remarks:

(1) By the words "almost everywhere" ("at a point of generic position", "at a generic point") we mean usually "on a set of full measure"; i.e., the set on which the condition is violated has measure zero with respect to some measure $d\mu$ which in a neighborhood

corresponding momenta of inertia, then making the substitution

$$M = (-Ap, -Bq, -Cr), \quad p = (\gamma, \gamma', \gamma''),$$
$$(a_1, a_2, a_3) = \Big(\frac{1}{A}, \frac{1}{B}, \frac{1}{C}\Big), \quad l(p) = -mg(x_0\gamma + y_0\gamma' + z_0\gamma'')$$

in the equation

$$\dot{M} = \{M, H\}, \quad \dot{p} = \{p, H\}$$

we obtain the equations of motion in terms typical for a monograph on analytic mechanics. Here $\omega = (p, q, r)$ is the angular velocity vector, $m$ is the mass of the body, $g$ is the free fall acceleration, $Q = (x_0, y_0, z_0)$ are the coordinates of the center of mass of the body. For $Q = 0$ the equations $\dot{M} = \{M, H\}$ coincide with the usual Euler equations (see §9.3), since the gravitational force is not taken into account in the Hamiltonian.

of every point takes the form $d\mu = \rho(x)dx^1 \wedge \dots \wedge dx^{2n}$ where $\rho(x)$ is a positive continuous function. Under the conditions of the theorem, on the common level surface $\{I_1 = c_1, \dots, I_n = c_n\}$, the functions $I_1, \dots, I_n$ are functionally independent and the surface itself is a submanifold (by the implicit function theorem).

(2) The Hamiltonian system is (*Liouville*) *integrable* (or *completely integrable*) if it satisfies the conditions of the theorem. In classical mechanics of the 19th century one usually considered the case of *analytically integrable systems* in which the first integrals $I_1, \dots, I_n$ are real-analytic functions.

(3) The theorem is ineffective without the procedure of construction of the action-angle coordinates which is available, but will be only briefly mentioned in the proof. With this procedure we can explicitly solve Hamilton's equations in a neighborhood $U$ of the torus $X$: on each invariant torus, it reduces to a linear equation with constant coefficients and the motion over the torus is conditionally periodic:

$$\varphi(t) = \varphi(0) + \psi(s)t.$$

*Proof.* (1) First of all prove we the following assertion:

**Lemma 9.7.** *$X$ is diffeomorphic to the $n$-dimensional torus.*

*Proof of the lemma.* The Hamiltonian flows $v_1 = v_{I_1}, \dots, v_n = v_{I_n}$ touch the level surface $X$, since all values of the integrals $I_1, \dots, I_n$ are conserved upon translation along these flows. Moreover,

$$[v_{I_j}, v_{I_k}] = -v_{\{I_j, I_k\}} = 0, \quad j, k = 1, \dots, n;$$

consequently, these flows commute. Since the manifold $X$ is compact, every ordinary differential equation of the form $\dot{x} = v_i(x), i = 1, \dots, n$, has a global solution, i.e., a solution for all values of $t$.

The translations along the trajectories of these flows generate the action of the group $\mathbb{R}^n$ on $X$ which relates to each point $x \in X$ its translation $\varphi_t(x)$ for a unit time along the trajectory of the flow $t_1 v_1 + \dots + t_n v_n$, $t \in \mathbb{R}^n$. For small $t$ this map has the form

$$x \to x + t_1 v_1 + \dots + t_n v_n + O(|t|^2);$$

therefore, the Jacobian of the map $t \to \varphi_t(x)$ at $t = 0$ is nonzero for each point $x \in X$. Consequently, by the inverse function theorem, for $|t| < \varepsilon(x)$ the map $t \to \varphi_t(x)$ is a diffeomorphism of the open ball $|t| < \varepsilon(x)$ onto its image.

If $x$ and $y$ lie in $X$, then join them by a curve $\gamma$ and for each point $z$ of the curve take a neighborhood which is covered diffeomorphically under the map $t \to \varphi_t(z)$. The intersections of these neighborhoods with $\gamma$ constitute an open covering of $\gamma$ from which we can choose a finite subcovering $U_1, \dots, U_k$, $x \in U_1$, $y \in U_k$, $U_i \cap U_{i+1} \neq \emptyset$, $i = 1, \dots, n-1$. Put $x_0 = x$ and $y = x_k$ and choose points $x_i \in U_i \cap U_{i+1}$. We see that $x_{i+1} = \varphi_{t_i}(x_i))$ at some $t_i$, $i = 1, \dots, k-1$. Consequently, $y = \varphi_{t_{k-1}} \dots \varphi_{t_0}(x)$. Since the points $x, y \in X$ were chosen arbitrary, we conclude that

$$\{\varphi_t(x) \mid t \in \mathbb{R}^n\} = X \quad \text{for every point } x \in X;$$

i.e., $X$ is the homogeneous space of the group $\mathbb{R}^n$:

$$X = \mathbb{R}^n / \Gamma.$$

Since the map $t \to \varphi_t(x)$ is a diffeomorphism for small $t$, for nonzero vectors $\gamma$ in $\Gamma$, the value $|\gamma|$ is bounded from below and we take $e_1$ to be the vector in $\Gamma$ with the least norm. Take $e_2$ to be the vector in $\Gamma$ lying at the least nonzero distance to the straight line $\mathbb{R}e_1$ and continue this process, taking at each step the vector $e_i$ to be the vector in $\Gamma$ lying at the least distance to the subspace $\mathbb{R}e_1 + \dots + \mathbb{R}e_{i-1}$. Eventually, we construct a basis $e_1, \dots, e_n$ for the group $\Gamma = \mathbb{Z}^n$. Therefore,

$$X = \mathbb{R}^n / \mathbb{Z}^n = T^n.$$

The lemma is proven. □

(2) Prove that some small neighborhood $U$ of the torus $X$ foliates into invariant tori over the $n$-dimensional disk.

Since the covectors grad $I_1, \dots,$ grad $I_n$ are linearly independent at each point $x \in X$, by the implicit function theorem, there is a neighborhood $U_x$ diffeomorphic to the direct product $V_x \times D_x$ with coordinates $(y, z)$ such that $y$ are local coordinates on $X$ and $\mathcal{M}(y, z) = z$. Cover $X$ with the neighborhoods $V_x$ and choose a finite subcovering $V_{x_1}, \dots, V_{x_l}$. Let $D = \cap_i D_{x_i} \subset \mathbb{R}^n$. The union $U = \cup_i V_{x_i} \times D$, where $V_{x_i} \times D \subset V_{x_i} \times D_{x_i} = U_{x_i}$, is the sought neighborhood whose closure is compact and therefore each component of $\mathcal{M}^{-1}(a)$, $a \in D$, is compact and consequently homeomorphic to the torus.

We obtain a natural diffeomorphism of $U = X \times D$ under which the moment map is the projection to the second factor.

(3) Construct the action–angle coordinates. We introduce the angular variables $\varphi_1, \dots, \varphi_n$ on the invariant tori in $U$ defined modulo $2\pi$ and take the values of the first integrals $I_1, \dots, I_n$ to be the coordinates in $D$.

As in the proof of Darboux's theorem, we can take the coordinates on the torus to be the times of translations along the trajectories of the Hamiltonian flows $v_1, \dots, v_n$. Denote these times by $t_1, \dots, t_n$. Obviously, we obtain

$$\{t_j, I_k\} = \delta_{jk}, \quad \{I_j, I_k\} = \{t_j, t_k\} = 0. \tag{9.27}$$

On each torus, the angular variables $\varphi$ are obtained from the time variables $t$ by a linear change depending on the torus. Therefore, it follows from (9.27) that in the variables $I$ and $\varphi$ the symplectic form has the shape

$$\omega = \sum_{j<k} a_{jk}(I, \varphi) dI_j \wedge dI_k + \sum_{j,k} b_{jk}(I) dI_j \wedge d\varphi_k. \tag{9.28}$$

In particular, the restrictions of $\omega$ to the invariant tori are zero and in the neighborhood $U$ we have a 1-form $\alpha$ such that

$$d\alpha = \omega.$$

This assertion follows from the following topological reasons: the restriction of $\omega$ to $U$ realizes the zero cohomology class and the exterior derivative $d$ on $\omega$ is invertible: we do not prove this, referring the reader to [4]. For $M = T^*N$ take $\alpha = p_i dx^i$ (in this form the derivation of formulas for the action–angle variables often appears in physical literature). Now, put

$$s_k = \frac{1}{2\pi}\int_{\gamma_k} \alpha, \quad k = 1, \dots, n.$$

Here $\gamma_k$ is the contour on the torus traversed as $\varphi_k$ varies from 0 to $2\pi$ (the other variables $\varphi_i$ do not vary). For the cotangent bundle we obtain the following classical formula from physics literature:

$$s_k = \frac{1}{2\pi}\int_{\gamma_k} p\, dx, \quad k = 1, \dots, n.$$

Show that the functions $s_k$, $k = 1, \dots, n$, are independent of $\varphi$. Take another contour $\gamma'$ which together with $\gamma_k$ bounds a surface $S$ on the torus. Then, using the Stokes theorem, we obtain

$$\int_{\gamma_k} \alpha - \int_{\gamma'} \alpha = \int_{\partial S} \alpha = \int_S d\alpha = \int_S \omega = 0.$$

It follows from (9.28) that

$$\alpha = \sum_k \left( \int \sum_j b_{jk}(I) dI_k \right) d\varphi_k + \sum_k f_j dI_j + df_0.$$

Only the first summand has nonzero contribution in the integral over $\gamma_k$ and we obtain

$$ds_k = \sum_j b_{jk} dI_j, \quad \frac{\partial s_k}{\partial I_j} = b_{jk}(I_1, \dots, I_n).$$

Since the form $\omega$ is nondegenerate, the matrix $b_{jk}$ is invertible and, by the inverse function theorem, there is a change of coordinates

$$I_1, \dots, I_n \to s_1, \dots, s_n, \quad s = s(I)$$

(here we might need to replace $U$ with a smaller domain which still foliates into invariant tori). We obtain

$$\omega = \sum_{j<k} \hat{a}_{jk} ds_j \wedge ds_k + \sum_{j,k} \hat{b}_{jk} ds_j \wedge d\varphi_k.$$

But in the new coordinates we have

$$\alpha = \sum_k \Big( \sum_j \hat{b}_{jk} ds_j \Big) d\varphi_k + \sum_j g_j ds_j + dg_0$$

and obtain the relation

$$ds_k = \sum_j \hat{b}_{jk} ds_k,$$

which implies that $\hat{b}_{jk} = \delta_{jk}$. Since the form $\omega = \sum_{j<k} \hat{a}_{jk} ds_j \wedge ds_k + \sum_k ds_k \wedge d\varphi_k$ is closed, the coefficients $\hat{a}_{jk}$ are independent of the angular variables and the form $\sum_{j<k} \hat{a}_{jk} ds_j \wedge ds_k$ is closed, too:

$$\sum_{j<k} \hat{a}_{jk} ds_j \wedge ds_k = d\Big( \sum_k h_k(s) d\varphi_k \Big).$$

The change of the angular variables $\varphi_k \to \varphi_k + h_k(I)$ reduces the symplectic form to the sought shape

$$\omega = \sum_k ds_k \wedge d\varphi_k.$$

In these coordinates the Hamiltonian function is $H = H(s_1, \dots, s_n)$ and Hamilton's equations take the form

$$\dot{s}_j = \{s_j, H\} = 0, \quad \dot{\varphi}_k = \{\varphi_k, H\} = \frac{\partial H}{\partial s_k} = \psi_k(s_1, \dots, s_n), \quad j, k = 1, \dots, n.$$

The theorem is proven. □

There are two natural generalizations of the Liouville integrability:

(1) *Integrability on an energy level surface* or, in general, on a common level surface of first integrals $\{I_1 = \dots = I_k = 0\}$.

Indeed, for many systems only some special energy levels are filled with tori with conditionally periodic motions. In this case we say that the system is *integrable on a level surface* $N = \{I_1 = c_1, \dots, I_k = c_k\}$ if functions $I_1, \dots, I_n$ (including the Hamiltonian $H$) are defined in a neighborhood of this level surface and the following conditions are satisfied:

(1) the functions $I_1, \dots, I_n$ are functionally independent almost everywhere on $N$ and their Hamiltonian flows touch the level surface $N$;

(2) the Hamiltonian flows with the Hamiltonians $I_1, \dots, I_n$ commute on $N$.

Then

- *the moment map is defined on* $N$:

$$\mathcal{M}\colon N \to \mathbb{R}^{n-k}, \quad \mathcal{M}(x) = (I_{k+1}(x), \dots, I_n(x)).$$

*If $X$ is a compact component of the set $\mathcal{M}^{-1}(\tau)$ and $I_1, \dots, I_n$ are functionally independent on $X$, then $X$ is diffeomorphic to the torus and some neighborhood of $X$ in $N$ foliates into invariant tori on which the Hamiltonian flow is conditionally periodic.*

Actually, the proofs of these assertions repeat that of Theorem 9.10.

Note also that integrability is possible in the more general case when the function $I$ is not a first integral, but some level surface $\{I = c\}$ is preserved by the Hamiltonian flow and the flow is integrable on it. Such cases appear in rigid body dynamics.

(2) *Integrability by means of noncommuting integrals.*

Let $\mathcal{A}$ be an algebra of first integrals linearly generated by functions $I_1, \dots, I_{n+k}$. Suppose that these functions are functionally independent almost everywhere. Suppose that these integrals do not commute everywhere: the identities $\{I_j, I_k\} = 0$ need not hold everywhere. Suppose that $\mathcal{A}$ contains a commutative subalgebra $B$, $H \in B$, and

$$\dim \mathcal{A} + \dim B = \dim M = 2n.$$

Then

- *the common level surfaces*

$$N_c = \{I_1 = c_1, \dots, I_{n+k} = c_{n+k}\}$$

*are $(n-k)$-dimensional tori on which the Hamiltonian flows corresponding to the functions in $B$ determine conditionally periodic motions.*

The proof of this assertion is similar to that of Theorem 9.10 which corresponds to the limit case $\mathcal{A} = B$.

This method for integration can be generalized to the case when the subalgebra $B$ depends on the level surface, i.e., $B = B_c \subset \mathcal{A}$, and the Hamiltonian flows corresponding to the functions in $B_c$ commute on $N_c$. In this case the surfaces $N_c$ are again tori with conditionally periodic motions.

### Examples of integrable systems

**1 Geodesic flow on a surface of revolution.** Let $\Sigma$ be the surface in $\mathbb{R}^3$ obtained by revolution of the graph of a function $f(x)$ around the $Ox$-axis (see Problem 2.2). It is given by the parametric equations

$$x = u, \quad y = f(u)\cos v, \quad z = f(u)\sin v,$$

and the metric takes the form

$$ds^2 = (1 + f'^2)du^2 + f^2 dv^2.$$

The geodesic flow on $T^*\Sigma$ has the Hamiltonian

$$H = \frac{1}{1+f'^2} p_u^2 + \frac{1}{f^2} p_v^2.$$

Since $\dim T^*\Sigma = 4$, for the Liouville integrability it suffices to find one more first integral which is functionally independent of $H$ almost everywhere. By the Clairaut theorem (see Problem 2.13), the sought integral has the form

$$I = f(u) \cos \psi,$$

where $\psi$ is the angle between the geodesic and the parallel. This fact is a consequence of the momentum conservation law (Lemma 9.4): the metric and hence the Lagrangian of the geodesic flow do not depend explicitly on $v$ and therefore the corresponding momentum $p_v = I\sqrt{H}$ is preserved.

**Problem 9.4.** Prove that the geodesics on a surface of revolution are given by the formula

$$v(u) = \int c \frac{\sqrt{1+f'^2}}{\sqrt{f(f^2-c^2)}} du,$$

where $c$ is the value of the integral $I$ on the geodesic: $I = c$.

**2 Liouville metrics on the two-dimensional torus.** Suppose that a metric

$$ds^2 = (f(u) + g(v))(du^2 + dv^2)$$

is given on a surface with coordinates $u$ and $v$ (this metric is called a *Liouville metric*). Since the surface is two-dimensional, to integrate the geodesic flow, it suffices to find one more first integral (besides the Hamiltonian $H = \frac{p_u^2+p_v^2}{f+g}$). It is easy to verify that the following function is a first integral:

$$I = \frac{g p_u^2 - f p_v^2}{f+g}.$$

**Problem 9.5.** Prove that for a Liouville metric the geodesic equation has the form

$$\frac{du}{dv} = \pm \frac{\sqrt{f(u)+c}}{\sqrt{g(v)-c}},$$

where $c$ is the value of the integral $I$ on the geodesic: $I = c$.

If the variables $u$ and $v$ are defined modulo $\mathbb{Z}$, $f(u+1) = f(u)$, and $g(v+1) = g(v)$ then we obtain a Liouville metric on the two-dimensional torus $\mathbb{R}^2/\mathbb{Z}^2$.

**3 Euler's equations of motion of a rigid body around a fixed point.** These are the equations

$$\dot{x} = \{x, H\}$$

on the dual space of the Lie algebra of the group SO(3) (see §9.3). This dynamics reduces into symplectic leaves, two-dimensional spheres $|x| = \text{const}$. To integrate the Hamiltonian system on a two-dimensional manifold, it suffices to know one integral which is already given, the Hamiltonian

$$H = \frac{1}{2}\left(\frac{x_1^2}{A} + \frac{x_2^2}{B} + \frac{x_3^2}{C}\right).$$

Let $A > B > C$. Then the level surface $\{H = \frac{\lambda}{2} > 0\}$ foliates into the intersections of the surface with the spheres $\{|x| = \mu\}$. Moreover, for the restriction of Euler's equations to the level surface of the Hamiltonian we have the following picture (see Figure 9.3):

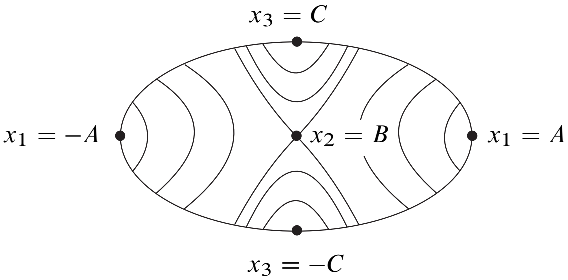


Figure 9.3. The integral curves of Euler's equations ($H = \frac{1}{2}$).

- $\lambda C \leq \mu \leq \lambda A$;
- the stationary points of the Euler equations are $(\pm\sqrt{\lambda A}, 0, 0)$, $(0, \pm\sqrt{\lambda B}, 0)$, and $(0, 0, \pm\sqrt{\lambda C})$, and they correspond to rotations of the body around the axes of inertia (in this case $\mu = \lambda A, \lambda B, \lambda C$, respectively);
- the intersection of the sphere $\{|x| = \mu\}$ and the level surface $\{H = \frac{\lambda}{2}\}$ is
  (a) homeomorphic to the circles for $\mu \neq \lambda A, \lambda B, \lambda C$,
  (b) consists of a pair of stationary points for $\mu = \lambda A, \lambda C$,
  (c) consists of a pair of stationary points and a pair of separatrices joining these points for $\mu = \lambda B$.

From this picture we easily see that the rotations around the axes of inertia corresponding to the maximal and minimal moments of inertia ($A$ and $C$) are stable, while the rotation around the axis corresponding to the moment of inertia $B$ is unstable.

**4 Integrable cases in rigid body dynamics.** Since the equations of motion of a heavy rigid body around a fixed point and the Kirchhoff equations are Hamiltonian with respect to the Lie–Poisson bracket on $e(3)^*$, they reduce onto the common level surfaces of two Casimirs of the algebra $e(3)$:

$$I_1 = p^2 = p_1^2 + p_2^2 + p_3^2, \quad I_2 = ps = M_1 p_1 + M_2 p_2 + M_3 p_3.$$

These surfaces (for $p \neq 0$) are diffeomorphic to the cotangent bundle of the two-dimensional sphere $S^2$ and the symplectic form is

$$\omega = \omega_0 + sF,$$

where $\omega_0$ is the canonical symplectic form on the cotangent bundle and $F$ is a closed 2-form on $S^2$ ("the magnetic field") (see §9.7). We have already one first integral on this four-dimensional symplectic manifold, the Hamiltonian. Therefore, to integrate the equations, it suffices to know one more independent first integral. It turns out that such integral exists only for very special values of the parameters entering in the Hamiltonian.

There are three famous integrable cases.

(a) *The Euler–Poinsot case:*

$$x_0 = y_0 = z_0 = 0, \text{ i.e., } l(p) = 0$$

(the body is fixed at the center of mass). In this case $H = H(M)$ and the dynamics of the vector $M$ is described by the usual Euler equations on the algebra SO(3). Obviously, the following casimir of this algebra is a first integral:

$$I = M_1^2 + M_2^2 + M_3^2.$$

(b) *The Lagrange–Poisson case:*

$$A = B, \quad x_0 = y_0 = 0$$

(the body is axially symmetric and the fixed point lies on the axis of symmetry). In this case the Hamiltonian is invariant under revolutions around the axis of symmetry. This action is Poisson and has an obvious linear integral which generates this action, namely the momentum with respect to the axis of symmetry:

$$I = M_3.$$

(c) *The Kovalevskaya case:*

$$A = B = 2C, \quad y_0 = z_0 = 0.$$

The extra integral has the form

$$I = (a^2 M_1^2 + a^2 M_2^2 - 2al_1 p_1)^2 + (2a^2 M_1 M_2 - 2al_1 p_2)^2,$$

where $2a_1 = 2a_2 = a_3 = a$ or, with the notation accepted in analytic mechanics,

$$I = (p^2 + q^2 + c\gamma)^2 + (2pq + c\gamma')^2, \quad c = \frac{mgx_0}{C}.$$

The last case was found by Kovalevskaya while proving the following theorem:

- *the three cases indicated above constitute all* (*general*) *integrable cases of the equations of motion of a heavy rigid body around a fixed point such that the solutions are unique meromorphic functions of time as a complex parameter.*

Here we say that an integrable case is called general when the problem is integrable everywhere (for all values of the initial data).

Other general integrable cases for this problem are unknown, but there are interesting particular cases of integrability. For example,

(d) *the Goryachev–Chaplygin case:*

$$A = B = 4C, \quad y_0 = z_0.$$

The problem has the following extra first integral on the zero level surface $I_2 = ps$ (i.e., for $s = 0$):

$$I = a^2 M_3(M_1^2 + M_2^2) + l_1 p_3 = Cr(p^2 + q^2) + mgx_0\gamma'',$$

where $a = a_1 = a_2 = A^{-1}$.

(e) *The Hess case:*

$$y_0 = 0, \quad x_0\sqrt{A(B-C)} + z_0\sqrt{C(A-B)} = 0.$$

The function

$$I = l_1 M_1 + l_2 M_2 = mg(Ax_0 p + Cz_0 r)$$

is not conserved by the motion: $\dot{I} \neq 0$ almost everywhere. However, for $I = 0$ we obtain $\dot{I} = 0$. Thus, the problem is integrable on the level surface $\{I = 0\}$ on which the function $I$ is a first integral.

There are also three known general integrable cases for the Kirchhoff equations: the Kirchhoff, Clebsch, and Steklov cases. No analog of the Kovalevskaya theorem has been established for these equations.

# Bibliography


[1] Arnold, V. I., *Mathematical Methods of Classical Mechanics.* 2nd edition. Grad. Texts in Math. 60, Springer-Verlag, New York 1989.

[2] Dubrovin, B. A., Fomenko, A. T., and Novikov, S. P., *Modern Geometry. Methods and Applications.* I. The geometry of surfaces,transformation groups, and fields. Second edition. Grad. Texts in Math. 93, Springer-Verlag, New York 1992; II. The geometry and topology of manifolds. Grad. Texts in Math. 104, Springer-Verlag, New York 1985. III. Introduction to homology theory. Grad. Texts in Math. 124, Springer-Verlag, New York 1990.

[3] Novikov, S. P., *Topology*. Encyclopaedia Math. Sci. 12, Springer, Berlin 1996.

[4] Novikov, S. P., and Taimanov, I. A., *Modern Geometric Structures and Fields.* Grad. Stud. Math. 71, Amer. Math. Soc., Providence, RI, 2006.

[5] Pontryagin, L. S., *Topological Groups.* Gordon and Breach Science Publishers, Inc., New York–London–Paris, 1966.

[6] Seifert, H., and Threlfall, W., *Topology.* Pure Appl. Math. 89, Academic Press, Inc., New York–London 1980.


### Supplementary reading sources on Riemannian geometry


[7] do Carmo, M. P., *Riemannian Geometry*. Math. Theory Appl., Birkhäuser, Boston, 1992.

[8] Gallot, S., Hulin, D., and Lafontaine, J., *Riemannian Geometry.* 3rd edition. Universitext. Springer-Verlag, Berlin 2004.

[9] Montiel, S., and Ros, A., *Curves and Surfaces.* Grad. Stud. Math. 69, Amer. Math. Soc., Providence, RI, 2005.

[10] Willmore, T. J., *Riemannian Geometry.* Oxford Science Publ., Clarendon Press, Oxford, 1993.


### Supplementary reading sources on Lie groups and representation theory


[11] Lang, S., $SL_2(\mathbb{R})$. 2nd edition. Grad. Texts in Math. 105, Springer-Verlag, New York 1985.

[12] Serre, J.-P., *Linear Representations of Finite Groups.* Grad. Texts in Math. 42, Springer-Verlag, New York, Heidelberg 1977.

[13] Warner, F. W., *Foundations of Differentiable Manifolds and Lie Groups.* Grad. Texts in Math. 94, Springer-Verlag, New York, Berlin 1983.


### Supplementary reading sources on symplectic geometry and integrable systems


[14] Arnold, V. I., Kozlov, V. V., and Neishtadt, A. I., *Mathematical Aspects of Classical and Celestial Mechanics.* 3rd edition, Encyclopaedia Math. Sci. 3, Springer-Verlag, Berlin 2006.

[15] *Dynamical Systems, IV*. Encyclopaedia Math. Sci. 4, Springer-Verlag, Berlin 2001.

[16] McDuff, D., and Salamon, D., *Introduction to Symplectic Topology.* 2nd edition. Oxford Mathematical Monographs. Oxford University Press, New York, 1998.

# Index