A.D. ALEXANDROV SELECTED WORKS

PART II

Intrinsic Geometry of Convex Surfaces

© 2006 by Taylor & Francis Group, LLC

A.D. ALEXANDROV SELECTED WORKS

PART II

Intrinsic Geometry of Convex Surfaces

Edited by
S.S. Kutateladze
Novosibirsk University, Russia

Translated from Russian by
S. Vakhrameyev

Chapman & Hall/CRC
Taylor & Francis Group
Boca Raton London New York Singapore

© 2006 by Taylor & Francis Group, LLC

Published in 2006 by
Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2006 by Taylor & Francis Group, LLC
Chapman & Hall/CRC is an imprint of Taylor & Francis Group

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-415-29802-4 (Hardcover)
International Standard Book Number-13: 978-0-415-29802-5 (Hardcover)
Library of Congress Card Number 2004049387

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Alexsandrov, A.D. (Aleksandr Danilovich), 1912–
[Vnutrenniaia geometriia vypuklykh poverkhnostei. English]
Intrinsic geometry of convex surfaces / [edited by] S.S. Kutateladze ; translated by Sergei A. vakhrameyev.
p. cm. – (A.D. Alexandrov, selected works ; pt. 2)
Includes index.
ISBN 0-415-29802-4 (alk. paper)
1. Convex surfaces. 2. Geometry, Differential. I. Kutateladze, S.S. (Semen Samsonovich) II. Title
QA3.A358413 2004
[QA643]

510 s–dc22
[516.3′6] 2004049387

T&F informa
Taylor & Francis Group
is the Academic Division of T&F Informa plc.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

© 2006 by Taylor & Francis Group, LLC

CONTENTS

Foreword **ix**

Preface **xi**

Chapter I **Basic Concepts and Results** **1**
1. The General Concept and Problems of Intrinsic Geometry 1
2. Gaussian Intrinsic Geometry 8
3. A Polyhedral Metric 13
4. Development 17
5. Passage from Polyhedra to Arbitrary Surfaces 22
6. A Manifold with an Intrinsic Metric 23
7. Basic Concepts of Intrinsic Geometry 28
8. Curvature 34
9. Characteristic Properties of the Intrinsic Metric 38
10. Some Singularities of the Intrinsic Geometry of Convex Surfaces 46
11. Theorems of the Intrinsic Geometry of Convex Surfaces 52

Chapter II **General Propositions About Intrinsic Metrics** **57**
1. General Theorems on Rectifiable Curves 57
2. General Theorems on Shortest Arcs 64
3. The Nonoverlapping Condition for Shortest Arcs 71
4. A Convex Neighborhood 73
5. General Properties of Convex Domains 80
6. Triangulation 83

Chapter III **Characteristic Properties of the Intrinsic Metric** **91**
1. Convergence of the Metrics of Convergent Convex Surfaces 91
2. The Convexity Condition for a Polyhedral Metric 99
3. The Convexity Condition for the Metric of a Convex Surface 108
4. Consequences of the Convexity Condition 113

Chapter IV **The Angle** **121**
1. General Theorems on Addition of Angles 121
2. Theorems on Addition of Angles on Convex Surfaces 128
3. The Angle of a Sector Bounded by Shortest Arcs 131
4. On Convergence of Angles 136
5. The Tangent Cone 141
6. The Spatial Meaning of the Angle between Shortest Arcs 147

© 2006 by Taylor & Francis Group, LLC

Chapter V **Curvature** **157**
1. Intrinsic Curvature 157
2. The Area of a Spherical Image 163
3. Generalization of the Gauss Theorem 174
4. The Curvature of a Borel Set 181
5. Directions in Which It Is Impossible to Draw a Shortest Arc 186
6. Curvature as Measure of Non-Euclidicity of the Metric 188

Chapter VI **Existence of a Convex Polyhedron with a Given Metric** **197**
1. On Determining a Metric from a Development 197
2. The Idea of the Proof of the Realization Theorem 204
3. Small Deformations of a Polyhedron 210
4. Deformation of a Convex Polyhedral Angle 213
5. The Rigidity Theorem 218
6. Realizability of the Metrics Close to Realized Metrics 222
7. Smooth Passage from a Given Metric to a Realizable Metric 225
8. Proof of the Realizability Theorem 233

Chapter VII **Existence of a Closed Convex Surface** **235**
1. The Result and the Method of Proof 235
2. The Main Lemma on Convex Triangles 241
3. Corollaries of the Main Lemma on Convex Triangles 249
4. The Complete Angle at a Point 252
5. Curvature and Two Related Estimates 258
6. Approximation of a Metric of Positive Curvature 262
7. Realization of a Metric of Positive Curvature Given on a Sphere 269

Chapter VIII **Other Existence Theorems** **277**
1. The Gluing Theorem 277
2. Application of the Gluing Theorem to the Realization Theorems 281
3. Realizability of a Complete Metric of Positive Curvature 284
4. Manifolds on Which a Metric of Positive Curvature Can Be Given 288
5. The Question of the Uniqueness of a Convex Surface with a Given Metric 295
6. Various Definitions of a Metric of Positive Curvature 298

Chapter IX **Curves on Convex Surfaces** **301**
1. The Direction of a Curve 301
2. The Swerve of a Curve 308
3. The General Gluing Theorem 316
4. Convex Domains 320
5. Quasigeodesics 326
6. A Circle 332

Chapter X **Area** **341**
1. The Intrinsic Definition of Area 341
2. The Extrinsic–Geometric Meaning of Area 350
3. Extremal Properties of Pyramids and Cones 356

Chapter XI **The Role of Specific Curvature** **365**
1. Intrinsic Geometry of a Surface 365
2. Intrinsic Geometry of a Surface of Bounded Specific Curvature 376
3. The Shape of a Convex Surface Depending on Its Curvature 386

© 2006 by Taylor & Francis Group, LLC

Chapter XII **Generalization** **393**
1. Convex Surfaces in Spaces of Constant Curvature 393
2. Realization Theorems in Spaces of Constant Curvature 398
3. Surfaces of Indefinite Curvature 402

Appendix **Basics of Convex Bodies** **409**
1. Convex Domains and Curves 409
2. Convex Bodies. A Supporting Plane 411
3. A Convex Cone 414
4. Topological Types of Convex Bodies 415
5. A Convex Polyhedron and the Convex Hull 418
6. On Convergence of Convex Surfaces 421

© 2006 by Taylor & Francis Group, LLC

FOREWORD

Aleksandr Danilovich Alexandrov (1912–1999) is usually ranked as second to Gauss in surface theory. This appraisal rests primarily on his masterpiece whose English translation lies before the reader.

There is no need in expatiating upon the content of this book since the main ideas of geometry "in the large" have been placed in the treasure trove of modern science half a century ago. The reader will be pleased to taste the original presentation of such fundamental concepts as shortest arc, intrinsic metric, angle, curvature, etc.

For this reason my introductory words are kept at a minimum, confining merely to those of pride and apology.

It was in the mid 1990s that Alexandrov asked me to help him in editing this translation he chose as the second volume of his *Selected Works*. We agreed with him from the very beginning that no modernizing comments will be added to the book since the area of geometry in the large is impossible to embrace these days; thus, the reader can feel the original flavor of invention which is undoubtedly spoiled in translation. I am proud of this opportunity to return a trifle to the brilliant person whose genius, charm, and sympathy are listed as decoration and consolation of about two decades of my life.

I also offer my sincere words of apology to the reader for the crushing defeats I have suffered as editor in the severe battle against mistranslation and solecism. Unfortunately, my inspired teacher and long-time interlocutor had no chance to look through this translation but I am pretty sure that no blunders of mine nor the translator's nor anybody else's can ever conceal the keen mind and farsighted vision of Alexandrov, a geometrical giant of the 20th century.

S.S. Kutateladze
Novosibirsk, 2002

© 2006 by Taylor & Francis Group, LLC

PREFACE

Intrinsic geometry investigates the properties of a surface and of figures on it within the limits of the surface similar to the way plane geometry studies a plane without regard to the fact that it may be located in some space. The start of intrinsic geometry was made by Gauss' paper "*Disquisitiones generales circa superficies curvas,*" which appeared in 1827. Since that time, intrinsic geometry has advanced so far that, at present, all of its major issues can be considered solved, at least those that deal with the geometry of small pieces of regular surfaces. However, in their conventional form the differential-geometry methods apply only to regular surfaces, i.e., the surfaces set up by equations with functions differentiable sufficiently many times, usually, at least three times. Meanwhile, irregular surfaces merit no less consideration, as they often occur in real life and can be made from, say, a sheet of paper. For example, any polyhedron or cone, or the surface of a lens with sharp edges are not regular. It is no wonder then that there is a need to study irregular surfaces, too. Further, the differentiability conditions imposed on surfaces in differential geometry are not always justified by the geometric essence of a problem and are there mostly for the convenience of analytical tools. Finally, in recent years, attention of geometers has been attracted to problems of geometry "in the large;" in solving these problems, analytical tools lose their usual automatism and often prove inefficient, thus requiring the use of purely geometrical or topological arguments.

It certainly is hopeless to study all possible surfaces, as we cannot expect any far-reaching general results for them. We restrict our consideration to convex surfaces. In a very general sense, a convex surface is the whole boundary or part of the boundary of a convex body, i.e., a body including each line segment with endpoints in this body. According to Minkowski, "theorems on convex bodies are of especial fascination, since, as a rule, they are true for the whole category of these objects without any exceptions." Convex surfaces, too, turn out to be, so to say, a very successfully defined object. Recent research into them yielded general results not constrained by any preconceived regularity conditions. These results, however, do not belong to their intrinsic geometry. It seemed, therefore, essential to study the intrinsic geometry of convex surfaces, too, and, first of all, to find the conditions that completely characterize their intrinsic metric. It turned out that the intrinsic geometry of arbitrary convex surfaces could be no less rich in content than that of regular surfaces. I also succeeded in clarifying the geometric conditions of using the basic formulas of the classical intrinsic geometry. This book expounds the major aspects of an extensive theory that emerged here; it took me several years to construct the theory.

As regards the basic concepts and methods, I completely leave the framework of common differential geometry. Coordinates on a surface are introduced only

© 2006 by Taylor & Francis Group, LLC

in special cases. A metric of a surface is given directly by the distance between points measured on the surface rather than by a line segment. The fundamental concepts of length, angle, area, integral curvature, etc. are given intrinsic geometric definitions. In this connection, the method of research is of almost exclusively geometric character. The essence of the method is that, first, convex polyhedra are studied, and then the results obtained for them are extended to any convex surfaces by passing to the limit. Second, the figures on a curved surface are replaced by corresponding figures on a plane, and the resulting distortions are studied. This pertains primarily to triangles. In accordance with this, we approximate a "curved" metric by a "polyhedral metric" given by a set of plane triangles glued together by their sides. Therefore, in particular, an important role is played by the theorem yielding the conditions under which we can glue a closed convex polyhedron together from polygons cut out of a sheet of paper. Finally, an extremely fruitful method is cut and paste, which consists in constructing a surface "pasted" together from pieces of other surfaces.

Although our primary goal is to study convex surfaces without imposing additional regularity conditions, some results prove new for regular surfaces, too, especially when we are talking about problems "in the large." It is also worth noting that our methods prove applicable to studies of nonconvex surfaces. The emerging abstraction of our theory is considered, though in little detail, in the last chapter.

I should thank D.A. Raikov, the editor of this book, and V.A. Zalgaller for a number of remarks that enabled me to correct some essential inaccuracies of exposition. I would like to remember here my friends I.M. Liberman and S.P. Olovyanishnikov, who perished at the front in 1941 at the height of their geometric talent. Constant close contacts with them contributed to my elucidating many points of the theory expounded here when it was just conceived. The proof of existence of the one-sided tangent at each point of a geodesic on a convex surface, first given by Liberman and reproduced in Sec. 5 of Chapter IV, is an undisputable specimen of beautiful geometric argumentation; the method used by Liberman admits many other essential applications. Section 5 of Chapter VIII presents an excellent theorem by Olovyanishnikov on bending infinite convex surfaces – the first general result in this old problem dealt with by Darboux in his time.

A few words on the order and character of exposition.

I tried to make the book intelligible not only for experts. In view of this and also because I leave the framework of common differential geometry, I begin with the very basic notions: definition of the intrinsic metric of a surface, statement of intrinsic-geometry problems, etc. The required minimum of information on the theory of convex bodies is given in Appendix. The mathematical tools I use are not numerous. Some topological minimum is essential: Jordan curve theorem; Euler theorem; the concepts of closed and open sets, a boundary, a continuous mapping; etc. Chapters V and X essentially use the basic facts of the Lebesgue measure theory, and Chapters X and XI use the Lebesgue integral. (Using not the very first theorems of these theories, I refer to the popular courses, where their proof is given.) Chapter VI uses the popular implicit function theorem. Any deep knowledge of differential geometry is absolutely not required.

Every effort was made to explain and prove all basic concepts and theorems in as great detail as possible. Many interesting problems, which are however not

© 2006 by Taylor & Francis Group, LLC

basic, are presented more briefly, in the form of a review. Moreover, in the course of exposition I allowed myself to formulate a number of yet unsolved problems; their probable difficulty certainly varies within wide limits.

Chapter I introduces the basic concepts and presents the major results without proof. It is sort of a survey of the theory as a whole. Chapter II presents the familiar general theorems on rectifiable curves (Sec. 1) and the general theorems on curves of minimal length (Secs. 2 and 3). This part of Chapter II is required for subsequent exposition; the next sections (4–6) contain slightly more profound results, which are first used only in Chapter V and especially in Chapters VII and X. Chapter III establishes some rather general properties of the intrinsic metric of convex surfaces. After that, the exposition proceeds along two almost independent lines: (1) further study of intrinsic geometry of convex surfaces and (2) proof of existence of a surface with a given metric.

The first line travels as follows: Chapter IV studies the basic properties of the angle between two shortest arcs. Chapter V deals with curvature theory (Secs. 1–4) and some of its applications (Secs. 5 and 6). Sections 1 and 2 of Chapter IX study the foundations of the theory of curves on convex surfaces, and Secs. 3–6 deal with applications and special problems of curve theory, which are mostly presented as a review. Sections 1 and 2 of Chapter X present the foundation of the theory of area on convex surfaces, and Sec. 3 considers certain maximization problems concerning the area. Finally, Chapter XI discusses the surfaces at which the ratio of the curvature of a domain and the area of this domain are subject to some constraints.

The second line traverses Chapters VI, VII, and VIII and also Secs. 3 and 4 of Chapter IX; Chapters VI, VII, and VIII use only the results of Sec. 1 of Chapter IV and Sec. 2 of Chapter V (besides the results of Chapters II and III).

Chapter XII is a review-like presentation of the generalization of the whole theory to convex surfaces in Lobachevskii space and in spherical space; the last section of this chapter outlines the theory of nonconvex surfaces which are certainly subject to some necessary conditions.

A.D. Alexandrov
Leningrad, 1946

© 2006 by Taylor & Francis Group, LLC

Chapter I

BASIC CONCEPTS AND RESULTS

1. The General Concept and Problems of Intrinsic Geometry

A *surface* is a figure (a set of points) in three-dimensional Euclidean space that has the following properties: (1) each of its points has a neighborhood homeomorphic to a disk and (2) each of its two points can be connected by a continuous curve.[1] Herewith, a neighborhood of a point O on a surface is any part of this surface that contains all of its points lying in a ball centered at O.[2] Any continuous image of a segment is called a *continuous curve*.

Let F be a surface having the property that each of its two points can be connected by a curve of a bounded length. Then for each pair of points X and Y of the surface F, there exists the greatest lower bound of lengths of the curves lying on the surface F and connecting these points.

This greatest lower bound is called the *distance between X and Y on F* and is denoted as $\rho_F(XY)$.

The distance so defined satisfies the three major conditions that are usually imposed on the concept of distance in general. These conditions are as follows:

1. $\rho(XY) = 0$ if and only if X and Y coincide;
2. $\rho(XY) = \rho(YX)$, i.e., the distance from X to Y is the same as that from Y to X;
3. $\rho(XY) + \rho(YZ) \geq \rho(XZ)$. This last condition is called the *triangle inequality*.

[1] By virtue of this definition, a surface has no self-intersections, and we assume that the boundary of a surface (if it exists) does not belong to the surface; in this way, say, a hemisphere will be a surface in the sense of this definition only if we exclude the bounding equator, since for a point in the equator there is no circular neighborhood in the hemisphere. We introduce these restrictions for convenience in order to avoid various stipulations connected with the fact that points at which a surface intersects itself, as well as boundary points, play a special role.

[2] The concept of the neighborhood in a set (a topological space) M is usually defined as follows: (1) the set M itself is a neighborhood of a point O in the set M; (2) the point O belongs to the neighborhood; (3) if U and V are two neighborhoods of O, then there exists a neighborhood of O in U and V; (4) if a point A belongs to a neighborhood U of the point O, then there exists a neighborhood of A that lies in U (i.e., a neighborhood is an open set relative to M). From this definition of the neighborhood of a point on a surface, the first three conditions hold and the last, generally speaking, does not, because by a neighborhood we mean *any* part of the surface that contains all of its points confined in some ball. This relaxation of the conditions determining the concept of neighborhood is inessential, and we introduce it for simplicity. Further on, we always bear in mind our simplified definition of neighborhood. The same digression from the usual way of definition has been made, e.g., in Ch. II, Sec. 5 of *Topology* by H. Seifert and W. Threlfall; they show such a definition to be quite sufficient.

© 2006 by Taylor & Francis Group, LLC

These conditions imply that $\rho(XY) \geq 0$, since $\rho(XY) + \rho(YX) \geq \rho(XX)$, i.e., $2\rho(XY) \geq 0$.

The fact that the above conditions do hold for the defined distance on a surface is so obvious from the definition that there is no need to give their proof.

Any set in which the number $\rho(XY)$ (the function of a pair of elements) is defined for each pair of its elements (points) X and Y, satisfying the above three conditions, is called a *metric space*; the function $\rho(XY)$ is called the *metric* of this space, and its value for a pair of points X and Y is called the distance between these points.[3] Using this notion, we can say that by the above definition of the distance on a surface any surface becomes a metric space. The goal of intrinsic geometry is to study this metric space by itself, irrespective of the fact that it is a figure in a three-dimensional space or any other ambient space. The above-defined metric on a surface is called *intrinsic* in contrast to the "extrinsic" metric that yields spatial distances between points of the surface. A surface has intrinsic metric if and only if each two of its points can be connected by a curve of finite length; therefore, we can speak of intrinsic geometry only as regards such surfaces.

Only those concepts and theorems, which can be described and formulated such that no properties of the surface, except for the properties of its metric, are ultimately used, belong to the intrinsic geometry of a given surface.

Examples of these concepts, besides the very concept of the distance on a surface, are, first and foremost, the concept of neighborhood (the neighborhood of a point X is each set of points of this surface containing all points of the surface that are distant from X by no less than some $r > 0$ in the sense of the distance of the surface) and all related topological concepts, i.e., the notions of closed and convex sets, the condensation points, the boundary of a set, etc. In what follows, we shall show that, using only the intrinsic metric of a surface, we can introduce the concepts of length, angle, and area, as well as an analog of a line segment, a shortest arc connecting two points on a surface; after that, there naturally appears the concept of rectangle, etc. In short, we will see that intrinsic geometry acknowledges the notions and problems as substantive as plane geometry does; the latter is none other than the intrinsic geometry of a plane.

If we have two metric spaces, and there is a correspondence between their points such that the distance between each two points in one space is equal to the distance between the corresponding two points in the other, these spaces are called *isometric*; the above correspondence or mapping from one space onto the other is said to be *isometric*. If we can establish an isometric correspondence between two surfaces, their intrinsic geometry is the same. The propositions of intrinsic geometry are invariant under isometric mappings.

[3] We do not assume the reader to be familiar with the theory of metric spaces. It is sufficient to know that a neighborhood of a point X in a metric space is each set that contains all points lying at a distance of less than some $r > 0$ from X. After that, the definitions of the closed and open sets, boundary, continuous mapping, condensation point, etc. are given by repeating verbatim the definitions of these notions for the case of Euclidean space. From the properties of distance, we need the following: if points X_n and Y_n converge to points X and Y, i.e., if $\rho(X_nX) \to 0$ and $\rho(Y_nY) \to 0$, then $\rho(X_nY_n) \to \rho(XY)$. This follows immediately from the triangle inequality. Indeed, $\rho(X_nY_n) \leq \rho(X_nY) + \rho(YY_n) \leq \rho(XY) + \rho(X_nX) + \rho(Y_nY)$ and, in the same way, $\rho(XY) \leq \rho(X_nY_n) + \rho(X_nX) + \rho(Y_nY)$, i.e., $|\rho(XY) - \rho(X_nY_n)| \leq \rho(X_nX) + \rho(Y_nY)$ and, for $\rho(X_nX)$ and $\rho(Y_nY)$ tending to zero, we have $\rho(X_nY_n) \to \rho(XY)$.

© 2006 by Taylor & Francis Group, LLC

Studies of the intrinsic geometry of all surfaces in general are, certainly, a matter of great difficulty, and it is hardly possible to obtain far-reaching general results here, since the concept of surface in general is very broad and includes, so to say, arbitrarily "bad" irregular surfaces. Therefore, we restrict ourselves to a more or less broad class of surfaces satisfying additional conditions. For example, classical differential geometry investigates only those surfaces that can be given in rectangular coordinates by equations involving functions that are differentiable sufficiently many times, generally speaking, no less than thrice. We call these surfaces *regular*.

In this book, we consider convex surfaces expecially, without any restrictions of smoothness, existence of curvature, etc. The definition of convex surface is as follows. A *convex body* is a closed set having interior points and possessing the property that, together with each two of its points, it contains the whole line segment connecting these two points. A *convex surface* is any domain (i.e., a connected and open set) lying on the boundary of the convex body. A surface that is the whole boundary of a convex body is called a *complete* convex surface.

Further on, we shall constantly use some basic information on convex surfaces. This information is given in a small appendix at the end of the book. The reader familiar with the basics of convex body theory will hardly find there anything new; the reader unfamiliar with this theory can read the appendix first or turn to it when necessary, following the references we will make.

In particular, Sec. 4 of Appendix proves that complete convex surfaces can be of only three topologically different types:

1. closed convex surfaces homeomorphic to the sphere; they bound convex bodies of finite size;

2. infinite convex surfaces homeomorphic to the plane, as, e.g., an elliptic paraboloid;

3. convex cylindrical surfaces homeomorphic to the circular cylinder.

The property that a complete convex surface is homeomorphic to the sphere, plane or cylinder obviously implies that any convex surface, i.e., a domain on a complete convex surface, is a surface in the sense of the initial definition, i.e., each of its points has a neighborhood homeomorphic to a disk, and each two of its points can be connected by a continuous curve lying in this domain.

Moreover, *each two points of a convex surface can be connected by a curve of finite length lying on this surface*; *therefore, each convex surface has the intrinsic metric* (were this otherwise, the issue of the intrinsic geometry of all convex surfaces would have had no meaning).

Indeed, let F be some convex surface; it is a domain on the boundary of some convex body H. Take some point A on F and draw a supporting plane P to the surface F through this point (see Appendix, Sec. 2). If a point O lies in the interior of H, then each ray emanating from this point intersects the boundary of H at no more than a single point. Therefore, in a sufficiently small neighborhood of A, the projection from the point O yields a one-to-one bicontinuous mapping of the surface F onto the plane P. Let U be a neighborhood of A whose projection to the plane P is the disk U' centered at A (Fig. 1). Take two points X and Y in the

© 2006 by Taylor & Francis Group, LLC

neighborhood U; let X' and Y' be the projections of these points to the plane P. The segment $X'Y'$ lies in the disk U' and is the projection of the intersection curve $\widehat{XY}$ of the plane angle $OX'Y'$ and the neighborhood U. The curve $\widehat{XY}$ is convex, as the intersection of the plane $OX'Y'$ with the convex body H is a convex plane domain. Since each convex curve has length, by our argument, we have proved the following: each point A of a convex surface has a neighborhood in which each two points can be connected by a curve of finite length.

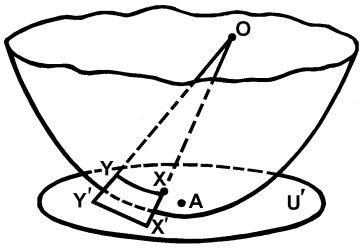


Fig. 1

Now let X and Y be two arbitrary points of the surface F. Connect these points by an arbitrary continuous curve L lying on F; this curve exists by virtue of the very definition of surface but can have no finite length. According to what we have proved, each point of the curve L can be surrounded by a neighborhood in which each two points can be connected by a curve of finite length. Using the famous Borel lemma, we can refine a finite subcover of L from these neighborhoods. Then, taking a finite sequence of points on the curve L, each of which belongs to two intersecting neighborhoods, and connecting these points by curves of finite length, we obtain the curve that also has finite length and connects two given points X and Y. Thus, each convex surface does have intrinsic metric.

What are the general problems of the intrinsic geometry of surfaces? These problems can be somewhat conventionally divided into five groups. Suppose that we consider a class of surfaces Φ, e.g., the class of convex surfaces. From the intrinsic-geometrical point of view, each surface of the class Φ is some metric space. The first group of problems reduces to the search for those properties that single out the surfaces of the class Φ from all possible metric spaces. In other words, we speak about conditions that are necessary and sufficient for a given metric space to be isometric to some surface from the class Φ. A surface F is isometric to a space R if there is a correspondence between points of space R and the surface F such that, for each pair of points of the surface F, the distance in the sense of the intrinsic metric is equal to the distance between the corresponding points of space R. We say that this surface *realizes* the metric of space R.

The problem considered can be described as that of axiomatic justification of the intrinsic geometry of the surfaces of class Φ. Indeed, adding the conditions necessary and sufficient for a metric space to be isometric to a surface of class Φ to the conditions (axioms) defining the concept of a metric space, we obtain a complete set of conditions, axioms that define surfaces of class Φ from the viewpoint of their intrinsic geometry. The simplest example of such a set of axioms can be the axiomatics of Euclidean plane geometry based on the concept of distance; in this case, the class Φ consists of a single surface, the plane.

The first group of problems is followed by another, also belonging to the foundations of intrinsic geometry and completely analogous to the known problem of the foundations of plane geometry. The problem is to define and study the main

© 2006 by Taylor & Francis Group, LLC

properties of the following basic geometric magnitudes: length, angle, and area. Besides, the intrinsic geometry of surfaces has a new concept that has no analogs in plane geometry: the concept of the intrinsic curvature of surface. These concepts will be defined below; here, we only formulate the problem – define and study the properties of the magnitudes indicated.

After these concepts are all defined, they need to be brought into action, as any mathematical theory acquires intensionality only if its basic propositions lead to a sufficient number of contensive theorems. Now, the third group of problems of intrinsic geometry consists of developing, based on the solutions of the first two problems, the intrinsic geometry of surfaces of class Φ and obtaining various general and sufficiently contensive theorems. For instance, general theorems on triangles whose sides are shortest arcs on a surface; or the solution of the important problem of finding necessary and sufficient conditions under which a convex surface proves so regular from the viewpoint of intrinsic geometry that the method of classical differential geometry is applicable. Of course, the boundaries of the third group of problems are absolutely indefinite and, in fact, infinite, as even the number of general theorems that can be proved is beyond any limits.

The fourth group of problems, also rather indefinite by nature, consists of pinpointing and elaborating some, as general as possible, methods for solving various problems of intrinsic geometry. Of course, nobody can give a universal method applicable for solving any problem, in the same way as there is no universal medicine for all diseases. But pinpointing a general way to solving a large number of problems is certainly reasonable. Such a general way for us will be that of passage to the limit from convex polyhedra to arbitrary convex surfaces.

Finally, the fifth group of problems, which is beyond the framework of intrinsic geometry proper, is to elucidate the relation of the intrinsic geometry of a surface to its "extrinsic" geometry, i.e., to elucidate how the intrinsic geometric properties of a surface and figures on it affect their spatial properties. For instance, what are the conditions imposed on the intrinsic metric of a surface of class Φ that ensure its smoothness? What properties of a surface of class Φ as a figure in space depend only on its intrinsic metric and are, thus, preserved under any isometric transformations sending this surface again to a surface of class Φ?

For us, the class of all convex surfaces shall be the class Φ. However, the solution of the problem of necessary and sufficient conditions for a given metric space to be isometric to a convex surface has one difficulty, which is worth discussing now.

Imagine some surface F, say, a sphere. Take two points A and B on it. The distance $\rho_F(AB)$ between them is determined by the length of the shortest of the arcs AB of the great circle. Now, cut out a domain F_1 containing the points A and B but disjoint from not only some segment of the arc AB but also from all lines arbitrarily close to this segment. For F_1 we can take, e.g., part of the sphere that remains after deleting some disk centered at the midpoint of the arc AB. We obtain a new surface F_1, which is part of the previous one. The lengths of all curves that connect the points A and B on F_1 are greater than the length of the arc AB by some positive ε. Therefore, it will turn out that the distance between A and B on F_1 is greater than on F; namely,

$$\rho_{F_1}(AB) \geq \rho_F(AB) + \varepsilon.$$

© 2006 by Taylor & Francis Group, LLC

Herewith, it is clear that $\rho_{F_1}(AB)$ depends on the manner in which we cut out the domain F_1; it can even be made arbitrarily large. For this, we should take as F_1 the domain that contains the points A and B and wraps around the sphere many times as a narrow band (see Fig. 2).

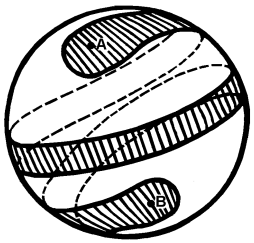


Fig. 2

This simple example shows that, by cutting out various domains from a surface, we can obtain new surfaces whose intrinsic metrics differ greatly from the metric of the initial surface. In this way, we obtain a hardly controllable set of different intrinsic metrics, and it is not clear if it is too difficult to search for, or if there is really any sense in searching for, properties that characterize completely each of the metrics obtained in this way. For example, even the answer to the following simple question – what are necessary and sufficient conditions for the metric of a surface that is isometric to a domain on the plane – proves rather complicated. The condition for a surface to be "developable," i.e., for its Gaussian curvature to be zero, is not sufficient, since, e.g., a cylinder is not isometric to any domain on the plane; it can be developed on the plane only if it is cut up in a suitable manner, but this would violate its intrinsic metric.

However, all surfaces that are parts of the same surface have much in common. Take any point O on a surface F and consider the part of this surface formed by points lying on O at a distance of less than a given $r > 0$ in the sense of the distances on F. This is a neighborhood of the point O in the sense of the intrinsic metric of the surface F; in short, an "intrinsic neighborhood." If we take an arbitrary domain F_1 that contains the point O, there always exists some $r > 0$ such that the corresponding intrinsic neighborhood of the point O is so distant from the boundary of the domain F_1 that the curve, which connects each two points X and Y on the surface F and has the minimum length or at least the length arbitrarily close to the greatest lower bound of the lengths of curves between X and Y, passes away from the boundary of the domain F_1 at some distance. Therefore, the distance between X and Y on F_1 is the same as on F. Thus, we arrive at the following conclusion:

The metric on an arbitrarily small intrinsic neighborhood of any given point of a surface does not depend on the fact which domain containing this point we cut out from the surface and consider as a new surface. In exactly the same way, the metric on an arbitrarily small neighborhood of a given point of a surface F does not change if we extend the surface F up to some larger surface F_1.

With regard to convex surfaces, this remark leads to the following result. Let F be a convex surface, and let O be a certain point on F. By definition, F is a domain on the boundary of some convex body H. Take the convex body H_1 to be the intersection of H with some ball centered at the point O and containing an *a priori* given bounded part of the surface F. The surface of the body H_1 is a closed convex surface F_1 on which the point O has a neighborhood coinciding with its neighborhood on the initial surface F. If this neighborhood is sufficiently small, then the metric on this neighborhood does not depend on whether we consider this metric on the surface F or on the surface F_1. Thus, *any bounded part of a convex surface can be completed to a closed convex surface, and the properties of*

© 2006 by Taylor & Francis Group, LLC

the metric in small neighborhoods can be studied only for closed convex surfaces. In what follows, we shall often use this remark, since it is simpler to deal with closed surfaces in many respects.

The properties that hold in an arbitrarily small neighborhood are called local or the properties "in the small." We have shown that the local properties of the intrinsic metric of a surface are stable in the operation of distinguishing various domains from this surface and remain common for these domains. Therefore, it is natural that the local properties are the first to consider in studies of the intrinsic geometry of surfaces; thus, the initial and simplest chapter, intrinsic geometry "in the small," is singled out first from intrinsic geometry "in the large."

The problem of local characterization of those metric spaces that are convex surfaces is formulated as follows: what are necessary and sufficient conditions for a metric in an arbitrarily small neighborhood of any point of a given metric space to be such that there exists a convex surface whose arbitrarily small part is isometric to the given neighborhood? In this book, we give an answer to this question.

Aside from the local properties common for all convex surfaces, the properties of complete convex surfaces, i.e., the whole boundaries of convex bodies, are also of interest. Namely, we will answer the following question: what are necessary and sufficient conditions for a metric space to be such that there exists a complete convex surface that is isometric to this space? This is said to be the statement of the problem "in the large," in contrast to the previous problem "in the small." Since the definition of a convex surface implies that it can be completed to the whole boundary of a convex body, the statement of this question is quite natural.

It turns out, however, that we may combine these two problems, "in the small" and "in the large," into one problem.

A *shortest arc* XY, also called a shortest join of the points X and Y, is a curve on a surface that connects two given points X and Y and has a minimal length among all curves lying on the surface and connecting these two points. A shortest arc is an analog of a line segment and is the main object of intrinsic geometry. However, it is well known that the properties of a shortest arc do not precisely inherit all properties of a line segment; e.g., there are pairs of points on a sphere that are connected not by a single arc but by infinitely many shortest arcs. Since the distance $\rho(XY)$ on a surface is equal to the greatest lower bound of the lengths of curves connecting the points X and Y, the length of the shortest arc XY is equal precisely to the distance between these points. We can prove that if a surface has intrinsic metric in principle, then each point of this surface has an intrinsic neighborhood, each two points of which can be connected by a shortest arc. Also, we can prove that each two points of a complete convex surface can be connected by a shortest arc. Hence, shortest arcs always exist. (These assertions will be proved in Sec. 2 of Chapter II.)

A domain G on a surface F is called *convex* (in the sense of the intrinsic metric of F) if each two points of the domain G can be connected by a shortest arc lying in this domain. Cut out a convex domain G from a surface F and consider this domain itself as a surface. If X and Y are two points of the domain G, the distance between these points does not depend on whether we consider the domain G as an independent surface or as part of the surface F. Thus, a convex domain on a surface has an important property: its metric does not depend on which part of

© 2006 by Taylor & Francis Group, LLC

the surface including this domain we distinguish and consider as an independent surface. Besides, the fact that each two points in a convex domain can be connected by a shortest arc makes possible many constructions in such domains that require shortest arcs to be drawn. In Chapter II, we consider some basic properties of convex domains and prove, in particular, that each point on a convex surface has an arbitrarily small convex neighborhood. Also, since two points of a complete convex surface can always be connected by a shortest arc, such a surface itself is a convex domain. Thus, the study of the intrinsic metric of convex domains combines, in particular, the study of the intrinsic metric "in the small" and "in the large." In connection with what we have said above, it becomes clear that main attention can be focused on the study of the intrinsic geometry of convex domains; it is possible to completely solve the first problem of intrinsic geometry for these domains, i.e., to find a necessary and sufficient condition under which a given metric space is isometric to a convex domain on a complete convex surface; in particular, such a domain itself can be a complete surface.

The next section considers in general how the basic problems of classical differential geometry were stated and solved. We do not endeavor to say anything new in this respect; our aim is only to make clearer the interplay and distinction between the classical approach to these problems and the approach we pursue in this book. We will never use the results and methods of classical differential geometry, but they show the goals to reach in order to preserve, if possible, the main classical achievements in a more general theory.

In the following sections of this chapter, we introduce the basic concepts of intrinsic geometry and formulate the main results in order to present a clearer general picture of the theory whose details will be developed step by step in the subsequent exposition.

2. Gaussian Intrinsic Geometry

The foundations of intrinsic geometry were laid by Gauss in his remarkable paper, "*Disquisitiones generales circa superficies curvas*," which appeared in 1827. The first two basic ideas of this paper, the introduction of coordinates on a surface and a line element of this surface, can be summarized in modern terms as follows. Let F be a certain surface; take some point of this surface and a neighborhood U of this point which is homeomorphic to a square D. Introduce the usual Cartesian coordinates u, v on this square D. Then since there is a bijective and bicontinuous correspondence between the points of the neighborhood U and those of the square D, we can represent the neighborhood U by the equation

$$x = x(u, v);$$

this equation means that a vector x going from a chosen origin to a point of the domain U corresponds to each pair of values of the coordinates u, v. In other words, x is a function of u and v. Assume that this function is continuously differentiable, i.e., there exist the partial derivatives $x_u(u, v)$ and $x_v(u, v)$, which continuously depend on u and v; also, assume that the vectors x_u and x_v are linearly independent at each point. If a surface admits such a representation in a neighborhood of each

© 2006 by Taylor & Francis Group, LLC

point, then it is said to be *smooth* or *differentiable*. At each point, it has the tangent plane, i.e., the plane that passes through the vectors $x_u(u,v)$ and $x_v(u,v)$. A small neighborhood of a point on this surface is uniquely projected to the tangent plane at this point, and the surface in such a small neighborhood can be replaced by a part of the plane to within the second order infinitesimals.

If $x(u,v)$ and $x(u+du, v+dv)$ are two close points on a surface, then

$$x(u+du, v+dv) - x(u,v) \cong x_u(u,v)du + x_v(u,v)dv$$

to within higher order infinitesimals, i.e., (neglecting second order infinitesimals), the vector connecting two close points on the surface lies in the tangent plane. The square of the length of the vector $dx = x_u du + x_v dv$ is

$$dx^2 = ds^2 = Edu^2 + 2Fdu\,dv + Gdv^2, \tag{1}$$

where, according to Gauss, we put

$$E = x_u^2, \quad F = x_u x_v, \quad G = x_v^2. \tag{2}$$

Equation (1) is precisely the square of the length of the vector dx on the plane, which is expanded into constituents with respect to the vectors du and dv and which has the components (constituents) equal to du and dv with respect to these vectors.

Thus, roughly speaking, the meaning of the Gaussian expression for the square of the length element is that, infinitesimally, a surface is isometric to a small domain on the tangent plane to within higher order infinitesimals, and the quadratic form (1) – the first fundamental form of the surface – yields the representation of the length on the tangent plane in the coordinates du, dv with the principal vectors x_u and x_v. The coefficients E, F, and G depend on the values of u and v and completely determine the intrinsic metric of the surface. Namely, the length of a curve $u = u(t)$, $v = v(t)$ (i.e., $x = x(u(t), v(t))$ on a surface is expressed by the integral

$$s = \int ds = \int \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\, dt. \tag{3}$$

If the length is defined, the distance is also defined as the greatest lower bound of the lengths of curves. Hence the intrinsic metric of a smooth surface is given by its first quadratic form. The implied concept of intrinsic geometry is defined by Gauss in Sec. 13 of his paper as follows:

"The proposition deduced in the previous section leads to the consideration of curved surfaces from a new point of view, which deserves most careful study by geometers. Namely, if we consider surfaces not as the boundaries of bodies but as bodies one dimension of which is infinitely small and which are, besides, completely flexible but not extensible, then the properties of a surface partly depend on the shape it takes at a given moment and are partly absolute, i.e., remain invariable however strongly it is bent. The latter properties whose study opens a fruitful new field for geometry include the measure of curvature and total curvature in the sense that we have established; the theory of shortest arcs and something else, which we shall consider later, also belong here." [Then Gauss studies, in particular, triangles

© 2006 by Taylor & Francis Group, LLC

on the plane.] "From this point of view, the plane and the surface that can be developed onto the plane, for example, a cone, are considered as essentially the same. The true initial point for the general expression of a surface from this point of view is the formula $\sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}$, expressing the dependence of an arc element on two auxiliary variables" (i.e., u and v).

The length of curve, the angle between curves, the area of domain, the geodesic curvature of curve, and finally, the measure of curvature of a surface at a given point, or, as Gauss states, the specific curvature of surface (in contrast to total curvature equal to the area of the spherical image) are expressed through coefficients E, F, and G by the well-known formulae given in every differential geometry course and stemming essentially from Gauss. Gauss also proved the main theorem on the relation of the intrinsic metric of a surface to its spatial shape; the theorem asserts that the product of main curvatures at each point of a surface, which is precisely the measure of curvature of this surface, depends only on the intrinsic metric and is expressed through coefficients E, F, and G by the formula given by Gauss.

By introducing coordinates and the quadratic form ds^2, Gauss gave a general analytical method for studying the intrinsic geometry of regular surfaces. Finally, Gauss himself and his successors proved many theorems of intrinsic geometry, thus showing that a new field open for geometry by Gauss is, indeed, exceptionally fruitful. This was how those general problems of intrinsic geometry discussed above in Sec. 1 were solved for regular surfaces. What remained was the first of the problems posed there. From the point of view of Gaussian theory, this problem is formulated as follows.

Consider a domain D in which coordinates u, v are introduced and three continuous functions $E(u, v)$, $F(u, v)$, and $G(u, v)$ are given such that the quadratic form

$$ds^2 = E du^2 + 2F du\, dv + G dv^2 \tag{4}$$

is positive definite everywhere. It is common parlance to say that the line element ds is given by these data at each point of the domain D. Then the length of a curve $u = u(t)$, $v = v(t)$ in the domain D is defined by equation (4), and so the distance between each two points of the domain D turns out to be defined as the greatest lower bound of lengths of curves connecting these points. The question is: Does there exist a surface whose first quadratic form coincides with a given form in the domain D for a suitable choice of coordinates? The following theorem by Darboux gives an answer to this question "in the small:"[4]

If the coefficients $E(u, v)$, $F(u, v)$, and $G(u, v)$ are analytic functions of u and v, i.e., if they expand in power series in a neighborhood of each point, then the metric given by the line element with these coefficients is locally realizable by an analytic surface, i.e., for each point of D, there exists a neighborhood of this point which can isometrically be mapped onto some analytic surface. Moreover, the coordinates u and v themselves are transferred to this surface (the corresponding points have the same coordinates!), and the coefficients E, F, and G, calculated by the formulas in

[4] This theorem can be found in every comprehensive course of differential geometry; see S. P. Finnikov, Surface theory [in Russian]. As far as we know, this theorem is not proved in the general form under weaker assumptions on the coefficients E, F, and G. If the Gaussian curvature does not change sign, it is likely that the assumption on triple differentiability of the functions E, F, and G is sufficient.

© 2006 by Taylor & Francis Group, LLC

Equation (2) at each point of the surface, coincide with the given values of E, F, and G at the corresponding point of the domain D.

This Darboux theorem shows that the possibility of determining a metric from a line element with analytic coefficients completely characterizes the intrinsic metric of an analytic surface in the small. Thus, in any case, since we speak about the local properties, all reduces to a given arbitrary line element with analytic coefficients. Everything that can be defined and proved by using only this element is meaningful for surfaces in space and will belong to their intrinsic geometry.

The possibility of determining a metric from a line element means that this metric is Euclidean. Strictly speaking, we can prove the following: Let a line element with continuous coefficients E, F, and G be given in a domain D. Let E_0, F_0, and G_0 be the values of these coefficients at the point (u_0, v_0). We take two vectors $\mathbf{a}$ and $\mathbf{b}$ on a plane P such that

$$\mathbf{a}^2 = E_0, \quad \mathbf{b}^2 = G_0, \quad \mathbf{ab} = F_0. \tag{5}$$

Since $E_0G_0 - F_0^2 > 0$, these vectors exist and are linearly independent. We take these vectors as the basis vectors of the coordinate system; denote by p and q the coordinates of this system. Then the distance between two points (p, q) and $(p + \Delta p, q + \Delta q)$ on the plane P is expressed by the equation

$$\rho_0 = \sqrt{E_0\Delta p^2 + 2F_0\Delta p\Delta q + G_0\Delta q^2}. \tag{6}$$

We put the point (p, q) of the plane P into correspondence with each point (u, v) of the domain D such that

$$p = u - u_0, \quad q = v - v_0. \tag{7}$$

Then the point (u_0, v_0) passes to the origin in the plane and $\Delta p = \Delta u$ and $\Delta q = \Delta v$. By the continuity of the coefficients E, F, and G, the expression for the length in the form of integral (3) leads to the fact that the distance ρ between the points (u, v) and $(u + \Delta u, v + \Delta v)$ of the domain D near the point (u_0, v_0) can be represented by the equation

$$\rho = \sqrt{E_0\Delta u^2 + 2F_0\Delta u\Delta v + G_0\Delta v^2} + \delta, \tag{8}$$

where δ is an infinitesimal order higher than the distances from the points (u, v) and $(u + \Delta u, v + \Delta v)$ to the point (u_0, v_0). Comparing equations (6) and (8), we obtain

$$|\rho - \rho_0| < \varepsilon\rho_1,$$

where ρ_1 is the maximum of the distances from the points (u, v) and $(u+\Delta u, v+\Delta v)$ to the point (u_0, v_0) and ε is an infinitesimal together with ρ_1. In brief, our mapping of the domain D onto the plane P near the point (u_0, v_0) turns out to be isometric to within higher order infinitesimals. Hence the geometric meaning of prescribing a line element is that we define a metric in the domain D which coincides in the infinitely small with the metric on the plane, and the coefficients E, F, and G define the mapping of a small neighborhood from D into the plane P by equations (5) and (7), which turns out to be isometric in the infinitely small. The geometric content of the analyticity requirement of the Darboux theorem and the geometric meaning

© 2006 by Taylor & Francis Group, LLC

of the concept of analytic surface are not clear. This requirement depends not on the geometric essence of the problem but rather on the fact that the solvability of the equations to which Darboux reduces the problem is available in general for analytic functions.

Convex surfaces are characterized by the property that their principal radii of curvature cannot have different signs, and so the Gaussian curvature of a convex surface is never negative. Therefore, we can conclude from the Darboux theorem that *for a metric given by the line element ds to be locally realized by a convex surface, it is necessary and sufficient that the line element ds has nonnegative curvature everywhere.* "The curvature of the line element" is defined in terms of the coefficients E, F, and G by the well-known Gaussian equation. Hence, for regular convex surfaces, the question of their intrinsic metrics is completely solved in the small. In general, the intrinsic geometry of regular surfaces in the small is elaborated rather thoroughly.

However, this is not the case for intrinsic geometry in the large. Only two problems on the realization of a given metric that arise here, other than the problems whose solution is almost trivial, were solved. The first of them – the problem of the realization of an arbitrary complete metric of constant negative curvature – was solved negatively. Namely, Hilbert proved that there is no complete regular surface of constant negative curvature.[5] The second problem is to prove that any metric given on a sphere by a line element with analytic coefficients and everywhere positive curvature is realizable by an analytic closed convex surface.[6] This problem was posed by Hermann Weyl in 1915; Weyl sketched a solution, very simple conceptually but so difficult to realize that did not complete it himself. The gap in Weyl's arguments was removed by Hans Lewy in 1938 based on general theorems on partial differential equations of the Monge–Amperè type he obtained, and the theorem formulated by Weyl was proved.[7] Thus, the complete characterization of the intrinsic metric of analytic closed surfaces of curvature positive everywhere was given. (It is easy to prove that each closed surface of a continuous positive curvature is a convex surface.)

In Chapter VII, we shall prove the theorem whose particular case is the following assertion:

[5] See W. Blaschke, Differential geometry, Sec. 96. A metric is called complete if any infinite (and bounded in the sense of this metric) set of points in the domain of the definition of the metric has a condensation point.

[6] The fact that the metric is given on the sphere means that the sphere is accepted as the domain D. However, it is not possible to introduce coordinates u and v such that there is a bijective bicontinuous correspondence between the pairs u, v and the points of this sphere. Therefore, we have to cover the sphere by finitely many domains D_i, each of which with their own coordinates; in the common part of two domains, passage from one coordinate system to the other must be given by analytic functions whose Jacobian does not vanish.

[7] H. Weyl, Über die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement, Vierteljahresschrift der naturforschenden Gesellschaft in Zürich, Vol. 61 (1915), pp. 40–72. H. Lewy, On the existence of a closed convex surface realizing a given Riemannian metric. Proc. Nat. Acad. Sci. USA, Vol. 24 (1938), pp. 104–106.

We can also note that Blaschke and Herglotz indicate another approach to the solution of the Weyl problem by reducing it to an extremal problem. But no solution is mentioned. See W. Blaschke und G. Herglotz, Über die Verwirklichung einer geschlossenen Fläche mit vorgeschriebenem Bogenelement im Euklidischen Raum, Sitz-Berichte Bayer. Akad. Wiss. 1937, Heft 2, pp. 229–230.

© 2006 by Taylor & Francis Group, LLC

Each metric on the sphere given by a line element of nonnegative curvature can be realized by a closed convex surface. This theorem is slightly more general than the Weyl theorem, since the analyticity assumption is omitted here while the curvature can vanish and have discontinuities. However, this theorem does not embrace the Weyl theorem, since there is no assertion on the analyticity of the surface realizing an analytic metric. We can prove that this surface is smooth on condition of the existence and boundedness of curvature, but it may fail to be twice differentiable. Thus, the existence of intrinsic curvature does not ensure the existence of definite principal curvatures at each point of a surface.

3. A Polyhedral Metric

Even for such simple nonsmooth surfaces as polyhedra, there is no first quadratic form and, consequently, no Gaussian curvature at vertices, though both exist at all other points. Therefore, it becomes more convenient to use other notions here. Of course, we can define the first form everywhere, except for some exceptional points, and characterize these points by some additional conditions, e.g., in the case of a convex polyhedron by the condition that the length of a circle centered at a vertex is less than $2\pi r$. Herewith, the length of the circle can be calculated from the first quadratic form. If we return to general convex surfaces, the matter becomes more intricate, and these additional conditions are not obvious here at all. Besides, the set of points at which there is no tangent plane and, consequently, the points where a surface has no curvature, becomes very complex. The discontinuities and the absence of the coefficients E, F, and G and their derivatives force us to use the tools of real function theory, Lebesgue and Stieltjes integrals, Radon integrals, etc. Fearing to be bogged down in the morass of analytical difficulties and to lose any bit of visuality, we thus turn from analytical generalizations and methods to other geometrical analogs on assuming that the concepts of distance, continuous mapping, etc. belong to geometry. In any case, a systematically geometric theory must start from geometrical foundations rather than adapt itself to the technique of analysis.[8]

Any surface, in any case, a sufficiently good one, so to say, can be approximated by a polyhedron with any accuracy. In particular, a closed convex surface can be approximated by closed convex polyhedra (see Appendix, Sec. 6). A polyhedron is a surface composed of plane polygons, and, therefore, the polyhedra are in a certain sense the simplest surfaces. We need more elementary methods for their study, and the results concerning them are absolutely elementary and visual. At the same time, since we can approximate surfaces by polyhedra in general, it is natural to expect that many results abstracted for polyhedra can be further extended to the surfaces presented as limits of the already-inspected polyhedra by passage to the limit. However, this natural method has been elaborated in surface theory to practically no extent, since it was possible to solve the problems of this theory by

[8] The well-known example of discrepancy between an analytical formula and a geometric notion is an expression of the length of a plane curve as the integral $\int \sqrt{\dot{x}^2 + \dot{y}^2}\, dt$. This integral is not always equal to length, and its existence does not imply the existence of the length of curve $x = x(t), y = y(t)$ although the existence of length implies the existence of this integral, the latter being understood in the Lebesgue sense.

© 2006 by Taylor & Francis Group, LLC

analytical methods.[9] But in the case where these methods become overcomplicated, a direct application of geometric methods can often not only lead to the desired aim faster but also bring about a clearer and geometric understanding of the facts. Starting from these remarks, we assign ourselves the task of, first, studying the intrinsic metric of convex polyhedra and, second, of abstracting the so-obtained results to arbitrary convex surfaces by passage to the limit.

Polyhedra can be defined in different ways. The following definition is convenient for our purposes: a polyhedron is a surface composed of a finite or infinite number of plane polygons, the faces; moreover, for each point of a polyhedron, there exists a neighborhood belonging only to a finite number of faces. Since a polyhedron is defined as a surface, its every point has a neighborhood homeomorphic to a disk.

Points of a polyhedron can be of the following three kinds: the points lying in the interiors of faces, the points lying in the interiors of edges, and the vertices. Each edge belongs only to two faces since otherwise the points of their edge would have no neighborhood homeomorphic to a disk. Therefore, a neighborhood of a point that lies in the interior of some edge is a piece of a dihedral angle. Such an angle can be straightened onto a plane; this makes it clear that the neighborhoods of points lying in the interior of edges or faces are the same in the sense of intrinsic geometry. A neighborhood of a vertex A is obtained if we take all points of the polyhedron that are at a distance less than a given $r > 0$ from A; moreover, r can be taken so small that the so-obtained neighborhood contains no vertices but A. This neighborhood is a disk of radius r on the dihedral angle at the vertex A. If the sum of angles on the faces meeting at the vertex A is equal to θ, then the length of the circle of this disk is equal to θr. Therefore, if $\theta \neq 2\pi$, then neighborhood of this vertex is not isometric to any part of the plane.

We now introduce a cone, i.e., a surface formed by the rays, the cone generators, going from the center of some sphere to all points of a closed rectifiable curve, the base line of this cone, lying on this sphere. The directrix has the property that all its points are equidistant from the vertex. Therefore, if we cut the cone along its ruling, then it becomes possible to unfold this cone onto the plane; with this procedure, the rulings pass to half-lines emanating from the image of vertex, and the directrix unfolds onto the circle centered at this point, keeping length. As a result, this cone covers a certain angle on the plane; of course, this angle can be arbitrarily greater than 2π. This angle is called the *complete angle at the vertex of a cone*. If this angle is equal to 2π, then the cone unfolds onto the whole plane in a one-to-one fashion, and so this cone is isometric to the plane. A polyhedral angle, a dihedral angle, and a plane are particular cases of a cone. On the other hand, it is clear that every cone admits bending, i.e., it can be mapped isometrically onto a polyhedral angle, a dihedral angle, or a plane. However, we prefer to speak about a cone in general in problems of intrinsic geometry in order to abstract from any special spatial shape.

[9] From the papers studying surfaces by polyhedral approximation but not related to the concept of surface area, we are aware of only the following: the article by Maxwell (a well-known creator of electromagnetic theory) "Transformation of surfaces by bending," Scientific papers, Vol. I, p. 80; Minkowski's paper "Volumen und Oberfläche," Gesammelte Abhandlungen, Vol. II, p. 230; and also some papers related to Minkowski's paper, and the paper by Fischer, "Herleitung von drei symmetrischen Gleichungen der Flächentheorie...," Deutsche Math., Vol. 3 (1938), pp. 402–437.

© 2006 by Taylor & Francis Group, LLC

A polyhedron is characterized by the property that the sum of the angles meeting at each of its vertices is always less than 2π. This immediately implies the theorem that characterizes the intrinsic geometry of convex polyhedra in the small. *A metric space is locally isometric to a convex polyhedron if and only if for each point of this space there exists a neighborhood isometric to a cone whose complete angle at the vertex is less than or equal to* 2π. Obviously, the angle equal to 2π relates to the points other than vertices.

As we see, the solution of the problem of the intrinsic geometry of polyhedron in the small is very simple, but this solution is infinitely distant from that of the problems of even the local intrinsic geometry of general surfaces, not mentioning the problems in the large. In fact, if we approximate a certain curved surface by polyhedra, then the number of the vertices of these polyhedra increases indefinitely, and even arbitrarily near each point of the surface this number tends to infinity. Hence, from the consideration of only those neighborhoods that contain no more than one vertex of polyhedra we obtain no information on the whole surface on passing to the limit. Roughly speaking, in order to pass to a surface, it is necessary to study at least a part of polyhedra "in the large."

For a metric space to be isometric to a polyhedron not only in a small neighborhood, it is necessary in any case that every two points of this space can be connected by a curve of finite length and the distance between the points equal the greatest lower bound of the lengths of such curves. This condition is necessary since the distance on a surface is defined precisely in this way. In more detail, this condition should be understood as follows. Let R be a metric space whose every point has a neighborhood isometric to a cone. Let X and Y be two arbitrary points of space R, and let L be a continuous curve connecting these points.[10] The length of the curve is defined as follows: each point of the curve is surrounded by a neighborhood isometric to a cone; using the Borel lemma, we refine finitely neighborhoods that also cover the whole curve. In each of these neighborhoods, the length of a segment lying in this neighborhood is defined in the same way as on the cone, for example, by developing this cone onto the plane. If a segment of the curve has no length, we can replace it by another segment that already has length. The length of the whole curve connecting the points X and Y is then defined as the sum of the lengths of all segments. If the length of a curve was defined, then it becomes clear what the greatest lower bound of the lengths of the curves connecting X and Y is, and thus the meaning of the imposed condition is revealed.[11]

[10] Assume that a parameter t ranges over the closed interval $[0, 1]$; to each t, we put in correspondence the point $X(t)$ of the metric space R such that (1) $X(0) = 0$, $X(1) = 1$; (2) $X(t)$ depends continuously on t, i.e., for all t and $\varepsilon > 0$, there exists $\delta > 0$ such that $\rho(X(t)X(t')) < \varepsilon$ whenever $|t-t'|, \delta$; here, ρ stands for the distance on R. This defines a convex curve in R. More precisely, by a continuous curve $X(t)$ we should mean the set of pairs consisting of the point $X(t)$ and the value t rather than the set of points $X(t)$. These are not equivalent when a curve has multiple points, i.e., if $X(t_1) = X(t_2)$ for a certain $t_1 \neq t_2$; in this case, we have the same point but different pairs.

[11] We see that the length of a curve is defined automatically if each point has a neighborhood isometric to a cone. Nevertheless, the indicated condition is not superfluous. This is seen from the following example. Introduce a new distance $\rho(XY)$ instead of the usual Euclidean distance $\rho_0(XY)$ as follows: if $\rho_0(XY) \leq 1$, then we put $\rho_0(XY) = \rho(XY)$; if $\rho_0(XY) > 1$, then we put $\rho(XY) = 1$. The plane becomes a metric space whose every point has a neighborhood isometric to a disk, namely, to the disk of radius $1/2$. Meanwhile, the distance in this space is not equal to the greatest lower bound of the lengths of curves in this space in general.

© 2006 by Taylor & Francis Group, LLC

A metric space in which (1) each point has a neighborhood isometric to a cone and (2) the distance between two points is equal to the greatest lower bound of the lengths of the curves connecting these points shall be called a *space with polyhedral metric*. A polyhedral metric shall be called a polyhedral metric of *positive curvature* if each point has a neighborhood isometric to a cone with complete angle of, at most, 2π at the vertex. The meaning of the term "positive curvature" will be explained in Sec. 8. For a space to be isometric to a convex polyhedron, it is necessary that this space has a polyhedral metric of positive curvature. As we have indicated, this condition is also sufficient in the small.

A closed convex polyhedron is homeomorphic to the sphere, and so, for a metric space to be isometric to a closed convex polyhedron, it is necessary that this space is also homeomorphic to the sphere and has a convex polyhedral metric. It turns out that these conditions are also sufficient, but with only one stipulation. Namely, we also include plane convex polygons to the set of convex polyhedra assuming however that these polygons are "doubly covered." This means that each polygon has two sides, say, upper and lower, and we assume that the points on the upper and lower sides are different even if these points coincide. Since we need an analogous idea in a more general case, we will give its precise definition once for any arbitrary domain on the plane.

Assume a finite convex domain G bounded by a curve L, adjoined also to this domain, so that G is closed. Take a sphere S and the equator E on this sphere and map S in a continuous way onto the domain G so that the equator E is bijectively and bicontinuously mapped onto the boundary curve L, and each of the two spheres is also bijectively and bicontinuously mapped onto the domain G. As a result, the interior of the domain G turns out to be the doubly-covered image of the sphere S. This mapping can be conventionally considered as bijective and bicontinuous, i.e., as a homeomorphism. It is convenient to visualize some points as lying on one side of the domain G and the others, as lying on the other side. If two points X and Y lie on one side, then they can be connected by a segment (since the domain is convex), and this segment lies on the same side as the points X and Y. If the points X and Y lie on different sides, a curve between these points traveling from the point X along the side on which this point lies must cross the boundary and pass to the other side where the point Y lies. In this case, we define the distance between the points X and Y as the greatest lower bound of the lengths of such curves passing through the boundary of the domain G. Then even if the points X and Y are really identical, the distance between them is other than zero whenever they lie on the opposite sides of the domain G. We thus transform the domain G into a closed convex surface. This seems to be natural if we consider a closed convex surface that is flattened almost to a plane domain. For example, we can take a surface composed of two cones based on the boundary of the domain G. When the heights of these cones tend to zero, each of these cones tends to one of the sides of the domain G. Thus, the doubly-covered domain G is a limit case of a closed convex surface in which the latter is completely flatten.

As an example, we consider a doubly-covered equilateral triangle. An interior point of this triangle has a circular neighborhood. If we take a point inside a side, then its every neighborhood consists of two superposed half-disks; if we unfold this neighborhood along the diameter lying on the side of the triangle, then it becomes a

© 2006 by Taylor & Francis Group, LLC

usual disk. Each neighborhood of a vertex of the triangle consists of two superposed sectors that are pasted together along the radii from the vertices along the sides of the triangle; if we press down these radii, then these sectors bend and, going away from each other, form a cone with a complete angle at the vertex equal to 120°. Hence we see that a doubly-covered triangle does satisfy the necessary conditions indicated above; that is, this triangle is homeomorphic to the sphere (in the sense in which this conditional homeomorphism was defined), and each point of this triangle has a neighborhood isometric to a cone with complete angle of, at most, 2π at the vertex. It is easy to prove that there is no other convex polyhedron isometric to a doubly-covered triangle except for this triangle itself. Therefore, this is the only way to consider an abstractly defined metric space isometric to a doubly-covered triangle as a convex polyhedron. Hence if we do not want to eliminate such cases by special conventions, then the doubly-covered triangles should be included into a list of convex polyhedra.

Considering the doubly-covered triangles as polyhedra, we formulate the theorem that characterizes the intrinsic metric of a closed convex polyhedron.

A metric space is isometric to a closed convex polyhedron if and only if (1) *this space is homeomorphic to the sphere;* (2) *for each of its points, there exists a neighborhood isometric to a cone with complete angle* $\leq 2\pi$ *at the vertex;* (3) *the distance in this space is equal to the greatest lower bound of the lengths of curves.*

Since we have seen that the necessity of these conditions is obvious, this theorem can be viewed as a theorem on the existence of a closed convex polyhedron with an *a priori* given polyhedral metric of positive curvature. If we take a point inside a closed convex polyhedron P, draw a sphere S with center at this point, and project the polyhedron to the sphere S from this point, then we obtain a bijective and bicontinuous correspondence between the points of the sphere and those of the polyhedron. If the points x' and Y' on P correspond to the points X and Y on S, then assigning the distance between the points X' and Y' on the polyhedron to the points X and Y, we so translate the metric of the polyhedron to the sphere S. In general, every metric space homeomorphic to the sphere can be mapped onto the sphere, and we can view the metric of this space as given directly on the sphere. This understanding is convenient, since we obtain here a completely definite domain, the sphere, on which the metric is given. Precisely this understanding will be constantly implied. Using it, we can formulate the theorem on the existence of a convex polyhedron with given metric in the following form: *for each polyhedral metric of positive curvature on the sphere, there exists a closed convex polyhedron* (*or a doubly-covered convex polygon*) *realizing this metric.*

This theorem is one of the main theorems of the entire theory set forth in this book. Chapter VI is devoted to the proof of this theorem.

4. Development

It is perhaps somewhat strange that, when speaking about polyhedra, we use such abstract concepts as metric space, homeomorphisms, etc. This is done deliberately in order to take into account these concepts by examining a simple example of a polyhedron and to set the study of polyhedra along with that of any convex surfaces,

© 2006 by Taylor & Francis Group, LLC

where these concepts turn out to be not very general because of the complexity of the problem. However, when turning to polyhedra, we have especially stipulated the elementary and illustrative character of results that refer to them. Besides, speaking about a metric space that is isometric, say, to a convex polyhedron, we do not know at all how this space can be assigned in reality. Naturally, this abstractness should be eliminated and everything should be reduced to a form, which can indeed be considered elementary. To this end, we shall use the assignment of a polyhedron by its development.

A *development* is nothing else than a set of polygons that is assigned the way in which they should be glued to each other along sides and vertices. If we even exclude the gluing of vertices that are not caused by that of sides (e.g., this gluing occurs in a polyhedron composed of two tetrahedra that are adjoined to each other at a vertex), we cannot avoid the gluing of vertices, since for two sides, there always exist two possibilities of gluing, that is, in one opposite direction or another. Two sides AB and CD can be glued in such a way that the following vertices coincide: A with C and B with D, or A with D and B with C; of course, these two possibilities should be considered as different ones.

It is well known that the cube can be glued from one cross-like polygon. Therefore, there is no necessity to assume that all sides and vertices being glued belong to distinct polygons. Meanwhile, we assume that the gluing of polygons is performed along the whole sides. This is not a restriction, since, if necessary, any point inside a side can be considered a vertex. Further, it is convenient to assume that each of the polygons of the development is bounded by one closed broken line without multiple points. This can be attained if we cut polygons bounded by several broken lines or by a broken line with multiple points in an appropriate way.

Polygons of any development are called *faces*; sides of these polygons are called *edges*; moreover, the sides that are glued to each other are certainly considered as one edge; finally, vertices of a polygon are called *vertices* of the development; moreover, the vertices that are glued to each other are again considered as one vertex. Herewith, the gluing is certainly not assumed to be realized in reality; only the law of gluing is stated and, say, the ideal identification of sides and vertices being glued is indicated.

Since the sides being glued are identified, a neighborhood of a point lying on a side is assumed to be composed of parts of those faces to which this side belongs. In exactly the same way, a neighborhood of a vertex is formed by sectors that are cut out from all faces to which this vertex belongs. For each point of a polyhedron, there exists a neighborhood that is homeomorphic to a disk. This leads to the following conditions that a development should satisfy.

1. All angles of polygons approaching one vertex should form a cyclic sequence when they are glued to each other along sides according to the gluing law. In more detail, if O is a vertex, then taking a polyhedron P_1 approaching this vertex by its sides a_1 and a_2, we glue the polygon P_2 to the side a_2; there is a side a_3 of the polygon P_2 that approaches the vertex O; we glue the polygon P_3 to this side, etc., according to the gluing law. The condition is that eventually we approach the polygon P_n, which should be glued to the first polygon P_1 along the free side a_1 after gluing with the polygon P_{n-1} along the side a_n. Thus, this sequence is closed,

© 2006 by Taylor & Francis Group, LLC

and all polyhedra approaching the vertex O should be exhausted. As a result, a neighborhood of the vertex O consists of sectors glued to one another in a cyclic order so that a figure homeomorphic to a disk is obtained.

It is not to be ruled out that some polygons approach the vertex O several times. In the same way, it is possible that only one polygon approaches the vertex O by its angle whose sides are glued to each other in such a way that a neighborhood of the vertex O is something like a "package." In this case, the whole cyclic sequence of sectors is reduced to a single sector. The possibilities mentioned here can be observed from the development of the cube (Fig. 3), where a neighborhood of the vertex O consists of a single angle, and a neighborhood of the vertex A consists of three angles of the same polygon. The glued vertices are denoted by the same letters in this figure.

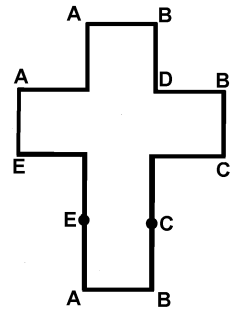


Fig. 3

The formulated condition obviously implies that each side of one polygon should be glued with only one side of another.

2. We can pass from one polygon to another going along the polygons that are glued with each other, i.e., which have sides and vertices identified with each other.

This second condition ensures the connectedness, since otherwise the development falls into parts from which no connected polygon could be glued.

For completeness, it is necessary to add the following condition to these two conditions. Its necessity is obvious.

3. The identified sides should be of equal length, and the coinciding segments of sides being glued should be of equal length.

Now let X and Y be two points of the development. As is clear from condition 2, we can construct a broken line starting at the point X and ending at the point Y whose links lie subsequently on the faces of the development that have common (identified) sides or vertices; if a segment of this broken line arrives at the boundary of some face at a certain point Z, then we continue this segment in the face that is glued to the first one at the point Z. The length of this broken line is defined automatically, since its links lie on polygons, where the length is the usual length of a segment in Euclidean geometry. The greatest lower bound of the length of broken lines connecting the points X and Y in the development is accepted as the distance between these points. Thus, the development assigns the intrinsic metric of a polyhedron in reality. However, a development can be given *a priori* not assuming that one can glue a certain polyhedron of this development, and hence we obtain a method for prescribing a polyhedral metric space.

We can prove the converse assertion, namely that any polyhedral metric can be given by a development. Since we do not need this assertion in such general form, we restrict ourselves to the necessary case of a polyhedral metric on a sphere, when this assertion has a very simple proof.

Let a polyhedral metric be given on a sphere. Then each point of this sphere can be surrounded by a neighborhood that is isometric to a cone, and then this neighborhood can be divided into triangles having a common vertex at this point

© 2006 by Taylor & Francis Group, LLC

which are isometric to ordinary plane triangles. As a result, the sphere S turns out to be covered by such triangles. By the Borel lemma, we can choose a finite number of triangles from these triangles, which also cover the whole sphere S. Overlapping each other, these triangles are themselves divided into polygons that have no common interior points. These polygons cover the whole sphere, and since each of them can be developed onto the plane, these polygons form the development that represents the metric given on S.

The edges of the development form a net of segments. This net is connected, i.e., we can pass from one segment to another proceeding only along the segments of this net. Indeed, let a and b be two edges, and let P and Q be the faces containing these edges. By the connectedness property of the development, we can pass from P to Q going along the faces, which are glued with each other. This passage can be realized by going only along the boundaries of the faces, i.e., along the edges; therefore, we assume that the boundary of each face consists of only one closed broken line. Consequently, we can pass from the edge a to the edge b going along the edges all the time.

If f is the number of faces in the development, k is the number of edges, and e is the number of vertices, then the sphere is divided into k domains with e vertices by a connected net consisting of k segments. In this case, by the Euler theorem,

$$f - k + e = 2.$$

Hence, every development homeomorphic to the sphere must satisfy this Euler formula. Also, it is known that, under the above three conditions, every development satisfying the Euler formula is homeomorphic to the sphere.[12]

Finally, we have the last condition for a convex polyhedron: the complete angle at the vertex of the cone, to which a neighborhood of each point of this polyhedron is isomorphic, must be no more than 2π. This condition holds automatically for the points inside the faces or edges of the development, i.e., the complete angle at each of them is precisely 2π. Therefore, it remains to require the fulfillment of this condition for vertices. The complete angle at a vertex is composed of the angles of the faces of the development which are contiguous to this vertex. Therefore, our condition reduces to the following: the sum of the angles meeting at each vertex should be $\leq 2\pi$.

Thus, the theorem on convex polyhedra, which is formulated in Sec. 3, can be rephrased in terms of developments as follows:

The development of each convex polyhedron satisfies the three above-formulated necessary conditions on each development, the Euler condition, and, finally, the condition that the sum of the angles at each of its vertices is $\leq 2\pi$. *At the same time, a closed convex polyhedron* (*or a doubly-covered convex polygon*) *originates by gluing from each development satisfying these conditions.*

It is noteworthy that only two such polyhedra can be glued; one is obtained from the other by turning out the seamy side or, which is the same, by reflection in the plane. (Of course, if a polyhedron has a plane of symmetry, these two polyhedra

[12] A perfectly elementary proof of this assertion and even a general topological classification of developments can be found in the books by H. Seifert and W. Threlfall, Topologie, Chap. VI, or P. S. Aleksandrov and V. A. Efremovich, Essay on the Basic Concepts of Topology, Chap. 1.

© 2006 by Taylor & Francis Group, LLC

do not differ from each other. On the other hand, we can glue arbitrarily many nonconvex polyhedra from the development of a convex polyhedron: it suffices to pass the polyhedron down inside itself in a neighborhood of a vertex.)

The recruitment of doubly-covered polygons becomes very comprehensible here. For example, if a development consists of two equal triangles whose equal sides, together with vertices, are identified pairwise, then the polyhedron resulting from this development represents these two triangles passed together in a single triangle so that the latter is doubly covered in a literal sense.

In the form presented here our theorem turns out to be perfectly elementary; that is, if we are given a real development, say, in the form of polygons cut out from a sheet of paper, with sides and vertices to be glued denoted by the same letters, then we can always verify the fulfillment of the conditions of the theorem for this development in finitely many steps. So that we can give an answer to the question whether or not we can obtain a convex polyhedron from this development. If we obtain such a polyhedron, then this polyhedron is unique up to motion and reflection. Therefore, if we begin to glue this polyhedron, it will take its shape in a sense automatically. However, we can say nothing other than make a few obvious but completely insufficient remarks about this shape and how the faces of the development are to be flexed when producing the faces and edges of the polyhedron we glue.

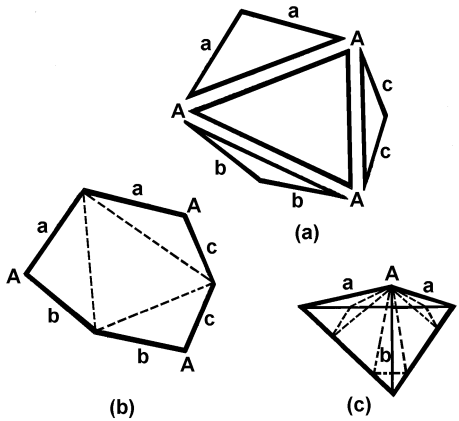


Fig. 4

We illustrate this fact with a curious example that will show us how little the faces and edges of the development may have to do with those of the polyhedron. Figure 4(a) shows the development of a tetrahedron; Fig. 4(b) presents the lines of the development, which will become the edges of the tetrahedron in this development when we glue the elements of the latter into a single polygon; finally, Fig. 4(c) shows the already-glued tetrahedron with the indication of the lines along which this tetrahedron should be cut to give the development of Fig. 4(a); besides the dotted lines, the tetrahedron should be cut along the edges meeting at the vertex A. This development is also curious in that it demonstrates rather intricate possibilities of gluing the sides and vertices of the same polygon of a development (in this case, with regard to each of the four triangles).

We imagine a development as given directly by polygons with instructions for the rules of gluing. Although this point of view, reminding us of a sheet of paper, scissors, and glue, corresponds most perfectly to the gist of the matter, it might not satisfy a mathematician. In this connection, we note that if all polygons of the development are triangulated, then we can define the structure of this development by incidence matrices and assign some length to each edge. This approach to the

© 2006 by Taylor & Francis Group, LLC

abstract definition of a development allows us to determine each polyhedral metric from a simple scheme avoiding the visual ideas of polygon gluing, etc.

5. Passage from Polyhedra to Arbitrary Surfaces

After the problem of the intrinsic metric of closed convex polyhedra has been solved, a way to solving the problem of the intrinsic metric of an arbitrary closed convex surface is charted on the basis of the following "theorem on the convergence of metrics" which will be proved in Sec. 1 of Chapter III.

If a sequence of closed convex surfaces F_n $(n = 1, 2, \ldots)$ converges to a closed convex surface F and a sequence of pairs of points X_n and Y_n on the surface F_n converges to a pair of points X and Y on F $(X_n \to X$ and $Y_n \to Y)$, then the distances on F_n between X_n and Y_n converge to the distance on F between the points X and Y, i.e.,

$$\lim_{n\to\infty} \rho_{F_n}(X_nY_n) = \rho_F(XY).$$

Let F be a closed convex surface, and $P_1, P_2, \ldots$ be a sequence of convex polyhedra converging to this surface. Take a point O inside F and also inside all P_n; this is possible whenever P_n are sufficiently close to F. Circumscribe a sphere S around O, and project the surface F and all polyhedra P_n to this sphere from the point O. Then the metrics $\rho_F(XY)$ and $\rho_{P_n}(XY)$ are transferred from F and P_n to the sphere S if the distance between the corresponding points on the surface F and on the polyhedra P_n, respectively, is put in correspondence to a pair of points A and B on S as the distance between A and B. The just stated theorem on the convergence of distances implies that for each pair of points A and B on the sphere S we have the relation

$$\rho(AB) = \lim_{n\to\infty} \rho_n(AB),$$

where $\rho(AB)$ and $\rho_n(AB)$ are the distances transferred from F and P_n to the sphere S by the above construction. Hence the metric of a convex surface, transferred to the sphere, is the limit of the metrics of convex polyhedra.

Assume a metric $\rho(XY)$ on the sphere that is the limit of polyhedral metrics $\rho_n(XY)$ of positive curvature on the same sphere; i.e., for each pair of points X and Y on the sphere S, the following equality holds:

$$\rho(XY) = \lim_{n\to\infty} \rho_n(XY).$$

Supposing that the realization theorem for a polyhedral metric of positive curvature on the sphere is already proved, we can construct a sequence of closed convex polyhedra P_n realizing these metrics ρ_n. We can choose a convergent subsequence from this sequence.[13] Some closed convex surface F is the limit of this sequence.

Since, on the one hand, the metrics ρ_n converge to the given metric ρ on the sphere S; and, on the other hand, by the above theorem, the metrics of polyhedra

[13] This can be done, e.g., by the well-known "selection theorem" by Blaschke. In Sec. 7 of Chapter VII, which gives a rigorous justification of the assertion sketched here, we provide a direct proof of this property without any reference to the Blaschke theorem.

© 2006 by Taylor & Francis Group, LLC

converging to the surface F converge to the metric on F, we obtain that the surface F has precisely the metric $\rho(XY)$.

Thus, we arrive at the following theorem.

Theorem. *A metric given on a sphere can be realized by a closed convex surface if and only if this metric is the limit of a sequence of polyhedral metrics of positive curvature that are given on the same sphere.*

This theorem solves the problem of characterizing the intrinsic metric of closed convex surfaces. But the theorem is not rich in content, because it says nothing about the intrinsic geometry of convex surfaces except for the fact that the metric of a convex surface is the limit of metrics of convex polyhedra. We remain perfectly ignorant of what properties of a metric characterize the fact that this metric is or is not a limit of polyhedral metrics of positive curvature. However, this theorem yields a method for the study of the intrinsic metric of a convex surface. Suppose we want to prove that a metric on the sphere is realized by a closed convex surface if and only if this metric satisfies some set of conditions A. To prove the necessity of A, we first can prove these conditions for closed convex polyhedra, and then, by passage to the limit precisely as described in the above theorem, we transfer these conditions to all closed convex surfaces. To prove the sufficiency of conditions A, we can prove first that a metric given on the sphere and satisfying these conditions can be represented as a limit of polyhedral metrics of positive curvature; then this theorem will imply the sufficiency of conditions A.

This is a particular case of polyhedral approximation, which in the general form looks as follows. Assume that we want to prove property B of convex surfaces. To this end, first we prove that this property holds for convex polyhedra. Then we prove a convergence theorem of the following type: if property B holds for convex surfaces F_n converging to a convex surface F, then this property is true for the surface F. Since convex surfaces can be approximated by polyhedra, the combination of both results demonstrates that property B holds for all convex surfaces. This way of arguing will be used for proving quite a few theorems.

6. A Manifold with an Intrinsic Metric

Here, we prove two of the most important properties of the intrinsic metric of convex surfaces. Let F be a certain convex surface, and let O be a point of this surface. We can define a neighborhood of the point O in two ways, "intrinsic" and "extrinsic." In the former case, a neighborhood of the point O of the surface F is a part of F lying in a ball that is circumscribed around O, or in general, each set on the surface F, which includes such a part of this surface F. In the latter case, a neighborhood of the point O is each set of points of F, which includes some geodesic disk centered at the point O.[14] A geodesic disk is the set of points that lie from a given point at a distance less than some given distance in the sense of the intrinsic metric on F. Let

[14] See the remark at the beginning of Sec. 1 on the distinction of this definition from the conventional definition.

© 2006 by Taylor & Francis Group, LLC

us prove that both definitions are equivalent; i.e., each "extrinsic" neighborhood is "intrinsic," and vice versa.[15]

Let U be an "extrinsic" neighborhood of the point O on the convex surface F. By definition, there exists $r > 0$ such that each point on F lying in the space at a distance of less than r from O belongs to U. But the distance from O to each point X measured on the surface is obviously no less than the distance from O to X in the space. Consequently, U contains all points of the surface F which lie in the geodesic disk of radius r centered at the point O. Hence U is also an "intrinsic" neighborhood.

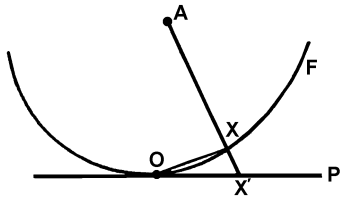


Fig. 5

Conversely, let U be a given "intrinsic" neighborhood of the point O, and let r_0 be the radius of the geodesic disk centered at O, which is contained in this neighborhood. Show that U is also an "extrinsic" neighborhood, i.e., U contains all points of the surface F lying at a distance of, at most, some $r > 0$ from the point O, provided that the distance is taken in the space. To prove this, draw a supporting plane P to our surface at the point O (Fig. 5 shows the section of the surface F by the plane passing through O). The surface F is a domain on the boundary of a certain convex body. Take a point A inside this body and project the surface F to the plane P from this point. The projection of the surface F onto the plane P is bijective and bicontinuous near the point O. Therefore, the projection of F covers a certain neighborhood of the point O on the plane P. Take the disk K of radius r centered at the point O on the plane P such that the projection of the surface F covers this disk. Let V be the part of the surface F whose projection is this disk K. This V is nothing more than a certain "extrinsic" neighborhood of the point O. Let h be the maximum of the distances from points X of the neighborhood V to the corresponding points X' of the disk K counted in the direction of the projection, so that $XX'' \leq h$. If the radius r of the disk K tends to zero, then obviously h also tends to zero. Therefore, we can choose a small r such that

$$r + h < r_0,$$

where r_0 is the radius of the geodesic disk lying in the given "intrinsic" neighborhood U of the point O. We assert that for this choice of the radius R the "extrinsic" neighborhood V lies in U. Indeed, take some point X in the neighborhood V and

[15] This fact does not hold for every surface. For example, take a closed Weierstrass curve L on the plane which has no tangent anywhere and, so, has no length. (We can say that the length of each arc of this curve is infinite.) Connecting all points of this curve by line segments with a certain point O not lying in the same plane, we obtain the cone with vertex O and base line the curve L. If two points X and Y lie on different generators then the shortest arc connecting them on this cone consists of the segments OX and OY, since the length of a curve intersecting the generators is necessarily infinite. This cone has the intrinsic metric, but a small intrinsic neighborhood of the point X of a ruling consists of only a segment of this ruling, since the point X is "far" from the points of other rulings; the shortest join goes through the vertex O. The point O is a unique point with a circular intrinsic neighborhood.

© 2006 by Taylor & Francis Group, LLC

draw the plane Q through the points A, O, and X. This plane intersects the disk K along its diameter. This diameter is the projection of the arc of the convex curve along which the plane Q intersects the surface F. The length of the convex arc $\overset{\frown}{OX}$ is no greater than the union of the segments $\overline{OX'} + \overline{XX'}$.[16] But the length of $\overset{\frown}{OX}$ is at lest the distance $\rho(OX)$ on the surface F, and $\overline{OX'} \leq r$ and $\overline{XX'} \leq h$. Therefore, $\rho(OX) \leq r + h < r_0$. This means that the point X belongs to the geodesic disk of radius r_0. Hence the "extrinsic" neighborhood V lies in the given "intrinsic" neighborhood U, so that U is also an "extrinsic" neighborhood; this is what we have to prove.

When we claim that each point of a surface has a neighborhood homeomorphic to a disk, we certainly speak about an "extrinsic" neighborhood since the intrinsic metric was not even mentioned. The introduction of the intrinsic metric allows us to introduce the concept of an intrinsic neighborhood; in intrinsic geometry, by a neighborhood, we always mean exactly an intrinsic neighborhood. Since both definitions of neighborhoods are equivalent for convex surfaces, we do not distinguish them in what follows. Further in the book we can assume an intrinsic neighborhood in all assertions on convex surfaces, which involve the concept of neighborhood. First of all, the concept of continuity is meant; as is known, the definition of this concept involves only the concept of neighborhood. Thus, we obtain the following assertions:

1. each point of a convex surface has an intrinsic neighborhood homeomorphic to a disk;

2. each two points of a convex surface can be connected by a curve that is continuous in the intrinsic sense of the metric;

3. each closed convex surface admits bijective and bicontinuous mapping onto a sphere (with continuity understood in the sense of the intrinsic metric of the surface);

4. each convex surface is homeomorphic to a domain on the sphere. (This fact is obviously implied by item 3 for a convex surface that is closed or that is only a part of a closed surface. Also, the plane and cylinder are homeomorphic to a domain on the sphere; therefore, convex surfaces that are parts of infinite complete surfaces are also homeomorphic to domains on the sphere.)

Consequently, *a metric space isometric to a convex surface is homeomorphic to a domain on the sphere.*

The intrinsic metric of a surface is characterized by the property that the distance is defined as the greatest lower bound of the length of curves. This condition should also be extended to the metric spaces under consideration. To this end, we define an object that will be called the length of a curve in a metric space.

[16] The convex curve $\overset{\frown}{OX} + \overline{OX}$ composed of the arc $\overset{\frown}{OX}$ and the segment $\overline{OX}$ lies in the triangle OXX', and so (by Theorem 3 of Sec. 3 in the Appendix)

$$s(\overset{\frown}{OX}) \leq s(\overline{OX'}) + s(\overline{XX'}).$$

© 2006 by Taylor & Francis Group, LLC

Let $X(t)$ $(0 \le t \le 1)$ be a continuous curve in a metric space R with metric $\rho(XY)$. Partition the closed interval $[0, 1]$, the range of the parameter t, by points $0 = t_0 < t_1 < \ldots < t_n = 1$. Consider the sum

$$\sum_{i=1}^{n} \rho(X(t_{i-1}), X(t_i))$$

of the distances between all pairs of consecutive points ($X(t_0)$ and $X(t_1)$, $X(t_1)$ and $X(t_2)$, etc.). The greatest lower bound of these sums over all possible partitions of the closed interval $[0, 1]$ is called the length of the curve $X(t)$. If this greatest lower bound does not exist or, which is the same, equals infinity, then this curve has no length or, say, has infinite length. This definition of length literally repeats the definition of length accepted in Euclidean geometry. (More often, the length is defined as the limit of the indicated sums when the range of the parameter t is indefinitely refined; nevertheless, from the triangle inequality, the well-known fact follows easily that both these definitions are equivalent.) Also, it is easy to prove that the so defined length has all properties of the usual length in Euclidean space; this will be done in Sec. 1 of Chapter II.

If we have some surface F and a continuous curve L on this surface, the length of this curve can be defined by the following two methods: first, as the length in the space, and second, as the length in the sense of the intrinsic metric of the surface F. In the former case, the length of this curve is

$$s_0(L) = \sup \sum \rho_0(X(t_{i-1})X(t_i)) \tag{1}$$

(the symbol sup stands for the greatest upper bound), where ρ_0 is the distance in the space; in the latter case, the length is

$$s(L) = \sup \sum \rho(X(t_{i-1})X(t_i)), \tag{2}$$

where ρ is the distance on the surface.

Let us prove that *both lengths coincide, i.e., for each surface, the fact that a given curve L has length in one sense implies that this curve has length in the other sense, and both lengths are equal.*

First of all, we note that by the definition of distance on a surface, this distance is the greatest lower bound of the lengths of curves on this surface; and moreover, these lengths are defined in the space, i.e., by the formulas similar to (1). Therefore, if L_i is the segment of the curve L between the points $X(t_{i-1})$ and $X(t_i)$, the length $s_0(L_i)$ of this segment is at least the distance on the surface between these points; i.e.,

$$s_0(L_i) \ge \rho(X(t_{i-1})X(t_i)). \tag{3}$$

Summing these inequalities and noting that the sum of the lengths of L_i is equal to the length of the whole curve L, we obtain

$$s_0(L) \ge \sum \rho(X(t_{i-1})X(t_i)). \tag{4}$$

Since all these sums are at most $s_0(L)$, the greatest lower bound of them is at most $s_0(L)$; i.e.,

$$s_0(L) \ge s(L). \tag{5}$$

© 2006 by Taylor & Francis Group, LLC

Therefore, the existence of the finite length $s_0(L)$ implies the existence of the finite length $s(L)$.

On the other hand, the distance in the space is at most the distance on the surface (since the distance in the space is taken along the line that is shortest in the whole space rather than only on the surface!), i.e.,

$$\rho_0(X(t_{i-1})X(t_i)) \le \rho(X(t_{i-1})X(t_i)). \tag{6}$$

This and formulae (1) and (2) make it clear that

$$s_0(L) \le s(L). \tag{7}$$

Therefore, the existence of the finite length $s(L)$ implies the existence of the finite length $s_0(L)$.

Inequalities (5) and (7) yield

$$s_0(L) = s(L);$$

as required.

Since the lengths of curves on a surface coincide for both definitions, *the distance on a surface is also equal to the greatest lower bound of the lengths of curves in the sense of the definition of length via the intrinsic metric of this surface.* This formulation is free of any reference to the distance and length in the space; it belongs entirely to intrinsic geometry and, therefore, can be abstracted to metric spaces.

The metric of an arbitrary metric space is called intrinsic *if the distance between each two points of this space is equal to the greatest lower bound of the lengths of curves connecting these points*; *moreover, the length of each curve is defined in terms of the metric by formula* (2). We have proved that *the metric of each surface is intrinsic in the sense of this definition.*

Now, summarizing the conclusions of this section, we can say that *a metric space isometric to a convex surface is homeomorphic to a domain on the sphere and has intrinsic metric.*

A metric space is called a *two-dimensional manifold* if this space is connected and its every point has a neighborhood homeomorphic to a disk (an n-dimensional manifold possesses a neighborhood of its every point which is homeomorphic to an n-dimensional ball).[17] Since we shall deal only with two-dimensional manifolds, they shall be called just manifolds. Each surface in the sense of the general definition of Sec. 1 is a two-dimensional manifold if we measure the distance between its points in the space. In the sense of the intrinsic metric, not every surface is a manifold since intrinsic neighborhoods on this surface are not simultaneously "extrinsic" in general. However, we have shown that all convex surfaces turn out to be manifolds also from the point of view of the intrinsic metric, namely, manifolds homeomorphic to domains on the sphere. In exactly the same way, each polyhedron or each regular surface is a manifold from the point of view of the intrinsic metric. In general, it is appropriate to distinguish especially between those surfaces which are manifolds in

[17] In topology, it is accepted that a manifold is any topological space that has the properties mentioned above and, in addition, admits a partition into simplexes. In our case, this requirement turns out to be extraneous.

© 2006 by Taylor & Francis Group, LLC

the sense of their intrinsic metrics. This property is enjoyed by each surface with intrinsic metric such that the convergence of its points X_n to X as points in the space implies that the distance $\rho(X_nX)$ on the surface tends to zero.[18] Consequently, we have proved that the metric of each surface is intrinsic in the sense that the distance between two points is equal to the greatest lower bound of the lengths of connecting curves which are measured in the metric of this surface. The comparison of these two results generates the concept of a *manifold with the intrinsic metric* as a general object for study, i.e., *a metric space that is a manifold and has intrinsic metric in the sense of the above general definition.*[19] Some general properties of manifolds with intrinsic metric will be considered in Chapter II. Convex surfaces are a particular class of these manifolds.

7. Basic Concepts of Intrinsic Geometry

We give here the definitions of some basic concepts of intrinsic geometry. An arbitrary manifold with intrinsic metric $\rho(XY)$ is considered; in particular, this manifold can be a convex surface. But since our definitions are based only on the concept of distance, they do not thus depend on whether or not this metric is realized by some surface.

First of all, together with the concept of distance, we have the concepts of neighborhood convergence of a sequence of points and other topological concepts, which are defined for each metric space in general. In Sec. 6, we defined the length of curve using a given metric $\rho(XY)$. A curve between two given points that has the minimal length among all curves connecting these points is called a *shortest arc* (a shortest join of these points). If the distance $\rho(XY)$ is equal to the greatest lower bound of the lengths of the curves joining the points X and Y, then the length of a shortest arc XY is obviously equal to the distance $\rho(XY)$. However, there can be no shortest join of two points of a surface; only curves of the length arbitrarily close to the distance should exist. For example, if the surface under consideration is simply a nonconvex plane domain then it always has infinitely many pairs of points like this. However, we shall show below that in each manifold with intrinsic metric two sufficiently close points can always be connected by a shortest arc, and on each closed and even each complete convex surface, each two points can be connected by a shortest arc.

A *geodesic* is each continuous curve that is a shortest arc on its every sufficiently small segments. More precisely, a geodesic is a continuous curve $X(t)$ $(0 \le t \le 1)$ such that for each t we can find a closed interval $[t_1, t_2]$ containing t in its interior

[18] If ρ_0 is the distance in the space, then, obviously, $\rho_0 \le \rho$, and, therefore, if $\rho(X_nX) \le \varepsilon$, then all the more, $\rho_0(X_nX) \le \varepsilon$. Therefore, an extrinsic neighborhood is always intrinsic. If $\rho_0(X_nX) \to 0$ implies $\rho(X_nX) \to 0$, then conversely, any intrinsic neighborhood is extrinsic.

[19] All metric spaces that are actually studied in geometry (Riemannian and Finsler spaces) are n-dimensional manifolds with intrinsic metric. Metric spaces of functional analysis also have intrinsic metrics. Finally, the following theorem holds: each metric space is isometric to a subspace of a space with intrinsic metric. The proof of this theorem is trivial: with each pair of points X, Y of a given space R, we associate a segment of length $\rho(XY)$, and the endpoints of this segment are identified with the points X and Y; we so obtain a space with intrinsic metric, which includes R. The following is an important property of spaces with intrinsic metric: if two spaces R_1 and R_2 are mapped homeomorphically onto one another so that this mapping is isometric in arbitrarily small corresponding neighborhoods, then this mapping is an isometry in the large. The intrinsic metric is determined in the large from its values in arbitrarily small domains.

© 2006 by Taylor & Francis Group, LLC

(or on the boundary if $t = 0$ or 1) such that the curve $X(t)$ on the closed interval $[t_1, t_2]$ is a shortest join of the points $X(t_1)$ and $X(t_2)$. By this definition and by the Borel lemma, we can cover a geodesic by a finite number of shortest arcs, dividing it into a finite number of shortest arcs.

For example, an arc of a great circle on a sphere that is greater than the half-disk is a geodesic but not a shortest arc. Shortest arcs and geodesics play the role of lines in intrinsic geometry.

A *triangle* is a figure that is homeomorphic to a disk and is bounded by three shortest arcs. These shortest arcs themselves are called the sides, and the meeting points of the sides are called the vertices of this triangle. In contrast to an ordinary triangle, a *geodesic triangle* is a figure bounded by three geodesics.

A *polygon* is a figure enjoying the following two properties: (1) the boundary of this figure consists of at most finitely many shortest arcs (in particular, this figure can have no boundary at all) and (2) each sequence of points of this figure has a cluster point belonging to this figure (i.e., the figure is compact in common parlance). A sphere, a half-sphere, a band on a cylinder bounded by two base lines can serve as examples of polygons.

Also, we can define a *disk* and a *circle* as follows: a disk is the locus of points lying at a distance less than or equal to a given number from a given point and a circle is the locus of points equidistant from a given point.

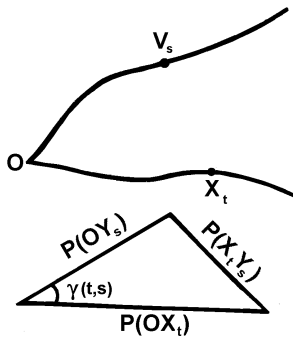


Fig. 6

We now define the concept of angle between two curves emanating from a common point of a manifold with intrinsic metric. Let $X_t = X(t)$ $(0 \le t \le 1)$ and $Y_s = Y(s)$ $(0 \le s \le 1)$ be two continuous curves emanating from a point O; i.e., $X_0 = Y_0$. We assume that O is not a multiple point; i.e., the points X_t and X_s are different from O for t and s different from zero. Draw a triangle with sides equal to $\rho(OX_t)$, $\rho(OY_s)$, and $\rho(X_tY_s)$ on the plane. The angle opposite the side equal to $\rho(X_tY_s)$ is denoted by $\gamma(t, s)$ (see Fig. 6). Since the sum of each two of the numbers $\rho(OX_t)$, $\rho(OY_s)$, and $\rho(X_tY_s)$ is at least the third number by the triangle inequality, such a triangle always exists. Of course, this triangle can degenerate into a segment if, e.g., $\rho(X_tY_s) = \rho(OX_t) + \rho(OY_s)$; this is not to be excluded from consideration. Also, we assume that t and s differ from O, and hence $\rho(OX_t)$ and $\rho(OY_s)$ are positive.

The angle between two given curves is the limit of the angle $\gamma(t, s)$ as t and s tend to zero, i.e., X_t and Y_s tend to O. Of course, this limit may fail to exist in which case we assume that the given curves make no angle at the point O.

This notion is applied in particular to shortest arcs emanating from a common point O. Let L and M be such arcs, and let X and Y be variable points on these arcs which are distinct from O. There is no necessity to introduce the parameters t and s here; we can take the arc lengths OX and OY and since the arc length of each shortest arc is equal to the distance, $\rho(OX)$ and $\rho(OY)$ themselves are substituted for t and s which we will always denote by x and y. Further, according

© 2006 by Taylor & Francis Group, LLC

to the general definition, we construct the plane triangle with sides $x = \rho(OX)$, $y = \rho(OY)$, and $z = \rho(OX)$ and take the angle $\gamma(x, y)$ opposite the side equal to z. The angle between the shortest arcs L and M is the limit α of these angles $\gamma(x, y)$ as x and y tend to zero.

We shall prove (in Sec. 4 of Chapter III) that *the so defined angle between two shortest arcs on a convex surface always exists.*

Of course, the angle can formally be defined without using triangles on the plane at all. Indeed,

$$z^2 = x^2 + y^2 - 2xy\,\cos\gamma(xy)$$

in these triangles. Therefore, we can merely define the angle α between two shortest arcs L and M by the formula

$$\cos\alpha = \lim_{t,s\to 0} \frac{x^2 + y^2 - z^2}{2xy}.$$

In the case of two arbitrary curves, we put

$$\cos\alpha = \lim_{x,y\to 0} \frac{\rho(OX_t)^2 + \rho(OY_s)^2 - \rho(X_tY_s)^2}{2\rho(OX_t)\rho(OY_s)};$$

moreover, we take the value of α in the closed interval $[0, \pi]$. But the definition using plane triangles shows us the true essence of these formulas.

The general properties of the angle will be studied in detail in Chapter IV. In many respects, these properties are analogous to the properties of the angle between two half-lines on the plane. The above definition of angle coincides with the definition of differential geometry for smooth surfaces; i.e., the angle between two curves emanating from a point O on a smooth surface turns out to be the angle between the tangents to these curves at the point O. In order to reveal an analogous spatial meaning of the angle between two curves on any convex surface, we introduce the concept of tangent cone which is a natural generalization of the concept of tangent plane.

Let F be some closed surface, and let O be an arbitrary point on this surface. Consider a homothetic dilation of this surface with the center of homothety at the point O. If the coefficient of homothety tends to infinity, then we can show that the surfaces resulting from F converge to some convex cone with vertex at the point O (see Sec. 5 of Chapter IV). *This cone is called the tangent cone of F at O.* If our surface has tangent plane at the point O, then this plane is the tangent cone at this point. The tangent cone reduces to a dihedral angle at the points lying on the edges of the surface; only at the peaks, or the so-called *conical points*, can this cone be an arbitrary convex cone. Also, we can show that the tangent cone can be defined as the locus of points lying on all half-lines that are the limits of the half-lines traveling from the point O to some points X of the surface F as these points X converge to O.

Let a curve L be drawn from the point O of the surface F. The limit of the half-line traveling from O to the variable points X of the curve O as X tends to O is called the *half-tangent* to L at O. The half-tangent is a generator of the tangent cone whenever this half-tangent exists.

© 2006 by Taylor & Francis Group, LLC

We shall prove (in Sec. 1 of Chapter IV) the following theorem: *Two curves emanating from a common point O on a convex surface F form the angle at this point in the sense of the intrinsic definition given above if and only if these curves have half-tangents at the point O. In this case, the angle between these curves is equal to the angle between their half-tangents which is measured on the tangent cone.*

The latter assertion should be understood as follows: the half-tangents, presenting generators of the tangent cone, divide this cone into two sectors; the angles of these sectors are defined on unfolding them onto the plane; the smallest of these angles is the angle between the two half-tangents measured on the tangent cone. If the tangent cone is a plane at the point O, then this angle is the usual angle between two half-tangents.

In particular, the following theorem, which was proved by I. M. Liberman, holds for shortest arcs.

A shortest arc on a convex surface has the left and right tangents at each point (and hence the left and right half-tangents).[20]

Therefore, the angle between two shortest arcs emanating from a common point is equal to the angle between their half-tangents at this point which is measured on the tangent cone.

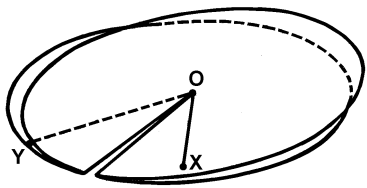


Fig. 7

The intrinsic definition of angle for curves on convex surface agrees perfectly with the natural concept of angle; but in other cases this definition can have one peculiarity which is worth mentioning here. Take, e.g., a cone with complete angle at the vertex equal to 4π; such a cone can be obtained if we cut two planes along half-lines and glue them together along the sides of these cuts (Fig. 7).[21] This cone can be divided into four sectors with angle equal to π, i.e., into four half-planes. If two points X and Y lie in the opposite sectors, a shortest arc connecting them should pass through the vertex O of this cone, and the segments OX and OY of this arc divide a neighborhood of the vertex into two sectors, each of which has angle greater than π. It is natural to assume that the angle between the segments OX and OY is equal to the smallest of the angles of these sectors. Meanwhile, according to our definition of angle, this angle can never be greater than π, since this angle is the limit of the angles of some plane triangles, while the angles of a plane triangle are never greater than π. For the segments OX and OY, the angle in our definition is obviously equal to π. Hence we encounter here some disagreement between this and the natural concepts of angle. Nevertheless, our definition turns out to be very convenient. The above mentioned disagreement is eliminated by introducing a new concept of the *angle of a sector*.

[20] I. Liberman, Geodesic lines on convex surfaces, Dokl. Akad. Nauk SSSR, Vol. XXXII, No. 5 (1941), pp. 310–313. This theorem will be proved in Sec. 6 of Chapter IV. The right (left) tangent to a curve $X(t)$ at a point $X(t_0)$ is the limit of the secants passing through the points $X(t_0)$ and $X(t)$ of the curve as t tends to t_0 from the right (left), i.e., always $t > t_0$ ($t < t_0$).

[21] In Fig. 7, one plane lies entirely over the other. The edges of the cuts along which the gluing was carried out are slightly moved apart.

© 2006 by Taylor & Francis Group, LLC

Let L and M be two shortest arcs emanating from a point O in a manifold with intrinsic metric, which have no other common points near O. (As we show in Sec. 4 of Chapter III, every two distinct shortest arcs on a convex surface enjoy this property.) Then the shortest arcs L and M divide a sufficiently small neighborhood of the point O into two "sectors." Let U be one of these sectors; we shall draw shortest arcs $N_1, N_2, \ldots, N_n$ in this sector and number them in the order of their location between L and M. Denote by $\alpha_0, \alpha_1, \ldots, \alpha_n$ the angles between L and N_1, N_1 and $N_2, \ldots, N_n$ and M. The greatest lower bound of the sums $\alpha_0+\alpha_1+\ldots+\alpha_n$ over all possible choices of the shortest arcs N_i and the number of these arcs is called the *angle of the sector* U. In Sec. 3 of Chapter IV, we shall justify this definition and study the properties of the angle of a sector in more detail; here we restrict discussion to some general remarks.

It is easy to verify that, on any cone and in particular on a cone with complete angle 2π, the angles of sectors between generators which are defined from unfolding onto the plane, coincide with the angles obtained from our definition. Thus, the introduction of the concept of angle of a sector removes the above disagreement.

The following assertion holds for convex surfaces.

Assume that shortest arcs $L_1, L_2, \ldots, L_n$ with half-tangents $T_1, T_2, \ldots, T_n$ emanate from a point O on a convex surface. Then the angles of the sectors into which these shortest arcs divide a neighborhood of the point O are equal to the angles of the sectors into which their half-tangents divide the tangent cone at the point O. (The proof is given in Sec. 6 of Chapter IV.)

In particular, this implies that *the angles of sectors on a convex surface arc adjoined exactly in the same manner as the angles of sectors on the cone, and the sum of the angles of sectors which form a full neighborhood of the point O does not depend on these sectors and is equal to the complete angle of the tangent cone at this point.* Therefore, this sum is naturally called the *complete angle at the point* O. The complete angle at every point does not exceed 2π on a convex surface; it is always equal to 2π on a smooth surface, but it can be greater than 2π on a nonsmooth surface, e.g., as the angle at the vertex of the cone considered above. If the complete angle at a point O is no greater than 2π, then the angle between two shortest arcs emanating from the point O coincides with the minimum of the angles of those sectors to which these shortest arcs partition a neighborhood of the point O. If the complete angle around the point O is greater than 2π, there can exist shortest arcs emanating from the point O which divide its neighborhood into sectors with angles greater than π, and none of the angles of these sectors is equal to the angle between the shortest arcs.

The concept of angle of a sector is similar to that in plane geometry. In exact analogy with plane geometry, the angle of a polygon P at the vertex A is the angle of the sector bounded by the sides meeting at A in which lies P. For example, if we take three points A, B, and C on the sphere and connect them by shortest arcs, then we divide this sphere into two triangles ABC; if in one of these triangles the angle at the vertex A is equal to the angle between the shortest arcs AB and AC, then this angle complements to 2π the angle at A in the other triangle.

Perhaps, the only essential distinction between the angles between shortest arcs and those between half-lines in the plane is that it is not always possible to draw an angle equal to a given angle so that one side of the former is a given shortest arc. For example, take a cone of revolution together with its base. Let O be the

© 2006 by Taylor & Francis Group, LLC

vertex of this cone, and let A be a point of the base circle. There are no shortest arcs on this cone tangent to the base circle; the fact that there are no such arcs on the base is obvious, and we confirm that they are absent on the lateral surface by unfolding the latter on the plane. Therefore, there is no shortest arc making the right angle with OA at point A. In other words, in the direction of the base circle, there are no shortest arcs on a cone of revolution. Hence, in contrast to the case of regular surfaces, it is not possible in general to draw a shortest arc in any direction on a convex surface. However, we shall show below in Sec. 5 of Chapter V that the set of such specific directions at each point has angle measure zero.

In view of the fact that there are no shortest arcs tangent to the base circle on a cone of revolution, it is not convenient to introduce the intrinsic concept of tangent to a curve as the limit of shortest secants. Instead, we introduce the concept of direction of a curve at a given point. Namely, we shall say that *a curve emanating from a point O has a direction at this point if this curve makes some angle with itself.* At first glance, this definition seems absurd, but it does make sense. Indeed, let $X_t = X(t)$ be a curve emanating from a point $O = X(t_0)$ in a manifold with intrinsic metric ρ. We shall consider this curve as two coinciding curves $X_t = X(t)$ and $X'_t = X'(T)$; the angle between these curves is defined by the formula

$$\alpha = \lim_{t,t'\to 0} \gamma(t, t').$$

It is absolutely obvious that for some curves $X(t)$ this limit fails to exist for *all* ways in which t and t' vanish. Considering this question for a curve on the plane, we easily see that the limit $\lim_{t,t'\to 0} \gamma(t, t')$ exists if the curve $X(t)$ has tangent at the point $X(0)$. This is also true on every convex surface; namely, we prove in Sec. 1 of Chapter IX that *a curve on a convex surface has tangent* (*the limit of secant*) *at its initial point if and only if this curve has direction at this point in the sense of our definition.*

In Chapter IX, we shall introduce the concept of a "swerve of the direction of a curve" or simply a "swerve of curve," which is a generalization of the concept of the integral of the curvature of a curve with respect to the arc length. Thus, we shall have at our disposal the basic concepts of the intrinsic-geometric theory of curves and detail a number of their applications.

To complete the definition of all basic concepts of intrinsic geometry which are analogous to the familiar concepts of plane geometry, it remains only to introduce the area of a polygon in intrinsic way, and so to introduce the area of each domain on a surface. Let P be a polygon on a surface, i.e., a figure bounded by finitely many shortest arcs. Partition this polygon into triangles; to each of these triangles we put in correspondence the plane triangle with sides of the same length; then we sum the areas of all these plane triangles. *The limit of these sums*, which is obtained by indefinitely refining partitions of the polygon P, if this limit exists, *is taken as the area of the polyhedron P.* In Sec. 6 of Chapter II, we shall prove that polygons on a convex surface can always be partitioned into arbitrarily small triangles; in Chapter X, we shall prove that for each polygon P on a convex surface, the so-defined area exists and coincides with its area in the sense of the usual definition as the limit of the areas of polyhedral surfaces converging to P. When the area of polygons is available, the extension of the definition of area to other domains is carried out by the well-known methods of measure theory.

© 2006 by Taylor & Francis Group, LLC

Thus, we see that all concepts analogous to the basic concepts of plane geometry can be introduced on every convex surface. We may expect that the intrinsic geometry of arbitrary convex surfaces can be developed as rich in content as plane geometry. But there is one more basic concept for curved surfaces, namely, the concept of their intrinsic curvature; this concept will be considered separately in the next section.

8. Curvature

The following Gaussian theorem serves as a starting point for the abstraction of the concept of curvature to all convex surfaces: the area of the spherical image of a triangle is equal to its "excess," i.e., to the excess of the sum of angles of this triangle over π. In this theorem, we equate two values; the first is the area of the spherical image which characterizes the curving of a surface in the space, and the second, the "excess of a triangle," characterizes the deviation of the intrinsic metric of a surface from the metric of the plane, i.e., in a way, it is a measure of non-Euclidicity of the intrinsic metric of this surface. In this connection, we have two curvatures, "extrinsic" and "intrinsic."

Let G be a certain domain on a convex surface F. We shall draw all supporting planes to the surface F at all points of the domain G and also we will draw all radii parallel to the outer normals of these planes from the center of the unit sphere S. The set of points on the sphere S consisting of the endpoints of these radii is called the *spherical image* of G. This definition differs from the usual only by the fact that instead of tangent planes that do not exist in general at some points of the domain G, we take all possible supporting planes. The area of this spherical image will represent the "extrinsic curvature" of the domain.

If X is a conical point of the surface F, then the spherical image of this single point occupies a whole domain on the sphere S. If L is some curved edge of the surface, then its spherical image also covers a whole domain on the sphere S; e.g., the spherical image of the base circle of a cylinder of revolution covers a whole half-sphere.

Similar to the area of the spherical image, the intrinsic curvature is defined as a set function on a surface; i.e., to each set M in some class of sets (to each Borel set in the general case[22]), we put in correspondence the number $\omega(M)$, the curvature of the set M. In accordance with the terminology accepted in differential geometry, we should speak here about the total (or integral) intrinsic curvature, but for brevity we omit both adjectives; this leads to no confusion, since we will never mean anything else by the term "curvature." In the definition of curvature, we can start from the excess of a triangle $\alpha + \beta + \gamma - \pi$, where α, β, and γ are the angles of the triangle. We shall prove that this excess is always nonnegative. (This is a direct consequence of Theorem 3 in Sec. 4 of Chapter III.)

[22]The class of Borel sets is the class of sets which is generated from closed (or open) sets by infinitely many operations of countable union and intersection. More precisely, this means the following. A system of sets is called a ring if this system contains the union $\sum_{n=1}^{\infty} E_n$ and the intersection $\prod_{n=1}^{\infty} E_n$ whenever this system contains a sequence of sets $E_1, E_2, \ldots, E_n$. A set is called Borel if this set belongs to each ring containing all closed sets. Open sets are representable as union of closed sets and so are Borel sets. The union and intersection of countably many Borel sets are Borel sets. Finally, if A and B are two Borel sets, then $A - B$ is also a Borel set.

© 2006 by Taylor & Francis Group, LLC

Consider all possible finite or countable covers of a given set M on a surface by the interiors of triangles. *The curvature of the set M on this convex surface can be defined as the greatest lower bound of the sums of excesses of the triangles whose interiors cover M.*

This definition reminds us of the definition of the outer measure of a set according to Lebesgue; curvature enjoys the properties that are analogous to the properties of a measure. It turns out that each set on a convex surface has finite curvature. Since the excess of a triangle on a surface is no less than zero, *the curvature of every set on a convex surface is nonnegative.* For Borel sets, the curvature turns out to be completely additive or, as it is called sometimes, totally additive; i.e., if $M = \sum_{i=1}^{\infty} M_i$ and the sets M_i are pairwise disjoint, then $\omega(M) = \sum_{i=1}^{\infty} \omega(M_i)$. The curvature of the interior of triangle proves equal to the excess of this triangle, and, in general, the curvature of the interior of an n-gon homeomorphic to a disk is equal to the excess of the sum of its angles as compared with the sum of a plane n-gon; as is known, the latter is equal to $(n-2)\pi$. The curvature of a complete closed surface is always equal to π. The curvature of a singleton is equal to 2π minus the complete angle at this point, and hence curvature is not equal to zero for conical points. This is an essential distinction of the curvature from Lebesgue measure. We can also see from this that, speaking about the curvature, it is necessary to distinguish the curvature of a polygon itself from the curvature of its interior since this polygon can have conical points as vertices. For example, the curvature of the interior of a face of a cube is equal to zero, and the curvature of the face together with its vertices is 2π. As for the sides of a polygon, they can add nothing to curvature, since the curvature of each shortest arc without endpoint on a convex surface is equal to zero. The deletion of endpoints is essential for the same reason that the endpoints of a shortest arc can be conical points; however, it is easy to show that a shortest arc cannot pass through a conical point.

Although the above definition of curvature is the best possible from a general point of view, recalling our aims, we can replace this definition by an equivalent but more elementary definition, which will be given in Chapter V devoted especially to the theory of curvature. Also, we reveal in this chapter the role of curvature as the "measure of non-Euclidicity" of the intrinsic metric of a surface. We cite here only the following among all theorems belonging to this theory: *if the curvature of a domain G on a convex surface is equal to zero, then each point of G has a neighborhood isometric to a part of the plane.* The example of the lateral boundary of a cylinder shows that the whole domain G may fail to be isometric to part of the plane.

The connection between the curvature and area of the spherical image of a set on a convex surface is stated in the following generalized Gaussian theorem: *the curvature of each set on a convex surface is equal to the area of its spherical image* (provided that this spherical image has area, i.e., is Lebesgue measurable). In particular, the area of the spherical image of the interior of a triangle is equal to the excess of this triangle, and the area of the spherical image of a singleton is equal to 2π minus the complete angle of this point. The proof of this generalized Gaussian theorem in Chapter V is carried out by using the method of polyhedral approximation according to the scheme given at the end of Sec. 5.

The significance of the results formulated here is twofold. First, we can see that sets on convex surfaces are of nonnegative curvature. As is known, this property

© 2006 by Taylor & Francis Group, LLC

distinguishes regular convex surfaces from all regular surfaces in general; we shall prove that this property is also characteristic of all convex surfaces. Roughly speaking, we shall prove that each point of a manifold with intrinsic metric in which the curvature of every set is nonnegative has a neighborhood isometric to a convex surface. Here, we are rushing ahead; a precise formulation of this result will be given in the next section.

Second, the generalized Gaussian theorem relates the intrinsic metric to the spatial shape of a convex surface; thus, it opens a way for a more detailed study of this relation. Some relevant results will be considered in detail in Chapter XI.

Our concept of curvature is a generalization of the well-known differential-geometric concept of the total or integral curvature, which is often defined as the integral of the Gaussian curvature with respect to the surface area. Hence the starting point here is the concept of curvature as a point function rather than a set function. In contrast, we take precisely the integral notion of curvature as a starting point, and the curvature at a point is then defined as the specific curvature, i.e., as the limit of the ratio of the curvature of a domain and its area under the condition that this domain shrinks to a given point. This limit does not necessarily exist at each point of a convex surface. Moreover, the curvature of a domain on a convex surface may fail to be equal to the integral of the specific curvature with respect to the area. Polyhedra give the simplest example of this: the specific curvature exists everywhere except for vertices and is obviously equal to zero. Therefore, this specific curvature is integrable, but the integral of this curvature over the whole surface of the polyhedron vanishes whereas the curvature of a convex polyhedron having at least one vertex is different from zero.

The curvature of a set can be defined in an especially simple way if this set lies on a polyhedron or on a manifold with a polyhedral metric since the curvature on a polyhedron is a discrete set function supported only in the set of its vertices. Let a set M on a polyhedron contain the vertices A_1, A_2, ..., A_n with complete angles at these vertices equal to $\theta_1, \theta_2, \ldots, \theta$; the curvature of this set can be defined as

$$\omega(M) = \sum_{i=1}^{n} (2\pi - \theta_i),$$

i.e., as the sum of the curvatures of the vertices belonging to M. With this definition, it is obvious that each set M on a polyhedron is of negative curvature if and only if the complete angle at each vertex does not exceed 2π. The term "polyhedral metric of positive curvature" introduced above originates from this observation.

In order to connect this definition of the curvature of a polyhedron with the general definition mentioned above, we shall prove the following theorem.

Let a polygon P with angles $\alpha_1, \alpha_2, \ldots, \alpha_n$ which is homeomorphic to a disk be given on a polyhedron, and let $\theta_1, \theta_2, \ldots, \theta_m$ be the complete angles at the vertices of P lying in the interior of P. Then

$$2\pi - \sum_{i=1}^{n} (\pi - \alpha_i) = \sum_{i=1}^{n} \alpha_i - (n-2)\pi = \sum_{j=1}^{m} (2\pi - \theta_j). \tag{1}$$

Briefly, this means that *the excess of a polygon is equal to the sum of curvatures of the vertices contained inside this polygon.*

© 2006 by Taylor & Francis Group, LLC

To prove this theorem, we divide the polygon P into triangles T_P, each of which lies on the same face. In this process, it can happen that some vertices of these triangles lying on the boundary of the polygon P do not coincide with vertices of this polygon. However, we can consider these vertices as vertices of the polygon P; this does not change the sum on the right-hand side of (1) since the angle at such a vertex is equal to π; and, hence, the summand $\pi - \alpha_i$ corresponding to this vertex is equal to zero. In exactly the same way, certain vertices of the triangles T_P lying in the interior of the polygon P may fail to coincide with vertices of the polyhedron. However, adding these vertices to the vertices of the polyhedron, we do not change the sum on the right-hand side of (1), since the complete angle at such a vertex is equal to 2π, and hence the summand $2\pi - \theta_j$ corresponding to this vertex vanishes. Therefore, we can assume that the vertices of the triangles T_P lying on the boundary of the polygon P are vertices of the polygon itself, while the vertices of these triangles lying in the interior of P can be assumed to be vertices of the polyhedron. Since each triangle T_P lies on one face, the sum of the angles of T_P is equal to π, and if the number of all triangles is equal to f, the sum of their angles is equal to $f\pi$. On the other hand, the sum of these angles is equal to the sum of the angles at all vertices of the triangles, i.e., to the sum of the angles of the polygon P plus the sum of the complete angles at interior vertices. Consequently,

$$f\pi = \sum_{i=1}^{n} \alpha_i + \sum_{j=1}^{m} \theta_j. \tag{2}$$

If k is the number of sides of the triangles T_P and e is the number of vertices, then $f - k + e = 1$ by the Euler theorem. Consequently, the number of all vertices is equal to the sum of the numbers of vertices on the boundary and in the interior, i.e., $e = m + n$; therefore,

$$f - k + m + n = 1. \tag{3}$$

Finally, each triangle has three sides, but there are only n sides lying on the boundary, each of which belongs to one triangle, and each of the sides belongs to two triangles. Therefore,

$$3f = 2k - n. \tag{4}$$

Multiplying equation (3) by equation (2) and putting $2k = 3f + n$ in (3), we obtain

$$f = 2m + n - 2. \tag{5}$$

Substituting this expression for f into formula (2), we obtain

$$2\pi m + \pi n - 2\pi = \sum_{i=1}^{n} \alpha_i + \sum_{j=1}^{m} \theta_j$$

or

$$\sum_{i=1}^{n} \alpha_i - (n-2)\pi = \sum_{j=1}^{m} (2\pi - \theta_j);$$

as required.

Applying this result to convex polyhedra, we easily see that for them the definition of the curvature of a set M as the sum of curvatures of vertices lying in M is equivalent to the definition of curvature as the greatest lower bound of the sums of the excesses of the triangles whose interiors cover M.

© 2006 by Taylor & Francis Group, LLC

9. Characteristic Properties of the Intrinsic Metric of a Convex Surface

When speaking about the conditions characterizing the intrinsic metric of a closed convex polyhedra in Secs. 3 and 4, we enlist doubly-covered convex polygons in the set of polyhedra; this is necessary in order to avoid additional stipulations in the theorem on the existence of a polyhedron with a given metric. For exactly the same reason, it is appropriate to enlist doubly-covered bounded domains in the set of complete convex surfaces. Plane convex domains not covering the whole plane can be of the following three types: (1) bounded domains homeomorphic to a disk; (2) infinite domains homeomorphic to a half-plane; (3) bands between pairs of parallel lines. The definition of what should be meant by a doubly-covered convex bounded domain was given in Sec. 3 by mapping the sphere onto this domain under which the interior of the domain turns out to be covered by the images of both half-spheres, and the boundary turns out to be covered by the image of the equator that separates these half-spheres. By exact analogy, we should consider a doubly-covered domain homeomorphic to the plane as a continuous image of the plane whose two half-planes cover the interior of this domain. Finally, a doubly-covered band between two parallel lines is nothing more than a flattened cylinder. Such a doubly-covered domain can be folded into an ordinary cylinder. Thus, convex domains considered as doubly-covered are transformed to complete convex surfaces, which are, respectively, closed, homeomorphic to the plane, and cylindrical.

If two points X and Y lie on one side of a domain, then a shortest arc connecting these points is the line segment XY. If these points lie on opposite sides of this domain, then a shortest arc connecting them consists of two line segments meeting at the boundary of the domain.

Having enlisted the above limit cases into the set of complete convex surfaces, we naturally have to enlist all domains on such degenerate convex surfaces into the set of all convex surfaces. Moreover, we must bear in mind that the interior points of a doubly-covered domain are considered as two points. For example, we can form a nonclosed surface from a surface that is twice covered by a disk on cutting out a domain only on one side of the disk and keeping the other intact. All properties of convex surfaces discussed in the preceding sections also hold for these degenerate surfaces. Without further stipulation, we shall assume in the sequel that these degenerate surfaces are enlisted into the set of all convex surfaces.

We can solve the problem of the properties characterizing the intrinsic metric of a convex surface, i.e., the problem of those conditions under which a metric space is isometric to a convex surface, not for all convex surfaces in general but rather for surfaces that are "convex in themselves," i.e., for those surfaces that are convex domains on the boundaries of convex bodies. These are surfaces whose every two points can be connected by a shortest arc. The reasons behind this restriction were revealed in Sec. 1: every two points can be connected by a shortest arc not on every convex surface; but, however, each point on a convex surface has a small convex neighborhood. The lateral surface of a right cylinder or a cone of revolution, a half-sphere, etc. are examples of incomplete convex surfaces whose every two points can be connected by a shortest arc.

© 2006 by Taylor & Francis Group, LLC

As each question about necessary and sufficient conditions, the problem here is twofold. First, we are interested in necessary conditions, i.e., the conditions that the intrinsic metric of convex surfaces satisfies in reality. In this case, it is desirable to find possibly stronger conditions easily yielding various corollaries of interest. Second, we also want to distinguish among these necessary conditions those that are also sufficient; and, in this case, the opposite is desirable, namely we want to find possibly weaker sufficient conditions. The confrontation of these two aims forces us to divide these two aspects of the problem and to consider them separately.

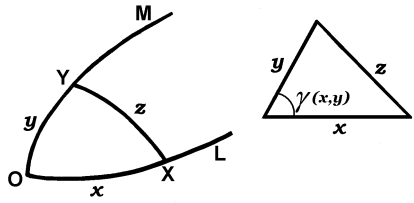


Fig. 8

We have proved the following two general properties of convex surfaces in Sec. 6: each convex surface has intrinsic metric and is homeomorphic to a domain on the sphere. Below in Sec. 3 of Chapter III, we shall show that *the intrinsic metric of a convex surface satisfies one more condition which is called the convexity condition.* Let L and M be two shortest arcs emanating from a common point O, and let X and Y be variable points on these arcs which are different from O.

Let x and y denote the distance from O to X and Y, and let $z(x, y)$ be the distance between X and Y. In accordance with the accepted notation, $\gamma(x, y)$ is the angle in the plane triangle with sides equal to x, y, and z, which is opposite to the side equal to z (Fig. 8). *The convexity condition consists of the fact that* $\gamma(x, y)$ *is a nonincreasing function of* x *and* y[23] *on each pair of closed intervals* $0 < x \le x_0$, $0 < y \le y_0$ *such that the points* X *and* Y *corresponding to arbitrary two values* x *and* y *from these closed intervals can be connected by a shortest arc.* (The shortest arc XY exists on a surface that is "convex in itself" for every two points X and Y, and, therefore, the indicated interval extends to the whole lengths of the shortest arcs L and M.)

Take two pairs of points X_1, Y_1 and X_2, Y_2 such that $x_1 < x_2$ and $y_1 < y_2$ (Fig. 9). Consider two plane triangles $O_i'X_i'Y_i'$ $(i = 1, 2)$ with sides $O_i'X_i = x_i$, $O_iY_i' = y_i$, and $X_i'Y_i' = z_i = z(x_i, y_i)$. The angles at their vertices O_1' and O_2' are $\gamma_1 = \gamma(x_1, y_1)$ and $\gamma_2 = \gamma(x_2, y_2)$; by the convexity condition, $\gamma(x_1, y_1) \ge \gamma(x_2, y_2)$. If we take two points X_1'' and Y_1'' on the sides $O_2'X_2'$ and $O_2'Y_2'$ of the triangle $O_2'X_2'Y_2'$ such that $O_2'X_1'' = O_1''X_1' = x_1$ and $O_2'Y_1'' = O_1'Y_1' = y_1$, the sides of the triangle $O_2''X_1''Y_1''$, which subtend the angle $\gamma_2 = \gamma(x_2, y_2)$, are equal to the sides of the triangle $O_1'X_1'Y_1'$, which subtend the angle $\gamma_1 = \gamma(x_1, y_1)$, and since $\gamma_1 \ge \gamma_2$, we have $X_1'Y_1' \ge X_1''Y_1''$. By construction, the length of the segment $X_1'Y_1'$ is equal to the length of the shortest arc X_1Y_1. Therefore, the obtained result means that the distance between the points of the lateral sides on the triangle OX_2Y_2 is no less than the distance between the corresponding points on the plane triangle $O_2'X_2'Y_2'$.

It seems natural that this should be so in the case for a convex surface: on this surface, the sides of a triangle with sides of given lengths are moved apart a greater distance than on the plane. The more convex the surface, the greater is

[23] This means that $\gamma(x_1, y_1) \ge \gamma(x_2, y_2)$ whenever $x_1 \le x_2$ and $y_1 \le y_2$.

© 2006 by Taylor & Francis Group, LLC

this movement. An equilateral triangle on the sphere can serve as an illustrative example. This remark reveals the reason why the above-mentioned condition is called the convexity condition. The proof of the fact that this condition holds for any convex surface will be performed on using the method of polyhedra approximation according to the general scheme of Sec. 5.

The convexity condition turns out to be rather strong and easily implies many important theorems of the intrinsic geometry of convex surfaces. For example, since the angle $\gamma(x, y)$ does not decrease when x and y decrease, by the convexity condition, the limit $\lim_{x,y\to 0}\gamma(x, y)$ exists; by definition, this limit is the angle between the shortest arcs L and M. Thus, the existence of the angle made by two shortest arcs is a direct consequence of the convexity condition. In such a simple way, we obtain the following theorem: each angle of a triangle on a convex surface, as well as angle between its sides, is at least the corresponding angle of the plane triangle with the sides of the same length.[24] This theorem serves not only as the basis for studying triangles on convex surfaces but also has many other applications. In particular, this theorem implies that the sum of the angles of a triangle on a convex surface is always at least π; this is the result, which was mentioned in the preceding section in connection with curvature theory.

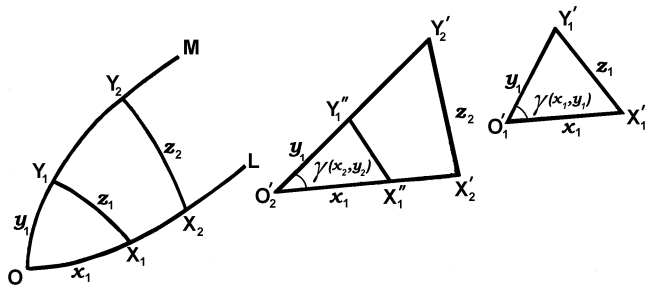


Fig. 9

Here we mention one more condition among other consequences of the convexity condition. Let L and M be two shortest arcs emanating from a point O, and let X and Y be variable points of these shortest arcs. Let $\gamma(x, y)$ be as above. *If at least one of the points X and Y tends to O and the other varies in an arbitrary way but so that the shortest arc XY converges to a segment of one of the shortest arcs L or M, then the angle $\gamma(x, y)$ tends to some limit.* Of course, this limit is exactly the same as in the case where both points X and Y tend to O, and hence it is equal to the angle between the shortest arcs L and M. This result can also be expressed as follows: if the product xy tends to zero in such a way that the shortest arcs XY tend to a part of the line $L + M$ composed of these shortest arcs, then there exists the limit $\lim_{x,y\to 0}\gamma(x, y)$. Giving the name "angle in the strong

[24] *Proof.* Let ABC be a triangle on a convex surface. We restrict consideration to the case where every two points on the sides of this triangle can be connected by a shortest arc. Let $\gamma(x, y)$ be the angle of the plane triangle defined according to the convexity condition for the points X and Y on the sides AB and AC. By the definition of the angle α between the sides AB and BC, we have $\alpha = \lim_{x,y\to 0}\gamma(x, y)$. The corresponding angle α_0 of the plane triangle with sides of the same length is the value of $\gamma(x, y)$ corresponding to $x = AB$ and $y = AC$. Since $\gamma(x, y)$ does not increase when x and y increase, we obviously have $\alpha \geq \alpha_0$; as required. Moreover, since the angle of the sector lying in the triangle is no less than the angle between the sides, then the angle of this sector, i.e., the angle at the vertex of the triangle, is at least α_0.

© 2006 by Taylor & Francis Group, LLC

sense" to this limit, we can say that there always exists some "angle in the strong sense" between two shortest arcs on a convex surface. This result is well known for regular surfaces and is based on the fact that in a geodesic neighborhood the metric of a surface coincides with the metric of the line on the plane with accuracy up to within infinitesimals of the order higher than the width of this neighborhood. A sufficiently small neighborhood of a shortest arc on a convex polyhedron can be developed onto the plane,[25] and, so, the existence of the angle in the strong sense between shortest arcs on a convex polyhedron is obvious.

The convexity condition turns out to be characteristic of the intrinsic metric of a convex surface. Namely, we can prove the following theorem.

A given metric space R is isometric to a convex surface of which every two points can be connected by a shortest arc if and only if the following conditions hold: (1) *R is homeomorphic to a domain on the sphere;* (2) *every two points of R can be connected by a curve whose length is equal to the distance between these points;* (3) *the convexity condition holds in R.*

Hence we can develop the intrinsic geometry of convex surface in a purely intrinsic way on the basis of only the convexity condition, together with the general properties of manifolds with intrinsic metric, i.e., without any arguments connected with the properties of a convex surface as a figure in the space. The convexity condition turns out to be very convenient for this purpose, since this condition contains a very strong requirement. At the same time, precisely because of this strength, this condition cannot be considered a successful sufficient condition. For example, the fact that this condition holds for every manifold with polyhedral metric of positive curvature is a theorem, which is elementary in essence, but not obvious. In exactly the same way, the fact that this condition holds for manifolds with a metric given by a line element of positive curvature is also a theorem that is not proved "in two words." (The fact that this theorem holds seems to remain unnoticed at all by now.) These remarks imply that we have to search for other weaker sufficient conditions.

We can test the following two conditions.

1. *There is some angle in the strong sense* (*as defined above*) *between every two shortest arcs emanating from a common point.*

2. *The sum of the angles between the sides of every triangle is at least π.*

As we have mentioned, these conditions follow from the convexity condition and hold for each convex surface. The fact that these conditions hold in particular for a manifold with polyhedral metric of positive curvature was proved here for the first condition and for the second in the preceding section. The fact that these conditions hold for a manifold with metric given by a line element of positive curvature is a fact known since the age of Gauss.

The condition of the existence of the angle in the strong sense can be weakened if we replace this condition by that of the existence of the angle in the sense of our definition of Sec. 7, i.e., the angle is $\lim_{x,y\to 0}\gamma(x,y)$. But the proof of this possibility,

[25] If this shortest arc emanates from a vertex, then in order to unfold this neighborhood onto the plane, it is necessary to make a cut along a segment emanating from this vertex.

© 2006 by Taylor & Francis Group, LLC

which is at our disposal, involves an additional although simple condition and is too difficult for presentation in the framework of this book.

However, the condition of the existence of an angle in any sense turns out to be superfluous in general! Of course, without the existence of an angle, the condition on the sum of the angles of the triangle becomes meaningless, but it possibly defines the angle between shortest arcs in so that this angle always exists and the condition on the sum of so defined angles of the triangle is automatically sufficient. This definition of angle can be introduced in the following way. Let L and M be two shortest arcs emanating from a common point O in a manifold with intrinsic metric ρ. Let X and Y be variable points different from O on these shortest arcs, and let $x = \rho(OX)$, $y = \rho(OY)$, and $z = z(x, y) = \rho(XY)$. According to the accepted condition, we denote by $\gamma(x, y)$ the angle opposite the side of length z in the plane triangle with sides equal to x, y, and z. *The lower angle (in the strong sense) between the shortest arcs L and M is the lower limit of the angle $\gamma(x, y)$ which is obtained when at least one of the points X and Y tends to O and the shortest arc XY becomes infinitely close to one of the shortest arcs L or M.* (In Sec. 2, Chapter II, we shall prove that the shortest arc XY exists in a manifold with intrinsic metric whenever the points X and Y are sufficiently close to the point O.) This definition differs from the definition of angle in the strong sense only by the fact that instead of the limit of $\gamma(x, y)$ we take the lower limit which always exists since $0 \leq \gamma(x, y) \leq \pi$. We can consider this lower angle (in the strong sense) instead of the angle in the strong sense; this angle reduces to the angle in the strong sense on a convex surface because of the existence of the latter. Using this definition we can formulate the following theorem.

Theorem A. *A metric space R is isometric to a convex surface whose every two points can be connected by a shortest arc, if and only if the following conditions hold:*

1. *R is homeomorphic to the sphere.*
2. *Every two points in R can be connected by a curve of length equal to the distance between these points. (Or, in a weaker form, for every two points X and Y, there exists a point Z such that $\rho(XZ) = \rho(YZ) = \rho(XY)/2$, where ρ is the metric on R.)*
3. *Each point of R has a neighborhood such that, for each triangle lying in this neighborhood, the sum of the lower angles between its sides is at least π.*

The first and third conditions hold for all convex surfaces. But the second condition holds only for those surfaces, whose every two points can be connected by a shortest arc; for arbitrary convex surface, we can say only that its metric is intrinsic; i.e., for every two points, there exists a curve connecting them such that its length is arbitrarily close to the distance between these points. However, this condition, together with the first and third conditions, is not sufficient in general for the existence of any convex surface isometric to the space R enjoying all three conditions. For example, take two equal spheres, cut them along their meridians, and glue them so as to obtain a doubled sphere. If we delete the poles of this doubled sphere, the endpoints of these cuts, we obtain a manifold that is homeomorphic to

© 2006 by Taylor & Francis Group, LLC

a domain on the sphere[26] and has metric of positive curvature; this manifold is not realizable *a priori* in the form of a convex surface since the total curvature of this manifold is 8π while the curvature of a convex surface cannot exceed 4π.

Since we have already mentioned that the metric of a convex surface satisfies conditions stronger than condition 3 of Theorem A (namely, the convexity condition and its consequences presented above), the essence of this theorem lies in the sufficiency of these conditions.

Let us compare condition 3 with the conditions defining a polyhedral metric and a metric given by a line element. We assume *a priori* in the definition of a manifold with polyhedral metric that this manifold is locally isometric to a surface in the space, namely, to a cone. As we revealed in Sec. 2, the essence of determining a metric from a line element is that a manifold with such a metric is assumed to be infinitesimally isometric to the plane. We have nothing similar to that in our condition precisely for that reason that we have excluded the hypothesis of the existence of the angle; as is easily seen, this hypothesis is very close to the hypothesis of infinitesimal isometry to the plane. However, since this condition, together with the other two, is sufficient for the existence of a surface with a given metric, this condition implies now the existence of the angle, as well as infinitesimal isometry to the plane everywhere except for countably many conical points. This fact makes this condition especially interesting, so we take it as the basis. Some other conditions that can replace this condition will be considered in general terms in Sec. 6 of Chapter VIII.[27]

In the framework of our book, Theorem A cannot be proved completely, and we are forced to restrict consideration to the most essential particular cases of this theorem which give a local characterization of the metric of a complete convex surface on the one hand and a characterization of the metric of an arbitrary convex surfaces on the other hand. Moreover, in order to simplify the proofs, we will slightly change the definition of lower angle given here; but now we abstain from this in order to avoid a conglomeration of definitions. For brevity, an intrinsic metric satisfying condition 3 of Theorem A will be called the *metric of positive curvature* in what follows.[28]

We shall formulate the theorems that characterize the metric of a convex surface as existence theorems for a surface with given metric; since, as Theorem A, the necessity of the conditions of these theorems is of no interest.

When speaking about a metric space homeomorphic to a domain on the sphere or to some other specific manifold, it is convenient for the sake of visuality to consider

[26] Namely, a domain homeomorphic to the sphere with excluded poles; this can be directly verified by mapping each of these cut spheres on a half-sphere.

[27] Unfortunately, our notion of lower angle is very far from the natural idea of angle. With our definition, the lower angle between shortest arcs depends on these arcs as a whole, while in the natural definition, the angle should depend only on arbitrarily small segments of these shortest arcs starting from their common origin. Therefore, the lower angle can change in general if we replace two shortest arcs L_1 and L_2 by their segments or, in contrast, if we extend these arcs. This can be avoided by introduction of another notion of lower angle; its definition will be given in Sec. 6 of Chapter VIII. We do not introduce this notion of lower angle immediately because this notion itself poses essential inconveniences. The question of various definitions of angle will be considered in Sec. 6 of Chapter VIII.

[28] Since the Euclidean metric satisfies, in particular, this definition, it is better to speak about a metric of nonnegative curvature.

© 2006 by Taylor & Francis Group, LLC

this space R as the last manifold with a continuous function $\rho(XY)$ of the pair of points of this manifold, the metric of the space R.[29] In this sense we can speak about a manifold with metric given on this manifold.

A characterization of the intrinsic metric of closed convex surfaces is given by the following theorem.

Theorem B. *A metric space homeomorphic to the sphere and with a metric of positive curvature is isometric to a closed convex surface.*

This is a particular case of Theorem A since in each space with intrinsic metric homeomorphic to the sphere (and in each compact space with intrinsic metric in general), every two points can be connected by a shortest arc. The proof of Theorem B is given in Chapter VII; this proof is based on approximation of general metric of positive curvature by polyhedral metrics and on the theorem on the existence of a convex polyhedron with given metric which was formulated in Sec. 3. Namely, we partition a space with metric of positive curvature that is homeomorphic to the sphere into small triangles; each of these triangles is replaced by the plane triangle with sides of the same length. These plane triangles form a development; we prove that this development determines a polyhedral metric of positive curvature; the finer the partition is, the closer this metric is to the given. Therefore, we can use the argument whose scheme was sketched in Sec. 5. Namely, refining triangulation indefinitely and gluing closed convex polyhedra from the corresponding developments, we obtain a sequence of polyhedra converging to a convex surface isometric to a given space.

The conditions of Theorem A are no longer sufficient for characterization of the metric of an infinite complete surface. For example, a half-sphere is homeomorphic to a plane, and its every two points can be connected by a shortest arc; but of course this half-sphere is not isometric to any convex surface. Hence an additional condition, namely, the condition of "completeness" of a metric, is necessary here.

A metric of a space R is called complete if the Weierstrass theorem holds for R, i.e., if each infinite bounded set in R has a cluster point.[30]

The necessity of the completeness condition is implied by the following theorem.

A convex surface has a complete metric if and only if this surface is complete, i.e., is the whole boundary of a convex body.

Proof. Let F be a complete convex surface. Take an infinite set of points M on this surface, which is bounded in the sense of its intrinsic metric. Since the distance

[29] Since the metric $\rho(XY)$ is a continuous function on the metric space, it is necessary to claim that this metric is continuous on the corresponding specific manifold.

[30] In the theory of metric spaces, a complete metric space is a metric space in which every Cauchy fundamental sequence converges to a certain point, i.e., $\rho(X_nX_m) \to 0$ as $n, m \to \infty$ implies that there exists a point X such that $\rho(X_nX) \to 0$. Cohn-Vossen has proved that a space with intrinsic metric (and even a locally compact space) is complete in this sense if and only if the metric of this space is complete in the sense of our definition (see S. Cohn-Vossen, On existence of shortest paths. Dokl. Akad. Nauk SSSR, Vol. III, No. 8 (1935), pp. 339–342). Cohn-Vossen used the notion that is slightly different from that of intrinsic metric; however, this does not play an essential role. The notion of complete metric itself was introduced into differential geometry by Hopf and Rinow. However, this notion was considered earlier under another name. For example, E. Cartan considered spaces with complete metrics and called them "normal spaces." See E. Cartan, Geometry of Riemannian Spaces, Ch. III.

© 2006 by Taylor & Francis Group, LLC

in the space is not greater than the distance on the surface, this set is necessarily bounded in the sense of the distance in the space and hence has at least one cluster point in the space. This point A belongs to the surface F (since the boundary of each set is a closed set, i.e., contains all its cluster points).

The point A is also a cluster point of the set M in the sense of the intrinsic metric, since we proved in Sec. 6 that the convergence of points in the sense of the distance in the space is equivalent to their convergence in the sense of the distance on the surface itself. This proves that a complete convex surface has a complete metric.

If the convex surface F is not complete, then this surface is a part of some complete surface F_0. Take two points A and B on F_0 such that A lies on F and B is outside F. Draw a shortest arc AB on F_0; let C be a point of this shortest arc which is nearest to A and does not belong to F. (The possibility of connecting two points on a complete surface by a shortest arc is proved in Sec. 2 of Chapter II.) The segment AC of our shortest arc with the deleted point C is a bounded subset on F, but its limit point C does not belong to this set. A sequence of points lying on the segment AC and converging to C has no cluster points on F. Consequently, the metric of an incomplete convex surface is not complete.

The concept of complete metric is one of the basic concepts of geometry "in the large," since intrinsic geometry in the large is in essence the theory of manifolds with complete intrinsic metric.

Infinite complete convex surfaces can be of two types: surfaces homeomorphic to a plane and cylinders. The intrinsic metric of the latter is characterized in an obvious way; for the former, we have the following theorem.

Theorem C. *A metric space homeomorphic to a plane and having complete metric of positive curvature is isometric to an infinite convex surface.*

The local characterization of the metric of an arbitrary convex surface consists of the following theorem.

Theorem D. *In a manifold with metric of positive curvature, each point has a neighborhood isometric to a convex surface.*

The proof of this theorem, given in Sec. 2 of Chapter VIII, is based on the following. We prove that in a manifold with a metric of positive curvature, each point that has a neighborhood is a convex polygon. Cut out such a polygon P from this manifold, take one more copy P' of this polygon, and identify their corresponding sides. The polygon P is homeomorphic to a disk, and, so, we obtain the manifold $P + P'$ homeomorphic to the sphere as a result of such an identification. In this manifold, the metric is defined in a natural way; i.e., the distance between two points X and Y is the greatest lower bound of the lengths of the curves connecting these points; moreover, the length is defined via the metric which we already have in P and P'. On the basis of the fact that the angles of a convex polygon do not exceed π, and hence the complete angles at the glued vertices of the polygons P and P' turn out to be no less than 2π, it is easy to prove that the metric in the manifold $P + P'$ is of positive curvature. According to Theorem B, this metric can be realized by a convex closed surface; cutting out the part corresponding to the

© 2006 by Taylor & Francis Group, LLC

polygon P from this surface, we obtain a convex surface that is isometric to this polygon. Thus, we have proved not only Theorem B but also the fact that each convex polygon homeomorphic to a disk and cut out from a manifold of positive curvature is isometric to a convex surface.

The proof of Theorem C is performed in Sec. 4 of Chapter VII by reduction to Theorem C via an analogous method. The generalization of this method of gluing a new manifold from parts of given manifolds leads us to the general "gluing theorem" which has many interesting applications. The statement and sketch of the proof of this theorem are given in Sec. 3 of Chapter IX, but a complete proof cannot be presented in the framework of this book. This gluing theorem states very general conditions under which we can "glue" a new manifold with metric of positive curvature from pieces of given manifolds with metric of positive curvature by identification of parts of boundaries of these pieces. The "gluing" of a convex surface from pieces of other convex surfaces is a particular case of this theorem; this is a far-reaching generalization of the theorem of gluing a convex polyhedron from parts of the plane.

The general gluing theorem allows us in particular to reveal which convex surfaces have the property that every two points can be connected by a shortest arc. It turns out that, e.g., such bounded surfaces can be only one of the following types: (1) a closed surface; (2) a surface isometric to a convex cap, i.e., to a convex surface with a plane boundary that is projected in a one-to-one fashion onto the plane of the boundary; (3) a surface isometric to the lateral surface of a cylinder of revolution, and a surface that can be obtained from the surface of the first or second type by deleting any closed set of points such that no shortest arcs pass through them (in particular, all conical points enjoy this property). All surfaces mentioned here and only these surfaces are bounded and possess shortest joins of all pair of points.

10. Some Singularities of the Intrinsic Geometry of Convex Surfaces

There is an essential qualitative difference between the intrinsic geometry of arbitrary convex surfaces and plane geometry or even the intrinsic geometry of regular convex surfaces. This distinction is connected with singularities that can be represented by an arbitrary convex surface. Understanding these singularities is very important, as it warns a geometer against constructing graphic patterns by analogy with plane geometry, which may fail to be in correspondence with reality, and therefore can lead to rather hasty false conclusions. In this connection, it is useful to reveal here the basic qualitative singularities of the intrinsic geometry of convex surfaces by examining a number of examples. These singularities are basically related to the distinction of the properties of shortest arcs on convex surfaces from the properties of line segments on a plane.

First of all, we list "nice" properties of shortest arcs on convex surfaces, which will be proved in later parts of the book (Sec. 2 of Chapter II and Sec. 4 of Chapter III).

1. On each complete convex surface, every two points can be connected by a shortest arc. On an arbitrary convex surface, each point has a neighborhood,

© 2006 by Taylor & Francis Group, LLC

whose every two points can be connected by a shortest arc.

2. A shortest arc is homeomorphic to a line segment. Each arc of a shortest arc is also a shortest arc. The length of a shortest arc is equal to the distance between its ends.

3. Only one shortest arc can emanate from a given point in a given direction; or, in other words, two tangent shortest arcs overlap.

4. For every two shortest arcs only those cases of mutual location are possible that take place for arcs of large disks on the sphere. Namely, two shortest arcs either have no common points or have only one common point or two common points that turn to be their common endpoints likewise two great half-circles on the sphere connecting two diametrically opposite (antipodal) points. Finally, two shortest arcs can overlap on a whole segment, so that one of these shortest arcs is a part of the other or so that one endpoint of their common segment or the endpoint of one of these arcs and the other endpoint of this segment is the endpoint of the other shortest arc likewise two overlapping arcs of great circle. No other mutual arrangements of shortest arcs are possible (see Fig. 16).

5. The limit of shortest arcs is a shortest arc, i.e., if shortest arcs converge to some curve, this curve is also a shortest arc.

Next, we list the main singularities of shortest arcs and then demonstrate them by a number of examples. Moreover, we can assume that the possible existence of two points with two shortest joins on some convex surface is not a singularity since each nonconvex domain on the plane enjoys the same property.

1. There can be pairs of points on a convex surface which can be connected not by one but by many shortest arcs; such points can be even arbitrarily close to each other.

2. There can be points from which no shortest arcs emanate in certain directions. We shall give an example of a surface whose every point enjoys this property.

3. There can be points such that no shortest arc passes through them.

4. If we connect a given point A with a variable point X by a shortest arc, the shortest arc AX can vary discontinuously while the point X moves continuously.

Singularities 2 and 3 imply that in the general case not every shortest arc can be prolonged beyond one of its ends A, so as to obtain a curve that is a shortest arc at least on an arbitrarily small part containing the point A in its interior. In other words, a geodesic line on a convex surface may fail to admit an infinite prolongation as in the case of regular surfaces.

Singularities 1, 3, and 4 exist on each convex cone (with complete angle $< 2\pi$). Let O be the vertex of such a cone, and let A and B be certain points of this cone which are different from O. Draw the generators OA and OB, cut the cone along

© 2006 by Taylor & Francis Group, LLC

these generators into two sectors, and unfold these sectors onto the plane. Since the sum of the angles of these sectors is $\theta < 2\pi$, the angle of one of them is less than or equal to $\theta/2 < \pi$. The shortest arc AB corresponds to the line segment AB drawn in this sector. We see from this that no shortest arcs can pass through the vertex of this cone. Further, if the angles of both sectors are equal, then the segments AB in these sectors are equal; and, hence, in this case the points A and B arbitrarily close to the vertex O are connected by two shortest arcs.

Let us circumscribe a circle, on our cone, centered at the vertex and passing through point A. Move point B along this circle starting from point A. When the angle ϕ between the generators OA and OB remains less than $\theta/2$, the shortest arc AB varies continuously. But when passing through the value $\phi = \theta/2$, we have a jump of the shortest arc from one side of O to the other. Consequently, the shortest arc AB varies discontinuously when the point B moves continuously.

These considerations are of a very general nature. Indeed, let no shortest arc pass through point O of some surface. We circumscribe a closed curve L about this point O, take an arbitrary point A on this curve, and move the variable point X along the curve L in a definite direction starting from point A. During this motion, the shortest arc AX cannot pass through O, and hence this shortest arc should jump through O for a certain position X_0 of point X. In this case, there exist at least two shortest arcs AX_0; these arcs are the limits of the shortest arcs AB that are obtained when point X approaches X_0 from each of the two sides. This argument does not pretend to be rigorous, although it can be made perfectly sound; it is important for us only to comprehend that the existence of points lying on no shortest arc imply other singularities of the behavior of shortest arcs.

We shall prove in Sec. 3 of Chapter IV that there are no shortest arcs passing through a conical point of a convex surface (i.e., a point with complete angle $< 2\pi$). By the way, this is sufficiently obvious by itself. A more interesting fact is that there may be noncritical points with this property (this was communicated to me by V.A. Zalgaller). For example, let G be a convex domain on the plane, and let L be the curve bounding this domain. If the curvature of the curve L is infinite at a certain point A, then there are no shortest arcs passing through point A on the doubly-covered domain G as is shown by a straightforward calculation. In this example, the point A lies on an edge of the surface, but the same property can have points at which there is a tangent plane. The next construction provides us with a surface of revolution such that no shortest arcs pass through its pole.

Let K_0 be the lateral surface of a cylinder of revolution. The shortest arcs that connect points of its base circle cover a certain annular domain H_0 on it.[31] Cut out this domain and adjoin a cone K_1 with complete angle at the vertex greater than that of K_0 to the smaller circle of this domain. Again, we distinguish a domain H_1 on this cone, which is covered by shortest arcs between the points of its base circle and adjoining a cone K_2 to the smaller circle of the domain H_1. Repeating this construction, we arrive in the limit at a convex surface of revolution; it is easy to verify that no shortest arc passes through the pole of this surface. If the complete

[31] If θ is the complete angle at the vertex of the cone K_0 and r_0 is the length of the cone generator, then the domain H_0 is a truncated cone with generator $r = r_0(1 - \cos\theta/2)$; this can be verified if we cut the cone K_0 along the generator, develop the result onto the plane, and note that the chords of the obtained sector are shortest arcs.

© 2006 by Taylor & Francis Group, LLC

angles of the cones K_1, $K_2, \ldots$ converge to 2π sufficiently fast, then this surface has tangent plane at the pole. If instead of the domains H_0, $H_1, \ldots$, we take slightly larger domains, then the edges between them can be smoothed, and we obtain a completely smooth surface of revolution such that no shortest arc passes through its pole. Also, it is worth observing that the set of points such that there are no shortest arcs passing through them has cardinality of the continuum, while the set of conical points is almost countable (this will be proved in what follows).

An example of a surface with points such that shortest arcs emanate from them in some but not all directions was already presented in Sec. 7. Namely, there are no shortest arcs on a cone of revolution that touch the base circle. The same property is characteristic, e.g., of the base circle of a cylinder or the edge of a doubly-convex lens. But all these surfaces are nonsmooth. A smooth surface such that no shortest arc emanates from some point in some direction can be obtained by the following construction. Take sphere S, choose a pole O on this sphere, draw a meridian L through O, and construct a sequence of points $A_1, A_2, \ldots$ on this meridian, which converges to O. Circumscribe disjoint circles around the points A_n. Construct cones K_n that touch the sphere S along these circles. There are no shortest arcs on the surface obtained in this way that emanate from O in the direction of the meridian L. This then is proved by a simple argument given a few lines below. If we smooth the vertices of the cones K_n we obtain a smooth surface on which there are no shortest arcs emanating from O in the direction of the meridian L. If we draw several meridians L from point O and perform the same construction for each of them then there are no shortest arcs emanating from point O in the directions of all these meridians. Also, we can take a countable everywhere-dense set of meridians L and perform the same construction for each of them, guaranteeing only that all circles C_n are disjoint (this can easily be attained if the radii of these circles decrease sufficiently fast when passing from one meridian to another; moreover, the meridians L should be enumerated in advance).

Finally, we can find a convex surface at whose every point there is an everywhere dense set of directions such that shortest arc passes through this point in any direction in this set. A surface with an everywhere dense set of conical points has this property. We can construct such a surface as follows. We begin with a regular tetrahedron P_0. Let $A_1, \ldots, A_n$ be the barycenters of the faces of this tetrahedron. We move the point A_i outside the tetrahedron and construct the convex hull of the tetrahedron P_0 and these points. If the displacements of A_i are sufficiently small then we obtain the polyhedron P_1 having all vertices of the tetrahedron P_0 and the displaced points A_i as vertices. Taking the barycenters of the faces of the polyhedron, we repeat the same construction, and proceed likewise to infinity. If the displacements of barycenters of the faces decrease sufficiently fast at each step then the complete angles at the already obtained vertices of the polyhedra P_n increase sufficiently slowly and do not tend to 2π. Therefore, in the limit we obtain a surface with a countable everywhere dense set of conical points.

Now let O be an arbitrary point of such a surface. There are conical points that are arbitrarily close to O, and without shortest arcs passing through these points. Therefore, according to the general argument above, we can conclude that there is a point B_1 connected with O by two shortest arcs. These two shortest arcs bound some digon D_1 (Fig. 10) on this surface. Since the set of conical points is everywhere

© 2006 by Taylor & Francis Group, LLC

dense, there are conical points in this digon arbitrarily close to point O. We can conclude from this that there is a point B_2 in this digon D_1 which is connected with O by two shortest arcs. These shortest arcs bound digon D_2. Repeating this argument, we obtain a sequence of digons $D_1, D_2, \ldots$, shrinking to point O. There is no shortest arc emanating from this point in each direction lying in these digons. To be more precise, let K be the tangent cone to our surface at point O. Let T_n and T_n' be generators of this cone tangent to two shortest arcs OB_n bounding the digon D_n, and let V_n be the sector bounded by them in the cone K. Since the digon D_{n+1} is in D_n, the sector V_{n+1} lies in the sector V_n. There exists a generator T lying in all sectors V_n; we assert that there is no shortest arc on our surface which touches this generator. If this shortest arc were available then this arc would necessarily intersect the shortest arcs OB_n and would have common points with them other than O (at least for large n). But this is impossible, since if two shortest arcs on a convex surface have two common points, these points are the endpoints of these arcs. The same argument is applicable to the previous example. A rigorous proof of this assertion is inessential for us since we are interested mostly in a virtual aspect of the problem.

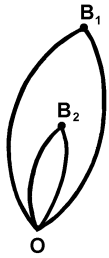


Fig. 10

Among other aspects, here we reveal a relation between the existence of singular directions in which there are no shortest arcs emanating from the point O on the one hand and the existence of points that are arbitrarily close to O and can be connected with O by two shortest arcs. This relation can be traced in the opposite direction. Indeed, if there are no shortest arcs emanating from O in the direction T, then as the point X moves around O, the shortest arc OX must jump over the direction T. There are two shortest arcs OX at the time of this jump which bound some digon containing the direction T.

Singularities of other figures on convex surfaces and, in particular, those of a circle and a disk are related to the singularities just revealed. First of all, there can be arbitrarily small disks on a convex surface which are not homeomorphic to a plane disk. This phenomenon takes place on every cone with complete angle less than π. Such a cone that is cut out along the ruling and developed on the plane is shown in Fig. 11; a family of circles with a center at point A is drawn on this cone. When the radius increases, the circle touches itself at a point B and then falls into two closed curves; one of these curves shrinks to the vertex of the cone and vanishes. The closer point A is to the vertex, the shorter radii invoke these singularities. In Sec. 6 of Chapter IX, which is devoted to the circle, we shall show that there can be circles on a convex surface which are homeomorphic to an arbitrary closed subset of a plane circle which includes at least one whole arc. These remarks show that the following theorem cannot be considered as trivial:

To each point A on a convex surface, there is some $r > 0$ such that each disk centered at point A with radius less than r is homeomorphic to a plane disk.

If two shortest arcs go from point A to point B, then the circle centered at point A and passing through B has two radii AB. In Sec. 6 of Chapter IX, we shall prove the following theorem.

Let B be a point on a circle centered at A, and X_n be a sequence of points on this circle which converges to B so that the radii AX_n converge to the radius AB. Then the angle between AB and the shortest arcs BX_n converges to a right angle.

© 2006 by Taylor & Francis Group, LLC

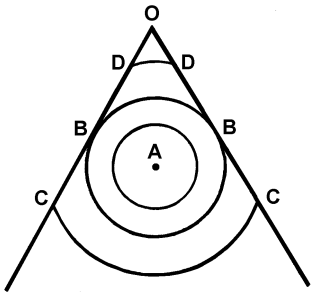


Fig. 11

Roughly speaking, this means that the circle is perpendicular to the radius. However, several radii can go to point B; in this case, the different arcs of the circle, issuing from point B, are perpendicular to different radii. Point B is a corner of the circle. This phenomenon can be observed by examining our example of a circle on the cone: in Fig. 11, the points C and D are corners of the circle. Since there can exist points B, that are arbitrarily close to point A and are connected with A by shortest arcs, even an arbitrarily small circle with a given center can have these corners.

We mention one more singularity of triangles on convex surfaces. Every triangle on the plane and every sufficiently small triangle on a regular surface is convex, i.e., a shortest arc connecting its every two points lies in this triangle. There can exist arbitrarily small nonconvex triangles on irregular convex surfaces. (For "large" triangles, this assertion is trivial since, e.g., if we take three points A, B, and C on the sphere which do not lie on the same great circle and connect these points by shortest arcs, then the surface is divided into two triangles, one of which is *a priori* nonconvex.)

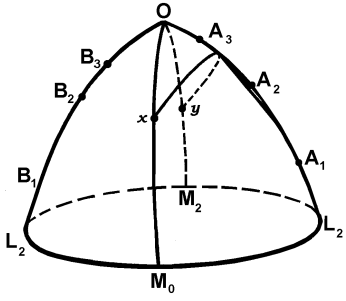


Fig. 12

In order to construct the corresponding example, we begin with a spindle-shape surface of revolution F_0 (Fig. 12). Draw two meridians OL_1 and OL_2 from its pole O that divide this surface into two equal sectors. Construct two sequences of points $A_1, A_2, \ldots$ and $B_1, B_2, \ldots$ on OL_1 and OL_2, respectively, in such a way that the distances OA_{2n} and B_{2n+1} are equal to $1/2^n$. Taking the tangent planes at the points A_{2n} and B_{2n+1}, we slightly move them to the interior of the surface F_0 and cut the resulting caps off this surface, replacing them by some pieces of the plane. We obtain a convex surface F.

Draw two more meridians OM_1 and OM_2 that bisect the angles between OL_1 and OL_2. These meridians divide the surface F into two sectors, U_1 and U_2; namely, U_1 includes the meridian OL_1, and U_2 includes the meridian OL_2. Let A and B be two points on M_1 and M_2 equidistant from the pole O. I assert that the triangle OAB is always nonconvex for all points A and B arbitrarily close to O. Indeed, take points X and Y on OA and OB equidistant from O. The shortest arc XY cannot pass through the pole O and lies in one of the sectors U_1 and U_2. If we move the points X and Y to the pole O, then the shortest arc XY will jump from time to time from the sector U_1 to the sector U_2 and vice versa. Indeed, assume, e.g., that the shortest arc always stays in the sector U_1. When the points X and Y move continuously, this shortest arc should pass the point A_{2n+1}. However, this is not possible, since no cap with the point A_{2n+1} is cut off in contrast to the point B_{2n+1}; therefore, the path from X to Y through

© 2006 by Taylor & Francis Group, LLC

A_{2n+1} is longer. Hence the triangle OAB is nonconvex irrespective of in which of the sectors U_1 and U_2 this triangle lies.

The same phenomenon can be obtained on a smooth surface. To this end, it suffices to start with a smooth surface of revolution F_0 such that there are no shortest arcs passing through its pole; moreover, we must smooth the edges of the cuts.

This singularity of triangles shows that, e.g., we cannot consider as trivial the theorem asserting that each closed convex surface can be partitioned into arbitrarily small convex triangles. And, indeed, the proof of this theorem, which is proposed in Sec. 6 of Chapter II, turns out to be rather complicated.

We also note some singularities of the angle made by shortest arcs. As we have already shown in Sec. 7, it is not always possible to draw an angle that is equal to a given angle and has a given shortest arc OA as one of its sides since there can be no shortest arc emanating from point O in the corresponding direction. Further, if we take a point X different from A and move this point continuously then the angle between the shortest arcs OA and OX can vary discontinuously since so does the shortest arc OX itself. However, we shall prove (Sec. 4 of Chapter IV) that if the shortest arcs OX converge to a shortest arc OB, then the angle between OX and OA converges to the angle between OB and OA. At the same time, it turns out that if points O_n converge to O and the shortest arcs L_n and M_n emanating from O_n converge to shortest arcs L and M emanating from O, then the angles between L_n and M_n may fail to converge to the angle between L and M. This phenomenon can be observed on each polyhedron, e.g., on a cube (Fig. 13). Let O be a vertex of this cube, and let O_n be points on one of its edges contiguous to the vertex O. Let L and M be two other edges of this cube meeting at the vertex O, and let L_n and M_n be some segments on the faces which are parallel to the edges L and M and are drawn from the points O_n. The angle between the edges L and M is obviously equal to $\pi/2$, and the angle between the segments L_n and M_n is equal to π (this is the angle measured on the cube!). Therefore, the angle between L_n and M_n does not *a priori* converge to the angle between L and M. If we consider the sequence of pairs of the shortest arcs L_1 and L and M, L_2 and M_2, L and M, etc., the angles between them are $\pi, \pi/2, \pi, \pi/2, \ldots$, and hence do not converge to any limit at all. The conditions under which convergence of the angles takes place despite the above will be revealed in Sec. 4 of Chapter IV.

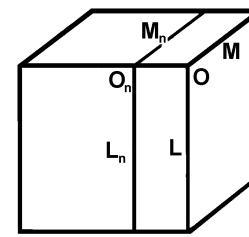


Fig. 13

11. Theorems of the Intrinsic Geometry of Convex Surfaces

Everything discussed above refers to the foundations of intrinsic geometry of convex surfaces. We now want to observe in general terms some results that can be obtained in the framework of this geometry and to demonstrate the characteristic features of these results. If we follow the scheme of presentation of elementary geometry, then it is necessary to begin with the theory of triangles. We mention some relevant theorems.

© 2006 by Taylor & Francis Group, LLC

1. The angles of a triangle on a convex surface are no less than the corresponding angles of the plane triangle having sides of the same length (Theorem 3 in Sec. 4 of Chapter III).

2. If the curvature of the interior of a triangle, i.e., the sum of its angle minus π, is equal to zero, then this triangle is isometric to a plane triangle (Theorem 3 in Sec. 6 of Chapter V).

3. If a, b, and c are the sides of a triangle, ω is the curvature of its interior, γ is the angle opposite the side c, and $\pi \geq \gamma \geq \omega$, then

$$a^2 + b^2 - 2ab\cos\gamma \geq c^2 \geq a^2 + b^2 - 2ab\cos(\gamma - \omega). \tag{1}$$

If $\gamma < \omega$ (or $> \pi$), the inequality on the right-(left-)hand side is replaced by the trivial inequality

$$c \geq |a - b| \quad (c \leq a + b).$$

This theorem is an immediate corollary of the first. Inequalities (1) turn out to be sharp in the sense that for a given a, b, c, γ, and ω there exists a triangle, and even a triangle on a convex polyhedron, for which c is arbitrarily close to the left- or right-hand side of formula (1).

4. If σ is the area of a triangle T on a convex surface and σ_0 is the area of the plane triangle T_0 with sides of the same length, then $0 \leq \sigma - \sigma_0 \leq \omega d^2/2$ where ω is the curvature of the interior of the triangle T and d is its diameter. Moreover, $\sigma = \sigma_0$ if and only if the triangle T is isometric to the triangle T_0. (See Sec. 1 of Chapter X.)

The list of the theorems of this type can be expanded. However, the general character of these theorems is now clear. We can continue the comparison of a triangle on a convex surface with the corresponding triangle on the plane; moreover, the difference of the former from the latter will be estimated through the curvature of the interior of the triangle. For example, to each trigonometrical formula connecting the sides and the angles of a plane triangle, we can put in correspondence some inequalities that give an estimate of deviation of the triangle on this convex surface from such a formula independent of the curvature of the interior of the triangle. The point is that we study the intrinsic geometry of an arbitrary rather than given convex surface; owing only to this fact, our theorems must be mostly of a qualitative character and simple quantitative relations must appear mostly as inequalities. The sharp relations will inevitably become rather involved in most cases.

Note two more theorems of the same type.

1. We mentioned in Sec. 10 that, for each point O on a convex surface F, there exists $r_0 > 0$ such that each circle on F of radius less than r_0 centered at O is a simple closed curve. Let $r < r_0$, l be the length of the circle of radius r centered at O, let θ be the complete angle around the point O, and ω be the curvature of the interior of the correspondent disk with the center O deleted. The ratio $\dfrac{l}{r}$ does not increase as r increases and satisfies the inequalities $\theta \geq \dfrac{l}{r} \geq \theta - \omega$; moreover, $l = \theta r$ if and only if this disk is isometric to the

© 2006 by Taylor & Francis Group, LLC

disk on a cone with complete angle θ. (The proof is given in Sec. 6 of Chapter IX.)

2. The cone of revolution has maximal area among all convex surfaces bounded by a simple closed curve of given length and given curvature less than 2π. This theorem generalizes the well-known maximality property of a disk. (The proof is given in Sec. 3 of Chapter X.)

Some interesting general problems of intrinsic geometry arise in the theory of curves on convex surfaces; Chapter IX is devoted to this theory. The study of curves bounding convex domains or the study of geodesics "in the large" on convex surfaces can serve as an example of these problems. The following theorem is an example: if there is at least one curve on a convex surface which can be prolonged indefinitely to both sides and whose every segment is a shortest arc, then this surface is isometric to the plane or a cylinder. For regular surfaces, this theorem was proved by S. Cohn-Vossen, one of a number of beautiful results concerning the properties of geodesics "in the large" on infinite complete surfaces.[32] The problem of abstracting the Cohn-Vossen results to more general convex surfaces deserves attention and is, probably, not very difficult.

A specific class of theorems of intrinsic geometry arises if we consider those convex surfaces on which the ratio of the curvature and the area of an arbitrary domain is subject to various restrictions (see Chapter XI). This ratio is equal to $1/R^2$ for all domains on the sphere of radius R. It turns out that if this ratio lies in the interval between $1/R_1^2$ and $1/R_2^2$ on a surface F then the intrinsic properties of the surface F turn out to be intermediate in some respect between the properties of two spheres of radii R_1 and R_2. We can present the following theorems as examples, in which we assume that $R_1 \geq R_2$ and denote by S_1 and S_2 the spheres of radius R_1 and R_2, respectively.

1. Let T be a triangle on such a surface F, and let T_1 and T_2 be triangles on the spheres S_1 and S_2 with sides of the same length as T. If α, α_1, and α_2 are the corresponding angles of these triangles then $\alpha_1 \leq \alpha \leq \alpha_2$. If σ, σ_1, and σ_2 are the areas of these triangles then $\sigma_1 \leq \sigma \leq \sigma_2$.

2. Let l, l_1, and l_2 be the lengths of circles of the same radius on the surface F and the spheres S_1 and S_2, and let σ, σ_1, and σ_2 be the areas of the disks of the same radius. Then we have the inequalities $l_1 \geq l \geq l_2$ and $\sigma_1 \geq \sigma \geq \sigma_2$.

3. The upper bound l of the lengths for which a geodesic on the surface F is a shortest arc, as compared with all curves arbitrarily close to it, satisfies the Bonnet estimates[33] $\pi R_1 \geq l \geq \pi R_2$. Here πR_1 and πR_2 are the lengths of

[32] S. Cohn-Vossen, Totalkrümmung und geodätische Linien auf einfachzusammenhängen offenen vollständingen Flächenstücken. Mat. Sb. (new ser.), Vol. 1 (1936), pp. 139–163. Cohn-Vossen considers the question from a purely intrinsic point of view; that is, he considers the plane with a given line element of positive curvature. However, in view of our realization theorem for such a metric this is perfectly immaterial whether we speak about an abstract metric or about a complete convex surface.

[33] For regular surfaces, these estimates were found by Bonnet in 1855; these estimates can be found in every comprehensive work of differential geometry, e.g., "Differential Geometry" by W. Blaschke (Secs. 98–100).

© 2006 by Taylor & Francis Group, LLC

the great half-circles of the spheres S_1 and S_2, i.e., the maximal lengths of shortest arcs on them.

Finally, it is natural to mention the following theorem. If the ratio of the curvature and area of domain of a surface F is equal to $1/R^2$, then each point on F has a neighborhood isometric to a part of the sphere of radius R. In addition, if F is a closed convex surface then this surface itself is the sphere of radius R.

The limit of the ratio of the curvature and area of a domain as this domain shrinks to some point X, if this limit exists, is the Gaussian curvature at the point X. This limit always exists on regular surfaces. At the same time, its existence also proves sufficient for a convex surface to be so regular from the intrinsic point of view that we can apply the basic tools of Gaussian intrinsic geometry to this surface. Namely, the following theorem holds.

Let the ratio of the curvature of a triangle on a surface F to the area of this triangle always tend to a definite limit $K(X)$ as this triangle shrinks to a point X. (It is not necessary that this triangle contains the point X; the limit $K(X)$ can certainly depend on the point X, but this limit turns out to be a continuous function of this point.) Then we can introduce the geodesic polar coordinates r, ϕ in a neighborhood of each point O on the surface F; the metric of this surface can be given by the line element $ds^2 = dr^2 + Gd\phi^2$; moreover, the coefficient G is a continuous function of r and ϕ and has continuous partial derivatives $\partial G/\partial r$, $\partial^2 G/\partial r^2$. The following Gaussian formula for the curvature holds:

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

Although the coefficient G can have no derivative $\partial G/\partial\phi$, we can transform the usual formula for the geodesic curvature and the differential equation of geodesics in such a way that they become applicable in our neighborhood, where the coordinates r, ϕ are introduced.

Changing the formulation of this theorem, we can transform it into a theorem that provides a definition of Gaussian intrinsic metric given by line element in terms mentioning neither coordinates no line element and, moreover, without the assumption of positive curvature. The latter arises from the following very important fact. It turns out that the concepts and methods, which are discussed in regard to convex surfaces in Euclidean space, can be abstracted almost without changes to convex surfaces in Lobachevskii space; and, for these surfaces, we can obtain results that are absolutely similar to those formulated in this chapter for convex surfaces of Euclidean space. For example, for this surface we have a theorem completely analogous to Theorem A of Sec. 9 with the only difference that in Condition 3 of this new theorem we must assume that the sum of the lower angles between the sides of a triangle is at least the sum of the angles of the corresponding triangle with sides of the same length in the corresponding Lobachevskii space. For example, this implies that each surface homeomorphic to the sphere and of Gaussian curvature greater than some K can be embedded as a convex surface in Lobachevskii space of curvature at most K. Thus, the intrinsic geometry of convex surfaces of Lobachevskii space covers the intrinsic geometry of all surfaces whose curvature is bounded below. This generalization of our theory will be considered in more detail

© 2006 by Taylor & Francis Group, LLC

in Chapter XII; in particular, there we give a precise formulation of the general theorem, yielding the definition of Gaussian intrinsic metric which uses neither the concept of coordinates no that of line elements. Right now we confine discussion to the above general remarks.

© 2006 by Taylor & Francis Group, LLC

Chapter II

GENERAL PROPOSITIONS ABOUT INTRINSIC METRICS

1. General Theorems on Rectifiable Curves

The general theorems on rectifiable curves, i.e., curves of finite length in an arbitrary metric space, which will be proved here, literally repeat the well-known theorems on curves in Euclidean space. However, it is useful to present their proofs in order to confirm that in fact, in these theorems, we use nothing except for three basic properties of an arbitrary metric[1] and the definition of length.

First of all, we make more precise the concept of a curve. This should be done for two reasons. First, the definition of a curve as a set of points that is a continuous image of a line segment is not appropriate, since we use some parametric representation of a curve when defining the length. For example, a line segment which is transversed twice in one direction first and then in the opposite direction has length twice that of the same segment transversed once. It is appropriate to consider a twice transversed segment as a curve other than a once transversed segment. But, on the other hand, we find it inconvenient to define a continuous curve as a set of points $X(t)$ with a given parameterization, since the same line segment transversed with different velocities, i.e., such that different points correspond to equal values of the parameter, will be considered as two different curves, although it is natural to consider this segment as the same curve. These reasons force us to give the following definition of equal and different curves.

Let R be a metric space with metric ρ. Assume that to each value of the parameter t on the closed interval $[0, 1]$, there corresponds a unique point $X(t)$ of the space R such that for every $\varepsilon > 0$ there is some $\delta > 0$ satisfying $\rho(X(t_1)X(t_2)) < \varepsilon$ whenever $|t_1 - t_2| < \delta$. We obtain a continuous mapping of the closed interval $[0, 1]$ into the space R and the set L of points $X(t)$ as the result of this mapping. Suppose that we have another mapping of a closed interval $[a, b]$ into the space R; the parameter s is varied on the closed interval $[a, b]$, and to each s, we put in correspondence a point $Y(s)$, and, moreover, this mapping is also continuous. We say that both mappings $X(t)$ and $Y(t)$ define the same curve if the following condition holds: the parameter s can be represented as a monotone function of the parameter t; i.e., $s = f(t)$ so that $X(t) = Y(f(t))$ for any t. Then, conversely, the parameter t turns out to be a monotone function of the parameter s; i.e., $t = f^{-1}(s)$ and for all S we have $Y(s) = X(f^{-1}(s))$. Hence, this definition is

[1] That is, (1) $\rho(XY) = 0$ if and only if $X = Y$, (2) $\rho(XY) = \rho(YX)$, and (3) $\rho(XY) + \rho(YZ) \geq \rho(XZ)$.

© 2006 by Taylor & Francis Group, LLC

symmetric; obviously, this definition also satisfies the reflexivity and transitivity requirements. It follows from this definition that the points $X(t)$ and $Y(t)$ form one and the same set.[2]

It is important to keep in mind the following fact: our monotone function $s = f(t)$ can have a constant value s_0 on some closed interval $[t_1, t_2]$. Then the point $X(t)$ remains fixed and coincides with $Y(s_0)$ when t runs over the closed interval $[t_1, t_2]$. Then the inverse function $t = f^{-1}(s)$ for $s = s_0$ assumes all values in the closed interval $[t_1, t_2]$. Thus, none of the functions $f(t)$ and $f^{-1}(s)$ are assumed to be single-valued; they are only assumed to be monotone, i.e., for $t_1 < t_2$, we always have $f(t_1) \leq f(t_2)$ or $f(t_1) \geq f(t_2)$ independent of the values of $f(t_1)$ and $f(t_2)$ (there can be a whole segment of such values).

It is convenient to use the following kinematic representation: when the parameter t increases from 0 to 1, the point $X(t)$ moves and continuously traverses the curve L. A monotone transformation of the parameter is reduced to the fact that we again run over the same points $X(t)$ in the same or opposite order. The velocity of motion can be arbitrary; even the stoppings are admissible (the parameter varies, while the point is fixed). But no turn is admissible anywhere.

Let L be a curve in a metric space with metric $\rho(XY)$, and let us have some parameter t $(0 \leq t \leq 1)$ on the curve L. Then we can denote by $X(t)$ $(0 \leq t \leq 1)$ this curve L assuming that $X(t)$ corresponds to t. However, this curve can be parameterized in finitely many ways as we have just seen.

The length of a curve $X(t)$ is the least upper bound of the sums

$$\sum_{i=1}^{n} \rho(X(t_{i-1})X(t_i)), \tag{1}$$

where $0 = t_0 < t_1 < t_2 \leq \ldots < t_n = 1$.

The length of a curve does not depend on the choice of a parameter on this curve. Indeed, let this curve be represented as $Y(s)$ $(a \leq s \leq b)$. Then instead of sum (1) we have

$$\sum_{i=1}^{m} \rho(Y(s_{i-1})Y(s_i)),$$

where $s_i = f(t_i)$. By condition, $f(t)$ should be a monotone function, and since $t_0 < t_1 < \ldots < t_n$, we have $s_0 \leq s_1 \leq \ldots \leq s_n = 1$, or, otherwise, $s_0 \geq s_1 \geq \ldots \geq s_n = 1$. Therefore, if we slightly generalize several sums (1) admitting equal values of the parameter in them, it becomes obvious that these sums are the same for different choices of parameters. Hence the least upper bound of these sums, i.e., the length of the curve, does not depend on the choice of the parameter.[3]

Meanwhile, it is obvious that in allowing nonmonotonic transformations of the parameter, we introduce other sums and can obtain another value of the length. Consequently, the length of a curve does not depend on parameterization by the condition of monotonicity of admissible transformations of the parameter.

[2] The function $f(t)$ is assumed to map the whole domain of t onto the whole domain of s.

[3] If $t_{i-1} = t_i$, then $X(t_{i-1}) = X(t_i)$, and, therefore, $\rho(X(t_{i-1})X(t_i)) = 0$, i.e., the corresponding summand is nonessential. Therefore, we can always assume that $t_{i-1} < t_i$. When we pass to another parameter s, some different values of s can correspond to the same t. However, the same point of the curve corresponds to these values, and, therefore, $\rho(Y(s_{i-1})Y(s_i)) = 0$ again.

© 2006 by Taylor & Francis Group, LLC

In the later sections of the book, we will speak about curves in a given metric space R with metric ρ that have finite length. The following main assertions can be proved for these curves.

Theorem 1. *If a curve of finite length is divided into finitely many arcs, then each of these arcs has finite length and the sum of the lengths of these arcs is equal to the length of the whole curve.*

Theorem 2. *The arc length $s(t)$ of a curve between $X(0)$ and $X(t)$ is a monotonically nondecreasing function of t.*

Theorem 3. *On each curve, we can introduce the arc length counted from $X(0)$ as a parameter.*

We say that a sequence of curves L_n converges to a curve L if we can choose a parameter t $(0 \le t \le 1)$ on these curves L_n and L such that for all $\varepsilon > 0$ there exists N such that $\rho(X_n(t)X(T)) < \varepsilon$ for $n > N$ and for all t, where $X_n(t)$ and $X(t)$ are points on the curve L_n and L.

Theorem 4. *We can choose a convergent subsequence from each infinite set of curves in a compact domain of length not exceeding a given one.*

Here by a compact domain we mean a set with the property that each infinite sequence of its points has a condensation point in this set.

Theorem 5. *If curves L_n converge to a curve L, then the length of L is not greater than the lower limit of the lengths of L_n.*

Proof of Theorem 1. Assume that a curve L with endpoints A and B is divided into two parts L_1 and L_2 by a point C. The length of L is the least upper bound of sums of the form (1) for an arbitrary choice of the points $X(t_i)$. When we define the lengths of its arcs L_1 and L_2, the point C should be one of the points of the partition, since this point is the end of these arcs. Hence, the arbitrariness in the choice of points of the partition in the former case is greater than that in the latter case. Therefore, the least upper bound of sums (1) in the former case cannot be less than in the latter case, and thus,

$$s(L) \ge s(L_1) + s(L_2), \tag{2}$$

where s stands for the arc length.

We now fix any $\varepsilon > 0$; let $X(t_i) = x_i$ be points of a partition of the curve L such that sum (1) differs from the length $s(L)$ by less than ε, i.e.,

$$\sum_{i=1}^{n} \rho(X_{i-1}X_i) > s(L) - \varepsilon. \tag{3}$$

Assume that the point C lies between the points X_{k-1} and X_k (i.e., the value of t corresponding to C lies between t_{k-1} and t_k). By the triangle inequality,

$$\rho(X_{k-1}X_k) \le \rho(X_{k-1}C) + \rho(CX_k), \tag{4}$$

© 2006 by Taylor & Francis Group, LLC

and, therefore,

$$\sum_{i=1}^{n} \rho(X_{i-1}X_i) \le \left[\sum_{i=1}^{k-1} \rho(X_{i-1}X_i) + \rho(X_{k-1}X_k)\right] + \left[\sum_{i=k+1}^{n} \rho(X_{i-1}X_i) + \rho(X_kC)\right]. \tag{5}$$

But it is clear from the definition of the length that the sums in the square brackets are not less than the length of L_1 and the length of L_2, respectively, and, therefore,

$$\sum_{i=1}^{n} \rho(X_{i-1}X_i) \le s(L_1) + s(L_2). \tag{6}$$

Comparing this inequality with (3), we obtain

$$s(L) - \varepsilon \le s(L_1) + s(L_2); \tag{7}$$

since ε is arbitrary, this implies

$$s(L) \le s(L_1) + s(L_2). \tag{8}$$

Comparing this inequality with (2), we conclude that

$$s(L) = s(L_1) + s(L_2). \tag{9}$$

The theorem is proved for partition of a curve into two arcs. This easily implies the same result for a partition into any finite number of arcs.

Proof of Theorem 2. It is clear from Theorem 1 that the length $s(t)$ of the arc between the points $X(0)$ and $X(t)$ is a monotonically nondecreasing function of t. Therefore, in order to prove the continuity of $s(t)$, it is sufficient to show that for any t and $\varepsilon > 0$ there is $\delta > 0$ such that

$$s(t + \Delta t) - s(t) < \varepsilon, \tag{10}$$

whenever $0 < \Delta t < \delta$. (This will prove the right continuity of $s(t)$, i.e., that for $t_i > t$ and $t_i \to t$ we have $s(t_i) \to s(t)$; but interchanging the endpoints of the curve and introducing the parameter $1 - t$ instead of t, in exactly the same way we obtain the left continuity of $s(t)$.)

Assume that for some $t = T$ this property fails. This means that there is $\varepsilon > 0$ such that for an arbitrarily small $\Delta t > 0$, the inequality

$$s(T + \Delta t) - s(T) \ge \varepsilon \tag{11}$$

holds. By the monotonicity of $s(t)$, this inequality holds for all $\Delta t \ge 0$. Take the partition points $X(t + i) = X_i$ so that the length of the curve $s(L) = s(1)$ differs from sum (1) by less than $\varepsilon/2$, i.e.,

$$s(1) < \sum_{i=1}^{n} \rho(X_{i-1}X_i) + \frac{\varepsilon}{2}. \tag{12}$$

© 2006 by Taylor & Francis Group, LLC

Since adding new partition points can only increase the sum of distances (by the triangle inequality), we can assume that there are two neighboring partition points,

$$X_k = X(T) \text{ and } X_{k+1} = X(T + \Delta t),$$

adding, if necessary, two partition points; moreover, we can assume that Δt is chosen so that

$$\rho(X_k X_{k+1}) \le \frac{\varepsilon}{2}. \tag{13}$$

This is possible by the continuous dependence of $X(t)$ on t.

Now the curve is divided into three arcs corresponding to the closed intervals $[0, T]$, $[T, T + \Delta t]$, and $[T + \Delta t, 1]$. By Theorem 1, their lengths are equal to

$$s(t), \quad s(T + \Delta t) - s(T), \quad S(1) - s(T + \Delta t),$$

respectively, and their sum is equal to the length $s(1)$ of the whole curve. Moreover, by the definition of length,

$$s(T) \ge \sum_{i=1}^{k} \rho(X_{i-1} X_i),$$

$$s(1) - s(T + \Delta t) \ge \sum_{i=k+2}^{n} \rho(X_{i-1} X_i).$$

Therefore, inequality (12) implies

$$s(1) < s(T) + [s(1) - s(T + \Delta t)] + \rho(X_k X_{k+1}) + \frac{\varepsilon}{2},$$

i.e.,

$$s(T + \Delta t) - s(T) \le \rho(X_k X_{k+1}) + \frac{\varepsilon}{2};$$

by inequality (13), this implies

$$S(T + \Delta t) - s(T) < \varepsilon.$$

However, this inequality contradicts inequality (13); therefore, inequality (11) is not possible, and the theorem is proved.

Proof of Theorem 3. Since the arclength $s(t)$ of a curve $X(t)$ is a continuous function of t, this function assumes all values from zero up to the length S of the whole curve, and a single point corresponds to each value of the length s of the arc; this is the endpoint of the arc of this length. Therefore, with each number s (the value of the arclength) in the closed interval $[0, S]$, we associate a single point $Y(s)$ on our curve. Meanwhile, the definition of the arclength easily implies

$$\rho(Y(s + \Delta s) Y(s)) \le |\Delta s|,$$

i.e., whenever $|\Delta s|$ is sufficiently small, the distance between the points $Y(s + \Delta s)$ and $Y(s)$ is also small. This means that the point $Y(s)$ depends continuously on s, and, thus, $Y(s)$ is a continuous curve when we take arclength as parameter. This is the same continuous curve $X(t)$, since $s(t)$ is a monotone function, and, therefore, this transformation of the parameter is admissible.

© 2006 by Taylor & Francis Group, LLC

Proof of Theorem 4. Consider a sequence of curves L_n in a compact domain G of length less than S. Introduce the parameter t on each of these curves, which is equal to the arclength divided by the length of the whole curve. Then each curve L_n is represented as $X_n(t)$, where $0 \le t \le 1$. Take all rational values of t and enumerate them in an arbitrary order: $t_1, t_2, \ldots$. Choose a subsequence $L_{11}, L_{12}, \ldots$ from the curves L_n so that the points of these curves corresponding to the value t_1 converge to some point $X(t_1)$; this is possible, since the domain G is compact. We then choose a sequence of curves $L_{21}, L_{22} \ldots$ from this sequence such that the points of these curves corresponding to t_2 converge to some point $X(t_2)$. We proceed in the same manner up to infinity; after that, we take the "diagonal" sequence $L_{11}, L_{22}, \ldots$ for which we have $X_{nn}(t_i) \to X(t_i)$ for all t_i as $n \to \infty$.

Now let t be some value of the parameter between zero and unity, and let ε be a given arbitrarily small positive number. Take a rational number t_i such that

$$\rho(X_{nn}(t_i)X_{nn}(t)) \le \varepsilon \tag{14}$$

for all n. This is always possible. Indeed, if $s_{nn}(t, t_i)$ is the arclength of L_{nn} between $X_{nn}(t)$ and $X_{nn}(t_i)$, then by the definition of length and by the choice of t, we have

$$\rho(X_{nn}(t_i)X_{nn}(t)) \le s_{nn}(t, t_i) = |t - t_i| s(L_{nn}) < |t - t_i| S,$$

where $s(L_{nn})$ is the length of curve L_{nn} (and by condition, all $s(L_{nn}) < S$). Therefore, if we take t_i so that $|t - t_i|S < \varepsilon$, then inequality (14) holds.

Further, take N so large that for $k, m > N$ we have

$$\rho(X_{kk}(t_i)X_{mm}(t_i)) < \varepsilon. \tag{15}$$

This is possible, since the points $X_{nn}(t_i)$ converge. Then by the triangle inequality, it follows from (14) and (15) that

$$\rho(X_{kk}(t)X_{mm}(t)) < 3\varepsilon,$$

whenever k and $m > N$. Hence, the points $X_n n(t)$ form a Cauchy sequence and converge to some point $X(t)$ by the compactness of the domain under study.

The points $X(t)$ $(0 \le t \le 1)$ comprise a continuous curve. Indeed, fix $\varepsilon > 0$; then by the choice of the parameter t, for all t' and t'' such that $|t' - t''| < \varepsilon/S$ and for all n, we have

$$\rho(X_{nn}(t')X_{nn}(t'')) < |t' - t''| S < \varepsilon.$$

Since $X_{nn}(t')$ and $X_{nn}(t'')$ converge to $X(t')$ and $X(t'')$, we then find in the limit that

$$\rho(X(t')X(t'')) \le |t' - t''| S < \varepsilon. \tag{15a}$$

This means that $X(t)$ depends continuously on t, and hence it is a continuous curve.

It remains to show that the curves L_{nn} converge to this curve. We have demonstrated only that $X_{nn}(t) \to X(t)$ for every t but we have to prove that for all $\varepsilon > 0$, there is N such that $\rho(X(t), X_{nn}(t)) < \varepsilon$ for all t whenever $n > N$. To prove this, we take values $t_1, t_2, \ldots, t_k$ of the parameter such that for all t there exists t_i with

© 2006 by Taylor & Francis Group, LLC

$|t - t_i| < \varepsilon/(3S)$. Since $X_{nn}(t_i) \to X(t_i)$ for all t_i, there is N such that for $n > N$ and for all i $(i = 1, 2, \ldots k)$, we have

$$\rho(X(t_i)X_{nn}(t_i)) < \frac{\varepsilon}{3}. \tag{16}$$

Now let t be an arbitrary value of the parameter, and let t_i be such that

$$|t - t_i| < \frac{\varepsilon}{3S}. \tag{17}$$

The triangle inequality implies

$$\rho(X(t)X_{nn}(t)) \le \rho(X(t)X(t_i)) + \rho(X(t_i)X_{nn}(t_i)) + \rho(X_{nn}(t_i)X_{nn}(t)).$$

By (15a) and (17),

$$\rho(X(t)X(t_i)) < |t - t_i|S < \frac{\varepsilon}{3},$$

and in exactly the same way,

$$\rho(X_{nn}(t)X_{nn}(t_i)) < \frac{\varepsilon}{3}.$$

Consequently, taking (16) into account, we have $\rho(X(t)X_{nn}(t)) < \varepsilon$ $(n > N)$, i.e., the curves $X_{nn}(t)$ converge to $X(T)$; as required.

Proof of Theorem 5. Let the curves L_n converge to L, and let $s(L_n)$ and $s(L)$ be the lengths of these curves; moreover, assume that

$$S = \liminf_{n\to\infty} s(L_n) < \infty,$$

i.e., the lower limit of the lengths of L_n is finite. Let L_n $(n = 1, 2, \ldots)$ be a subsequence of the sequence L_n such that

$$\lim_{n\to\infty} s(L_{nn}) = S.$$

Obviously, the curves L_{nn} also converge to L; by the definition of convergence, we can choose the parameter t on them such that $X_{nn} \to X(t)$. Let $0 = t_0 < t_1 < \ldots < t_n = 1$, and let ε be an arbitrary positive number. Then there is N such that for all $n \ge N$ and for all chosen t_i, we have

$$\rho(X_{nn}(t_i), X(t_i)) \le \frac{\varepsilon}{2m},$$

and then, by the triangle inequality,

$$\rho(X(t_{i-1}X(t_i)) \le \rho(X_{nn}(t_{i-1}), X_{nn}(t_i)) + 2\frac{\varepsilon}{2m};$$

this implies

$$\sum_{i=1}^{m} \rho(X(t_{i-1})X(t_i)) \le \sum_{i=1}^{m} \rho(X_{nn}(t_{i-1})X_{nn}(t_i)) + \varepsilon.$$

© 2006 by Taylor & Francis Group, LLC

But

$$\sum_{i=1}^{m} \rho(X_{nn}(t_{i-1})X_{nn}(t_i)) \le S(L_{nn})$$

and $s(L_{nn}) < S + \varepsilon$ whenever n is sufficiently large. Therefore,

$$\sum_{i=1}^{m} \rho(X(t_{i-1})X(t_i)) < S + 2\varepsilon,$$

and since ε is arbitrary, we have

$$\sum_{i=1}^{m} \rho(X(t_{i-1})X(t_i)) \le S.$$

The length of L is the least upper bound of the sums on the left-hand side of this inequality, and, therefore, this length was not less than S; this is the claim.

2. General Theorems on Shortest Arcs

The properties of shortest arcs proved here will be needed in various sections of the subsequent presentation; here we list them in order to assume them known in what follows. We will speak about shortest arcs in a manifold with an arbitrary intrinsic metric $\rho(XY)$. By definition, a shortest arc is an arc of minimal length among all curves joining two given points.

A shortest arc between two points X and Y (A and B, etc.) will be denoted by XY (AB, etc.). Of course, not every two points can be connected by a shortest arc in an arbitrary manifold, and, meanwhile, it can happen that two given points X and Y can be connected by several shortest arcs. In this case, if the contrary is not specified, XY stands for any of these shortest arcs.

The definition of shortest arc obviously implies the following properties.

1. Each segment of a shortest arc is a shortest arc, since when we replace this segment by a shorter curve, the shortest arc will be shortened, which is impossible.

2. A curve joining two given points is a shortest arc if and only if the length of this curve is equal to the distance between these points. This is a straightforward consequence of the fact that the distance is the greatest lower bound of the lengths of the curve in the intrinsic metric.

3. Two shortest arcs XY and YZ together form a shortest arc XZ if and only if $\rho(XY) + \rho(YZ) = \rho(XZ)$. This remark follows from the fact that the lengths of the shortest arcs XY and YZ are equal to $\rho(XY)$ and $\rho(YZ)$ and the curve $XY + YZ$ is a shortest arc if and only if its length is equal to $\rho(XZ)$.

4. A shortest arc is homeomorphic to a line segment. In fact, if s is the arclength of a shortest arc counted from one of its ends, then according to Theorem 3 of the previous section, this shortest arc can be given as $X(s)$, i.e., the points of this shortest arc depend continuously on s when s runs over all values from

© 2006 by Taylor & Francis Group, LLC

0 to s, where s is the length of the whole shortest arc. Hence, we have a continuous mapping of the closed interval $[0, 1]$ onto this shortest arc. This mapping is one to one, since when the same point corresponds to two values s' and s'', we shorten the shortest arc by removing the arc corresponding to the interval (s', s''), which is impossible. Finally, by the properties of a shortest arc indicated above, for any two points $X(s')$ and $X(s'')$ of our shortest arcs, we have $\rho(X(s')X(s'')) = |s' - s''|$. This obviously implies that the mapping of the closed interval $[0, s]$ onto a shortest arc is a homeomorphism.

Theorem 1. *If a sequence of shortest arcs converges then the limit of this sequence is a shortest arc.*

Proof. Let shortest arcs X_nY_n converge to a curve L; then the points X_n and Y_n converge to the endpoints X and Y of this curve L (this follows from the definition of the convergence of curves of the preceding section). In this case,

$$\rho(XY) = \lim_{n\to\infty} \rho(X_nY_n).$$

On the other hand, if $s(L)$ is the length of L then

$$s(L) \geq \rho(XY).$$

At the same time, the lengths of the shortest arcs X_nY_n are equal to the distances $\rho(X_nY_n)$, and by Theorem 5 of the preceding section, the length of the limit curve cannot be greater than the limit of their length; i.e.,

$$s(L) \leq \lim_{n\to\infty} \rho(X_nY_n).$$

Comparing the formulae obtained, we see that the length of L is equal to the distance between its ends, and hence L is a shortest arc.

Theorem 2. *If two points can be joined by a curve of a given length in a given compact domain, then there is a shortest arc of these two points in the domain under study.*

Proof. Let A and B be two points of a compact domain G, and let s be the length of curve joining these points. Finally, let s_0 be the greatest lower bound of lengths of all curves joining the points A and B in G. Obviously, $s_0 \leq s$, and there exists a sequence of curves L_n joining A and B in G whose lengths $s(L_n)$ tend to s_0. (In the exceptional case, all the curves L_n can coincide with one another.) By Theorem 4 of Sec. 1, we can choose a subsequence from the curves L_n which converges to some curve L. By Theorem 5 of Sec. 1, the length of L does not exceed the lower limit of the lengths $s(L_n)$, and, thus, $s(L) \leq s_0$; since s_0 is the greatest lower bound of the lengths of all curves connecting A and B, we have $s(L) = s_0$; thus, the curve L is a shortest arc in G.

Since a closed convex surface is homeomorphic to the sphere, this surface is compact. There is a curve of finite length joining every two points X and Y on this surface, e.g., a convex curve that is obtained in the section of the surface by a plane passing through the points X and Y. Therefore, Theorem 2 implies that every two points on a closed convex surface can be connected by a shortest arc.

Theorem 4 easily implies even the following stronger assertion.

© 2006 by Taylor & Francis Group, LLC

Theorem 3. *There is a shortest arc of every two points on a manifold with complete intrinsic metric.*

Proof. Let A and B be two points of a manifold with complete intrinsic metric $\rho(XY)$. By the definition of intrinsic metric, there exists a sequence of curves L_n connecting the points A and B whose lengths converge to the distance $\rho(AB)$. Obviously, we can assume that the lengths of all L_n are not greater than some s_0. It is clear that all these curves lie in the disk of radius s_0 centered at the point A (or B). But by the definition of complete metric, each bounded infinite set in our manifold has accumulation points; this makes it clear that every bounded and closed set is compact. Hence the indicated disk is compact. Applying Theorem 2 to this disk, we see that there is a shortest arc of A and B in this disk. The length of this shortest arc does not exceed the lower limit of the lengths of L_n, i.e., it is not greater than $\rho(AB)$. But any curve connecting the points A and B cannot have length less than $\rho(AB)$. Consequently, the length of our curve is equal to $\rho(AB)$, and, thus, this curve is a shortest arc not only in the chosen disk, but in the whole manifold. This proves the theorem.

Further, Theorem 2 implies the following important theorem.

Theorem 4. *Each neighborhood U of a manifold with intrinsic metric includes a neighborhood V such that there is a shortest arc of every two points of V lying entirely in U.*

Proof. Let O be a given point in a manifold with intrinsic metric $\rho(XY)$, and let U be a given neighborhood of this point. The point O has a neighborhood homeomorphic to a disk, and, therefore, we can take a neighborhood U' of the point O which is contained in U and is homeomorphic to a disk[4] and such that the closure $\overline{U'}$ of this neighborhood, which is homeomorphic to a disk with its boundary, also lies in U.

Let r_0 be the greatest lower bound of distances from the point O to points that lie outside $\overline{U'}$. It is clear from the definition of our neighborhood that $r_0 > 0$. Take a geodesic disk V of radius

$$r = \frac{r_0 - \varepsilon}{2} \tag{1}$$

centered at O, where ε is any number between zero and r_0. If two points X and Y lie in the disk V, then $\rho(OX), \rho(OY) \leq r$, and using the triangle inequality and (1), we have

$$\rho(XY) \leq \rho(OX) + \rho(OY) < 2r = r_0 - \varepsilon. \tag{2}$$

By the definition of intrinsic metric, there exists a curve L joining X and Y whose length $s(L)$ differs from the distance between these points by less than ε, i.e.,

$$s(L) < \rho(XY) + \varepsilon. \tag{3}$$

[4] The intersection of two neighborhoods is again a neighborhood. Therefore the intersection of U with a neighborhood homeomorphic to a disk is a neighborhood; at the same time, this intersection obviously contains a neighborhood homeomorphic to a disk.

© 2006 by Taylor & Francis Group, LLC

If a point Z lies on L, then by the very definition of length,

$$s(L) \geq \rho(XZ) + \rho(ZY). \tag{4}$$

Comparing (4), (3), and (2), we obtain

$$\rho(XZ) + \rho(ZY) < r_0. \tag{5}$$

If the point Z lies outside $\overline{U'}$, then by the definition of r_0, we have

$$\rho(OZ) \geq r_0, \tag{6}$$

and since $\rho(OX) \leq r = (r_0 - \varepsilon)/2$, (6) and the triangle inequality implies

$$\rho(XZ) \geq \rho(OZ) - \rho(OX) \geq r_0 - r = \frac{r_0 + \varepsilon}{2}. \tag{7}$$

By the same reason,

$$\rho(YZ) \geq \frac{r_0 + \varepsilon}{2}. \tag{8}$$

Adding (7) and (8), we obtain

$$\rho(XZ) + \rho(YZ) \geq r_0 + \varepsilon;$$

which contradicts (5). Hence, no points of a curve L of length satisfying (3) can lie outside $\overline{U'}$.

Since our metric is intrinsic, there exists a sequence of curves L_n joining the points X and Y whose lengths converge to $\rho(XY)$; we can assume that all these curves satisfy inequality (3). In this case, we have proved that these curves lie in $\overline{U'}$. Since the set $\overline{U'}$ is homeomorphic to a closed disk, this set is compact, and, therefore, we can choose a convergent subsequence from the curves L_n (by Theorem 4 of Sec. 1). The limit curve L joins the points X and Y and has a length not greater than the limit of the lengths of the curves L_n (by Theorem 5 of Sec. 1), i.e., by the choice of L_n, the length of L does not exceed $\rho(XY)$. But since the distance is the greatest lower bound of the lengths of curves, the length of the curve L should be equal to $\rho(XY)$, i.e., the curve L is a shortest arc. This curve lies in $\overline{U'}$, and moreover, in the initial neighborhood U. Thus, the proof of the theorem is complete.

Among other things, Theorem 4 implies that if we have a continuous curve $X(t)$, then taking sufficiently dense points $X(t_i)$ on this curve, we can connect these points by shortest arcs, and then we obtain a broken line composed of shortest arcs, the "geodesic broken line" inscribed in the curve $X(t)$. The length of this broken line is equal to $\sum_{i=1}^{n} \rho(X(t_{i-1})X(t_i))$, since the length of each shortest arc is equal to the distance. Hence, we can say that the length of a curve is the least upper bound of the lengths of geodesic broken lines inscribed in this curve. Of course, we can prove in the usual way that this length is the limit of the lengths of these broken lines provided that the greatest diameter of arcs between the points $X(t_{i-1})$ and $X(t_i)$ tends to zero.

© 2006 by Taylor & Francis Group, LLC

Theorem 5. *Let a submanifold R be distinguished on a manifold R_0 with intrinsic metric $\rho_0(XY)$. Then a unique intrinsic metric $\rho(XY)$ is induced in R, which coincides with the metric $\rho_0(XY)$ in a sufficiently small neighborhood of an arbitrary point of R.*

First of all, we have to define what is meant by a metric "induced" in R. The meaning of this term is as follows. Let X and Y be two points in R. Take all curves connecting X and Y and lying in R. Define the lengths of these curves using the given metric ρ_0 in R_0. Then take the greatest upper bound of the lengths of these curves; this bound is taken as the distance $\rho(XY)$ in R between X and Y.

This definition is similar to the definition of the intrinsic metric on a surface. From this definition, we see that the intrinsic metric of a surface is the intrinsic metric induced by the metric of the ambient space of this surface.

Proof of Theorem 5. Let A and B be two points in R. Since R is a manifold by condition, there exists a continuous curve L in this manifold joining A and B. By Theorem 4, each point of the curve L can be surrounded by a neighborhood, whose every two points are connected by a shortest arc. Since R is an open set in R_0, its every point has a neighborhood lying entirely in R. Therefore, we can take neighborhoods of points of the curve L lying in R so small that shortest arcs of pairs of points of one neighborhood do not leave R. This is possible by Theorem 4. We can choose a finitely many neighborhoods from the set of these neighborhoods that cover the whole curve L. Connecting the points of the curve L lying in the same neighborhood by shortest arcs,[5] we obtain a new curve connecting the points A and B. Since this curve is composed of finitely many shortest arcs, this curve has length. Hence, we can connect the points A and B in R by a curve of finite length, and, therefore, there exists the greatest lower bound of the lengths of these curves. This bound is accepted as the distance $\rho(AB)$ between A and B in R.

It is easy to conclude from the definition of length that this metric enjoys all three main properties. The fact that this metric is intrinsic is proved by the argument in Sec. 6 of Chapter I. Namely, in the same way, we prove that the lengths of curves in R defined by the metrics ρ_0 and ρ are equal. Therefore, the distance $\rho(XY)$ in R turns out to be the greatest lower bound of the lengths of the curves as in the definition of length by the metric ρ in R instead of the initial metric ρ_0.

Given two points A and B in R, we have

$$\rho(AB) \geq \rho_0(AB).$$

Indeed, $\rho(AB)$ is the greatest lower bound of the lengths of the curves lying in R while $\rho_0(AB)$ is the same bound of the lengths of the curves lying in the whole R_0. The set of the latter curves is larger, and, therefore, the greatest lower bound of their lengths does not exceed the greatest lower bound of the lengths of those curves that lie in R only.

Now let O be an arbitrary point in R. By Theorem 4, we can take a neighborhood V in each neighborhood U of this point, whose every two points can be connected

[5] Let X_k and X_{k+1} be the centers of two successive neighborhoods U_k and U_{k+1} on the curve L. Let L be a common point of these neighborhoods. We join X_k with Y_k and Y_k with X_{k+1} by shortest arcs.

© 2006 by Taylor & Francis Group, LLC

by a shortest arc lying in U. We take a neighborhood U so that this neighborhood lies entirely in R, and a shortest arc is taken in the initial metric ρ_0. The length of the shortest arc AB of A and B is equal to $\rho_0(AB)$. Since the lengths of curves in the metrics ρ and ρ_0 coincide, the length $s(AB)$ of the shortest arc AB in the metric ρ is the same, i.e.,

$$s(AB) = \rho_0(AB).$$

The distance $\rho(AB)$ is the greatest lower bound of the lengths of the curves joining A and B, and, therefore,

$$\rho(AB) \leq s(AB) = \rho_0(AB).$$

But since we always have

$$\rho(AB) \geq \rho_0(AB),$$

we have, therefore, $\rho(AB) = \rho_0(AB)$, i.e., the metrics ρ and ρ_0 coincide in a small neighborhood of O. It is clear that there is only one metric in R with this property, since the length of a curve is defined by the distances between points of this curve which are arbitrarily close to each other. Thus, the proof of the theorem is complete.

The intrinsic metric exists on a complete convex surface, since every two points of this surface can be connected by a curve of finite length, namely, by convex curve obtained in the section of this surface by some plane. Now Theorem 5 implies that each convex surface also has intrinsic metric, and in a small neighborhood of each point this metric is the same both on the complete surface and a given domain containing this point. Thus, we arrive at the same conclusions, which were obtained in Chapter I in a different way.

Theorem 6. *If there is a shortest curve in a closed domain G of a manifold with intrinsic metric*[6] *whose boundary consists of a finitely many shortest arcs, this curve is either a geodesic or a geodesic broken line with vertices at the vertices of the domain G, i.e., this curve consists of a finitely many geodesics having common successive endpoints at some vertices of the domain G. Of course, the case is not excluded in which this curve never passes through the vertices of the domain.*

Proof. Let L be a shortest curve in the domain G whose boundary consists of a finite number of shortest arcs. These shortest arcs will be called the sides of G, and the points of junctions will be called the vertices of G.

Let O be a point on L lying in the interior of G. Then this point has a neighborhood U lying in G. By Theorem 4, we can take a neighborhood V of the point O which is contained in U such that its every two points can be connected by a shortest arc lying in U and, thus, in G. If a segment of the curve L lying in V is not a shortest arc, then by the property of the neighborhood V we can replace this arc by a shortest arc lying in G. Thus, we shorten the line L, which is impossible since this line is the shortest line in G. Consequently, each point of the curve L lying in the interior of the domain G is contained in a segment of the curve L which is a shortest arc.

[6] That is, a curve lying in G such that this curve is the shortest curve among all curves lying in G and connecting two given points.

© 2006 by Taylor & Francis Group, LLC

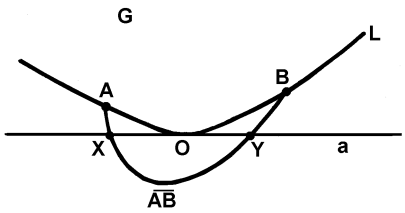


Fig. 14

Now let O be a point on the curve L lying in the interior of a side a of the domain G (Fig. 14). Take a neighborhood U about the point O so small that this neighborhood does not intersect any other sides of the domain G. In this neighborhood, we again take a neighborhood V such that its every two points can be connected by a shortest arc lying in U. Take two points A and B on the curve L lying in the neighborhood V such that the arc AB of L contains the point O. If the arc AB of L is not a shortest arc, we replace this arc by a shortest arc $\overline{AB}$. This shortest arc is in U but can not lie in G, and then this arc intersects the side A. For example, let X and Y be two subsequent points of the intersection of the shortest arc $\overline{AB}$ with the side a. Since both $\overline{AB}$ and a are shortest arcs, their segments $\overline{XY}$ and XY are also shortest arcs. Therefore, they have equal length, and replacing the segment $\overline{XY}$ of the line $\overline{AB}$ by the segment XY of the side a, we obtain a line of the same length. Performing such changes, we obtain a shortest arc of A and B in the domain G (in the interior or on the boundary of this domain). Since the line L cannot be shortened by condition, i.e., it is not possible to replace this line by a shorter line that also lies in G, this line should have the same length on the segment AB as this shortest arc. Hence, each point of the curve L lying in the interior of a side of the domain G belongs to the segment of this curve that is a shortest arc.

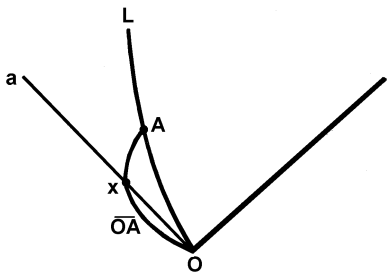


Fig. 15

Now let a point O of the curve L be a vertex of the domain G (Fig. 15). We take a neighborhood U of the point O, which does not intersect the sides of G, except for those that meet at the point O. In this neighborhood, we again take a neighborhood V, whose every two points can be connected by a shortest arc in U. Let A be a point on L lying in the neighborhood V. If the segment OA of the line L is not a shortest arc, then we again draw a shortest arc $\overline{OA}$. If $\overline{OA}$ is in G, we replace the segment OA of the line L by the shortest arc $\overline{OA}$. If the shortest arc $\overline{OA}$ is not in the domain G, this arc intersects its sides, and these are obviously only the sides that meet at O. For example, let X be intersection point of the line $\overline{OA}$ with the side a. Since a is a shortest arc approaching the point O, the segment $\overline{OX}$ of the line $\overline{OA}$ can be replaced by the segment OX of the side a. Since the line L cannot be shortened, the segment of this line between O and X cannot be shorter than the segment OX of the side a. Hence L is a shortest arc on the segment OX.

This implies that if a point O of the line L is a vertex of the domain G, each branch of L emanating from O is a shortest arc on each sufficiently small segment.

It follows from what we have said above that if a line L never passes through the vertices, its every point belongs to an arc that is a shortest arc, i.e., L is a geodesic. If L passes through some vertices, then this line is a geodesic between two vertices

© 2006 by Taylor & Francis Group, LLC

lying consecutively on this curve; since there is a finite number of vertices, L is a geodesic broken line. This completes the proof.

3. The Nonoverlapping Condition for Shortest Arcs

Let us introduce the following requirement, which will be called the *nonoverlapping condition for shortest arcs. If two shortest arcs AB and AC connecting a point A with two points B and C coincide on some segment AD, one of them is included in the other.* Roughly speaking, two essentially different shortest arcs emanating from the same point can overlap on no segment; this formulation explains the meaning of the name of this condition.

For illustration purposes, we consider a cone K with the complete angle at the vertex greater than 2π. Draw three shortest arcs on this cone from its vertex O, i.e., three segments OA, OB, and OC in such a way that the angles that OA makes with OB and OC are equal to π. In this case, it is easily seen that the lines $AO + OB$ and $AO + OC$ are shortest arcs connecting the point A with the points B and C. These shortest arcs coincide on the segment AO and then diverge. Hence, the nonoverlapping condition for shortest arcs does not hold on a cone whose complete angle at the vertex is greater than 2π. This implies that this condition holds for every polyhedral metric of positive curvature. However, this assertion will be proved in Sec. 2 of Chapter III; we present this condition here for illustration purposes only. In exactly the same way, we will prove in Sec. 4 of Chapter III that the nonoverlapping condition for a shortest arc holds on every convex surface. Thus, in fact, only intrinsic metric manifolds satisfying this condition are of interest to us. Also, we note that this condition always holds for each regular surface, since shortest arcs on such a surface are defined by differential equations that assume a unique solution for given initial data. Hence, the nonoverlapping condition for shortest arcs is of a sufficiently general significance; at the same time, its content is very simple. Therefore, it seems natural to formulate this condition and to state consequences that are implied by this condition. We will use these consequences throughout the book. In subsequent chapters, it is assumed that we discuss shortest arcs on a manifold for which our condition holds.

Theorem. *Only the following possibilities are open for the mutual location of two shortest arcs (Fig. 11): (1) these shortest arcs have no common points; (2) they have only one common point; (3) they have only two common points, and then these points are their common endpoints; (4) one of these shortest arcs includes the other; (5) these shortest arcs coincide on some segment; one endpoint of this segment is an endpoint of one of these arcs, and the other endpoint of this segment is an endpoint of the other arc; these segments contain all common points of these arcs (Fig. 16).*

Proof. Obviously, we must prove that if two shortest arcs AB and CD have two common points, at least one of which is not their common endpoint, one of the two latter possibilities holds. Therefore, assume that the shortest arcs AB and CD have two common points X and Y, and, moreover, X is not the endpoint of the shortest arc AB. For definiteness, we shall assume that the point X belongs to the segment AY of the shortest arc AB and to the segment $\overline{CY}$ of the shortest arc CD (Fig. 17). The segments of the shortest arc CD will be marked by an overline.

© 2006 by Taylor & Francis Group, LLC

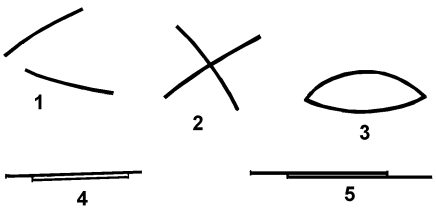


Fig. 16

The segment $\overline{XY}$ of the shortest arc CD is equal to the segment XY of the shortest arc AB, since both these segments are shortest arcs between the points X and Y. Therefore, the line $AX + \overline{XY}$ is a shortest arc between the points A and Y. Thus, there are two shortest arcs AY and $AX + \overline{XY}$ between these points. They coincide on the segment AX, and since, in addition, they have common endpoints, they should entirely coincide according to the nonoverlapping condition for shortest arcs. Hence our shortest arcs AB and CD coincide on the segment XY.

Now consider the segments AY and $\overline{CY}$ of our shortest arcs AB and CD. By condition, they contain the point X and hence coincide on the segment XY. In this case, by the nonoverlapping condition, one of them is contained in the other. For exactly the same reason, one of the segments BX and $\overline{DX}$ of the shortest arcs AB and CD is contained in the other. Hence, the following two cases are possible: (1) $\overline{CY}$ is included in AY and $\overline{DX}$ is included in BX; then it is obvious that CD is included in AB, or otherwise, AY and BX are included in $\overline{CY}$ and $\overline{DX}$, and then AB is contained in CD; (2) $\overline{CY}$ is contained in AY and BX is contained in $\overline{DX}$, and then AB and CD coincide on the segment CB, or otherwise, AY is contained in $\overline{CY}$ and $\overline{DX}$ is contained in BX; then AB and CD coincide on the segment AD.

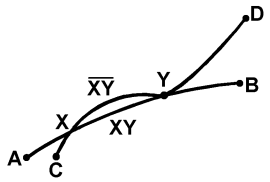


Fig. 17

It remains to prove that if, e.g., AB and CD coincide on the segment BC, then this segment contains all their common points. However, assume that they contain one more common point Z that does not belong to the segment BC (Fig. 18). Then we obtain two shortest arcs AB and $AZ + \overline{ZB}$, which coincide on the segment AZ. Since they have common endpoints, they should coincide by the nonoverlapping condition. This leads us to the fact that the shortest arc CD should overlap itself on the segment $\overline{CB}$; of course, this is not possible. Hence, the segment CB contains all common points of the shortest arcs AB and CD. The proof of the theorem is complete.

This theorem shows that under the nonoverlapping condition, two shortest arcs have the same behavior as arcs of the great circles on a sphere in the sense of their mutual location.

Let us list some obvious consequences of this theorem.

1. If at least one of the two points X and Y of a shortest arc L is an interior point, the segment XY of this shortest arc is a unique shortest arc of X and Y.

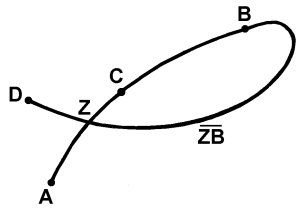


Fig. 18

© 2006 by Taylor & Francis Group, LLC

2. If three points A, B, and C do not lie on one shortest arc, the shortest arcs AB, AC, and BC have no common points, except for the endpoint; therefore, these shortest arcs bound a domain that is homeomorphic to a disk in a domain homeomorphic to a disk, i.e., an ordinary triangle.

3. A finite set of shortest arcs can be divided into a finite set of segments having no common points, except for their endpoints. For two shortest arcs, this assertion follows directly from the theorem. For any number of shortest arcs, this assertion is proved on using the same theorem by obvious induction. This conclusion is especially important for the study of intersections of polygons; this assertion implies that the intersection of two or finitely many polygons forms only finitely many polygons.

4. *Let C be an interior point of a shortest arc AB, and let points A_n and C_n converge to A and C, respectively. Then the shortest arcs A_nC_n (if they exist) can converge only to the segment AC of the shortest arc AB.*
Indeed, the limit AC of the shortest arcs A_nC_n (if this limit exists) is a shortest arc; if this limit does not coincide with AC, we have two different shortest arcs connecting the points A and C; this contradicts the first of the indicated consequences of the theorem.

Note that it is not difficult to prove the following stronger assertion:

Let C be an interior point of a shortest arc AB, and let points A_n and C_n converge to A and C, respectively. Then at least for large n there exist shortest arcs A_nC_n and these shortest arcs converge to the segment AC of the shortest arc AB.

(Without the nonoverlapping condition, the existence, as well as the convergence, of the shortest arcs A_nC_n are not ensured in a manifold with incomplete metric.) We skip the proof of this theorem, since we do not use it in the subsequent chapters.

The above results considerably simplify the arguments with shortest arcs on manifolds satisfying the nonoverlapping condition. Because of simplicity of these results, we will often use them without special references. In the next sections, we deduce deeper consequences of the nonoverlapping conditions for shortest arcs, which will be needed first in Chapter V, and then in Chapters VII and IX. Therefore, the reader who wishes to move on to the specific study of convex surface can skip the remaining three sections of this chapter.

4. A Convex Neighborhood

A set of points G in an intrinsic metric manifold is called convex if there is a shortest arc of every two points of G lying entirely in G. Let G be an open connected set in a manifold R_0 with intrinsic metric ρ_0. By Theorem 5 of Sec. 2, the metric ρ_0 induces a new intrinsic metric ρ in G; this metric $\rho(XY)$ is the greatest lower bound of the lengths of curves lying in G and joining two points X and Y; moreover, the length is measured in the metric ρ_0. Then for any X and Y in G, we have $\rho(XY) \geq \rho_0(XY)$. If G is convex, then the points X and Y can be connected in G by a line that is a shortest arc in the sense of the metric ρ_0; the length of this line is equal to $\rho(XY)$.

© 2006 by Taylor & Francis Group, LLC

But by definition, $\rho(XY)$ is not greater than the length of the line connecting the points X and Y in G, and thus $\rho(XY) \leq \rho_0(XY)$. Hence $\rho(XY) = \rho_0(XY)$, i.e., the metric induced in a convex set coincides with the initial one; cutting out a convex set from a manifold, we do not change the distance between the points of this set.

In this section, we shall prove the following theorem.

If the nonoverlapping condition for shortest arcs holds in an intrinsic metric manifold, then each point of this manifold has an arbitrary small neighborhood which is a convex polygon homeomorphic to a disk.

The significance of this theorem is clear from the property of convex sets just revealed. Restricting ourselves to a convex neighborhood U of a point O, we do not change the metric in this neighborhood if we exclude from consideration the other part of the manifold. Moreover, each point in U can be connected by a shortest arc. Therefore, in all arguments "in the small," it is especially appropriate to use a convex neighborhood.

The proof of the above theorem rests first of all on the following general lemma. We will say that a closed curve L surrounds some domain M in a domain U if it is not possible to contract this curve to a point in the domain U without intersecting the domain M.[7]

Lemma 1. *Assume that we have two domains U and V homeomorphic to a circle in an intrinsic metric manifold, and, moreover, V lies in U and has the property that there is a shortest arc of its every two points lying entirely in the domain U. Suppose that we have a connected domain M in V and a curve L surrounding M in V and shortest among all curves lying in U and surrounding M in U. Then the closed domain bounded by the curve L is convex and homeomorphic to a disk. Moreover, if the domain M is a geodesic polygon, then L is a geodesic broken line, and hence the domain bounded by this line is a polygon. The vertices of the broken line L can lie only at the vertices of the polygon M. (In particular, L has no vertices at all whenever this curve is a closed geodesic (Fig. 19).)*

Proof. Suppose that a curve L satisfying the conditions of the lemma has a double point A. Then the curve L can be decomposed into two closed curves L_1 and L_2. If both L_1 and L_2 do not surround the domain M, we can contract them to A not intersecting N and keeping A fixed,[8] and thus contract the whole curve to A. But since L surrounds M, this is impossible and, therefore, one of the curves L_1 and L_2 surrounds M. But, in this case, the other curve can be excluded;

[7] A closed curve is a continuous image of circle. Two closed curves $X(t)$ and $Y(t)$ are considered equal if there exists a monotone mapping of this circle t onto the circle s: $s = s(t)$ such that $X(t) = Y(s(t))$. Hence the notion of a closed curve differs from the notion of a curve with coinciding endpoints by the fact that we fix a point on the latter one which is its coinciding endpoints. One says that a closed curve $X_0(t)$ is deformable to a closed curve $X_1(t)$ in a domain U if there is a function $X(t,s)$ $(0 \leq s \leq 1)$ such that (1) $X(t,s) \in U$ for all t and s; (2) $X(t,s)$ forms a closed curve for any s; (3) $X(t,s)$ depends continuously on y and s; (4) $X(t,0) = X_0(t)$ and $X(t,1) = X_1(t)$. If X_i is reduced to a singleton, then we say that the curve $X_0(t)$ contracts to a point.

[8] It is easy to prove the well-known fact that if a curve can be contracted to a point, then we may contract this curve to its every point, keeping this point fixed.

© 2006 by Taylor & Francis Group, LLC

moreover, we obtain a curve that is shorter than L,[9] which is impossible, since L is the shortest curve among all curves that surround M. Consequently, L has no double points and so this curve bounds a domain homeomorphic to a circle.

Let now A and B be two points in the domain bounded by the curve L. By condition, there exists a shortest arc AB, since L lies in V and V is homeomorphic to a disk, and thus the whole domain bounded by L lies in V. If the shortest arc AB does not enter the domain bounded by L, then this arc leaves this domain and then returns to it again, intersecting L at two points C and D such that the segment CD lies outside of L. Replacing one of the arcs $\overset{\frown}{CD}$ of the curve L by the segment CD of the shortest arc AB, we then obtain a curve that again surrounds M.[10] Doing this, we will not obtain a curve that is shorter than L. Therefore, the arc $\overset{\frown}{CD}$ itself should be a shortest arc. Applying the same argument to all segments of the shortest arc AB lying outside of the domain bounded by the curve L, we arrive at the fact that the points A and B can be connected by a shortest arc lying in the indicated domain or on the boundary of this domain. This proves the convexity of the domain under study.

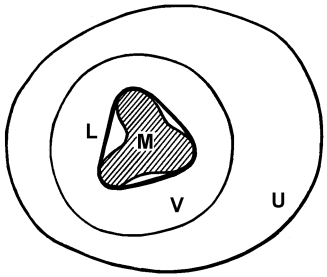


Fig. 19

It remains to prove that if the domain M is a geodesic polygon then the curve L is a geodesic broken line. We proved in Sec. 2 (Theorem 6) that a line that is the shortest line in a geodesic polygon is exactly a geodesic broken line with vertices at the vertices of the polygon. Apply this theorem to the domain N, which is the exterior domain for M. Take two points A and B on L that divide L into two parts of equal length. If both segments AB of the curve L are not shortest arcs in N, then replacing one of these segments by a shorter line, we obtain a curve that also surrounds M but is shorter than L. But L is the shortest line among all lines surrounding M. Therefore, its segments AB are shortest arcs in N and, by the indicated lemma, they are geodesic broken lines. Hence L is also a geodesic broken line. As follows from Theorem 6 of Sec. 2, the vertices of this line can lie only at the vertices of the polygon M.

From this lemma, it is easy to obtain a result that is in essence more general than the theorem formulated at the beginning of this section; which then becomes a consequence of this result in virtue of some lemma to be proved below.

[9] The only exception is the case where one of the curves L_1 and L_2, which does not surround M, degenerates into a point. However, this case can be excluded if we take the arc length as the parameter on L; this is possible, since L has length by condition.

[10] The segment CD and each of the arcs $\overset{\frown}{CD}$ form closed curves without multiple points. Therefore, they bound domains homeomorphic to a circle. One of these domains contains the whole domain bounded by the curve L. The shortest arc AB can intersect L at many points, and we take any pair of points neighboring on L at which AB leaves the domain bounded by L as C and D.

© 2006 by Taylor & Francis Group, LLC

Theorem 1. *Assume that a point O in a manifold R with intrinsic metric ρ can be surrounded by an arbitrarily small polygon of an arbitrarily small perimeter.*[11] *Then the point O can be surrounded by the same convex polygon, i.e., by an arbitrarily small polygon of an arbitrarily small perimeter.*

Proof. Take a neighborhood U around the point O which is homeomorphic to a disk, and let V be a neighborhood in U which is also homeomorphic to a disk such that its every two points can be connected by a shortest arc lying in U (see Theorem 4 in Sec. 2). Let l be the greatest lower bound of the lengths of the curves that lie in U, have at least one point outside the neighborhood V or on its boundary, and surround the point O, i.e., the curves that cannot be contracted into a point without intersection of O when remaining in U. (In particular, these curves can pass through O.) This greatest lower bound exists and is positive.

Indeed, there exists a point A on the boundary of V which is nearest to O. A shortest arc AO does exist by the choice of the neighborhood V and, obviously, lies in U. If we view this shortest arc as doubly traversed then this arc transforms into a closed curve surrounding the point O and having the point A on the boundary of V. Hence the considered curves exist, and the greatest lower bound l of their lengths does not exceed the doubled length of the shortest arc AO. Assume, however, that $l = 0$. This means that there exists a sequence of points $X_1, X_2, \ldots$ lying in U outside V or on the boundary of V such that a curve L_n passes through each point X_n which surrounds the point O and has length l_n; moreover, $l_n \to 0$ as $n \to \infty$ (Fig. 20). Obviously, we can assume that the points X_n converge to some point X. Each point X_n has a neighborhood W_n homeomorphic to a disk that does not contain the point O. Let r_n be the "radius" of such a neighborhood, i.e., the greatest lower bound of the distances from X_n to the points then outside W_n. It is clear from the very definition of the neighborhood that $r_n > 0$. If a curve passing through the point X_n has length $< r_n$, then this curve lies in the neighborhood W_n, as is clear from the definition of r_n. Since W_n is homeomorphic to a disk and does not contain the point O, such a curve can be contracted to the point X_n without intersection with the point O, i.e., such a curve of length $< r_n$ cannot surround O. Hence the curves L_n should have length $\geq r_n$, i.e., $l_n \geq r_n$.

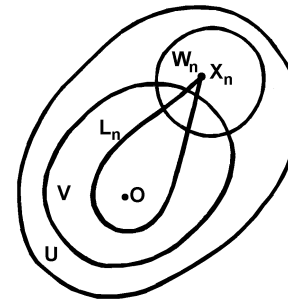


Fig. 20

Now let r be the "radius" of a neighborhood W of the point X, which is a limit point of all points X_n; herewith, assume that W does not contain O and is homeomorphic to a disk. Let m be so large that for $n > m$, the distance from X_n

[11] The fact that the polygon P surrounds O means that this polygon is a neighborhood of O. In an intrinsic metric manifold, any point can be surrounded by an arbitrarily small polygon. To this end, it is sufficient to circumscribe a polygon into a certain closed curve that bounds a small neighborhood of the point O. The possibility of circumscribing a broken line into any closed curve was proved in Sec. 2. But the fact that the polygon P is small does not imply that its perimeter is small. For example, consider the metric defined in a disk of radius 1 by the line element $ds^2 = dr^2 + d\varphi^2$. The center of such a disk cannot be surrounded by a polygon of perimeter less than 2π. This metric is obtained when one identifies all points of the base of a cylinder.

© 2006 by Taylor & Francis Group, LLC

to X is less than $r/2$. Then we can take neighborhoods W_n of "radius" $r_n = r/2$ for X_n when $n > m$. Thus, we obtain the neighborhoods W_n such that r_n does not tend to zero. But $l_n \geq r_n$, and, therefore, l_n does not tend to zero either. Hence our assumption that $l = 0$ is false; i.e., $l > 0$.

We now take a geodesic polygon P containing the point O in its interior which lies in the neighborhood V and has perimeter less than l. Such a polygon exists by the condition of the theorem. Consider all curves lying in U that surround this polygon and have length $< l$. The boundary of the polyhedron P itself is obviously such a curve, and hence these curves exist. Since the neighborhood U is homeomorphic to a closed disk, U is compact, and by Theorem 4 of Sec. 1, we can choose a convergent sequence from all the curves considered. By Theorem 5 of Sec. 1, this can be done so that the length of the limit curve L is equal to the greatest lower bound of the lengths of all curves under consideration. Hence the curve L is the shortest curve among those that surround the polygon P, and the length of this curve is less than l. (It is easy to prove that the limit curve also surrounds P.)

Since the polygon P contains the point O, the curve L also surrounds the point O. The number l is the greatest lower bound of the lengths of the curves that have points outside V and surround O. Since the length of L is less than l, the curve L has no points outside V, i.e., this curve lies entirely in V. Hence we have the neighborhood U homeomorphic to a disk and the other neighborhood V in U, which is also homeomorphic to a disk and is such that its every two points can be connected by a shortest arc lying in U. Further, we have the geodesic polygon P in V and the curve L surrounding this polygon, which lies in V and is the shortest curve among those surrounding this polygon and lying in U. Hence we are in the conditions of Lemma 1, and applying this lemma, we see that the curve L bounds a convex geodesic polygon. Since the point O lies in the interior of this polygon, this polygon is a neighborhood of this point. The theorem is proved.

We now prove the following lemma, which, together with Theorem 1, immediately leads to the theorem formulated at the beginning of this section.

Lemma 2. *If the nonoverlapping condition for shortest arcs holds in a manifold R with an intrinsic metric, then each point of R has an arbitrarily small neighborhood that is a triangle.*

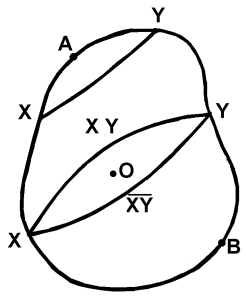


Fig. 21

Proof. Let O be a given point of the manifold R. This point O has a neighborhood U homeomorphic to a disk. Therefore, the point O can be surrounded by a simple closed curve L such that this curve bounds an arbitrary small neighborhood around O. Take two points A and B on the curve L. The curve L is divided into two arcs AB. We represent each of these arcs in the parametric form: $X_t = X(t), Y = Y(t)$ $(0 \leq t \leq 1)$; when t increases from 0 to 1, the points X_t and Y_t run over their own arc AB from point A to point B (Fig. 21). If the curve L passes sufficiently close to the point O, for any t there exists a shortest arc X_tY_t lying in the neighborhood U. Consider the following two possibilities for these shortest arcs separately.

© 2006 by Taylor & Francis Group, LLC

1. There exists a shortest arc X_tY_t passing through the point O.

2. There is no such shortest arc.[12]

We first assume that the second possibility takes place. Let M_t be the closed curve formed by the shortest arc X_tY_t and the arcs AX_t and AY_t of the curve L. When t is small, the points X_t and Y_t lie sufficiently close to the point A, and the curve M_t can be contracted to a point without intersection with O. Let T be the greatest lower bound of those t for which this is possible. If t is close to 1, the points X_t and Y_t lie near the point B, and then the line M_t cannot be contracted into a point without intersection with O (and, of course, remaining in the neighborhood U). Hence $T > 0$ and < 1, and, therefore, the points X_T and Y_T differ one from the other.

Take a sequence of values $t_n < T$ converging to T such that the curves M_{t_n} can be contracted to a point without intersection with O. We can choose a convergent sequence from the shortest arcs $X_{t_n}Y_{t_n}$, and the limit of this sequence is some shortest arc X_TY_T. The curve M_T we have is composed of this arc and the arcs AX_T and AY_T of the curve L. Since all the curves M_{t_n} can be contracted into the point A without intersection with O, the limit curve M_T cannot be contracted into a point without intersection with O if and only if this curve itself passes through the point O. However, this is not possible, since the shortest arc X_TY_T cannot pass through the point O by assumption. Consequently, the curve M_T can be contracted to a point without intersection with O.

On the other hand, we take a sequence of values $t'_n > T$ converging to T. Choose a convergent sequence from the shortest arcs $X_{t'_n}Y_{t'_n}$; the limit of this sequence is some shortest arc $\overline{X_TY_T}$ connecting the points X_T and Y_T. The curve $\overline{M_T}$ we have is composed of this shortest arc and the arcs AX_T and AY_T of the curve L. By the definition of the number T, none of the curve $M_{t'_n}$ can be contracted into a point without intersection with O. Therefore, the limit curve $\overline{M_T}$ also cannot be contracted to a point without intersection with O.

The curves M_T and $\overline{M_T}$ are composed of the same arcs of the curve L and the shortest arcs X_TY_T and $\overline{X_TY_T}$. Since M_T can be contracted to a point without intersection with O, and this is not so for $\overline{M_T}$, these shortest arcs are different. They form the digon D with vertices X_T and Y_T, and since none of the indicated shortest arcs can pass through the point O by condition, the digon D contains the point O in its interior. Indeed, otherwise we can deform the shortest arc $\overline{X_TY + T}$ into X_TY_T without intersection with O; thus, the curve $\overline{M_T}$ is deformable to M_T and can be contracted to the point A without intersection with O; this is not possible as we have shown.

Thus, we have obtained the digon D containing the point O in its interior. Taking any point Z on one of the sides of this digon, we transform this digon into the triangle $X_TY_TZ_T$. Hence there exists a triangle containing the point O in its interior.

[12]This second possibility takes place, e.g., in the case where the point O is the vertex of a convex cone.

© 2006 by Taylor & Francis Group, LLC

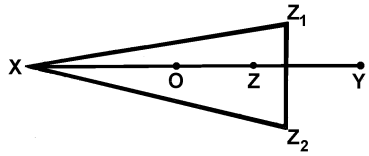


Fig. 22

It remains to study the case in which there are shortest arcs passing through the point O among the shortest arcs X_tY_t. Let XY be one of these shortest arc (Fig. 22). Take a point Z on this shortest arc which lies between the points O and Y. The shortest arc XY is homeomorphic to a straight line segment, and, therefore, this shortest arc divides all points that do not lie on this arc and are close to the point Z into the following two classes: the points of one class lie to one side of the shortest arc XY, and the other points lie on the other side.[13] Let Z_1 and Z_2 be two points that are close to the point Z and lie on the different sides of the shortest arc XY. Since the nonoverlapping condition for shortest arcs holds in our manifold, the shortest arcs XZ_1 and XZ_2 for Z_1 and Z_2 converging to Z converge to the segment XZ of the shortest arc XY. Hence, for Z_1 and Z_2 sufficiently close to Z, these shortest arcs belong to the opposite sides of the shortest arc XY, and, together with the shortest arc Z_1Z_2, they bound the triangle, which contains the point O in its interior. The lemma is proved.

We now easily prove the theorem that is formulated at the beginning of this section.

Theorem 2. *If the nonoverlapping conditions for shortest arcs hold in a manifold with an intrinsic metric, then each point of such a manifold has an arbitrarily small neighborhood that is a convex geodesic triangle.*

Proof. Let O be a given point of a manifold that satisfies the condition of the theorem. By Lemma 2, the point O can be included into the interior of an arbitrarily small triangle ABC. By the "triangle inequality," the length of each side of this triangle does not exceed the sum of distances from the ends of this side to the point O; e.g., $\rho(AB) \le \rho(AO) + \rho(BO)$. Hence, if a triangle ABC is sufficiently small then its perimeter is also small. Then by Theorem 1, there exists a sufficiently small convex polygon that contains this triangle. The vertices of this polygon P can lie only at the vertices of the triangle ABC as follows from Lemma 1; therefore, the polygon P can have at most three vertices. We can take, if need be, some of its boundary points as vertices and, thus, transform this triangle into a geodesic triangle. Of course, the sides of such a triangle may not be shortest arcs *a priori; therefore, according to our terminology, this triangle is called a "geodesic triangle." Each geodesic can be divided into finitely many shortest arcs; considering each such shortest arc as a side, we transform the geodesic triangle into a polygon whose sides are shortest arcs.*

[13] The shortest arc XY, together with some of its neighborhoods, can be homeomorphically mapped onto the plane so that this arc passes to a line segment. Then the notions "on one and on the other side of the shortest arc XY" have a clear meaning for the points sufficiently close to interior points of this shortest arc.

© 2006 by Taylor & Francis Group, LLC

5. General Properties of Convex Domains

Let R be a manifold with certain intrinsic metric satisfying the nonoverlapping condition for shortest arcs. By a convex domain, we mean a set of points of R satisfying the following conditions: (1) this set is convex; (2) this set is compact; (3) this set has interior points; (4) the boundary of this set consists of at most finitely many simple closed curves.

A convex polygon is a convex domain whose boundary consists of at most finitely many shortest arcs; in particular, this convex domain can be boundaryless, and then this domain is the whole manifold which thus turns out compact. Conversely, each compact manifold is a convex "polygon" since its every two points can be connected by a shortest arc by Theorem 1 of Sec. 2.

The properties of convex domains, which will be proved here, are completely analogous to the well-known properties of convex domains on the plane. However, it is useful to keep in mind that any convex set of not only an abstract manifold with intrinsic metric but also of a convex surface may fail to have such good properties in general. For example, the intersection of two convex domains can be nonconvex. Since a shortest arc is obviously a convex set, the intersection of two different shortest arcs connecting two given points consists of these two points and, therefore, is not convex. It was shown in Sec. 10 of Chapter I that there can be points on a convex surface such that there is no shortest arc of them. Therefore, removing these points from a convex set, we obtain again a convex set. Also, in Sec. 10 of Chapter I, we noted that the set of such points can be dense in a closed convex surface, and, removing these points, we obtain a convex set that differs by far from convex sets on the plane.

Theorem 1. *If two points A and B belong to a convex domain G, and, moreover, the point A lies in the interior of G, then every shortest arc AB lies entirely in the interior of the domain G (of course, except for a point B, if this point lies on the boundary of G).*

Proof. Let the point A lie in the interior of the convex domain G, and let the point B belong to the interior of G or its boundary. Assume that the points A and B can be connected by a shortest arc AB partially running outside of G. Take a point C in the interior of this arc so close to A that this point is also in the interior of G. Then by the nonoverlapping condition for shortest arcs, the segment CB of our shortest arc is a unique shortest arc of C and B. But this segment does not lie entirely in G, while the points C and B belong to G, i.e., we obtain a contradiction to the convexity of the domain G. Therefore, every shortest arc AB lies in G.

Assume now that there is a point D inside the shortest arc AB that lies on the boundary of the domain G. Then D lies on one of the curves that comprise the boundary of G by assumption (there are finitely many of these curves); let L be such a curve. Since this shortest arc is homeomorphic to a line segment, for lucidity, we map this shortest arc, together with one of its neighborhoods, into the plane so that this shortest arc passes to a line segment (Fig. 23). The curve L is also mapped into the plane at least in some neighborhood of D.

© 2006 by Taylor & Francis Group, LLC

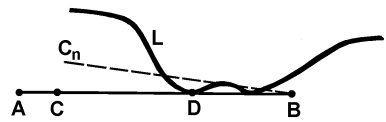


Fig. 23

Take a point C inside the shortest arc AB of A and D so this point lies in the interior of the domain G. Then the segment CB of our shortest arc is a unique shortest arc of C and B, and if a sequence of points C_n converges to C then the shortest arcs C_nB converge to CB (by Corollary 4 of Theorem in Sec. 3). If the points C_n do not lie on AB but are closer to C than A, the shortest arcs C_nB cannot intersect AB. In fact, if C_nB and AB intersect each other, then they have common interior points. By the consequences deduced from the nonoverlapping condition for shortest arcs in Sec. 3, we conclude that the latter should overlap, which is impossible if C_n does not lie on AB and, at the same time, is closer to C than the point A. Hence, whenever the point C_n is sufficiently close to C, the shortest arc C_nB passes near AB and does not intersect AB, i.e., lies on one side of AB. Therefore, if we take a point C_n lying on the same side of AB as L, C_nB intersects L. But this is impossible, since L divides a neighborhood of the shortest arc AB into two parts, one of which does not belong to G while C_nB should pass in G, since C_n lies in the interior of G. Hence there are no points on AB but, possibly, B that lie on the boundary of G.

Theorem 1 implies obviously that the interior of every convex domain is convex and each shortest arc of two points of the interior lies in the latter.

Theorem 2. *If the intersection of two convex domains has interior points then it is convex.*

Proof. Let G_1 and G_2 be two convex domains, and let G be their common part. Let a point A lie in the interior of G, and let a point B lie in the interior or the boundary of the domain G. The point A lies in the interior of G_1, as well as in the interior of G_2, and, therefore, by Theorem 1, every shortest arc AB passes in the interior of G_1, as well as in the interior of G_2, i.e., this arc passes in the interior of G. This implies that each point B in G is a limit of some interior points, namely, of the points lying on the shortest arc AB.

Now let A and B be two points in G. Assume that points A_n and B_n lying in the interior of G converge to A and B, respectively. As follows from what we have just proved, the shortest arcs A_nB_n belongs to G. Since G_1 and G_2 are compact, G is also compact, and hence there is a limit of shortest arcs A_nB_n in G. (More precisely, it is possible to choose a sequence from these shortest arcs which converges to a curve lying in G.) By Theorems 1 and 2, this limit is a shortest arc of A and B.

Theorem 3. *If the intersection of two polygons has interior points, then this intersection is a convex polygon.*

Proof. A convex polygon is a set characterized by the following properties: (1) this set is convex; (2) it has interior points; (3) it is compact; and (4) its boundary consists of at most finitely many shortest arcs. If the intersection P of two convex polygons P_1 and P_2 has interior points, this intersection has the second property. Then, the convexity of this intersection follows from Theorem 2. The compactness is obvious. Finally, the boundary of P consists of a part of the boundary of P_1

© 2006 by Taylor & Francis Group, LLC

lying in P_2 and a part of the boundary of P_2 lying in P_1. The boundaries of P_1 and P_2 consist of finitely many shortest arcs. The nonoverlapping condition for shortest arcs implies that two shortest arcs either overlap one another or have no more than two common points. Therefore, the number of segments of the shortest arcs bounding P_1 (or P_2) and lying in P_2 (or in P_1) is finite. Hence P also has the fourth property.

Theorem 4. *If a shortest arc in a convex domain divides this domain, then each of the parts of this domain is convex; moreover, this is so without the overlapping condition for shortest arcs.*

Proof. Let H be one of the parts into which the domain G is divided by a shortest arc L. We include L itself into H. Let A and B be two points in H, and let AB be a shortest arc of these points lying in G. Then, if AB has points outside H, then this arc should pass from H to $G - H$ (i.e., to the complement of H) and should return to H. But since G and $G - H$ are separated by the shortest arc L, AB intersects L at least twice. Let X and Y be the most distant point from the other points common for AB and L. Since L is a shortest arc, we again obtain a shortest arc of A and B when replacing the segment XY of AB by the same segment of L. Since L is included into H, the obtained shortest arc lies in H. Hence H is convex.

Theorem 4a. *If a shortest arc passing in a convex polygon divides this polygon, then each part of this polygon is a convex polygon.*

Proof. This follows from Theorem 4 and the arguments used in the proof of Theorem 3 in an obvious way.

Theorem 5. *The boundary of a convex polygon* (*and of every convex domain*) *has no multiple points.*

Proof. If more than two sides meet at a vertex O of a polyhedron P then at least two angles of the polygon P are contiguous to O. Let the sides a, b, and c meet at O; moreover, assume that a and b belong to one angle, while c belongs to the other one (Fig. 24). Take points A, B, and C on the sides a, b, and c, respectively. If these points are sufficiently close to O, then the shortest arcs AC and BC passing from one angle to the other inevitably transverse the point O. Hence these shortest arcs have not only the common end C but also the common interior point O, and by Theorem 1 of Sec. 2, they overlap. But since the points A and O, B and O lie on the shortest arcs a and b, respectively, we have that AC goes along a and BC along b by the same reason. Hence the sides a and b overlap, which is impossible. This completes the proof of theorem.

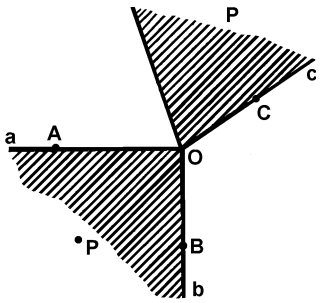


Fig. 24

© 2006 by Taylor & Francis Group, LLC

6. Triangulation

In this section, we prove that each polygon in a manifold with intrinsic metric satisfying the nonoverlapping condition for shortest arcs can be partitioned into arbitrarily small triangles. To "partition a polygon P into arbitrarily small triangles" means to represent this triangle P as the sum of finitely many triangles that have no common interior points and whose diameters are less than an arbitrarily given positive number. Recall that a polygon is a compact set with interior points whose boundary consists of at most finitely many shortest arcs, which are called the sides of this polygon; in particular, a "polygon" can have no boundary at all and then this polygon is the whole manifold which is compact in this case. Hence, the theorem mentioned also provides the possibility of dividing each of these manifolds into arbitrarily small triangles.

This theorem plays a fundamental role in our theory, since the method for approximating a given metric by polyhedral metrics, which was generally described in Sec. 5 of Chapter I, is based on this theorem. The definition of area given in Sec. 7 of Chapter I is also based on this theorem. Unfortunately, we have no sufficiently simple proof of this important theorem, although this theorem looks sufficiently obvious. In our proof, we will use the results of Secs. 3–5 and a number of lemmas, which now will be proved sequentially. In what follows, we bear in mind without further speculation that all our arguments are applied to some manifold with intrinsic metric satisfying the nonoverlapping condition for shortest arcs.

Lemma 1. *A convex polygon homeomorphic to a circle can be partitioned into convex triangles such that in each of these triangles, the sum of every two sides is strictly greater than the third.*

Proof. Let $A_1, A_2, \dots, A_n$ be vertices of a given convex polygon Q homeomorphic to a disk, which are numbered in the order of their location on the boundary of this polygon. If two neighboring sides, e.g., the sides A_1A_2 and A_2A_3 form together one shortest arc, then we exclude the vertex A_2 and unite these sides into a single side. Hence we can assume *a priori* that no two neighboring sides of our polygon form one shortest arc.

Of course, the polygon Q has at least two vertices. If Q has exactly three vertices, this polygon is already a triangle. If Q has two vertices then, taking another arbitrary point on its boundary as one more vertex, we transform this polygon into a triangle. If Q has more than three vertices then its vertices A_1 and A_3 separated by the vertex A_2 are not neighboring. Since the polygon Q is convex, we can draw a shortest arc A_1A_3 in this polygon. By Theorem 1 of the preceding section, this shortest arc either runs in the interior of Q or lies on its boundary. However, the latter case is impossible, since by the condition of the choice of vertices A_i, the segment of the boundary of the polygon Q between two nonneighboring vertices is not a shortest arc. Hence A_1A_2 lies in the interior of Q, and, by Theorem 4a of Sec. 5, this arc divides Q into two convex polygons, one of which is the triangle $A_1A_2A_3$. If the other triangle has more than three vertices, then we do the same for this polygon, and so on.

© 2006 by Taylor & Francis Group, LLC

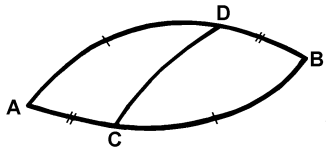


Fig. 25

Thus, we divide the polygon Q into convex triangles bounded by shortest arcs. But, in general, there can be triangles in this partition for which the sum of some two sides is equal to the third (each side cannot be greater than the sum of the other two, since this side is a shortest arc). In particular, if the polygon Q has only two vertices A_1 and A_2, then taking some point A_3 on one side of this triangle as the third vertex, we obtain a triangle for which the sum of the sides A_1A_3 and A_2A_3 is equal to the side A_1A_2. Let ABC be some triangle for which the sum of two sides is equal to the third, i.e., $AC + CB = AB$ (Fig. 25). Since AB is a shortest arc, $AC + CB$ is also a shortest arc. Take a point D on the side AB such that $AD = BC$, and hence $AC = BD$; then draw the shortest arc CD in the triangle ABC. The following two cases are possible: (1) either $CD < AD + AC = BD + BC$ or (2) $CD = AD + AC = BD + BC$.

In the first case, the shortest arc CD cannot lie entirely on the boundary of the triangle ABC, and hence, according to Theorem 1 of Sec. 5, this arc lies entirely in the interior of this triangle. Then this arc divides this triangle into two triangles ACD and BCD. In both these triangles, the sum of any two sides is greater than the third side. Indeed, (1) $CD < AD + AC$ by assumption; (2) if, say, $AC = AD + CD$, $AD + CD$ is a shortest arc emanating from the point A and having the point D as a common point with the shortest arc AB; but since D lies in the interior of AB, this is impossible by the nonoverlapping condition for shortest arcs; hence $AC < AD + CD$. In exactly the same way, we find that $AC + CD$, and that analogous inequalities hold for the triangle BCD.

We now consider the second case where $CD = AD + AC = BD + BC$. Since CD is a shortest arc, $DB + BC$ is also a shortest arc. Take a point E on the side AB, between the points D and B. Since AB is a shortest arc, $EC = EB + BC$ is also a shortest arc. Finally, AC is a shortest arc by condition. Therefore, taking the point E instead of the point B as the vertex, we obtain the triangle AEC with shorter sides, which coincides with the triangle ABC. In this triangle, the sum of any two sides is less than the third. For example, by the choice of point E, we have $EC < DB + BC$, but $DB + BC = AC + AD < AC + AE$, and hence, $EC < AC + AE$. It is also easy to prove that $AC < CE + AE$ and $AE < AC + CE$. Hence, in the second case, it is sufficient to replace one vertex by the other. Thus, the lemma is proved.

When a polygon is partitioned into convex triangles, each of these triangles can be partitioned into smaller triangles, e.g., by connecting the middles of their sides with each other. It is possible to prove that we so arrive at a partition into arbitrarily small triangles. However, the possibility of partitioning into convex triangles was stated in Lemma 1 only for convex polygons homeomorphic to a disk. This is not sufficient for our purposes since, e.g., the partition of a polygon homeomorphic to a sphere will be especially important to us. Therefore, the problem consists of ensuring the possibility of application of Lemma 1 by the covering of a given polygon by small convex polygons without common interior points. The solution of this problem is prepared by three lemmas on convex polygons which will be

© 2006 by Taylor & Francis Group, LLC

proved here. In these lemmas, we speak about convex polygons in some domain homeomorphic to a disk.

Lemma 2. *A figure composed of two convex polygons P and Q that are homeomorphic to a disk can be partitioned into finitely many convex polygons homeomorphic to a disk, having no common interior points pairwise, and such that one of them is contained in P and the others are contained in Q. The only exception is the trivial case in which P itself lies in Q.*

Proof. Let P and Q be two convex polygons homeomorphic to a disk. If they have no common interior points or if one of them is contained in the other then the claim of the lemma is obvious. Therefore, we assume that P and Q have common interior points and neither of them is contained in the other. Then a part of the boundary of each of them lies in the interior of the other. (This follows from the fact that P and Q are in the domain homeomorphic to a disk by the above assumption.)[14]

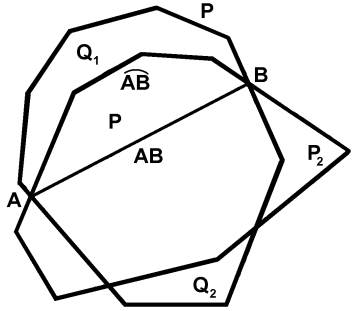


Fig. 26

Let the boundary of the polygon P intersect the boundary of the polygon Q at points A and B, and let this boundary be in the interior of Q on the segment $\overset{\frown}{AB}$ (Fig. 26). The points A and B belong to P and Q simultaneously, and since the intersection of P and Q is convex in accordance with Theorem 2 of the preceding section, there exists a shortest arc AB lying in P, as well as in Q. Draw this shortest arc AB.

Let us prove the following assertion.

The shortest arc AB is cut off a polygon Q_1 lying in Q from the figure $P + Q$ composed of P and Q. The remaining figure is either a polygon included in P or consists of two polygons P_2 and Q_2, which are included in P and Q, respectively. Moreover, the number of segments of the boundary of P_2 that pass in the interior of Q_2 is less than the number of segments of the boundary of P lying in the interior of Q (namely, the segment AB is excluded). All polygons mentioned are convex and homeomorphic to a disk.

If we look at the figure then this assertion becomes absolutely obvious. Before proving this assertion, we show how this assertion implies our lemma.

If the remainder of the cutting off the polygon Q_1 is a polygon lying in P, then the lemma is proved. If two polygons P_2 and Q_2 remain after the cutting off then we can apply the same assertion to them, i.e., to cut off from $P_2 + Q_2$ a polygon contained in Q_2 and hence in Q and to arrive at the figure $P_3 + Q_3$. Under this procedure, the number of segments of the boundary of P which are contained in Q_2 decreases. The number of segments of the boundary of P contained in Q is finite, since P and Q are bounded by a finite number of shortest arcs, and nonoverlapping

[14] For example, if P and Q cover a sphere, then the whole boundary of P can be contained in Q and vice versa.

© 2006 by Taylor & Francis Group, LLC

shortest arcs can have only finitely many common points. Therefore, applying our argument further, after a finite number of steps, we arrive at the case in which there is only one polygon contained. This will prove the lemma.

Therefore, we have to prove the above-formulated proposition concerning the shortest arc AB. According to Theorem 1 of the preceding section, a shortest arc connecting two points of a convex polygon either has no common points, except for the endpoints, with the boundary of this polygon or lies entirely on the boundary of this polygon. Therefore, the shortest arc AB either lies in the interior of Q and then divides Q into two parts, since Q is homeomorphic to a disk, or the shortest arc AB goes along the boundary of Q. In the latter case, we also say that this arc "divides Q" but one such part degenerates into a line. Since the shortest arc AB also lies in P, the same is true for the polygon P. According to Theorem 4 of the preceding section, the parts into which the shortest arc AB divides P and Q are also convex.

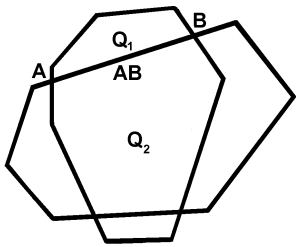


Fig. 27

The following two possibilities are open for the shortest arc AB: (1) this shortest arc has common points with the segment $\overset{\frown}{AB}$ of the boundary of P different from A and B; (2) there are no common points with this segment, except for A and B.

Consider the first case in which AB and $\overset{\frown}{AB}$ have common points different from A and B. This means that the shortest arc AB has common points with the boundary of the polygon P that are not the ends of this arc, and therefore this arc lies on the boundary of this polygon (Fig. 27). Then the shortest arc AB is obviously the segment $\overset{\frown}{AB}$ of the boundary of P. Hence AB lies in the interior of Q and divides Q into two convex polygons Q_1 and Q_2. At the same time, the shortest arc AB lies on the boundary of P, and, therefore, near AB, there are interior points of P in one of the polygons Q_1 and Q_2, while the other polygon has no such interior points.

Assume for definiteness that Q_1 has no interior points of P near AB. Then there are no interior points of P in Q_1 at all. Indeed, otherwise, the intersection of P and Q consists of two disconnected parts; one lies in Q_1 and the other lies in Q_2. This contradicts the fact that the intersection of P and Q is a convex polygon.

Now let us cut off the polygon Q_1. The remaining polygons are P and Q_2. The segment $\overset{\frown}{AB}$ of the boundary of P is not longer in the interior of Q_2. Hence the interior of the polygon Q_2 contains less segments of the boundary of P than the initial polygon P.

Now consider the second case in which the shortest arc AB and the segment $\overset{\frown}{AB}$ of the boundary of the polygon P have no common points, except for A and B (Fig. 26). In this case, AB and $\overset{\frown}{AB}$ form a closed curve that bounds one of the parts into which AB divides the polygon P. Denote this part by P_1 and the other by P_2.

© 2006 by Taylor & Francis Group, LLC

Since AB and $\overset{\frown}{AB}$ are contained in Q, the whole part P_1 of the polygon P bounded by them is contained in Q.[15]

The shortest arc AB also divides the polygon Q into two parts. Since P_1 is contained in Q, one of these parts contains P_1 entirely. Denote this part by Q_1 and the other part by Q_2; Q_1 is a convex polygon, since Q_1 is cut off from Q by the shortest arc AB and contains interior points (since $AB \neq \overset{\frown}{AB}$, and hence P_1 contains interior points).

Let us cut off the polygon Q_1 from the figure $P + Q$ formed by the polygons P and Q. There remain P_2 and Q_2, since P_1 is contained in Q_1. If P does not reduce to P_1, and thus, is not contained in Q, then P_2 is a polygon. As for Q_2, then this polygon can degenerate into a segment. In this case, our construction is over: the figure $P + Q$ is divided into the convex polygons Q_1 and P_1 having no common interior points. If Q_2 does not reduce to a segment then we obtain a figure composed of two convex polygons P_2 and Q_2. Moreover, the number of segments of the boundary of P_2 that are contained in Q_2 is less than the number of segments of the boundary of P inside Q, since the segment $\overset{\frown}{AB}$ is excluded and no segments arise.

Thus, in both cases of location of the shortest arc AB, we have proved that this shortest arc cut off a convex polygon Q_1 contained in Q from the figure $P + Q$ such that the remaining figure $P_2 + Q_2$ again consists of two convex polygons homeomorphic to a disk, but the number of segments of the boundary of P_2 passing in the interior of Q_2 is less than the number of segments of the boundary of P lying inside Q. The number of segments of the boundary of Q contained in P is finite, since P and Q are bounded by a finite number of shortest arcs and nonoverlapping shortest arcs can have only a finite number of common points. Therefore, applying our construction further and cutting off new polygons at each step, after a finite number of steps, we arrive at the case where only a polygon contained in P remains. The lemma is proved.

Lemma 3. *Let $P_1, \ldots, P_n$ and $Q_1, \ldots, Q_n$ be two systems of polygons, each of which is homeomorphic to a disk, and, furthermore, let the polygons of the same system have no common interior points. Then the figure composed of all these polygons can be partitioned into a finite number of polygons that are also convex, homeomorphic to a disk, and have no common interior points.*

Proof. We proceed by induction on n and m. If $n = m = 1$, this lemma reduces to the previous one. Assume that our lemma holds for one polygon P and m polygons $Q_1, \ldots, Q_n$; let us prove that this lemma also holds for $m + 1$ polygons $Q_1, \ldots, Q_{m+1}$.

According to Lemma 2, the figure $P + Q_{m+1}$ can be divided into polygons that have no interior points such that only one polygon P' among them is contained in P, and other polygons are contained in Q_{m+1}. Since Q_{m+1} has no common points with $Q_1, \ldots, Q_m$, the polygons $Q_1, \ldots, Q_m$ are not changed in the process of this

[15] We consider a domain homeomorphic to a disk. In such a domain, the closed curve $\overset{\frown}{AB} + AB$ bounds one compact domain. This is precisely P_1, and since Q is homeomorphic to a disk, the inclusion $Q \supset \overset{\frown}{AB} + AB$ implies $Q \supset P_1$.

© 2006 by Taylor & Francis Group, LLC

construction. By the induction hypothesis, the figure composed of these polygons and one polygon P' can be divided into polygons that have no common interior points. Hence the whole figure $P + (Q_1 + \ldots + Q_{m+1})$ admits such a construction. This proves the lemma for $n = 1$ and any m.

Assume now that this lemma holds for a given n and any m; then we shall prove that this lemma holds for $n + 1$ and any m. Suppose that we have polygons $P_1, \ldots, P_{n+1}$ and $Q_1, \ldots, Q_m$ satisfying the conditions of the lemma. By what we have proved above, the figure $P_{n+1} + (Q_1 + \ldots + Q_m)$ can be divided into polygons without common interior points. After that, n polygons $P_1, \ldots, P_n$ and a certain number of other polygons that have no common interior points remain. By the induction hypothesis, the figure composed of all these polygons can be divided in the required way; after that, the whole figure $(P_1 + \ldots + P_{n+1}) + (Q_1 + \ldots + Q_m)$ turns out to be divided; as required.

Lemma 4. *A figure composed of a finite number of convex polygons, each of which is homeomorphic to a disk, can be divided into a finite number of analogous polygons without common interior points.*

Proof. Let $P_1, P_2, \ldots, P_n$ be polygons forming the figure under study. Take pairs of these polygons that have common points with one another and enumerate these pairs in an arbitrary way. Take the first such pair and, using Lemma 2, divide it into a finite number of convex polygons without common interior points. Then we take the second pair and apply the same procedure to this pair, and so on.

Assume that we have arrived at some (P_k, P_l). It can happen that we met one or both polygons of this pair in the preceding pairs. Then these polygons are divided into polygons $P_{k1}, P_{k2}, \ldots, P_{km_k}, P_{l1}, P_{l2}, \ldots, P_{lm_l}$, and moreover, the polygons P_{ki}, as well as the polygons P_{lj}, have no common interior points with each other. Applying Lemma 3, we can divide all these polygons into those without common interior points.

Thus, passing to the pair (P_k, P_l) nearest with respect to the order, every time we take into account all the operations, which were already realized, that is, instead of the polygons P_k and P_l, we take those polygons that were obtained as a result of the previous divisions. Moreover, every time we use Lemma 3. Inspecting all pairs of the polygons $P_1, \ldots, P_n$, we obtain the required partition of the figure composed of these polygons.

Now it is easy to prove the theorem on triangulation.

Theorem. *Each polygon can be divided into arbitrarily small triangles such that the sum of two sides of every triangle is greater than the third. If this polygon is convex then these triangles can be chosen convex.*

Proof. Let P be a polygon that should be divided into triangles of diameter less than a given ε. By the theorem in Sec. 4, each point of the polygon P can be surrounded by a convex polygon of diameter less than ε homeomorphic to a disk. Moreover, we can take these polygons so small that no one of them contains more than one vertex of the polygon P. By the Borel lemma, we can refine a finite subcover of P from these polygons.

© 2006 by Taylor & Francis Group, LLC

According to Lemma 4, the figure formed by the chosen polygons[16] can be divided into convex polygons homeomorphic to a disk without common points pairwise. Obviously, these polygons $Q_1, \ldots, Q_n$ cover P, their diameters are less than ε, and none of them contains more than one vertex of the polygon P. Of course, we omit those polygons that have no common interior points with P.

The polygons $Q_1, \ldots, Q_n$ can be of the following three types: (1) polygons that have no points of the boundary of P in their interiors; (2) polygons that contain only segments of sides of P and do not contain the vertices of the polygon P; (3) polygons that contain one vertex of P in the interior. Let us consider the polygons of each type separately.

The polygons of the first type are contained in P, and we divide them into convex triangles according to Lemma 1.

The common part of every polygon Q_i of the second type and the polyhedron P consists of convex polygons homeomorphic to a disk that have no common interior points. In fact, by condition, Q_i does not contain the vertices of P, and, therefore, any side of the polygon P entering Q_i also goes away from Q_i. Thus, this side divides Q_i into two convex polygons homeomorphic to a disk. If Q intersects several sides of P then repeating this argument sequentially for all of them, we make sure that they divide Q_i into convex polygons homeomorphic to a disk. A part of these polygons $Q_{i1}, \ldots, Q_{im}$ will represent the intersection of Q_i with P.

Also, we divide each of the polygons $Q_{i1}, \ldots, Q_{im}$ into convex triangles according to Lemma 1.

We are left with the polygons of the third type. Let Q_k be one of them. If some side of the polygon P enters Q_k and leaves this polygon then this side divides Q_k into two convex polygons homeomorphic to a disk. Therefore, all these sides of P (if they exist) partition Q_k into polygons homeomorphic to a disk. Those polygons among these polygons that have no common points with P are excluded; denote the other polygons by $Q_{k1}, \ldots, Q_{kl}$.

The polygon Q_k contains one vertex A of the polygon P in its interior. Suppose that this vertex is contained in the interior of Q_{k1}. Hence the other polygons $Q_{k1}, \ldots, Q_{kl}$ already have no parts of the boundary of P in their interiors. Therefore, they are included in P. According to Lemma 1, we divide each of them into convex triangles.

If the polygon P is convex, then its common part with the polygon Q_{k1} is a convex polygon obviously homeomorphic to a disk, since there are no other parts of the boundary of the polyhedron P in the interior of Q_{k1}, except for segments of two sides meeting at the vertex A.

Hence, if the polyhedron P is convex, then performing the same operation for all polygons of the third type, we arrive at the total partition of P into convex triangles.

If the polygon P is not convex, then we have to proceed in a slightly different way for the polygon Q_{k1}.The polygon Q_{k1} contains the vertex A of the polygon P in its interior. Let us connect this vertex with the vertices of the polygon Q_{kl} (Fig. 28) by shortest arcs. These shortest arcs lie in the interior of the polygon

[16] Although the figure can be nonhomeomorphic to a disk, Lemma 4 is applicable if ε is sufficiently small; then pairs of disjoint small polygons lie in neighborhoods homeomorphic to a disk.

© 2006 by Taylor & Francis Group, LLC

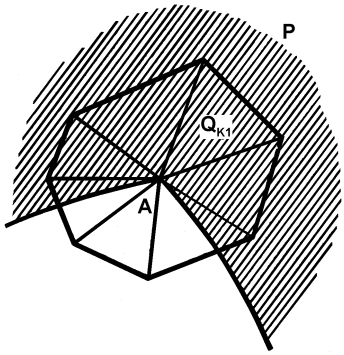


Fig. 28

Q_{k1}, since this polygon is convex and the point A lies in its interior. We add sides of the polygon P which meet each other at A to these shortest arcs. As a result, the polygon Q_{k1} is divided into triangles with one common vertex A; we take those triangles that are contained in P. If for some of these triangles the sum of the lengths of two sides is equal to the length of the third, then applying the method of the proof of Lemma 1, we partition these triangles into triangles for which this property fails. Proceeding in the same way for all polygons of the third type, we arrive at a triangulation of the polygon P which has the required properties; this completes the proof of the theorem.

If the polygon P is not convex, then the triangles into which this polygon is divided can be nonconvex. For a polygon on a convex surface, we can prove that this polygon can always be divided into convex triangles even if this polygon itself is not convex. It is unknown if the same is true of any intrinsic metric manifold satisfying the nonoverlapping condition for shortest arcs.

© 2006 by Taylor & Francis Group, LLC

Chapter III

CHARACTERISTIC PROPERTIES OF THE INTRINSIC METRIC

1. Convergence of the Metrics of Convergent Convex Surfaces

The method for the approximation by convex polyhedra is our main tool for the study of the intrinsic metric of a convex surface. This method is based on the well-known theorem, which asserts that for each closed convex surface there exists a sequence of closed convex polyhedra converging to this surface. The second main point of this method consists of the theorem on the convergence of metrics of convergent convex surfaces. These two theorems imply the possibility of passing to the limit from the metrics of convex polyhedra to the metric of an arbitrary convex surface. Therefore, we begin the study of the characteristic properties of the intrinsic metric of a convex surface with the following theorem on the convergence of metrics.

Theorem 1. *If a sequence of closed convex surfaces F_n converges to a closed convex surface F and if two sequences of points X_n and Y_n on F_n converge to two points X and Y on F, respectively, then the distances between the points X_n and Y_n measured on the surfaces F_n converge to the distance between the points X and Y measured on F, i.e., $\rho_F(XY) = \lim_{n\to\infty} \rho_{F_n}(XY)$.*

We also include doubly-covered convex domains for the set of closed convex surfaces; on such a "surface," each point that does not lie on the boundary is considered as two points. Therefore, if a surface F degenerates into a doubly-covered plane domain, then it is necessary to refine the notion of convergence. We distinguish two sides of the plane domain into which the surface F degenerates and assume that some points of the surface F lie on one side and other points lie on the other side. Let a point X lie in the interior of the domain G. Consider all oriented straight lines p which are orthogonal to the plane containing the domain G. Let us aim these lines to the half-space into which the side of the domain G containing the point X is turned. These lines intersect the surfaces F_n. We say that points X_n on the surfaces F_n converge to X if (1) these points converge to X as points in the space, and (2) for each sufficiently large n the point X_n is placed on the surface F_n in such a way that the line p passing through this point aims outside the surface F_n, i.e., is directed to the half-space in which the outer normal of some supporting plane of the surface F_n at the point X_n goes. If this condition fails for some arbitrarily large n, the sequence of points X_n is not assumed convergent to

© 2006 by Taylor & Francis Group, LLC

the point X. If X lies on the boundary of the domain G, then we can assume that this point lies on any side, and this additional condition turns out superfluous. For the given definition of the convergence only, Theorem 1 turns out to be true even in the case where the limit surface F degenerates into a doubly-covered plane domain.

To prove this theorem, we need two lemmas.

Let F be some closed convex surface, and let A be a point not lying in the interior of the body H bounded by F. There is a point A' on F nearest to A. As is easily seen, this point has the property that the supporting plane to F which is orthogonal to the segment AA' passes through this point (only if A itself does not lie on F; otherwise, $A = A'$). Incidentally, we see from this that the point A' is unique. This point is called the projection of A to F.

We say that a curve L lies outside a surface F if this curve has no points in the interior of the body H. If a curve L lies outside F, then the geometric locus of projections of all points of the curve L is called the projection L' of this curve to F.

Busemann and Feller have proved the following important lemma.[1]

Lemma 1. *If a curve L lying outside a closed convex surface F has finite length, then its projection to F is a curve whose length does not exceed the length of L.*

The proof is based on the following observation:

Let A' and B' be two points lying on two half-planes that make a convex dihedral angle V. Let two points A and B lie outside the angle V on the perpendiculars erected from the points A' and B'. Then the distance between A and B is not less than the distance between A' and B'. We leave the proof to the reader. (See Fig. 29, where the angle V is made by support planes to the surface F at the projections A' and B' of the points A and B.)

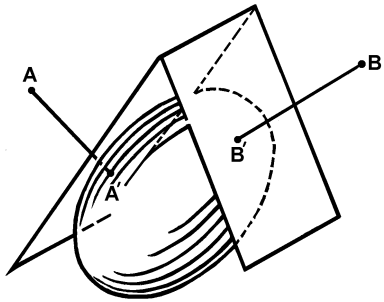


Fig. 29

Now let the points A and B lie outside F (or on F), and let A' and B' be the projections of these points to F. The support planes at the points A' and B', orthogonal to the segments AA' and BB', make a dihedral angle, and the points A and B lie outside this angle. Therefore, applying the previous observation, we conclude that the distance between A' and B' does not exceed the distance between A and B, i.e., *the distance between the projections of points to a convex surface is always not greater than the distance between these points themselves.*

Let a curve $X(t)$ $(0 \le t \le 1)$ lie outside F, and let $X'(t)$ be the projections of the points $X(t)$ to F; these projections form the projection of the curve $X(t)$ to F. If ρ denotes the distance in the space then, as we have shown, for all t_1 and t_2,

$$\rho(X'(t_1)X'(t_2)) \le \rho(X(t_1)X(t_2)) \tag{1}$$

[1] H. Busemann und W. Feller, Krümmungseigenschaften konvexer Flächen, Acta Math., Vol. 66 (1936), p. 41 (Hilfssatz 1).

© 2006 by Taylor & Francis Group, LLC

Since the points $X(t)$ comprise a continuous curve, for any $\varepsilon > 0$, there is some $\delta > 0$ such that $\rho(X(t_1)X(t_2)) < \varepsilon$ whenever $|t_1 - t_2| < \delta$. But then (1) implies that $\rho(X'(t_1)X'(t_2)) < \varepsilon$, too; this means that the points $X'(t)$ comprise a continuous curve.

Now we take arbitrary values t_i such that $0 = t_0 < t_1 < \ldots < t_n = 1$. Then by inequality (1),

$$\sum_{i=1}^{n} \rho(X'(t_{i-1})X'(t_i)) \leq \sum_{i=1}^{n} \rho(X(t_{i-1})X(t_i)). \tag{2}$$

The sum on the right-hand side does not exceed the length s of the curve $X(t)$, and hence

$$\sum_{i=1}^{n} \rho(X'(t_{i-1})X'(t_i)) \leq s. \tag{3}$$

But the values t_i are chosen arbitrarily, and this inequality holds for any choice of values. Therefore, the least upper bound of the sums in inequality (3) is also not greater than s; but this is, by definition, the length of the curve $X'(t)$. Consequently, the length of the curve $X'(t)$ does not exceed the length of the curve $X(t)$; this is what we were required to prove.

The proof of Lemma 1 immediately implies the following lemma.

Lemma 2. *If a curve L lies outside a closed convex surface F, then the length of this curve is not less than the distance on F between the projections of its endpoints to the surface F. In particular, if the ends A and B of the curve L lie on F, then the length of the curve L is not less than the length of a shortest arc AB on the surface F.*

Indeed, by Lemma 1, the length of L is not less than the length of its projection to F, while the length of the projection is, of course, no less than the length of a shortest arc connecting the ends of this projection.

This lemma is useful not only for the proof of the theorem, which is given at the beginning of this section, but, as Busemann and Feller have shown, it proves very useful for the study of properties of shortest arcs on a convex surface. It will be used for this purpose in Sec. 6 of Chapter IV.

Lemma 3. *If a sequence of closed convex surfaces F_n converges to a nondegenerate surface F and if points X_n and Y_n lying on F_n converge to the same point X on F, then the distance between X_n and Y_n on F_n converges to zero:*

$$\lim_{n\to\infty} \rho(X_n Y_n) = 0.$$

Proof. Since the surface F is nondegenerate, the surfaces F_n are also nondegenerate at least for a sufficiently large n. Restricting ourselves to large n only, we can take a point O around which we may circumscribe the ball of some radius $r > 0$ that lies inside all surfaces F_n and inside F.

Choose some large n and draw the plane through the points O, X_n, and Y_n (Fig. 30). This plane intersects the surface F_n along a convex curve. Draw the supporting lines to this curve at the points X_n and Y_n; let Z be a point

© 2006 by Taylor & Francis Group, LLC

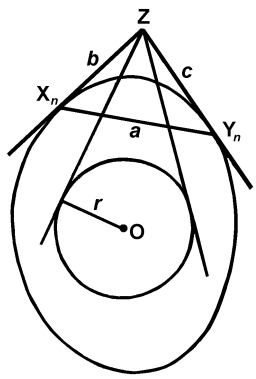


Fig. 30

resulting in the cross-section of the sphere of radius r centered at O by our plane. Let β be the angle made by these tangents. Then

$$\sin\frac{\beta}{2} = \frac{r}{OZ}. \tag{4}$$

If D is the diameter of the surface F_n, then $OX_n < D$, and if $X_N Z = b$ and $Y_n Z = c$, then, obviously, $OZ < OX_n + X_n Z < D + b$, and, moreover,

$$OZ < D + b + c. \tag{5}$$

If α is the angle at Z in the triangle $X_n Y_n Z$, then, obviously,

$$\alpha \geq \beta. \tag{6}$$

Inequalities (4), (5), and (6) imply

$$\sin\frac{\alpha}{2} > \frac{r}{D + b + c}. \tag{7}$$

If we denote the length of $X_n Y_n$ by a, then[2]

$$\sin\frac{\alpha}{2} \leq \frac{a}{b + c}. \tag{8}$$

Therefore, (7) implies

$$\frac{a}{b + c} > \frac{r}{D + (b + c)};$$

Hence, if $r - a > 0$, then

$$b + c < \frac{aD}{r - a}. \tag{9}$$

The segments $X_n Z$ and $Y_n Z$ lie outside the surface F_n, and so by the previous lemma,

$$\rho_{F_n}(X_n Y_n) \leq X_n Z + Y_n Z = b + c. \tag{10}$$

Therefore, (9) implies the inequality[3]

$$\rho_{F_n}(X_n Y_n) < \frac{AD}{r - a}. \tag{11}$$

[2]For given a and $a + b$, the angle α is maximal if the triangle $X_n Y_n Z$ is equilateral, in which case $a = (b + c)\sin(\alpha/2)$.

[3]This inequality is of interest in its own right, since it yields the bound of the intrinsic distance between the points X_n and Y_n depending on their distances a in the space, the diameter D of the surface, and the radius r of the ball included in this surface. This bound is very rough. If ρ is the distance on a convex surface, a is the distance between the same points in the space, r is the radius of the ball contained in this surface, and R is the radius of the ball including the surface, then it is possible to prove that the following sharp bound holds:

$$\frac{\rho}{a} \leq \sqrt{\left(\frac{R}{r}\right)^2 - 1} + \frac{R}{2}\arcsin\frac{r}{R} < \frac{R}{r} - +1.$$

It is an interesting problem to prove this and to find a surface at which the maximum of the ratio ρ/a is attained (given R and r).

© 2006 by Taylor & Francis Group, LLC

Here, $\rho_{F_n}(X_nY_n)$ is the distance between X_n and Y_n. By condition, these points converge to the same point X as $n \to \infty$. Therefore, $a \to 0$ while D remains bounded, since the surfaces F_n converge. Therefore, (1) implies

$$\rho_{F_n}(X_nY_n) \to 0;$$

this is as claimed.

We now prove the theorem that was formulated at the beginning of this section.

Let closed convex surfaces F_n converge to a surface F, and let points X_n and Y_n on F_n converge to points X and Y on F, respectively. All surfaces F_n can be included into a sufficiently large ball S. There always exist points A_n and B_n in this ball whose projections to F_n are the points X_n and Y_n. These points A_n and B_n are obtained if we draw the outer normals to the supporting planes of the surface F_n at the points X_n and Y_n up to their intersection with the surface of the ball S. Therefore, by Lemma 2, the distances on F_n between the points X_n and Y_n do not exceed the length of the great semi-circle of the ball S, i.e., they are all bounded.

We have proved in Sec. 2 of Chapter II that every two points on a closed convex surface can be connected by a shortest arc. Let L_n be shortest arcs connecting the points X_n and Y_n on the surfaces F_n. The length of each of these arcs is equal to the distance on F_n between X_n and Y_n, i.e.,

$$s(L_n) = \rho_{F_n}(X_nY_n). \tag{12}$$

It follows from what we have just proved that the lengths $s(L_n)$ are uniformly bounded, and, therefore, by Theorem 4 in Sec. 4 of Chapter II, we can choose a sequence of the curves L_n which converges to some curve L. Obviously, this curve lies on F and connects the points X and Y. By Theorem 5 in Sec. 1 of Chapter II, the length of this curve satisfies the inequality

$$s(L) \le \lim_{n\to\infty} s(L_n). \tag{13}$$

At the same time, it is clear from the definition of the distance on a surface that

$$s(L) \ge \rho_F(XY), \tag{14}$$

where ρ_F is the distance on F. Formulas (12), (13), and (14) imply

$$\rho_F(XY) \le \liminf_{n\to\infty} \rho_{F_n}(X_nY_n). \tag{15}$$

If we now demonstrate that the reverse inequality also holds, then we prove the theorem. Here, we consider the following two cases separately: (1) the surface F is nondegenerate and (2) this surface is degenerate.

Let the surface F be nondegenerate. Take a point O inside this surface and perform the homothety transform with the center at O of the surfaces F_n so that all these surfaces turn out to be inside F. Since the surfaces F_n converge to F, the coefficients of homothety λ_n can be taken closer and closer to 1 as n increases and $\lambda_n \to 1$ as $n \to \infty$. The surfaces and points, which are obtained from the surfaces F_n and the points X_n, Y_n as a result of this transformation, will be denoted by

© 2006 by Taylor & Francis Group, LLC

$\lambda_n F_n$, $\lambda_n X_n$, and $\lambda_n Y_n$. Since $\lambda_n \to 1$ and the points X_n and Y_n tend to X and Y, the points $\lambda_n X_n$ and $\lambda_n Y_n$ also converge to X and Y, respectively.

Let X_n' and Y_n' be the projections of the points X and Y to the surfaces $\lambda_n F_n$. By Lemma 2,

$$\rho_{\lambda_n F_n}(X_n' Y_n') \le \rho_F(XY). \tag{16}$$

Obviously, the points X_n' converge to X as $n \to \infty$, and at the same time, the points $\lambda_n X_n$ also converge to X. Therefore, by Lemma 3,

$$\rho_{\lambda_n F_n}(X_n' \lambda_n X_n) \to 0, \tag{17}$$

and, by the same arguments,

$$\rho_{\lambda_n F_n}(Y_n', \lambda_n Y_n) \to 0. \tag{18}$$

By the "triangle inequality,"

$$\rho_{\lambda_n F_n}(\lambda_n X_n, \lambda_n Y_n) \le \rho_{\lambda_n F_n}(\lambda_n X_n, X_n') + \rho_{\lambda_n F_n}(X_n' Y_n') + \rho_{\lambda_n F_n}(Y_n' \lambda_n, Y_n). \tag{19}$$

Using inequality (16) and relations (17) and (18) and passing to the limit as $n \to \infty$, we obtain

$$\limsup_{n\to\infty} \rho_{\lambda_n F_n}(\lambda_n X_n \lambda_n Y_n') = \rho_F(XY). \tag{20}$$

But under the homothety with coefficient λ_n, all distances change by λ_n times, and, therefore,

$$\rho_{\lambda_n F_n}(\lambda_n X_n \lambda_n Y_n) = \lambda_n \rho_{\lambda_n F_n}(X_n Y_n); \tag{21}$$

since $\lambda_n \to 1$, formula (20) implies

$$\limsup_{n\to\infty} \rho_{\lambda_n F_n}(X_n Y_n) \le \rho_F(XY). \tag{22}$$

Now, comparing this formula with formula (15), we obtain

$$\lim_{n\to\infty} \rho_{F_n}(X_n Y_n) = \rho_F(XY);$$

this is what we were required to prove.

Assume now that the surface F degenerates into a doubly-covered plane domain. Consider the following two cases: (1) the points X and Y lie on one side of F (moreover, they can lie on the boundary) and (2) the points X and Y lie on different sides of F.

In the first case, we draw the plane P_n through the points X_n and Y_n that is orthogonal to the plane including F. The planes P_n intersect the surfaces F_n along closed convex curves.[4] We take those segments L_n on these curves between the points X_n and Y_n which converge to the segment XY. This is always possible, since F_n converge to F and X_n and Y_n converge to X and Y.

[4] These curves degenerate into doubly-covered segment whenever F_n degenerate. If the plane P_n is a supporting plane to F, then we take the segment $\overline{X_n Y_n}$ as L_n.

© 2006 by Taylor & Francis Group, LLC

The curves L_n are plane convex curves that converge to the segment $\overline{XY}$, and, therefore, the lengths of these curves converge to the length of this segment.[5] The length of the curve L_n is not less than the distance between the points X_n and Y_n on F, while the length of the segment $\overline{XY}$ is the distance between the points X and Y on F. Therefore,

$$\rho_F(XY) \geq \lim_{n\to\infty} \rho_{F_n}(X_nY_n). \tag{23}$$

Assume now that the points X and Y lie on the different sides of F. Then a shortest arc of these points on F consists of two line segments meeting at some point Z of the boundary of F. Therefore,

$$\rho_F(XY) = \rho_F(XZ) + \rho_F(YZ). \tag{24}$$

Take a sequence of points Z_n on the surfaces F_n which converges to Z. Since X and Y, as well as Y and Z, lie now on one side of F_n, we have

$$\rho_F(XY) \geq \lim_{n\to\infty} \rho_{F_n}(X_nZ_n), \quad \rho_F(XZ) \geq \lim_{n\to\infty} \rho_{F_n}(Y_nZ_n). \tag{25}$$

At the same time,

$$\rho_{F_n}(X_nZ_n) + \rho(Z_nY_n) \geq \rho_{F_n}(X_nY_n). \tag{26}$$

Again, formulas (24), (25), and (26) imply

$$\rho_F(XY) \geq \lim_{n\to\infty} \rho_{F_n}(X_nY_n). \tag{23}$$

Thus, inequality (23) is proved under both hypotheses on the location of the points X and Y. Combining this inequality with inequality (15), we obtain

$$\rho_F(XY) = \lim_{n\to\infty} \rho_{F_n}(X_nY_n);$$

this is what we were required to prove.

In exactly the same way, we may prove that if F_n converge to a segment and the points X_n and Y_n lying on F_n converge to X and Y, then $\rho_{F_n}(X_nY_n)$ converge to the distance between X and Y. If this segment degenerates into a point, then $\rho_{F_n}(X_nY_n) \to 0$.

Theorem 1 can be strengthened by proving that the convergence of metrics on convergent convex surfaces turns out to be uniform in some sense. Namely, the following assertion holds.

Theorem 2. *For every convex closed surface F and for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever the deviation*[6] *of a closed convex surface S from F is less*

[5] Indeed, let us construct a rectangle with vertices at the points X_n and Y_n that contains the curve L_n. If h is the height of this rectangle then $s(L_n) \leq \overline{X_nY_n} + 2h$. Since $h \to 0$ and $X_n \to X$, $Y_n \to Y$, we have $\limsup_{n\to\infty} s(L_n) \leq XY$. On the other hand, by Theorem 5 in Sec. 1 of Chapter II, $\overline{XY} \leq \liminf_{n\to\infty} s(L_n)$, since $L_n \to XY$. Consequently, $XY = \lim_{n\to\infty} s(L_n)$.

[6] The distance $\rho(X, F)$ from a point X to a set F is the greatest lower bound of the distances from X to the points of the set F. The deviation of a set F_1 from a set F_2 is the least upper bound of the distances from points of F_1 to F_2. The deviation of F_1 and F_2 from one another is the maximum of the deviations of F_1 from F_2 and F_2 from F_1. If the deviation of the surface F_n from F tends to zero then F_n converge to F.

© 2006 by Taylor & Francis Group, LLC

than δ and the distances of some points X and Y on F from some points A and B on S are also less than δ, we have

$$|\rho_F(XY) - \rho_S(XY)| < \varepsilon,$$

where ρ_F and ρ_S are the distances on F and S, respectively.

Proof. Assume the contrary. Then there exists a sequence of surfaces S_n converging to F such that it is possible to choose points X^n and Y^n on F and X_n and Y_n on S_n so that the distances from the points X^n and Y^n to X_n and Y_n tend to zero, but for an arbitrarily large n,

$$|\rho_{S_n}(X_nY_n) - \rho_F(X^nY^n)| > \varepsilon > 0. \tag{27}$$

We can choose convergent sequences from the points X^n and Y^n, and, therefore, we can simply assume that the sequences X_n and Y_n of points on S_n converge to some points X and Y on F. Then, instead of (27), we obtain

$$|\rho_{S_n}(X_nY_n) - \rho_F(XY)| > \varepsilon > 0. \tag{28}$$

But, by Theorem 1, $\rho_{S_n}(X_nY_n) \to \rho_F(XY)$. This contradicts inequality (28); the theorem is proved.

Theorem 1 also implies the following assertion.

Theorem 3. *Let a sequence of closed convex surfaces F_n converge to a surface F, and let a sequence of pairs of points X_n and Y_n lying on F_n converge to a pair of points X and Y lying on F; then we can choose a subsequence converging to a shortest arc XY on the surface F from each sequence of shortest arcs X_nY_n on the surfaces F_n.*

Proof. Under the hypotheses of the lemma, by Theorem 1 we have

$$\rho_F(XY) = \lim_{n\to\infty} \rho_{F_n}(X_nY_n), \tag{29}$$

where $\rho_F(XY)$ is the distance on F, i.e., the length of the shortest arc XY, and $\rho_{F_n}(X_nY_n)$ is the distance on F_n, i.e., the length of the shortest arc X_nY_n. On the other hand, we can choose a convergent subsequence from each sequence of rectifiable curves lying in a bounded domain of the space and having the lengths uniformly bounded; moreover, the length of the limit curve is not greater than the lower limit of the lengths of the curves of this sequence. It is clear from formula (29) that the lengths of the shortest arcs X_nY_n are uniformly bounded and their limit is equal to $\rho_F(XY)$. Therefore, we can take a sequence $X_{n_i}Y_{n_i}$ converging to some curve XY whose length does not exceed $\rho_F(XY)$. But this curve obviously lies on the surface F and, therefore, the length of this curve cannot be less than $\rho_F(XY)$. Hence this length is equal to $\rho_F(XY)$, i.e., the curve XY is a shortest arc on F.

© 2006 by Taylor & Francis Group, LLC

2. The Convexity Condition for a Polyhedral Metric

In Chapter I, we have introduced the "convexity condition" for a metric which reads as follows:

Let L and M be two shortest arcs emanating from a point O of a manifold with intrinsic metric. Let X and Y be variable points on these shortest arcs, let x and y be the distances from X and Y to O, and let $z = z(x, y)$ be the distance between X and Y. Finally, let $\gamma(x, y)$ be the angle in the plane triangle with sides equal to x, y, and z, subtended by the side equal to z. The convexity condition requires that $\gamma(x, y)$ is a nondecreasing function of x and y on each interval of values $0 < x \le x_0$, $0 < y \le y_0$, where there exist shortest arcs XY.[7] (The values $x = 0$ and $y = 0$ are excluded, since $\gamma(x, 0)$ and $\gamma(0, y)$ are not defined.)

Our first goal is to show that the intrinsic metric of any convex surface satisfies the convexity condition; on the basis of this condition, it is possible to develop the intrinsic geometry of convex surfaces not using other properties of their intrinsic metrics except for general properties of every intrinsic metric. First, we will show that the convexity condition holds for convex polyhedra; passing to the limit, we then extend this property to all convex surfaces. The whole derivation of the convexity condition on polyhedra is of a purely intrinsic geometric character. We need to know only that the metric of each of them is a polyhedral metric of positive curvature. Therefore, instead of polyhedra, we will consider an arbitrary manifold R with such a metric.

Before proving that the convexity condition holds on R, we deduce some properties of shortest arcs in R. These properties will be needed for the following.

(1) By the definition of a polyhedral metric of positive curvature, each point of the manifold on which this metric is given has a neighborhood that is isometric to a cone with complete angle at the vertex $\le 2\pi$. Therefore, a shortest arc emanating from an arbitrary point O is a line segment in a neighborhood of this point lying on such a cone and emanating from the vertex of this cone. This also implies that *there is a definite angle made by shortest arcs emanating from one point O; this is the angle made by segments to which shortest arcs reduce in the neighborhood of the point O.*

(2) *A shortest arc cannot pass through a point such that the complete angle around this point is $< 2\pi$.*

Indeed, two branches of a shortest arc emanating from such a point O divide a neighborhood of this point into two sectors. The sum of angles of these sectors is $< 2\pi$, and therefore the angle of at least one of them is $< \pi$. Developing this sector onto the plane, we see that such an angle can be cut off by drawing the segment connecting every two points X and Y on the sides of this sector. This segment XY is shorter than the sum of the segments OX and OY of the shortest arc; this contradicts the very definition of a shortest arc.

(3) *A shortest arc has a neighborhood that can be developed onto the plane; moreover, this shortest arc becomes a line segment under this development.*

If the endpoint of a shortest arc lies at the point O at which the complete angle is $< 2\pi$, then this neighborhood cannot be developed onto the plane. But we will draw,

[7] It is useful to keep in mind that by the continuous dependence of the distance $\rho(XY)$ on the points X and Y, $\gamma(x, y)$ is a continuous function of x and y.

© 2006 by Taylor & Francis Group, LLC

from the point O, a segment on the cone to which this neighborhood is isometric and will cut off the neighborhood along the segment.[8] Then the neighborhood can be developed onto the plane. The possibility of developing neighborhoods of the endpoints of a shortest arc is understood in exactly this sense.

Since we have proved that a shortest arc cannot pass through a point with the complete angle $< 2\pi$, then the complete angle around each interior point of a shortest arc is equal to 2π. Therefore, a neighborhood of such a point can be developed onto the plane. The definition of a shortest arc obviously implies that when a neighborhood of a point of it is developed onto the plane, the shortest arc itself goes to a line segment. Now, given a shortest arc L, we can surround each point of this arc by a neighborhood that can be developed onto the plane. We can choose a finite number of neighborhoods from these neighborhoods that cover the whole shortest arc L. Developing these neighborhoods onto the plane sequentially, we develop the whole neighborhood of the shortest arc L, which is formed by them, and since L itself is developed onto a line segment on each small part of it, the whole L also goes to a line segment.

(4) *If two shortest arcs AB and AC emanating from the same point A coincide on some segment AD, one of these arcs is a part of another, i.e., the nonoverlapping condition for shortest arcs holds.*

If two shortest arcs AB and AC coincide on some segment but not one of them is contained in the other, then, obviously, there is a segment AD on which these shortest arcs coincide, and then, on the segments DB and DC, they go away from one another. If we develop a neighborhood of the shortest arc AB onto the plane, then this shortest arc goes to a line segment. At the same time, the shortest arc AC also becomes a line segment (at least that part of this arc, which lies in the neighborhood developed). Of course, two line segments cannot coincide on some part AD and then diverge. Hence this is not possible for the shortest arcs AB and AC either as claimed.

Along with these simple remarks on shortest arcs, we need a more detailed result.

Lemma 1. *Let a shortest arc AB and a point C not lying on AB be given in some compact domain G of a manifold with polyhedral metric of positive curvature such that each point X on AB can be connected with C by a shortest arc lying in the domain G. The shortest arcs CX can go from the point X to one or both sides of AB. Then the following assertions hold for these shortest arcs*:

1. *if points X converge to B, remaining distinct from B, then the angles ξ between CX and AX converge to some limit, and this limit does not exceed the angle α between AB and each given shortest arc BC;*[9] *that is,* $\lim_{X\to B} \xi \geq \alpha$ *(Fig. 31);*

[8] The operation of cutting off some domain along some curve will be often used in what follows; therefore, we here give its rigorous definition. Let L be a curve of finite length without multiple points in a domain G. The cutting off of the domain G along the curve L means the replacing of the domain G by another domain G' containing two curves L_1 and L_2 with common endpoints such that if we identify the points on the curves L_1 and L_2, which are equidistant from the endpoints (counting along the curves L_1 and L_2), G' transforms into a domain isometric to G.

[9] There can exist several shortest arcs of B and C.

© 2006 by Taylor & Francis Group, LLC

2. *if a sequence of points X_n converges to B and the shortest arcs CX_n go from X_n to the same side of AB then these shortest arcs converge to some arc $\overline{BC}$ connecting the points B and C. (Hence there exist at most two such shortest arcs $\overline{CB}$ and $\overline{\overline{CB}}$, which are the limits of the shortest arcs CB as $X \to B$ $(X \neq B)$. One of the shortest arcs $\overline{CB}$ and $\overline{\overline{CB}}$ is the limit of the shortest arcs CX, which go from AB to one side, and the other is the limit of the shortest arcs CB, which go from AB to the other side. The sides of the shortest arc AB differ from each other in an obvious way, since a neighborhood of a shortest arc can be developed onto the plane, and, moreover, this shortest arc itself becomes a line segment.*)

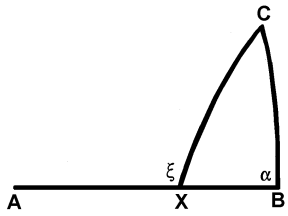


Fig. 31

Proof. Take a neighborhood U of the shortest arc AB that can be developed onto the plane. Under this procedure, we do not distinguish between the neighborhood U itself and its points on the one hand and the developed neighborhood and its points on the other hand. If the complete angle around the point B (or A) is less than 2π, we cut the neighborhood U along the shortest arc emanating from B and making equal angles with AB to both sides. In this case, although the neighborhood U developed onto the plane does not surround the point B, this neighborhood is symmetric with respect to a line segment onto which the shortest arc AB is developed.

Consider a sequence of points X_n other than B and convergent to B; let each point X_n be connected with C by a shortest arc CX_n in such a way that these shortest arcs CX_n converge to some shortest arc $\overline{BC}$. Segments of shortest arcs CX_n and $\overline{BC}$ contained in the neighborhood U go to line segments under the development onto the plane; the convergence of shortest arcs implies the convergence of these segments. This obviously implies that if ξ_n is the angle between CX_n and AX_n and $\overline{\alpha}$ is the angle between $\overline{BC}$ and AB, we have

$$\overline{\alpha} = \lim_{n\to\infty} \xi_n. \tag{1}$$

Now let BC be a shortest arc of B and C, and let α be the angle between this arc and AB. Let us prove that

$$\alpha \geq \overline{\alpha} = \lim_{n\to\infty} \xi_n. \tag{2}$$

Assume the contrary, i.e., $\alpha < \overline{\alpha}$. Then there exists $\delta > 0$ such that for all sufficiently large N, we have

$$\xi_n > \alpha + \delta.$$

Let l and l_n be segments of the shortest arcs BC and CX_n which lie in the neighborhood U. The line segments l and l_n making the angles α and ξ_n with AB and AX_n correspond to them in the neighborhood U developed onto the plane. There are two possibilities for these segments:

© 2006 by Taylor & Francis Group, LLC

1. they lie to one side of AB (Fig. 32);
2. they lie to different sides of AB (Fig. 33).

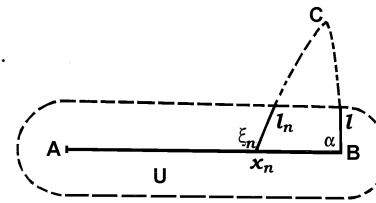


Fig. 32

In the first case, the fact that $\xi_n > \alpha + \delta$ implies obviously that whenever the point X_n is sufficiently close to B, the segments l and l_n intersect each other. This means that the shortest arcs BC and CX_n intersect each other. But we have shown above that the nonoverlapping condition for shortest arcs holds for a polyhedral metric of positive curvature; by this condition (by Theorem in Sec. 3 of Chapter II), two shortest arcs emanating from a common C cannot intersect one another. Hence, in the first case of location of the segments l and l_n, the assumption that $\alpha < \lim_{n\to\infty} \xi_n$ leads to a contradiction. Thus, it should be $\alpha \geq \lim_{n\to\infty} \xi_n$.

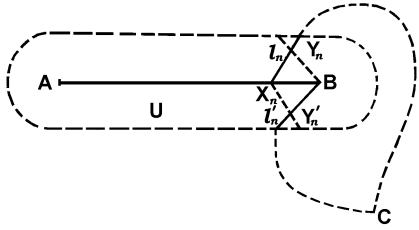


Fig. 33

We now consider the second case in which the segments l and l_n lie to different sides of AB.

Construct the segments l' and l'_n that are symmetric to the segments l and l_n with respect to AB (Fig. 33). Then the angle between AB and l' is also equal to α, and since $\xi_n > \alpha + \delta$, the segments l' and l_n intersect each other whenever the point X_n is sufficiently close to B. Show that this is impossible.

If the segments l' and l_n intersect each other at the point Y_n, then by symmetry, the segments l and l'_n intersect each other at the symmetric point Y'_n. Each of the shortest arcs CX_n and CB is divided by these points into two segments, i.e.,

$$CX_n = CY_n + Y_nX_n, \quad CB = CY'_n + Y'_nB.$$

Meanwhile, consider the curve $CY_n + Y_nB$ formed by the segment CY_n of the shortest arc CX_n and the segment Y_nB. This curve connects the points C and B, and since CB is a shortest arc, we have

$$CY_n + Y_nB \geq CB = CY'_n + Y'_nB.$$

Since $Y_nB = Y'_nB$ by symmetry, we have

$$CY_n \geq CY'_n.$$

But the points Y_n and Y'_n play exactly the same role, and, therefore, it also should be that $CY'_n \geq CY_n$ (it is sufficient to consider the line $CY'_n + Y'_nX_n$ instead of the line $CY_n + Y_nB$ and to compare this line with the shortest arc CX_n). Hence the segments CY_n and CY'_n of the shortest arcs CX_n and CB are equal. Thus,

$$CB = CY'_n + Y'_nB = CY_n + Y_nB,$$

© 2006 by Taylor & Francis Group, LLC

i.e., the line $CY_n + Y_nB$ is also a shortest arc of B and C. But this shortest arc overlaps the shortest arc CX_n along its segment CY_n, which contradicts the nonoverlapping condition for shortest arcs. This contradiction indicates that the segments l and l' cannot intersect one another. Hence, our assumption that $\alpha < \lim_{n\to\infty} \xi_n$ is not possible, and, thus, in the second case, we also have

$$\alpha \geq \lim_{n\to\infty} \xi_n.$$

Now let points X (which lie on AB and are different from B) converge to B in an arbitrary way. Assume that the angles ξ made by the shortest arcs CX and AX do not converge to any limit. Then we can choose two sequences of points X: X_n and X'_n, in such a way that (1) there exist the limits of the corresponding angles ξ_n and ξ'_n, and $\lim_{n\to\infty} \xi_n > \lim_{n\to\infty} \xi'_n$ and (2) the shortest arcs CX_n and CX'_n converge to some shortest arcs $\overline{CB}$ and $\overline{\overline{CB}}$, respectively. (The second condition is satisfied, since, by the compactness of the domain considered, we can choose a convergent subsequence from each sequence of shortest arcs.) If $\overline{\alpha}$ and $\overline{\overline{\alpha}}$ are the angles made by the shortest arcs $\overline{CB}$ and $\overline{\overline{CB}}$ with AB, then by formula (1),

$$\overline{\alpha} = \lim_{n\to\infty} \xi_n, \quad \overline{\overline{\alpha}} = \lim_{n\to\infty} \xi'_n.$$

And since $\overline{\overline{\alpha}}$ is the angle between the shortest arc $\overline{\overline{BC}}$ and AB, by formula (2), we have

$$\overline{\overline{\alpha}} \geq \lim_{n\to\infty} \xi_n,$$

i.e., $\lim_{n\to\infty} \xi'_n \geq \lim_{n\to\infty} \xi_n$; at the same time, $\lim_{n\to\infty} \xi'_n < \lim_{n\to\infty} \xi$ by condition. The so obtained contradiction shows that the angles converge to some limit. At the same time, by formula (2) $\lim_{X\to B} \xi$ does not exceed the angle α between AB and any shortest arc BC. Hence, the first assertion of the lemma is proved. Moreover, this also proves that each shortest arc BC presenting the limit of the sequence of shortest arcs CX makes the same angle $\alpha = \lim_{X\to B} \xi$ with AB. But, obviously, there are no more than two such shortest arcs; one of them makes the angle α to one side of AB, and the other one makes the same angle to the other side. Therefore, if the shortest arcs CX go to one side from AB, then they converge to the corresponding shortest arc BC.[10] This also proves the second assertion of our lemma.

Also, we use the following simple lemma of plane geometry.

Lemma 2. *If the corresponding sides of two quadrangles on the plane are equal, then either all their corresponding angles are equal or, in each of them, the angles greater than the corresponding angles in the other are separated by angles less than the corresponding angles in the other* (*Fig.* 34).

We include the limit cases of quadrangles in the set of all quadrangles, where some of their angles are equal to π or zero; i.e., we admit the degeneration of a quadrangle into a triangle with one point at the side taken as a vertex or even into

[10] They cannot diverge since, in view of the compactness of the domain G, we can choose a convergent subsequence from each sequence of shortest arcs CX_n.

© 2006 by Taylor & Francis Group, LLC

a segment, which should be considered here as consisting of four sides: one pair of sides overlaps the other.

Proof. Let $A_1B_1C_1D_1$ and $A_2B_2C_2D_2$ be two convex quadrangles on the plane. The vertices of these quadrangles are denoted in cyclic order. Let their corresponding sides be equal, i.e., $A_1B_1 = A_2B_2$, and so on. If, say, the angle A_1 in the first quadrangle is greater than the angle A_2 in the second quadrangle, then the diagonal B_1D_1 is greater than the diagonal B_2D_2. This is implied by the following theorem: If for two triangles, two sides of one of them are equal to two sides of the other and the angles between these sides are not equal, then the greater side subtends the greater angle. And if B_1D_1 is greater than B_2D_2, then by the converse theorem, the angle C_1 is greater than the angle C_2.

The sum of angles of a quadrangle is always equal to 2π. Therefore, at least one of the angles B_1 or D_1 in the first quadrangle should be less than the corresponding angle in the second quadrangle. But if, say, the angle B_1 is less than the angle B_2, then applying the same theorem on triangles, we obtain that the angle D_1 is also less than the angle D_2; the lemma is proved.

Theorem. *The convexity condition holds in every compact domain in a manifold with polyhedral metric of positive curvature,* i.e., if two shortest arcs L and M emanating from the same point O lie in some compact domain G and if there is a shortest arc in G of every two of L and M, then the angle $\gamma(x, y)$ for these shortest arcs L and M turns out to be a nonincreasing function of x and y.

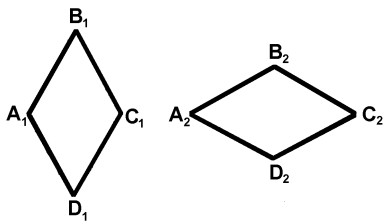


Fig. 34

(In fact, the requirement of a compact domain is superfluous: the convexity condition holds always in a polyhedral metric of positive curvature; but this more general theorem is not proved here.)

Proof. Obviously, it is sufficient to prove that for each constant y the quantity $\gamma(x, y)$ turns out to be a nonincreasing function of x, i.e., $\gamma(x_1, y) \geq \gamma(x_2, y)$ for $x_1 < x_2$. Then, since x and y play similar roles, we conclude that $\gamma(x, y_1) \geq \gamma(x, y_2)$ for $y_1 < y_2$, and hence, for any $x_1 < x_2, y_1 < y_2$, we have $\gamma(x_1, y_1) \geq \gamma(x_2, y_1) \geq \gamma(x_2, y_2)$, i.e., $\gamma(x, y)$ is a nonincreasing function of x and y.

Therefore, we choose an arbitrary point Y inside the shortest arc M, and fixing this point, assume that only the point X varies. Then, as y is constant, for the sake of brevity we can omit it as an argument and write $\gamma(x)$ instead of $\gamma(x, y)$; the length of the shortest arc XY will be denoted by $z(x)$. By the shortest arc XY, we mean any shortest arc connecting the points X and Y and passing in the given compact domain G.

Let $\overline{\alpha}(x)$ be the angle between the shortest arc XY and the segment OX of the shortest arc L. Since there can be several shortest arcs XY for a fixed position of X, $\overline{\alpha}(x)$ is, in general, a multivalued function of x; unless it is stated otherwise, by $\overline{\alpha}(x)$ we can mean the angle between one of the shortest arcs XY and the segment

© 2006 by Taylor & Francis Group, LLC

OX of the shortest arc L (Fig. 35). The corresponding angle in the plane triangles with the sides x, y, and $z(x)$, i.e., the angle between the sides equal to x and $z(x)$ will be denoted by $\alpha(x)$. This $\alpha(x)$ is a single-valued continuous function of x, since the distance $z(x) = XY$ is a single-valued continuous function of X.[11]

After these preliminary remarks, we proceed directly to the proof of the theorem. We will simultaneously prove the following two assertions:

(A) $\gamma(x)$ is a nonincreasing function of x;

(B) $\alpha(x) \le \overline{\alpha}(x)$ for all x, i.e., $\alpha(x)$ is not greater than any of the values of $\overline{\alpha}(x)$.

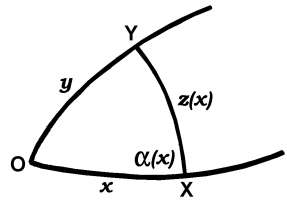


Fig. 35

This second assertion will be needed in the proof of the first.

We start with proving both assertions for a sufficiently small x.

If the point X is sufficiently close to O, then the shortest arcs OX, OY, and XY bound a triangle that can be developed onto the plane. Indeed, since the nonoverlapping condition for shortest arcs is valid, the segment OY of the shortest arc M is a unique shortest arc of O and Y, and as the point X tends to O, the shortest arc XY converges to OY. At the same time, a neighborhood of the shortest arc M can be developed onto the plane, and then the whole triangle OXY is developed onto the plane. This is the plane triangle on the plane with the sides x, y and $z(x)$, and its angle opposite the side z is $\gamma(x)$, while the angle opposite the side y is $\alpha(x)$. Until the triangle OXY can be developed onto the plane, it is obvious that $\gamma(x)$ remains constant (γ is equal to the angle at O in the triangle OXY), while $\alpha(x)$ is equal to the angle $\overline{\alpha}(x)$ made by OX and XY. Hence, for a sufficiently small x,

$$\gamma(x) = \text{const}, \quad \alpha(x) = \overline{\alpha}(x),$$

i.e., both assertions (A) and (B) are true for small x. Assume that x_0 is the maximum number for which both assertions (A) and (B) hold for the interval $0 < x < x_0$, and let us prove that x_0 should be equal to the length a of the shortest arc L, i.e., we prove that both assertions hold for the whole shortest arc L.

To prove this, we assume the contrary, i.e., assume that $x_0 < a$, so that the corresponding point X_0 lies inside the shortest arc L. Since $\gamma(x)$ is a continuous function of x, this function is also nonincreasing on the whole right-closed interval $0 < x \le x_0$.

Inequality (A) also holds for $x = x_0$. Indeed, since $\alpha(x)$ is a continuous function, we have

$$\alpha(x_0) = \lim_{x \to x_0} \alpha(x). \tag{3}$$

[11] It is possible that the shortest arc XY overlaps L and M. Then $\gamma = \pi$ and $\alpha(x) = \overline{\alpha}(x) = 0$. In exactly the same way, it is even possible that L and M overlap each other. (However, in this case, obviously, $\gamma(x)$ is always equal to zero, and thus, the theorem is true.) Not excluding these singularities, however, we do not focus our attention on them. Although our arguments are presented in such a way as if these cases do not hold, they remain valid in these specific cases.

© 2006 by Taylor & Francis Group, LLC

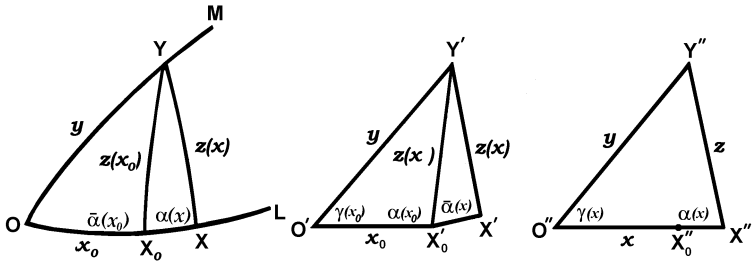


Fig. 36

On the other hand, by Lemma 1, when the points X converge to X_0 from the left, the angles $\overline{\alpha}(x)$ converge to the limit, which is not greater than any of the values $\overline{\alpha}(x_0)$, i.e.,

$$\lim_{x \to x_0} \overline{\alpha}(x) \le \overline{\alpha}(x_0). \tag{4}$$

Since $\alpha(x) \le \overline{\alpha}(x)$ for $x < x_0$ by assumption, (3) and (4) imply

$$\alpha(x_0) \le \overline{\alpha}(x_0).$$

Thus, on the whole right-closed interval $0 < x \le x_0$, $\gamma(x)$ is a nonincreasing function; and the angles $\alpha(x)$ and $\overline{\alpha}(x)$ satisfy inequality (B), i.e., $\alpha(x) \ge \overline{\alpha}(x)$.

Now let us prove that if x is sufficiently close to x_0 and is greater than x_0, then in exactly the same way,

$$\gamma(x_0) \ge \gamma(x) \quad \text{and} \quad \alpha(x) \ge \overline{\alpha}(x_0).$$

To prove this, we consider a point X lying on L to the right of X_0. According to Lemma 1, there exist at most two shortest arcs X_0Y, which are limits of the shortest arcs XY when X tends to X_0 from the right. The shortest arcs XY lying on one side of the shortest arc L converge always to the corresponding shortest arc X_0Y. Therefore, if the point X is sufficiently close to X_0 from the right, then the shortest arc XY is sufficiently close to the corresponding shortest arc X_0Y. Since a neighborhood of this shortest arc X_0Y can be developed onto the plane, the shortest arcs X_0Y and XY bound the triangle X_0XY which can be developed onto the plane. The angle at the vertex X_0 in this triangle is equal to $\pi - \overline{\alpha}(x_0)$.

Let us construct the triangle $O'X_0Y'$ in the plane with sides equal to x_0, y, and $z(x_0)$:

$$O'X_0' = OX_0 = x_0, \quad O'Y' = OY = y, \quad X_0'Y' = X_0Y' = z(x_0).$$

Adjoin the triangle X_0XY to this triangle along the side $X_0'Y'$, which is developed onto the plane, in such a way that its vertex X_0 coincides with X_0' and the vertex Y coincides with Y_0'. We obtain the quadrangle $O'X_0'X'Y'$ (Fig. 36).

Let us prove that all angles of this quadrangle do not exceed π. This is obvious for the angles at the vertices O' and X'. The angle at the vertex X_0' consists of the angles of the triangles $O'X_0'Y'$ and $X'X_0'Y'$. These angles are equal to $\alpha(x_0)$ and

© 2006 by Taylor & Francis Group, LLC

$\pi - \alpha(x_0)$, respectively. But since $\alpha(x_0) \le \overline{\alpha}(x_0)$, the sum of these angles is not greater than π. (It is possible that this sum is equal to π; then our quadrangle has one angle equal to π.) Finally, since $O'X_0' + X_0'X' = OX$, $O'Y' + Y'X' = OY + YX$, and OX is a shortest arc, we have

$$O'X_0' + X_0' \le O'Y' + Y'X'.$$

Therefore, the angle at the vertex Y' is also not greater than π, since otherwise the reverse inequality holds.

Hence all angles of the quadrangle $O'X_0'X'Y'$ do not exceed π, and this quadrangle is convex.

Also, we consider the triangle $O''X''Y''$ with the sides

$$O''X'' = x, \quad O''Y'' = y, \quad X''Y'' = z(x).$$

Since $x = OX = O'X_0' + X_0X'$, we can take a point X_0'' on the side of this triangle such that

$$O''X_0' = O'X_0', \quad X_0''X'' = X_0'X'.$$

Then our triangle passes into the quadrangle $O''X_0''X''Y''$ with sides equal to those of the quadrangle $O'X_0'X'Y'$. The angle at the vertex X_0' in the quadrangle $O''X_0''X''Y''$ is equal to π, and the angle at the vertex X_0' in the quadrangle $O'X_0'X'Y'$ is no greater than π. Therefore, by Lemma 2, for the angles at the vertices O and O'', X' and X'', we have the inequalities

$$\angle O' \ge \angle O'', \quad \angle X' \ge \angle X''.$$

But the angle $\angle O'$ is the angle in the triangle with sides x_0, y, and $z(x_0)$, i.e., is $\gamma(x_0)$, while $\angle O''$ is the angle in the triangle with sides x, y, and $z(x)$, i.e., is $\gamma(x)$. Hence

$$\gamma(x_0) \ge \gamma(x). \tag{5}$$

The angle $\angle X'$ is the angle in the triangle $X_0'X'Y'$, i.e., in the triangle X_0XY developed onto the plane, and so this angle is equal to $\overline{\alpha}(x)$. The angle $\angle X''$ is the angle in the triangle $O''X''Y''$ with sides x, y, and $z(x)$, i.e., the angle $\alpha(x)$. Consequently,

$$\overline{\alpha}(x) \ge \alpha(x). \tag{6}$$

Take now some $x_1 > x_0$ such that (6) holds for all $x < x_1$. Let us show that $\gamma(x)$ is a nonincreasing function on the whole interval from 0 to x_1. Since this is assumed to be true on the interval from 0 to x_0, it suffices to prove that $\gamma(x)$ does not increase on the closed interval from x_0 up to x_1. To this end, we take two values $x' < x''$ in this closed interval. By inequality (6), $\overline{\alpha}(x') \ge \alpha(x')$. Therefore, substituting x' for x_0 and x'' for x in the argument presented above, we obtain the following inequality which is completely analogous to inequality (5):

$$\gamma(x') \ge \gamma(x'').$$

Hence the function $\gamma(x)$ does not increase on the interval $(0, x_1)$ which is greater than the interval $(0, x_0)$, and $\overline{\alpha}(x) \ge \alpha(x)$ by inequality (6). Thus, x_0 cannot

© 2006 by Taylor & Francis Group, LLC

be the least upper bound of those intervals on which both assertions are true. Consequently, they are true on the whole shortest arc l; the theorem is proved.

Since a closed convex polyhedron is compact, has the polyhedral metric of positive curvature, and there is a shortest arc of its every two points, we can, applying the proved theorem to this polyhedron, say that *the convexity condition in the large holds on every closed convex polyhedron.* In contrast to "in the small," the words "in the large" mean that in order to apply the convexity condition, it is not necessary to restrict consideration to small segments of shortest arcs L and M, since there is a shortest arc of every two points on L and M.

3. The Convexity Condition for the Metric of a Convex Surface

We now prove that *the metric of every convex surface satisfies the convexity condition in the large.* To this end, we first prove a slightly weaker assertion.

Lemma 1. *Let F be a closed convex surface, and let $P_1, P_2, \ldots, P_n, \ldots$ be a sequence of closed convex polyhedra converging to F. Assume that two shortest arcs L and M on F emanating from a point O are limits of shortest arcs L_n and M_n on the polyhedra P_n emanating from points O_n, and, naturally, converge to O. Then the convexity condition holds for the shortest arcs L and M, i.e., the angle $\gamma(x, y)$, defined as in the convexity condition, is a nonincreasing function of x and y.*

(This lemma is weaker than the required result, since the shortest arcs L and M in this lemma are not arbitrary but rather the limits of shortest arcs on the polyhedra P_n.)

Proof. Let X and Y be two points inside the shortest arcs L and M, and let $\rho_F(OX) = x$, $\rho_F(OY) = y$, and $\rho_F(XY) = z(x, y)$. Since the shortest arcs L_n and M_n on the polyhedra P_n converge to L and M, at least for a sufficiently large n, there are points X_n and Y_n on them such that

$$\rho_{P_n}(O_nX_n) = x, \quad \rho_{P_n}(O_n, Y_n) = y.$$

These points converge to X and Y, respectively,[12] and, therefore, by the theorem on the convergence of metrics,

$$\lim_{n\to\infty} \rho_{P_n}(X_nY_n) = \rho_F(XY)$$

or

$$\lim_{n\to\infty} z_n(x, y) = z(x, y).$$

Therefore, the angle $\gamma_n(x, y)$ in the plane triangle with sides x, y and $z_n(x, y)$ converges to the angle $\gamma(x, y)$ in the triangle with sides x, y, and $\gamma(x, y)$. In other words, the functions $\gamma_n(x, y)$ converge to $\gamma(x, y)$. Since the convexity condition holds on polyhedra, the functions $\gamma_n(x, y)$ are nonincreasing. Therefore, the limit of these functions is also a nonincreasing function; as required.

[12] The limits of the points X_n and Y_n lie on L and M by convergence of L_n to L and M_n to M. The theorem on convergence of metrics implies that the distances of these limits to the point O are equal to x and y. Hence these limits are the points X and Y.

© 2006 by Taylor & Francis Group, LLC

We see that the convexity condition is easily extended to those shortest arcs L and M that are limits of shortest arcs on polyhedra. The difficulty as a whole lies in proving the fact that for *every* two shortest arcs L and M emanating from one point O on the surface F, there exists a convergent sequence of pairs of shortest arcs emanating from the point O_n on polyhedra P_n.

To prove this, we first prove three lemmas on shortest arcs satisfying the conditions of Lemma 1.

Lemma 2. *There is a definite angle between two shortest arcs L and M satisfying the conditions of Lemma 1; this angle is equal to zero if and only if these shortest arcs overlap each other, i.e., if one of them is a part of the other or they coincide.*

Proof. By the definition of the angle between the shortest arcs L and M given in Sec. 7 of Chapter I, this angle is the limit of the angles $\gamma(x, y)$ as $x, y \to 0$. But since $\gamma(x, y)$ is a nonincreasing function of x and y by Lemma 1, this limit does exist.

Assume that the shortest arcs L and M do not overlap. Then for these arcs, we can find two distinct points X_1 and Y_1 that are equidistant from O, i.e., points such that $\rho_F(OX_1) = x_1 = \rho_F(OY_1) = y_1$. Since the points X_1 and Y_1 are distinct, we have $\rho_F(X_1Y_1) = z(x_1, y_1) \neq 0$, and therefore the corresponding angle $\gamma(x_1, y_1)$ is different from zero. Since the angle $\gamma(x, y)$ decreases as x and y decrease, we have $\lim_{x,y\to 0} \gamma(x, y) \geq \gamma(x_1, y_1) > 0$. This means that the angle made by the shortest arcs L and M is different from zero. The lemma is proved.

Lemma 3. *Under the conditions of Lemma 1, the angle α between the shortest arcs L and M does not exceed the lower limit of the angles α_n between the shortest arcs L_n and M_n, that is,*

$$\alpha \leq \liminf_{n\to\infty} \alpha_n.$$

Proof. By the definition of angle, $\alpha = \lim_{x,y\to 0} \gamma(x, y)$, where $\gamma(x, y)$ means the same as above. Therefore, taking an arbitrary $\varepsilon > 0$, for sufficiently small x and y, we have

$$\alpha < \gamma(x, y) + \varepsilon. \tag{1}$$

At the same time, when proving Lemma 1, we have verified that the angles $\gamma_n(x, y)$ defined for the shortest arcs L_n and M_n converge to $\gamma(x, y)$. Therefore, for sufficiently large n and for the given x and y, we have

$$\gamma(x, y) < \gamma_n(x, y) + \varepsilon. \tag{2}$$

But each $\gamma_n(x, y)$ is a nonincreasing function of x and y, and, therefore,

$$\gamma_n(x, y) \leq \lim_{x,y\to 0} \gamma_n(x, y) = \alpha_n. \tag{3}$$

Comparing inequalities (1), (2), and (3), we find for a sufficiently large n that

$$\alpha < \alpha_n + 2\varepsilon.$$

Since ε is arbitrarily small, this implies

$$\alpha \leq \liminf_{n\to\infty} \alpha_n,$$

as required.

© 2006 by Taylor & Francis Group, LLC

Lemma 4. *Assume that three shortest arcs L, M, and N emanate from a point O on a convex surface F and they are the limits of shortest arcs L_n, M_n, and N_n emanating from points O_n on convex polyhedra P_n converging to F. Then the sum of angles between the shortest arcs L, M, and N cannot be greater than 2π.*

Proof. The sum of the angles α_n, β_n, and γ_n between L_n and M_n, M_n and N_n, and N_n and L_n cannot be greater than 2π, since the complete angle at a point on a convex polyhedron is always $\leq 2\pi$.[13] Consequently,

$$\alpha_n + \beta_n + \gamma_n \leq 2\pi. \tag{4}$$

By the previous lemma,

$$\liminf_{n\to\infty} \alpha_n \geq \alpha, \quad \liminf_{n\to\infty} \beta_n \geq \beta, \quad \liminf_{n\to\infty} \gamma_n \geq \gamma,$$

where α, β, and γ stand for the angles between the shortest arcs L and M, M and N, and N and L. Therefore, passing to the limit in inequality (4), we obtain

$$\alpha_n + \beta + \gamma \leq 2\pi;$$

as required.

We now prove the last lemma, which allows us to abstract the convexity condition from polyhedra to all convex surfaces.

Lemma 5. *Let a sequence of closed convex polyhedra P_n converge to a closed convex surface F. Let OA be a shortest arc on F that connects two points O and A, and let a point B lie inside the shortest arc OA. If two sequences of points O_n and B_n lying on the polyhedra P_n converge to the points O and B, then the shortest arcs O_nB_n on these polyhedra converge to the segment OB of the shortest arc OA.*

Proof. Assume by way of contradiction that the shortest arcs O_nB_n do not converge to the segment OB of the shortest arc OA. Then we can choose a subsequence $O_{n_i}B_{n_i}$ from them that contains no subsequence convergent to OB. At the same time, according to Theorem 3 of Sec. 1, we can choose a subsequence from the sequence of shortest arcs $O_{n_i}B_{n_i}$ that converges to a certain shortest arc on F connecting the points O and B. This shortest arc is hence different from OB; denote this shortest arc by $\overline{OB}$. Excluding those shortest arcs that are not included in the chosen subsequence converging to $\overline{OB}$, we can simply assume that the shortest arcs O_nB_n on the polyhedra P_n converge to $\overline{OB}$.

Let C be some point on OB not belonging to $\overline{OB}$, and let $\overline{CB}$ be a shortest arc on the surface F that connects the points C and B and is the limit of some sequence chosen from the shortest arcs C_nB_n on the polyhedra P_n. (The points B_n are the endpoints of those shortest arcs whose limit is $\overline{OB}$, while the points C_n are arbitrary points converging to C.) Again, excluding those shortest arcs that do not belong to the chosen subsequence, we can assume that the shortest arcs C_nB_n on the polyhedra P_n converge to $\overline{CB}$.

[13] $\alpha_n + \beta_n + \gamma_n$ can be different from the complete angle around the point O_n, since, e.g., the shortest arc N_n can go in the angle made by L_n and M_n. In this case, $\alpha_n = \beta_n + \gamma_n$, while $\alpha_n \leq \pi$; therefore, $\alpha_n + \beta_n + \gamma_n \leq 2\pi$.

© 2006 by Taylor & Francis Group, LLC

Finally, let $\overline{BA}$ be a shortest arc on F of B and A, presenting the limit of a certain subsequence chosen from the shortest arcs B_nA_n on the polyhedra P_n. (The points B_n are as above, while A_n are arbitrary points converging to A.) Again, excluding superfluous shortest arcs, we can assume that the shortest arcs B_nA_n on the polyhedra P_n converge to BA.

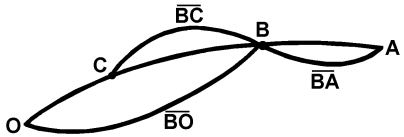


Fig. 37

Thus, we arrive at the following situation (Fig. 37).

Three shortest arcs $\overline{BA}$, $\overline{BC}$, and $\overline{BO}$ on the surface F emanate from the point B, which are the limits of the shortest arcs on the polyhedra P_n emanating from the points B_n. By Lemma 4, the sum of angles made by these shortest arcs cannot be greater than 2π. However, we shall show now that this sum should be greater than 2π. The contradiction obtained proves that our assumption that the segment OB of the shortest arc OA is not the limit of the shortest arcs O_nB_n on the polyhedra P_n is false.

Let us show that the line $\overline{OB}+\overline{BA}$ composed of the shortest arcs $\overline{OB}$ and $\overline{BA}$ is itself a shortest arc. Indeed, the lengths of $\overline{OB}$ and OB are equal, since both these lines are shortest arcs connecting the same points. (Segments of the shortest arcs OA are denoted without overbar.) By the same reason, $\overline{BA}$ and BA have equal lengths. Consequently, the length of $\overline{OB}+\overline{BA}$ is equal to the length of OA, i.e., to the length of the shortest arc OA; this means that the line $\overline{OB}+\overline{BA}$ itself is a shortest arc.

If two shortest arcs are prolongations of each other and together form a shortest arc, then the angle between these shortest arcs is equal to π, as, in this case, the angle $\gamma(x,y)$ in an auxiliary plane triangle is always equal to π, since $z=x+y$. Therefore, the angle made by $\overline{OB}$ and $\overline{BA}$ is equal to π.

In exactly the same way, we prove that the angle between $\overline{CB}$ and $\overline{BA}$ is also equal to π (it is sufficient to replace O by C in the above argument).

Finally, by Lemma 2, the angle between the shortest arcs $\overline{OB}$ and $\overline{CB}$ is different from zero, since they do not overlap each other by construction.

Thus, we have obtained that the sum of all three angles made by all shortest arcs $\overline{OB}$, $\overline{CB}$, and $\overline{AB}$ should be greater than 2π. This contradiction proves that our initial assumption is not true; the lemma is proved.

We now prove our main theorem.

Theorem 1. *The convexity condition in the large holds for every closed convex surface.*

Proof. Let L and M be two shortest arcs emanating from a point O on a closed convex surface F. Take two points X_0 and Y_0 inside them. Then by the above lemma, the segments OX_0 and OY_0 of the shortest arcs L and M can be represented as the limits of shortest arcs emanating from points O_n on some closed convex polyhedra P_n converging to F. But, in this case, by Lemma 1, the convexity condition holds for the shortest arcs OX_0 and OY_0. In other words, the angle $\gamma(x,y)$ defined for the shortest arcs L and M turns out to be a nonincreasing function on the interval $0<x\le x_0, 0<y\le y_0$ corresponding to the segments OX_0 and OY_0.

© 2006 by Taylor & Francis Group, LLC

But the points X_0 and Y_0 can be chosen arbitrarily close to the ends of the shortest arcs L and M. Therefore, $\gamma(x, y)$ is a nonincreasing function on the whole of these shortest arcs. Since $\gamma(x, y)$ is a continuous function, this is so up to the endpoints of the shortest arcs L and M. Thus, the theorem is proved.

If F is an arbitrary convex surface and O is a point on this surface, then taking a bounded part of F containing the point O and completing this part to a closed convex surface F_0, we can take an arbitrarily small neighborhood of the point O such that the distances on this neighborhood, measured on F and F_0, coincide with each other. Therefore, Theorem 1 implies the following theorem.

Theorem 2. *For each point on a convex surface, there exists a neighborhood in which the convexity condition holds.*

Now let F be an infinite complete convex surface, and let L and M be two shortest arcs on this surface emanating from point O. We can intersect F by a sphere S containing the shortest arcs L and M which is, at the same time, so large that the distance from this sphere to the shortest arcs L and M is arbitrarily large. (It is sufficient to take the point O as the center of this sphere and assume that the radius of the sphere is greater than the lengths of these shortest arcs by an arbitrarily given value.) The part of the surface F, that is cut off by the sphere S, together with the part of the sphere S, that is cut out by the surface F, forms the closed convex surface F_0; and if the sphere S is sufficiently large then all distances between the points of the shortest arcs L and M are the same on F_0 and F. Therefore, Theorem 1 also implies the following theorem.

Theorem 3. *The metric of an infinite complete convex surface satisfies the convexity condition in the large.*

It may happen that the shortest arcs emanating from the same point on an infinite complete convex surface prolong to infinity still remaining shortest, as, e.g., the meridians of every infinite surface of revolution. The angle $\gamma(x, y)$ is obviously a nonincreasing function on the whole of these infinitely extended shortest arcs.

In the proof of the fact that the convexity condition holds for a closed convex surface as given above, only the following two facts are essential:

1. the fulfillment of the convexity condition for a polyhedral metric of positive curvature;

2. the possibility of approximation of the metric of a convex surface by the metrics of convex polyhedra, i.e., by polyhedral metrics of positive curvature.

Therefore, generalizing our conclusions, we arrive at the fact that the convexity condition holds in each manifold with intrinsic metric which is the limit of polyhedral metrics of positive curvature. The most general result obtainable here can be formulated precisely as follows.

Let the intrinsic metric $\rho(XY)$ be given on some manifold R. Assume that this metric is such that in this manifold, there are polyhedral metrics of positive curvature $\rho_n(XY)$ converging to this metric, i.e., $\rho(XY) = \lim \rho_n(XY)$ for all X and Y. Finally, let L and M be two shortest arcs in the sense of the metric ρ

© 2006 by Taylor & Francis Group, LLC

emanating from the same point O such that any two points X and Y on them can be connected by a shortest arc. Then the angle $\gamma(x, y)$ defined for them in the same way as it was done in the formulation of the convexity condition is a nonincreasing function of x and y.

Since in every manifold with intrinsic metric two sufficiently close points, or any two points if the metric is complete, can be connected by a shortest arc, the theorem formulated above implies the following.

The convexity condition in the small holds in each manifold with intrinsic metric that presents the limit of polyhedral metrics of positive curvature. If this metric is complete, then the convexity condition in the large is valid.

These theorems especially stress a purely geometric character of all our preceding conclusions. Their complete proofs are obtainable rather easily on the basis of the arguments similar to those applied above. If desired, the reader can provide these proofs by himself or herself. As we are dealing with convex surfaces, these general theorems are beyond the scope of our discussion and will not be used further.

4. Consequences of the Convexity Condition

We have already mentioned that the convexity condition is characteristic of just convex surfaces and that their intrinsic geometry can be developed solely on the basis of this condition if not mentioning the general properties of every intrinsic metric. But this is not always the simplest and shortest way, and, therefore, we do not take it sometimes. Here, we derive a number of important theorems of intrinsic geometry of convex surfaces which immediate follow from the convexity condition.

Theorem 1. *There is a definite angle between two shortest arcs emanating from the same point on a convex surface which is other than zero, except for the trivial case in which these shortest arcs overlap.*

Proof. By the definition of angle between two shortest arcs, this angle is the limit of the angle $\gamma(x, y)$ in the convexity condition as x and y tend to zero. By the convexity condition, $\gamma(x, y)$ does not decrease as x and y decrease, and, therefore, $\lim_{x,y\to 0} \gamma(x, y)$ exists.

Assume that the angle between two shortest arcs L and M is equal to zero; i.e., $\lim_{x,y\to 0} \gamma(x, y) = 0$. Then, since $\gamma(x, y)$ does not increase as x and y increase, we have $\gamma(x, x) = 0$ for any x. But $\gamma(x, x)$ is the angle in the plane equilateral triangle with a lateral side x and the base equal to the distance between the two points on the shortest arcs L and M each at the distance x from the origin. Therefore, these points coincide for each x, i.e., the shortest arcs L and M overlap.

Theorem 2. *The nonoverlapping condition for shortest arcs holds on a convex surface, i.e., if two shortest arcs L and M emanating from a common point O coincide on some segment OA, then one of these arcs is a part of the other.*

Proof. Assume that the shortest arcs L and M emanating from a common point O coincide on some segment. Let OA be the maximal segment on which these arcs coincide; this segment obviously exists since shortest arcs are closed sets. If OA is one of the shortest arcs L and M, then the theorem is true.

© 2006 by Taylor & Francis Group, LLC

Assume that the segment OA coincides with neither of the shortest arcs L and M. Then there are two distinct points X and Y on these arcs that are equidistant from O. If the points X and Y are sufficiently close to the point A, then these points can be connected by shortest arcs, so we can use the convexity condition. Put $OA = a$ and $OX = OY = x$; moreover, $x > a$ by the choice of X and Y. Then by the convexity condition, $\gamma(x, y) \le \gamma(a, a)$. But $\gamma(a, a) = 0$; and, therefore, $\gamma(x, x) = 0$; by the definition of the angle γ, this implies $\rho(XY) = 0$. Consequently, all points X and Y of the shortest arcs L and M that are equidistant from O coincide. This means that the segment OA cannot be the maximum segment on which the shortest arcs L and M overlap each other if this segment is not one of these arcs. The theorem is proved.

Since the nonoverlapping condition for shortest arcs holds for convex surfaces, all consequences of this condition obtained in Sec. 3 of Chapter II also hold; the results of Secs. 3–6 will be needed later.

Theorem 3. *The angles between the sides of a triangle on a convex surface are not less than the corresponding angles of the plane triangle with sides of the same length.*[14] (The corresponding angles are those between the corresponding equal sides.)

We prove this theorem under the assumption that it is possible to apply the convexity condition to sides of this triangle. Since we have proved that this condition holds on an arbitrary convex surface "in the small" and on a complete surface "in the large," this theorem will thus be proved for sufficiently small triangles on an arbitrary convex surface and for all triangles on complete surfaces. In fact, this theorem holds without any restrictions, but its complete proof requires additional arguments.

Proof. Let ABC be a triangle on a convex surface with sides $AB = c$, $AC = b$, and $BC = a$. Let α be the angle between the sides AB and AC of this triangle, and let α_0 be the corresponding angle of the plane triangle with sides of the same length. Take variable points X and Y on the sides AB and AC and define the angle $\gamma(x, y)$ of the convexity condition. Then, obviously,

$$\alpha_0 = \gamma(b, c), \tag{1}$$

and, at the same time,

$$\alpha = \lim_{x,y\to 0} \gamma(x, y). \tag{2}$$

By the convexity condition, the angle $\gamma(x, y)$ is a nonincreasing function of x and y. Therefore,

$$\gamma(b, c) \le \lim_{x,y\to 0} \gamma(x, y); \tag{3}$$

this implies $\alpha_0 \le \alpha$; this is what we were required to prove.

[14] The angle between two sides of a triangle is between shortest arcs, and so this angle exists. This angle should be distinguished from the angle of the triangle, i.e., the angle of the sector that is bounded by sides and lies in the triangle. This angle will be defined in Chapter IV.

© 2006 by Taylor & Francis Group, LLC

In the proof of the following theorems, we lean on Theorem 3 rather than the convexity condition, since in what follows, in the study of an abstract metric of positive curvature, we can use precisely the result of Theorem 3.

Theorem 4. *Let a sequence of convex surfaces F_n converge to a convex surface F. Let shortest arcs L_n and M_n emanating from points O_n on the surfaces F_n converge to two shortest arcs L and M emanating from a common point O on the surface F (of course, in addition, the points O_n converge to the point O). If α is the angle made by L and M and α_n is the angle between L_n and M_n then*

$$\alpha \leq \liminf_{n\to\infty} \alpha_n.$$

(In particular, all the surfaces F_n can coincide with F, and then we obtain the theorem on convergence of angles on a convex surface. The fact that $\lim_{n\to\infty} \alpha_n$ may fail to exist, and even if it does, it can be greater than α, is shown in Sec. 10 of Chapter I by the example of a cube.)

Proof. Without loss of generality, we can assume that the surfaces F_n and F are closed. To this end, it is sufficient to cut out finite domains from these surfaces F_n and F which contain the points O_n and O and to complete these domains in an appropriate way up to closed surfaces. Under this procedure, the metric near the points O_n and O does not change, so the angles between shortest arcs do not change either.

Let $\gamma(x, y)$ and $\gamma_n(x, y)$ be the angles defined for the shortest arcs L, M and L_n, M_n the way it was done in the convexity condition. By the definition of angle, the angle between L and M is

$$\alpha = \lim_{x,y\to 0} \gamma(x, y).$$

Therefore, given $\varepsilon > 0$, we can take x and y so small that

$$\alpha < \gamma(x, y) + \varepsilon. \tag{4}$$

Let X and Y be those points on L and M for which $\rho(OX) = x$ and $\rho(OY) = y$, where ρ is the distance on the surface F. Let X_n and Y_n be points on L_n and M_n that converge to the points X and Y. Then by the theorem on convergence of metrics,

$$\begin{aligned} x_n = \rho_n(OX_n) &\to \rho(OX) = x, \\ y_n = \rho_n(OY_n) &\to \rho(OY) = y, \\ \rho_n(X_nY_n) &\to \rho(XY), \end{aligned} \tag{5}$$

where ρ_n is the distance on the surface F_n.

Since $\gamma_n(x, y)$ is the angle in the plane triangle with sides equal to $\rho_n(OX)$, $\rho_n(OY)$, and $\rho_n(XY)$ while $\gamma(x, y)$ is the angle in the plane triangle with sides $\rho(OX)$, $\rho(OY)$, and $\rho(XY)$, the convergences of (5) imply

$$\gamma(x, y) = \lim_{n\to\infty} \gamma_n(x_n, y_n). \tag{6}$$

© 2006 by Taylor & Francis Group, LLC

Connecting the points X_n and Y_n by a shortest arc, we obtain the triangle $O_nX_nY_n$ with the angle α_n made by the sides O_nX_n and O_nY_n which are segments of the shortest arcs L_n and M_n. Since $\gamma_n(x_n, y_n)$ is the angle in the plane triangles with sides equal to those of the triangle $O_nX_nY_n$, we have

$$\alpha_n \geq \gamma_n(x_n, y_n) \tag{7}$$

by Theorem 3. Using inequality (6), we conclude that

$$\liminf_{n\to\infty} \alpha_n \geq \lim_{n\to\infty} \gamma_n(x_n, y_n) = \gamma(x, y). \tag{8}$$

Comparing this with (4), we obtain

$$\liminf_{n\to\infty} \alpha_n + \varepsilon > \alpha.$$

Since ε is arbitrary,

$$\alpha \leq \lim_{n\to\infty} \alpha_n; \tag{9}$$

as required.

The result obtained is the first and main theorem on convergence of angles; other results will be obtained in Sec. 4 of Chapter IV, which studies the convergence of angles.

Theorem 5. *There is an angle in the strong sense between every two shortest arcs on a convex surface that emanate from a common point.*

According to the definition of Sec. 9 of Chapter I, this means the following.

Let L and M be two shortest arcs emanating from a common point O on a convex surface. Assume that there is a shortest join of every two points X and Y on these arcs. Let X_n and Y_n be a sequence of points on L and M which are other than O and such that (1) $X_n \to O$; (2) the shortest arcs X_nY_n converge to some part of the shortest arc M. Then, if $\gamma(x, y)$ is the angle in the convexity condition, $\lim_{n\to\infty} \gamma(x_n, y_n)$ exists and is equal to the angle α between the shortest arcs L and M.

It is possible that the points Y_n converge to O so that not only $x_n \to 0$ but also $y_n \to 0$. Then the shortest arcs X_nY_n converge to the point O, which is also a part of the shortest arc M. This makes it clear that this theorem is a strengthening of the theorem on the existence of angle. This theorem shows that the limit of $\gamma(x, y)$ exists not only when the points x and y tend to zero but also in the case where at least one of the variables x and y tends to zero whenever the shortest arcs XY converge to a part of one of the shortest arcs M or L.

If the points Y_n converge to some interior point A of the shortest arc M, then the shortest arcs X_nY_n converge to the segment OA of the shortest arc M, i.e., condition (2) holds automatically for them. (This follows from Corollary 4 of the theorem in Sec. 3 of Chapter II, which is based on the overlapping condition for shortest arcs, and we have shown that the latter condition holds.) If the points Y_n converge to the end of the shortest arc M, then the shortest arcs X_nY_n may fail to converge to M; hence, in this case, condition (2) may fail to hold automatically. It is

© 2006 by Taylor & Francis Group, LLC

easy to verify that if this condition does not hold, the assertion of the theorem does not hold, either.[15] Thus, this condition is not only sufficient but is also necessary for $\lim_{n\to\infty}$ to be equal to α.

We now continue with the proof of our theorem.

Proof. We can assume that the points Y_n converge to a point other than O; since if they converge to O, our theorem reduces to the theorem on the existence of angle as we have just mentioned. Then the shortest arcs X_nY_n converge to some segment of the shortest arc M. Also, we can assume that the shortest arcs L and M do not overlap since the theorem is trivial in this case.

Since each shortest arc X_nY_n has the common point Y_n with M, the nonoverlapping condition for shortest arcs opens only the following two possibilities: either that the shortest arcs X_nY_n and M have no other common points or that X_nY_n overlaps M on the whole segment OY_n. In the latter case, denoting by ρ the metric of the surface, we have $\rho(X_nY_n) = \rho(OX_n) + \rho(OY_n)$. Therefore the plane triangle with such sides degenerates into a segment, and $\gamma(x_n, y_n) = \pi$. If this is true for at least one n, the angle α between the shortest arcs L and M is obviously equal to π. Then, taking those n_i for which $\gamma(x_{n_i}, y_{n_i}) = \pi$, we have

$$\lim_{n\to\infty} \gamma(x_{n_i}, y_{n_i}) = \alpha.$$

Therefore, it remains to consider a sequence of shortest arcs X_nY_n such that each of them has no common points with M but Y_n. Then, whenever the shortest arc X_nY_n gets sufficiently close to M, it makes sense to say that this shortest arc runs on a particular side of M.[16] We restrict exposition to the shortest arcs X_nY_n running on the same side of M. To avoid abusing notation, we will denote these arcs by the same subscript n (Fig. 38).

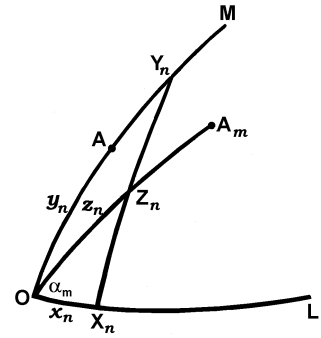


Fig. 38

Take some point A inside M and construct a sequence of points A_N that converge to A and lie on the same side of M as all X_nY_n. Since the point A is inside M, the shortest arcs OA_m converge to the segment OA of the shortest arc M, and for sufficiently large m, in an arbitrarily small neighborhood of M they run on to the same side as the shortest arcs X_nY_n.

Let α_m be the angle made by L and OA_m; then, by Theorem 4, which was proved above,

$$\liminf_{m\to\infty} \alpha_m \geq \alpha.$$

Therefore, given an arbitrary positive ε, we can find m satisfying

$$\alpha_m + \varepsilon > \alpha. \tag{10}$$

[15] Indeed, if $X_nY_n \to M'$ then by our theorem, $\lim \gamma(x_n, y_n)$ is equal to the angle between L and M'. If M' is not included in M, it is easy to show that the angle between L and M' differs from the angle between L and M.

[16] Since a shortest arc is homeomorphic to a line segment, we can map this arc, together with its neighborhood, into a plane in such a way that this shortest arc goes to the segment, and then the notion of the location on a particular side of this arc has a concrete sense.

© 2006 by Taylor & Francis Group, LLC

Specify such an m and fix it in what follows. For a sufficiently large n, the shortest arc X_nY_n gets arbitrarily close to M, and therefore the point A_m runs outside the triangle OX_nY_n. And since the shortest arc OA_m goes on the same side of M as the shortest arcs X_nY_n, the latter intersect OA_m for a sufficiently large n. Let Z_n be the point of intersection of the shortest arc X_nY_n with OA_m. Since the shortest arcs X_nY_n converge to a segment of the shortest arc M and OA_m has no common points with M but O, the points Z_n converge to O.

Put $\rho(OZ_n) = z_n$, and let $\gamma_m(x_n, z_n)$ be the angle defined for the points X_n and Z_n in the same way as in the convexity condition. Then, since X_n and Z_n converge to O, we have

$$\lim_{n\to\infty} \gamma_m(x_n, z_n) = \alpha_m. \tag{11}$$

Choose some n and construct the plane triangles $O'X'_nZ'_n$ and $O'Z'_nY'_n$ with sides

$$\begin{aligned} O'X'_n = OX_n, \quad O'Z'_n = OZ_n, \quad X'_nZ'_n = X_nZ_n, \\ O'Y'_n = OY_n, \quad Z'_nY'_n = Z_nY_n, \end{aligned} \tag{12}$$

where, for the sake of brevity, the symbols OX_n, etc. stand for the lengths of the corresponding shortest arcs. We join these triangles in such a way that the side $O'Z'_n$ is their common part and they lie on the opposite sides of it. Then these triangles form the quadrangle $O'X'_nZ_nY'_n$.

Formulas (12) imply

$$O'X'_n + O'Y'_n = OX_n + OY_n, \quad X''_nZ'_n + Z'_nY'_n = X_nZ_n + Z_nY_n.$$

The point Z_n lies on the shortest arc X_nY_n, and, therefore,

$$X_nZ_n + Z_nY_n = X_nY_n \le OX_n + OY_n.$$

Hence

$$O'X'_n + O'Y'n \ge X'_nZ'_n + Z'_nY'_n.$$

This implies that the angle δ at the vertex O' of our quadrangle does not exceed π (otherwise, we have the reverse inequality). Since this is true, the angle δ is also the angle of the triangle $O'X'_nY'_n$. The angle at O' in the triangle $O'X'_nZ'_n$ is nothing else but $\gamma_m(x_n, z_n)$. Therefore,

$$\gamma_m(x_n, y_n) \le \delta. \tag{13}$$

At the same time, we construct the triangle $O''X''_nY''$ with sides

$$O''X''_n = OX_n, \quad O''Y''_n = OY_n, \quad X''_nY''_n = X_nY_n. \tag{14}$$

The angle at the vertex O'' of this triangle is nothing else but $\gamma(x_n, y_n)$.

We see from (14) and (12) that the sides that make up this angle are equal to the sides that make up the angle δ in the triangle $O'X'_nY'_n$. For the opposite sides, the inequality

$$X''_nY''_n \ge X'_nY'_n \tag{15}$$

© 2006 by Taylor & Francis Group, LLC

holds (since $X_n''Y_n'' = X_nY_n = X_nZ_n + Z_nY_n = X_n'Z_n' + Z_n'Y_n' \geq X_n'Y_n'$). This implies that for the angles $\gamma(x_n, y_n)$ and δ, the following analogous inequality holds:

$$\gamma(x_n, y_n) \geq \delta. \tag{16}$$

Comparing these inequalities with inequality (13), we obtain

$$\gamma(x_n, y_n) \geq \gamma_m(x_n, z_n). \tag{17}$$

This inequality holds for any n. Therefore, passing to the limit and using the fact that, by (11), the limit of the angles $\gamma_m(x_n, z_n)$ is equal to the angle α_m, we obtain

$$\liminf_{n\to\infty} \gamma(x_n, y_n) \geq \alpha_m.$$

Since $\alpha_m + \varepsilon > \alpha$, by inequality (10) we have

$$\liminf_{n\to\infty} \gamma(x_n, y_n) \geq \alpha - \varepsilon.$$

But ε is chosen arbitrarily, therefore,

$$\liminf_{n\to\infty} \gamma(x_n, y_n) \geq \alpha. \tag{18}$$

Since α is the angle made by the sides of the triangle OX_nY_n and $\gamma(x_n, y_n)$ is the corresponding angle in the plane triangle with sides of the same length, $\gamma(x_n, y_n) \leq \alpha$ by Theorem 4. Consequently,

$$\limsup_{n\to\infty} \gamma(x_n, y_n) \leq \alpha. \tag{19}$$

Comparing this inequality with (8), we see that the limit of the angle $\gamma(x_n, y_n)$ exists and equals α, as required.

Note that inequality (19) can be proved without reference to Theorem 3. Indeed, since $\gamma(x_n, y_n)$ is the angle of the plane triangle with sides x_n, y_n, and $z_n = X_nY_n$, we have

$$z_n^2 = x_n^2 + y_n^2 - 2x_ny_n \cos\gamma(x_n, y_n)$$

or

$$\cos\gamma(x_n, y_n) = \frac{y_n^2 - z_n^2 + x_n^2}{2x_ny_n} = \frac{y_n - z_n}{x_n}\frac{y_n + z_n}{2y_n} + \frac{x_n}{2y_n}. \tag{20}$$

And since $x + n \to 0$ and $y_n \to z_n$, we have

$$\liminf_{n\to\infty} \cos\gamma(x_n, y_n) = \liminf_{n\to\infty} \frac{y_n - z_n}{x_n}. \tag{21}$$

We now take a sequence of points Y_n' on the shortest arc M which converges to O so that $x_n/y_n' \to 0$, where $y_n' = OY_n'$. Then the limit of the angles $\gamma(x_n, y_n')$ exists and equals the angle α between L and M. Since $\gamma(x_n, y_n')$ is the angle of the plane triangle with sides x_n, y_n', and $z_n' = X_nY_n'$, for this triangle, by analogy with (20), we have

$$\cos\gamma(x_n, y_n') = \frac{y_n' - z_n}{x - n}\frac{y_n' + z_n'}{2y_n'} + \frac{x_n}{2y_n'}. \tag{22}$$

© 2006 by Taylor & Francis Group, LLC

By condition, $x_n/2y'_n \to 0$; at the same time, by the triangle inequality, $|z'_n - y'_n| \leq x_n$, i.e., have $(y'_n + z_n)/2y'_n \to 1$. Therefore, (22) implies

$$\cos\alpha = \lim_{n\to\infty} \cos\gamma(x_n, y'_n) = \lim_{n\to\infty} \frac{y'_n - z'_n}{x_n}. \tag{23}$$

We now note that by the triangle inequality, $OY_n - OY'_n = Y_nY'_n \geq X_nY_n - X_nY'_n$; i.e., $y_n - y'_n \geq z_n - z'_n$ or $y_n - z_n \geq y'_n - z'_n$. By this inequality, the comparison of formulas (21) and (23) leads to the conclusion that

$$\liminf_{n\to\infty} \cos\gamma(x_n, y_n) \geq \cos\alpha$$

or

$$\alpha \geq \limsup_{n\to\infty} \gamma(x_n, y_n).$$

This is exactly inequality (19); this inequality is hence proved on the basis of only the triangle inequality. Therefore, the theorem itself rests on the following three facts: (1) the existence of the angle; (2) the nonoverlapping condition for shortest arcs; (3) the possibility of drawing a shortest arc OA_m that makes with L the angle that is arbitrarily close to α. This remark will be useful in Sec. 1 of Chapter VIII, where we consider abstract manifolds.

© 2006 by Taylor & Francis Group, LLC

Chapter IV

THE ANGLE

1. General Theorems on Addition of Angles

In the previous chapter, we have proved that there is a definite angle between two arbitrary shortest arcs emanating from a common point on a convex surface. Here, we want to study the properties of the angle made by shortest arcs and, first of all, to prove the theorem on addition of angles. It turns out that this theorem is based only on the fact of existence of the angle itself; therefore, there is no necessity to assume that we speak about shortest arcs on a convex surface: we can in general deal with such curves in a manifold with intrinsic metric, between which there exist definite angles.[1] This general statement of the problem is of interest at least because the concept of angle is one of the basic concepts of geometry.

Recall the definition of angle between curves. Let L and M be two curves emanating from a point O in a certain manifold with intrinsic metric ρ. Let $X_t = X(t)$ $(0 \le t \le 1)$ and $Y_s = Y(s)$ $(0 \le s \le 1)$ be parameterizations of these curves; moreover, assume that $X_0 = Y_0 = 0$, and for a sufficiently small t and s, the points X_t and Y_s differ from zero if t and s do not vanish. Construct the plane triangle with sides equal to $\rho(OX_t)$, $\rho(OY_s)$, and $\rho(X_tY_s)$; denote by $\gamma(t, s)$ the angle of this triangle subtended by the side equal to $\rho(X_tY_s)$. The limit of this angle as t and s tend to zero, i.e., $\lim_{t,s\to\infty} \gamma(s, t)$, is called the angle between the curves L and M at the point O.

We will consider curves L_i emanating from the same point O of a manifold with intrinsic metric restricting exposition to a small neighborhood of the point O in which there is a shortest arc of every two points. We shall assume that these curves L_i are given in the parametric form $X_i = X(t_i)$ $(0 \le t_i \le 1)$ so that $X(0) = 0$ and $X_i(t_i) \ne 0$ for t_i other than zero. The condition that the point X_i tends to O is equivalent to the assertion that $t_i \to 0$. Further, $\gamma_{ij}(t_i, t_j)$ denotes the angle in the plane triangle with the sides equal to OX_i, OX_j, and X_iX_j, which lies opposite the side equal to X_iX_j; for brevity, we omit the symbol ρ. Finally, α_{ij} stands for the angle made by the curves L_i and L_j.

The result to be proved below reduces mainly to the following.

1. *If three curves L_1, L_2, and L_3 emanate from a point O and make angles α_{12}, α_{23}, and α_{31}, the sum of each two of these angles is no less than the third.*

2. *Moreover, if shortest arcs X_1X_3 intersect L_2 for X_1 and X_3 arbitrarily close to O, then $\alpha_{12} + \alpha_{23} = \alpha_{13}$.*

[1] Namely, the theorems to be stated are true in every metric space.

© 2006 by Taylor & Francis Group, LLC

These properties of the angle are completely analogous to those of the angle between half-lines on the plane. In order to prove these properties, it is convenient first to establish a few lemmas on the angles $\gamma_{ij}(t_i, t_j)$.

Lemma 1. *For every three curves L_1, L_2, and L_3 emanating from a common point O, we have*

$$\limsup_{t_1,t_3\to 0} \gamma_{13}(t_1, t_3) \le \limsup_{t_1,t_2\to 0} \gamma_{12}(t_1, t_2) + \limsup_{t_2,t_3\to 0} \gamma_{23}(t_2, t_3).$$

(The existence of the angles is thus not assumed.)

Proof. Set

$$\limsup_{t_i,t_j\to 0} \gamma_{ij}(t_i, t_j) = \overline{\alpha}_{ij}.$$

We have to prove that $\overline{\alpha}_{13} \le \overline{\alpha}_{12} + \overline{\alpha}_{23}$. If $\overline{\alpha}_{13} = 0$ then this inequality is trivial, and so we can assume that $\overline{\alpha}_{13} > 0$.

Take a sequence of pairs of points $X_1^{(n)}$ and $X_3^{(n)}$ on the curves L_1 and L_3 such that the angles $\gamma_{13}(t_1^{(n)}, t_3^{(n)})$ taken for these points converge to $\overline{\alpha}_{13}$. Since $\overline{\alpha}_{13} > 0$, we can assume that all $\gamma_{13}(t_1^{(n)}, t_3^{(n)}) > 0$ (it is sufficient to remove those pairs $X_1^{(n)}, X_3^{(n)}$ for which $\gamma_{13} = 0$). Moreover, we first assume that there are infinitely many angles different from π among the angles $\gamma_{13}(t_1^{(n)}, t_3^{(n)})$ and restrict consideration to only those n for which $\gamma_{13}(t_1^{(n)}, t_3^{(n)}) \ne \pi$. Finally, we take any point M on the curve L_2 that is different from O and restrict consideration to only those pairs $X_1^{(n)}$, $X_3^{(n)}$ for which $OX_1^{(n)}$ and $OX_3^{(n)} < OM$.

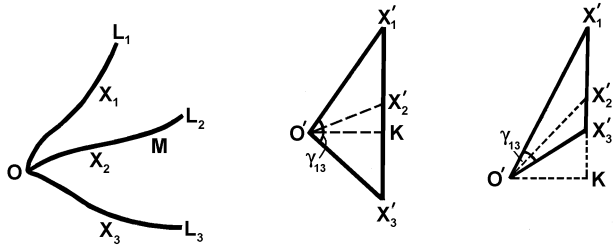


Fig. 39

Let X_1, X_3 be a certain pair of points of the chosen sequence. As usual, we construct the plane triangle $O'X_1'X_3'$ with the sides $O'X_1' = OX_1$, $O'X_3' = OX_3$, and $X_1'X_3' = X_1X_3$ (Fig. 39). The angle at the vertex O' of this triangle is precisely γ_{13}, and since γ_{13} is different from zero and π by assumption, this triangle does not degenerate to a segment. For definiteness, we assume that $O'X_1' \ge O'X_3'$. Let $O'K$ be the height of the triangle $O'X_1'X_3'$ dropped from the vertex O'; obviously, $0 < O'K < O'X'$.

Since the point $X_2(t_2)$ of the curve L_2 depends continuously on the parameter t_2, the distance OX_2 is a continuous function of t_2. This distance takes the values O and OM; since we assume that $OX_1 < OM$, there is a minimal value $t_2 = \tau_1$ for which

$$OX_2(\tau_1) = OX_1 = O'X_1'.$$

© 2006 by Taylor & Francis Group, LLC

Since $O'K < OX_1$, in exactly the same way we find that there is a maximal value $t_2 = \tau_2$ among those $t_2 < \tau_1$ for which

$$OX_2(\tau_2) = O'K.$$

To each value of the parameter t_2 in the closed interval $[\tau_2, \tau_1]$, we can put in correspondence the point $X_2'(t_2)$ on the segment KX_1' such that[2]

$$OX_2(t_2) = O'X_2'(t_2).$$

If the point X_3' lies between X_1' and K, then there is a maximal value of the parameter $t_2 = \tau_3$ in the closed interval $[\tau_2, \tau_1]$ for which $OX_2(\tau_3) = O'X_3'$. Then, when t_2 varies continuously from τ_1 up to τ_3, the point X_2' moves continuously from X_1' up to X_3'. In contrast, if the point K lies between X_1' and X_3' then there is a minimal value of the parameter $t_2 = \tau_4$ in the closed interval $[\tau_2, \tau_1]$ for which $OX_2(\tau_4) = O'X_3$. Then we can change the parameter t_2 from τ_1 to τ_2; under this procedure, the point $X_2'(t_2)$ moves continuously along the segment X_1K from X_1' to K, and then, if we again increase the value of t_2 from τ_2 to τ_4 taking the point $X_2'(t_2)$ on the ray KX_3', the point $X_2(t_2)$ moves continuously on the segment KX from K to X_3'.

Thus, in both cases of location of K, we obtain a process of continuous change of the parameter t_2 such that point X_2' moves continuously drawing the segment $X_1'X_3'$.

Consider the ratios

$$\frac{X_1'X_2'}{X_1'X_2' + X_2'X_3'} \quad \text{and} \quad \frac{X_1X_2}{X_1X_2 + X_2X_3}.$$

Each is a continuous function of t_2, and for the range of t_2 in question, the first of these ratios takes all values from 0 up to 1; the second remaining between 0 and 1. Therefore, both relations are equal for some value of t_2, and we have those points X_2 and X_2' for which

$$\frac{X_1'X_2'}{X_1'X_2' + X_2'X_3'} = \frac{X_1X_2}{X_1X_2 + X_2X_3}, \quad O'X_2' = OX_2. \tag{1}$$

But, by construction, $X_1'X_3' = X_1X_3$, and by the triangle inequality,

$$X_1'X_2' + X_2'X_3' = X_1'X_3' = X_1X_3 \le X_1X_2 + X_2X_3;$$

which, together with (1), yields

$$X_1'X_2' \le X_1X_2, \quad X_2'X_3' \le X_2X_3$$

[2]This is impossible if for some $t_2 \in [\tau_2, \tau_1]$, either $OX_2(t_2) < O'K$ or $OX_2 > O'X_1$. But both these possibilities are excluded. For example, if $OX_2(t_2) < O'K$ for $t_2 = \tau$, then by the continuous dependence of the distance $OX_2(t_2)$ on t_2 and by the fact that $OX_2(\tau_1) = OX_1 > O'K$, there is some $t_2 \in [\tau, \tau_2]$ such that $OX_2(t_2) = O'K$, so that τ_2 is not the maximum of those values $t_2 < \tau_1$ for which $OX_2(t_2) = O'K$; which would contradict the definition of τ_2. For a similar reason, it is impossible that $OX_2(t_2) > O'X_1'$ for $t_2 \in [\tau_2, \tau_1]$.

© 2006 by Taylor & Francis Group, LLC

Therefore, when we construct the plane triangles with sides OX_1, OX_2, X_1X_2 and OX_2, OX_3, X_2X_3, these triangles differ from the triangles $O'X_1'X_2'$ and $O'X_2'X_3'$ only by that the sides X_1X_2 and X_2X_3 are greater (not less) than $X_1'X_2'$ and $X_2'X_3'$; the other sides are the same, since

$$O'X + 1' = OX_1, \quad O'X_2' = OX_2, \quad OX_3' = OX_3.$$

At the same time, the angles of these triangles opposite the sides X_1X_2 and X_2X_3 are nothing else but γ_{12} and γ_{23}. Therefore,

$$\angle X_1'O'X_2' \le \gamma_{12}, \quad \angle X_2'O'X_3' \le \gamma_{23}.$$

Since the angle at O' in the triangle $O'X_1'X_3'$ is γ_{13}, these inequalities yield

$$\gamma_{13} \le \gamma_{12} + \gamma_{23}.$$

By the choice of the sequence of pairs X_1, X_2 we have $\lim \gamma_{13} = \overline{\alpha}_{13}$. Therefore, passing to the limit on the left-hand side and to the upper limit on the right-hand side, we obtain the required inequality

$$\overline{\alpha}_{13} \le \overline{\alpha}_{12} + \overline{\alpha}_{23}.$$

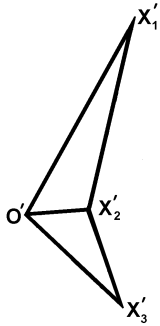


Fig. 40

We have assumed that there are infinitely many values different from π among the values $\gamma_{13}(t_1^{(n)}, t_2^{(n)})$. Assume now that all $\gamma_{13} = \pi$ (a finite number of values γ_{13} can be excluded). This means that

$$X_1^{(n)}X_3^{(n)} = X_1^{(n)}O + OX_3^{(n)}. \tag{2}$$

In this case, if the inequality

$$\overline{\alpha}_{13} = \pi \le \overline{\alpha}_{12} + \overline{\alpha}_{23}$$

fails, then there exists $\varepsilon > 0$ such that for all sufficiently large n and for small t_2, i.e., for $X_(t_2)$ sufficiently close to O, we have

$$\gamma_{12} \le \gamma_{23} < \pi - \varepsilon.$$

But then for a given n, we can take t_2 so small that the plane triangles $O'X_1'X_2'$ and $O'X_2'X_3'$ with sides equal to $OX_1^{(n)}$, $OX_2^{(n)}$, $X_1^{(n)}X_2^{(n)}$ and $OX_2^{(n)}$, $OX_3^{(n)}$, $X_2^{(n)}X_3^{(n)}$ taken together comprise a concave triangle as is displayed in Fig. 40 (the angle at O' of this triangle is equal to $\gamma_{12} + \gamma_{23}$, and since this angle is $< \pi - \varepsilon$, the angle at X is $> \pi$ for X_2' close to O').

But, in this case

$$X_1'X_2' + X_2'X_3' < O'X_1' + O'X_3',$$

i.e.,

$$X_1^{(n)}X_2^{(n)} + X_2^{(n)}X_3^{(n)} < OX_1^{(n)} + OX_3^{(n)},$$

or, by Eq. (2),

$$X_1^{(n)}X_2^{(n)} + X_2^{(n)}X_3^{(n)} < X_1^{(n)}X_3^{(n)}.$$

© 2006 by Taylor & Francis Group, LLC

But this contradicts the triangle inequality. Hence the assumption that in the case under consideration the inequality $\pi \le \overline{\alpha}_{12} + \overline{\alpha}_{23}$ does not hold is not true. Thus, in all the cases,

$$\overline{\alpha}_{13} \le \overline{\alpha}_{12} + \overline{\alpha}_{23},$$

and the lemma is proved.

Lemma 1 immediately implies the following theorem.

Theorem 1. *If three curves L_1, L_2, and L_3 emanate from a common point O and make definite angles α_{12}, α_{23}, and α_{31} with each other, then the sum of each two of these angles is not less than the third.*

Indeed, $\alpha_{ij} = \lim_{t_i, t_j \to 0} \gamma_{ij}(t_i, t_j)$, and, therefore, Lemma 1 implies $\alpha_{13} \le \alpha_{12} + \alpha_{23}$, and the same for α_{23} and α_{31}.

We can give Lemma 1 in a slightly more general form.

Lemma 1^***.** *For arbitrarily many curves $L_1, \dots, L_n$ emanating from a common point,*

$$\limsup_{t_1, t_n \to 0} \gamma_{1n}(t_1, t_n) \le \limsup_{t_1, t_n \to 0} \gamma_{12}(t_1, t_2) + \dots + \limsup_{t_{n-1}, t_n \to 0} \gamma_{n-1,n}(t_{n-1}, t_n).$$

Indeed, by Lemma 1,

$$\limsup_{t_1, t_n \to 0} \gamma_{12}(t_1, t_n) \le \limsup_{t_1, t_{n-1} \to 0} \gamma_{1,n-1}(t_1, t_{n-1}) + \limsup_{t_{n-1}, t_n \to 0} \gamma_{n-1,n}(t_{n-1}, t_n).$$

Estimating $\lim_{t_1, t_{n-1} \to 0} \sup \gamma_{1,n-1}(t_1, t_{n-1})$ in the same way, and proceeding likewise, we obtain the required inequality.

Lemma 2. *Let curves $L_1, \dots, L_n$ emanate from a common point O, and let X_1 and X_n be variable points on the curves L_1 and L_n. Assume that for X_1 and X_2 arbitrarily close to O, the shortest arc X_1X_2 intersects the curves $L_2, \dots, L_{n-1}$ at points $X_2, \dots, X_{n-1}$ that are different from O and lie between X_1 and X_n in the order of their indices. We restrict consideration to only those points X_1 and X_n for which this is the case. Then, if t_i are the values of the parameters on the curves L_i corresponding to the above indicated points X_i,*

$$\liminf_{t_1, t_n \to 0} \gamma_{1n}(t_1, t_n) \ge \liminf_{t_1, t_2 \to 0} \gamma_{12}(t_1, t_2) + \dots + \liminf_{t_{n-1}, t_n \to 0} \gamma_{n-1,n}(t_{n-1}, t_n)$$

(*where the lower limits are taken only over the admissible values of t_i*).

Proof. Let X_1 and X_n be those points on the curve L_1 and L_n for which at least one of the shortest arcs connecting these points intersects $L_2, \dots, L_{n-1}$. We shall consider only these points X_1 and X_n without further specification.

Since there are such points arbitrarily close to the point O by condition, we can consider them as variable points tending to O. By the shortest arc X_1X_n, we mean precisely the shortest arc that intersects $L_2, \dots, L_{n-1}$. The points of this intersection are denoted by X_i (see Fig. 41, where $n = 4$). (It is important that X_i is different from O, and the fact whether this point is unique or not plays no role in what follows.)

We put

$$OX_i = x_i \quad (i = 1, \dots, n).$$

© 2006 by Taylor & Francis Group, LLC

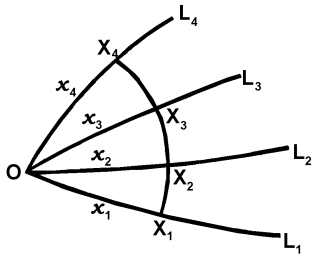


Fig. 41

Consider some location of the points X_1 and X_n and the corresponding points $X_2, \ldots, X_{n-1}$. Construct the plane triangles $O'X'_iX'_{i+1}$ with sides equal to the sides of the triangles OX_iX_{i+1}, i.e., for example, $O'X'_1 = OX_1 = x$, and so on. We adjoin sequentially these triangles along the sides $O'X'_i$ so that the triangles $O'X'_{i-1}X'_i$ and $O'X'_iX'_{i+1}$ have the common vertices O' and X'_i. As a result, we have the polygon $O'X'_1 \ldots X'_n$ (Fig. 42).

The angles at the vertex O' in the triangles $O'X'_iX'_{i+1}$ are $\gamma_{i,i+1}(t_i, t_{i+1})$.

Further, we construct the plane triangle $O''X''_1X''_{i+1}$ with sides equal to the sides of the triangle OX_1X_n. The angle at its vertex O'' is equal to $\gamma_{1n}(t_1, t_n)$.

Let us show that the angle at the vertex O'' in the triangle $O''X''_1X''_n$ is not less than the angle at the vertex O' in the polygon $O'X'_1 \ldots X_n$, i.e.,

$$\gamma_{1n}(t_1, t_n) \geq \gamma_{12}(t_1, t_2) + \cdots + \gamma_{n-1,n}(t_{n-1}, t_n).$$

To this end, we first note that $\gamma_{12} + \ldots + \gamma_{n-1,n} \leq \pi$.

Indeed, by the construction of the polygon $O'X'_1 \ldots X'_n$,

$$X'_1X'_2 + \ldots + X'_{n-1}X'_n = X_1X_2 + \cdots + X_{n-1}X_n = X_1X_n$$

and

$$O'X'_1 = OX_1, \quad O'X'_n = OX_n.$$

And since the line X_1X_n is a shortest arc, we have $X_1X_n \leq OX_1 + OX_n$, and, therefore,

$$O'X'_1 + O'X'_n \geq X'_1X'_2 + \cdots + X'_{n-1}X'_n.$$

However, if the angle at the vertex O' is greater than π, then the reverse inequality holds.

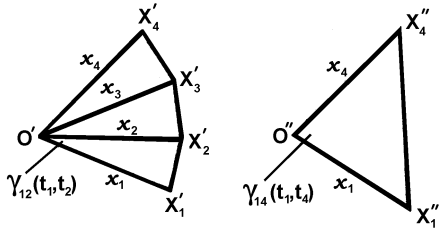


Fig. 42

Thus, the angle $\gamma_{12} + \ldots + \gamma_{n-1,n}$ at the vertex O' in our polygon is not greater than π. Therefore, this angle is also the angle in the triangle $O'X'_1X'_n$. Comparing this triangle with the triangle $O''X''_1X''_n$, we see that

$$O'X'_1 = O''X''_1, \quad O'X'_n = O''X''_n$$

by construction, i.e., the corresponding sides of these triangles meeting at the vertices O' and O'' are equal. At the same time, obviously,

$$X'_1X'_n \leq X'_1X'_2 + \cdots + X'_{n-1}X'_n,$$

while, by construction,

$$X''_1X''_n = X_1X_n = X_1X_2 + \cdots + X_{n-1}X_n = X'_1X'_2 + \cdots X'_{n-1}X'_n;$$

© 2006 by Taylor & Francis Group, LLC

therefore,

$$X_1'X_n' \le X_1''X_n''.$$

Hence, for the angles at the vertices O' and O'', we have the same inequality, i.e.,

$$\gamma_{12}(t_1, t_2) + \cdots + \gamma_{n-1,n}(t_{n-1}, t_n) \le \gamma_{1n}(t_1, t_n). \tag{3}$$

If the points X_1 and X_n tend to O, then all points X_i also tend to O, i.e., $t_i \to 0$ as $t_1, t_n \to 0$. Passing to the limit in inequality (3) as $t_1,\ t_n \to 0$ for the admitted values of t_1 and t_n, we obtain

$$\liminf_{t_1,t_n\to 0} \gamma_{1n}(t_1, t_n) \ge \liminf_{t_1,t_2\to 0} \gamma_{12}(t_1, t_2) + \cdots + \liminf_{t_{n-1},t_n\to 0} \gamma_{n-1,n}(t_{n-1}, t_n);$$

as required.

Theorem 2. *Let curves $L_1, \ldots, L_n$ emanate from a common point O and make definite angles $\alpha_{12}, \ldots, \alpha_{n-1,n}$, and α_{n1} with each other. Let X_1 and X_2 be variable points on L_1 and L_2. Assume that for X_1 and X_2 arbitrarily close to O, the shortest arc X_1X_n intersects the curves $L_2, \ldots, L_{n-1}$ at the points $X_2, \ldots, X_{n-1}$ that are different from O and lie between X_1 and X_2 in the order of their indices. Then*

$$\alpha_{1n} = \alpha_{12} + \alpha_{23} + \cdots + \alpha_{n-1,n}.$$

Proof. Since the angles $\alpha_{i,i+1}$ exist, the limits

$$\lim_{t_i,t_{i+1}\to 0} \gamma_{i,i+1}(t_i, t_{i+1}) = \alpha_{i,i+1}$$

exist, too. Therefore, by Lemma 1*,

$$\alpha_{1n} \le \alpha_{12} + \cdots + \alpha_{n-1,n}.$$

And since it is assumed that the conditions of Lemma 2 hold in the theorem, we have

$$\alpha_{1n} \ge \alpha_{12} + \cdots + \alpha_{n-1,n}.$$

Hence,

$$\alpha_{1n} = \alpha_{12} + \cdots + \alpha_{n-1,n}.$$

If $L_1, \ldots, L_n$ are shortest arcs in a manifold satisfying the nonoverlapping conditions for shortest arcs, then the shortest arcs $L_1, \ldots, L_n$, nonoverlapping and emanating from a common point O, have no other common points near O and hence partition a neighborhood of this point into sectors. Therefore, there is sense in saying that the shortest arcs $L_2, \ldots, L_{n-1}$ are located between L_1 and L_n in the order of their numbers from L_1 to L_n. In this case, if the shortest arc X_1X_n intersects $L_2, \ldots, L_n$, the intersection points $X_2 \ldots, X_n$ are located on $L_2, \ldots, L_n$ in the order of their indices from X_1 to X_n. Without the nonoverlapping condition for shortest arcs $L_1, \ldots, L_n$, it is not possible to assert that they lie in a definite order, and therefore we have to speak about the order of the points $X_1, \ldots, X_n$.

© 2006 by Taylor & Francis Group, LLC

Theorem 3. *Let curves $L_1 \ldots, L_n$ emanate from a common point O and make sequentially defined angles $\alpha_{12} \ldots, \alpha_{n-1,n}$ with each other that are different from zero. Let X_1 and X_n be variable points on L_1 and L_n. Assume that if X_1 and X_n are sufficiently close to O; then the shortest arc X_1X_n intersect $L_2, \ldots, L_n$ at the points $X_2, \ldots, X_{n-1}$ that are different from zero and are located between X_1 and X_n in the order of their numbers. Then there is a definite angle made by the shortest arcs L_1 and L_n, and this angle is equal to the sum of the angles $\alpha_{12}, \ldots \alpha_{n-1,n}$.*

Proof. The conditions of this theorem differ from those of Theorem 2 by the following two facts: the existence of the angle α_{1n} is not assumed, but it is assumed that the shortest arcs X_1X_n intersect $L_2, \ldots, L_{n-1}$ for all X_1 and X_n sufficiently close to O. By the latter condition, we can use Lemma 2 without any restrictions on the variables t_1 and t_2. Since the angles $\alpha_{i,i+1}$ are assumed to exist, then the limits

$$\lim_{t_i, t_{i+1}} \gamma_{i,i+1}(t_i, t_{i+1}) = \alpha_{i,i+1}$$

exist, and, therefore, it follows from Lemma 2 that

$$\lim_{t_1, t_n} \gamma_{1n} t_i, t_{i+1} \geq \alpha_{12} + \cdots + \alpha_{n-1,n}.$$

On the other hand, for the same reason, Lemma 1* implies

$$\limsup_{t_1, t_n \to 0} \gamma 1n(t_1, t_n) \leq \alpha_{12} + \cdots + \alpha_{n-1,n}.$$

Consequently, $\lim_{t_1,t_n \to 0} \gamma_{1n}(t_1, t_n)$ exists and equals $\alpha_{12} + \cdots + \alpha_{n-1,n}$, as required.

Certainly, this theorem plays no role for the shortest arcs on a convex surface, since the angle α_{1n} always exists, but it proves useful for other curves and also for shortest arcs in an abstract manifold with intrinsic metric.

This concludes our study of angles between any curves; in the next sections, we will speak of the angles between shortest arcs on convex surfaces. Then, in Chapter IX, we will consider some properties of the angle between any curves on convex surfaces in more detail.

2. Theorems on Addition of Angles on Convex Surfaces

For angles on convex surfaces, the general addition theorems proved in the previous section can be complemented by the theorem, analogous to the well-known theorem of plane geometry, that the sum of adjacent angles is equal to two right angles.

Theorem 1. *If a shortest arc M emanates from a point O lying inside a shortest arc L, the sum of angles that M makes with the branches of the shortest arc L is equal to π.*

Proof. The proof of this theorem is based on the two facts that refer precisely to convex surfaces: first, on the nonoverlapping condition for shortest arcs and, second, on the theorem on convergence of angles (Theorem 4 in Sec. 4 of Chapter III), that is, if shortest arcs N'_n and N''_n converge to N' and N'', then the angle between N'

© 2006 by Taylor & Francis Group, LLC

and N'' is not less than the lower limit of the angles between N_n' and N_n''. We shall use the consequences of the nonoverlapping condition for shortest arcs, which are deduced in Sec. 3 of Chapter II.

The branches L_1 and L_2 of the shortest arc L into which the point O divides this shortest arc are shortest arcs themselves. Denote by α_1 and α_2 the angles that these branches make with the shortest arc M. The angle between L_1 and L_2 is equal to π, and, therefore, by Theorem 1 of Sec. 1,

$$\alpha_1 + \alpha_2 \geq \pi. \tag{1}$$

If we prove that, at the same time, $\alpha_1 + \alpha_2 \leq \pi$, the theorem will be proved.

Let us restrict ourselves to a small neighborhood of O. The shortest arc L divides such a neighborhood into two domains U and V. Of course, we can assume that the shortest arc M does not overlap L, and then, by the nonoverlapping condition for shortest arcs (Theorem 1 in Sec. 3 of Chapter II), this shortest arc has no common points with L at all, but O. Therefore, M lies entirely in one of the domains U or V, e.g., in the domain U (Fig. 43).

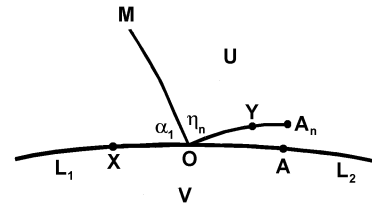


Fig. 43

Take a point A on L_2 and construct a sequence of points A_n converging to A inside the part of the domain U which is bounded by L_2 and M. The shortest arcs OA_n converge to the segment OA of the shortest arc L, since (by the nonoverlapping condition) OA is the unique shortest arc connecting the points O and A.

Take two points X and Y on L_1 and one of the shortest arcs OA, respectively. If these points are sufficiently close to the point O, then the shortest arc connecting these points lies in the domain U. Indeed, if the shortest arc XY "enters" the domain V, them this arc should intersect the shortest arc L at one more point except for X. But then, by the nonoverlapping condition for shortest arcs, this shortest arc should overlap L. However, this is impossible, since then the point Y, together with all points of the shortest arc OA_n, lies on L; this contradicts our assumption. Thus, the shortest arc XY has no common points with L but X and lies in the domain U.

The point X lies on L_1, and, by the construction of the shortest arcs OA_n, the point Y lies between the shortest arcs M and L_2. Therefore, the shortest arc XY intersects M whenever the points X and Y are sufficiently close to O.

In view of Theorem 2 of Sec. 1, this implies that the angle ξ_n made by OA_n and L_1 is equal to the sum of the angles α_1 and η_n made by L_1 and M and M and OA_n, that is,

$$\xi_n = \alpha_1 + \eta_n. \tag{2}$$

The shortest arcs OA_n converge to OA, and, therefore, by Theorem 4 in Sec. 4 of Chapter III,

$$\liminf_{n\to\infty} \eta_n \leq \alpha_2. \tag{3}$$

© 2006 by Taylor & Francis Group, LLC

On the other hand, $\xi_n \leq \pi$, and so

$$\limsup_{n\to\infty} \xi_n \leq \pi. \tag{4}$$

Therefore, passing to the limit as $n \to \infty$ in Eq. (2), we obtain from inequalities (3) and (4):

$$\alpha_1 + \alpha_2 \leq \pi.$$

Since we have already shown that $\alpha_1 + \alpha_2 \geq \pi$, we thus have $\alpha_1 + \alpha_2 = \pi$; the theorem is proved.

In contrast to the general theorems of the previous section, the above-obtained property of adjacent angles is not true in an arbitrary manifold with intrinsic metric. For example, if the nonoverlapping condition for shortest arcs does not hold, then the following case is possible: a shortest arc M is a prolongation of L_1 and comprises a new shortest arc together with L_1 but does not coincide with L_2. In this case, the angle α_1 between M and L_1 is equal to π, and $\alpha_1 + \alpha_2 > \pi$ whenever the angle α_2 between L and M is different from zero. For example, this is the case for each cone whose complete angle at the vertex is greater than 2π.

The theorem on adjacent angles is easily generalized to the case in which several shortest arcs emanate from a point O on a shortest arc L.

Theorem 1^***.** *Let shortest arcs $M_1, \ldots, M_n$ be drawn from a point O lying on a shortest arc L which runs near the point O on one side of L. Let them be enumerated in the order of their location from one branch L_1 of the shortest arc L to the other branch L_2 and let $\alpha_0, \alpha_1, \ldots, \alpha_n$ be the angles made by L_1 and M_1. Then*

$$\alpha_0 + \alpha_1 + \cdots + \alpha_n = \pi.$$

Proof. Indeed, it is easy to conclude from the nonoverlapping condition for shortest arcs that if two points X and Y lying on L_1 and M are different from O and are sufficiently close to this point, then the shortest arc XY intersects $M_1, \ldots, M_{n-1}$ at points different from O. Therefore, by Theorem 2 of the previous section, the angle made by L_1 and M is equal to $\alpha_0 + \cdots + \alpha_{n-1}$. At the same time, this angle is adjacent to α_n, and so $\alpha_0 + \cdots + \alpha_n = \pi$.

This result is an essential complement to Theorem 2 of the previous section. In this theorem we require that the shortest arc X_1X_n of two points X_1 and X_n on the shortest arcs L and L_n, respectively, intersects the other shortest arcs $L_2, \ldots, L_{n-1}$ at points different from O. If this condition does not hold then the shortest arcs L_1 and L_n are a prolongation of each other; since the shortest arc X_1X_n passes through the point O, the segments X_1O and X_nO of the shortest arcs L_1 and L_n form together one shortest arc, and the angle between L_1 and L_n turns out equal to π. Moreover, if the shortest arcs $L_2, \ldots, L_{n-1}$ lie on one side of the line $L_1 + L_n$ in the order of their indices from L_1 to L_n, then we are in the conditions of Theorem 1^*. Therefore, for a convex surface, we can unite Theorem 2 of Sec. 1 with Theorem 1^* in the following single theorem.

© 2006 by Taylor & Francis Group, LLC

Theorem 2. *Let shortest arc $L_1, \ldots, L_n$ enumerated in the order of their location around a point O emanate from this point on a convex surface.*[3] *Let $\alpha_{12}, \ldots, \alpha_{n-1,n}$, α_{1n} be the angles between the neighboring arcs of these shortest arcs. If, for X_1 and X_2 arbitrarily close to O, the shortest arc X_1X_n has common points with all shortest arcs $L_2, \ldots, L_{n-1}$ then*

$$\alpha_{1n} = \alpha_{12} + \cdots + \alpha_{n-1,n}.$$

3. The Angle of a Sector Bounded by Shortest Arcs

Two shortest arcs L and M emanating from a common point O on a convex surface divide a neighborhood of this point into two *sectors*. The precise definition of these sectors can be given, e.g., as follows. The point O has a neighborhood homeomorphic to a disk. Take this neighborhood U so small that the shortest arcs L and M intersect its boundary. Let A and B be the nearest points to A of intersections of the shortest arcs L and M with the boundary of the neighborhood U. Since a shortest arc is homeomorphic to a line segment, the segments OA and OB of the shortest arcs L and M divide the neighborhood U into two parts, the sectors. By homeomorphic mapping, we may map the neighborhood U onto a disk, the segments OA and OB becoming the radii of this disk; herewith, the disk proves to be divided into two sectors. In what follows, when speaking about sectors, we shall keep in mind so small neighborhood of the point O that each shortest arc connecting points from different sectors intersects one of the shortest arc L and M. The shortest arcs that bound a sector are also included into it. For formal reasons, it seems convenient to assume that a sole shortest arc is also a sector. Such "zero" sector can be said to be bounded by two overlapping shortest arcs.

The neighborhood divided into sectors by shortest arcs can, certainly, be chosen in various ways, and then the sectors bounded by two given shortest arcs L and M are different sets of points. However, we shall consider them as the same sector. To be more exact, two sets V_1 and V_2 bounded by the same pair of shortest arcs L and M in two neighborhoods U_1 and U_2 are considered as a single sector if in a neighborhood U_3 contained simultaneously in U_1 and U_2 the shortest arcs L and M isolate the set (sector) V_3 simultaneously contained in V_1 and V_2. Thus, a sector is an arbitrary representative of the whole class of sets or, still better, the very class of sets isolated from neighborhoods of a given point by two shortest arcs.

We will say that a sector V is contained in a polygon P if V is bounded by two sides of the polygon P meeting at one of its vertices O and if, for a sufficiently small neighborhood of the point O, one of the representatives of the class of sets which is the sector V, lies in P.

With each sector bounded by two shortest arcs emanating from a common point, we associate a definite angle according to the following rule. Let a sector V be bounded by two shortest arcs L and M emanating from a point O. Let us draw some more shortest arcs $N_1, \ldots, N_n$ from this point O, which are enumerated in the order of their location from L to M. Denote by $\alpha_0, \alpha_1, \ldots, \alpha_n$ the angles between the neighboring shortest arcs, i.e., between L and N_1, N_1 and N_2, and finally, between N_n and M. We will vary both the shortest arcs N_i and their number in an

[3] By the nonoverlapping condition for shortest arcs, this assertion has a meaning.

© 2006 by Taylor & Francis Group, LLC

arbitrary way, still retaining them in the sector V. With each set of shortest arcs L_i, we associate the sum $\alpha_0 + \alpha_1 + \ldots + \alpha_n$. *The least upper bound of these sums is taken as the angle of the sector* V. If this least upper bound does not exist, then we can assume that the angle of this sector is equal to infinity. This remark plays no role for us, since we will prove below that the indicated least upper bound does not exceed 2π for sectors on convex surfaces. However, up to the moment when this will be proved, we admit implicitly infinite values for angles of sectors.[4]

In this definition of the angle of a sector, the empty set of shortest arcs N_i is also admitted, i.e., we can simply draw no shortest arcs, and then the sum $\alpha_0 + \cdots + \alpha_n$ is reduced to one angle between the shortest arcs L and M. Since the angle of a sector is defined as the least upper bound of these sums, the previous remark implies that the angle of a sector bounded by two shortest arcs is always not less than the angle between these shortest arcs.

Let us find some conditions that ensure the equality of both angles.

Take two variable points X and Y on two shortest arcs L and M emanating from a common point O. Assume that the shortest arcs L and M do not overlap. By the nonoverlapping condition for shortest arcs, two shortest arcs having a common point either have no other interior common points or overlap. Therefore, the shortest arc XY either has no common points with L and M but the points X and Y or overlaps them, so that on the part between X and Y, the shortest arcs L and M combine into a single shortest arc OX. In the first case, the shortest arc XY lies entirely inside one of the sectors. In the second case, we have the same grounds to assume that this shortest arc lies in one or the other sector.

Let the shortest arc XY lie in a sector U bounded by L and M for X and Y sufficiently close to O. Draw some shortest arcs $N_1, \ldots, N_n$ from the point O in the sector U, which are enumerated in the order of their location from L to M. Let $\alpha_0, \alpha_1, \ldots, \alpha_n$ be the angles between L and N_1, N_1 and N_2,...,N_n and M. If the shortest arcs XY go inside the sector U, then they intersect the shortest arcs $N_1, \ldots, N_n$, and, therefore, Theorem 2 of Sec. 1 implies that the sum of angles $\alpha_0 + \cdots + \alpha_n$ is equal to the angle α between L and M. Consequently, the least upper bound of these sums, i.e., the angle of the sector U, is equal to the angle α.

If the shortest arc XY overlaps L and M, then the segments OX and OY of these shortest arcs constitute one shortest arc. Therefore, first, the angle α between L and M is equal to π. Second, by Theorem 1^* of Sec. 2, the sum of angles $\alpha_0 + \cdots + \alpha_n$ is equal to π. Consequently, the least upper bound of these sums, i.e., the angle of the sector U, is again equal to the angle π.

[4]It is useful to compare this definition of the angle of a sector with the definition of the length of a curve. There is a full formal analogy between both definitions, which is strengthened by the fact that the requirements which the distance satisfies hold for angles between shortest arcs; in particular, similar to the triangle inequality, Theorem 1 of Sec. 1 holds: $\alpha(L_1L_2) + \alpha(L_2L_3) \geq \alpha(L_1L_3)$, where $\alpha(L_iL_j)$ is the angle between two curves L_i and L_j. This analogy is not very interesting for two-dimensional manifolds, since the set of shortest arcs emanating from a common point in them is one-dimensional. This is not so for higher-dimensional manifolds. Let a family of shortest arcs $L(t)$ emanating from a common point O be given in a manifold R $(0 \leq t \leq 1)$. Take values $t_0 = 0 < t_1 < t_2 \ldots < t_n = 1$ and form the sum $\sum \alpha(L(t_{i-1}, L(t_i))$. It is natural to take the least upper bound of these sums as the "angular length" of the family $L(t)$. The family $L(t)$ connecting two given shortest arcs $L = L(0)$ and $M = L(1)$ and having the shortest angular length is an analog of the plane sector bounded with L and M.

© 2006 by Taylor & Francis Group, LLC

The above results can be summarized as follows.

Theorem 1. *The angle of the sector U bounded by two shortest arcs L and M emanating from a common point O is always not less than the angle between these shortest arcs themselves. If the shortest arcs XY connecting two points X and Y on L and M go in the sector U for X and Y arbitrarily close to O, then the angle of the sector U is equal to the angle between these shortest arcs. Hence, in all the cases, the angle of one of the two mutually complementary sectors is equal to the angle between the shortest arcs bounding these sectors.*

Hence, if the shortest arcs XY for X and Y arbitrarily close to O go in the sector U, as well as in the complemented sector U', or simultaneously in both sectors, then the angles of both sectors are equal to the angle between the shortest arcs L and M.

In particular, if XY overlaps L and M, the angles of both sectors U and U' are equal to π.

(The angle of zero sector is equal to zero, and, therefore, the theorem also holds even in the case where the shortest arcs L and M overlap each other.)

Now let us prove that angles of sectors are added in the same way as angles of sectors on the plane.

Theorem 2. *If a sector U is composed of two sectors V and W, the angle of U is equal to the sum of angles of V and W provided of cause that the sectors V and W have no common points, except for points of the shortest arc separating these sectors and emanating from the vertex of U.*

Proof. Assume that the sector U is bounded by shortest arcs L and M emanating from a point O. Assume also that U is divided into two sectors V and W emanating from the same point O; and, moreover, let V denote the sector bounded by L and N.

Let u, v, and w be angles of the sectors U, V, and W.

Draw shortest arcs $N_1, \ldots, N_n$ from the point O inside the sector U, which are numbered in the order of their location from L to M. As above, let $\alpha_0, \alpha_1, \ldots, \alpha_n$ denote the angles made by L and N_1, N_1 and N_2, $\ldots$, N_n and M. By the definition of the angle of a sector,

$$u = \sup \sum_{i=0}^{n} \alpha_i.$$

Add one more shortest arc N' to the set of shortest arcs N_i that passes, e.g., between N_k and N_{k+1}. Then the angle α_k made by N_k and N_{k+1} in the sum $\alpha_0+\alpha_1+\cdots+\alpha_n$ should be replaced by the sum of angles made by N_k and N' and N' and N_{k+1}. But, by Theorem 1 of Sec. 1, the sum of these angles is not less than the angle made by N_k and N_{k+1}. Hence the sum of angles of the form $\alpha_0+\alpha_1+\cdots+\alpha_n$ does not decrease when we add a new shortest arc to the set of those drawn. Therefore, when seeking the least upper bound of these sums, we can assume that there is a shortest arc N among the shortest arcs N_i dividing the sectors V and W. For example, let the shortest arc N_m be this shortest arc N. Then the shortest arcs $N_1, \ldots, N_{m-1}$ lie in the sector V, while the shortest arcs $N_{m+1}, \ldots, N_n$ lie in the

© 2006 by Taylor & Francis Group, LLC

sector W. These shortest arcs are arbitrary, and, therefore, in accordance with the definition of angles of sectors, we have

$$v = \sup \sum_{i=1}^{m-1} \alpha_i, \quad w = \sup \sum_{i=m}^{n} \alpha_i.$$

And since $u = \sum_{i=1}^{n} \alpha_i$, therefore, $u = v + w$; as required. Of course, this implies that the angle of the sector composed of arbitrary many sectors is equal to the sum of angles of those sectors.

Further, the above theorem implies that *the sum of angles of two mutually complementary sectors does not depend on these sectors and depends only on the point that is their vertex.*

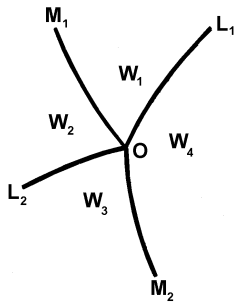


Fig. 44

Indeed, let us have two pairs of mutually complementary sectors U and U' and V and V' with a common vertex O. The first is bounded by shortest arcs L_1 and L_2, and the second is bounded by shortest arcs M_1 and M_2. Assume, e.g., that the shortest arc M_1 lies in the sector U, M_2 lies in the sector U', the shortest arc L_1 lies in the sector V, and L_2 lies in the sector V'. Then a neighborhood of the point O is decomposed into some sectors W_1, W_2, W_3, and W_4 that are bounded by the shortest arcs L_1 and M_1, M_1 and L_2, etc. (Fig. 44). The initial sectors are represented as the sums

$$U = W_1 + W_2, \quad U' = W_3 + W_4,$$
$$V = W_1 + W_4, \quad V' = W_2 + W_3.$$

Therefore, the above theorem implies that the sum of angles of the sectors U and U', as well as the sum of the angles of the sectors V and V', is equal to the sum of the angles of the sectors $W_1, \ldots, W_4$. Therefore, these sums are equal.

The same result, of course, is obtained for another interlocation of the shortest arcs bounding these sectors.

The above result leads to the following important definition: *The sum of two mutually complementary sectors with vertex O is called the complete angle at O.*

However, the complete angle can be defined directly without appeal to the concept of the angle of a sector. Assume that shortest arcs $L_1, \ldots, L_n$ emanate from a point O and are numbered in the order of their cyclic location around O. Let α_{ij} denote the angle between L_i and L_j. Arrange the sum $\alpha_{12} + \cdots + \alpha_{n-1,n} + \alpha_{n1}$. *The least upper bound of these sums over each finite number of all possible shortest arcs emanating from O is equal to the complete angle at the point O.*

In fact, when proving the theorem on addition of angles of sectors, we show that the addition of new shortest arcs to the set of those drawn does not decrease the sum of angles between neighboring shortest arcs. Therefore, when determining the least upper bound of these sums, we can assume that there are two given shortest arcs L and M among the shortest arcs L_i, which divide the neighborhood of the point

© 2006 by Taylor & Francis Group, LLC

O into sectors U and U'. Then each sum $\alpha_{12} + \cdots + \alpha_{n-1,n} + \alpha_{n1}$ splits into two sums; one for shortest arcs passing in the sector U, and the other for shortest arcs passing in the complementary sector U'. The least upper bounds of these partial sums are equal to the angles of the sectors U and U' by definition, while their sum yields the sum of angles of these sectors, i.e., we come to the complete angle at the point O.

Using the concept of the complete angle at a point, we can formulate the following theorem.

Theorem 3. *The sum of angles of sectors with a common vertex O, which cover the whole neighborhood of O and have no common interior points, is equal to the complete angle at O.*

Indeed, let $U_1, U_2, \ldots, U_n$ be such sectors with vertex O. The sectors U_1 and $U_2 + \cdots + U_n$ are complementary. Therefore, the sum of their angles is equal to the complete angle at the point O. At the same time, the angle of the sector $U_2 + \cdots + U_n$ is equal to the sum of angles of the sectors composing this sector. Hence the sum of the angles of all sectors $U_1, \ldots, U_n$ is equal to the complete angle at the point O.

It is hardly worth mentioning that the complete angle at the vertex of a cone as defined in Chapter I is the complete angle in the sense of this definition.

Theorem 4. *The complete angle at a point on a convex surface is always less than or equal to 2π.*

This theorem is obvious for convex polyhedra, while for all convex surfaces it can easily be proved by passing to the limit from polyhedra. However, we do not present this proof here, since in the next section it appears as a simple corollary of the general theorem on convergence of the angles of sectors.

It is clear from the very definition of complete angle that the angle of each sector with vertex O is not greater than the complete angle at this point. Therefore, the angle of each sector on a convex surface has a *finite* value, and moreover, it is less than 2π. The exception consists in the sector complementary to the zero sector: the angle of such sector with vertex a point O is equal to the complete angle at this point and, therefore, is equal to 2π in general. In the sequel, we will show that on a convex surface those points at which the complete angle is less than 2π comprise at most a countable set.

The following theorem states an important particularity of these points.

Theorem 5. *There are no shortest arcs passing through a point at which the complete angle is less than 2π.*

Indeed, if a shortest arc passes through a point O then the branches of this shortest arc into which the point O divides it make the angle equal to π. Therefore, the angles of both sectors bounded by these branches are not less than π and, hence, the complete angle at the point O is no less than 2π. This argument is true in any manifold. But there are no points on a convex surface at which the complete angle is greater than 2π, and therefore *the complete angle at a point belonging to a shortest arc on a convex surface is equal to 2π.*

We now have two angles: the angle between shortest arcs and the angle of a sector. *By the angle of a polygon P at a vertex A, we shall always mean the angle*

© 2006 by Taylor & Francis Group, LLC

of the sector that is bounded by the sides meeting at A and lies in the polygon P. This angle can be equal to the angle between the sides or complement it to the complete angle at the point A. All angles of a convex polygon are equal to the angles between its sides. Each triangle on the plane is convex, and, therefore, both angles coincide for plane triangles. But on convex surfaces, even an arbitrarily small triangle may fail to be convex, and far from always the angle between sides of a triangle will be equal to the angle of a sector contained in this triangle and bounded by its sides. For example, on a closed convex surface, three shortest arcs connecting three points A, B, and C divide it into two triangles ABC. And if in one of them the angle at the vertex A is equal to the angle made by the sides AB and AC, then in the other triangle this angle turns out to be the complement of this angle to the complete angle at the point A.

4. On Convergence of Angles

In Sec. 4 of Chapter III, we proved the following general theorem on convergence of angles on convex surfaces.

Theorem 1. *Let two shortest arcs L and M emanate from a point O on a convex surface F and make an angle α with each other. Let shortest arcs L_n and M_n converging to L and M be drawn from points O_n on convex surfaces F_n converging to F, and let the points O_n converge to O. If α_n stands for the angle between L_n and M_n, then*

$$\alpha \leq \liminf_{n\to\infty} \alpha_n.$$

In Sec. 10 of Chapter I, we have shown by inspecting an example that the limit of the angles α_n may fail to exist, and if it does exist, then it can be greater than α. Now we want to find the conditions under which the angles α_n always converge to the angle α and to study the convergence of angles for the sectors bounded by the shortest arcs L and M.

Let a sector U be given on a convex surface F which is bounded by two shortest arcs L and M emanating from a point O, and let sectors U_n be given on convex surfaces F_n converging to F which are bounded by shortest arcs L_n and M_n and emanate from the points O_n. We shall say that the sectors U_n converge to the sector U if the following three conditions hold:

(1) the points O_n converge to O; (2) the shortest arcs L_n and M_n converge to L and M, respectively; (3) each interior point of the sector U is the limit of interior points of the sectors U_n.

Two shortest arcs L and M emanating from a common point bound two mutually complementary sectors U and U'. It is easy to prove that *if the points O_n converge to O and the shortest arcs L_n and M_n converge to the shortest arcs L and M, then some of the mutually complementary sectors bounded by the shortest arcs L_n and M_n converge to U and the rest of them converge to U'. If the sectors U_n converge to U, then the sectors U'_n complementary to U_n converge to the sector U' complementary to U.*

This assertion is so obvious that we do not prove it here in order to keep our main theme; the proof will be given at the end of this section in the form of a supplement. We now turn to studying the convergence of angles of convergent sectors.

© 2006 by Taylor & Francis Group, LLC

Theorem 2. *Let sectors U_n on convex surfaces converging to F converge to a sector U. If α and α_n stand for the angles of the sectors U and U_n then*

$$\alpha \le \liminf_{n\to\infty} \alpha_n.$$

Proof. Let the sector U be bounded by two shortest arcs L and M emanating from a point O, and let the sectors U_n be bounded by shortest arcs L_n and M_n emanating from points O_n. By the very definition of convergence of sectors, O_n converge to O, while L_n and M_n converge to L and M; we assume that L_n converge to L and M_n converge to M.

The definition of the angle of a sector implies that for each $\varepsilon > 0$ we can draw shortest arcs $N_1, \ldots, N_m$ in the sector U such that on enumerating them in the order of their location from L to M and denoting the angles between L and N_1, N_1 and N_2,..., N_m and M by $\alpha^0, \alpha^1, \ldots, \alpha^m$, we obtain

$$\alpha - (\alpha^0 + \cdots + \alpha^n) < \varepsilon. \tag{1}$$

Drawing shortest arcs $N_1, \ldots, N_m$ in the sector U such that this condition holds, we take points $A_1, \ldots, A_m$ inside these arcs. Then the segments $OA_1, \ldots, OA_m$ of these shortest arcs are unique shortest arcs connecting the point O with the points A_i. We take the points A_i so close to O that, first, they lie inside the sector U, and, second, it is possible to take points A_{in} inside the sectors U_n which converge to the points A_i, respectively. This is possible by the very definition of convergence of sectors.

Since the shortest arcs OA_i are unique shortest arcs connecting O with A_i, the shortest arcs OA_{in} converge to them, and at least for sufficiently large i they are located between L_n and M_n in the same order.[5] Let $\alpha_n^0, \alpha_n^1, \ldots, \alpha_n^m$ be the angles made by the shortest arc L_n and the shortest arcs O_nA_{0n}, O_nA_{1n} and O_nA_{2n}, and so on. Then the definition of the angle of a sector implies that the angle of U_n is not less than the sum of the angles $\alpha_n^0, \ldots, \alpha_n^m$, i.e.,

$$\alpha_n \ge \alpha_n^0 + \ldots + \alpha_n^m, \tag{2}$$

whenever the shortest arcs O_nA_{in} are located in the order of their indices between L_n and M_n for sufficiently large n.

Since the shortest arcs $L_n, O_nA_{1n}, \ldots, M_n$ converge to the shortest arcs $L, OA_1, \ldots, M$, applying Theorem 1 to the angles between these shortest arcs, we obtain

$$\alpha^i \le \liminf_{n\to\infty} \alpha_n^i \quad (i = 0, 1, \ldots, m). \tag{3}$$

Therefore, formula (2) implies

$$\liminf_{n\to\infty} \alpha_n \ge \alpha^0 = \ldots + \alpha^m;$$

applying inequality (1), we obtain from this that $\liminf_{n\to\infty} \alpha_n > \alpha - \varepsilon$.

[5] Indeed, we can assume that all OA_i do not overlap. First, if two given points O_n and A_{in} can be connected by several shortest arcs then we take one of these arcs. For large i, the shortest arcs O_nA_{in} do not overlap and divide the sector U_n into smaller sectors which converge to the sectors bounded by the shortest arcs OA_i by the assertion formulated above without proof.

© 2006 by Taylor & Francis Group, LLC

Since ε is arbitrary, this implies

$$\liminf_{n\to\infty} \alpha_n \geq \alpha;$$

as required.

Corollary 1. *Let points O_{n1} lying on convex surfaces F_n converging to a convex surface F converge to a point O on F. If θ_n and θ stand for the complete angles at the points O_n and O, then $\theta \leq \liminf_{n\to\infty} \theta_n$.*

Since the complete angle at a point is equal to the sum of two mutually complementary sectors, this assertion is a direct consequence of Theorem 2.

Corollary 2. *The complete angle at a point on a convex surface does not exceed 2π.*

This assertion is obvious for points on smooth surfaces. Therefore, if in the previous corollary by convex surfaces we mean polyhedra, then $\theta_n \leq 2\pi$ and so $\theta \leq 2\pi$. This proves Theorem 4 of the preceding section.

Theorem 3. *Let sectors U_n on convex surfaces X_n converging to a convex surface F converge to a sector U on this surface, and let the complete angles θ_n at the vertices of U_n converge to the complete angle θ at the vertex of U. If α_n and α stand for the angles of U_n and U, then* (1) *the limit $\lim_{n\to\infty} \alpha_n$ exists and* (2) *$\alpha = \lim_{n\to\infty} \alpha_n$.*

Proof. Since the sectors U_n converge to the sector U, the complementary sectors U_n' converge to the sector U' complementary to U. The angles of the sectors U_n' and U' are equal to $\theta_n - \alpha_n$ and $\theta - \alpha$, respectively. Therefore, by Theorem 2,

$$\theta - \alpha \leq \liminf_{n\to\infty}(\theta_n - \alpha_n),$$

and since $\lim_{n\to\infty} \theta_n = \theta$ by condition, we have

$$\alpha \leq \limsup_{n\to\infty} \alpha_n.$$

On the other hand, by Theorem 1, it should be

$$\alpha \leq \liminf_{n\to\infty} \alpha_n.$$

Hence there exists the limit $\lim_{n\to\infty} \alpha_n$, and it is equal to α. The theorem is proved.

Note that if $\theta \neq \lim_{n\to\infty} \theta_n$, the angle of at least one of the sectors U and U' is not equal to the limit of angles of the sectors U_n and U_n' or these limits do not exist. Therefore, Theorem 3 can be formulated as follows: the angles of both mutually complementary sectors U and U' are equal to the limits of angles of mutually complementary sectors U_n and U_n' converging to them if and only if the complete angles at the vertices of the sectors U_n converge to the complete angle at the vertex of the sector U.

© 2006 by Taylor & Francis Group, LLC

Theorem 4. *Let sectors U_n on convex surfaces F_n converging to F converge to a sector U on F such that the complete angle at its vertex is equal to 2π. If α_n and α stand for the angles of the sectors U and U_n, then* (1) *the limit* $\lim_{n\to\infty}\alpha_n$ *exists and* (2) $\alpha = \lim \alpha_n$.

Proof. Let θ_n be the complete angle at the vertex of the sector U_n. Since the complete angle at the vertex of the sector U is equal to 2π, we have

$$\liminf_{n\to\infty} \theta_n \geq 2\pi$$

in accordance with Corollary 1 to Theorem 2. But, by Corollary 2 to the same theorem, $\theta_n \leq 2\pi$. Therefore, the limit $\lim_{n\to\infty}\theta_n$ exists and is equal to 2π, i.e., to the complete angle at the vertex of the sector U. This means that in this case, the condition of Theorem 3 holds. Therefore, the assertion of this theorem also holds, i.e., the limit $\lim_{n\to\infty}\alpha_n$ exists and is equal to α.

Theorems 3 and 4 easily imply analogous properties of the convergence of angles between shortest arcs.

Theorem 5. *Let convex surfaces F_n converge to a convex surface F, and let shortest arcs L_n and M_n emanating from points O_n on the surfaces F_n converge to two shortest arcs L and M emanating from a point O on F; moreover, assume that the points O_n converge to O. If the conditions of Theorems* 3 *or* 4 *hold, i.e., if the complete angles at the points O_n converge to the complete angle at the point O, and, in particular, if the complete angle at the point O is equal to 2π, then the angles between the shortest arcs L_n and M_n converge to the angle between L and M.*

Proof. In the previous section, we have shown that the minimum of the angles of two mutually complementary sectors is equal to the angle between shortest arcs that bound this sector; if the angles of both sectors are equal, they are equal to the angle between the bounding shortest arcs. Hence, if α and α' are the angles of the two mutually complementary sectors U and U' bounded by the shortest arcs L and M and if $\alpha \leq \alpha'$ then α is the angle between the shortest arcs L and M. Let α_n and α'_n be the angles of two mutually complementary sectors U_n and U'_n bounded by the shortest arcs L_n and M_n. If the conditions of Theorem 3 or 4 hold, then

$$\lim_{n\to\infty}\alpha_n = \alpha, \quad \lim_{n\to\infty}\alpha'_n = \alpha'.$$

If $\alpha = \alpha'$, the angles α_n and α'_n play the same role, and since one of them is equal to the angle between the shortest arcs L_n and M_n, then the limit of these angles is also equal to α, i.e., to the angle between L and M.

If $\alpha < \alpha'$, then we have $\alpha_n < \alpha'_n$ for large n, and so α_n is the angle between the shortest arcs L_n and M_n. Therefore, in this case, the relation $\alpha = \lim_{n\to\infty}\alpha_n$ means that the angle between the shortest arcs L_n and M_n converges to the angle between L and M. The theorem is proved.

In all Theorems 1–5 proved here, the surfaces F_n converging to F can be certainly assumed equal to F. Then these theorems transform into propositions on convergence of angles on convex surfaces. Further, if on a given surface we consider

© 2006 by Taylor & Francis Group, LLC

sectors U_n with a common vertex, the condition of Theorem 3 holds trivially: the vertices of all sectors U_n and, therefore, the complete angles around them coincide. Therefore, Theorems 4 and 5 imply the following theorem.

Theorem 6. *Let shortest arcs L_n and M_n emanating from a common point on a convex surface converge to two shortest arcs L and M. Then* (1) *the angles between the shortest arcs L_n and M_n themselves converge to the angle between L and M and* (2) *the angles of both sectors bounded by the shortest arcs L_n and M_n converge to the angles of sectors bounded by the shortest arcs L and M, respectively.*

This theorem expresses the continuous dependence of the angle on shortest arcs and essentially supplements the main properties of the angle which are proved in the previous sections. However, this does not complete the study of main properties of the angle since, e.g., we have not proved even such a simple (at first glance) theorem that a neighborhood of each point on a convex surface can be divided into sectors with arbitrarily small angles. This theorem will be proved only in Sec. 5 of the next chapter. Then the general theory of the angle between shortest arcs on a convex surface will be completed from the standpoint point of intrinsic geometry.

Supplement

In the beginning of this section, we have formulated the assertion that if points O_n and shortest arcs L_n and M_n emanating from O_n on convex surfaces F_n converging to F converge to a point O and two shortest arcs L and M on F, respectively, then some of two mutually complementary sectors bounded by the shortest arcs L_n and M_n converge to one of the sectors U and U' bounded by L and M, while the rest of them converge to the other sectors. If we take two points A and B on L and M and connect these points by a broken line composed of shortest arcs and running in the sector U, we obtain a polygon that contains this sector. It is clear thence that the above-formulated proposition concerning convergence of sectors is an immediate corollary to the following general lemma:

Lemma. *Assume given a closed broken line L without multiple points on a convex surface F. This broken line divides F into two polygons G and H. Assume that we are given closed broken lines L_n without multiple points on closed convex surfaces F_n converging to F, and the sides of these broken lines converge to the sides of the broken line L, respectively. Then the domains into which the broken lines L_n divide the surfaces F_n converge to the domains G and H, respectively.*[6]

Proof. Take an arbitrary point A inside the domain G, and let $A_1, A_2, \ldots$ be a sequence of points on surfaces F_1, $F_2, \ldots$ converging to A. For an arbitrarily large n, the points A_n cannot lie on the curves L_n, since the point A lies inside the domain G and the curves L_n converge to the curve L, that is the boundary of this domain. Therefore, at least for a large n, the point A_n lies inside one of the domains bounded by the curve L_n on the surface. Denote this domain by G_n. We need to prove that the domains G_n converge to the domain G while the domains H_n complementing the domains G_n to the surface F_n converge to the domain H.

[6] Here we speak about closed domains, since the limit of a sequence of arbitrary sets is a closed set.

© 2006 by Taylor & Francis Group, LLC

Assume first that the surface F does not degenerate into a doubly-covered plane domain. Then at least for a large n, the surfaces F_n are also nondegenerate, and we can take a point O lying inside these surfaces F_n and inside F. Circumscribing a sphere S around the point O, we project the surface F and F_n from the point O to the sphere S. Then the domains G, H, G_n, and H_n and the curves L and L_n are projected on the sphere S, and the projections of L_n converge to the projection of L. By a homeomorphic transformation of the sphere S onto itself, the projection of the curve L transforms into a circle; moreover, the projections of the curve L_n go to the curves converging to this circle. Then the theorem reduces to the assertion that if simple closed curves on the sphere converge to a circle then the domains bounded by them converge to the domain bounded by this circle. The proof of this assertion is obvious, and we do not give it here.

If the surface F degenerates into a plane domain F^*, then taking a point O inside this domain and circumscribing a unit disk around this point, we transform the surface F into a doubly-covered disk K. Namely, if a point X lies in the "upper" ("lower") domain and the ray OX intersects the boundary of the domain F^* at the point Y, then to the point X we put in correspondence a point Z such that this point lies "upper" ("lower") on the disk K and $OZ = OX/OY$. After that, circumscribing the unit sphere S around the point O, we transform each upper (lower) point of the disk K into a point of the upper (lower) hemisphere lying over this point. As a result, the surface F turns out to be homeomorphically mapped onto the sphere S. The surface F_n can first be projected to the plane containing F, and then we can map them onto the sphere in the same way (the singularities connected with possible noninjectivity of the projection play no role). Thus, the preceding argument can be repeated.

5. The Tangent Cone

By now, we have studied the properties of the angle between shortest arcs from a purely intrinsic point of view. But when we speak about a convex surface rather than an abstract manifold, the shortest arcs are curves in the space, and there naturally arises the question about the spatial, or say, extrinsic geometric meaning of the angle between shortest arcs. The answer we will give is as follows.

Let O be a point on a convex surface F. We will homothetically dilate the surface F with the center of homothety at the point O. Then, as the coefficient of homothety increases indefinitely, the surfaces obtained will converge to some convex cone with vertex at the point O. This cone K is called the *tangent cone* to the surface F at the point O.[7] At the same time, under this infinite homothetic dilation of the surface F, any shortest arc L emanating from a point on the surface F converges to some generator of the tangent cone K. This generator T is nothing else but the limit of half-lines emanating from O and going to the variable point X of the shortest arc L provided that the point X tends to the point O. Hence, the generator T is a tangent half-line or, say, a *half-tangent* to the shortest arc L. If

[7] More often, the tangent cone is defined as a cone formed by the limit of rays OX going from the point O to the points X of the surface F under the condition that X tends to O. It is easy to show that both definitions are equivalent. Our definition is only in greater correspondence with our purposes.

© 2006 by Taylor & Francis Group, LLC

two shortest arcs L_1 and L_2 emanate from the point O with the half-tangents T_1 and T_2, these half-tangents divide the cone K into two sectors V and V'. Let U and U' be two sectors into which the shortest arcs L_1 and L_2 divide a neighborhood of the point O on the surface F. Under the indefinite dilation, one of these sectors converges to V, and the other converges to V'. It turns out that the angle of the sector among U and U' which converges to V is equal to the angle of the sector V, and the angle of the sector that converges to V' is equal to the angle of the sector V'. This result completely reveals the spatial sense of the angle of a sector and, hence, that of the angle between shortest arcs. In particular, this immediately implies that the complete angle at the point O on the surface F is nothing else but the complete angle of the tangent cone at the point O. The theorem on addition of angles of sectors on the surface F reduces to the addition of angles of sectors bounded by the cone generator. Some other properties of the angle also acquire a corresponding spatial interpretation.

If the tangent cone reduces to a plane then this plane is the tangent plane to the surface F at the point O. In this case, the angle made by shortest arcs reduces to the angle between the half-tangents to them. If the tangent cone is not a plane, then the angle between half-tangents measured in the space is in general different from the angle measured in the tangent cone. But exactly this latter angle is equal to the angle made by shortest arcs. Hence, generally speaking, the spatial angle made by shortest arcs, i.e., the angle made by their half-tangents, is not equal to the angle made by them in the sense of intrinsic geometry of the surface.

Each convex cone can be a tangent cone to a convex surface. In turn, convex cones can be of the following three types: a plane, a dihedral angle, and a cone with complete angle $< 2\pi$. In the first case, a point of the surface is, say, "smooth," in the second case, it is "edge-like," and in the third case, "conical."

In this section, we study the relation between the metric of a convex surface in a neighborhood of a point O on the one hand, and that of the tangent cone at the point O on the other. The result obtained here is of independent significance for understanding the properties of the intrinsic metric of a convex surface. The next section will prove the existence of half-tangents to shortest arcs and then will prove all we have said here about the intrinsic-geometric nature of the angle between shortest arcs.

Lemma. *A convex surface has the tangent cone at its every point, i.e., if we homothetically dilate a convex surface F at one of its points O, then this surface will converge to some convex cone with vertex at O.*

Proof. Assume that the surface F lies entirely in one plane P. If, moreover, a neighborhood of the point O on the surface is its neighborhood in this plane P, then the lemma is obvious: under an indefinite dilation of the surface, this surface in the limit covers the whole plane P which is consequently the tangent cone at the point O. If a neighborhood of the point O on the surface is not a neighborhood of it on the plane, this means that the point O lies on the boundary of the domain on the plane P which is doubly-covered by the surface F near the point O. In this case, a segment of this boundary near the point O is a convex curve L (Fig. 45). The left and right half-tangents T_1 and T_2 to L at the point O cut out a convex

© 2006 by Taylor & Francis Group, LLC

angle V from P that contains the whole surface F.[8] (V can be a half-plane.) The doubly-covered angle V is the tangent cone to F at the point O. Indeed, we can take two points X and Y on L arbitrarily close to O such that the half-lines OX and OY are arbitrarily close to T_1 and T_2; moreover, the triangle OXY belongs to F. Then under an indefinite dilation of the surface F, this triangle in the limit covers the whole angle made by OX and OY. But since the half-lines OX and OY can be taken arbitrarily close to the half-tangents T_1 and T_2, it is clear that the whole angle made by T_1 and T_2 turns out to be covered in the limit by the increasing surface F. Thus, this angle is the limit of surfaces homothetic to F.

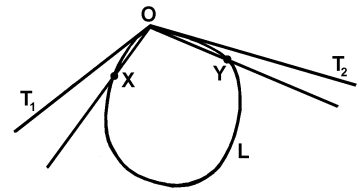


Fig. 45

Assume now that the surface F does not lie in any plane. If we complement this surface so as to make it the boundary of a convex body H then the resulting body has interior points.

Let us draw rays from the point O through all interior points of the body H and join all rays that are the limits of drawn rays to these rays. As a result, we obtain a convex solid cone V that includes the body H.[9] The surface of this cone is the tangent cone to the surface F at the point O.

Indeed, let X be some point on the surface of the cone V that differs from the point O. The ray OX is the limit of rays OX_n passing through interior points X_n of the body H. The plane P_n passing (for a given n) through the points O, X, and X_n intersects the surface F near O along a convex curve L_n one of whose branches lies in the angle made by the rays OX and OX_n, since the ray OX_n passes through interior points of the body H and L_n is a part of its boundary. (Figure 46 shows the section of the body H by the plane P_n.) Under a dilation of the surface F, this branch of the curve L also increases and remains in the angle made by OX and OX_n. But since this angle can be arbitrarily small for large n, this makes it clear that the point X is the limit of the points lying on the dilated curves L_n, i.e., is the limit of points lying on the dilated surfaces F.

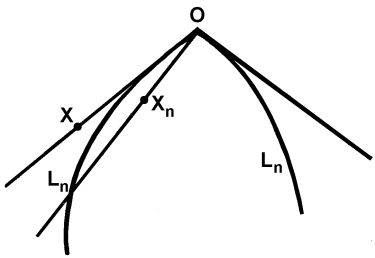


Fig. 46

On the other hand, if the point X lies outside the cone V then this point cannot be the limit

[8] Under an indefinite dilation of the curve L at the point O, this curve remains convex and, therefore, converges to two half-lines emanating from O. These lines are the left and right half-tangents.

[9] If two points A and B lie on the rays going from O through two points A_0 and B_0 of the body H, then the segment AB consists of the points lying on the rays going from O through the points of the segment A_0B_0. But this segment is contained in H by the convexity of H. Hence, projecting H from O, we obtain a convex cone. The closure of this cone is taken as K. The closure of a convex set is convex, since as $A_n \to A$ and $B_n \to B$, the segment AB is the limit of the segments A_nB_n.

© 2006 by Taylor & Francis Group, LLC

of points lying on the dilated surfaces F_n, since all these surface obviously remain in the cone V. If the point X lies inside the cone V, then the ray OX passes through interior points of H. Otherwise, a support plane passes through the ray OX that cuts out a part of the cone V which cannot contain rays that are the limits of the rays going from O through interior points of the body H. But since the ray OX passes through interior points of body H, there is a cone around this ray that consists of similar rays. Therefore, this cone contains neither points of the surface F nor points of the surfaces homothetic to F; except, certainly, for the point O. Therefore, the point X lying inside the cone V cannot be the limit of points of surfaces homothetic to F.

Consequently, all points of the surface of the cone V and these points only are the limits of points of surfaces homothetic to F; this means that the surface of the cone V is a limit of surfaces homothetic to F.

We will prove now that a convex surface in an infinitely small neighborhood of its every point O is isometric (to within infinitesimals of higher order) to the tangent cone at the point O. But before doing this, it is necessary to precisely define the meaning of the words "isometric to within infinitesimals of higher order."

Let F and F' be two surfaces, and let O and O' be two points on these surfaces. We say that in neighborhoods of the points O and O', the surfaces F and F' are isometric to within infinitesimals of higher order or "*infinitesimally isometric*" if there exists a mapping h of a neighborhood of the point O onto the neighborhood of the point O' having the following properties:

1. h is one-to-one and bicontinuous;
2. $H(O) = O'$; i.e., the mapping h transforms O into O';
3. for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|\rho_F(XY) - \rho_{F'}(X'Y')| < \varepsilon \max[\rho_F(OX), \rho_F(OY)] \tag{1}$$

whenever $\rho_F(OX), \rho_F(OY) < \delta$, where ρ_F and $\rho_{F'}$ are the distances on F and F', the points X and Y lie in the indicated neighborhood of the point O, and X' and Y' are the images of these points under the mapping h.

(We can write

$$|\rho_F(XY) - \rho_{F'}(X'Y')| < \varepsilon' \max[\rho_{F'}(OX), \rho_{F'}(OY)]$$

instead of (1), where ε' is connected with ε by the equation $\varepsilon' = \varepsilon/(1-\varepsilon)$. This makes it clear that the surfaces play the same role here.)

Each mapping having the above properties is called an *"infinitesimally isometric"* mapping of a neighborhood of the point O onto a neighborhood of the point O'.

Theorem 1. *Each convex surface is infinitesimally isometric to the tangent cone at an arbitrary point of this surface in a neighborhood of this point.*

In other words, a neighborhood of each point O of a convex surface F admits infinitesimally isometric mapping h onto a neighborhood of the same point in the tangent cone.

© 2006 by Taylor & Francis Group, LLC

Consequently, in accordance with this definition, the mapping h is such that (1) h is a homeomorphism; (2) $h(O) = O'$; (3) for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|\rho_F(XY) - \rho_K(X'Y')| < \varepsilon \max[\rho_F(OX), \rho_F(OY)]$$

whenever $\rho_F(OX), \rho_F(OY) < \delta$, where $\rho_K(X'Y')$ is the distance on the tangent cone between the images X' and Y' of the points X and Y.

If the cone K does not degenerate into a doubly-covered angle on the plane, then we can take the projection along each direction passing inside the cone K as the mapping h.

If the cone K degenerates into a doubly-covered angle V on the plane P, then the surface F itself lies in P. We can assume that F is closed, since we consider the properties of this surface in the small. Then F doubly covers some convex domain on the plane P. The point O lies on the boundary of this domain. Take some direction passing inside the angle V in the plane P and draw the line parallel to this direction through each point of the surface F. Let one of these lines l leave the surface F at the point X and pass through one of the legs of the angle V at the point X_1. Then we displace the whole segment of the line l lying in F so that the point X attains the point X_1. This operation defines a mapping of the surface F to the cone K if we agree to send each side of the surface F to the corresponding side of the cone K. It is easy to see that under this condition this mapping is bijective and bicontinuous in a neighborhood of the point O. This mapping can be taken as the mapping h in the theorem.

Now we prove our theorem.

Proof. Since we speak about the properties of the surface in the small, we can assume that F is closed in order to avoid extra stipulations. Let O be a point on F, and let K be the tangent cone to F at the point O. We will denote by λM the image of M under the homothety with coefficient λ and with the center at the point O.

Let ρ_λ be the distance on the surface λF; in particular, ρ_1 is the distance on F itself, let ρ_K be the distance on the cone K, and let ρ_0 be the distance in the space.

We restrict ourselves to the points of the surfaces λF whose distances from the point O are less than 1, i.e.,

$$\rho_\lambda(OX) < 1. \tag{2}$$

Circumscribe a ball S of a large radius R, say, $R > 2$, around the point O. The piece of the surface of the ball S that is cut out by the cone K, together with the part of the cone lying in the ball S, comprises a closed convex surface K_1. Similarly, the part of the surface λF lying in the ball S, together with the piece of the ball itself lying inside λF, forms a closed convex surface F_λ. And since the ball S is of a large radius, the distances on λF and F_λ are the same for the points satisfying condition (2). Analogously, the distances on K and K_1 are the same for the points whose distance from O is less than a half of the radius of the ball S. Therefore, if we restrict ourselves to only these points, then instead of distances on F_λ and K_1, we can take directly the distances on λF and K. Meanwhile, since the surfaces λF converge to the tangent cone K as $\lambda \to \infty$, the surfaces F_λ also converge to K_1. The surfaces F_λ and K_1 are closed, and so Theorem 2 of Sec. 1 on the convergence

© 2006 by Taylor & Francis Group, LLC

of metrics is applicable to them. Using this theorem and restricting consideration to points that are not very distant from the point O, we can assert the following: for any $\varepsilon > 0$ there exist λ_0 and $\delta > 0$ such that

$$|\rho_F(XY) - \rho_K(X'Y')| < \varepsilon, \tag{3}$$

$$(\rho_\lambda(OX), \rho_\lambda(OY) < 1, \quad \rho(XX'), \rho(YY') < \delta, \quad \lambda > \lambda_0)$$

whenever $\lambda > \lambda_0$ and the distances from the points X and Y on λF to the points X' and Y' on K are less than δ.

Assume now that the cone K does not degenerate. Let H stand for the projection along some direction going inside K. Then, since λF converge to K, there exists λ_1 such that for any $\lambda > \lambda_1$, the distance from the point X on λF to its projection on K is less than δ whenever the distance from X to O is less than 1, that is,

$$\rho(X, h(X)) < \delta \quad (\rho_\lambda < 1). \tag{4}$$

Now if we take any $\lambda > \lambda_0$ and $> \lambda_1$, then by (3) and (4), we have

$$|\rho_\lambda(XY) - \rho_K(h(X)h(Y))| < \varepsilon \tag{5}$$

if

$$|\rho_F(OX), \rho_\lambda(OY)| < 1 \quad \text{and} \quad \lambda > \max(\lambda_0, \lambda_1) = \lambda_2.$$

Now let A and B be those points on the initial surface F that pass to the points X and Y under the homothetic dilation of F by λ times, i.e., $X = \lambda A$ and $Y = \lambda B$. The distance on the surface also increases by λ times under this homothetic dilation, and, therefore,

$$\rho_\lambda(XY) = \lambda\rho_1(XY). \tag{6}$$

At the same time, under each homothety, the projection becomes a projection; therefore,

$$h(X) = \lambda h(A), \quad h(Y) = \lambda h(b).$$

Finally, under our homothety transformations with the center O, the cone K goes to itself, and if the points $h(A)$ and $h(B)$ go to the points $h(X)$ and $h(Y)$, then the distance between these points becomes λ times greater, i.e.,

$$\rho_K(h(X)h(Y)) = \lambda\rho_K(h(A)h(B)). \tag{7}$$

Now, using Eqs. (6) and (7), instead of (5), we obtain

$$|\rho_1(AB) - \rho_K(h(A)h(B))| < \frac{\varepsilon}{\lambda}$$

if

$$\rho_1(OA), \rho_1(OB) < \frac{1}{\lambda} \leq \frac{1}{\lambda_2}.$$

Taking

$$\frac{1}{\lambda} = \max[\rho_1(OA), \rho_1(OB)] \quad \text{and} \quad \delta = \frac{1}{\lambda_0},$$

we see that these formulas yield the assertion of the theorem.

© 2006 by Taylor & Francis Group, LLC

If the cone K degenerates into a doubly-covered plane angle, then instead of projection we take the above-described mapping as h. Moreover, all arguments remain the same almost literally, so that it is wasteful to repeat them in this case.

Of course, the above theorem does not belong to intrinsic geometry since it involves the concept of tangent cone which is defined in a purely extrinsic manner. However, in this theorem, we proved a very important property of the intrinsic metric of a convex surface and if we slightly change the statement of this theorem, then this theorem acquires the following purely intrinsic form.

For each point O of a convex surface there exists a cone K such that a neighborhood of the point O admits an "infinitesimally isomorphic" mapping onto a neighborhood of the cone K.

Here, the cone K itself is, of course, considered from the point of view of its intrinsic metric. Since the tangent cone to a convex surface is convex, then from the point of view of intrinsic geometry this cone is characterized by the fact that the complete angle at its vertex is $\leq 2\pi$.

In conclusion, we note that considering a surface, as well as the tangent cone, only from the point of view of intrinsic geometry, we arrive naturally at the following concept. Let a point O in a manifold R with some metric admit an "infinitesimally isomorphic" mapping onto a neighborhood of the vertex of some cone K. In this case, we say that this cone K is the "tangent cone of the manifold R at the point O." This concept proves useful in the theory of manifolds with intrinsic metric; this is a generalization of the concept of tangent Euclidean plane accepted in the theory of two-dimensional Riemann manifolds, i.e., in the final analysis, the concept of a line element. However, we would not use it in the general form in what follows, and would not dwell on it here.

6. The Spatial Meaning of the Angle between Shortest Arcs

The existence of a half-tangent to a shortest arc on a convex surface was proved by I. M. Liberman in his beautiful paper "Geodesics on convex surfaces."[10] The proof is based on the lemma by Busemann and Feller, which we have used in the proof of the theorem on convergence of metrics (Lemma 2 in Sec. 1 of Chapter III). Using this lemma, Liberman applies an elegant geometrical argument and obtains not only the existence of a half-tangent but also deduces a number of important extrinsic-geometrical properties of shortest arcs on convex surfaces. We reproduce here the proof of a part of the Liberman results. Since we shall speak about the properties of shortest arcs in the small, we can consider closed convex surfaces, because each finite convex surface can be completed to a closed surface without changing its metric in sufficiently small domains.

Lemma 1. *Let F be a convex surface that does not degenerate into a plane domain, and let L be a shortest arc on F. Take an arbitrary point O on L and draw parallel*

[10] Dokl. Akad. Nauk SSSR, Vol. XXXII, No. 5 (1941), pp. 310–313. Owing to the generality and simplicity of the method, Liberman obtained many important theorems on shortest arcs on convex surfaces in the space of arbitrary dimension in the framework of this small note. These theorems, together with their proofs, are abstracted word by word to convex surfaces in n-dimensional spaces of constant curvature. The proof of existence of a tangent to a shortest arc at those points where the convex surface under study has a tangent plane was given by Busemann and Feller.

© 2006 by Taylor & Francis Group, LLC

lines through all points of L that are sufficiently close to O so that these lines pass through interior points of the convex body bounded by the surface F. If we develop the cylinder obtained in such a way onto the plane, then a segment of the shortest arc that lies on this cylinder goes to a convex curve. This curve is convex with respect to the side corresponding to the part of the cylinder C issuing from the body bounded by the surface F.

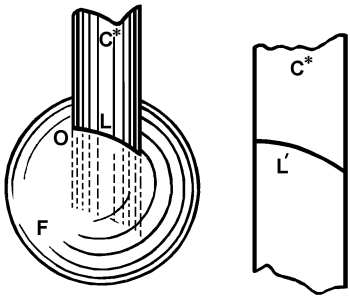


Fig. 47

Proof. Let L' be the curve that is the image of the shortest arc L under the development of the cylinder C onto the plane (Fig. 47). Let C^* be the part of the cylinder C which issues from the body and is bounded by the surface F. Assume that the curve L does not have the property claimed in the lemma. Then we can find two points A' and B' on this curve such that the chord $\overline{A'B'}$ lies on the part C^* of the developed cylinder C. If $\overset{\frown}{A'B'}$ is the corresponding arc of the curve L' then, obviously,

$$\overset{\frown}{A'B'} > \overline{A'B'}. \tag{1}$$

On the cylinder C itself, we associate a line $\overline{AB}$ and an arc $\overset{\frown}{AB}$ of the shortest arc L with the chord $\overline{A'B'}$ and the shortest arc $\overset{\frown}{A'B'}$ (the points A and B are the images of A' and B' under the development of the cylinder).

Herewith,

$$\overline{A'B'} = \overline{AB}, \quad \overset{\frown}{A'B'} = \overset{\frown}{AB}\ . \tag{2}.$$

The line $\overline{AB}$ lies on the part C^* of the cylinder C issuing from the body bounded by the surface F. Hence, by Lemma 2 in Sec. 1 of Chapter III, this line is not shorter than the shortest arc $\overset{\frown}{AB}$, i.e.,

$$\overset{\frown}{AB} \leq \overline{AB}. \tag{3}.$$

But, by inequality (2), this inequality contradicts inequality (3). Therefore, our assumption is impossible, and the lemma is proved.

Lemma 2. *Let O be a point on a convex surface F, and let X_i $(i = 1, 2, \ldots)$ be a sequence of points on the same surface which converges to O. Then the ratio of the distance from O to X_i in the space and that on the surface F tends to* 1:

$$\lim_{X_i \to O} \frac{\rho(OX_i)}{\rho_F(OX_i)} = 1, \tag{4}$$

where ρ is the distance in the space and ρ_F is the distance on the surface.

Proof. Let us project a neighborhood of the point O on the tangent point K to the surface F at the point O along some direction going inside the cone K.

© 2006 by Taylor & Francis Group, LLC

Let X_i' be the projections of the points X_i, and let ρ_K be the distance on K. Obviously,

$$\lim_{X_i \to O} \frac{\rho_K(OX_i')}{\rho(OX_i)} = 1, \tag{5}$$

while, by the theorem on the tangent cone which is proved in Sec. 5,[11]

$$\lim_{X_i \to O} \frac{\rho_K(OX_i')}{\rho_F(OX_i)} = 1. \tag{6}$$

Equations (5) and (6) imply (4) as required.

Theorem 1. *Let the end of the vector $x(s)$ drawn from some origin circumscribe a shortest arc on a convex surface, and, moreover, let s be the arclength of this shortest arc which is measured from one of its ends. Then the function $x(s)$ has the right and left derivatives, $x_r'(s)$ and $x_l'(s)$, at each point, and, moreover, $|x_r'|^2 + |x_l'|^2 = 1$. (Of course, there is only one of these derivatives at the ends of this shortest arc.)*

Proof. The shortest arc L is one or two line segments on the surface generating into a doubly-covered plane domain. Therefore, in this case, the theorem is obvious, and we can assume that we deal with a nondegenerate surface. It is sufficient to restrict exposition to considering the right derivative, since both derivatives play the same role. By definition,

$$x_r'(s) = \lim_{h \to 0} \frac{x(s+h) - x(s)}{h}.$$

Obviously, $|x(s+h) - x(s)|$ is the length of the chord and h is the length of the corresponding arc of the shortest arc. Therefore, Lemma 2 implies that if $x_r'(s)$ exists, then $|x_r'(s)|^2 = 1$. Therefore, it remains to prove the existence of $x_r'(s)$. To this end, we take some point O on the shortest arc L, isolating, if need be, some finite domain containing this point from our convex surface and completing this domain to a convex surface that does not degenerate into a plane domain. Take three points X, Y, and Z inside this surface F such that these points do not lie in one plane with the point O. Drawing three lines OX, OY, and OZ, we take them as the coordinate axes; the point O is the origin of this coordinate system. Let $x(s)$, $y(s)$, and $z(s)$ be the orthogonal projections of the vector $x(s)$ drawn from O to a variable point on the shortest arc L. If we prove that each of the functions $x(s)$, $y(s)$, and $z(s)$ has the right derivative at the point O, then we come to the desired result.

For example, we consider $x(s)$. We draw lines parallel to the line OX through all points of the shortest arc L which are close to O. Since the point X lies inside the surface F, the lines drawn pass through interior points of the surface F, i.e., inside the body bounded by this surface. These lines comprise some cylinder.

If we develop this cylinder onto the plane, then according to Lemma 1 the segment of the shortest arc L lying on this cylinder goes to a convex curve L'. The values $x(s)$ are nothing else but the coordinate of a point of the shortest arc L along

[11] According to this theorem, for $|\rho_F(OX_i)| < \delta$, we have $|\rho_F(OX_i) - \rho_K(OX_i')| < \varepsilon\rho_F(OX_i)$ or $\left|\frac{\rho_K(OX_i')}{\rho_F(OX_i)} - 1\right| < \varepsilon$, so that we obtain (6) as $\rho_F(OX_i) \to 0$.

© 2006 by Taylor & Francis Group, LLC

the axis OX.[12] Since the curve L' is convex, there is a right derivative of $x(s)$ with respect to the arclength at each point of this curve,[13] i.e., with respect to s, since length is preserved under development. The theorem is proved.

The geometric meaning of this theorem consists in the following two assertions: (1) *a shortest arc on a convex surface has the right and left tangents at each point;* (2) *the ratio of a chord and an arc of each shortest arc tends to* 1 *when one of the endpoints of this arc is fixed while the arclength tends to zero.*

If we take a point O on a shortest arc L and dilate the surface at this point to infinity, then the surface transforms to the tangent cone, while the right and left branches of the shortest arc L go to the right and left half-tangents. The right (left) half-tangent is the limit of half-lines drawn from the point O through a point infinitely close to O and lying on the curve to the right (left) from O. This easily implies that this tangent is also the limit of the right (left) branch of the curve under an indefinite dilation of this curve at O.

Theorem 2. *The angle made by two shortest arcs emanating from a point O on a convex surface F is equal to the angle made by their half-tangents which is measured on the cone K tangent to F at the point O.*

Proof. Take a point O on a convex surface F and project a neighborhood of this point to the tangent cone at this point along a direction inside this cone, or in the case the cone degenerates to a doubly-covered plane angle, construct the mapping of a neighborhood of the point O which is described in Sec. 5. Under an indefinite dilation of the cone at its vertex, the projections of shortest arcs emanating from O converge to the half-tangents.

Let L and M be two shortest arcs emanating from the point O, L' and M' be the half-tangents to these arcs, and let X and Y be variable points on L and M converging to O in such a way that the ratio $\rho_F(OY)/\rho_F(OX)$ remains within positive bounds. Then we prove that *the angle $\gamma(x, y)$ in the plane triangle with sides $x = \rho_F(OX)$, $y = \rho_F(OY)$, and $z = \rho_F(XY)$ converges to the angle made by the half-tangents L' and M'.* Since the limit of the angle $\gamma(x, y)$ is the angle made by these shortest arcs, the theorem will be proved.

Choose some sequence of points X_n and Y_n such that the ratio $\rho_F(OY_n)/\rho_F(OX_n)$ tends to some limit. Project the shortest arcs L and M on the cone K, and let X'_n and Y'_n stand for the projections of the points X_n and Y_n. To each position of the points X_n and Y_n, we put in correspondence the homothety transformation with the center of homothety at the point O and coefficient of homothety equal to $1\rho_K(OX'_n)$. The points X''_n to which the points X'_n pass under this homothety lie at the constant distance 1 from the point O, i.e., $\rho_K(OX''_n) = 1$. The projections of the shortest arcs L and M converge to the half-tangents L' and M', since the coefficients of this homothety $1/\rho_K(OX'_n)$ increases infinitely. Hence

[12] The coordinate is determined by the projection to the axis X along the plane orthogonal to the axis OX, but this plane is not parallel to the plane OYZ in general.

[13] A convex curve has the right tangent at each point. If $\eta = f(\xi)$ is the equation of this curve in the Cartesian coordinates, then there exists $(d\eta/d\xi)_{\text{right}}$. At the same time, the ratio of the length of the chord $\sqrt{\Delta\xi^2 + \delta\eta^2}$ and the arclength tends to 1 as $\Delta\xi \to 0$ by Lemma 2. This makes it clear that there exists the right derivative of η with respect to the arc.

© 2006 by Taylor & Francis Group, LLC

the points X_n'' converge to some point X'' on L'. The points Y_n'' corresponding to the points Y_n' also converge to some point Y'' on M'' different from the point O. Indeed, the ratio $\rho_F(OY_n)/\rho_F(OX_n)$ tends to a positive limit, and thus, the ratio $\rho_K(OY_n')/\rho_K(OX_n')$ tends to the same limit as follows from the theorem of the previous section.[14] But the ratio of distances does not change under a homothetic enlargement, and hence, $\rho_K(OY_n'')/\rho_K(OX_n'') = \rho_K(OY_n'')$ tends to the same limit. Besides, since the points Y_n'' approach the half-tangent M', they converge to some point Y'' on M' different from O.

Compare the ratio of the distances

$$\rho_F(OX_n) : \rho_F(OY_n) : \rho_F(X_nY_n)$$

and

$$\rho_K(OX_n') : \rho_K(OY_n') : \rho_K(X_n'Y_n').$$

As we have already mentioned, the theorem of the preceding section implies that the limits of the ratios $\rho_F(OY_n) : \rho_F(OX_n)$ and $\rho_K(OY_n') : \rho_K(OX_n')$ are equal. By the same theorem,

$$\frac{\rho_F(X_nY_n) - \rho_K(X_n'Y_n')}{\max[\rho_F(OX_n), \rho_F(OY_n)]} \to 0. \tag{7}$$

Since the ratio $\rho_F(OY_n) : \rho_F(OX_n)$ is bounded, we can simply take $\rho_F(OX_n)$ in the denominator. Moreover, the ratio $\rho_K(OX_n') : \rho_F(OX_n)$ tends to 1, and, therefore, we can write

$$\frac{\rho_F(X_nY_n)}{\rho_F(OX_n)} - \frac{\rho_K(X_n'Y_n')}{\rho_K(OX_n')} \to 0$$

instead of (7). Therefore, the limits of all ratios

$$\rho_F(OX_n) : \rho_F(OY_n) : \rho_F(X_nY_n)$$

and

$$\rho(OX_n') : \rho_K(OY_n') : \rho_K(X_n'Y_n')$$

should be equal to each other.

But the ratios of distances do not change under dilations. Therefore, in the latter ratios, the points X_n' and Y_n' can be replaced by the points X_n and Y_n, respectively. But the points X_n' and Y_n' converge to the points X'' and Y'', and so we finally conclude that the limits of the ratios

$$\rho_F(OX_n) : \rho_F(OY_n) : \rho_F(X_nY_n)$$

are equal to the ratios

$$\rho_K(OX'') : \rho_K(OY'') : \rho_K(X''Y'').$$

[14] In footnote 11 on page 149 (in the proof of Lemma 2), we have already proved that $\frac{\rho_K(OX_n')}{\rho_F(OX_n)} \to 1$, and similarly, $\frac{\rho_K(OY_n')}{\rho_F(OY_n)} \to 1$; this implies the assertion just formulated.

© 2006 by Taylor & Francis Group, LLC

This implies that the angles of the plane triangles T_n with sides equal to $x_n = \rho_F(OX_n)$, $y_n = \rho(OY_n)$, and $z_n = \rho_F(X_nY_n)$ converge to the angles of the triangle $OX''Y''$ (this is a triangle on the cone, and so it is isometric to the plane triangle with sides $\rho_K(OX'')$, etc.) The angle at the vertex O of the triangle $OX''Y''$ is the angle α' made by the half-tangents L' and M', and the angle of the triangle T_n corresponding to this angle is $\gamma(x_n, y_n)$. Hence

$$\lim_{n\to\infty} \gamma(x_n, y_n) = \alpha'.$$

This is proved for each sequence X_n, Y_n such that the ratio $\rho_F(OY_n) : \rho_F(OX_n)$ has a positive bound. Hence the same is true if the points X and Y tend to O in such a way that the ratio $\rho_F(OY) : \rho_F(OX)$ remains within positive bounds. Since $\lim_{x,y\to 0} \gamma(x, y)$ is the angle made by the shortest arcs L and M, this proves the theorem.

The fact that at each stage of the proof we do not use the existence of the limit $\lim_{x,y\to 0} \gamma(x, y)$ is worthy of consideration. The existence of such a limit was proved for all sequences such that the ratio $y/x = \rho_F(OY)/\rho_F(OX)$ remains within positive bounds. Therefore, if we weaken the definition of angle between shortest arcs by admitting only such sequences x and y, then the existence of the angle itself follows from the theorems on the tangent cone and the existence of half-tangents to shortest arcs. Of course, this result is weaker than the existence of $\lim_{x,y\to 0} \gamma(x, y)$ for all sequences $x, y \to 0$ and cannot be strengthened, since we can give examples of surfaces, certainly nonconvex, for which the theorems on the tangent cones and the existence of half-tangents are valid, whereas $\lim_{x,y\to o} \gamma(x, y)$ exists not for all sequences $x, y \to 0$. Nevertheless, this opens up a possibility of proposing a theory on the angle between shortest arcs based on the above theorems without attracting the convexity condition.

Theorem 3. *The complete angle at a point O on a convex surface F is equal to the complete angle of the tangent cone K of the surface F at the point O.*

Proof. Let θ be the complete angle at the point O of the surface F. As is shown in Sec. 3, this angle is equal to the least upper bound of the sums of the angles between neighboring shortest arcs drawn from O (the number of these arcs is arbitrary). Therefore, for each $\varepsilon > 0$, there exist shortest arcs $L_1, \ldots, L_n$ emanating from O such that

$$\alpha_{12} + \alpha_{23} + \cdots + \alpha_{n-1,n} + \alpha_{n1} > \theta\varepsilon; \tag{8}$$

herewith, the shortest arcs $L_1, \ldots, L_n$ are assumed to be enumerable in the order of their location around the point O, and α_{ij} stands for the angle between L_i and L_j.

As we have already proved, the angles α_{ij} are equal to the angles made by the half-tangents to the shortest arcs L_i and L_j. Obviously, these half-tangents are located on the tangent cone K in the same order. The complete angle θ' at the

© 2006 by Taylor & Francis Group, LLC

vertex of the cone K is also equal to the least upper bound of the sums of angles made by its generators.[15] Therefore,

$$\theta' \geq \alpha_{12} + \cdots + \alpha_{n1}.$$

Comparing this inequality with inequality (8), we obtain $\theta' > \theta - \varepsilon$; since ε is arbitrary, we have

$$\theta' > \theta. \tag{9}$$

Under the dilation of the surface F with coefficient λ, this surface goes to the surface λF with the same complete angle of the point O; the surface λF converges to the cone K as $\lambda \to \infty$. We have proved in Sec. 4 that the lower limit of complete angles is no less than the complete angle on the limit surface; therefore, $\liminf_{\lambda\to\infty} \theta_\lambda \geq \theta'$, and since $\theta_\lambda = \theta$, we have $\theta \geq \theta'$. Comparing this inequality with inequality (9), we see that $\theta = \theta'$; as required.

Theorem 4. *Let two shortest arcs L and M emanating from a point O on a convex surface F divide a neighborhood of this point into two sectors U and V. As we have shown, under an indefinite dilation at the point O, the shortest arcs L and M converge to their half-tangents L' and M', while the sectors U and V converge to the sectors U' and V' which are bounded by the half-tangents L' and M' on the cone K tangent to F at the point O. The angles of the sectors U' and V' are equal to the angles of the sectors U and V.*

Proof. Since under an indefinite dilation the surface F converges to the cone K and the shortest arcs L and M converge to the half-tangents L' and M', the sectors U and V converge to the sectors bounded by the generators L' and M' on K. Angles are not changed under homothetic transformations. Therefore, the complete angles around the point O on the surfaces λF obtained from F by the homothetic enlargement are constantly equal to the complete angle of the cone K. Then we can apply Theorem 3 of Sec. 4, which asserts that in this case, the limit of the angles of sectors is equal to the angle of the limit sector. The angles of all sectors obtained from U (or V) by the homothetic enlargement are equal to each other. Consequently, these angles are equal to the angle of the limit sectors U' (or V'); this is what was required to prove.

Thus, the angles of sectors on a surface reduce to the angles of sectors on the tangent cone, and so they obtain a simple visual meaning. The angles of sectors on a surface could be defined as the angles of the corresponding sectors on the tangent cone; and the complete angle at a point, as the complete angle of the tangent cone. Then we would immediately obtain all properties of the angles of sectors which are deduced in Sec. 3. Comparing this remark with what was said after Theorem 2 on the angle between shortest arcs, we see that the theory of angles can in fact be constructed on the basis of the theorem on the tangent cone and the existence of half-tangents. Such a way of presentation is possibly even more visual in the first sections of this chapter. The deficiency of this presentation consists in the fact, also

[15] Of course, this angle is equal to the sum of angles made by the generators if they divide the cone into sufficiently small sectors; but it is possible that $\theta' > \alpha_{12} + \ldots + \alpha_{n1}$ without this condition.

© 2006 by Taylor & Francis Group, LLC

mentioned after Theorem 2, that the theorem on the existence of angle is obtained in a weakened form. Besides, the concepts of tangent cone and half-tangent involved here are external to the intrinsic metric of a surface.[16]

The essential distinction between the intrinsic and "extrinsic" points of view is revealed, e.g., in the theorems on convergence of angles. We have proved in Sec. 4 that if convex surfaces F_n converge to F and if the complete angles at points O_n on the surfaces F_n converge to the complete angle at the limit point O on the surface F, then the angles of sectors with vertices at O_n converge to the angles of the limit sectors with vertices at O (Theorem 3 of Sec. 4). Liberman proved the following theorem on the convergence of half-tangents to shortest arcs.

Theorem. *Let convex surfaces F_n converge to a convex surface F, let points O_n on the surfaces F_n converge to a point O, and let shortest arcs L_n emanating from the points O_n converge to a shortest arc L. If the tangent cones of F_n at O_n converge to the tangent cone of F at O, then the half-tangents to the shortest arcs L_n at O_n converge to the half-tangents to the shortest arc L at O.*[17]

This theorem makes Theorem 3 of Sec. 4 more precise if the tangent cones at O_n converge to the tangent cone at O. However, Theorem 3 of Sec. 4 is not implied by the Liberman theorem, since the convergence of the complete angles of the tangent cones does not imply the convergence of the cones themselves. Indeed, if the tangent cone at the point O reduces to a dihedral angle, then its complete angle is equal to 2π, and then, as was shown in Sec. 4, the complete angles at the points O_n converge to the complete angle at the point O. However, the tangent cones at the points O_n can reduce to planes and do not converge to the tangent cone at the point O as can be seen by examining very simple examples. Then the half-tangents to shortest arcs emanating from the points O_n *a priori* do not converge to the half-tangents of shortest arcs emanating from the point O. We can show that this is the only case where the convergence of the complete angles at points on convergent convex surfaces does not imply the convergence of the tangent cones.

From the intrinsic point of view, those points at which the tangent cone turns out to be a dihedral angle are not different from the points at which there is a tangent plane. Points inside edges of a polyhedron can serve as such examples. However,

[16] As was mentioned in Sec. 5, we can attach a purely intrinsic geometric meaning to the concept of tangent cone. A half-tangent can also be treated in a similar way. Let K be the tangent cone to a surface F at a point O (in the sense indicated at the end of Sec. 5). Let L be a curve on F emanating from O. If the image of L on the cone K has half-tangent L' at O then L' can be called the half-tangent of the curve L. Here, we speak about the image of L under some infinitesimally isometric mapping. This mapping is not uniquely defined, and under one mapping the image can have a half-tangent and can fail to have it under another (it is easy to give respective examples). Hence, this definition entails some inconvenience. In any case, when using tangents, we go away from the surface. Also, we note that neither the existence of half-tangents nor that of a tangent cone alone imply the existence of an angle between shortest arcs in the sense of intrinsic geometry. We can give examples of surfaces (of course, nonconvex) at some points of which only one of the following assertions holds: either the existence of the tangent cone or the existence of half-tangents to shortest arcs; in these examples, not all shortest arcs emanated from the indicated points make definite angles with each other even in the weaker sense that was defined after Theorem 2.

[17] See the cited note by Liberman in Dokl. Akad. Nauk SSSR, Vol. XXXII, No. 5 (1941), pp. 310–313. In that note, this theorem is even not formulated, but it is obviously implied by the theorems presented therein. Still, a barely outlined idea of the proof allows us to reconstruct the proof completely.

© 2006 by Taylor & Francis Group, LLC

e.g., the points on the base of a cylinder of revolution differ from other points of this cylinder, since even in their arbitrarily small neighborhood the cylinder (together with its base) cannot be developed onto the plane. However, this distinction between points of the cylinder does not forbid their neighborhoods to be isomorphic to within infinitesimals of order higher than 1; only this accuracy is necessary for the definition of angle.

In conclusion, we present one more theorem by Liberman which states the general properties of shortest arcs as spatial curves.[18] To this end, we return to the arguments that prove the existence of a half-tangent. Drawing parallel lines through points of a shortest arc L that go to the interior of the body bounded by a given convex surface, we obtain a cylinder. Under the development of this cylinder onto the plane, the shortest arc goes to a convex curve L'. The coordinate $x(s)$ along the generators is the same for points on L and L'. And since the curve L' is convex, the right and left derivatives $x'_r(s)$ and $x'_l(s)$ of this coordinate with respect to the arclength have the following properties:

1. For all s, $x'_r(s)$ is right continuous, $x'_l(s)$ is left continuous, and

$$x'_r(s) = \lim_{h\to 0} x'_r(s+h) = \lim_{h\to 0} x'_l(s+h),$$
$$x'_l(s) = \lim_{h\to 0} x'_l(s-h) = \lim_{h\to 0} x'_r(s-h).$$

2. Everywhere, except for countable many values of s, we have

$$x'_r(s) = x'_l(s).$$

3. $x'_r(s)$ and $x'_l(s)$ are monotone functions.

   As is known, a monotone function has a derivative almost everywhere; therefore, (3) implies the following

4. For almost all values of s there is the second derivative $x''(s)$.

Applying the same conclusions to the coordinates $y(s)$ and $z(s)$ in two other directions and passing from the projections $x(s)$, $y(s)$, and $z(s)$ to the vector $x(s)$ itself, we finally arrive at the following theorem.

Theorem. *Let the end of the vector $x(s)$ circumscribe a shortest arc on a convex surface, and let s be the arclength of this shortest arc. Let $x'_r(s)$ and $x'_l(s)$ be the right and left derivatives of $x(s)$ in s. Then*
(1) *for all s,*

$$x'_r(s) = \lim_{h\to 0} x'_r(s+h) \lim_{h\to 0} x'_l(s+h),$$
$$x'_l(s) = \lim_{h\to 0} x'_l(s-h) = \lim_{h\to 0} x'_r(s-h),$$

i.e., the right and left half-tangents are continuous from the right and from the left, respectively;

[18] However, this theorem is not closely related to our topic, and we do not use this theorem in the sequel.

© 2006 by Taylor & Francis Group, LLC

(2) *everywhere, except for at most countable many points, we have* $x'_r(s) = x'_l(s)$, *i.e., the right and left half-tangents coincide*; (3) $x'_r(s)$ *and* $x'_l(s)$ *are of bounded variation*; (4) *for almost all* s, *there exists* $x''(s)$, *i.e., each shortest arc has curvature almost everywhere.*

Finally, we can mention that Busemann and Feller had proved the following assertion:

If a convex surface has the tangent plane P *at a point* O, *then the projection to* P *of each shortest arc emanating from* O *has zero curvature at the point* O, i.e., at the points where there is a tangent plane, the behavior of every shortest arc is the same as on a regular surface.

The proof of this theorem and further results on shortest arcs can be found in the already cited paper by Liberman.

© 2006 by Taylor & Francis Group, LLC

Chapter V

CURVATURE

1. Intrinsic Curvature

In this chapter, we present the foundations of the theory of curvature of convex surfaces. Some aspects of this theory were already given in Chapter I (Secs. 9–11). This theory is very important for intrinsic geometry, and we will not only give a number of its applications in the last section of this chapter, but will also constantly use it in the most essential way in what follows. Suffice it to say that the axiomatic definition of the intrinsic metric of a convex surface whose justification will be given in Chapter VIII originated essentially from the concept of curvature.

In essence, a surface has two curvatures, intrinsic and extrinsic; the first is a measure of deviation of intrinsic geometry of a surface from plane geometry, and the second characterizes the deformation of a surface in space. The definitions of these two curvatures, the study of their properties, and the relation between them form the subject of the curvature theory. In the sequel, we shall merely say "curvature" instead of "intrinsic curvature."

The curvature is defined as a set function, i.e., to a set M on a convex surface, we put in correspondence some number $\omega(M)$, the curvature of this set M on the surface F under consideration.

We begin with the definition of curvature for the following three types of "basic" sets: open triangles, open shortest arcs, and points.[1] An open triangle is a triangle with vertices and sides excluded; an open shortest arc is a shortest arc with ends excluded.

Let T be an open triangle on a convex surface, and let α, β, and γ be the angles of this triangle (i.e., the angles of the corresponding closed triangles that include into themselves the sides and the vertices). We take the number

$$\omega(T) = \alpha + \beta + \gamma - \pi \tag{1}$$

to be the curvature of this triangle T.

The curvature of every shortest arc K on a convex surface is taken to be zero:

$$\omega(L) = 0. \tag{2}$$

Let X be a point on a convex surface, and let θ be the complete angle at this point. Then the number

$$\omega(X) = 2\pi - \theta \tag{3}$$

[1] A set consisting of a single point X will merely be called a point X and denoted by X. The curvature of the empty set is zero by definition.

© 2006 by Taylor & Francis Group, LLC

is taken to be the curvature of the point X.

The same point X of the space can be a point of many convex surfaces, and the complete angles at these points on these surfaces can be distinct. Hence, the curvature of a point X depends not only on this point itself, but also on a convex surface F on which it is considered. An analogous remark can also be referred to other sets for which we define curvature. Therefore, to be more precise, we have to speak about "the curvature of a set M on a surface F" and write $\omega_F(M)$. However, we will constantly omit this reference to a surface F, introducing it only in the case where the same set M is simultaneously considered on several surfaces; if we consider only one surface, omitting the reference that a curvature is taken just on this surface leads to no confusion.

Each angle of a triangle on a convex surface is no less than the corresponding angle of the plane triangle with sides of the same length. Therefore, the sum of angles of a triangle on a convex surface is no less than π, and hence *the curvature of every open triangle is nonnegative.* The theorem on angles of a triangle, which is used here, was proved in the case of an arbitrary convex surface only for sufficiently small triangles (see Theorem 3 in Sec. 4 of Chapter III). Therefore, we shall consider *only* such triangles, and *only* the open triangles corresponding to them will be admitted to the collection of basis sets.

The complete angle at a point on a convex surface is always $\leq 2\pi$; therefore, *the curvature of a point is nonnegative.* Thus, the curvatures of all basis sets are nonnegative.

Let us consider all sets on some convex surface that can be represented as unions of pairwise disjoint "basis" sets. These sets can be called "elementary." For example, each geodesic polygon is an elementary set. Indeed, according to the theorem of Sec. 6 of Chapter II, each geodesic polygon P can be partitioned into arbitrarily small triangles. Taking the interiors of these triangles, their sides with excluded ends, and finally, their vertices, we obtain a representation of the polygon P as the union of disjoint "basis" sets.

Let M be an "elementary" set, and let

$$M = \sum_{i=1}^{n} B_i$$

be a representation of this set as a union of disjoint "basis" sets. *The number*

$$\omega(M) = \sum_{i=1}^{n} \omega(B_i) \tag{4}$$

is called the curvature of M. To make this definition more sound, we prove the following assertion: if $M = \sum_i B_i = \sum_j A_j$ are two representations of the set M in the form of a union of disjoint "basis" sets, then $\sum_i \omega(B_i) = \sum_j \omega(A_j)$; i.e., $\omega(M)$ does not depend on the choice of partition of the set M into "basis" sets. This assertion will be proved below.

The essence of this definition is that the curvature proves an additive set function, i.e., if M_1 and M_2 are disjoint, then

$$\omega(M_1 + M_2) = \omega(M_1) + \omega(M_2). \tag{5}$$

© 2006 by Taylor & Francis Group, LLC

Indeed, let $M_1 = \sum_i A_i$ and $M_2 = \sum_j B_j$ be partitions of the sets M_1 and M_2 into basis sets; then

$$M_1 + M_2 = \sum_i A_i + \sum_j B_j$$

is a partition of the union $M_1 + M_2$ into basis sets, and hence, by formula (4),

$$\omega(M_1 + M_2) = \sum_i \omega(A_i) + \sum_j \omega(B_j) = \omega(M_1) + \omega(M_2).$$

However, this proof of additivity of the curvature will make no sense until the soundness of the very definition of the curvature is proved; therefore, we cannot use additivity yet.

In order to prove that our definition of the curvature of "elementary" sets is sound, we prove the following theorem, which is also of interest in its own right.

Theorem 1. *Let P be the interior of a geodesic triangle with angles $\alpha_1, \alpha_2, \ldots, \alpha_n$ whose Euler characteristic is $\chi(P)$. The curvature of P is equal to*

$$\omega(P) = 2\pi\chi(P) - \sum_{i=1}^{n}(\pi - \alpha_i). \tag{6}$$

In a more detailed form, this theorem means the following: the interior P of each geodesic polygon is obviously an elementary set (since a geodesic polygon itself is an elementary set). Let P be partitioned into f triangles, i.e., let P be represented as a sum of f open triangles $T_1,\ldots T_f$, their m vertices $X_1 \ldots, X_m$, and p sides $L_1,\ldots,$ L_p. Then, by formula (4), it should be

$$\omega(P) = \sum_{i=1}^{f} \omega(T_i) + \sum_{j=1}^{m} \omega(X_j),$$

where all $\omega(L_k)$ can be omitted since they are equal to zero by definition. Hence the theorem asserts that for every partition of the polygon P into triangles T_i we have

$$\sum_i \omega(T_i) + \sum_j \omega(X_j) = 2\pi\chi(P) - \sum_{i=1}^{n}(\pi - \alpha_i). \tag{7}$$

In particular, this means that the sum on the left-hand side is the same for all partitions, and hence the curvature is defined soundly for the interior of a polygon.

Proof. Consider the polygon $\overline{P}$ whose interior is P. The polygon $\overline{P}$ is partitioned into triangles $T_1,\ldots,$ T_f. Admit those vertices of the triangles T_i which lie on its sides into the set of vertices of the polygon $\overline{P}$. Since the angles at these vertices are equal to π, this does not change the right-hand side of formula (7). Therefore, we can assume that these additional angles belong to the set of angles $\alpha_1, \ldots, \alpha_n$.

If ϕ_1^i, ϕ_2^i, and ϕ_3^i stand for the angles of the triangle T_i, then

$$\omega(T_i) = \phi_1^i + \phi_2^i + \phi_3^i - \pi.$$

© 2006 by Taylor & Francis Group, LLC

Summing over all triangles, we obtain

$$\sum_{i=1}^{f} \omega(T_i) = \sum_{i,j} \phi_j^i - f\pi. \tag{8}$$

On the other hand, the sum of angles at each interior vertex X_i of the partition is equal to the complete angle at this vertex, i.e., to $2\pi - \omega(X_i)$, while the sum of angles at a vertex on the boundary is equal to the angle of the polygon. Therefore, the sum of all angles of the triangles T_i can be represented as follows:

$$\sum_{i,j} \phi_j^i = \sum_{i=1}^{m} [2\pi - \omega(X_i)] + \sum_{i=1}^{n} \alpha_i = 2\pi m - \sum_{i=1}^{m} \omega(X_i) + \sum_{i=1}^{n} \alpha_i. \tag{9}$$

Inserting this expression into formula (8), we obtain

$$\begin{aligned} \sum_{i=1}^{f} \omega(T_i) + \sum_{i=1}^{m} \omega(X_i) &= (2m - f)\pi + \sum_{i=1}^{n} \alpha_i \\ &= (2m - f + n)\pi - \sum_{i=1}^{n} (\pi - \alpha_i). \end{aligned} \tag{10}$$

The number of all vertices of the partition is equal to the number of the interior vertices and those lying on the boundary, that is,

$$e = m + n. \tag{11}$$

The number of sides n in the partition can be calculated in the following way: each triangle has $3f$ sides, but the sides lying in the interior are counted twice, while the sides lying on the boundary only once. The number of the latter is equal to the number of all vertices of the polygon, i.e., to n. Hence

$$3f = 2k - n. \tag{12}$$

By the definition of the Euler characteristic,

$$\chi(P) = f - k + e.$$

Multiplying by 2 and using Eq. (12), we infer from this that

$$2\chi(p) = 2e - n - f.$$

Now, using Eq. (12), we obtain

$$2\chi(P) = 2m + n - f, \tag{13}$$

and, therefore, formula (10) can be rewritten as

$$\sum_{i=1}^{f} \omega(T_i) + \sum_{i=1}^{m} \omega(X_i) = 2\pi\chi(P) - \sum_{i=1}^{n} (\pi - \alpha_i);$$

as claimed.

© 2006 by Taylor & Francis Group, LLC

Recall that a closed surface is also assumed to be a geodesic polygon, only without boundary, and hence without angles. Since such a convex surface is homeomorphic to the sphere, the Euler characteristic of this surface is equal to 2, and, therefore, the curvature of every closed convex surface is equal to 4π according to the theorem proved.

The Euler characteristic of an n-gon homeomorphic to a disk is equal to 1, and therefore, the curvature of the interior of this n-gon is expressed by the formula $\omega(P) = \sum_{i-1}^{n} \alpha_i - (n-2)\pi$. In particular, we have $\omega(T) = \alpha_1 + \alpha_2 + \alpha_3 - \pi$ for a triangle. The curvature of a polygon with boundary is the sum of the curvature of its interior and the curvature of the vertices; the curvature of the sides with ends excluded is equal to zero, since the curvature of a geodesic with ends excluded is always equal to zero. Indeed, a geodesic line can be partitioned into a finite number of open shortest arcs and isolated points which are common ends of these shortest arcs. But a geodesic cannot pass through a point the angle at which is $< 2\pi$, i.e., through a point of curvature > 0. Therefore, the sum of curvatures of shortest arcs and points into which a given geodesic is partitioned is always equal to zero.

Since a triangle includes its boundary by definition, its curvature is the sum of of the curvature of its interior and the curvatures of vertices. However, we will never consider the curvature of a triangle, but only the curvature of its interior. Therefore, for brevity, we will merely say the "curvature of a triangle" but always keep in mind the curvature of its interior unless the contrary is explicitly stated.

Using Theorem 1, we now easily prove the following theorem.

Theorem 2. *The definition of the curvature of "elementary" sets given above is sound, i.e., if a set M is represented as two sums of pairwise disjoint "basis" sets*

$$M = \sum_i A_i = \sum_j B_j,$$

then

$$\sum_i \omega(A_i) = \sum_j \omega(B_j).$$

Proof. Assume that both partitions of the set M are implemented simultaneously; then M is partitioned into the sets A_iA_j, the intersections of sets of the first and second partitions. Take some set A_i of the first partition; then

$$A_i = \sum A_iB_j,$$

where many sets among the sets A_iB_j can be empty (namely, those for which A_i is disjoint from B_j). Let us inspect a possible structure of a nonempty set A_iB_j.

If one of the sets A_i and B_j is a point then A_iB_j is the same point.

If A_i and B_j are open shortest arcs then A_iB_j is either an open arc or a singleton, since two shortest arcs either overlap or have more than a single point in their intersection.

If A_i is an open shortest arc and B_j is the interior of a triangle (or vice versa), then A_iB_j is either a single shortest arc or consists of two open shortest arcs. Indeed, the shortest arc A_i cannot intersect each of the sides of the triangle B_j more than

© 2006 by Taylor & Francis Group, LLC

once (since these sides are also shortest arcs). Therefore, if A_i intersects only one or two sides, A_iB_j is a single open shortest arc, and if A_i intersects all three sides (goes away from the triangle through one side, enters the triangle through the second side, and goes away from the triangle through the third), then A_iB_j consists of two open shortest arcs.

Finally, let A_i and B_j be open triangles. Since their sides cannot intersect each other more than at a finite set of points, the set A_iB_j can be only a union of finitely many interiors of polygons. But each polygon can be partitioned into triangles.

Now we calculate the curvatures of the sets A_iB_j, of course, not counting empty sets since their curvatures are equal to zero.

If A_i is a point X, then, obviously,

$$\omega(A_i) = \omega(X). \tag{14}$$

If A_i is an open shortest arc then the above consideration of the sets A_iB_j implies that the partition into the sets A_iB_j is a partition into a finite number of open shortest arcs L^p and points X^q. The curvatures of all points X^q are equal to zero, since a shortest arc cannot pass through a point of nonzero curvature. Therefore,

$$\omega(A_i) = \sum \omega(L^p) + \sum \omega(X^q) = 0. \tag{15}$$

Finally, if A_i is an open triangle, then its partition into the sets A_iB_j can be refined so that we obtain a partition into finitely many open triangles T^s, open shortest arcs M^r, and points Y^e. By Theorem 1,[2]

$$\omega(A_i) = \sum \omega(T^s) + \sum \omega(M^r) + \sum \omega(Y^e). \tag{16}$$

Summing formulas (14), (15), and (16), we see that the sum of curvatures of the sets A_i is represented as the sum of curvatures of sets that have the following properties: (1) each of these sets lies in some A_i and some B_j simultaneously; (2) the union of these sets is the given set M, and hence these sets cover all A_i as well as B_j. Therefore, it is clear that applying the same argument we have just used for the sets A_i to the sets B_j, we obtain the same expression for the sum of their curvatures. Hence,

$$\sum \omega(A_i) = \sum \omega(B_j);$$

as required.

Thus, we have justified the definition of curvature for elementary sets.

The properties of the curvature can be formulated in the form of the following theorem.

Theorem 3. *The curvature of a convex surface is additive nonnegative function of elementary sets.*

[2]This formula is a particular case of (7); we should take the triangle A_i for the polygon P in (7); then the right-hand side of (7) yields $\omega(A_i)$ (according to the definition of the curvature of a triangle), while the left-hand side of (7) yields the right-hand side of (16), since the curvatures of open shortest arcs M^r are equal to zero by definition, i.e., $\omega(M^r) = 0$.

© 2006 by Taylor & Francis Group, LLC

Indeed, the additivity of this function is immediate by definition (see page 158), while nonnegativity follows from the same property of the curvature of "basis" sets.

All above arguments are of a purely intrinsic character, and, therefore, they can be applied not only to convex surfaces but also to those manifolds with intrinsic metric whose every point has a neighborhood isometric to a convex surface. Hence *the definition of curvature, together with Theorems 1, 2, and 3, also holds for each manifold with intrinsic metric which is locally isometric to convex surface.*

Moreover, if we do not take into account the nonnegativity of curvature, only the following were involved in other conclusions: (1) the concept of angle and its additivity; (2) the fact that the angle between two branches of one shortest arc is equal to π; (3) the fact that two shortest arcs can intersect at most finite set. These properties hold, e.g., on every regular surface irrespective of whether it is convex or not. Thus, the above foundations of the theory of curvature have a very general nature and can be applied not only to convex surfaces.

Also, we note that using purely intrinsic-geometric methods, we can prove that the curvature of sets on convex surfaces is *completely additive*, rather than simply additive, i.e., if a set M is represented as the union of a sequence of pairwise disjoint sets M_1, M_2,... and if the curvature is defined for M as well as for all M_i, then

$$\omega(M) = \sum \omega(M_i).$$

Using the known methods of measure theory, we can then abstract the concept of curvature to open, closed, and in general, all Borel sets on a convex surface. However, these results are not so simple to obtain, and we will arrive at them as we move forward, when considering the interplay between the intrinsic and "extrinsic" curvatures.

2. The Area of a Spherical Image

Let some set M be given on a convex surface F. Take a sphere S of radius 1. Considering support planes to the surface F at the points of M, we draw the unit outer normals to these planes from the center of S. The geometric locus of the endpoints of these normals is called the *spherical image of* M. If our surface is a doubly-covered plane domain and if a point X lies inside this domain, then the normal is considered as outer if this normal aims at that side on which the point X lies.

Following Gauss, the area of the spherical image of M, if this area exists, is the measure of extrinsic curvature (the integral extrinsic curvature) of the set M on the surface F.

It is useful to recall that not every set on the sphere has area, since this set can fail to be measurable. For example, even the fact that the spherical image of the interior of a geodesic triangle has area, although this fact seems obvious, has in fact not a simple proof. We consider arbitrary convex surfaces, and, therefore, the spherical images of even, say, very simple, sets can have a very intricate structure, and the problem of the existence of their area proves not trival. This does not enable us to restrict ourselves with only the elementary theory of area, and we have to use the theory of Lebesgue measure. University courses usually consider only

© 2006 by Taylor & Francis Group, LLC

plane measure;[3] but measure theory on a sphere does not differ essentially, so we can assume it to be known. By the area of a spherical image, we mean the Lebesgue measure of this image everywhere.

The relation between the area of spherical image and intrinsic metric of a surface was discovered by Gauss. The famous Gauss theorem, which is one of the cornerstones of surface theory, is that the area of the spherical image of a domain on a regular surface depends only on the intrinsic metric of this surface. In particular, the area of the spherical image of a geodesic triangle is equal to the sum of its angles minus π. Our goal is to abstract the Gauss theorem to all convex surfaces; to this end, it is necessary first to study the properties of the spherical image of a convex surface.

With each set M on a convex surface F, we associate its spherical image M^*. The area of the spherical image of M, if this area exists, is denoted by $\psi(M)$; this is a set function on the convex surface F. Since M^* can have no area, i.e., can be nonmeasurable, $\psi(M)$ turns out to be defined only for some sets M. Our first goal is to show that this set function is defined for all closed and all open sets M and even for all Borel sets. In fact, we will deal only with closed or open sets and also with finite unions and intersections of these sets. However, it is convenient to define the function ψ for all Borel sets, since, as is known, we can freely apply to them the operation of taking the union of finite and countable families of sets and the operation of subtraction of one set from another.

In what follows, we will show that $\psi(M)$ is an additive set function, i.e., $\psi(M_1 + M_2) = \psi(M_1) + \psi(M_2)$ whenever the sets M_1 and M_2 have no common points. This is not so obvious as it may seem at first glance, since the fact that two sets M_1 and M_2 have no common points does not always imply that their spherical images have no common points, either. For example, the spherical images of two vertices of the tetrahedron have common points.

Finally, we will prove that the function $\psi(M)$ is not only additive but also completely additive (see page 163). As is known, complete additivity is not implied by additivity.

Strictly speaking, we do not need the complete additivity of $\psi(M)$ but rather some equivalent property that consists in the following. We say that sets M_i form a vanishing sequence if (1) they are sequentially nested, i.e., M_1 includes M_2, M_2 includes M_3, and so on; (2) the intersection of these sets is empty. The property of $\psi(M)$ we speak about reads: if the sets M_i at which ψ is defined comprise a vanishing sequence, then

$$\lim_{i\to\infty} \psi(M_i) = 0. \tag{1}$$

This property of a set function is called *continuity*.[4] We will prove precisely this property of the area of a spherical image, and namely this property turns out to be more convenient in applications. The fact that this property is equivalent to complete additivity ensues from the following general lemma.

[3] See, e.g., Ch. J. de la Vallée-Poussen, A Course of Infinitesimal Analysis, Vol. II, Chapter III, Sec. 2.

[4] This notion, introduced by Frechét, is not widespread. But it is, certainly, absolutely natural for additive functions.

© 2006 by Taylor & Francis Group, LLC

Lemma 1. *An additive set function $\phi(M)$ is completely additive if and only if $\phi(M)$ is continuous in the above sense.* (It is assumed that ϕ is defined over all Borel sets.)

Proof. Let $\phi(M)$ be an arbitrary continuous additive set function. Consider disjoint sets $M_1, M_2, \ldots$. By additivity,

$$\sum_{i=1}^{n} \phi(M_i) = \phi(\sum_{i=1}^{m} M_i) \tag{2}$$

for all n.

Obviously, the sets $\sum_{i=n+i}^{\infty} M_i$ comprise a vanishing sequence, and so

$$\lim_{n\to\infty} \phi(\sum_{i=n+1}^{\infty} M_i) = 0,$$

or, by additivity,

$$\phi(\sum_{i=1}^{\infty} M_i) - \lim_{n\to\infty} \phi(\sum_{i=1}^{n} M_i) = 0.$$

Comparing this equation with Eq. (2), we see that

$$\phi(\sum_{i=1}^{\infty} M_i) = \sum_{i=1}^{\infty} \phi(M_i),$$

i.e., $\phi(M)$ is a completely additive function.

Assume that $\phi(M)$ is completely additive. If sets M_i comprise a vanishing sequence then the sets $M_1 - M_2, M_2 - M_3, \ldots$ are pairwise disjoint and the union of these sets is M_1. Therefore, by complete additivity,

$$\phi(M_1) = \sum_{i=1}^{\infty} \phi(M_i M_{i+1}),$$

or

$$\phi(M_1) = \lim_{n\to\infty} \sum_{i=1}^{n} \phi(M_i - M_{i+1}). \tag{3}$$

Since $M_i \supset M_{i+1}$, we have $\phi(M_i - M_{i+1}) = \phi(M_i) = \phi(M_{i+1})$ by additivity, and hence the sum standing in (3) is equal to $\phi(M_1) - \phi(M_{n+1})$. Therefore, (3) implies $\lim_{n\to\infty} \phi(M_n) = 0$, i.e., the function $\phi(M)$ is continuous. The lemma is proved.

We will also prove some assertions on the convergence of areas of the spherical images of sets on convex surfaces, together with properties of the area of a spherical image mentioned above.

In the sequel, we will speak about closed convex surfaces; this simplifies the arguments but is, certainly, immaterial since each bounded convex surface can be completed to a closed surface. Herewith, its supporting planes and, hence, spherical images are certainly preserved. Below, M^* always stands for the spherical image of some set M, and $\psi(M)$ stands for the area of this spherical image.

© 2006 by Taylor & Francis Group, LLC

Lemma 2. *If a sequence of convex surfaces F_n converges to a convex surface F and if a sequence of points X_n on F_n converges to a point X on F, the limit of each convergent sequence of supporting planes to F_n at X_n is a supporting plane to F at X.*

Proof. Let supporting planes P_n to F_n at X_n converge to some plane P. Each surface F_n lies to one side of its supporting plane P_n. Therefore, the limit surface lies to one side of the limit plane P.[5] At the same time, the plane P obviously passes through the point X of the surface F, which is the limit point of X_n. Hence the plane P is a support plane to F at X.

Lemma 3. *The spherical image of a closed set on a convex surface is closed.*

Proof. Let M be a closed set on a convex surface F, and let M^* be its spherical image. Let a sequence of points N_i in M^* converge to a point N. The definition of spherical image makes it clear that the points N_i are the ends of normals n_i to the surface F at some points X_i belonging to the set M. We can choose a convergent subsequence X_{ij} from the points X_i, and since M is closed, the limit point X of this sequence belongs to M. If we take the same surface F in Lemma 2 for all surfaces F_n, then it is seen from this lemma that the limit n of the normals n_{ij} at the points X_{ij} is a normal to the surface F at the point X. And since the point X belongs to M, the end N of this normal n belongs to the spherical image M^* of M. Hence the limit of every sequence of points in M^* belongs to M^*, i.e., M^* is closed.

A supporting plane to a convex surface F is called *singular* if this plane contains more than one point of F. If two points X and Y of a closed convex surface F lie in a supporting plane P then the whole segment XY belongs to F.

Lemma 4. *The sets of ends of normals to singular supporting planes of a convex surface have measure zero.* (It is assumed that these normals are drawn from the center of the unit sphere, and we speak about the sets of their ends on this sphere.)

Proof. Let F be a convex surface, and let M be the set of ends of singular normals, i.e., normals to singular supporting planes of the surface F. Take two lines L_1 and L_2 passing through the center of unit sphere S. Define the following three classes of singular supporting planes to F: a singular support plane P belongs to the first class if this plane contains a segment of F which is not perpendicular to the line L_1; P belongs to the second class if this plane includes a segment of the surface F which is not perpendicular to the line L_2; finally, P belongs to the third class if this plane contains no segment of F which is not perpendicular to L_1 and L_2. A plane P can belong to the first and second classes simultaneously, but only the fact that each supporting plane belongs to one of these classes is important for us.

In accordance with the fact whether a given plane P belongs to the first, second, or third classes, we say that the end of the normal to this plane (on the unit sphere S) belongs to one of the sets M_1, M_2, or M_3. The set of ends of all singular normals is

[5] This is proved in Sec. 6 of the Appendix.

© 2006 by Taylor & Francis Group, LLC

the union of these sets, i.e., $M = M_1 + M_2 + M_3$. Therefore, if we prove that each of the sets M_1, M_2, and M_3 has measure zero, then the set M is of measure zero.

By the definition of M_3, the endpoint of the normal to the plane P belongs to M_3 if there is a segment in P perpendicular to both lines L_1 and L_2. This means that all these planes P are parallel to a sole line, namely to the line perpendicular to L_1 and L_2. Therefore, the endpoints of normals to these planes lie on one large disk, and, therefore, the set of them, i.e., P_3, has measure zero.

Since the sets M_1 and M_2 play a similar role as is implied by their definitions, it is sufficient to consider one of these sets, say, M_1.

First, we prove that the set M_1 is measurable. Let M_1^{km} be the set of normals to those support planes that have common segments of length $\geq 1/k$ with the surface F and the angle between this segment and L_1 differs from the right angle by at least π/m. It is clear that M_1 is the union of all these M_1^{km} $(k, m = 1, 2, \ldots)$; that is,

$$M_1 = \sum_{k,m=1}^{\infty} M_1^{km}.$$

At the same time, every M_1^{km} is closed. Indeed, let N be a limit point of the set M_1^{km}, and let $N_1, N_2, \ldots$ be a sequence of points in M_1^{km} which converges to N. Then by Lemma 2, the limit of a convergent sequence chosen from the support planes $P_1, P_2, \ldots$ with the endpoints $N_1, N_2, \ldots$ of normals is a supporting plane P with the endpoint N of the normal. A segment of the surface F having the length $\geq 1/k$ and making the angle $\leq (\pi/2) - (\pi/m)$ with the line L_1 lies in each plane P_i. Since the surface F is closed by condition, we can choose a convergent sequence from the indicated segments whose limit is the same segment lying in the plane P. Hence P also contains such segment; this means that the point N also belongs to M_1^{km} and proves that M_1^{km} is closed. Consequently, the set M_1 is the sum of countable many closed sets, and so M_1 is measurable.

Take now some direction L perpendicular to the line L_1 and project the surface F to the plane along this direction. Parallel supporting planes to F envelop the projecting cylinder, and the endpoints of normals to these planes lie on the large disk C perpendicular to L. Since L is perpendicular to the line L_1, this disk C passes through the intersection points N and N' of the line L_1 and the unit sphere S.

The projection of the surface F is some convex domain G. If a singular support plane belongs to the first class and is parallel to the direction of the projection, then the line segment l' on the boundary of the domain G corresponds to this plane. This segment l' is the projection of the segment l of the surface F lying in the plane P; since l is not perpendicular to L_1, i.e., is not parallel to L, l' does not degenerate into a point. The number of segments lying on the boundary of the convex domain G is obviously at most countable. Therefore, the number of singular support planes of the first class that are parallel to the direction L is also at most countable. The ends of normals to these planes form the intersection of the set M_1 with the large disk C, and, therefore, this intersection consists of at most countably many points.

However, L is an arbitrary direction perpendicular to the line L_1, and thus, the disk C is an arbitrary large disk that passes through the points N and N'. Hence

© 2006 by Taylor & Francis Group, LLC

the intersection of such a disk with the set M_1 is countable and always has measure zero. But since M_1 is measurable, this set is of measure zero.[6] The lemma is proved.

The following main theorem on the area of a spherical image is easy from Lemmas 3 and 4:

Theorem 1. *The area of a spherical image is a completely additive set function on a convex surface which is defined for all Borel sets.*

Proof. First of all, we prove that the spherical image of each Borel set on a convex surface has area, or, in other words, the spherical image of every Borel set is measurable. To this end, we sequentially prove the following three assertions: (1) the spherical image of every closed set is measurable; (2) if the spherical image M^* of M on a surface F is measurable then the spherical image $(F-M)^*$ of the complement of M is also measurable; (3) if the spherical images of sets $M_1, M_2, \ldots$ are measurable then the spherical image of their union is also measurable.

(1) By Lemma 3, the spherical image of a closed set is closed and hence measurable.

(2) Let the spherical image M^* of a set M on a surface F be measurable. Consider the spherical image $(F-M)^*$ of the complement of this set. Since the spherical image of the closed convex surface F covers the whole sphere S, we have

$$(F-M)^* = (S-M)^* + M^*(F-M)^*, \tag{4}$$

i.e., the spherical image of $F-M$ consists of the complement of the spherical image of M and the common part of the spherical images of M and $F-M$.

If a point N belongs to M^* and $(F-M)^*$ simultaneously, then this means that there are points of M, as well as $F-M$ in the supporting plane whose normal has the endpoint N, i.e., this supporting plane is singular. Therefore, Lemma 4 implies that the set $M^*(F-M)^*$ is of measure zero. At the same time, M^* is measurable by assumption, and so its complement $(S-M^*)$ is also measurable. Thus, formula (4) yields the representation of $(F-M)^*$ as the union of two measurable sets. Hence this set itself is measurable; as claimed.

Moreover, the measure of $M^*(F-M)^*$ is equal to zero, while the area of the whole unit sphere is equal to 4π; therefore, formula (4) implies

$$\psi(F-M) = 4\pi - \psi(M).$$

Since the complement of a closed set is an open set, this implies in particular that the spherical image of every open set G has area and $\psi(G) = 4\pi - \psi(F-G)$.

[6] Here we use the following theorem: If M_1 is a measurable set and its intersection with each large disk passing through a given point is of linear measure zero, then the set M_1 is of (surface) measure zero. This theorem is an immediate corollary of the famous Fubini theorem on reducing a double integral to a repeated integral:

$$\iint f(x,y)\,dx\,dy = \int \left(\int f(x,y)\,dx \right) dy.$$

It is sufficient to take spherical coordinates as x and y and the characteristic function of M_1 (i.e., the function equal to 1 on M_1 and 0 outside M_1) as $f(x,y)$. The proof of the Fubini theorem is given, e.g., in the book by Ch. J. de la Valée-Poussin, A Course of infinitesimal analysis, Vol. II, Chapter III, Sec. 5 (1933).

© 2006 by Taylor & Francis Group, LLC

(3) Let the spherical images M_n^* of sets M_n $(n = 1, 2, \ldots)$ be measurable. The spherical image of a union of sets is, obviously, the sum of its spherical images, i.e.,

$$(\sum_{n=1}^{\infty} M_n)^* = \sum_{n=1}^{\infty} M_n^*.$$

Since the union of countable many measurable sets is measurable, the set $(\sum_{n=1}^{\infty} M_n)^*$ is measurable, i.e., the spherical image of $\sum_{n=1}^{\infty} M_n$ has area.

The Borel sets are generated from closed sets by taking unions and intersection of sequences of sets. Therefore, in order to prove that the spherical images of not only a closed set but also every Borel set has area, it remains to prove that if the spherical images of $M_1, M_2 \ldots$ have area, then the intersection $\prod_{n=1}^{\infty} M_n$ also has area.

To this end, we note that the following obvious relation holds:

$$\prod_{n=1}^{\infty} M_n = F - \sum_{n=1}^{\infty} (F - M_n),$$

where F is a closed surface on which the sets M_n are considered. We have proved that if the spherical images of M_n have area, their complements $F - M_n$ also have area; then by what we have proved above, the spherical image of the union $\sum_{n=1}^{\infty}(F-M_n)$ and the complement $F-\sum_{n=1}^{\infty}(F-M_n)$ have area, i.e., the spherical image of $\prod_{n=1}^{\infty} M_n$ has area.

Therefore, we have proved that the spherical images of every Borel set on a convex surface has area.

Now let us prove that the area of the spherical image $\psi(M)$ is an additive set function.

Let M_1 and M_2 be two disjoint sets on a surface F. Let their spherical images have areas $\psi(M_1)$ and $\psi(M_2)$. Since M_1 and M_2 are disjoint, the common points of their spherical images are the ends of normals to singular supporting planes. Therefore, Lemma 4 implies that the common part of the spherical images of the sets M_1 and M_2 has measure zero and so the area of the union of the spherical images of M_1 and M_2 is equal to the sum of areas of these spherical images. And since the union of spherical images is obviously the spherical image of this union of sets, we have

$$\psi(M_1 + M_2) = \psi(M_1) + \psi(M_2);$$

i.e., $\psi(M)$ is an additive function.

It remains to prove that this function is completely additive; to this end, it is sufficient to show that this function is continuous, i.e., for every vanishing sequence of sets M_n, we have

$$\lim_{n\to\infty} \psi(M_n) = 0.$$

Let some sets M_n comprise a vanishing sequence, i.e., M_{n+1} lies in M_n and the intersection of all M_n is empty. The spherical images M_n^* of these sets are also nested in each other. If the intersection of all M_n^* is empty then, as is known from measure theory, the limit of measures of these sets is equal to zero; $\lim_{n\to\infty} \psi(M_n) = 0$. If the intersection of all M_n^* is not empty, then each point N of this intersection is the endpoint of the normal corresponding to some points X_n of each of the sets M_n.

© 2006 by Taylor & Francis Group, LLC

These points X_n cannot coincide with each other, since otherwise the intersection of all M_n would be nonempty. Therefore, this point N is the end of the normal corresponding to distinct points of the surface. According to Lemma 4, the set of all these points has measure zero, and so the intersection of all M_n^* is of measure zero. But since M_n^* are nested, the limit of their measures is equal to the measure of their intersection, and so $\lim_{n\to\infty} \psi(M_n) = 0$.

By now we have kept in mind only sets on bounded convex surfaces. However, it is not difficult to generalize Theorem 1 to the case of infinite convex surfaces. First of all, we note that the area of the spherical image of every infinite convex surface is at most 2π. Indeed, an infinite convex surface can be complemented to a complete infinite convex surface F. The body, bounded by such surface, contains at least one half-line L (see Sec. 4 in the Appendix). Therefore, for any supporting plane to F there exists a plane, parallel to this support plane, that is supporting to the half-line L. This implies that the spherical image of the surface F is contained in the spherical image of this half-line, and, therefore, its area is not greater than 2π. (Also, it is easy to show that the spherical image of a complete convex surface is always a spherically convex set.)

Let F be an infinite convex surface. Construct a sequence of indefinitely dilated balls S_n and consider the surfaces F_n, which are cut off from F by the balls S_n.

Let M be a Borel set on F, M_n be a part of this set contained in the ball S_n, and let M^* and M_n^* be the spherical images of these sets. Then obviously, $M_n \subset M_{n+1}$ and $M = \sum_{n=1}^{\infty} M_n$, and also $M_n^* \subset M_{n+1}^*$ and $M^* = \sum_{n=1}^{\infty} M_n^*$. The sets M_n are obviously Borel, and hence, by what we have proved in Theorem 1, the sets M_n^* are also measurable. Therefore, by the well-known theorem of measure theory, M^* is measurable and its measure is equal to the limit of measures of the sets M_n^*. This means, first, that *the spherical image of every infinite Borel set M has area* and, second, this area is equal to the limit of areas of the spherical images of bounded sets M_n, i.e., $\psi(M) = \lim_{n\to\infty} \psi(M_n)$. Therefore, the additivity of the area of the spherical image $\psi(M)$ for unbounded sets is simple on passing to the limit. The continuity of this area is also easily proved. Indeed, let sets M^i comprise a vanishing sequence. Given $\varepsilon > 0$, we take a surface F_n that is cut off from F by the ball S and is such that $\psi(F - F_n) < \varepsilon$. (This is possible, since $\psi(F) = \lim_{n\to\infty} \psi(F_n)$ by what was proved above.) By additivity of ψ, we have $\psi(M^i) = \psi(M^i F_n) + \psi(M^i(F - F_n))$. Here, the sets $M^i F_n$ are bounded, and for a fixed n, these sets comprise a vanishing sequence, so that $\lim_{n\to\infty} \psi(M^i F_n) = 0$ by Theorem 1. In turn, the sets $M^i(F - F_n)$ are included in $F - F_n$, and, therefore, $\psi(M^i(F - F_n)) < \varepsilon$. Hence $\limsup_{i\to\infty}(M^i) \le \varepsilon$; but since ε is arbitrary, we have $\lim_{i\to\infty} \psi(M^i) = 0$; this also proves the continuity of ψ in the case of infinite convex surfaces.

Theorem 2. *If a sequence of closed convex surfaces F_n converges to a surface F and if a sequence of closed sets M_n on F_n converges to a closed set M on F then*

$$\psi(M) \ge \lim_{n\to\infty} \psi(M_n).$$

Note that $\psi(M)$ can *a priori* be different from $\lim_{n\to\infty} \psi(M_n)$. For example, if M is a vertex of a polyhedron, then $\psi(M) \ne 0$, and if M_n are points converging to

© 2006 by Taylor & Francis Group, LLC

M on smooth surfaces, then $\psi(M_n) = 0$ for all n; thus, $\psi(M) > \lim_{n\to\infty} \psi(M_n)$. in exactly the same way, the limit $\lim_{n\to\infty} \psi(M_n)$ can fail to exist. For example, if the sets M and M_n are defined as in the previous example then the sequence M_1, M, M_2, M,... converges to M, but there is no limit of the areas of the spherical images of these sets.

Proof of Theorem 2. Let M^* be the spherical image of a closed set M on a convex surface F. Take an open set G including M^* on the unit sphere S such that the area of this set exceeds the area of M^* by less than some given $\varepsilon > 0$, i.e.,

$$\sigma(G) < \sigma(M^*) + \varepsilon, \tag{5}$$

where σ denotes the area of a set.

Let sets M_n on convex surfaces F_n converging to F converge to M. We prove that for a sufficiently large n, the spherical images M_n^* of the sets M_n are contained in G. To prove this, we assume the contrary. Then we have a sequence of indices $n_1, n_2, \ldots$ such that points $N_{n_1}, N_{n_2}, \ldots$ lying outside G are contained in the sets $M_{n_1}, M_{n_2}, \ldots$. Without loss of generality, we can assume that all these points converge to a certain point N; since all N_{n_i} lie outside G; therefore, N never belongs to M^*.

The points N_{n_i} are the endpoints of normals that correspond to some points X_{n_i} of the surface F_{n_i} which belong to the set M_{n_i}. Since these sets converge to the set M, we can assume that the points X_{n_i} converge to a certain point X of this set; otherwise, we can choose a convergent sequence and restrict consideration to this sequence. Then by Lemma 1, the limit of each convergent sequence of normals at the points X_{n_i} is a normal at the point X. Hence the point N, as the limit of the endpoints N_i of the normals at the point X_{n_i}, is the end of the normal at the point X. Since X belongs to M, we see that N belongs to M^*; but this contradicts what we have proved above. Hence our assumption is not true, and thus, for n greater than some n_0, all M_n lie in the open set G.

But then the area of each of these M_n is not greater than the area of G, i.e.,

$$\sigma(M_n^*) \le \sigma(G) \quad (n \ge n_0).$$

Therefore, (5) implies $\sigma(M_n^*) < \sigma(M^*) + \varepsilon$ for $n \ge n_0$, in accordance with our notation for the area of a spherical image,

$$\psi(M_n) < \psi(M) + \varepsilon \quad (n \ge n_0).$$

This formula makes it clear that

$$\limsup_{n\to\infty} \psi(M_n) < \psi(M) + \varepsilon,$$

and since ε is arbitrary, we have

$$\limsup_{n\to\infty} \psi(M_n) \le \psi(M);$$

© 2006 by Taylor & Francis Group, LLC

as claimed.[7]

Theorem 2 implies the following result which will be needed in the proof of the generalized Gauss theorem.

Theorem 3. *Let a sequence of closed convex surface F_n converge to a surface F. Let G and G_n be domains on F and F_n, and let $\overline{G}$ and $\overline{G_n}$ be the same domains together with boundaries. Assume that the closed domains $\overline{G_n}$ converge to $\overline{G}$ and the complements of the domains G_n, i.e., $F - G_n$, converge to the complement of the domain G, i.e., to $F - G$. If the areas of the spherical images of the boundaries of the domains G_n converge to the area of the spherical image of the boundary of G, then the areas of the spherical images of the domains G_n also converge to the area of the spherical image of the domain G.*

Proof. The complements of the domains G_n and G are closed sets, and so the previous theorem implies

$$\psi(F - G) \geq \limsup_{n\to\infty} \psi(F_n - G_n). \tag{7}$$

The spherical image of a closed convex surface covers the whole unit sphere so that $\psi(F) = \psi(F_n) = 4\pi$. At the same time, by the additivity of ψ, we have

$$\psi(F - G) = \psi(F) - \psi(G) \quad \text{and} \quad \psi(F_n - G_n) = \psi(F_n) - \psi(G_n).$$

Therefore, (7) implies

$$\psi(G) \leq \lim_{n\to\infty} \psi(G_n). \tag{8}$$

[7] The significance of the so-obtained property of the area of a surface image is not only that it implies many important consequences; it also includes a complete, and at the same time simplest, characteristic of the convergence of areas of spherical images of sets on convergent convex surfaces. This can be seen, e.g., from the following. Let a closed convex surface F be not degenerated into a doubly-covered plane domain. Take a point O inside F, circumscribe a sphere S around O, and project F to S. With each set M on S, we then associate the number $\psi(M)$, the area of the spherical image of the set on F whose projection is M. Consider closed convex surfaces F_n that also contain O inside themselves; and, projecting these surface on the same sphere S, we define the numbers $\psi_n(M)$ similarly. Then Theorem 2 makes it clear that if F_n converge to F, then for each closed set M on the sphere S, we have

$$\psi(M) \geq \limsup_{n\to\infty} \psi_n(M). \tag{6}$$

Obviously, the numbers $\psi_n(M)$ do not change under homothety of the surfaces F_n with center O. Therefore, we normalize all surfaces F_n so that all of them, e.g., become intersecting F. Then it turns out that if for each closed set M on S, we have $\psi(M) \geq \limsup_{n\to\infty} \psi_n(M)$, then the surfaces F_n converge to F; i.e., this property is necessary and sufficient for the convergence of surfaces. The analytical meaning of this property becomes clear owing to the following theorem: If completely additive set functions $\psi_n(M)$ and $\psi(M)$ are defined at each set of the sphere S, then $\psi_n(M)$ converge weakly to $\psi(M)$, i.e., for every continuous function $f(X)$,

$$\int f(X)\psi(dM) = \lim \int \psi_n(dM),$$

if and only if the corresponding relations (6) and the relation

$$\lim_{n\to\infty} (S) = \psi(S)$$

hold for each closed set M.

© 2006 by Taylor & Francis Group, LLC

The area of the spherical image of a domain is equal to the difference between the areas of the spherical images of the closed domains and its boundary; therefore, we have

$$\psi(G) = \psi(\overline{G}) - \psi(\text{the boundary of } G),$$
$$\psi(G_n) = \psi(\overline{G_n}) - \psi(\text{the boundary of } G_n). \tag{9}$$

By condition,

$$\psi(\text{the boundary of } G) = \lim_{n\to\infty} \psi(\text{the boundary of } G_n),$$

and since $\overline{G_n}$ converge to $\overline{G}$, we have

$$\psi(G) \geq \limsup_{n\to\infty} \psi(\overline{G_n})$$

by the previous theorem. Therefore, formulas (9) imply

$$\psi(G) \geq \limsup_{n\to\infty} \psi(G_n).$$

Comparing this inequality with (8), we see that

$$\psi(G) = \lim_{n\to\infty} \psi(G_n);$$

the proof of the theorem is complete.

Corollary. *If the boundaries of domains G_n converge to the boundary of a domain G and the area of the spherical image of the boundary of the domain G is zero, then the areas of the spherical images of the boundaries of G_n converge to zero, and hence (according to Theorem 3) the areas of the spherical images G_n converge to the area of the spherical image of G.*

Proof. Let the boundaries of G_n converge to the boundary of G, and let the area of the spherical image of the boundary of G be equal to zero. Since the boundary is a closed set, we have

$$\psi(\text{the boundary of } G) = \limsup_{n\to\infty} \psi(\text{the boundary of } G_n) \tag{10}$$

by Theorem 2; since $\psi(\text{the boundary of } G) = 0$ and $\psi(\text{the boundary of } G_n) \geq 0$ (the area cannot be negative), (1) implies

$$\lim_{n\to\infty} \psi(\text{the boundary of } G_n) = 0.$$

Thus, the second part of the theorem is proved.

Finally, we prove one more simple lemma.

© 2006 by Taylor & Francis Group, LLC

Lemma 5. *Let F be a convex surface, and let A be a point on this surface. Let a sequence of convex closed polyhedra P_n inscribed in the surface F converge to this surface, and, moreover, let the point A belong to all these polyhedra. Then the surfaces of the spherical images of the point A on the polyhedra P_n converge to the area of the spherical image of the point A on the surface F, i.e., if $\psi_n(A)$ is the area of the spherical image of the point A considered as a point of the polyhedra P_n and if $\psi(A)$ is the area of its spherical image as a point of the surface F, then*

$$\psi(A) = \lim_{n\to\infty} \psi_n(A). \tag{11}$$

Proof. Since the polyhedra P_n are inscribed into the surface F, each supporting plane to the surface F at the point A is also a support plane to the polyhedra P_n. Therefore, the spherical image of the point A as a point of the surface F lies in the spherical image of this point A considered as a point of the polyhedron P_n. Hence, for any n,

$$\psi(A) \le \psi_n(A). \tag{12}$$

But since a point is a closed set, we have

$$\psi(A) \ge \limsup_{n\to\infty} \psi_n(A) \tag{13}$$

by Theorem 2. Comparing (12) and (13), we obtain (11).

We can note that not only $\psi_n(A)$ converge to $\psi(A)$, but also under conditions of the lemma, the polyhedral angles at A of the polyhedra P_n also converge to the tangent cone of the surface F at this point.

3. Generalization of the Gauss Theorem

The Gauss theorem asserts that the area of the spherical image of a domain on a regular surface is equal to the intrinsic integral curvature of this domain. In Sec. 1, we have defined the curvature of "basis" sets on convex surfaces, open triangles, open shortest arcs, and points. Then we have introduced the notion of an "elementary" set which is a union of finitely many basis sets that are pairwise disjoint. The curvature of an elementary set was defined as the sum of the curvatures of those basis sets that compose this set.

An open triangle is an open set on a surface, a point is a closed set, and an open shortest arc is the difference between two closed sets, i.e., the shortest arc itself minus the set of its endpoints. Therefore, each of these basis sets and, thus, the spherical image of each elementary set has area. Owing to this fact, the generalization of the Gauss theorem to arbitrary convex surfaces should consist in the proof that *the area of the spherical image of an elementary set on a convex surface is equal to the curvature of this set.* But since the area of the spherical image as well as the curvature are additive functions, it is therefore sufficient to prove the theorem for basis sets. Thus, the generalization of the Gauss theorem reduces to the following three theorems.

Theorem 1. *The area of the spherical image of a point on a convex surface is equal to the curvature of this point, i.e., to 2π minus the complete angle at this point.*

© 2006 by Taylor & Francis Group, LLC

Theorem 2. *The area of the spherical image of an open shortest arc on a convex surface is equal to its curvature, i.e., to zero.*

Theorem 3. *The area of the spherical image of an open triangle on a convex surface is equal to its curvature, i.e., to the sum of its angles minus* π.

It is these three theorems that will be proved consecutively. The proof of the first of the theorems rests on the following lemma on a convex cone.

Lemma 1. *The area of the spherical image of a convex cone is equal to* 2π *minus the complete angle at the vertex of this cone.*

Proof. First of all, we prove that the spherical image of a convex cone is a convex subset of the unit sphere S. It is convenient to take the center of this sphere as the vertex of the cone. If P_1 and P_2 are two supporting planes to the cone, then this cone is contained in the dihedral angle between one half of these planes, so that each plane P that supports this dihedral angle is also a supporting plane to the cone. The endpoints of the outer normals to these planes P yield the smallest of the arcs of a large disk on the sphere that connects the ends of the exterior normals with the planes P_1 and P_2; this means that the spherical image of the cone is convex.

Now let us prove our lemma in the case where our cone turns out to be a polyhedral angle. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be plane angles of this polyhedral angle. If we revolve the support plane to the polyhedral angle along one of the edges, then the end of the outer normal draws an arc of the large disk, which connects the ends of normals with the faces meeting at this edge. We see from this that the spherical image of a convex polyhedral angle is a convex polygon on the sphere. The ends of normals to faces of the polyhedral angle are vertices of this polygon, and its angles are equal to $\pi - \alpha_i$, where α_i are angles on the corresponding faces; this is clear from the fact that under the revolution of the supporting plane around some edge the normal to this plane moves in the plane orthogonal to this edge. As is known, the area of an n-gon on the unit sphere is equal to the sum of its angles minus $(n-2)\pi$,[8] i.e., is equal to

$$\sum_{i=1}^{n}(\pi - \alpha_i) - (n-2)\pi = 2\pi - \sum_{i=1}^{n}\alpha_i.$$

But the sum of all α_i is the complete angle at the vertex of the polyhedral angle, and, therefore, our theorem is proved for the polyhedral angle.

Let us now take an arbitrary cone K and, choosing several cone generators, inscribe a polyhedral angle V with edges at these generators. If we circumscribe the sphere S of radius 1 around the vertex of K, then K intersects this sphere along some convex curve K_1, and the polyhedral angle intersects this sphere along the convex polygon V_1 inscribed into this curve. (If the cone K reduces to a doubly-covered plane angle, then the curve K_1 degenerates into a doubly-covered arc of the

[8] This is proved in a very simple way for a spherical triangle; if we partition a polygon into triangles, a simple computation of angles yields the same result for this polygon. Of course, we deal with a polygon that is homeomorphic to a disk.

© 2006 by Taylor & Francis Group, LLC

large disk; but we may not consider this case, since such a cone is a dihedral angle with two faces, while the theorem is already proved for polyhedral angles.) The perimeter of the polygon K_1 is equal to the sum of plane angles of the polyhedral angle V, and the length of the curve K_1 is equal to the complete angle at the vertex of the cone K. If we increase the number of edges of the angle V decreasing its plane angles, the polygon V_1 converges to the curve K_1 and its perimeter tends to the length of this curve. Therefore, the complete angle θ at the vertex of the cone K is equal to the limit of the complete angles ϕ of polyhedral angles V inscribed in this cone.

On the other hand, the polyhedral angle obviously lies in the cone K, and, therefore, each supporting plane to the cone K supports the angle V. Hence the spherical image of the cone K is contained in the spherical image of the angle V. As the number of edges of the angle V increases indefinitely, its spherical image decreases and converges to the spherical image of the cone K. The spherical images of the angles V and the cone K are convex and, by definition, the area of a convex domain is equal to the greatest lower bound of the areas of the convex polygons including this domain. By what we have proved above, the area of the spherical image of a polyhedral angle is equal to 2π minus the complete angle ϕ at its vertex, and the limit of angles ϕ is the complete angle θ at the vertex of the cone K. Therefore, the area of the spherical image of the cone K is equal to $2\pi - \theta$; as required.

Proof of Theorem 1. If we draw a supporting plane P through a point O of a convex surface F then under a dilation at the point O, the plane P remains a supporting plane, and in the limit, this plane turns out to be a supporting plane to the tangent cone at this point. On the other hand, the tangent cone includes the surface F, and so each of its tangent planes is a supporting plane to the surface F at the point O. Hence the tangent plane and the surface have the same support planes at the point O, and thus, the spherical image of the point O coincides with the spherical image of the tangent cone. Therefore, Lemma 1 just proved implies that the area of the spherical image of the point O is equal to 2π minus the complete angle θ of the tangent cone. But we have proved in Sec. 6 of Chapter IV that the complete angle of the tangent cone at the point O is equal to the complete angle at the point O on the surface itself. Thus, the theorem is proved.

Theorem 1 implies that the set of all points on a convex surface such that the complete angle at each of them is $< 2\pi$ is at most countable. Indeed, $2\pi - \theta$ is equal to the area of the spherical image of a point, while the area of the spherical image of the whole surface cannot exceed 4π. Hence the sum of the numbers $2\pi - \theta$ is finite; this is possible only in the case where they comprise at most a countable set.

The proof of the other two theorems formulated in the beginning of this section is not so simple; we prove these theorems by the method of passing to the limit from polyhedra to arbitrary closed surfaces. Therefore, it is necessary to begin with considering the spherical image of a convex polyhedron.

The spherical image of the interior of a face of a polyhedron is a point, the spherical image of an edge without endpoints is an arc of the large disk, and the spherical image of a vertex is a convex polygon. If the spherical images of two

© 2006 by Taylor & Francis Group, LLC

vertices have common points, then these points belong to the spherical images of edges or faces at which both these vertices are located. We see from what was said above that the area of the spherical image of each set on a convex polyhedron is equal to the sum of the areas of the spherical images of the vertices lying on this set. But we know that a shortest arc on a convex polyhedron cannot pass through a vertex, and, therefore, the area of the spherical image of a shortest arc without endpoints is equal to zero. Incidentally, this proves the second of the theorems formulated above for polyhedra; the third theorem consists in the following lemma.

Lemma 2. *The area of the spherical image of the interior of a geodesic polygon on a convex polyhedron is equal to the curvature of this interior.*

Proof. Let Q be a geodesic polygon on a convex polyhedron P. Obviously, this polygon can be partitioned into triangles each of which lies entirely on the same face of the polyhedron P. Then the curvatures of interior domains of all these triangles are equal to zero. By Theorem 1 of Sec. 1, the curvature of the polygon Q, or, more precisely, the curvature of its interior domains is equal to the sum of curvatures of the interior domains of triangles of this partition and the curvatures of vertices of these triangles lying inside Q. But the curvatures of these vertices that are not the vertices of the polyhedron are equal to zero, and, therefore, the curvature of the interior domain of the polyhedron Q is equal to the sum of the curvatures of the vertices of the polyhedron lying inside Q. The curvature of a vertex is equal to the area of its spherical image, and the area of the spherical image of each domain on a polyhedron is equal to the sum of the areas of the spherical images of the vertices of this polyhedron which lie in this domain. Hence the curvature of the interior of the polyhedron Q is equal to the area of the spherical image of this interior.

Now that we have proved the Gauss theorem for convex polyhedra, it remains to pass to the limit to all convex surfaces. Two quantities, the area of the spherical image and the curvature, participate in the theorem; moreover, the curvature is defined through the angles between shortest arcs. Therefore, in order to pass to the limit, we have to use, first, the results of Sec. 2 concerning the convergence of the areas of the spherical images of sets on convergent convex surfaces, and, second, the theorems on convergence of angles between shortest arcs which were proved in Sec. 4 of Chapter IV.

Proof of Theorem 2. It is sufficient to prove that we can draw two segments with the zero areas of spherical images from each point inside a shortest arc to both sides of this shortest arc. Then a shortest arc without endpoints can be covered by countable many of these segments, and the area of its spherical image also proves to equal zero. Besides, without loss of generality, we can consider shortest arcs lying on a closed convex surface.

Take a shortest arc L on a closed convex surface F and two points A and B inside this shortest arc. A neighborhood of the segment AB can be mapped onto the plane in such a way that the part of the shortest arc in this neighborhood goes to a line.[9] Therefore, for the points near the segment AB it makes sense to say whether they lie on the same side or on different sides from the shortest arc L.

[9] The possibility of such a mapping is implied by the fact that a shortest arc is homeomorphic to a line segment (see Sec. 2 of Chapter II).

© 2006 by Taylor & Francis Group, LLC

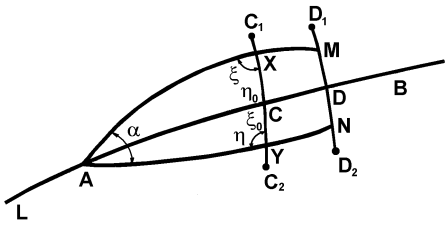


Fig. 48

Take a point Z inside the segment AB and two points D_1 and D_2 that lie on different sides of the shortest arc L. If the points D_1 and D_2 converge to the point Z, then the shortest arc D_1D_2 also converges to Z, and, therefore, we can choose the points D_1 and D_2 so that the shortest arc D_1D_2 intersects the segment AB at some point D (Fig. 48). The shortest arcs AB and D_1D_2 cannot have more than one common point, and if $D-1$ and D_2 are sufficiently close to the point Z, the shortest arc D_1D_2 goes near this point. This makes it clear that the segments D_1D and D_2D of the shortest arc D_1D_2 lie on the opposite sides from the shortest arc L.

In exactly the same way, we can draw the shortest arc C_1C_2 so that it intersects the segment AD at some point C and the segments C_1C and C_2C lie on the opposite sides of the shortest arc L. We shall assume that the segments C_1C and D_1D lie on different sides of L.

Now let us take two variable points M and N on the shortest arcs D_1D and D_2D and move these points to the point D. Then the shortest arcs AM and AN converge to AD (by Corollary 4 of Theorem in Sec. 3 of Chapter II). None of these shortest arcs can intersect the shortest arc L, since two nonoverlapping shortest arcs emanating from a common point cannot intersect each other. Hence, if the points M and N are sufficiently close to the point D, then the shortest arcs AM and AN go to their own side of the shortest arc L. Since these shortest arcs converge to the segment AD, therefore, AM intersects C_1C and AN intersects C_2C whenever the points M and N are sufficiently close to D. Denote by X and Y the intersection points. Then M and N converge to D, while X and Y converge to C.

We obtain the triangle AXY with the shortest arcs AX, AY, and XY as sides.[10] Let α, ξ, and η be the angles of this triangle at the vertices A, X, and Y, respectively. The fact that the shortest arcs AX and XY extend beyond the point X makes it clear that the angle ξ is less than π and is the angle between these shortest arcs (but not the complementary angle to the angle 2π at the point X). In exactly the same way, the angle η is the angle made by AY and XY.

As the point X converges to C, we see that AX and XC_2 converge to AC and CC_2. There are shortest arcs passing through the point C, and so the complete angle at this point is equal to 2π. Hence, by Theorem 5 in Sec. 4 of Chapter IV, the angle ξ made by AX and XY converges to the angle ξ_0 made by AC and CC_2. For the same reason, the angle η converges to the angle η_0 made by AC and CC_1. But $\xi_0 + \eta_0 = \pi$, so we can take the points X and Y so close to the point C that

$$\xi + \eta < \pi + \frac{\varepsilon}{2}, \tag{1}$$

where ε is an *a priori* given positive number. (More exactly, we choose not the

[10] The whole closed surface is, in fact, divided into two triangles, and we take the triangle, which contains the segment AC of the shortest arc L.

© 2006 by Taylor & Francis Group, LLC

points X and Y but the points N and N so close to A that AM and AN intersect CC_1 and CC_2 at the points X and Y sufficiently close to C.)

Since the shortest arcs AX and AY converge to the shortest arc AC, the angle α made by them also converges to zero (by the same Theorem 5 of Sec. 4.) Therefore, if the points X and Y are sufficiently close to the point C, then

$$\alpha < \frac{\varepsilon}{2}. \tag{2}$$

Since the points X and Y are taken so that both inequalities (1) and (2) are valid, the curvature of the triangle AXY is

$$\omega(AXY) = \alpha + \xi + \eta - \pi < \varepsilon. \tag{3}$$

(For brevity, we speak about the curvature of a triangle while always keeping in mind the curvature of the interior of this triangle.)

Now let us construct a sequence of closed convex polyhedra P_n converging to our surface F. We take points A_n, X_n, and Y_n on these polyhedra which converge to the points A, X, and Y. Since the shortest arcs AX, AY, and XY are parts of other shortest arcs AM, AN, and C_1C_2, these shortest arcs are unique shortest arcs that connect the points A, Y, and X. Therefore, the shortest arcs connecting the points A_n, X_n, and Y_n on the polyhedra P_n converge to them. But then the triangles $A_nX_nY_n$ converge to the triangle AXY (the proof is given in the Supplement to Sec. 4 of Chapter IV.) More exactly, three shortest arcs with pairwise common endpoints on any close surface bound two triangles. The triangles $A_nX_nY_n$ are those triangles which converge precisely to the triangle AXY considered above.

Let α_n, ξ_n, and η_n be the angles in the triangles $A_nX_nY_n$ at the vertices A_n, X_n, and Y_n, respectively. Since these shortest arcs pass through the points A, X, and Y (e.g., the shortest arc C_1C_2 passes through X and Y), the complete angles at these points are equal to 2π. Therefore, by Theorem 3 in Sec. 4 of Chapter IV, the angles α_n, ξ_n, and η_n converge to the angles α, ξ, and η, respectively, and thus, the curvatures of the triangles $A_nX_nY_n$ converge to the curvature of the triangle AXY, i.e.,

$$\omega(AXY) = \lim_{n\to\infty} \omega(A_nX_nY_n);$$

since $\omega(AXY) < \varepsilon$ by formula (3), we have

$$\lim_{n\to\infty} \omega(A_nX_nY_n) < \varepsilon. \tag{4}$$

Now let $\psi(A_nX_nY_n)$ and $\psi(AXY)$ be the areas of the spherical images of the interiors of the triangles $A_nX_nY_n$ and AXY, respectively. Since the Gauss theorem is proved for polyhedra, we have $\psi(A_nX_nY_n) = \omega(A_nX_nY_n)$, and inequality (4) implies

$$\lim_{n\to\infty} \psi(A_nX_nY_n) < \varepsilon. \tag{5}$$

The areas of the spherical images of the closed domains complementary to the triangles $A_nX_nY_n$ and AXY are equal to

$$4\pi - \psi(A_nX_nY_n) \quad \text{and} \quad 4\pi - \psi(AXY),$$

© 2006 by Taylor & Francis Group, LLC

respectively. The closed domains complementary to the triangles $A_nX_nY_n$ converge to the closed domain complementary to the triangle AXY. Therefore, applying Theorem 2 of Sec. 2, we obtain

$$4\pi - \psi(AXY) \geq \limsup_{n\to\infty}\{4\pi - \psi(A_nX_nY_n)\}$$

or

$$\psi(AXY) \leq \lim_{n\to\infty} \psi(A_nX_nY_n).$$

In view of inequality (5), this implies

$$\psi(AXY) < \varepsilon.$$

The triangle AXY contains the segment AC of the shortest arc L in its interior if we exclude the endpoints of the segment AC. Therefore, the area of the spherical image of the segment AC with ends excluded is no greater than ε. But since a shortest arc passes through the points A and C, the complete angles at them are equal to 2π, and so the areas of their spherical images are zero. Therefore,

$$\psi(AC) \leq \psi(AXY)$$

also for the whole segment AC, and inequality (6) implies

$$\psi(AC) < \varepsilon.$$

But since ε is arbitrary, we have

$$\psi(AC) = 0.$$

Thus, we obtain the segment AC which is drawn from the point A and whose area of the spherical image is equal to zero. We can cover the whole shortest arc with ends excluded by a countable set of such segments that are drawn from distinct points; therefore, the area of the spherical image of this shortest arc is also equal to zero.

Theorem 3. *The area of the spherical image of the interior of a triangle is equal to the curvature of this interior.*

Proof. Let T be a triangle on a convex surface F. If we divide each side of this triangle into two equal parts, then the so obtained halves are unique shortest arcs between their endpoints. Therefore, considering the midpoints of the sides of the triangle T as vertices, we transform this triangle into the hexagon whose sides are unique shortest arcs between their endpoints.

Let A_1, A_2, ..., A_6 be the vertices of the hexagon T, which are numbered in the order of their location, and let α_1, α_2, ..., α_6 be the angles at these vertices. Here $\alpha_2 = \alpha_4 = \alpha_6 = \pi$, and the curvature of T is equal to

$$\omega(T) = \alpha_1 + \alpha_3 + \alpha_5 - \pi = \sum_{l=1}^{6} \alpha_i - 4\pi. \tag{7}$$

© 2006 by Taylor & Francis Group, LLC

Let us construct a sequence of convex polyhedra P_n inscribed into the surface F in such a way that the points $A_1, \dots, A_6$ lie on these polyhedra (generally speaking, they are vertices of these polyhedra). Then by Lemma 5 of Sec. 2, for each point A_j, we have

$$\psi(A_j) = \lim_{n\to\infty} \psi_n(A_j^n), \tag{8}$$

where the index n shows that the point A_j is considered as a point of an nth polyhedron.

Since the sides of the hexagon T are unique shortest arcs between their endpoints, the shortest arcs $A_1^nA_2^n, \dots, A_6^nA_1^n$ converge to $A_1A_2, \dots, A_6A_1$. Therefore, the hexagons Q_n bounded by the shortest arcs $A_1^nA_2^n, \dots, A_6^nA_1^n$ on the polyhedra P_n converge to the hexagon T.[11]

Since $\psi_n(A_j^n)$ converge to $\psi(A_j)$, the complete angles at the points A_j^n converge to the complete angles at the points A_j. Therefore, Theorem 3 in Sec. 4 of Chapter IV implies that the angles of the hexagons Q_n converge to the corresponding angles of the hexagon T.

Thus, the curvatures of the hexagons Q_n converge to the curvature of the hexagon T, that is,

$$\omega(T) = \lim_{n\to\infty} \omega(Q_n). \tag{9}$$

The spherical image of the boundary of a polyhedron is composed of the spherical images of vertices and sides. The areas of the spherical images of sides without endpoints are equal to zero, while by formula (8), the areas of the spherical images of vertices of the hexagons Q_n converge to the areas of the spherical images of the vertices of the hexagon T. Hence the areas of the spherical images of the boundaries of the hexagons Q_n converge to the area of the spherical image of the hexagon T. By Theorem 3 of Sec. 2, this implies

$$\psi(T) = \lim_{n\to\infty} \psi(Q_n). \tag{10}$$

Since the Gauss theorem is proved for polyhedra, we have

$$\psi(Q_n) = \omega(Q_n),$$

and so (9) and (10) imply

$$\psi(T) = \omega(T),$$

which completes the proof of the theorem.

4. The Curvature of a Borel Set

Until now, curvature was defined only for elementary sets; we now define it for arbitrary Borel sets on a convex surface and prove that the so-defined curvature of each Borel set is equal to the area of the spherical image of this set. It is important to keep in mind that the notion of curve is an intrinsic-geometric concept, and so its definition must be of intrinsic-geometric nature. First, we define curvature for

[11] The broken lines $A_1^n \dots A_6^n$ bound two hexagons each on closed polyhedra. Some of them converge to T. The existence of such hexagons converging to T is proved in the Supplement to Sec. 4 of Chapter IV.

© 2006 by Taylor & Francis Group, LLC

closed bounded sets, starting from the concept of the curvature of an elementary set which is already defined.

We define the curvature of a closed set on a convex surface to be equal to the greatest lower bound of the curvatures of elementary sets containing this set. The curvature of any Borel set on a convex surfaces is defined to be the least upper bound of curvatures of bounded closed subsets that lie in this set.[12]

To guarantee that this definition of curvature for an arbitrary Borel set does not contradict its definition for closed sets, we must prove that both definitions give the same for a Borel set that is closed, i.e., that the curvature of a closed set is equal to the least upper bound of the curvatures of the closed sets lying in this set. And to guarantee that this definition does not contradict the early-introduced definition of curvature for elementary sets, we must prove that both definitions give the same for a Borel set that is elementary. These two facts will be obtained as a corollary to the following theorem which is an abstraction of the Gauss theorem to all Borel sets.

Theorem 1. *The curvature of every Borel set on a convex surface is equal to the area of the spherical image of this set.*

Keeping in mind the above definition of curvature, we split this theorem into the following two.

Theorem 1a. *For each bounded closed set on a convex surface, its curvature, i.e., the least upper bound of the curvatures of the elementary sets including this set is equal to the area of its spherical image.*

Theorem 1b. *For each Borel set on a convex surface, the curvature of this set, i.e., the least upper bound of the curvatures of the bounded closed subsets lying in this set is equal to the area of its spherical image.*

Since the area of the spherical image of a set is uniquely defined for all Borel sets and is equal to the curvature for elementary sets as it was proved in Sec. 3, these theorems in fact imply the soundness of the above definition of curvature.

Proof of Theorem 1a. Let M be a bounded closed set on a convex surface F. If a point X of the surface F does not belong to M then we can circumscribe a disk around this point whose radius r is so small that the disk does not intersect M. Therefore, conversely, circumscribing the disks of radius r around all points of the set M, we obtain some neighborhood U_r of M that does not contain X. This implies that a closed set is the intersection of all these "r-neighborhoods."

Each point Y of the set can be surrounded by a polygon lying in the disk of radius r centered at this point Y. According to the Borel lemma, we can refine

[12] We can start from another definition: The curvature of an open set may be defined to be the least upper bound of the curvatures of the elementary sets that lie in this set. The curvature of a Borel set is then defined as the greatest lower bound of curvatures of the open sets including this set. The relation between both definitions is the same as the well-known relation between the inner and outer measures; and, as in the case of Lebesgue measure, they yield the same.

© 2006 by Taylor & Francis Group, LLC

finitely many polygons from those that cover the whole set M. The union of this polygons comprises an elementary set Q_r that includes M.[13]

Since M is the intersection of its r-neighborhoods U_r, the sets $U_r - M$ comprise a vanishing sequence as r tends to zero and, by the continuity of the function ψ (Lemma 1 and Theorem 1 of Sec. 2), we have $\psi(U_r - M) \to 0$. But since Q_r is contained in U_r and contains M, all the more so, $\psi(Q_r - M) \to 0$, and hence

$$\psi(M) = \lim_{r \to 0} \psi(Q_r). \tag{1}$$

On the other hand, if some set Q contains M, then $\psi(Q) \geq \psi(M)$. Therefore, (1) makes it clear that $\psi(M)$ is the greatest lower bound of areas of spherical images of elementary sets containing M. Comparing this result with the above-given definition of the curvature of a closed set, we confirm that the curvature of any bounded closed set is equal to the area of its spherical image; Theorem 1a is proved.

In order to prove Theorem 1b, we use the following lemma from the general measure theory.

Lemma. *If $\phi(M)$ is a completely additive nonnegative function defined for all Borel sets, then for each Borel set M, this function is the least upper bound of its values for bounded closed subsets included in M.*

Proof. First of all, we note that $\phi(M)$ is equal to the least upper bound of the values of ϕ for closed bounded sets lying in M if $\phi(M)$ is not greater than this bound. Indeed, if $N \subset M$, then $\phi(M) = \phi(N) + \phi(M - N)$, and by the nonnegativity of ϕ, we have $\phi(M - N) \geq 0$, i.e., $\phi(M) \geq \phi(N)$. Hence $\phi(M)$ cannot be less than the least upper bound of the values of ϕ for sets N in M. Therefore, if $\phi(M)$ is not greater than this bound, then $\phi(M)$ is equal to the latter.

Let A be the set of those Borel sets M for which $\phi(M)$ is no greater than the least upper bound of the values of ϕ for closed bounded sets lying in M. According to the above remark, these are Borel sets M such that $\phi(M)$ equals this least upper bound. Thus, to prove the lemma, we have to show that the set A contains all Borel sets.

First, we prove that A contains all closed bounded sets. Indeed, if M is closed, then it is a closed set in M. Therefore, $\phi(M)$ is obviously not greater than the least upper bound of the values of ϕ for all closed sets lying in M, but this means that M lies in A.

Let us prove that if sets $M_1, M_2, \ldots$ lie in A, then their union $\sum_{i=1}^{\infty} M_i$ also lies in A. Since the function ϕ is completely additive, we have

$$\phi(M) = \phi(\sum_{i=1}^{\infty} M_i) = \lim_{n \to \infty} \phi(\sum_{i=1}^{n} M_i). \tag{2}$$

(The sets $\sum_{i=1}^{\infty} M_i - \sum_{i=1i}^{n} M_i$ constitute a vanishing sequence. Therefore, by the continuity property (see Lemma 1 of Sec. 2), we have $\lim_{n\to\infty} \phi(\sum_{i=1}^{\infty} M_i - \sum_{i=1}^{n} M_i) = 0$; using the additivity, we obtain formula (2) from this.) Since the

[13] By the nonoverlapping condition for shortest arcs, finitely many polygons are partitioned into finitely many polygons that have no common interior points and abut on each other along finitely many shortest arcs and points. Hence, they in fact comprise an elementary set.

© 2006 by Taylor & Francis Group, LLC

sets M_i lie in A by condition, for each $\varepsilon > 0$ there exist closed bounded subsets $N_i \subset M_i$ such that

$$\phi(M_i) < \phi(N_i) + \frac{\varepsilon}{2^i}$$

or $\phi(M_i - N_i) < \varepsilon/2^i$. Then, for all n, we have[14]

$$\phi(\sum_{i=1}^{n} M_i - \sum_{i=1}^{n} N_i) \leq \phi(\sum_{i=1}^{n}(M_i - N_i)) \leq \sum_{i=1}^{n} \phi(M_i - N_i) < \varepsilon$$

or

$$\phi(\sum_{i=1}^{n} M_i) < \phi(\sum_{i=1}^{n} N_i) + \varepsilon. \tag{3}$$

By formula (2), this implies that for a sufficiently large n,

$$\phi(M) < \phi(\sum_{i=1}^{n} N_i) + \varepsilon. \tag{4}$$

The union of finitely many closed sets N_i is a closed set lying in M. Therefore, by the arbitrariness of ε, formula (4) implies that $\phi(M)$ is not greater than the least upper bound of the values of ϕ for the closed sets lying in M, i.e., M belongs to the set A.

Now let us prove that if sets $M_1, M_2, \ldots$ lie in A, then their intersection $\prod_{i=1}^{\infty} M_i$ also lies in A. Since the sets M_i belong to A by condition, for any $\varepsilon > 0$ there exist closed sets $N_i \subset M_i$ such that

$$\phi(M - I) < \phi(N_i) + \frac{\varepsilon}{2^i},$$

or $\phi(M_i - N_i) < \varepsilon/2^i$. Then[15]

$$\phi(\prod_{i=1}^{\infty} M_i - \prod_{i=1}^{\infty} N_i) \leq \sum_{i=1}^{\infty} \phi(M_i - N_i) < \varepsilon,$$

or

$$\phi(\prod_{i=1}^{\infty} M_i) < \phi(\prod_{i=1}^{\infty} N_i) + \varepsilon. \tag{5}$$

But $\prod_{i=1}^{\infty} N_i$ is the intersection of the closed sets lying in the sets M_i, and so this set itself is a closed set lying in their intersection $\prod_{i=1}^{\infty} M_i$. Thus, by the arbitrariness of ε, formula (5) implies that $\phi(\prod_{i=1}^{\infty} M_i)$ is not greater than the least upper bound of values of ϕ for the closed sets lying in $\prod_{i=1}^{\infty} M_i$, i.e., $\prod_{i=1}^{\infty} M_i$ belongs to A.

[14]The first inequality follows from the fact that $\sum_{i=1}^{n} M_i - \sum_{i=1}^{n} N_i \subset \sum_{i=1}^{n}(M_i - N_i)$; the second ensues from the fact that $\phi(\sum_{i=1}^{n}(M_i - N_i))$ is equal to the sum of ϕ at the sets $M_1 - N_1, (M_2 - N_2) - (M_2 - N_2)(M_1 - N_1)$, and so on.

[15]Since it is easy to verify that $\prod_{i=1}^{\infty} M_i - \prod_{i=1}^{\infty} N_i \subset \prod_{i=1}^{\infty}(M_i - N_i)$, or, as we have already shown, for any n,

$$\phi(\sum_{i=1}^{n}(M_i - N_i)) \leq \sum_{i=1}^{n} \phi(M_i - N_i).$$

© 2006 by Taylor & Francis Group, LLC

Thus, the set A has the following properties: (1) A contains all closed bounded sets; (2) A contains the intersection and union of the sets $M_1, M_2, \ldots$ provided that the latter belongs to A. But by the very definition of Borel sets, these are exactly the sets that belong to every set enjoying the indicated properties.[16] Hence, these sets belong to the set A, as required.

Now it is easy to prove Theorem 1b.

The curvature of every Borel set on a convex surface is equal to the area of the spherical image of this set.

Indeed, the curvature of a Borel set is by definition the least upper bound of the curvatures of the closed bounded subsets lying in this set. But, by Theorem 1a, the curvature of a bounded closed set is equal to the area of its spherical image. Hence the curvature $\omega(M)$ of each Borel set M is equal to the least upper bound of the areas of the spherical images of the bounded closed sets lying in M. On the other hand, we have proved in Sec. 2, that the area of the spherical image ψ is a nonnegative completely additive function defined over all Borel sets. Hence we can apply the above lemma to this function. If applied to ψ, this lemma tells us that for each Borel set M, the value $\psi(M)$ is equal to the least upper bound of the areas of the spherical images of the bounded closed sets lying in M. And since we have just proved that this bound is equal to the curvature $\omega(M)$ of M, we thus have $\omega(M) = \psi(M)$, i.e., for each Borel set, the curvature is equal to the area of its spherical image.

In Sec. 2, we have shown (Theorem 1) that the area of a spherical image is nonnegative and completely additive; since this area is equal to curvature, the latter has the same properties, i.e., the following theorem holds:

Theorem 2. *Curvature is a nonnegative completely additive function of Borel sets.*

Along with these properties of curvature, we can mention that the curvature cannot exceed 4π of any convex surface, while the curvature of an infinite convex surface is even not greater than 2π. Finally, the curvature of a single point is always less than 2π. The following natural question arises: are there any other characteristic properties? This question has an essentially negative answer as we may see, e.g., from the following theorem.

Let some set function $\phi(M)$ be given on a plane E and enjoy the following properties: (1) *$\phi(M)$ is defined for every Borel set;* (2) *$\phi(M)$ does not assume negative values;* (3) *$\phi(M)$ is completely additive;* (4) *if M is a singleton then $\phi(M) < 2\pi$;* (5) *the value of this function at the whole plane E does not exceed 2π. Then there exists an infinite complete convex surface F such that for each set M on the plane E, the value $\phi(M)$ is the curvature of the set on the surface F whose orthogonal projection is M.*[17]

[16] There is no boundedness requirement for closed sets in the definition of Borel sets. However, this does not change the matter, since any unbounded closed set is a sum of countable many bounded closed sets.

[17] See A. D. Alexandrov, Existence and uniqueness of a convex surface with a given integral curvature, Dokl. Akad. Nauk SSSR, Vol. XXXV, No. 5 (1942), pp. 143–147. For the proof of an analogous theorem for closed surfaces, see the paper A. D. Alexandrov, Application of the theorem on the invariance of a domain to the proofs of existence, Izv. Akad. Nauk SSSR, Ser. Matematika, 1939, No. 3, pp. 244–255.

© 2006 by Taylor & Francis Group, LLC

5. The Set of Directions in Which It Is Impossible to Draw a Shortest Arc

The complete additivity property of curvature, which was proved in the previous section, is essential not only in its own right but also due to the fact that this property implies important corollaries for intrinsic geometry of convex surfaces. Here, we deduce from this property one theorem on the angles between shortest arcs, which in a sense completes the study of the general properties o angle. However, as we have already mentioned, in applications it is often more convenient to use not complete additivity but the continuity property equivalent to complete additivity; i.e., if sets M_n with curvatures $\omega(M_n)$ comprise a vanishing sequence, i.e., if $M_n \subset M_{n+1}$ and the intersection of all M_n is empty, then $\lim_{n\to\infty} \omega(M_n) = 0$. More precisely, complete additivity is equivalent to the conjunction of the following two properties: additivity and continuity (see Lemma 1 of Sec. 2). Of course, additivity and nonnegativity of curvature are of no less importance.

The angle between half-lines on the plane has the following important property: we can draw an angle equal to a given angle on each side of an arbitrary half-line. Angles between geodesics on regular surfaces have the same property, since we can draw a shortest arc from each point of such a surface in every direction.

A shortest arc on a convex surface emanating from a point O has the tangent at this point which is a generator of the tangent cone at the point O, and the angle made by shortest arcs is equal to the angle made by their tangents measured on the tangent cone (see Sec. 6 of Chapter IV). We say that a shortest arc emanates from the point O in direction of the tangent generator to this arc. Since the angle made by shortest arcs cannot be equal to zero, only one shortest arc emanates from a given point in a given direction. If the corresponding existence theorem, together with this uniqueness theorem, holds, i.e., if each generator of the tangent cone is tangent to some shortest arc, then we can draw an angle equal to arbitrarily given angle with one side of an arbitrary shortest arc (of course, we ignore the trivial fact that the angle between shortest arcs emanating from the point O cannot be greater than one half of the complete angle at O).

As we revealed in detail by examples of Sec. 10 of Chapter I, we can draw a shortest arc in each direction not on every convex surface. Hence, *in the general case, not for every shortest arc L on a convex surface emanating from a given point O, there is a shortest arc emanating from O that makes the angle with L equal to a given angle.* This is the main peculiarity of angles between arcs, which makes them different from angles made by half-lines on the plane. Still, in a sense, those directions in which we can draw a shortest arc from a given point are an overwhelming majority. The precise formulation of this important fact is given by the following theorem.

Theorem. *For each point on a convex surface, the set of those in which no shortest arcs emanate from this point has angular measure zero* (i.e., the set of generators of the tangent cone that are tangent to no shortest arc has angular measure zero). *In other words, the set of angles that can not be laid off a given shortest arc has measure zero.*

Proof. Circumscribing a simple closed curve L around a point O on a convex surface, we shall move some point X along this curve. For convenience, we take a broken

© 2006 by Taylor & Francis Group, LLC

line bounding a small convex polygon P that contains the point O in its interior as this curve L (such polygon exists by the theorem of Sec. 4 of Chapter II). If the point X tends to some point X_0, then by the theorem on convergence of angles (Theorem 6 in Sec. 4 of Chapter IV), the angle made by the shortest arc OX and the limit shortest arc OX_0 tends to zero. Hence the tangent to OX converges to the tangent of the shortest arc OX_0. Therefore, the set M of those generators of the tangent cone that are tangent to the shortest arcs OX going to some points of the broken line L is closed.

Let d be a singular direction at the point O, i.e., a generator of the tangent cone which is tangent to no shortest arc emanating from O. Since the set M of generators tangent to the shortest arcs OX is closed and does not contain d, this set contains two generators d_1 and d_2 such that the sector (d_1, d_2) bounded by them is the shortest sector among all sectors which are bounded by generators from M and include the generator d. Let OX_1 and OX_2 be two shortest arcs tangent to these generators and going to the points X_1 and X_2 of the broken line L. Let us show that the points X_1 and X_2 coincide (Fig. 49).

Suppose the points X_1 and X_2 do not coincide. Since L bounds a convex polygon, OX_1 and OX_2 go inside this polygon and divide this polygon into two sectors. If the points X_1 and X_2 do not coincide, then we can draw shortest arcs going to the points on the broken line L inside these sectors. But one of these sectors corresponds to the sector (d_1, d_2) which includes the generator d and the tangents to the shortest arcs going inside this sector make the angles with d which are less than the angles made with d by the generators d_1 and d_2. Hence the points X_1 and X_2 are the same point X. The two shortest arcs OX going to these points have the tangents d_1 and d_2 and bound a digon that encloses the direction d. This means only that the corresponding sector (d_1, d_2) contains the direction d.

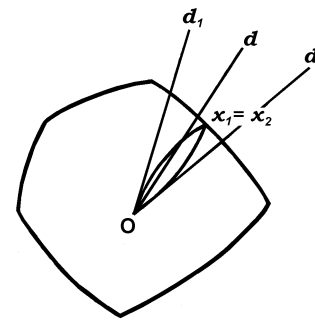


Fig. 49

Consequently, the result at which we arrived is that all singular directions emanating from the point O are contained in digons with one common vertex O and other vertices at some points X of the broken line L. Therefore, the angular measure of the set M of these singular directions does not exceed the sum of angles of these digons.[18]

But the sum of angles of a digon is equal to the curvature of its interior (taking a point on a side of a digon as its vertex, we transform this digon into a triangle with one angle equal to π; this implies that the curvature of the interior of every digon is equal to the sum of its angles). And the sum of the curvatures of the interiors of all digons does not exceed the curvature of the domain $P - O$, which includes them; the domain is obtained from the polygon P by deleting the point O.

[18] We can visualize the matter as follows: When the point X moves along the broken line L, the shortest arcs OX cannot be tangent to singular directions; they should jump over these directions and make the digons containing all singular directions at the points of jumps.

© 2006 by Taylor & Francis Group, LLC

Since the polygon P can be taken arbitrarily small, it is possible to construct a sequence of such polygons P_n contracted to the point O; obviously, then the domains $P_n - O$ comprise a vanishing sequence. Hence their curvatures tend to zero.

Taking the broken lines L_n that bound the polygons P_n and proceeding as above for each of them, we obtain a set of digons in each domain that contain all singular directions emanating from the point O. The sum of the curvatures of these digons, i.e., the sum of their angles tends to zero together with the curvature of the domains $P_n - O$. But these digons contain all singular direction for all n, and so the angular measure of the set of singular directions is zero.[19]

In particular, this implies that even if it could be impossible to draw an angle, equal to a given angle, from a given shortest arc, we can always draw an angle arbitrarily close to a given angle, and the neighborhood of an arbitrary point on a convex surface can always be partitioned into sectors with arbitrarily small angles. The conditions under which it is possible to draw a shortest arc from a given point on a convex surface in any direction will be found in Theorem 1 of Sec. 1 of Chapter XI.

6. Curvature as a Measure of Non-Euclidicity of the Metric of a Surface

When introducing the concept of curvature, we noted that the curvature serves as a measure of non-Euclidicity of the intrinsic geometry of a surface, i.e., the measure of its deviation from plane geometry. To some extent, this can be seen from the very definition of curvature, since the curvature of an open triangle is defined as the difference of the sum of its angles and the sum of angles of a plane triangle; and the curvature of a point is defined as the difference of the complete angle at a point on the plane, i.e., 2π and the complete angle at this point on the surface. The role of curvature as a measure of "non-Euclidicity" of the metric of a surface is revealed in many theorems of intrinsic geometry. In this section, we present a few of them and, in particular, we prove that Euclidean geometry holds for a convex surface of zero curvature.

Theorem 1. *If α, β, and γ are angles of a triangle on a convex surface, ω is the curvature of the interior of this triangle, and α_0, β_0, and γ_0 are angles of the plane triangle with sides of the same length as the length of the sides of the first triangle, then*

$$0 \le \alpha - \alpha_0 \le \omega, \quad 0 \le \beta - \beta_0 \le \omega, \quad 0 \le \gamma - \gamma_0 \le \omega. \tag{1}$$

Indeed, the angles of a triangle on a convex surface are no less than the angles of the plane triangle with sides of the same length. Therefore, all differences $\alpha -$

[19] Note that in proving the theorem on the set of singular directions, we have also clarified the structure of this set. We have proved that the set M_n of the tangent directions to the shortest arcs that go from O to the points X of the broken line L is closed; this is the case for all broken lines L_n. Therefore, the set M of all singular directions is the intersection of the open sets complementary to the closed sets M_n, and hence the set M is a G_δ-set; i.e., an intersection of countably many open sets. It seems probable that each G_δ set of measure zero can be the set of singular directions from some point on a convex surface. It is an interesting problem to prove (or disprove) this claim.

© 2006 by Taylor & Francis Group, LLC

$\alpha_0, \beta - \beta_0$, and $\gamma - \gamma_0$ are nonnegative. On the other hand, by the very definition of the curvature of the interior of a triangle, we have

$$(\alpha - \alpha_0) + (\beta - \beta_0) + (\gamma - \gamma_0) = \omega.$$

Hence each of these differences is not greater than ω. Since the angles between the sides of the triangle are not greater than its angles α, β, and γ but not less than α_0, β_0, and γ_0, we have the same bounds for them.

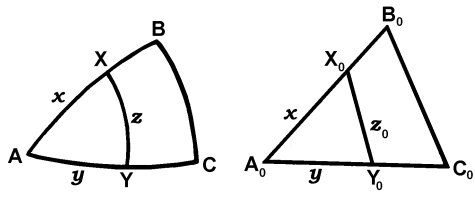


Fig. 50

It is of interest to make the estimate of the difference $\alpha - \alpha_0$ more exact, depending on the distribution of the curvature inside the triangle. The example of a triangle on a cone makes us sure that such a dependence must exist. However, we are not familiar with any, however exact, expression for it.

Theorem 2. *Let ABC be a convex triangle on a convex surface, and let X and Y be two points on its sides AB and CD. Let x and y be the lengths of the segments AX and AY of the sides AB and AC, and let z be the distance from X to Y* (Fig. 50)*. Draw the plane triangle $A_0B_0C_0$ with sides of the same length, i.e., $A_0B_0 = AB$, etc. Take two points X_0 and Y_0 on the sides of the latter triangle such that $A_0X_0 = x$, and $A_0Y_0 = y$; let $X_0Y_0 = z_0$. If ω is the curvature of the triangle ABX*[20] *and d is the length of its maximal side then*

$$|z - z_0| \le 4\sqrt{xy}\sin\frac{\omega}{4} \le \omega d.$$

Proof. Denote by γ_0 the angle at the vertex A in the triangle $A_0B_0C_0$, and let γ_1 denote the angle opposite z in the plane triangle with sides x, y, and z. Then

$$z^2 = x^2 + y^2 - 2xy\cos\gamma_1 = (x-y)^2 + 4xy\sin^2\frac{\gamma_1}{2},$$

$$z_0^2 = x^2 + y^2 - 2xy\cos\gamma_0 = (x-y)^2 + 4xy\sin^2\frac{\gamma_0}{2}.$$

First, these equations imply

$$z^2 - z_0^2 = 4xy(\sin^2\frac{\gamma_1}{2} - \sin^2\frac{\gamma_0}{2}), \tag{2}$$

and, second, they imply

$$z \ge 2\sqrt{xy}\sin\frac{\gamma_1}{2}, \quad z_0 \ge 2\sqrt{xy}\sin\frac{\gamma_0}{2};$$

thus,

$$z + z_9 \ge 2\sqrt{xy}\left(\sin\frac{\gamma_1}{2} + \sin\frac{\gamma_0}{-}2\right). \tag{3}$$

[20]Speaking about "the curvature of a triangle" and "the curvature of a polygon," here and in what follows we mean the curvatures of their interiors.

© 2006 by Taylor & Francis Group, LLC

Taking the modules on both sides of Eq. (2) and dividing them by (3), we obtain

$$|z - z_0| \le 2\sqrt{xy}\left|\sin\frac{\gamma_1}{2} + \sin\frac{\gamma_- 0}{2}\right|.$$

Since

$$\sin\frac{\gamma_1}{2} - \sin\frac{\gamma_0}{2} = 2\cos\frac{\gamma_1+\gamma_0}{4}\sin\frac{\gamma_1-\gamma_0}{4};$$

then all the more

$$|z - z_0| \le 4\sqrt{xy}\sin\frac{\gamma_1-\gamma_0}{4}. \tag{4}$$

According to inequalities (1),

$$\gamma - \omega \le \gamma_0 \le \gamma, \tag{5}$$

where γ is the angle at the vertex A of the triangle ABC. Further, since the triangle ABC is convex, a shortest arc XY goes in this triangle and cuts off the convex triangle AXY. If we denote by ω_1 the curvature of the triangle AXY, then comparing this triangle with the plane triangle with sides of the same length x, y, and z in the same way, we obtain

$$\gamma - \omega_1 \le \gamma_1 \le \gamma. \tag{6}$$

The curvature of the triangle ABC is equal to the sum of curvatures of the triangle AXY and the quadrangle $XYCB$, and since the curvature of this quadrangle is nonnegative, we have

$$\omega_1 \le \omega.$$

Therefore, (5) and (6) imply

$$|\gamma_1 - \gamma_0| \le \omega.$$

Since the angles of a convex triangle do not exceed π, we have $\omega \le 2\pi$, and, therefore,

$$\frac{|\gamma_1-\gamma_0|}{4} \le \frac{\omega}{4} \le \frac{\pi}{2};$$

this implies

$$\sin\frac{|\gamma_1-\gamma_0|}{4} \le \sin\frac{\omega}{4}.$$

Using this inequality, we obtain the inequality

$$|z - z_0| \le 4\sqrt{xy}\sin\frac{\omega}{4} \le \omega d \tag{7}$$

instead of inequality (4), since $\sqrt{xy} \le d$ if d is the length of the maximal side of the triangle ABC. The theorem is proved.

This result can be made more exact. By the convexity condition, $\gamma \ge \gamma_0$, and so $z \ge z_0$. Hence, instead of (7), we can write

$$0 \le z - z_0 \le 4\sqrt{xy}\sin\frac{\omega}{4} \le \omega_1.$$

Theorem 3. *If the curvature of a plane triangle on a convex surface is equal to zero, then this triangle is isometric to a plane triangle.*

© 2006 by Taylor & Francis Group, LLC

Proof. Let ABC be a triangle satisfying the conditions of the theorem, and let $A_0B_0C_0$ be a plane triangle with sides of the same length (Fig. 51). Take an arbitrary point D on the side BC and draw a shortest arc AD in the triangle ABC. At the same time, we take a point D_0 on the side B_0C_0 of the triangle $A_0B_0C_0$ such that the lengths of the segments BD and B_0D_0 are equal to each other. On the basis of the above theorem, we can estimate the difference between the length of the shortest arc AD and the length of the segment A_0D_0. Since the curvature of the triangle ABC is equal to zero by condition, it turns out that the lengths of AD and A_0D_0 are equal.

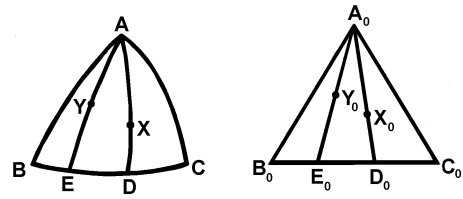


Fig. 51

Now let us construct the following mapping of the triangle $A_0B_0C_0$ into the triangle ABC. Take a point X_0 inside the triangle $A_0C_0C_0$ and draw the segment A_0D_0 through this point from the vertex A_0 to its intersection with the side B_0C_0 at the point D_0. Take the point D on the side BC of the triangle ABC so that the segment BD of this side is equal to the segment B_0D_0. This is always possible, since BC and B_0C_0 are equal. Draw the shortest arcs AD in the triangle ABC and take a point X on this shortest arc such that the segment AX of this shortest arc is equal to the segment A_0X_0. This is possible, since $AD = A_0D_0$ by what we have proved above. This point X is assigned to the point X_0.[21]

Let us prove that this mapping is an isometry. Let X_0 and Y_0 be two arbitrary points in the triangle $A_0B_0C_0$, and let A_0D_0 and A_0E_0 be the segments passing through these points which are extended to the side B_0C_0. Let X and Y be the points corresponding to them in the triangle ABC, and let AD and AE be the corresponding shortest arcs. By what we have proved above, $AD = A_0D_0$ and $AE = A_0E_0$. Moreover, $DE = D_0E_0$, since $BD = B_0D_0$ and $BE = B_0E_0$ by the definition of our mapping. Hence we obtain two triangles ADE and $A_0D_0E_0$ with equal sides. We have the points X and Y, X_0, and Y_0 on the sides AD and AE, A_0D_0, and A_0E_0 of these triangles, respectively, that cut off the equal segments $A_0X_0 = AX$ and $A_0Y_0 = AY$. The triangle ADE is convex, since this triangle is isolated from the convex triangle ABC by two shortest arcs (see Theorem 4 in Sec. 5 of Chapter II). The curvature of the triangle ADE is equal to zero, since the curvature of the whole triangle ABC is zero. Therefore, applying Theorem 1, we make sure that the segments XY and X_0Y_0 are equal. This proves that the mapping defined above is isometric.

In order to verify that the triangle ABC is isometric to the triangle $A_0B_0C_0$, it is necessary to prove in addition that our mapping acts onto the whole triangle, i.e., each point X of the triangle ABC is the image of some point of the triangle $A_0B_0C_0$. But this is obvious, since each isometry is a homeomorphism and, moreover, our mapping transforms the boundary of the triangle $A_0B_0C_0$ into the boundary of the

[21] There is only one shortest arc AD. If there were two shortest arcs AD in the triangle ABC, these shortest arcs would have bounded a digon contained in ABC. But a digon has a positive curvature equal to the sum of its angles. This contradicts the fact that the curvature of the whole triangle equals zero.

© 2006 by Taylor & Francis Group, LLC

triangle ABC. But as is known (and is easily proved), under a topological (and even continuous) mapping of one triangle onto the other which transforms the boundary of the first into the boundary of the second, all points of the first triangle prove to be the images of points of the second one.

The theorem is proved.

Theorem 4. *A domain G on a convex surface is locally isometric to the plane, i.e., each point of this surface has a neighborhood isometric to a part of the plane, if and only if the curvature of each triangle in the domain G is equal to zero, or, equivalently, the curvature of the whole domain is equal to zero.*

Proof. The necessity of the condition is obvious. To prove its sufficiency, we take a neighborhood around some point O of the domain G which is a convex geodesic polygon (the existence of such neighborhood was proved in Sec. 4 of Chapter II). This polygon can be partitioned into convex triangles by drawing diagonals in it. If the curvatures of these triangles equal zero, then, by the above theorem, all these triangles are isometric to plane triangles. If the point O lies inside one of these triangles, then this triangle is a neighborhood isometric to a part of the plane. If O lies on a common side of two triangles, then these triangles together make a neighborhood isometric to a plane quadrangle. The theorem is proved.

Note that the vanishing of the curvature of a domain G is not sufficient for this domain G to be isometric to a plane domain. This is shown by examining the lateral surface of a cylinder. Therefore, the vanishing of the curvature of a domain G implies only that Euclidean geometry "in the small" is realized.

We now present several examples of theorems without their proofs in which the curvature plays the role of the measure of "non-Euclidicity."

The main part of the theory of triangles in Euclidean geometry consists of the theorems that state some relations between the angles and the lengths of their sides. All trigonometrical formulas are of this kind. For example, the following theorems exemplify them: If the sides of two triangles are equal then the corresponding angles of these triangles are equal; the bisectrix divides the opposite side into two parts proportional to other two sides, etc.

The difference of the angles of a triangle on a convex surface from the angles of the plane triangle with the same sides is estimated according to formulas (1) via curvature. Therefore, with each theorem concerning relations between the angles and sides of plane triangles, we associate a general theorem on triangles on convex surfaces; this theorem should include the curvature of the triangles, since the latter is a magnitude that estimates the possible deviation of the relations in these triangles from the relations in plane triangles. For example, if the sides of two triangles on a convex surface are equal, then the module of the difference of every two of the corresponding angles does not exceed the sums of the curvatures of these triangles. This theorem is obviously implied by inequalities (1). The reader can take a random theorem on triangles in a course of geometry or trigonometry and try to formulate and prove the corresponding theorem for triangles on convex surfaces.

All these concern the triangles whose sides are shortest arcs. It turns out that the same results hold for a wider class of "ordinary" triangles. A triangle is said

© 2006 by Taylor & Francis Group, LLC

to be *ordinary* if its sides are geodesics, and for every pair of its vertices, there is no line between these vertices shorter than the side connecting them. In short, a normal triangle is a geodesic triangle whose sides are shortest arcs on this triangle (but not necessarily shortest arcs on the whole surface). For example, there are ordinary triangles on a cone with complete angle $< 2\pi$ whose sides are not shortest arcs. It is possible to prove the following theorem: Let X and Y be two points on the sides AB and AC of an ordinary triangle ABC, let x and y be the lengths of the segments AX and AY of these sides, and let $z = z(x, y)$ be the length of the line XY, shortest among all those connecting the points X and Y and lying in this triangle. Let $\gamma(x, y)$ be the angle opposite the side z in the plane triangle with sides x, y, and z. Then $\gamma(x, y)$ is a nonincreasing function of x and y. We leave the proof of this theorem to the reader. This proof can be performed by the method used for the proof of the convexity condition for a polyhedral metric of positive curvature. Now, when we have made good progress in developing the intrinsic geometry of convex surfaces, this is not difficult to implement. One difficulty consists in the fact that the line XY in this theorem can pass through vertices of the triangle if the angles at these vertices are $> \pi$.

The above theorem implies immediately that the angles of an ordinary triangle on a convex surface are not less than the angles of the plane triangle with sides of the same length. Thus, all results leaning on this property of the triangles with shortest sides extend to ordinary triangles. It is difficult to compare arbitrary geodesic triangles with plane triangles, since, e.g., the sum of two sides of a geodesic triangle can be greater than the third side. We can only say in general that the sum of angles of these triangles is not less than π, either.

To another section of elementary geometry, the theory of the circle, there also corresponds a number of theorems on the circle on a convex surface in which the degree of "non-Euclidicity" is also estimated by curvature. Examples of such theorems will be given in Sec. 6 of Chapter IX, devoted especially to the circle.

Also, let us consider in general terms the properties of the metric of a convex surface in small domains from the same point of view, i.e., as the measure of non-Euclidicity of this metric.

We have proved in Sec. 5 of Chapter IV that the metric of a convex surface F in a small neighborhood of a point O is approximately represented by the metric of the tangent cone K at this point. Namely, for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|\rho_F(XY) - \rho_K(X'Y')| < \varepsilon \max[\rho_F(OX), \rho_F(OY)] \tag{8}$$

whenever $\rho_F(OX), \rho_F(OY) < \delta$, where ρ_F and ρ_K are the distances on the surface F and the cone K and X' and Y' are the projections of two points X and Y to the cone K in an arbitrarily given direction passing inside this cone. The estimate of the difference between the distances in the triangles on a convex surface and in the plane, which is given by Theorem 2, together with the continuity property of curvature, leads to a deeper understanding of this relationship between the metric of a surface and the metric of the tangent cone.

By the continuity property of curvature, for each $\varepsilon > 0$, there exists a neighborhood U of the point O such that the curvature ω of the domain $U - O$ is less than ε. It is convenient to take a convex polygon as such a neighborhood. The

© 2006 by Taylor & Francis Group, LLC

angles of each triangle whose interior lies in the domain U differ by less than ε from the angles of the plane triangle with the same sides. According to the estimate of Theorem 2, the distances in these triangles differ from each other by less than εd, where d is the length of the greatest side.[22] Hence the metric turns out to be Euclidean in the domain U to within the magnitudes of order εd.

For example, take a triangle OAB with vertex the point O; let X and Y be some points on the sides OA and OB of this triangle. Take the plane triangle $O_0A_0B_0$ with sides of the same length, and let X_0 and Y_0 be the points on its sides O_0A_0 and O_0B_0 such that $\rho_0(O_0X_0) = \rho_F(OX)$ and $\rho_0(O_0Y_0) = \rho_F(OY)$, where ρ_0 and ρ_F be the metrics of the plane and the surface in question. Then by Theorem 2,

$$|\rho_F(XY) - \rho_0(X_0Y_0)| \leq \omega\sqrt{\rho_F(OX)\rho_F(OY)},$$

and thus,

$$|\rho_F(XY) - \rho_0(X_0Y_0)| \leq \omega \max[\rho_F(OX), \rho_F(OY)]. \tag{9}$$

Euclidean geometry holds on a cone without apex, and hence the relation between the metric of a surface and the metric of the tangent cone which is given by formula (8) is, essentially, the same relation that is given by formula (9). Moreover, the unknown ε in formula (8) is replaced by the value ω in (9), which is the curvature of the domain $U - O$.

From the point of view of the intrinsic metric, the tangent cone at the point O is characterized by the fact that its complete angle is equal to the complete angle θ at the point O. Formula (8) gives grounds to suppose that the neighborhood U of the point O can be mapped onto a neighborhood of the apex of the cone K with the complete angle θ in such a way that for all points X and Y in U, we have

$$|\rho_F(XY) - \rho_K(X'Y')| \leq \omega(U - O) \max[\rho_F(OX), \rho_F(OY)], \tag{10}$$

where X' and Y' are the images of the points X and Y on the cone K; moreover, such a mapping should map the point O into the apex of the cone. The angles of triangles in the domain $U - O$ and in the plane differ from each other by no more than $\omega(U - O)$; this fact leads to the conclusion that this mapping can be made such that the angles in the corresponding triangles XYZ and $X'Y'Z'$ on the surface and the cone differ from each other by not more than $\omega(U - O)$. It is interesting to give a complete proof of the fact that there exists a mapping with both properties.[23] If

[22] Theorem 2 deals with a convex triangle, but this is necessary only for the existence of the shortest arc XY and for the curvature of the triangle OXY to be not greater than ω. But under our conditions, the existence of the shortest arc XY is ensured by the smallness of the neighborhood U, while the curvature of the triangle OXY does not exceed ω since this triangle lies in U.

[23] This mapping is defined as follows: Draw some shortest arc L from the point O and assign the direction of O. Draw the circle C_X centered at O through the point X in a small neighborhood U of the point O. Let X_1 be the intersection point of the circle C_X with the shortest arc L, and let XX_1 be the arc of the circle C_X between X_1 and X_2 that goes from X_1 to X_2 in the given direction. Let λ be the ratio of the length of the arc XX_1 and the length of the whole circle C_X, and let r be the distance $\rho_F(OX)$. The numbers λ and r are coordinates in a neighborhood of the point O. We can introduce the same coordinates on the cone K (they are the polar coordinates to within the coefficient $1/\theta$ of the angle). If to each point X in the neighborhood U we now put in correspondence the point X' on the cone with the same coordinates λ, r, then we obtain the mapping for which formula (11) holds. The value ε_1 in this formula depends on the curvature ω of the domain $U - O$; but the question on the bound for ε_1 still remains open.

© 2006 by Taylor & Francis Group, LLC

the tangent cone at the point O reduces to the plane K then inequality (8) can be replaced by a stronger inequality

$$|\rho_F(XY) - \rho_K(X'Y')| \leq \varepsilon_1 \rho(XY), \tag{11}$$

where ε_1 is also an infinitely small value together with $\max[\rho_F(OX), \rho_F(OY)]$. Here X' and Y' stand for the projections of the points X and Y to the tangent plane K. This inequality implies that not only the distances but also the angles are slightly varied under the projection of a small neighborhood U of the point O on the surface to the tangent plane. If there is no tangent plane at the point O, then the projection to the tangent cone does not yield a similar result in general. However, V. A. Zalgaller indicated a mapping of a small neighborhood of the point O onto the tangent cone K for which formula (11) always holds for the whatever tangent cone K at the point O. In any case, it becomes clear that the estimate of deviation of the metric of a surface from the metric of the tangent cone is also an estimate of "non-Euclidicity" of the metric of a small domain $U - O$ and this deviation depends on the curvature of this domain.

© 2006 by Taylor & Francis Group, LLC

Chapter VI

EXISTENCE OF A CONVEX POLYHEDRON WITH A GIVEN METRIC

1. On Determining a Metric from a Development

This chapter is devoted to the proof of the following theorem:

Each polyhedral metric of positive curvature on the sphere is realized by a convex closed polyhedron.

The doubly-covered convex polyhedra are also reckoned among the set of convex polyhedra; we call them *degenerate polyhedra*. Since we deal only with closed convex polyhedra, the word "polyhedron" will always mean a closed convex polyhedron, degenerate or nondegenerate, unless otherwise specified. By a similar reason, the word "metric" will always mean a polyhedral metric of positive curvature on the sphere.

We have shown in Sec. 4 of Chapter I that each polyhedral metric on the sphere can be given by some development consisting of triangles, and, conversely, such a development homeomorphic to a sphere determines a polyhedral metric on a sphere. We will deal only with the developments composed of triangles. The *vertices of a metric* are the points such that the complete angle at each of them is not equal to 2π. A metric of positive curvature is characterized by the fact that the complete angle at each vertex is less than 2π. Of course, an *a priori* given development can have extra vertices with the complete angle equal to 2π, but we show that each metric can be assigned by some development without superfluous vertices.

Determining a metric from some development allows us to state our theorem in a perfectly elementary form as it was already mentioned in Sec. 4 of Chapter I. Instead of a function of two points of a sphere (and each metric is such a function), a development is an object defined by finite data, since each triangle of the development is determined by the length of its sides. However, determining a metric from some development is also fraught with some difficulties. First, the same metric can be given by infinitely many distinct developments. Second, if we can glue a polyhedron from a given development, it can still have little in common with the natural development of the polyhedron formed by its faces. The problem – of giving a method for finding the natural development of the polyhedron glued from a given development – seems hopelessly difficult. It hardly has any satisfactory general solution. The simple example of a development of a tetrahedron, given in Sec. 4 of Chapter I, confirms this to some extent; so what can we expect in the case of polyhedra with many vertices? Finally, when a polyhedron deforms continuously, its metric varies continuously, whereas the natural development may have a jump as seen from the following simple example.

© 2006 by Taylor & Francis Group, LLC

Take the polyhedron shown in Fig. 52,a, and continuously deform it by dropping the vertices A and B. Its structure remains the same until the vertices A, B, C, and D are in the same plane. After that, further dropping of the vertices A and B makes the polyhedron flex along the edge CD (otherwise, it would become nonconvex!) and take the shape shown in Fig. 52,b. Its natural development changes its structure when the vertices A, B, C, and D arrive at the same plane. There are even more possibilities for these jumps in the structure of the natural development for polyhedra with more vertices. By exact analogy, not every continuous deformation of the metric is realizable by a continuous deformation of the same development.

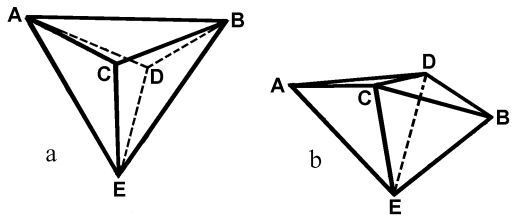


Fig. 52

These difficulties force us, first, to specify the relation between various developments that assign the same metric and choose those that are as simple as possible. Secondly, they also force us to elucidate how the deformation of one of the developments affects the others that realize the same metric. We proceed now to solving these problems to an extent necessary for the proof of the realization theorem given below.

Let a polyhedral metric of positive curvature with e vertices A_1, A_2, ..., A_e be given on sphere S. The number of these vertices is never less than three; this can be seen at least from the fact that the curvature of the whole sphere is equal to 4π for any given metric, while the curvature of a vertex is always less than 2π. Draw shortest arcs $A_1A_2, \ldots, A_1A_e$ from the vertex A_1 to all other vertices. Since a shortest arc cannot pass through any vertex and two shortest arcs emanating from a common point cannot intersect each other (as proved in Sec. 2 of Chapter III), the shortest arcs $A_1A_2, \ldots, A_1A_e$ have no common points but the vertex A_1. Therefore, cutting the sphere S along these shortest arcs, we obtain some polygon Q[1] whose $e-1$ vertices coincide with A_1, and the vertices $A_2, \ldots, A_e$ are other $e-1$ vertices of this polygon. The sides of this polygon are pairwise equal, since each pair of sides emanating from a common vertex A_i $(i \neq 1)$ is glued into one shortest arc A_1A_i. Since the polygon Q is homeomorphic to a disk and does not contain the vertices of the metric in its interior, we can develop this polygon onto the plane. Hence each metric of positive curvature on a sphere can be given by some development that consists of a single polygon with neighboring sides glued pairwise. But this is not essential for us, and we leave this remark without proof.[2]

We partition the polygon Q into triangles by diagonals. This partition can be realized as follows: Let $B_1, B_2, \ldots, B_b$ be vertices of the polygon Q, which are enumerated in the order of their location on the boundary of this polygon. Here, $n = 2e-2$ and the vertices B_i are the vertices of the metric: $A_1, A_2, A_1, A_3, \ldots, A_1, A_e$.

[1] To cut S along some line AB means to take a metric space S' having two lines $\overline{A'B'}$ and $\overline{\overline{A'B'}}$ under the identification of which one obtains S.

[2] We show now that the polygon Q can be partitioned by diagonals into triangles. Unfolding these triangles onto the plane, we develop the whole polygon Q. Of course, Q can cover itself under this development.

© 2006 by Taylor & Francis Group, LLC

Since the polygon Q does not contain the vertices of the metric in its interior, the sum of its angles β_i is expressed by the usual formula

$$\sum_{i=1}^{n} \beta_i = (n-2)\pi.$$

Therefore, the polygon Q has at least three vertices and the angles at these vertices are less than π. (Indeed, if the angles at $n-2$ vertices are $\geq \pi$, then the sum of these angles is $\geq (n-2)\pi$.) Hence the polyhedron Q has two nonadjacent vertices B_p and B_e with angles less than π.

Take two arbitrary vertices B_p and B_e that are separated by the vertices B_p and B_q. Draw the line B_kB_e that is a shortest arc in the polygon Q (Fig. 53). As was shown in Sec. 2 of Chapter II, this line exists and is a geodesic broken line with vertices at the vertices of the polygon Q. But if this line passes through vertices with angles less than π, then we can shorten this line by cutting the angle between its sides that meet each other at such a vertex. Therefore, our line cannot pass through the vertex B_p, as well as through the vertex B_q. Since these vertices are separated by the endpoints of the line B_kB_e, it passes even partially inside the polygon Q, i.e., when the line goes away from the boundary of this polygon at some vertex B_r, it arrives at the boundary again at some vertex B_s and is a geodesic on the segment B_rB_s. This geodesic B_rB_s is a diagonal of the polygon Q and divides this polygon into two polygons Q_1 and Q_2 each of which has less vertices than Q. Applying the same argument to the polygons Q_1 and Q_2, we partition them into polygons with fewer vertices and so on up to the moment when the whole polygon Q becomes triangulated.

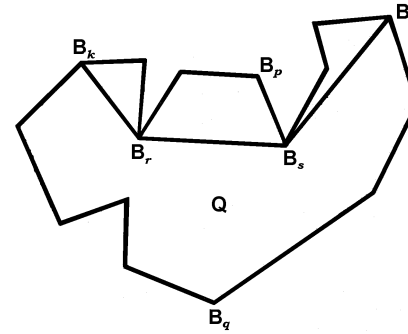


Fig. 53

Since these triangles do not contain the vertices of the metric in their interiors, and, therefore, their curvatures are equal to zero, each of them can be developed onto the plane. Indeed, let ABC be one of such triangles. Since the curvature of this triangle is zero, its angles are less than π. Therefore, the shortest arc AX of the vertex A and the point X on the side BC lie inside this triangle. Besides, for a given point X, there is only one such line, since otherwise two geodesics AX would form a digon, and the curvature of a digon is positive (it is equal to the sum of its angles as it follows from the general formula $\omega = \sum_{i=1}^{n} \alpha_i - (n-2)\pi$ for the curvature of an n-gon). But we have already proved in Sec. 2 of Chapter III that a neighborhood of a geodesic in a polyhedral metric of positive curvature can be developed onto the plane.[3] When the point X runs over the whole side BC, we develop sequentially neighborhoods of the geodesic AX onto the plane, and thus, we develop the whole triangle ABC. (Of course, this assertion should be made more

[3]We have proved this in Sec. 2 of Chapter III for shortest arcs, but since each geodesic can be covered by finitely many shortest arcs, this is also true for geodesics.

© 2006 by Taylor & Francis Group, LLC

precise; this can be done easily by using the Borel lemma on the choice of a finite number of neighborhoods.)

Thus, we have partitioned the polygon Q into triangles isometric to plane triangles. Therefore, we have obtained some development of the given metric. Considering an abstract metric in the sequel, we will define it from developments that are obtained by the above construction in contrast to the metric of a given polyhedron. Since such a development has no other vertices but the vertices of the given metric, we can formulate the following assertion.

Lemma 1. *Each polyhedral metric of positive curvature on a sphere can be given by some development composed of plane triangles and having no vertices but the vertices of the metric itself.*

Based on this lemma, we will deal in the sequel *only* with such developments. The sides of the triangles composing a development are called the edges of this development and the glued sides are considered as one edge. The glued vertices of triangles yield one vertex of the development or, which is the same for the development considered, one vertex of the given metric.

We have already noted that a polyhedral metric on a sphere cannot have less than three vertices. Consider the simplest case of a metric with three vertices A_1, A_2, and A_3. Connecting these vertices by shortest arcs, we divide the sphere into two triangles that do not contain the vertices of the metric and, therefore, can be developed onto the plane. Since these triangles have sides glued pairwise, they are equal. Therefore, developing these triangles onto the plane and overlapping them onto each other, we obtain a doubly-covered triangle. This prove the following assertion.

Lemma 2. *A polyhedral metric on the sphere that has only three vertices is realized by a double-covered triangle.*

Therefore, we can restrict further exposition to considering the metrics that have more than three vertices.

An important role in our arguments will be played by the concept of developments of the same structure. We say that two developments R_1 and R_2 has the *same structure* if there is a one-to-one correspondence between their elements, i.e., between triangles, edges, and vertices, which preserves the incidence (membership) relation between these elements. In more detail, this correspondence must enjoy the following properties.

1. With a triangle, an edge, or a vertex of the development R_1, we associate a triangle, an edge, or a vertex of R_2, and vice versa.

2. If an edge a_1 belongs to a triangle T_1 in the development R_1, then the corresponding edge a_1 belongs to the corresponding angle T_2 in the development R_2, and vice versa.

3. If a vertex A_1 belongs to an edge a_1 (or a triangle T_1) in the development R_1, then the corresponding vertex A_2 belongs to the corresponding edge a_2 (or the triangle T_2) in the development R_2.

© 2006 by Taylor & Francis Group, LLC

We can also change the lengths of edges of a development keeping the same structure. Strictly speaking, this means the following: we say that a development R is obtained from some development R_0 by changing the edge lengths if both developments have the same structure but the lengths of the corresponding edges are different. In general, there are several one-to-one correspondences between the elements of two developments witnessing that these developments have the same structure. However, in what follows we keep in mind only one correspondence. Then the corresponding elements of all developments of the same structure can be denoted by the same letters. Under this condition, we say that a development deforms continuously if the length of each of its edges varies continuously (i.e., the correspondence between edges, triangles, and vertices should be preserved in the process of changing the edge lengths). Development of a given structure is completely determined by the lengths of its edges, since each triangle is completely determined by the lengths of its sides.

Lemma 3. *If k is the number of edges and e is the number of vertices of a development, then $k = 3e-6$.* Hence the number of variables determining a development of a given structure is the same for all developments with a given number of vertices.

Indeed, if f is the number of triangles of the development, then by the Euler theorem,

$$f - k + e = 2.$$

Each triangle has three sides, and the sides are pairwise glued. Therefore,

$$3f = 2k.$$

Multiplying the first equation by 3 and substituting $2k$ for $3f$ in it, we obtain $k = 3e - 6$.

If two developments R_1 and R_2 assign the same metric, the triangles of one of them obviously consist of parts of triangles of the other, and vice versa. Vertices of both developments coincide, since they are vertices of the metric. But the edges of these developments may fail to coincide and can connect different vertices in the different developments. Some edge AB of the development R_1 emanating from the vertex A goes along the triangles of the development R_2 and intersects its edges in some order up to the moment when this edge arrives at the vertex B. In each triangle, a segment of the edge AB is linear, and, crossing an edge, this segment goes to a neighboring triangle in such a way that under the development onto the plane we also obtain a straight line segment. Therefore, under the sequential development on the plane of all triangles T_i traversed by the edge AB this edge transforms into a straight line segment. This is a diagonal of the polygon P covered by all consecutively unfolded triangles T_i (Fig. 54). Of course, the edge AB can intersect the same triangle many times, but each time this triangle appears in the polygon P.[4]

[4]The edge AB has finitely many intersection points with triangles. Indeed, if there were infinitely many intersection points of the edge AB with the triangle T_i, then the lengths of segments containing in T_i would tend to zero, since otherwise the edge AB would be of infinite length. But the segments between points on sides of a triangle can tend to zero only if they accumulate at a vertex of this triangle. Hence, the vertex of the triangle T_i should lie on the edge AB. But then this vertex is one of the endpoints of this edge, and the edge AB goes away from this vertex as one straight line segment; thus, there cannot be infinitely many of these segments.

© 2006 by Taylor & Francis Group, LLC

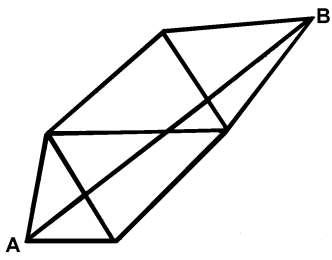


Fig. 54

We say that *the location of a development R_1 in a development R_2 is given* if we know which edges connect the vertices of the development R_1 and which edges of the development R_2 and in which order intersect each edge of the development R_1; in particular, this edge can coincide with some edge of the development R_2.

Lemma 4. *Let two developments R_0 and S_0 determine the same metric and have the same vertices. Changing the edge lengths of the development R_0, we obtain a new development R of the same structure. If the changes of the length of each edges are sufficiently small, then the metric given by the development R can be uniquely given by a development S that is located in R in the same way as S_0 is located in R_0. The developments S and S_0 have the same structure.*

Proof. Take some edge of the development S_0. If this edge coincides with some edge of the development R_0, then we take this edge in the development R as the corresponding edge of the development S.

Suppose that an edge AB of the development S_0 coincides with no edge of the development R_0. To this edge, there corresponds some segment in R_0 that goes along some triangles of the development R_0 and connects two vertices A and B of this development, since by condition the vertices of the developments R_0 and S_0 correspond to each other. Developing onto the plane all sequential triangles along which this segment goes, we obtain some polygon P with a diagonal AB (see Fig. 54). When the lengths of the edges of the development R continuously change, the polygon P also varies continuously. Therefore, there exists some $\varepsilon_{AB} > 0$ such that for each continuous change of the lengths of the edges which is less than ε_{AB}, the diagonal AB remains inside the polygon P. We can take the diagonal of the polygon P as the edge AB in the new development S. There is no other segment that connects the vertices A and B and intersects the same edges of the development R in the same order. Therefore, the edge AB in the development S is uniquely determined. Moreover, under a continuous deformation of the polygon P, the edge AB varies continuously. Arguing in this way for all edges of the development and taking ε less than all corresponding ε_{AB}, ε_{CD}, etc., we verify that for each change of the lengths of edges of the development R_0 which is less than ε we can take the corresponding edges of the development S_0. These edges connect the same vertices and traverse the same triangles of the development R. If ε is sufficiently small, then the lengths of these edges are close to their initial values. Cutting the development R along these edges and gluing the pieces of triangles of this development, we obtain the development S. It is clear by construction that the development S has the same structure as the initial development S_0. The lemma is proved.

Lemma 5. *Let the same metric be determined from two developments R and S whose vertices correspond to each other. Preserving the structure of the development R and the location of S in R, we change the edge lengths of the development R. Then by the previous lemma, the lengths of the development S turn out to be single-valued*

© 2006 by Taylor & Francis Group, LLC

functions of the length of the edges of the development R whenever the changes of edge lengths are sufficiently small. We assert that these functions are differentiable.

Proof. If an edge of the development S coincides with some edge of the development R, then the assertion is obvious since both edges are of the same length.

If an edge AB of the development S does not coincide with any edge of the development R, we again develop onto the plane this edge together with all triangles of the development R this edge traverses. The edge AB turns out to be a diagonal of the polygon P composed of these triangles. The sides of the polygon P are edges of the development R, and the angles of this polygon are the sums of the angles of triangles that compose P; therefore, they are differentiable functions of the lengths of the sides of these triangles, i.e., of the edge lengths of the development R. But the length of a diagonal is a differentiable function of the sides and angles of each polygon, and, therefore, this diagonal is also a differential function of the edges of the development R. This proves the lemma.

Thus, a polyhedral metric of positive curvature that has e vertices is given by developments that, for a given structure, are determined by the same number of variables $k = 3e - 6$, the edge lengths. The change of the edges of one development implies a single-valued and even differentiable change of the edges of the other. Therefore, when considering a given, as well as variable, metric, we can arbitrarily determine this metric from various developments. All these determinations are perfectly equivalent. Of course, we keep in mind the above condition that the vertices of each development lie at the vertices of the metric.

The set of all metrics with a given number of vertices can be transformed into a space or "manifold of metrics" in the following way. A neighborhood of a metric ρ_0 is assumed to be a set of all metrics ρ that can be given by those developments R that satisfy the following conditions: (1) all developments have the same structure as the structure of some development R_0 of the metric ρ_0; (2) the lengths of their edges differ from the lengths of the corresponding edges of the development R_0 by less than some given $\varepsilon > 0$. The edge lengths $r_1, \ldots, r_k$ of the developments R are coordinates in such a neighborhood. Passage from one set of developments to the other determines some transformation of the coordinates. Lemmas 4 and 5 assert that this transformation is one-to-one and differentiable. When speaking about metrics close to a given metric, we will keep in mind the metrics belonging to some neighborhood of the above type. Now the notion of convergence of metrics is defined as the convergence in the space of metrics, i.e., we say that metrics ρ_n converge to a metric ρ if, (1) starting from some index, all metrics ρ_n can be given by developments R_n of the same structure as some development R of the metric ρ; (2) the edge lengths of the developments R_n converge to the lengths of the corresponding edges of the development R_1.[5] The indicated geometric representation of the set of metrics that are close to a given one turns out to be convenient and will be used in what follows.

[5] A metric is a continuous function $\rho(XY)$ of points X and Y of the sphere S on which this metric is given. Therefore, the natural and general concept of the convergence of metrics is the concept of the convergence of functions. It is possible to show that the concept of convergence of polyhedral metrics with a given number of vertices introduced above is equivalent to the convergence of metrics as functions. But this is not necessary for us. Our goal is to consider sets of a finite number of parameters $(r_1, \ldots, r_n)$ instead of functions.

© 2006 by Taylor & Francis Group, LLC

2. The Idea of the Proof of the Realization Theorem

The proof of the existence of a polyhedron, which realizes a given metric, we will give here turns out to be very long if carried out in full detail. This is due to the following two reasons: First, this proof is based on one theorem on polyhedra, which is of significant interest in its own right (the rigidity theorem, Sec. 1 and 2), and second, there are a few difficulties in the very nature of the problem, which were already mentioned in the beginning of the chapter; many rather delicate arguments are required to overcome these difficulties. The idea of the proof itself is very simple, so we present it in general outline without elaborating the details yet. At the same time, by referring to the pages of this chapter where one or another assertion formulated is proved, we will give an outline of the whole proof.[6]

A convex polyhedron is completely determined from its vertices.[7] If we move the vertices of a polyhedron, then this polyhedron deforms in a unique fashion. Since a polyhedron itself can move as a rigid body, not every location of vertices produces a real deformation, because equal polyhedra can merely be considered as distinguishing copies of the same polyhedra. In order to exclude deformations which reduce to motions, we take three vertices A, B, and C on a face of a given polyhedron P and, choosing an arbitrary Cartesian coordinate system x, y, and z, displace the polyhedron P so that the vertex A coincides with the origin, and the vertex B lies on the positive semi-axis x; then we revolve the polyhedron so that the vertex C lies in the part of the plane $z = 0$ where $y > 0$. If the location of the vertices A, B, and C is always assumed to satisfy these conditions, then motion will be excluded. In what follows, a polyhedron whose three vertices of some of its faces satisfy these conditions will be called a *polyhedron with motion excluded.*

If a polyhedron has e vertices, then we have their $3e$ coordinates in total. But three coordinates of the vertex A are fixed, two coordinates of the vertex B are fixed, and one coordinate of the vertex C is fixed; therefore, we have $3e-6$ variable coordinates; changing these coordinates, we obtain new polyhedra. Hence a polyhedron with e vertices is determined by the assignment of $3e-6$ variables if motion is excluded. For sufficiently small values of these coordinates, the displacement of vertices can be arbitrary. Therefore, the coordinates of vertices assume arbitrary values in sufficiently small domains. However, for "large displacements" one or several vertices can enter the convex hull of the other vertices and cease to be vertices of the polyhedron. Hence, if we consider only polyhedra with e vertices, then their coordinates can vary only in some domain. However, the shape of this domain is completely immaterial.

Each polyhedron P has its definite metric; the edges of the polyhedron are the vertices of this metric. The latter can be given by some development R. The simplest way is, of course, to take the development formed by the polyhedron faces themselves; moreover, nontriangular faces can be partitioned into triangles by di-

[6] The first proof given by the author in the paper "Existence of a convex polyhedron and a convex surface with given metric," Mat. Sbornik, Vol. 11 (1941), pp. 15–61, is based on a similar but less elementary idea and comes across the same difficulties in a detailed presentation. Another proof was suggested by L. A. Lyusternik, but it is based on the Lewy–Weyl theorem on the existence of a convex surface with a given line element.

[7] See Sec. 5 of the Appendix where we prove that a convex polyhedron is the convex hull of its vertices.

© 2006 by Taylor & Francis Group, LLC

agonals. This development is said to be natural. We can also take some other development. However, we consider *only* those developments whose vertices correspond to the vertices of metrics, i.e., to the vertices of the polyhedron.

If the structure is given, then a development is determined from lengths of its edges $r_1, \ldots, r_k$. We have proved (Lemma 3 of Sec. 1) that the number of edges k is related to the number of vertices e by the equation

$$k = 3e - 6. \tag{1}$$

Hence *the number of variables determining a polyhedron is equal to the number of variables determining a development with the same structure.*

If a polyhedron is deformed by moving its vertices, then a given development changes. But it is sufficiently obvious and will be proved in Sec. 3 that under a small deformation of the polyhedron the structure of the development can be assumed to remain the same and only the edge lengths are varied. Hence, the lengths $r_1, \ldots, r_k$ turn out to be single-valued functions of the coordinates of the vertices $p_1, \ldots, p_k$ (there are $3e - 6$ coordinates of vertices when motion is excluded, i.e., there are k these coordinates by formula (1)); thus, in a neighborhood of the values $p_1^0, \ldots, p_k^0$ corresponding to the given polyhedron P, we have

$$\begin{aligned} r_1 &= f_1(p_1, ..., p_k), \\ &\ldots \\ r_k &= f_k(p_1, ..., p_k). \end{aligned} \tag{2}$$

What we want to prove is that, conversely, by giving a development, we can reconstruct the polyhedron. This problem can be treated as the problem of inverting functions (2) if we extend these functions to all polyhedra that admit developments of the same structure. Indeed, let us consider, on the one hand, all developments R of the same structure and, on the other hand, all polyhedra P that admit developments of this structure. Putting a definite development R in correspondence to each polyhedron P, we transform the edges of the development R into the functions of coordinates of the vertices of the polyhedron P, i.e., we obtain the set of functions (2). Conversely, if it would be possible to find a set of functions

$$\begin{aligned} p_1 &= g_1(r_1, ..., r_k), \\ &\ldots \\ p_k &= g_k(r_1, ..., r_k) \end{aligned} \tag{3}$$

such that (1) the functions g_i are defined for all values $r_1, \ldots, r_k$ corresponding to those developments that determine a metric of positive curvature and (2) the values of p_i obtained in this process are in fact the coordinates of vertices of a polyhedron,[8] then, to each development, we would put in correspondence a polyhedron, and, thus, our problem would be solved. Applying this result to developments of all possible structures, we would prove the realization theorem completely.

[8] That is, all points with these coordinates p_i, rather than only a part of these points, are vertices of their convex hull.

© 2006 by Taylor & Francis Group, LLC

However, it seems that this program cannot be carried out by a direct method for the following two reasons. First, as we have already mentioned in Sec. 1, the relation of a given development of a polyhedron with the natural development and, thus, with the coordinates of its vertices is very complicated. Second, if we consider the developments of an *a priori* given structure, then it is necessary to prove that, in addition, there exist polyhedra that admit developments of the same structure. Otherwise, there would be no functions (2) and nothing to invert.

In order to overcome these difficulties, we use the method of continuation whose essence is as follows.

1. First, we prove that functions (2) are invertible in a sufficiently small neighborhood of $(p_1^0, \dots, p_k^0)$. As will be explained in detail below, this proves that *all metrics sufficiently close to a realizable metric are also realizable.* (Recall that according to the definition of Sec. 1, metrics ρ are close to ρ_0 if these metrics can be given by developments of the same structure whose edge lengths are close to the edge lengths of the development R_0 determining the metric ρ_0.)

2. Second, we prove that each given metric ρ_1 can be connected with a realizable metric ρ_0 by a continuous family of metrics ρ_t $(0 \le t \le 1)$. (We say that metrics ρ_t form a continuous family if, to each t in the closed interval $[0, 1]$, there corresponds some metric ρ_t so that the metrics $\rho_{t'}$ converge to the metric ρ_t as $t' \to t$. The notion of convergence of metrics was defined in Sec. 1.

3. After that, passing from the metric ρ_0 to the neighboring metrics ρ_t and traversing the whole family ρ_t up to the given metric ρ_1, we prove the realizability of ρ_1.

Let us consider each item of this program in more detail. To prove the first assertion on the invertibility of functions (2), we use the following well-known inverse function theorem:

If a set of k equations with k variables of the form (2) *is satisfied for the values $p_1 = p_1^0, \dots, p_k = p_k^0$ and if the functions $f_i(p_1, \dots, p_k)$ are differentiable and their Jacobian is different from zero for $p_i = p_i^0$ $(i = 1, \dots, k)$, then these equations are solvable near the values $r_1^0, \dots, r_k^0$, and $p_1^0, \dots, p_k^0$, i.e., p_i turn out to be continuous functions of r_i that are defined for all values of r_i which are sufficiently close to r_i^0 and all p_i assume the values p_i^0 for $r = r_i^0$ $(i = 1, \dots, k)$.*

We will prove (in Sec. 3) that our functions $f_i(p_1, \dots, p_k)$ are in fact differentiable. Then we can write

$$\begin{aligned} dr_1 &= \frac{\partial f_1}{\partial p_1} dp_1 + \dots + \frac{\partial f_1}{\partial p_k} dp_k, \\ &\dots \\ dr_k &= \frac{\partial f_k}{\partial p_1} dp_1 + \dots + \frac{\partial f_k}{\partial p_k} dp_k. \end{aligned} \tag{4}$$

The determinant of this system of equations in the unknowns dp_i is the Jacobian. Therefore, the assertion that this Jacobian is different from zero is equivalent to the

© 2006 by Taylor & Francis Group, LLC

assertion that system (4) has no solutions but the trivial one when all $dp_i = 0$, in the case where all dr_i vanish. Therefore, in order to prove the invertibility of the set of function (2), it is sufficient to show that if all $dr_i = 0$ then all $dp_i = 0$.

This last assertion has a simple geometric meaning. Indeed, the fact that Eqs. (2) are satisfied for the values $p_i = p_i^0$ and $r_i = r_i^0$ means first that we have the polyhedron P_0 with coordinates of the vertices equal to p_i^0, and second, we have the development R_0 of this polyhedron (of a given structure) whose edge lengths are equal to r_i^0. Suppose that the polyhedron P_0 is continuously deformed in time t in such a way that the coordinates p_i vary with some speeds p_i. Then the edge lengths of the development also vary with some speeds r_i (this follows from the differentiability of the functions $f_i(p_1, \ldots, p_k)$). The assertion is that if all $\dot{r}_i = 0$, then all $\dot{p}_i = 0$, i.e., if all r_i are constant to within magnitudes of the second order, then all p_i are also constant to within magnitudes of the second order. Roughly speaking, this means that when motion is excluded, our polyhedron is not deformed if its development is not changed. If motion is admitted (i.e., if we omit the assumptions that are imposed on the location of the vertices A, B, and C), then the above-obtained result can be formulated as the following theorem.

If the edge lengths of a development are stationary, then a given polyhedron admits no deformation but infinitely small motions. Here a variable $x(t)$ is called stationary if

$$\dot{x}(0) \equiv \left(\frac{dx}{dt}\right)_{t=0} = 0.$$

If a figure admits no infinitely small deformations but motions, it is called *rigid*. Therefore, this theorem is a theorem on the rigidity of a closed convex polyhedron when the edge lengths of some development of it are stationary. We will prove this theorem for nondegenerate polyhedra; but it does not hold for degenerate polyhedra as is easily seen.[9] (The exception is a doubly-covered triangle for which it is trivial.)

The above rigidity theorem is a generalization of the well-known theorem already proved by Cauchy. Namely, Cauchy proved that *if the faces of a nondegenerate convex closed polyhedron are rigid*[10] *then the polyhedron itself is rigid.* The generalization is that we do not assume the rigidity of faces: in our case, they can flex; we assume only that all edge lengths of some development of the polyhedron are stationary. This requirement is weaker than the requirement of rigidity of faces even in the case where we deal with the natural development of the polyhedron. Indeed, if our polyhedron has, e.g., a quadrangular face, then this face can flex along each of its diagonals, although all its sides and diagonals preserve their lengths. Only in the case where all faces of a polyhedron are triangular and namely these faces are taken as the faces of the development, the stationarity of edges implies the

[9] Let $ABCD$ be a tetrahedron degenerating into a doubly-covered square with side a. Suppose that the diagonals AC and BD drawn on different sides of the plane of this square play the role of edges. Let us put forth the vertex C from the plane of the square by the height h. Then the edges AB, AD, and BD do not change, the edges BC and DC become equal to $\sqrt{a^2 + h^2}$, and the edge AC becomes equal to $\sqrt{2a^2 + h^2}$. We see from this that the change of edge lengths is of order h^2, i.e., of the second order with respect to the speed of motion of the vertex C if $h = \dot{h}t$. Therefore, for $t = 0$, the derivatives of the edge lengths in t are equal to zero, i.e., the edge lengths are stationary.

[10] That is, all deformations of faces, except for the motion of each of which as a whole, are of the second order of smallness.

© 2006 by Taylor & Francis Group, LLC

rigidity of faces. In this case, deformations of faces can consist only in the change of their edge lengths (since each triangle is completely determined by its sides), while the faces turns out to be rigid under the stationarity condition for the sides. Our generalized rigidity theorem will be proved in Secs. 4 and 5 by the same method Cauchy used to prove his theorem.

If the rigidity theorem for nondegenerate polyhedra is proved, then, as we have shown, this theorem implies that the Jacobian of the system of functions (2) is different from zero, and, hence, this system is invertible in some small neighborhood of the values $r_i = r_i^0$ and $p_i = p_i^0$ corresponding to some nondegenerate polyhedron.

This implies that metrics sufficiently close to the realizable ones are realizable. Indeed, let a given metric ρ_0 be realized by some polyhedron P_0, and let $p_1^0, \ldots, p_k^0$ be the coordinates of vertices of this polyhedron. The metric ρ_0 is given by some development R_0 with edge lengths $r_1^0, \ldots, r_k^0$; this is some development of the polyhedron P. When the vertices of the polyhedron P are subject to arbitrary but sufficiently small motions (however, keeping in mind the condition imposed on three vertices A, B, and C that excludes the motion of the polyhedron as a whole), we obtain new polyhedra P whose vertices have the coordinates $p_1, \ldots, p_k$. Hence the coordinates $p_1, \ldots, p_k$ can assume arbitrary values sufficiently close to $p_1^0, \ldots, p_k^0$. As was already mentioned, we will prove in Sec. 3 that polyhedra P that are sufficiently close to the given P_0 admit developments of the same structure as R_0 and have the edge lengths $r_1, \ldots, r_k$ close to $r_1^0, \ldots, r_k^0$. Thus, near the values $p_1^0, \ldots, p_k^0$, the edge lengths r_i turn out to be some functions of the coordinates p_i of vertices; these are functions (2). Assume now that we have proved these functions to be invertible, i.e., we have functions (3), $p_i = g_i(r_1, \ldots, r_k)$, where $r_1, \ldots, r_k$ are independent variables near the values $r_i = r_i^0$ $(i = 1, \ldots, k)$, and for $r_i = r_i^0$, the functions g_i assume the values $p_i = p_i^0$, which correspond to the initial polyhedron P_0.

Let ρ_1 be a metric sufficiently close to the given ρ_0; by the definition of the proximity of metrics, the metric ρ_1 can be given by some development R_1 of the same structure as R_0 with the edge lengths r_i^1 close to r_i^0. Let $p_i^1 = g_i(r_1^1, \ldots, r_k^1)$ be the values of functions (3) for $r_i = r_i^1$; these values are close to p_i^0, and thus, there exists a convex polyhedron P_1 whose vertices have the coordinates p_i^1. At the same time, this polyhedron admits a development R_1' of the same structure as R_0, and the edge lengths of this development are represented by functions (2), i.e., they are equal to $f_i(p_1^1, \ldots, p_k^1)$. But since the functions f_i are inverse to the functions g_i and $p_i^1 = g_i(r_1^1, \ldots, r_k^1)$, we have $f_i(p_1^1, \ldots, p_k^1) = r_i^1$; this means that the development ${R'}_1^1$ has the same edges as R_1. Hence the polyhedron P_1 has the development R_1, and so P_1, realizes the metric ρ_1.

The obtained result can be stated in the form of the following main lemma.

Lemma A. *If a closed convex nondegenerate polyhedron P_0 originates by gluing from some development R_0, then a closed convex polyhedron originates by gluing from each development of the same structure as R_0 whose edge lengths are sufficiently close to the edge lengths of the development R_0.*

In short: *Each metric close to a realizable one is also realizable.*

By this, the first item of our program is carried out. To implement the second item, we proceed by induction. We have proved in Sec. 1 (Lemma 2 of Sec. 1) that

© 2006 by Taylor & Francis Group, LLC

each metric with three vertices is realizable. Therefore, assuming that each metric with $e-1$ vertices is realizable, we will prove the realizability of metrics with e vertices ($e > 3$). Then we can demonstrate the following assertion.

Lemma B. *If a metric ρ has more than three vertices, it can be connected with a metric ρ_1 realized by a nondegenerate closed convex polyhedron by a continuous family of metrics; this family ρ_t can be chosen so that all the metrics ρ_t have the same number of vertices and none of them corresponds to a degenerate polyhedron (i.e., none of them can be realized by a degenerate polyhedron).*

This lemma will be proved in Sec. 7.

Finally, we will prove in Sec. 6 the following assertion.

Lemma C. *The limit of realizable metrics is a realizable metric.*

The plan of proving this lemma is obvious. If polyhedra P_n realize the metrics ρ_n that converge to a metric ρ, then we can extract a convergent sequence from them. The limit polyhedron of this sequence has the limit metric ρ. In fact, this is a particular case of the general theorem on the convergence of metrics of convergent convex surfaces. However, we have just defined the convergence of metrics via the convergence of developments, so it is necessary to prove Lemma C without reference to the general convergence theorem.

If Lemmas A, B, and C are proved, then the proof of the theorem on the existence of a polyhedron with a given metric is very easy to finish. Indeed, let ρ_0 be a given metric with e vertices ($e > 3$). Let ρ_t ($0 \le t \le 1$) be a continuous family of metrics that satisfies all the conditions indicated in Lemma B. This family connects ρ_0 with the metric ρ_1 realized by a nondegenerate polyhedron. In such a case, according to Lemma A, every metric close to ρ_1 is also realizable; in particular, all metrics ρ_t for t sufficiently close to 1 are realizable.

Now let T be the greatest lower bound of those t for which the metrics ρ_t are realizable. Since the family of metrics ρ_t is continuous, the metric ρ_T turns out to be the limit of the realizable metrics ρ_t ($t \to \infty$); therefore, according to Lemma C, the metric ρ_T is also realizable by some polyhedron P_T. Hence, if $T = 0$, then the theorem is proved. But T cannot be different from zero. Indeed, if $T > 0$, then the polyhedron P_T is nondegenerate, since there are no metrics corresponding to degenerate polyhedra in the family ρ_t for $t > 0$. Then, by Lemma A, all metrics close to ρ_T and, in particular, the metrics ρ_t for $t < T$ are realizable. This means that T is not the greatest lower bound of those t for which the metrics ρ_t are realizable. Consequently, $T = 0$, which completes the proof of the theorem.

Further presentation deals with the implementation of the above program. In Sec. 3, we give the proof of the existence and differentiability of functions (2) in a neighborhood of $p_1^0, \ldots, p_k^0$ corresponding to some given polyhedron. In Sec. 4, we prove some lemmas necessary for the proof of the rigidity theorem, which is then proved in Sec. 5. In Sec. 6, on the basis of the rigidity theorem, we prove Lemma A; here, Lemma C is also proved. Finally, Sec. 7 is devoted to the proof of Lemma B.

Except for the rigidity theorem, which should be considered as an independent result, this Lemma B is, in fact, the only difficult point of the whole program. After the proof of Lemmas A, B, and C, in Sec. 8 it remains only to repeat the above argument, which proves the realizability of any metric.

© 2006 by Taylor & Francis Group, LLC

3. Small Deformations of a Polyhedron

Let P_0 be a closed convex polyhedron. We will displace continuously the vertices of this polyhedron and construct convex polyhedra from them. If the location of vertices is given, then the polyhedron is completely determined, since each closed convex polyhedron is the boundary of the convex hull of its vertices (see the Appendix, Sec. 5, Theorem 2). We say that this polyhedron P_0 deforms as a result of displacement of its vertices.

If the displacement of vertices is sufficiently small, then the displaced vertices are the vertices of a new polyhedron. Indeed, a displaced vertex is no longer a vertex if the planes of faces passing through this vertex do not make a polyhedral angle when they unbend into a dihedral angle or even into a plane. But since the plane passing through the vertices and, thus, the faces of the polyhedron move continuously when the vertices continuously move, this can occur only when the displacement of vertices becomes sufficiently large. For the same reason, under sufficiently small displacements of vertices, the vertices not belonging to one face do not arrive at one plane, and thus, will not belong to one face (in other words, two faces will not unbend into one face). However, if there are more than three vertices in one plane, then even for arbitrarily small displacements these vertices do not stay in one plane in general. Therefore, under arbitrarily small displacements of vertices more than triangular faces can flex and form several faces. These flexes can go along the diagonals of faces; moreover, for some displacements, these flexes go along some diagonals while for other flexes they go along the others. Thus, although for small deformations induced by displacements of vertices the number of vertices is preserved and different faces do not unbend into one face, and, therefore, the edges do not vanish, there arise new edges that correspond to some diagonals of nontriangular faces on the polyhedron not all of whose faces are triangular in general. The structure of this polyhedron changes, and, moreover, this change varies depending on the character of the displacement of vertices. Considering the deformed polyhedron, we will call old those elements (faces, edges, angles) that were in the initial polyhedron, and we will call new those elements that appear as a result of deformation. Thus, when an "old" face flexes, it transforms into "new" faces separated by "new" edges.

If we divide nontriangular faces of our polyhedron into triangles by diagonals that do not intersect each other inside faces, then this polyhedron turns out to be composed of triangular "faces". The development comprising these "faces" is called natural. In general, a polyhedron has several natural developments, but their number is certainly finite. Under small displacements of vertices, the natural development may fail to remain natural if its triangles flex. But since flexes of faces go along diagonals, we always associate some natural development of the initial polyhedron with each natural development of the deformed polyhedron.

Lemma 1. *Let P_0 be a closed convex polyhedron and R_0 its development whose vertices coincide with those of the polyhedron. Assume that a polyhedron P is obtained from P_0 by small displacements of vertices. If the displacements are sufficiently small, P admits a unique development R of the same structure as the development R_0 located relative to a natural development of the polyhedron P in the same way as R_0 is relative to the corresponding natural development of the polyhedron P_0.*

© 2006 by Taylor & Francis Group, LLC

Proof. Consider polyhedra P that are close to P_0 and have natural developments S of the same structure. Let S_0 be the corresponding development of the polyhedron P_0. We have two developments R_0 and S_0 of the polyhedron P_0 and the development S of the polyhedron P. Under continuous displacement of vertices, the edge lengths of the development S are continuously varied, since its edges are edges and diagonals of faces of the polyhedron. Therefore, by Lemma 4 of Sec. 1, there exists $\varepsilon_S > 0$ such that whenever the displacement of vertices is less than ε_S (and is such that the resulting polyhedron P admits a natural development S), the development S can be replaced by unique development R which has the same structure as the given development R_0 and which is located relative the natural development S in the same way as R_0 is located relative to S_0.

Since the number of possible distinct developments is finite, for each of them, we take its own ε_S and then take the minimum of these ε_S as ε; then we obtain that for each displacement of vertices that are less than ε, the polyhedron P admits the unique development R with the required properties. The lemma is proved.

Assume that we have some nontriangular face Q_0 of the polyhedron P_0. Under the displacement of vertices, this face flexes along some diagonals and transforms into a flexed polygon Q. Each of the other diagonals is replaced by a flexed line; under the development of the changed face Q, this flexed line transforms into some diagonal of the resulting polygon. Let us prove the following

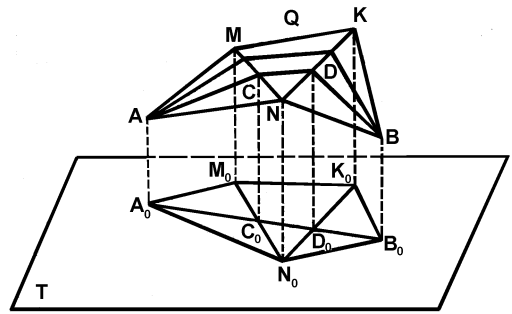


Fig. 55

Lemma 2. *The length of a flexed diagonal on the face Q differs from the distance between its endpoints by the second order of smallness relative to the displacement of vertices.*

Proof. Let T be the plane on which the face Q lies before the displacement; let AB be the flexed diagonal of this face (see Fig. 55 where AB is denoted by a bold line). Finally, let A_0 and B_0 be the projections of the points A and B to the plane T. If $\overline{AB}$ and $\overline{A_0B_0}$ stand for the straight line segments connecting the points A and B, and, respectively, the points A_0 and B_0, then we obviously have the following relation for the lengths of the flexed line AB and the straight line segments $\overline{AB}$ and $\overline{A_0B_0}$:

$$AB \geq \overline{AB} \geq \overline{A_0B_0}. \tag{1}$$

The face Q flexes along some diagonals (the diagonals MN and NK in Fig. 55). Let us project these diagonals to the plane T, and let $C_0, D_0, \ldots, G_0$ be the intersection points of these projections with the segment $\overline{A_0B_0}$.[11] If we erect perpendiculars to

[11] The projections of diagonals are slightly different from the initial position of the diagonals themselves in the plane T. Therefore, if displacements of vertices are sufficiently small, then the segment $\overline{A_0B_0}$ coincides with no projection of other diagonals, since this segment is also a diagonal at the initial position.

© 2006 by Taylor & Francis Group, LLC

the plane T at the points C_0, D_0, ..., G_0, then these perpendiculars intersect the corresponding diagonals at some points C, D, ..., G. Since the flexed diagonal AB is the shortest arc of the points A and B on the flexed face Q, we have

$$AC + CD + \cdots + GB \geq AB. \tag{2}$$

On the other hand, the segment $\overline{A_0B_0}$ is divided by the points C_0, D_0, ..., G_0 into the segments $\overline{A_0C_0}, \overline{C_0D_0}, \ldots, \overline{G_0B_0}$ and

$$\overline{A_0C_0} + \overline{C_0D_0} + \cdots + \overline{G_0B_0} = \overline{A_0B_0}. \tag{3}$$

Relations (1), (2), and (3) imply

$$AC + \cdots + GB \geq AB \geq \overline{AB} \geq \overline{A_0C_0} + \cdots + \overline{G_0B_0} \tag{4}$$

or

$$|AB - \overline{AB}| \leq |AC - \overline{A_0C_0}| + \cdots + |GB - \overline{G_0B_0}|. \tag{5}$$

Now, if we prove that each of the differences on the right-hand side of this inequality has the second order of smallness relative to the displacements of vertices, we prove our lemma.

Let us consider, e.g., the difference $AC - \overline{A_0C_0}$.

Obviously,

$$AC^2 = \overline{A_0C_0^2} + (\overline{AA_0} - \overline{CC_0})^2. \tag{6}$$

Let M and N be the endpoints of the diagonal that contains the point C and along which the face Q flexes. Let M_0 and V_0 be their projections to the plane T. Since the points M, C, and N lie on one line and the points M_0, C_0, and N_0 lie on the projection of this line to the plane T, we have

$$\frac{\overline{MM_0} - \overline{NN_0}}{\overline{MN}} = \frac{\overline{CC_0} - \overline{NN_0}}{\overline{CN}}.$$

This implies

$$\overline{CC_0} = \frac{\overline{CN}}{\overline{MN}}\overline{MM_0} + \frac{\overline{CM}}{\overline{MN}}\overline{NN_0}. \tag{7}$$

But the segments $\overline{MM_0}$ and $\overline{NN_0}$ are nothing else but the projections of the displacements of the vertices M and N to the direction perpendicular to the plane T. The segment $\overline{AA_0}$ is the same projection of the displacement of the vertex A. Therefore, using Eq. (7), we see from Eq. (6) that the difference $AC^2 - \overline{A_0C_0}^2$ and thus $AC - \overline{A_0C_0}$ are magnitudes of the second order relative to the displacements of vertices. Other differences $|CD - \overline{C_0D_0}|$, etc. are estimated in exactly the same way; but here, both points C and D lie on the diagonals, and the formulas similar to (7) should be applied to both points. The lemma is proved.

Lemma 3. *Let P_0 be a given polyhedron and R_0 be a development of this polyhedron whose vertices are those of the polyhedron. Let the polyhedron P_0 be deformed as a result of displacement of vertices and the development R corresponding to R_0 be constructed in accordance with Lemma 1 for the deformed polyhedron P. Then the edge lengths of the variable development R are differentiable functions of the coordinates of vertices.*

© 2006 by Taylor & Francis Group, LLC

Proof. Lemma 1 implies that the edge lengths of the development R are single-valued continuous functions of the coordinates of vertices which are defined in some range of these coordinates near their initial values. By Lemma 2, the diagonal of the flexed face differs from the distance between their endpoints by a magnitude of the second order. The distance between the vertices is, of course, a differentiable function of their coordinates, i.e.,

$$\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}.$$

Therefore, the length of the flexed diagonal at the initial moment of flex is also a differentiable function of the coordinates of vertices.

Take a natural development of the polyhedron P_0. The differentiability of diagonals just proved implies that all edges are differentiable functions of the coordinates of vertices. But by Lemma 5 of Sec. 1, the edge lengths of one development are differentiable functions of the other, and, therefore, the edge lengths of the development R are differentiable functions of the edge lengths of the natural development. Hence, they are also differentiable functions of the coordinates of vertices. (This was proved for the initial moment of deformation. But since each of its moments can be considered as initial, we have thus proved the differentiability of the edge lengths of the development for the whole range of coordinates of vertices for which the polyhedron P still admits the development R.)

In what follows, we will assume that the vertices of a polyhedron move with speeds defined at each point, i.e., their coordinates are represented as differentiable functions of some parameter t, which is convenient to consider as time. The initial polyhedron corresponds to t equal to zero. Then all angles on faces and between faces also vary with certain speeds, since they are differentiable functions of coordinates of vertices (this is easily verified). In exactly the same way, the edge lengths of each development of a polyhedron will vary with some speeds by Lemma 3.

4. Deformation of a Convex Polyhedral Angle

Before proving the rigidity theorem for nondegenerate closed convex polyhedra, we first consider the deformation of each of the polyhedral angles that a polyhedron deforms as a result of displacement of vertices. However, the fact that our polyhedral angle belongs to the polyhedron is immaterial. We can take an arbitrary polyhedral angle, draw new edges from the vertex on its faces, and revolve the old, as well as the new, edges around the vertex with certain speeds. Then the polyhedral angle deforms, its faces flex along new edges, and the polyhedral angles made by them vary with some speeds. Moreover, we will consider only those deformations under which the polyhedral angle remains convex. The terms "old" and "new" referring to edges and to plane and dihedral angles will be understood in the same way as they were defined in the previous section. Unless otherwise specified, we speak about the old as well as the new edges and angles between faces.

A polyhedral angle with n faces is determined by its $n - 1$ plane angles and all $n - 2$ dihedral angles made by these plane angles. Therefore, an n-hedral plane angle a_n is a function of these $n - 1$ plane angles $a_1, \ldots, a_n$ and $n - 2$ dihedral

© 2006 by Taylor & Francis Group, LLC

angles $\alpha_1, \ldots, \alpha_{n-2}$ made by them (dihedral angles at the new edges are equal to π in the beginning of the deformation). We say that a dihedral angle α_i lies opposite a plane angle a_i if the edge of the angle a_i does not belong to the old face at which the angle a_n lies at the initial moment.

Lemma 1. *The partial derivative of the plane angle a_n with respect to each of the dihedral angles α_i lying opposite a_n is positive.*[12]

Proof. Take the dihedral angle α_i that lies opposite a_n; let r be the edge of this dihedral angle. Let p and q be the edges of the plane angle a_n (Fig. 56). Draw the plane P and Q through the edges r and p, r and q. These planes, together with the plane of the angle a_n, bound the trihedral angle with edges p, q, and r (these edges do not lie in one plane since the angle α_i lies opposite a_n). Let b and c be the plane angles of this trihedral angle between the edges p and r , q and r, and let α be the dihedral angle made by these plane angles. If we circumscribe the unit sphere around the vertex of our polyhedral angle, then this angle cuts out a convex spherical polygon along this sphere. The angles b and c are diagonals of this polygon. Obviously, for constant angles $a_1, \ldots, a_n$ and $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_{n-2}$, we can vary the angle α_i by revolving the plane P and Q around the edge r and by revolving the polyhedral angles together with them which are cut off by these planes from the given n-hedral angle; in this process, all plane angles $a_1, \ldots, A - n - 1$ and also b and c and the dihedral angles $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_{n-2}$ remain the same. The increment of the angle α_i is equal to that of the angle α. Therefore,

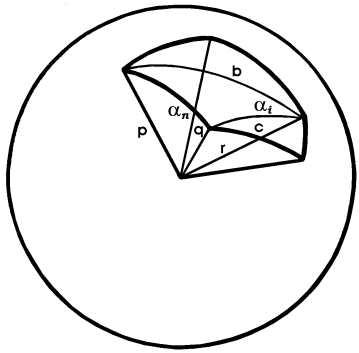


Fig. 56

$$\frac{\partial a_n}{\partial \alpha_i} = \frac{\partial a_n}{\partial \alpha}. \tag{1}$$

But the derivative of the plane angle with respect to the opposite dihedral angle is positive in the trihedral angle. This can be seen, e.g., from the following well-known formula of spherical trigonometry:

$$\cos a_n = \cos b \cos c + \sin b \sin c \cos \alpha;$$

by this formula,

$$\frac{\partial a_n}{\partial \alpha} = \sin b \sin c \frac{\sin \alpha}{\sin a_n} > 0, \tag{2}$$

since the angles b, c, a_n are obviously $< \pi$ and > 0, while $0 < \alpha\pi$, since the edge r is outside the old face that includes the angle a_n.

[12] We have already mentioned that $a_n = f(\alpha_1, \ldots, \alpha_{n-1})$; it is easy to show that this function is differentiable.

© 2006 by Taylor & Francis Group, LLC

Relations (1) and (2) imply

$$\frac{\partial a_n}{\partial \alpha_i} > 0,$$

as required.

Recall that a variable $x(t)$ is called stationary if its derivative with respect to t is equal to zero.

Lemma 2. *If there are new edges between two neighboring edges p and q of a convex polyhedral angle, and the plane angles into which these new edges divide the faces between p and q are stationary in the process of deformation, then the angle between the edges p and q is also stationary.*

Proof. Let $r_1, \ldots, r_m$ be new edges that lie between the edges p and q. Take points P, Q, $R_1, \ldots, R_m$ on the edges p, q, $r_1, \ldots$, R_m which are distant from the vertex O of our polyhedral angle by constant distances. Then we obtain the face V with vertices O, P, R_1, $\ldots$, R_m, Q that lies between the edges p and q.

The displacement of the edges p, q, $r_1, \ldots, r_m$ can certainly be considered as that resulting from the motion of the points P, Q, R_1, $\ldots$, R_m. Then the face V flexes due to the displacement of these points. The sides OP, OR_1, $\ldots$ are constant in the triangles OPR_1, OR_1R_2, etc., and the angles between them are stationary by the condition of the lemma. Therefore, the distance between the points P and Q on the face V is also stationary.[13] But by Lemma 2 of the preceding section, the spatial distance between the points P and Q differs from the distance between them on the face V by a magnitude of the second order of smallness. Therefore, the spatial distance between P and Q is also stationary.

Now, the sides OP and OQ are constant in the triangle OPQ, while the side PQ is stationary. Hence the angle opposite the side PQ, i.e., the angle between the edges p and q is also stationary, as required.

Let a convex polyhedral angle be deformed, and, moreover, let the emergence of new edges be admitted. We put the plus sign on an edge if the polyhedral angle on this edge increases with nonzero speed and the minus sign if this angle decreases with a nonzero speed. We call these edges distinguished. Other edges at which the polyhedral angles are stationary shall be undistinguished. This convention on the signs of edges will be used in the next section, and, therefore, should be kept in mind. Since faces flex along new edges under deformation, the angle at a new edge can only decrease since we consider deformations that yield convex polyhedral angles when these edges can move only outside these angles. Therefore, the new edges can have only the minus sign.

We prove here the following important lemma.

[13] The stationarity of the sides and angles between them implies the stationarity of other sides and angles of the triangles OPR_1, OR_1R_2, etc., i.e., the stationarity of all sides and angles of the faces V which are developed onto the plane and composed of these triangles. But if all sides and angles are stationary in a polygon, then its diagonals are also stationary. Hence, in particular, the diagonal PQ is stationary. This diagonal only flexes down, but its length is constant up to magnitudes of the second order.

© 2006 by Taylor & Francis Group, LLC

Lemma 3. *Let all plane angles of a deformed convex polyhedral angle be stationary. Putting the signs on its edges according to the above condition, we study the change of signs when going around the vertex of this polyhedral angle. The following three cases are only possible here:*

1. *The number of sign changes is equal to zero, in which case there are no distinguished edges, i.e., all dihedral angles are stationary.*

2. *The number of sign changes is not less than four.*

3. *The number of sign changes is equal to two. Then all old edges, except for any two neighboring edges p and q, are nondistinguished. The plus stands at the edges p and q. There are new edges between the edges p and q that are marked by the sign minus. All other edges remain undistinguished. In other words, only the face between p and q flexes with nonzero speed and the angles between p and q increase, while all other faces do not flex and the angles between them remain stationary.*[14]

Proof. Let us prove that there are no distinguished edges at all in the absence of sign changes. Indeed, if there are no sign changes, then the distinguished edges must have the same sign, say, plus. Take one of the old plane angles lying opposite at least one of the distinguished edges. Let a stand for this angle. Since all plane angles are stationary, we see that a is also stationary by Lemma 2, i.e., $da = 0$, although this angle can flex whenever some new edges appear on this angle. The angle a is a function of all other plane angles and the dihedral angles opposite to it. All plane angles are stationary and all dihedral angles, except for the angles $\alpha_1, \ldots, \alpha_m$ marked by plus, are also stationary, i.e., their differentials are equal to zero. We thus have

$$da = \frac{\partial a}{\partial \alpha_1} d\alpha_1 + \cdots + \frac{\partial a}{\partial \alpha_m} d\alpha_m.$$

Since all distinguished edges have the plus sign, $d\alpha_1, \ldots d\alpha_m > 0$. By Lemma 1 all derivatives $\frac{\partial a}{\partial \alpha_1} \ldots \frac{\partial a}{\partial \alpha_m}$ are also greater than zero. Therefore, $da > 0$ which, however, contradicts the already-proved stationarity of a. Hence, the first case of the lemma is clear: if there are distinguished edges then we have sign changes. Only an even number of sign changes is possible, since each sequence of pluses has two terminal entries at which the sign change occurs. The number of sign changes is thus equal to the doubled number of the sequences of pluses (or minuses). Therefore, if the number of sign changes is other than two, then there are at least four changes. This implies that the second case of the lemma can be omitted and only the case of two sign changes should be considered. Therefore, assume that we have exactly

[14] The first case takes place, e.g., when the polyhedral angle is not deformed at all. The simplest way to obtain the second case is to deform a tetrahedral angle preserving its faces unchanged. The third case can be seen in the simplest way by examining the following example. Take a trihedral angle with edges p, q, and r and draw the bisectrix s of the angle made by the edges p and q. We move this bisectrix outside this angle preserving the angles made by this bisectrix and the edges p and q. Then the angle at the bisectrix s decreases, while the angles at the edges p and q increase. The angle at the edge r also decreases, but the rate of this decay is zero at the initial moment, i.e., this angle is stationary. Indeed, by Lemma 2, the angle made by the edges p and q is stationary, and, therefore, the opposite dihedral angle is also stationary.

© 2006 by Taylor & Francis Group, LLC

two sign changes. Since only the minus sign can stand on new edges, there are old edges marked by plus sign.

A priori, only the following two possibilities can occur.

1. The plus sign stands on all old edges.

2. There are old edges marked by the minus sign.

Let us show that the second possibility is excluded. Since we have only two sign changes, there is only one sequence of edges with the plus sign and only one sequence of edges with the minus sign. Assume that the passage from the plus sign to the minus sign occurs on a face P, and let the reverse passage occur on a face Q. Draw a plane R through the vertex of the polyhedron angle that divides the angles at these faces into two halves. Since there are old edges among the edges marked by plus as well as by minus, the plane R passes between them inside our polyhedral angle and divides our angle into two angles V^+ and V^-. The polyhedral angle V^+ contains all edges that are marked by plus, and the angle V^- contains all edges marked by minus (see Fig. 57, where the polyhedral angle is replaced by a spherical polygon). By condition, all plane angles of these polyhedral angles, except for their common angle a on the plane R, are stationary. All dihedral angles α_i at the distinguished edges in the angle V^+ have positive differentials: $d\alpha_i > 0$. Since $\partial a/\partial\alpha_i > 0$ by Lemma 1, we have $da > 0$. But all dihedral angles α_i in the angle V^- have negative differentials: $d\alpha_i < 0$, and, therefore, we have $da < 0$. This is a contradiction which shows that the second case is impossible, i.e., there are no old edges marked by minus sign.

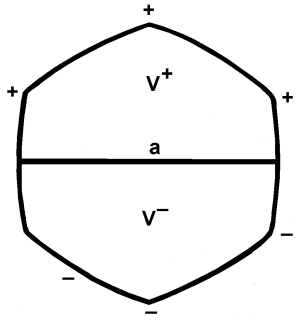


Fig. 57

We now let all distinguished old edges have the plus sign. Then there are new edges marked by the minus sign. Show that these new edges are included only in one old face.

Indeed, if we assume the contrary then the sequence of edges marked by minus is divided into two parts; each of these parts consists of new edges that lie on the same old face, and, moreover, they are divided by old edges. These old faces are not distinguished, since by condition the minus sign cannot stand on old edges, while the plus sign cannot stand on them by the condition that there are only two sign changes. Again, we draw a plane R through the bisectrices of those faces at which the passage from one sign to the other occurs. This plane goes inside our polyhedral angle, since it must separate old edges with plus signs from those that lie inside the sequence of new edges with minus signs. The plane R divides our polyhedral angle into two angles V^- and V^+; V^- contains only the edges with plus signs, and V^- contains only the edges with the minus signs. Therefore, as in the case just considered, the angle a on the plane R has positive differential in V^+ and negative differential in V^-. But this is a contradiction, so the new edges with minus signs can lie only on one old face.

© 2006 by Taylor & Francis Group, LLC

Assume that these edges lie on the old face between its edges p and q. Let us show that there cannot be plus signs on all old edges, except for p and q. Assume the contrary: let plus signs stand on old edges $r_1, \ldots, r_m$. By Lemma 2, the plane angle a made by p and q is stationary; if we draw an unflexed face through these edges, we obtain again a polyhedral angle with stationary plane angles.[15] If $\alpha_1, \ldots, \alpha_m$ are dihedral angles with edges $r_1, \ldots, r_m$, then $d\alpha_1, \ldots, d\alpha_m > 0$ according to our assumption. But since $\partial a/\partial \alpha_i > 0$ by Lemma 1 we have $da > 0$. However, this contradicts the stationarity of the angle A we have just proved, so there are no plus signs on all old edges but p and q.

Now, in order to complete the proof of the last assertion of our lemma, it remains only to show that the plus signs mark both edges p and q.

One of them must have the plus sign, since otherwise there is no sign changes at all. Assume that the plus sign stands only on the edge p. Then we draw a plane R through the bisectrices of the faces meeting at the edge p.

This plane separates the edge p from other edges and, therefore, divides the polyhedral angle into two angles V^+ and V^- again; one with the edge p marked by the plus sign and the other with the edges marked by only the minus sign. Repeating the argument which was twice applied before, we infer that the differential of the angle on the plane R would have different signs in V^+ and V^-; this is impossible since this angle is the same in each of these polyhedral angles. This contradiction shows that the plus sign cannot stand on one edge p. Thus, we have proved that if there are only two sign changes, then (1) new angles marked by the minus signs can be on only one old face, (2) only the edges of this old face can be marked by the plus signs, and (3) both of them are marked by the plus signs. This completes the proof of our lemma.

5. The Rigidity Theorem

We now prove the rigidity theorem for a nondegenerate closed convex polyhedron.

Theorem. *If a convex closed polyhedron nondegenerate into a polygon deforms as a result of the displacement of vertices so that the edge lengths of one of its developments are stationary, then all its vertices are also stationary, i.e., their speeds are equal to zero at the initial moment provided that the motion of this polyhedron is excluded as a solid.* Here, we speak about a development whose vertices coincide with vertices of that polyhedron.

Proof. Let P be a closed convex polyhedron nondegenerate into a polygon and R be a development of this polyhedron which does not contain other vertices except for the vertices of the polyhedron. Assume that the vertices of the polyhedron are displaced when the parameter t changes in such a way that at each moment the coordinates of vertices vary with some speeds. Then by Lemma 3 of Sec. 3, the edge lengths of the development also vary with certain speeds. With the value $t = 0$, we associate the initial position of the vertices of the polyhedron, and, by condition,

[15] That is, instead of the given deformation of the polyhedral angle, we consider another deformation under which the edges p and q move exactly in the same way as under the given deformation, and there arise no new edges between them at all. This means that we draw unflexed faces between p and q.

© 2006 by Taylor & Francis Group, LLC

the speeds of varying the edge lengths of the development R are equal to zero for $t = 0$.

Even for arbitrarily small deformations, the faces of the polyhedron P can flex in various ways. But we consider only one of these possible ways and, respectively, only those values of t for which the faces flex in the same manner; the other values of t are excluded from consideration. Then new definite faces will arise on the polyhedron P if the old faces do flex somehow. If these new faces are not triangular, then we can triangulate them by diagonals, so we can assume that the polyhedron P as a whole consists of triangular faces which never flex under deformation but only change their shape and rotate.

Thus, we have obtained a natural development of the polyhedron P. By Lemma 5 of Sec. 1, the edge lengths s_i of this development are differentiable functions of the edge lengths of the development R. Therefore, $ds_i = \sum(\partial s_i/\partial r_j)dr_j$, and for all $dr_j = 0$, we find that all $ds_i = 0$, i.e., when the edge lengths of the development R are stationary, so are the edge lengths of the development S. Therefore, in the sequel, we exclude the development R from consideration and deal only with the natural development S.

In order to prove the stationarity of vertices of the polyhedron P, we use the following general lemma.

Lemma. *Let some polyhedron Q (not necessarily closed and convex) composed of triangular faces deforms in such a way that* (1) *its faces do not flex*; (2) *the edge lengths are stationary;* (3) *the angles between the faces abutting on whole edges are also stationary. Then all vertices of the polyhedron Q are stationary if motion is excluded.*

In order to exclude motion, we take some face ABC and fix the location of the vertex A, the direction of the edge AB, and the position of the face ABC. Then the motion of the polyhedron Q as a solid becomes impossible.

Since the vertex A is fixed and the direction of the edge AB is constant, the vertex B can move only in the direction of this edge. If its speed is nonzero in this process, then the length of the edge AB varies with a nonzero speed. However, this length is stationary, and hence the vertex B is stationary.

Since the plane BC is fixed, the component of the speed of motion of the vertex C which is perpendicular to this plane is zero. The lengths of the edges AC and BC and the vertices A and B are stationary. Therefore, the components of the speed of the vertex C in the directions of these edges are equal to zero. Hence the vertex C is stationary.

Take some face that is adjacent to the face ABC along an edge. Let BCD be this face. Since the angle between the faces ABC and BCD is stationary and since the vertices B and C are stationary, the plane of the face BCD is stationary. Therefore, the component of the velocity of the vertex D, which is perpendicular to this plane, is equal to zero. Further, the lengths of the edges BD and CD and the vertices B and C are stationary. Therefore, the components of the velocity of the vertex D in the direction of these edges are also equal to zero, and hence, the vertex D is also stationary. Passing from one face to the adjacent faces in this way and repeating the same arguments, we see that all vertices of the polyhedron Q are stationary; the lemma is proved.

© 2006 by Taylor & Francis Group, LLC

We can now pass directly to the proof of the rigidity theorem. By the above remark, we can assume that the polyhedron P consists of triangular faces that do not flex under deformation and all edge lengths of the polyhedron P are stationary. We prove the rigidity theorem by induction on the number of vertices of the polyhedron.

The theorem is obvious for a tetrahedron. Indeed, let $ABCD$ be a tetrahedron that does not degenerate into a quadrangle. If motion is excluded by fixing some vertex A, the direction of the edge AB, and the plane of the face ABC, then repeating the arguments of the proof of the lemma, we see that the vertices B and C are stationary.

Since the vertices A, B, and C, and the lengths of the edges AD, BD, and CD are stationary, the components of the velocity of the vertex D in the direction of these edges must be equal to zero. The tetrahedron $ABCD$ is nondegenerate, and, therefore, the directions of these edges are not coplanar; hence the speed of the edge D is equal to zero, so that all vertices of the tetrahedron prove stationary.

Assume now that the rigidity theorem holds for all nondegenerate closed convex polyhedra such that the number of their vertices is less by 1 than the number of vertices of the polyhedron P; of course, we assume that the number of vertices of the polyhedron P is greater than 4. However, assume that the rigidity theorem fails for the polyhedron P itself, i.e., although its edge lengths are stationary and motion is excluded, its vertices are not stationary. If all dihedral angles of the polyhedron P are stationary, then the above lemma implies the stationarity of the vertices. Therefore, we have to assume that not all dihedral angles of the polyhedron P are stationary.

Following the rule introduced in the previous section, we mark each edges with the plus or minus sign depending on the sign of the speed of the varying dihedral angle at this edge; if this speed is zero, then the edge remains undistinguished.

First, we assume that at least one of the distinguished edges goes to each vertex of the polyhedron P. Then, according to Lemma 3 of the previous section, when going around each vertex of the polyhedron P, we have either no less than four sign changes or only two sign changes, and in this last case, the plus sign stands only on two neighboring old edges, while the minus sign stands only on new edges that are located between them. Hence the undistinguished edges go to the vertex at which there are only two sign changes. If we mark all undistinguished edges by the minus sign, then the number of sign changes around this vertex is not less than four. Obviously, the number of sign changes around vertices at which there are not less than four changes does not decrease. Hence we obtain a location of signs on edges of the polyhedron such that (1) all edges are distinguished and (2) there are not less than four sign changes around each vertex.

If the number of vertices of the polyhedron is equal to e, the total number N of sign changes is not less than $4e$, i.e.,

$$N \geq 4e. \tag{1}$$

The same number of sign changes can also be estimated in another way.

Namely, we count the number of sign changes traversing the faces. If two edges p and q incident to one vertex turn out neighboring in passage around this vertex, then they are also neighboring on the face and vice versa. Therefore, the number

© 2006 by Taylor & Francis Group, LLC

of sign changes in traversing all faces of the polyhedron is the same number N. By condition, all faces of the polyhedron P are triangular. Therefore, there cannot be two sign changes in traversing one face, and if the total number of faces is equal to f then

$$N \leq 2f. \tag{2}$$

If k is the number of edges, then $3f = 2k$, since each face has three edges and each edge is common for two faces. By the Euler theorem, $f-k+e = 2$. Multiplying this equation by 2 and substituting $2k = 3f$, we obtain $f = 2e - 4$ or

$$2f = 4e - 8. \tag{3}$$

Therefore, instead of (2), we can write

$$N \leq 4e - 8.$$

But this inequality contradicts inequality (1). Therefore, our assumption that at least one distinguished edge goes to each vertex of the polyhedron is false.

It remains to assume that no distinguished edges go to at least one vertex O. This means that all dihedral angles meeting at this vertex are stationary.

The faces contiguous to the vertex O form some polyhedron P_0, and the other faces form some polyhedron P_1. Both polyhedra are not closed and their union yields the given polyhedron P, i.e.,

$$P = P_0 + P_1.$$

The edge lengths and the angles of the polyhedron P are stationary. Therefore, if we exclude motion, then by our lemma, all vertices of the polyhedron P_0 are stationary. If the polyhedron P itself has no other vertices but those belonging to P_0, then its rigidity is proved. Therefore, assume that the polyhedron P also has other vertices not belonging to the polyhedron P_0. Exclude the vertex O and arrange the convex hull of the other vertices of the polyhedron P; the boundary of this convex hull is a certain convex polyhedron P^*. This polyhedron is nondegenerate. Indeed, if this polyhedron is degenerate, then this means that all vertices of the polyhedron P but the vertex O lie in one plane, so that P is a pyramid with vertex O. But all vertices of a pyramid belong to its lateral faces, i.e., to the faces meeting at O; by assumption, not all vertices of the polyhedron P belong to these faces. Hence, the polyhedron P^* is nondegenerate.

Let Q be some face of the polyhedron P that does not contain the vertex O. Then all its vertices belong to the polyhedron P^*, and, consequently, to P^*. Therefore, the face Q is also the face of the polyhedron P^*. All these faces of the polyhedron P that do not contain the vertex O, form the polyhedron P_1, and hence P_1 is a part of the polyhedron P^*. Denote by P_0^* the other part of the polyhedron P^*, so that

$$P^* = P_0^* + P_1.$$

On the other hand,

$$P = P_0 + P_1,$$

© 2006 by Taylor & Francis Group, LLC

and all vertices of the polyhedron P that do not belong to P_0 lie inside P_1. Therefore, only the vertices belonging to P_0 can be the vertices of the part P_0^* of the polyhedron P^*. We assume that the part P_0^* of the polyhedron P^* consists of triangular faces.

We have shown above that all vertices of the polyhedron P_0 are stationary. Therefore, the distances between them are stationary, i.e., all edges of the polyhedron P^* that belong to its part P_0^* are stationary. All other edges of the polyhedron P^* belong to P_1, i.e., are edges of the polyhedron P, and so their lengths are stationary. Moreover, all vertices belonging to P_0^* are stationary, and so the motion of the polyhedron P^* is excluded at least at the initial moment. Further, the polyhedron P^* is nondegenerate, and the number of its vertices is less by 1 than the number of vertices of P. Therefore, according to the induction hypothesis, the rigidity theorem can be applied to this polyhedron, and all its vertices turn out stationary. But these vertices are all vertices of the polyhedron P except for the vertex O, which, however, is also stationary since it belongs to P_0. Hence, all vertices of the polyhedron P are stationary. The rigidity theorem is proved.

6. Realizability of the Metrics Close to Realized Metrics

In this and the following sections, unless otherwise stated, the word "metric" always means a polyhedral metric on a sphere of positive curvature with a given number of vertices $e > 3$. The metrics that can be realized only by degenerate polyhedra will play a special role. They are called degenerate metrics. All developments, unless specified otherwise, are assumed to lack superfluous vertices with complete angle equal to 2π.

The rigidity theorem implies the following lemma which serves as a basis for the proof of the existence of a polyhedron with given metric.

Lemma 1. *The metrics close to a realizable nondegenerate metric are realizable, i.e., if a closed convex polyhedron nondegenerate into a polygon originates by gluing from a development R_0, then we can also glue a closed convex polyhedron from each development R of the same structure whose edge lengths are close to the edge lengths of the development R_0.*

Proof. Let P_0 be a polyhedron that originates by gluing from the development R_0. Putting one its vertex at the origin, the second to the axis x, and the third to the plane $z = 0$, we thus fix six coordinates of vertices. If the polyhedron P_0 has e vertices then we are left with $3e - 6$ variable coordinates. If the number of edges of the development R_0 is equal to k, then

$$k = 3e - 6.$$

As shown in Sec. 1, if we deform the polyhedron by a small displacement of its vertices, then by Lemma 1 of Sec. 1 the resulting polyhedron admits the development of the same structure as the development R_0. Hence, we have k variable coordinates of vertices,

$$p_1, p_2, \ldots, p_k,$$

and k variable edge lengths

$$r_1, r_2, \ldots, r_k$$

© 2006 by Taylor & Francis Group, LLC

of the development. Let $p_1^0, \ldots, p_k^0$ and $r_1^0, \ldots, r_k^0$ be the values of these variables for the polyhedron P_0. The variables $p_1, \ldots, p_k$ can assume arbitrary values in a small neighborhood of the values $p_1^0, \ldots, p_k^0$, and by Lemma 3 of Sec. 3, the edge lengths of the development are single-valued and differentiable functions of them, that is,

$$\begin{aligned} r_1 &= f_1(p_1, \ldots, p_k), \\ &\ldots \\ r_k &= f_k(p_1, \ldots, p_k). \end{aligned} \tag{1}$$

Differentiating these equations, we obtain

$$\begin{aligned} dr_1 &= \frac{\partial f_1}{\partial p_1} dp_1 + \cdots + \frac{\partial f_1}{\partial p_k} dp_k, \\ &\ldots \\ dr_k &= \frac{\partial f_k}{\partial p_1} dp_1 + \cdots + \frac{\partial f_k}{\partial p_k} dp_k. \end{aligned} \tag{2}$$

By the rigidity theorem, we have $dp_1 = dp_2 = \ldots = dp_k = 0$ for $dr_1 = dr_2 = \ldots = dr_k$. In other words, the system of equations (2) for the differentials dp_i has no solutions but the trivial one if the free terms vanish. As is known, this implies that the determinant of this system differs from zero. This determinant is the Jacobian of the system of functions (1). At the same time, Eqs. (1) are satisfied for the values $p_1^0, \ldots, p_k^0$ and $r_1^0, \ldots, r_k^0$. Therefore, by the familiar implicit function theorem, Eqs. (1) can be solved with respect to the variables $p_1, \ldots, p_k$ in a neighborhood of the values $r_1^0, \ldots, r_k^0$ and $p_1^0, \ldots, p_k^0$, that is,

$$\begin{aligned} p_1 &= g_1(r_1, \ldots, r_k), \\ &\ldots \\ p_k &= g_k(r_1, \ldots, r_k). \end{aligned} \tag{3}$$

These equations mean that the edge lengths of the development can arbitrarily be varied in a neighborhood of their values $r_1^0, \ldots, r_k^0$, and, moreover, every time we obtain the corresponding values of the coordinates of vertices. These are the coordinates of the vertices of a certain convex polyhedron P whenever all p_i are close to p_i^0.[16] Since system (8) is inverse to system (1), the polyhedron P has the development of the given structure with edge lengths close to the edge lengths of the development R_0 of the given polyhedron P_0. Thus, the lemma is proved.

We now prove one more lemma which is necessary for the proof of the existence of a polyhedron with given metric.

Lemma 2. *Let metrics ρ_n that are realized by polyhedra P_n converge to some metric ρ. Then we can choose a sequence from the polyhedra P_n which converges to a polyhedron realizing the metric ρ.*

[16] This polyhedron is the convex hull of the vertices with coordinates p_i

© 2006 by Taylor & Francis Group, LLC

Proof. By the definition of convergence of metrics of Sec.1, the metrics ρ_n and ρ can be given by developments R_n and R of the same structure in such a way that the edge lengths of the developments R_n converge to the corresponding edge lengths of the development R.

In this lemma we consider polyhedra up to motion, since otherwise this lemma is not true: we could take polyhedra P_n tending to infinity. In this connection, we may assume that all polyhedra P_n pass through a given point. The distance between every two vertices of the polyhedron P_n is always not greater than the sum of the lengths of all edges of its development. Since the edge lengths converge, the sums of them are bounded. Hence, all polyhedra P_n are included into some ball.

We mark by the same indices the vertices of the development R_n which correspond to each other. Then the vertices of each of the polyhedra P_n turn out to be also enumerated. Choose a sequence from the polyhedra P_n such that the first vertices of this sequence converge; then we choose one more sequence from this sequence such that the second vertices of this sequence converge, and so on. As a result, we obtain a sequence in which all vertices converge. For brevity, by P_n we also denote the polyhedra of this sequence; let R_n stand for the corresponding developments. The convex polyhedron P, spanned by the limits of vertices (i.e., the boundary of the convex hull of the limits of vertices), is the limit of the polyhedra P_n. Let us prove that this polyhedron has the development R.

To this end, we first prove that if some edge of the polyhedron P_n coincides with none of the edges of the development R_n, then the number of its intersections with all edges never exceeds some N_3 that is common for all edges of all polyhedra P_n.

Let L be the least upper bound of the edge lengths of the polyhedra P_n, and let h be the greatest lower bound of the heights of the triangles of the developments R_n ($h > 0$ since the developments R_n converge to the development R of the same structure). Finally, let m be the maximal number of the angles contiguous to one vertex of the development R_m; this m is the same for all R_n, since the latter have the same structure. We claim that

$$N_0 = \frac{2Lm}{h} + m. \tag{4}$$

Assume the contrary. Then there is an edge a on some polyhedron P_n which has

$$N > \frac{2Lm}{h} + m \tag{5}$$

intersections with the edges of the development R_n. We have m segments of length $h/2$ in succession on the edge a between the intersection points. Otherwise, there is at least one segment of length $\le h/2$ for each m of such segments, and we totally obtain the length $\ge [N/m]h/2$ ($[N/m]$ stands for the integral part of N/m) which is greater than L by inequality (5); this contradicts the fact that the length of the edge a is no greater than L.

Assume that the first of the m segments of length $< h/2$ going in succession lies in the triangle ABC of the development R_n (Fig. 58). The endpoints of this segments lie at a distance less than a half of the edges AB and CD from the vertex A, so they are closer to the vertex A than to the vertices B and C. Otherwise, this segment XY is no less than a half of the height of the triangle ABC. To

© 2006 by Taylor & Francis Group, LLC

verify this, we drop the perpendicular YH to the side AB from the point Y; since $YH \le XY$, XY is less than a half of the height dropped from the vertex C if and only if $AY < (1/2)AC$; an analogous argument is true for the point X.

Leaving the triangle ABC, the edge a enters the neighboring triangle ADC where the segment of this edge is again less than $h/2$. Therefore, this edge is closer to A than to the other vertices in this triangle. Since we have m such segments in succession and not more than m angles are contiguous to the vertex A, the edge a goes around the vertex A, then returns to the side AB, and intersects this side again. However, this is impossible. Indeed, A is the vertex of the development R_n and, therefore, the vertex of the polyhedron P_n. Therefore, some edge b would emanate from this vertex and go to some other vertex of the polyhedron P_n. This edge would intersect the edge a, which is impossible since the edges of the polyhedra meet only at vertices.

This proves that the number of intersections of each edge of the polyhedron P_n with the edges of the development R_n does not exceed the number N_0 defined by formula (4). This makes it clear that, conversely, no edge of the development R_n can intersect the edges of the polyhedron P_n more than a definite number of times.

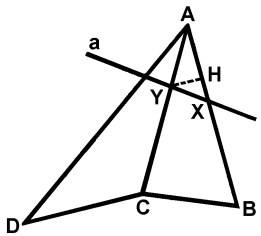


Fig. 58

We now enumerate the edges of each development R_n assigning the same index to the corresponding edges. Take the first edges of all these developments. If there are infinitely many edges coinciding with the edges of the polyhedron P_n among them, then we can choose a sequence P_{n_i} such that these edges converge. If infinitely many first edges of the developments only intersect the edges of the polyhedra P_n, then by the finiteness of the number of intersections, we can choose a sequence P_{n_i} in which these intersection points converge to some limit positions. The segments between neighboring intersection points of also converge, and hence the first edges of the developments R_n converge again. Further, applying this argument to the second edges, then to the third, and so on, we finally obtain a sequence of polyhedra P_{nn} in which all edges of the developments R_{nn} converge. The limits of these edges form a net of the same structure on the limit polyhedron and partition this polyhedron into triangles that do not contain the vertices of the polyhedra in their interiors. Thus, we have a certain development on P which has the same structure as the developments R_n and whose edge lengths are equal to the limits of the edge lengths of these developments. Thus, this is a required development.

7. Smooth Passage from a Given Metric to a Realizable Metric

According to the plan of Sec. 2, in order to prove the existence of a polyhedron with a given metric, it is necessary to show that if all metrics with $e-1$ vertices are realizable, then each metric ρ_0 with e vertices can be connected with a realizable metric ρ_1 by a continuous family of metrics ρ_t $(0 \le t \le 1)$ which contains no metrics realizable only by degenerate polyhedra for $t > 0$ (these metrics are called

© 2006 by Taylor & Francis Group, LLC

degenerate). We obtain this result not immediately, but only when we prove two lemmas that approximate it. In the sequel, unless specified otherwise, the term "metric" stands for a metric with e vertices ($e > 0$).

Lemma 1. *Each metric ρ_0 can be connected with a metric ρ_1 without one vertex by a continuous family ρ_t; in other words, the complete angle at one of the vertices of the metrics ρ_t tends to 2π as $t \to \infty$.*

Proof. First, we assume that there are two vertices among the vertices of the metric ρ_0 such that the sum of their curvatures is less than 2π. Let A_1 and A_2 be these vertices; the other vertices are denoted by $A_3, \dots, A_e$. Let ω_1 and ω_2 stand for the curvature of the vertices A_1 and A_2, respectively; by condition, we have $\omega_1 + \omega_2 < 2\pi$.

Draw a shortest arc A_1A_2 and make a cut along this arc; since the metric is given on the sphere, we speak about a cut of this sphere. Construct two equal triangles $A_1'A_2'A'$ and $A_1'', A_2''A''$ on the plane such that their bases are

$$A_1'A_2' = A_1''A_2'' = A_1A_2$$

and the angles at the vertices A_1', A_1'' and A_2', A_2'' equal $\omega_1/2$ and $\omega_2/2$, respectively. These triangles do exist, since $\omega_1 + \omega_2 < 2\pi$. Glue the bases of these triangles with both sides of the cut along the shortest arc A_1A_2 in such a way that the vertices A_1' and A_1'' coincide with A_1 and the vertices A_2' and A_2'' coincide with A_2. Furthermore, we glue together the lateral sides of our triangles, that is, we glue $A_1'A'$ with $A_1''A''$ and $A_2'A'$ with $A - 2''A''$. As a result, we have glued our cut A_1A_2, and thus we obtain some metric ρ_1 on the sphere.

The vertices A_1 and A_2 are absent in this metric. Indeed, when we glue the angles of our triangles to these vertices, we add the angle equal to $2 \cdot \omega_1/2 = \omega_1$ to the complete angle at A_1 and the angle $2 \cdot \omega_2/2 = \omega_2$ to the complete angle at A_2, i.e., we add precisely what they lack up to 2π. At the same time, the metric ρ_1 has a new vertex A that corresponds to the coinciding vertices A' and A'' of our triangles.

The above operation of passing from the initial metric ρ_0 to a metric ρ without one vertex can be realized in a continuous way. To this end, we take two variable points X' and X'' equidistant from A_2' and A_2'' on the sides $A_2'A'$ and $A_2''A''$ of our triangles. When the points X' and X'' run continuously along the sides $A_2'A'$ and $A_2''A''$ from the points A_2', A_2'' to the points A', A'', we glue the cut A_1A_2 by the triangles $A_1'A_2'X'$ and $A_1''A_2''X''$. Since the angles at the vertices A_2' and A_2'' of these triangles are equal to $\omega_2/2$, the vertex A_2 of the initial development disappear. The variable vertex X corresponding to the variable vertices X' and X'' replaces it. When these points run over the sides $A_2'A'$ and $A_2''A''$ from A_2' and A_2'' to A' and A'', the vertex X passes from A_2 to the vertex A of the finite metric ρ_1. In this way, we obtain a continuous family of metrics ρ_t which connects the initial metric ρ_0 with the metric ρ_1.

It is easy to verify that this family of metrics is in fact continuous in the sense of our definition of the convergence of metrics in Sec. 1. To this end, we draw shortest arcs $A_1X, A_1A_3, \dots, A_1A_e$ from the vertex A_1 that go to all other vertices of the metric ρ_t. Making cuts along these shortest arcs, instead of the sphere S we obtain a polygon Q that contains no vertices of the metric in its interior. When the vertex X varies continuously, the sides and the angles of this polygon

© 2006 by Taylor & Francis Group, LLC

vary continuously. The polygon Q can be divided into triangles by diagonals as was shown well back in Sec. 1. When the sides and angles of this polygon vary sufficiently slightly and continuously, this partition varies continuously. Hence we obtain a development whose continuous change realizes a sufficiently small change of the metric ρ_t. In accordance with our definition, this means that the family of metrics ρ_t is continuous.

We have assumed that there is a pair of vertices of the metric ρ_0 such that the sum of their curvatures is less than 2π. Assume now that there is no such pair, i.e., the sum of the curvatures of each two vertices is not less than 2π. In this case, the sum of curvatures of two pairs of vertices would be not less than 4π, i.e., not less than the curvature of the whole sphere. Hence there can be only four vertices, and the sum of the curvatures of each two of them must be equal to 2π, i.e., the curvature of each vertex is π.

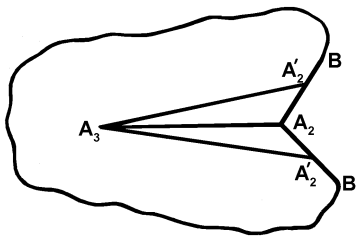


Fig. 59

Connect two vertices A_3 and A_2 by a shortest arc A_3A_2 and draw a shortest arc A_2B from the vertex A_2 such that these shortest arcs make equal angles with A_2A_3 to both its sides. These angles are equal to $\pi/2$, since the complete angle at A_2 is equal to π. Make a cut along A_2B and develop a neighborhood of the shortest arc A_3A_2 onto the plane (Fig. 59). This makes it clear that each point A_2' on A_2B sufficiently close to A_2 can be connected with A_3 by two equal geodesics going on the opposite sides of A_3A_2. Cut out both triangles $A-3A_2A'$ and glue both geodesics together. We obtain a new metric on the sphere such that the curvature of the vertex A_2' is less than π (since the complete angle at A_2' is greater than that at A_2 as is clear from the construction[17]). Therefore, the sum of curvatures of the vertex A_2' and, say, A_2 is less than 2π.

This operation of cutting two triangles is obviously inverse to the operation of gluing two triangles realized above. This operation can be realized in a continuous way by moving the point A_2' away from the vertex A_2. Then we obtain a continuous series of metrics that connects the initial metric ρ_0 with a metric that has two vertices such that the sum of their curvatures is less than 2π. By virtue of what we have said above, this latter metric can be connected with a metric ρ_1 whose number of vertices is less by one via a continuous series of metrics. Thus, we obtain a continuous series of metrics connecting the metric ρ_0 with this metric ρ_1. The lemma is proved.

Lemma 2. *If any metric with $e-1$ vertices is realizable, then any metric with e vertices can be connected with a nondegenerate realizable metric that also has e vertices by a continuous series of metrics.*

[17]See Fig. 59. The complete angle at A_2 is equal to the doubled angle at the vertex A_2 in the triangle $A_2A_2'A_3$, while the complete angle at A_2' is also equal to the doubled exterior angle of this triangle at the vertex A_2'. But the exterior angle is greater than the angle that is interior and nonadjacent to this angle; this implies our assertion.

© 2006 by Taylor & Francis Group, LLC

Proof. Let ρ_0 be a given metric with e vertices. By the preceding lemma, there exists a continuous series of metrics ρ_t that connects this metric with a metric ρ_1 having $e-1$ vertices. Assume that the metric ρ_1 is realizable by a polyhedron P_1, and let us prove the following two assertions in this case:

1. There are nondegenerate realizable metrics with e vertices that are arbitrarily close to the metric ρ_1.

2. The metric ρ_0 can be connected with one of these metrics by a continuous series of metrics.

To prove the first assertion, we consider in more detail the metric ρ_1 constructed in the previous lemma. This metric has the vertices $A_3, \ldots, A_e$ that correspond to the same vertices of the metric ρ_0; further, this metric has the vertex A which replaces the vertex A_2, and we can also isolate a point A_1 that corresponds to the vanishing vertex A_1 of the metric ρ_0. The points $A_3, \ldots, A_e, A$, and A_1 are points of the sphere S on which the metric ρ_1 is given. Draw shortest arcs AA_1, AA_3, $\ldots$, AA_e from the point A and make cuts along them; we thus transform the sphere S into a polygon Q. Dividing this polygon into triangles by diagonals, we obtain a development R that assigns the metric ρ_1. This development has one extra vertex A_1 that is not a vertex of the metric ρ_1.

Since the polygon Q is a result of cutting the sphere S along the shortest arcs AA_1, AA_3, $\ldots$, AA_e, its vertices are A, A_1, A, A_3,$\ldots$, A, A_e in the order of their location on its boundary. Thus, in particular, the "extra" vertex A_1 of the metric arises as a vertex of the polygon Q only one time. Therefore, among diagonals of the polygon Q, there are no diagonals that represent geodesics on the sphere S (in the sense of the metric ρ_1) connecting the vertex A_1 with itself. This implies that there are no triangles in the obtained development R_1 such that two their vertices arrive at the point A_1. This remark will be essentially used in what follows.

Since the polyhedron P_1 realizes the metric ρ_1, we can draw the development R_1 on this polyhedron. In other words, under an isometric mapping of the sphere S onto the polyhedron P_1, the development R_1 transforms into this polyhedron. All vertices of the development R_1 that are vertices of the metric are the vertices of the polyhedron P_1. One exception consists in the "extra" vertex A_1 of our development: since the complete angle around this vertex is equal to 2π, some point $\overline{A_1}$ corresponds to this vertex in the polygon P that lies inside its face or edge.

If we expose this point away from the polyhedron P_1, then constructing the convex hull of this displaced point and the vertices of the polyhedron P_1, we obtain a new convex polyhedron P. For a sufficiently small displacement of the point $\overline{A_1}$, the vertices of the polyhedron P_1 remain vertices and the point $\overline{A_1}$ also becomes a vertex. Hence the polyhedron P has e vertices. If the polyhedron P_1 degenerates into a polygon, then it is sufficient to take the point $\overline{A_1}$ out of its plane; then the polyhedron P is no longer degenerate.

If the displacement of the point $\overline{A_1}$ is sufficiently small, then according to what we have proved in Sec. 3, we can construct a development on the polyhedron P_1 that has the same structure as R_1 and the edge lengths close to the corresponding

© 2006 by Taylor & Francis Group, LLC

edge lengths of the development R_1.[18] Hence, we obtain a development R that is close to R_1; since all its vertices are vertices of the polyhedron P, the complete angle around each of them is $< 2\pi$. This means that *there are realizable metrics with e vertices that are arbitrarily close to the metric* ρ_1.

Now let us show that *the initial metric* ρ_0 *can be connected with one of such metrics.* To this end, we use the development R_1 again. Preserving its structure, we vary its edge lengths in some range. As a result, we obtain a set of developments R close to R_1. If we denote by k the number of edges of the development, then each development R is determined from k edge lengths $r_1, \ldots, r_k$. Therefore, the set of the so-obtained developments can be represented in the form of some domain G in the k-dimensional space with the coordinates $r_1, \ldots, r_k$. The relation between the developments R and the points of the domain R is one-to-one and continuous. Therefore, there is no difference for us whether we speak about the points of the domain G or the developments R.

Since the complete angle at the vertex A in the development R_1 is equal to 2π, this angle can become greater than 2π under an arbitrarily small change of edge lengths. Hence there are also developments R in the domain G, i.e., points corresponding to the developments R such that the complete angle at the vertex A_1 is greater than 2π. The continuous dependence of angle on edge lengths implies that the complete angles at other vertices remain less than 2π whenever the changes of the edges lengths are sufficiently small.

A development with the complete angle at the vertex A_1 will be called a development of the "second kind" in contrast to every development of the "first kind," whose complete angle at every point is less than 2π. The developments of the second kind as points are separated from the developments of the first kind as points by some hypersurface Φ defined by the equation

$$\sum_i \phi_i(r_1, \ldots, r_k) = 2\pi, \tag{1}$$

where $\phi_i(r_1, \ldots, r_k)$ stand for the angles of triangles of the development R which meet at the vertex A_1 and which are represented as functions of the edge lengths $r_1, \ldots, r_k$ (of course, each ϕ_i depends only on the three edges that form the sides of the corresponding triangle). The sum of all ϕ_i is thus the complete angle at the point A_1.

Take some triangle of the development R one of whose vertices approaches the point A_1. Let ϕ_1 be the angle of this triangle at the vertex A_1, and let r_1, r_2, and r_3 be its sides; moreover, assume that the side r_1 subtends the angle ϕ_1. We have mentioned above that there are no triangles in the development R whose two vertices arrive at the point A_1. Therefore, the edge r_1 of the development R is incident to the point A in no triangle.

By the generalized Pythagoras theorem,

$$r_1^2 = r_2^2 + r_3^2 - 2r_2r_3\cos\phi_1;$$

[18] In Sec. 3, we speak about developments all of whose vertices lie at the vertices of the polygon; this is not the case here, but it is easily seen that this does not violate the statement of Lemma 3 in Sec. 1 and its proof.

© 2006 by Taylor & Francis Group, LLC

which implies

$$\frac{\partial \phi_1}{\partial r_1} = \frac{r_1}{r_2 r_3 \sin \phi_1} > 0.$$

The edge r_1 can lie opposite the vertex A_1 in another triangle with the angle ϕ_1; then we also have $\partial \phi_1 / \partial r_1 > 0$. But this edge cannot be incident to the vertex A_1 anywhere. Therefore,

$$\frac{\partial}{\partial r_1} \sum_i \phi_i(r_1, \ldots, r_k) > 0.$$

By the familiar implicit function theorem, this implies that if Eq. (1) is satisfied for some values of edge lengths, then this equation is solvable in R_1 in a neighborhood of these values, and, moreover, r_1 turns out to be a single-valued differentiable function of the other edges. This means that the hypersurface Φ can be given by an equation of the form

$$r_1 = f(r_2, \ldots, r_k),$$

where f is a differentiable function. Hence the hypersurface Φ is smooth and divides the domain G into two parts. Since inequality (2) holds on the hypersurface Φ and moreover, $\sum_i \phi_i = 2\pi$ on this surface, we have

$$\sum_i \phi_i < 2\pi$$

to one side of this surface, and

$$\sum_i \phi_i > 2\pi$$

to the other side.

Points that lie to the same side of Φ can be connected with each other by a continuous curve not intersecting Φ. In terms of developments, this means that for every two developments of the first kind, i.e., those developments for which $\sum_i \phi_i < 2\pi$, there is a continuous family of similar developments connecting them under the obvious condition that they are taken sufficiently close to R_1. But among these developments, first, we have the development R that is realized by the polyhedron P and second, there are developments R_t determining those metrics ρ_t that comprise a continuous family connecting our initial metric ρ_0 with the metric ρ_1. Connecting one of such developments R_t with the development R, we obtain therefore a continuous family of metrics that connects the metric ρ_0 with the metric determined from the development R. Since this metric is realized by the nondegenerate polyhedron P, this completes the proof of our lemma.

The above result is not sufficient yet, since there can be degenerate metrics in the family of metrics ρ_t connecting ρ_0 with a realizable metric. We can assume that the metric ρ_0 is nondegenerate. Indeed, our final goal is to prove the realizability of the metric ρ_0. If this metric is degenerate, then, by definition, it can be realized by a degenerate polyhedron; therefore, there is nothing left to prove here. Therefore, we prove the last and final lemma on the possibility of connecting a given metric with a realizable metric.

© 2006 by Taylor & Francis Group, LLC

Lemma 3. *If some metric with $e-1$ vertices is realizable, then each nondegenerate metric ρ_0 with e vertices can be connected with a realizable metric that also has e vertices by a continuous family of metrics which contains no degenerate metrics.*

To prove this, we first consider degenerate metrics. Let a degenerate metric ρ be realized by a polyhedron P that is a doubly-covered convex polygon with vertices $A_1, A_2, \ldots, A_3$, in the order of their location on the boundary of the polygon (Fig. 60). Triangulate this polygon by diagonals from the vertex A_1. If each diagonal is assumed to be drawn on both sides of the polygon, then we obtain a development R of the polygon P which is naturally called a symmetric development. This development has the following properties.

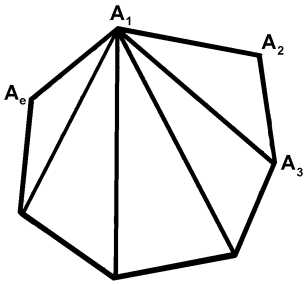


Fig. 60

1. The development R consists of pairwise equal triangles glued with each other along the edges $A_1A_2, A_2A_3, \ldots, A_eA_1$. These edges are naturally called exterior edges; their number is equal to the number of vertices e. These edges are unique shortest arcs that connect A_1 with A_2, A_2 with A_3, and so on, consecutively.

2. Since the total number of edges $k = 3e-6$, we are left with $2e-6$ edges that fall into two sets each of which contains $e-3$ edges. The edges of one of these sets lie to one side of the polygon P, and the edges of the other set lie to the other side of this polygon. With each edge A_1A_i of one set, we associate an equal edge of the other which connects the vertex A_1 with the same vertex A_i, except for the vertices A_2 and A_e. These edges are naturally called interior edges.

Hence, we have the following $e-3$ equations for interior edges:

$$r_1 = r_2, r_3 = r_4, \ldots, r_{2e-7} = r_{2e-6}. \tag{3}$$

Thus, each degenerate metric can be given by a symmetric development that satisfies the above two conditions. Conversely, each of these developments determines a degenerate metric. To verify this, we just glue two equal polygons from its triangles.

Let us prove that *if we violate at least one of the Eqs.* (3) *between the edges $A_1A_3, \ldots, A_1A_{e-1}$ of a degenerate metric ρ by a sufficiently small change, then we obtain a nondegenerate metric.* To this end, it is sufficient to show that if a sequence of degenerate metrics ρ_n converges to ρ, then for a sufficiently large n the corresponding edges of the metrics ρ_n are connected by the same Eqs. (3). By the definition of convergence of metrics, the metrics ρ_n are given by developments that have the same structure as a certain development of the metric ρ; therefore, we can speak about the corresponding vertices A_i of the metrics ρ_b and hence about the corresponding edges A_1A_i.

Since the metrics ρ_n are degenerate, they are realized by some polyhedra P_n. If the vertices $A_1, \ldots, A_n$ are located on the boundaries of these polyhedra in the

© 2006 by Taylor & Francis Group, LLC

same order as on the boundary of the polyhedron P that realizes the metric ρ, then Eqs. (3) connect the same edges in these metrics.

Assume, however, that for an arbitrarily large n the vertices $A_1, \dots A_e$ are located in another order on the boundaries of the polyhedra P_n. Then, choosing a sequence of polyhedra P_{ni} with the same order of vertices from the polyhedra P_n, we obtain a degenerate polyhedron $\overline{P}$ with the same order of vertices in the limit. Since this order is different from that of the polyhedron P, there are two vertices A_i and A_{i+1} on $\overline{P}$ that can be connected by two shortest arcs. At the same time, by the lemma on the realizability of the limit of realizable metrics (Lemma 2 in Sec. 6), the polyhedron $\overline{P}$ realizes the metric ρ, and, therefore, there is only one shortest arc A_iA_{i+1} on this polyhedron. This contradiction shows that for a sufficiently large n the vertices $A_1, \dots, A_e$ are located on the boundaries of the polyhedra P_n in the same order, and hence the same edges are connected by Eqs. (3).

We can now pass to the proof of Lemma 3 directly. Let ρ_0 be a given nondegenerate metric. We have to show that this metric can be connected with a nondegenerate realizable metric by a continuous family of nondegenerate metrics. By Lemma 2, the metric ρ_0 can be connected with a nondegenerate realizable metric by a continuous family of metrics ρ_t. If there are no degenerate metrics among these metrics, the required result follows.

Assume that there are degenerate metrics among the metrics ρ_t. Let T be the greatest lower bound of those t for which the metric ρ_t is degenerate. Then the metric ρ_T is also degenerate, since (this follows from Lemma 2 of Sec. 6) the limit of degenerate metrics is a degenerate metric. Since the metric ρ_0 is nondegenerate, $T > 0$. Determine the metric ρ_T by a symmetric development. Realize the metric ρ_T by a degenerate polyhedron P. If we displace one of the vertices of this polyhedron away from the plane of this polyhedron, we obtain a nondegenerate polyhedron. According to what we have proved in Sec. 3, this polyhedron has the metric close to ρ_T. Hence *there are nondegenerate realizable metrics arbitrarily close to the metric* ρ_T. Let us show that the metric ρ_0 can be connected with one of these metrics.

Determine the metric ρ_T from a symmetric development. There are $e-3$ pairs of equal edges

$$r_1 = r_2, \dots, r_{2e-7} = r_{2e-6}$$

in such a development. All metrics close to ρ_T can be given by developments of the same structure. Then each metric ρ close to ρ_T is determined from the $k = 3e-6$ edges $r_1, \dots, r_k$ and can be represented as a point in some domain G of the k-dimensional space with coordinates $r_1, \dots, r_k$. We have just proved that if at least one of the Eqs. (3) is violated by a small change of the metric, then we obtain a nondegenerate metric. Therefore, all degenerate metrics close to the metric ρ_T correspond to those points of the domain G for which Eqs. (3) hold. These $e-3$ equations define a linear subspace L of dimension $k-(e-3)$ in the k-dimensional space.

Here we have to distinguish the following two cases:

1. $e > 4$ and, therefore, $k-(e-3) < k-1$;

2. $e = 4$ and, therefore, $k-(e-3) = k-1$.

© 2006 by Taylor & Francis Group, LLC

In the first case, the linear subspace L has dimension less than $k-1$ and, therefore, does not split the k-dimensional domain G. Given every two points in a neighborhood of the subspace G not belonging to this subspace, we can join them with a continuous curve not intersecting L. This means that every two nondegenerate metrics can be connected by a continuous family of nondegenerate metrics. But we have proved that there are nondegenerate realizable metrics in a neighborhood of the metric ρ_T. Connecting one of these metrics ρ' with a certain metric in ρ_t $(t < T)$ avoiding all degenerate metrics, we obtain a continuous family of metrics that connects the initial metric ρ_0 with the metric ρ'. This family contains no degenerate metrics, since no metric ρ_t degenerates for $t < T$.

Now consider the second case in which $e = 4$ and the dimension of the subspace L is less by one than the dimension of the domain G. In this case, we have only one equation in (3), that is $r_1 = r_2$; the subspace L is a $(k-1)$-dimensional plane and, thus, divides the domain G into two parts. In fact, $r_1 = r_2$ on L itself, $r_1 > r_2$ to one side of L, and $r_1 < r_2$ to the other one. Since the development defining the metric ρ_t is symmetric, both interior edges play the same role. Therefore, if we change the notation, then the inequality $r_1 > r_2$ passes to the inequality $r_1 < r_2$. Therefore, to both sides of L, there are points that correspond to the same metrics (the distinction of these metrics lies only in the various ways of their representation by developments of the given structure). And since there are metrics ρ' in the neighborhood of the metric ρ_T that are realized by nondegenerate polyhedra, these metrics are available on both sides of L and, in particular, on that side from which the metrics for $t < T$ approach ρ_t. Hence one of these metrics ρ_t can be connected not intersecting L with one of the metrics ρ'. Then the initial metric ρ_0 turns out to be connected with such a metric ρ' by a family without degenerate metrics. The lemma is proved.

8. Proof of the Realizability Theorem

We are now ready to complete easily the proof of the theorem on the existence of a closed convex polyhedron with given metric.

In Sec. 1, we have proved the realizability of metrics with three edges. Therefore, we can prove the theorem by induction on the number of vertices. Assuming that the theorem is true for all metrics with $e-1$ vertices $(e > 3)$, we prove this theorem for metrics with e vertices.

Let ρ_0 be a given metric with e vertices. If the metric ρ_0 is degenerate, i.e., if it is realized by a degenerate polyhedron, then there is nothing to prove. Therefore, we can assume that the metric ρ_0 is nondegenerate. Then by Lemma 3 of the preceding section, the metric ρ_0 can be connected with a nondegenerate realizable metric ρ_1 by a continuous family of nondegenerate metrics ρ_t. By Lemma 1 of Sec. 6, the metrics close to ρ_1 are realizable, and in particular all metrics ρ_t are realizable for t sufficiently close to 1.

Let T be the greatest lower bound of those t for which the metrics ρ_t are realizable. Then, since the limit of realizable metrics is a realizable metric by Lemma 2 of Sec. 6, the metric ρ_T is realizable. Therefore, if $T = 0$ then the theorem is proved. But T cannot be other than zero. Indeed, the metric ρ_T is realizable and nondegenerate, since there are no degenerate metrics among the metrics ρ_t. Hence,

© 2006 by Taylor & Francis Group, LLC

all metrics that are close to ρ_T are also realizable according to Lemma 1 of Sec. 6. Therefore, were $T > 0$, all metrics ρ_t would be realizable for t close to T and less than T. But then T would not be the greatest lower bound of those t for which the metrics ρ_t are realizable. Consequently, $T = 0$; the theorem is proved.

© 2006 by Taylor & Francis Group, LLC

Chapter VII

EXISTENCE OF A CLOSED CONVEX SURFACE WITH A GIVEN METRIC

1. The Result and the Method of Proof

When all main properties of the intrinsic metric of a convex surface are studied and the existence of a closed convex polyhedron with a given polyhedral metric of positive curvature is proved, we can pass to the problem of the conditions that are not only necessary but sufficient for a given metric space to be isometric to a closed convex surface. This yields a possibility of solving the same problem for other types of convex surfaces in the next chapters. The following two necessary conditions were already deduced in Sec. 6 of Chapter I: *a metric space must be homeomorphic to the sphere, and its metric must be intrinsic* (in a more general case, we can speak about a manifold with an intrinsic metric). The simplest way is to take the convexity condition as the third necessary condition; however, as is explained in Sec. 9 of Chapter I, it is too strong to be considered as a successfully chosen sufficient condition. In Sec. 9 of Chapter I, we have indicated the condition which is much weaker in form (although it is certainly equivalent to the convexity condition, since it can replace the latter). We now recall this condition and the definitions of notions it involves but changing slightly the definition of "lower angle" given in Sec. 9 of Chapter I since this will considerably simplify the proof of sufficiency of our condition.

Let L and M be two shortest arcs emanating from a common point O on a manifold with intrinsic metric ρ. Take two variable points X and Y on these shortest arcs such that they are different from O and there are shortest arcs XY connecting these points; these shortest arcs always exist for all X and Y sufficiently close to O. As usual, we denote $x = \rho(OX), y = \rho(OY)$, and $z = \rho(XY)$; let $\gamma(x, y)$ be the angle lying opposite the side z in the plane triangle with the sides x, y, and z. For the shortest arcs L and M, we define the number $\tau(L, M)$ as follows. Let points X_n, Y_n comprise a sequence such that (1) $x_n y_n \to 0$ (i.e., at least one of the points X_n and Y_n approaches the point O); (2) the points X_n and Y_n can be connected by shortest arcs $X_n Y_n$ whose limit points as $n \to \infty$ lie on the line $L + M$ composed of the shortest arcs L and M, i.e., if, e.g., $x_{n_i} \to 0$ then the shortest arcs $X_{n_i} Y_{n_i}$ approach M. For such sequence of points $X_n Y_n$, we define the lower limit of the angle $\gamma(x_n, y_n)$. We denote by $\tau(L, M)$ the least of these limits; we can write

$$\tau(L, M) = \liminf_{xy \to 0} \gamma(x, y)$$

provided that the shortest arcs XY indefinitely approach the line $L + M$. This number $\tau(L, M)$ was called the lower angle between the shortest arcs L and M in Sec. 9 of Chapter I.

© 2006 by Taylor & Francis Group, LLC

There is nothing new in this number for us, since we have proved in Sec. 4 of Chapter III (Theorem 5) that there is an angle "in the strong sense" between shortest arcs emanating from a common point on a convex surface, i.e., under the condition imposed on the shortest arcs XY, $\lim_{xy\to 0}\gamma(x,y)$ exists and is equal to the angle between the shortest arcs L and M. This theorem ensues from the convexity condition in a sufficiently simple manner. Therefore, if we take $\tau(L,M)$ as the lower angle made by the shortest arcs L and M, then we see that the necessity of the condition that the sum of the lower angles of a triangle is no less than π is sufficiently easy to obtain.

However, it turns out inconvenient to take the number $\tau(L,M)$ itself as the lower angle made by shortest arcs for the following reason. Take two segments L_1 and L_2 of the shortest arcs L and M that also emanate from the point O; then in the process of determining the number $\tau(L_1,M_2)$, the number of admissible locations of the points X and Y decreases, and so the lower limit of angles $\gamma(x,y)$ might increase implying that $\tau(L_1,M_1)>\tau(L,M)$.

Theorem 5 in Sec. 4 of Chapter III mentioned above asserts that this cannot happen on a convex surface; but this is possible on an arbitrary manifold with an intrinsic metric, as can be verified by examining examples. Therefore, we define the lower angle between shortest arcs L and M as the magnitude that is obtained by the following rule.

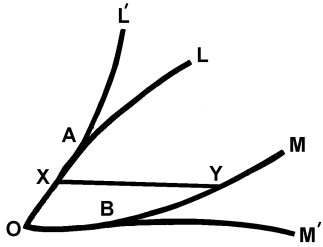


Fig. 61

Consider all those shortest arcs L' and M' emanating from O which coincide with L and M on some segments OA and OB (these segments are different for different shortest arcs in general). (Under the overlapping condition, the shortest arc L contains L', or vice versa; but without this condition, the shortest arc L' can deviate from L at some interior point; see Fig. 61.) *The greatest lower bound of the numbers $\tau(L',M')$ over all such shortest arcs L' and M' is defined to be the lower angle between the shortest arcs L and M.* For this definition, the lower angle is the same for all shortest arcs emanating from O and coinciding with L and M on some segments starting from the point O.[1]

The condition by which we want to characterize the intrinsic metric of convex surfaces consists in the following.

For each point, there is a neighborhood such that the sum of lower angles between the sides of every triangle lying in this neighborhood (these sides are shortest arcs) is not less than π. A metric of an arbitrary metric space is called a *metric of positive curvature* if this metric is intrinsic and satisfies the above condition.

[1] However, in this definition of lower angle, an essential role is played as before not only by the metric of an arbitrarily small neighborhood of the point O but by the metric of the whole domain which attains the shortest arcs L' and M'. In a remark of Sec. 9 in Chapter I, we have indicated that it is possible to introduce another notion of lower angle which is free of this deficiency. However, it was mentioned there that the new definition also has drawbacks that are more essential; but the main point is that we do use the above-introduced notion of lower angle in proofs. Moreover, we can restrict exposition to some neighborhood of the point O in the definition of lower angle.

© 2006 by Taylor & Francis Group, LLC

By a triangle in an arbitrary metric space we can mean only a figure that is formed by three shortest arcs pairwise connecting three given points. However, we deal not with arbitrary spaces but with two-dimensional manifolds. In such a manifold, three shortest arcs lying in a sufficiently small domain bound a closed set. It is this set that we will be call a triangle. If this set is convex, then we speak about a convex triangle.

In the case of a two-dimensional manifold, it turns out advantageous to slightly weaken the definition of a metric of positive curvature. Namely, it suffices to require that each point has a neighborhood such that the sum of angles of every *convex* triangle lying in this neighborhood is not less than π. In the sequel, we will use this weakened definition.

Each convex surface has a metric of positive curvature, since the lower angle on a convex surface reduces to the ordinary and the sum of the angles between sides of any triangle on a convex surface is not less than π. Since these results were obtained in Sec. 4 of Chapter III on the basis of the convexity condition, we can also formulate the following theorem: *each intrinsic metric satisfying the convexity condition is a metric of positive curvature.* Hence, when we prove the sufficiency of the requirement that a metric has positive curvature, we thus prove the sufficiency of the convexity condition.

We have shown in Sec. 9 of Chapter I via a simple example of a "doubled sphere" that the requirement of positive curvature itself is not sufficient yet for the spherical domain whose metric satisfies this requirement to be isometric to some convex surface. The most general theorem on the existence of a convex surface with a given metric of positive curvature, which we can prove, was formulated as Theorem A in Sec. 9 of Chapter I. We restrict exposition to three particular cases of this theorem, which are also formulated in Sec. 8 of Chapter I as Theorems A, B, and C.

Theorem 1. *A metric space whose metric is a metric of positive curvature and which is homeomerphic to a sphere is isometric to a convex closed surface.*

Theorem 2. *A metric space with a complete metric of positive curvature which is homeomorphic to a plane is isometric to an infinite complete surface.* (The necessity of the requirement that a metric is complete was proved in Sec. 9 of Chapter I.)

Theorem 3. *In each manifold with metric of positive curvature, each point has a neighborhood isometric to a convex surface.*

This chapter is devoted to the proof of Theorem 1; when we will have proved this theorem, Theorems 2 and 3 become easy to prove on using Theorem 1; this will be done in the next chapter.

The proof of Theorem 1 is based on the approximation of a general metric of positive curvature on the sphere by a polyhedral metric of positive curvature. This proof falls into three stages: first, we study the general properties of a metric of positive curvature and, using these properties, we try to prove that it is possible to approximate a general metric of positive curvature by a polyhedral metric. The second stage consists just in the proof of this possibility. The third stage is the proof of the theorem itself using the results of the second stage and the theorem on

© 2006 by Taylor & Francis Group, LLC

the existence of a polyhedron with a given metric which was proved in the preceding chapter.

Consider each of the stages in more detail.

The first stage of the proof is carried out in Secs. 2–4. We consider an arbitrary manifold with a metric of positive curvature or, in other words, a metric space that is a two-dimensional manifold having a metric of positive curvature. All considerations are "in the small," i.e., we consider only a domain U of this manifold which satisfies the following conditions:

1. U is homeomorphic to a disk;
2. there is a shortest arc of every two points of the domain U;
3. the sum of the shortest angles between sides of each triangle formed by the shortest arc of pairs of points of the domain U is not less than π.

Each point of a manifold with metric of positive curvature has a neighborhood in which all three conditions hold. Indeed, by the positivity of curvature, each point O has a neighborhood V in which the third condition holds. At the same time, we have proved (Theorem 4 in Sec. 2 of Chapter II) that in each neighborhood V of a point O in a manifold with intrinsic metric we can find a neighborhood U such that every two points in U are connected by a shortest arc, and each of these shortest arcs lies in the neighborhood V. Of course, this neighborhood U can be chosen homeomorphic to a disk. Then this neighborhood satisfies all three conditions mentioned above.

In what follows, the term "sufficiently small element" implies something that lies in such a neighborhood.

Three shortest arcs in a domain U, connecting some points A, B, and C, bound a closed set in U which will be called a *triangle* ABC. Such a triangle may not *a priori* be homeomorphic to a disk since the shortest arcs AB, BC, and CA can have other common points, together with the vertices A, B, and C, even in the case where they do not overlap. In fact, this is not possible in a manifold with a metric of positive curvature, and each sufficiently small triangle is homeomorphic to a disk if this triangle does not degenerate into a single shortest arc. But until this remains unproven, by a sufficiently small triangle we must mean merely a closed set bounded by three shortest arcs AB, BC, and CA.

The main object of this study comprises sufficiently small convex triangles. In Sec. 2, we prove the "main lemma on (sufficiently small) convex triangles" in a manifold with a metric of positive curvature. The proof of this lemma is the only essential difficulty of the first stage of our arguments. When we prove this lemma, the further facts will be obtained in a comparatively simple way. This lemma is not significant in its own right (so we do not even formulate it here), but it easily implies the important consequences of Sec. 3 (in what follows, it is assumed that we speak about a manifold with metric of positive curvature); namely,

Theorem A. *There exists an angle in the usual sense between sides of a convex triangle.* (This angle is equal to the least angle between the same sides.)

Theorem B. *Each angle of a sufficiently small convex triangle is not less than the corresponding angle of the plane triangle with sides of the same length.*

© 2006 by Taylor & Francis Group, LLC

Theorem C. *If two shortest arcs emanating from a common point overlap each other on some segment, then one of these arcs is a part of the other, i.e., the "nonoverlapping condition for shortest arcs" holds.*

One more theorem accompanies Theorems A, B, and C.

Theorem D. *Let shortest arcs $L_1, \dots, L_n$ emanating from a point O and enumerated in the order of their location around the point O partition a neighborhood of the point O into convex sectors. The sum of the angles $\alpha_{i,i+1}$ between neighboring shortest arcs L_i and L_{i+1} (and also between L_n and L_1) is always the same for the given point O and does not exceed 2π irrespective of the choice of the shortest arcs L_i.*

Since the sum of the angles $\alpha_{i,i+1}$ at the point O in Theorem D depends only on the point O itself, this sum can be called the complete angle at this point. The second assertion of Theorem D can thus be formulated as follows: *the complete angle at a point in a manifold with metric of positive curvature does not exceed 2π.*

Theorem C allows us to apply all results deduced from the nonoverlapping condition for shortest arcs in Chapter III to a manifold with a metric of positive curvature. The following is the most important among them.

Theorem E. *A convex triangle can be partitioned into arbitrarily small triangles.*

Theorems A–D allow us to extend all properties of the angle proved in Chapter IV to angles between sides of convex triangles. Furthermore, we can introduce the concept of curvature in a manifold with a metric of positive curvature in exactly the same way as in Sec. 1 of Chapter V with the only difference that, instead of arbitrary triangles, we will consider only sufficiently small convex triangles. The curvature of the interior of a convex polygon is defined on partitioning it into convex triangles, and, by the theorem proved in Sec. 1 of Chapter V, this curvature is expressed by the formula

$$\omega = 2\pi\chi - \sum_{i=1}^{n}(\pi - \alpha_i),$$

where χ is the Euler characteristic and α_i are the angles of the polygon (i.e., the angles between its neighboring sides). In particular, we have $\omega = 4\pi$ for the whole sphere with a metric of positive curvature.

This completes the first stage of the proof of Theorem 1. Further, in Sec. 5, we prove some auxiliary estimates that are used in Sec. 6 proving the possibility of approximation of a metric of positive curvature by polyhedral metrics. Let P be a convex polygon in a manifold with metric ρ of positive curvature. By Theorem E, this polygon can be partitioned into convex triangles T_i with sides not larger than some given d. To each triangle T_i, we put in correspondence the plane triangle T_i^0 with sides of the same length, and the triangle T_i^0 is mapped homeomorphically onto T_i in such a way that this mapping is an isometry (preserves length) if restricted to the sides. Then the triangles T_i^0 comprise a development P^0 composed of the triangles T_i^0 in the same way as the polygon P is composed of the triangles T_i. The mappings of the triangles T_i^0 onto T_i generate a homomorphism of the development P^0 onto P. This gives rise to some polyhedral metric ρ^0 inside the polygon P;

© 2006 by Taylor & Francis Group, LLC

i.e., if X and Y are two points in P, then $\rho^0(XY)$ is the distance between the corresponding points X^0 and Y^0 of the development P^0. We prove the following theorem about this polyhedral metric.

Theorem F. *The polyhedral metric $\rho^0(XY)$ is a metric of positive curvature. If the curvature $\omega < \pi$ for the polygon P, then for every two points X and Y in P we have*

$$|\rho(XY) - \rho^0(XY)| < Md,$$

where M is some constant and d is the greatest of the sides of all triangles T_i.

The fact that the metric ρ^0 is a metric of positive curvature is immediate from Theorems B and D. Indeed, by Theorem D, the sum of angles of the triangles T_i at each vertex does not exceed 2π, while by Theorem B the angles of the triangles T_i^0 are not greater than the angles of the triangles T_i. Hence the sum of angles of the triangles T_i^0 at each vertex does not exceed 2π; this means exactly that the polyhedral metric ρ^0 is a metric of positive curvature. The second part of Theorem F is proved by some elementary arguments that will be obtained in Sec. 5.

This completes the second stage of proving Theorem 1. We are left with the third stage, which is the proof of the theorem itself. Let R be a space with a metric of positive curvature which is homeomorphic to the sphere. According to Theorem E, we can construct a sequence of its partitions Z_n into convex triangles becoming smaller and smaller. Mapping each of these triangles onto the plane triangle with sides of the same length, to each partition Z_n we put in correspondence some development R_n that is composed of plane triangles. The inverse mapping from the development R_n onto the space R defines a polyhedral metric ρ_n of positive curvature in this space as we have just seen.

By the theorem proved in the preceding chapter, each of these metrics can be realized by a closed convex polyhedron P_n. This means that there exists a mapping of the space R onto the polyhedron P_n which preserves distances, i.e., $\rho_n(XY) = \rho_{P_n}\rho(\overline{X}_n\overline{Y}_n)$, where ρ_{P_n} is the distance on P_n and $\overline{X}_n, \overline{Y}_n$ are the images of the points X and Y in R.

If the polyhedron P_n is located in such a way that these points do not tend to infinity, then we can choose a sequence P_{n_i} such that all sequences of points $\overline{X}_{nn_i}$ that correspond to the same point X converge. Thus, with each point X in R, we associate a limit point $\overline{X}$ of the sequence $\overline{X}_{n_i}$. We prove that these points form some convex surface F to which the polyhedra P_{n_i} converge. Then, by the theorem on the convergence of metrics of convergent convex surfaces, we conclude that

$$\lim_{n\to\infty} \rho_{P_{n_i}}(\overline{X}_{n_i}\overline{Y}_{n_i}) = \rho_F(XY).$$

But since $\rho_{P_n}(\overline{X}_n\overline{Y}_n) = \rho_n(XY)$, we have

$$\rho_F(\overline{XY}) = \lim_{i\to\infty} \rho_{n_i}(XY). \tag{1}$$

On the other hand, using the second part of Theorem F, we can show that the polyhedral metrics ρ_{n_i} converge to the given metric ρ, i.e.,

$$\lim_{i\to\infty} \rho_{n_i}(XY) = \rho(XY). \tag{2}$$

© 2006 by Taylor & Francis Group, LLC

Comparing (1) and (2), we see that

$$\rho_F(\bar{X}\bar{Y}) = \rho(XY),$$

i.e., the distance between arbitrary points of the surface F is equal to the distance between the corresponding points of the space R. This means that the surface F is isometric to the space R; thus Theorem 1 is proved.

The way sketched is sufficiently long, and this is not surprising if we take into account that it leads us from an abstract metric space that satisfies a few general conditions to a rather specific object, a convex surface. The main difficulty is in the first stage, but this difficulty could be diminished if we assume *a priori* that our metric space satisfies stronger conditions, e.g., the convexity condition. As is shown in Sec. 4 of Chapter III, Theorems A, B, and C are direct consequences of this condition, and so it will remain to prove Theorem D. If we also postulate that the complete angle at each point does not exceed 2π then the whole first stage, together with the main difficulty of the proof, disappears. We have a right to proceed so since we know that the complete angle at every point on a convex surface in fact does not exceed 2π. Therefore, the reader who is not especially interested in keeping the number of postulates at a minimum can take Theorems A–D or the convexity condition or Theorem D as axioms and to begin the proof of Theorem 1 immediately from Sec. 5, skipping Secs. 2–4.

2. The Main Lemma on Convex Triangles

In this section, we will consider a sufficiently small convex triangle in a manifold with a metric of positive curvature. Such a triangle ABC was defined as a closed set bounded by three shortest arcs AB, BC, and CA; *a priori*, this triangle can have a rather bizarre shape. However, it is convenient for lucidity to exclude excessively "pathological" triangles from our consideration. This can be done in the following way.

Assume for example that the shortest arcs AB and AC have common points apart from A; then there is a point A' that is the most distant among these points from the point A. The segments AA' of both shortest arcs are equal, since otherwise, replacing one segment by a shorter one, we will shorten one of these shortest arcs. Therefore, if one of these segments is replaced with the other, then the shortest arcs AB and AC will coincide on the part AA', and have no other common points. The same operation can be performed for the other pairs of sides of the triangle. After that, we obtain a triangle ABC whose sides coincide up to certain points A', B', and C'. Of course, it can happen that this triangle degenerates into a single shortest arc. But this case can be excluded from consideration, since it is trivial.

If the sides AB and AC coincide on the segment AA', then the least angle between them is equal to zero. Since the sum of the angles of every triangle is not less than π, not all pairs of sides can overlap. This makes it clear that at least one of the points A', B', and C' coincides with A, B, and C, and in any case, the points A', B', and C' cannot coincide.

Hence the triangle ABC consists of the triangle $A'B'C'$ homeomorphic to a disk (since this triangle is bounded by three shortest arcs $A'B'$, $B'C'$, and $C'A'$ that have no common points but A', B', and C') such that segments that are prolongations of

© 2006 by Taylor & Francis Group, LLC

its sides can emanate from one or two of its vertices, e.g., the segment AA'. In the sequel, we may imagine this triangle exactly in this shape. In what follows, we will verify that the nonoverlapping condition for shortest arcs holds in every manifold of positive curvature, and, thus, a triangle cannot have such singularities in reality.

The Main Lemma. *Let ABC be a sufficiently small convex triangle in a manifold with the metric of positive curvature, and let $A_0B_0C_0$ be the plane triangle with sides of the same lengths, so that $A_0B_0 = AB$, etc. There are only the following two possibilities open for each pair of sides of the triangle ABC, e.g., for AB and AC: either the lower angle between them is not less than the corresponding angle of the triangle $A_0B_0C_0$ or these sides are prolongations of each other, so that some of their segments AB_1 and AC_1 together form a single shortest arc B_1C_1.* (In fact, the first case holds always, but we do not prove this fact.)

In proving this lemma, we will consider two variable points X and Y on the sides AB and AC of the triangle ABC and the shortest arcs XY of these points in the triangle ABC. Since the triangle ABC is convex, these shortest arcs do exist. For convenience, we assume that some directions from the vertex A to the vertices B and C are some directions from left to right, while the reverse directions are assumed to be some directions from right to left. For example, if a point X_1 is more distant from A than a point X_2, we say that this point lies to the right of this point. For a given location of the points X and Y, there can exists several or even infinitely many shortest arcs XY lying in the triangle ABC. As a preliminary, we will prove several lemmas about these shortest arcs.

Lemma 1. *There is the "leftmost" shortest arc among all shortest arcs XY of two given points X and Y on the sides of the convex triangle ABC which lies in this triangle, i.e., such that there are no other shortest arcs of the same points in the triangle AXY which is cut out from the triangle ABC by this shortest arc.*

Proof. We will consider only those shortest arcs XY that lie in the triangle ABC. If this shortest arc is unique, then it is "leftmost." Assume that there are two such shortest arcs XY and $\overline{XY}$. These shortest arcs cut out two triangles AXY and $A\overline{XY}$ from the triangle ABC. The intersection of these triangles is again a triangle bounded by some shortest arc $\overline{\overline{XY}}$ that coincides with one of the shortest arcs XY and $\overline{XY}$ or consists of their segments. Indeed, if XY and $\overline{XY}$ intersect at two points P and Q, then their segments PQ and $\overline{PQ}$ are equal. Therefore, the line composed of such segments is also a shortest arc. If the shortest arcs XY and $\overline{XY}$ intersect at infinitely many points, then we must argue rigorously as follows: each of the shortest arcs XY and $\overline{XY}$ can be mapped onto a line segment L with preservation of their lengths. The common part of both shortest arcs is a closed set; with this set, we associate some closed set M on this segment. The complement of this set is an open set, and hence it consists of countably many pairwise disjoint intervals. With each such interval (pq), we associate the segments PQ and $\overline{PQ}$ of the shortest arcs XY and $\overline{XY}$. We take a segment that is "leftmost," i.e., for example, the segment PQ if the latter lies in the triangle $A\overline{XY}$. Adjoining all such segments to the common part of the shortest arcs XY and $\overline{XY}$, we obtain some line $\overline{\overline{XY}}$. It is clear from the construction that, first, this line is mapped onto the

© 2006 by Taylor & Francis Group, LLC

same line segment L with preservation of lengths and, therefore, is a shortest arc; second, this line bounds the intersection of the triangles AXY and $A\overline{XY}$, i.e., lies "far left" than XY and $\overline{XY}$.

This makes it clear that we can always construct or choose the "leftmost" shortest arc from finitely many shortest joins of X and Y.

Assume now that there exist infinitely many shortest arcs of the points X and Y. We will say that a shortest arc XY lies far left than $\overline{XY}$ if this shortest arc lies in the triangle $A\overline{XY}$ that is cut out by the shortest arc $\overline{XY}$ and does not coincide with the latter. We have just proved that it is possible to choose or construct one of every two shortest arcs which lies far left than the other. Therefore, if some shortest arc $\overline{XY}$ is not leftmost, then there exist shortest arcs $\overline{XY}$ that lie far left than this arc. Each shortest arc of the points X and Y distinguishes a part of the triangle ABC complementary to the triangle $A\overline{XY}$. The interior of this part of the triangle ABC will be denoted by $G(\overline{XY})$. Take the union of all these sets $G(\overline{XY})$ for given X and Y. This is an open set G. In this set G, we take a countable everywhere dense set of points Z_i and consider the set of all circular neighborhoods with rational radii at each of these points. We obtain a countable set of neighborhoods. We choose from this set of neighborhoods only those each of which lies in some of the sets $G(\overline{XY})$. Enumerate these neighborhoods as follows: $U_1, U_2, \ldots$. Each set $G(\overline{XY})$ is the sum of the neighborhoods U_i contained in it. And since G is the union of all $G(\overline{XY})$, we have

$$G = \sum G(\overline{XY}) = \sum_{i=1}^{\infty} U_i.$$

By the choice of the neighborhoods U_i, each of them lies in some $G(\overline{XY})$. Therefore, among the sets $(\overline{XY})$, we can find a set that includes U_1, Denote by L_1 the corresponding shortest arc $\overline{XY}$. Further, if there are shortest arcs $\overline{XY}$ lying far left than L_1, then there is a shortest arc L_2 among them such that $G(L_2) \supset U_2$. Continuing this process of choosing shortest arcs L_n, we obtain a sequence of shortest arcs such that

$$G(L_n) \supset G(L_{n-1}) + U_n,$$

and, therefore,

$$\sum_{n=1}^{\infty} G(L_n) = G.$$

We can extract a convergent sequence of L_{n_i} from the shortest arcs L_n. The limit shortest arc L is the leftmost shortest arc among the shortest arcs XY. Otherwise, there would be a neighborhood U_m not lying in $G(L)$, so that for any $n_i > m$, the shortest arcs L_{n_i} and L bound the domain $G(L_{n_i} - G(L)$ that includes U_m; which is obviously impossible if the sequence of L_{n_i} converges to L. The lemma is proved.

Lemma 2. *If points X_n and Y_n converge to X and Y from the left, then the leftmost shortest arcs X_nY_n converge to the leftmost shortest arc XY.*

Proof. Assume that the leftmost shortest arcs X_nY_n do not converge to the leftmost shortest arc XY. Then we can extract a sequence $X_{n_i}Y_{n_i}$ from them which converges to some other shortest arc $\overline{XY}$, and for sufficiently large n_i, the shortest

© 2006 by Taylor & Francis Group, LLC

arcs $X_{n_i}Y_{n_i}$ intersect the leftmost shortest arc XY itself. For example, let $X_{n_k}Y_{n_k}$ intersect XY at two points P and Q so that the shortest arc XY goes far left than $X_{n_i}Y_{n_i}$ on the segment PQ (i.e., inside the triangle $AX_{n_i}Y_{n_i}$). Then, replacing the segment PQ of the shortest arc $X_{n_i}Y_{n_i}$ by a segment of the shortest arc XY, we obtain a shortest arc of the same points X_{n_i} and Y_{n_i} which lies far left than the shortest arc $X_{n_i}Y_{n_i}$. This contradicts the fact that the latter is the leftmost shortest arc from all shortest arcs of these points. The lemma is proved.

Lemma 3. *Denote by x, y, and $z = z(x, y)$ the distances AX, AY, and XY. The lower angles between the leftmost shortest arc XY and the segments AX and AY of the sides AB and AC are denoted by ξ and η. Furthermore, let $\partial z/\partial x$ and $\partial z/\partial y$ stand for the left and right upper derivatives.*[2] *These derivative are related to the angles ξ and η by the inequalities*

$$\cos\xi \geq \frac{\partial z}{\partial x} \geq -1, \quad \cos\eta \geq \frac{\partial z}{\partial y} \geq -1.$$

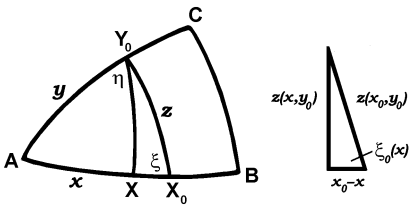


Fig. 62

Proof. Certainly, it is sufficient to consider one of these derivatives, for example, $\partial z/\partial x$. Therefore, we can choose and fix an arbitrary point Y_0.

By the triangle inequality,

$$|X_0Y_0 - XY_0| \leq X_0X,$$

i.e.,

$$|z(x_0, y_0) - x(x, y_0)| \leq |x - x_0|;$$

this obviously implies

$$\left|\frac{\partial z}{\partial x}\right| \leq 1.$$

Therefore, it remains to prove that $\partial z/\partial x \leq \cos\xi$. To this end, we take a point X lying to the left of the given point X_0 (Fig. 62) and construct the plane triangle with sides equal to

$$|x - x_0|, \quad z(x_0, y_0), \quad z(x, y_0).$$

The angle of this triangle opposite the third side is denoted by $\xi_0(x)$. By Lemma 2, the shortest arcs XY_0 converge to X_0Y_0 as $x \to x_0$. Therefore, from the definition of the lower angle between the shortest arcs X_0Y_0 and AX_0 it follows that

$$\lim \inf_{x \to x_0 - 0} \xi_0(x) \geq \xi. \tag{1}$$

At the same time, in the plane triangle, we have

$$z(x, y_0)^2 = z(x_0, y_0)^2 + (x - x_0)^2 - 2z(x_0, y_0)(x_0 - x)\cos\xi_0(x),$$

[2] That is, for example, $\frac{\partial z}{\partial x} = \limsup_{h \to -0} \frac{z(x+h,y)-z(x,y)}{h}$.

© 2006 by Taylor & Francis Group, LLC

implying

$$\frac{z(x_0y_0)-z(x,y_0)}{x_0-x}=\frac{2z(x_0,y_0)}{z(x_0,y_0)+z(x,y_0)}\cos\xi_0(x)-\frac{x_0-x}{z(x_0,y_0)+z(x,y_0)}.$$

But as $x\to x_0-0$, we also have $z(x,y_0)\to z(x_0,y_0)$, and so

$$\frac{\partial z(x_0,y_0)}{\partial x}=\limsup_{x\to x_0-0}\frac{z(x_0,y_0)-z(x,y_0)}{x_0-x}=\limsup_{x\to x_0-0}\cos\xi_0(x).$$

But by the inequality (1), we have

$$\limsup_{x\to x_0-0}\cos\xi_0(x)\le\cos\xi.$$

Consequently,

$$\frac{\partial z}{\partial x}\le\cos\xi;$$

as required.

Lemma 4. *Using the notation of Lemma* 3*, consider the plane triangle* $A_0X_0Y_0$ *with sides* x*,* y*, and* z*; as above, denote by* $\gamma(x,y)$ *the angle at the vertex* A_0*, and let* ξ_0 *and* η_0 *be the angles of this triangle at the vertices* X_0 *and* Y_0*, respectively* (*Fig.* 63)*. Assume that* $\xi>\xi_0$ *for given* x *and* y *and set* $\xi-\xi_0=\varepsilon$*. Then either* $\xi_0=0$ *so that the triangle* $A_0X_0Y_0$ *degenerates or there exists* $x'<x$ *such that*

$$\gamma(x',y)-\gamma(x,y)>\sin^2\frac{\varepsilon}{2}\log\frac{x}{x'}.$$

Proof. Assume that $\xi_0\ne0$. The functions $z(x,y)$ and $\gamma(x,y)$ can *a priori* be not differentiable. Therefore, by the symbols $\partial z/\partial x$ and $\partial\gamma/\partial x$ we mean their left upper derivatives with respect to x at a constant y. Then by Lemma 3, we can assert that

$$\frac{\partial z}{\partial x}\le\cos\xi. \tag{2}$$

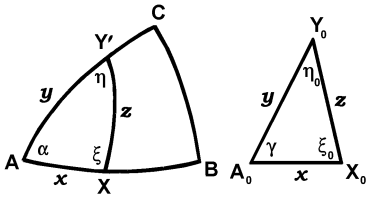


Fig. 63

On the other hand, we have

$$z^2=x^2+y^2-2xy\cos\gamma$$

in the triangle $A_0B_0C_0$, and, therefore,[3]

$$z\frac{\partial z}{\partial x}=x-y\cos\gamma+xy\sin\gamma\frac{\partial\gamma}{\partial x}.$$

And since $x=y\cos\gamma+z\cos\xi_0$ and $xy\sin\gamma=xz\sin\xi_0$, we have

$$\frac{\partial z}{\partial x}=\cos\xi_0+x\sin\xi_0\frac{\partial\gamma}{\partial x}. \tag{3}$$

[3]Certainly, it is important that $\gamma(x,y)$ is a continuous function of x and y and that $xy\sin\gamma\ge0$.

© 2006 by Taylor & Francis Group, LLC

Since $\xi_0 \neq 0$, this equation implies among other things that $\partial\gamma/\partial x$ assumes a finite value, because, by Lemma 4, $|\partial z/\partial x| \leq 1$. Together with inequality (2), Eq. (3) yields

$$\cos\xi \geq \cos\xi_0 + \sin\xi_0 \frac{\partial\gamma}{\partial x}.$$

Therefore,

$$\frac{\partial\gamma}{\partial x} \leq \frac{\cos\xi - \cos\xi_0}{x\sin\xi_0} \leq \frac{\cos\xi - \cos\xi_0}{x}, \tag{4}$$

since $\xi > \xi_0$, and hence $\cos\xi - \cos\xi_0 < 0$.

Since $\xi - \xi_0 = \varepsilon > 0$, we have[4]

$$\cos\xi_0 - \cos\xi = 2\sin\frac{\xi_0+\xi}{2}\sin\frac{\xi-\xi_0}{2} \geq 2\sin^2\frac{\xi-\xi_0}{2} = 2\sin^2\frac{\varepsilon}{2}.$$

Since, in addition, $1/x = d\log x/dx$, inequality (4) implies

$$\frac{\partial\gamma}{\partial x} \leq -2\sin^2\frac{\varepsilon}{2}\frac{d\log x}{dx}. \tag{5}$$

Here $\partial\gamma/\partial x$ is the upper derivative, i.e.,

$$\frac{\partial\gamma}{\partial x} = \lim \sup_{x'\to x-0} \frac{\gamma(x,y) - \gamma(x',y)}{x - x'}.$$

Therefore, for $x' < x$, we have

$$\gamma(x,y) - \gamma(x',y) < \frac{\partial\gamma}{\partial x}(x'-x) + \delta_1(x-x'), \tag{6}$$

where δ_1 is an $x - x'$. At the same time,

$$\log\frac{x}{x'} = \log x - \log x' = \frac{d\log x}{dx}(x-x') + \delta_2(x-x'), \tag{7}$$

where δ_2 is an $x - x'$.

Using formulas (6) and (7), from inequality (5), we obtain the following inequality for finite increments:

$$\gamma(x,y) - \gamma(x',y) < -\log\frac{x}{x'}2\sin^2\frac{\varepsilon}{2} + \delta\cdot(x-x'), \tag{8}$$

where δ is an infinitesimal magnitude with respect to $x - x'$. Of course, we can take $x - x'$ so small that

$$\delta\cdot(x-x') < \sin^2\frac{\varepsilon}{2}\log\frac{x}{x'}.$$

Inequality (8) obviously implies

$$\gamma(x',y) - \gamma(x,y) > \sin^2\frac{\varepsilon}{2}\log\frac{x}{x'}.$$

[4]Since $\xi \leq \pi$, we have $\cos\frac{\xi}{2} \geq 0$, and, therefore, $\sin\frac{\xi+\xi_0}{2} = \sin\frac{\xi}{2}\cos\frac{\xi_0}{2} + \sin\frac{\xi_0}{2}\cos\frac{\xi}{2} \geq \sin\frac{\xi}{2}\cos\frac{\xi_0}{2} - \sin\frac{\xi_0}{2}\cos\frac{\xi}{2} = \sin\frac{\xi-\xi_0}{2}$.

© 2006 by Taylor & Francis Group, LLC

The lemma is proved.

Now let us pass directly to the proof of the main lemma on convex triangles. Recall its content:

Let ABC be a sufficiently small convex triangle in a manifold with the metric of positive curvature, and let $A_0B_0C_0$ be the plane triangle with sides $A_0B_0 = AB$, $A_0C_0 = AC$, and $B_0C_0 = BC$. For each pair of sides of the triangle ABC, e.g., for AB and AC, the following two possibilities are only open: either the lower angle between these sides is not less than the corresponding angle of the triangle $A_0B_0C_0$ or these sides prolong one another, i.e., some segments AB_1 and AC_1 of these sides jointly comprise a single shortest arc.

Proof. Assume that the lower angle α between the sides AB and AC of the triangle ABC is less than the corresponding angle α_0 of the triangle $A_0B_oC_0$, so that

$$2\varepsilon = \alpha_0 - \alpha > 0. \tag{9}$$

We have to prove that this implies that the sides AB and AC prolong each other.

The side BC cannot be the unique shortest arc that goes in the triangle ABC and connects B and C. In this case, by Lemma 1, there is a leftmost shortest arc among them. We replace the side BC by this shortest arc and obtain a new triangle ABC which is also convex (this is obvious in its own right, and also follows from by Theorem 4 of Sec. 5 of Chapter II). In what follows, we will consider only this triangle. As before, we take two variable points X and Y on the sides AB and BC, and by XY, we always mean the leftmost shortest arc. We preserve the notation, which was introduced in Lemmas 3 and 4 (see Fig. 63). We will say that the pair of points X_1 and Y_1 lies more to the left than the pair X_2 and Y_2 if $x_1 \le x_2$ and $y_1 \le y_2$, and the strict inequality holds at least in one case.

If we denote the lengths of AB and AC by b and c, then in our notation $\gamma(b,c), \xi_0(b,c)$, and $\eta_0(b,c)$ stand for the angles of the triangle $A_0B_0C_0$, while $\xi(b,c)$ and $\eta(b,c)$ stand for the lower angles between the sides AB, AC and BC.

By assumption (9),

$$2\varepsilon = \gamma(b,c) - \alpha > 0. \tag{9}$$

The positivity of curvature yields

$$\alpha + \xi(b,c) + \eta(b,c) \ge \pi,$$

i.e.,

$$\alpha + \xi(b,c) + \eta(b,c) \ge \gamma(b,c) + \xi(b,c) + \eta(b,c).$$

By inequality (9), this implies that at least one of the angles ξ and η exceeds the corresponding angles ξ_0 and η_0 by at least ε. Let, e.g.,

$$\xi(b,c) \ge \xi_0(b,c) + \varepsilon.$$

Then by Lemma 2, only one of the following two cases is possible: either $\xi_0(b,c) = 0$ or there is some $x < b$ such that

$$\gamma(x,c) - \gamma(b,c) > \sin^2\frac{\varepsilon}{2}\ln\frac{b}{x} = \sin^2\frac{\varepsilon}{2}\ln\frac{bc}{xc}.$$

© 2006 by Taylor & Francis Group, LLC

By inequality (9), we have $\gamma(b,c) > 0$; therefore, if $\xi_0(b,c) = 0$, then the triangle $A_0B_0C_0$ degenerates into a segment and $A_0B_0 + A_0C_0 = B_0C_0$, i.e., $AB + AC = BC$, so that the sides AB and BC form a single shortest arc. Therefore, we can say that $\xi_0 \neq 0$ as well as $\eta_0(b,c) \neq 0$. If

$$\eta(b,c) \geq \eta_0(b,c) + \varepsilon \quad \text{and} \quad \eta_0(b,c) \neq 0,$$

then there exists $y < c$ such that

$$\gamma(b,y) - \gamma(b,c) > \sin^2 \frac{\varepsilon}{2} \log \frac{c}{y} = \sin^2 \frac{\varepsilon}{2} \log \frac{bc}{by}.$$

Therefore, in any case, to the left of the pair (b,c), there is a pair (x_1,y_1) with $y_1 = c$ or $x_1 = b$ such that

$$\gamma(x_1,y_1) - \gamma(b,c) > \sin^2 \frac{\varepsilon}{2} \log \frac{bc}{x_1y_1}. \tag{10}$$

Since $\gamma(x_1,y_1) > \gamma(b,c)$, by (9) we have

$$\gamma(x_1,y_1) - \alpha > 2\varepsilon.$$

Therefore, applying the same argument to the triangle AX_1Y_1 in which either $X_1 = B$ or $Y_1 = C$, we have the following two possibilities: either AX_1 and AY_1 form a single shortest arc or there exists a pair (x_2,y_2) that lies farther left than the pair (x_1,y_1) such that

$$\gamma(x_2,y_2) - \gamma(x_1,y_1) > \sin^2 \frac{\varepsilon}{2} \log \frac{x_1y_1}{x_2y_2}.$$

Summing this inequality with inequality (10), we obtain

$$\gamma(x_2,y_2) - \gamma(b,c) > \sin^2 \frac{\varepsilon}{2} \log \frac{bc}{x_2y_2}. \tag{11}$$

Since $\gamma(x_2,y_2) > \gamma(b,c)$, by (9), we obtain

$$\gamma(x_2,y_2) - \alpha > 2\varepsilon. \tag{12}$$

This process can be continued infinitely, yielding farther and farther "left" pairs (x,y) for which, first, the segments AX and AY do not form a single shortest arc and, second, the inequalities of the form (11) and (12) hold.

Let $(\overline{x},\overline{y})$ be the greatest lower bound of these pairs, i.e., the pair such that there exists such a pair (x,y) with x and y arbitrarily close to $\overline{x}$ and $\overline{y}$, but no pair (x,y) that lies farther left than the pair $(\overline{x},\overline{y})$.

None of the numbers $\overline{x}$ and $\overline{y}$ can be equal to zero. Indeed, if, e.g., $\overline{x} = 0$, then we would have a sequence of pairs (x_n,y_n) converging to $(\overline{x},\overline{y})$ for which inequalities (11) hold, and at the same time, $x_n \to \infty$. But this is obviously impossible, since as $x_n \to 0$, we have $\log(bc/(x_ny_n)) \to \infty$, and thus, $\gamma(x_n,y_n) \to \infty$, while $\gamma(x_n,y_n) \leq \pi$ (since $\gamma(x_n,y_n)$ is an angle of the plane triangle!). Hence, none of the numbers $\overline{x}$ and $\overline{y}$ can vanish.

© 2006 by Taylor & Francis Group, LLC

The functions $\gamma(x, y)$ and $\log(bc/(xy))$ are continuous on the half-open intervals $0 < x \le b$ and $0 < y \le c$. Therefore, inequalities (11) an (12) also hold for $x = \overline{x}$ and $y = \overline{y}$, since $(\overline{x}, \overline{y})$ is the greatest lower bound of those pairs (x, y) for which these inequalities hold.

Consequently, for the pair $(\overline{x}, \overline{y})$, we can apply the above arguments, i.e., we have the following two possibilities for this pair: (1) the segments $\overline{AX}$ and $\overline{AY}$ form a single shortest arc and (2) there exists a pair that lies farther left than $(\overline{x}, \overline{y})$ and satisfies (11) and (12). But, by definition, $(\overline{x}, \overline{y})$ is the greatest lower bound of the pairs satisfying these inequalities. Therefore, the second possibility is excluded and only the first possibility remains, as required.

3. Corollaries of the Main Lemma on Convex Triangles

The main lemma on convex triangles proven in the previous section easily implies a number of important corollaries:

Theorem 1. *Between the sides of each sufficiently small triangle in a manifold with a metric of positive curvature, there exist angles understood in the usual sense. If two sides do not prolong each other, then the usual angle and the lower angle between them are equal.*

In fact, these angles are equal also in the case where one of these sides prolongs the other. However, we will leave this assertion without proof.

Proof. Let ABC be a small convex triangle in a manifold with a metric of positive curvature. If the sides AB and AC prolong each other, then, obviously, there is an angle between them, understood in the usual sense and equal to π. Therefore, we assume that these sides AB and AC are not prolongations of each other; in this case, we will prove that the lower angle between these sides reduces to the angle in the usual sense.

As usual, let X and Y be variable points on AB and AC, and put $x = AX$, $y = AY$, and $z = XY$. Let $\gamma(x, y)$ have the usual meaning, i.e., be the angle opposite the side z in the plane triangle with sides x, y, and z. Finally, let α be the angle between AB and AC.

The definition of lower angle yields

$$\liminf_{x,y\to 0} \gamma(x, y) \ge \alpha.$$

On the other hand, applying the main lemma of the previous section to each of the triangles AXY, we obtain

$$\gamma(x, y) \le \alpha.$$

Hence, $\liminf_{x,y\to 0} \gamma(x, y)$, i.e., the angle between AB and AC exists and equals α.

Theorem 2. *The angles of a sufficiently small triangle ABC are not less than the corresponding angles of the plane triangle $A_0B_0C_0$ with sides of the same length.*

© 2006 by Taylor & Francis Group, LLC

Proof. If the sides AB and AC of the triangle ABC are prolongations of each other, then the angle between them equals π and is always not less than the corresponding angle in the plane triangle. If the sides AB and AC are not prolongations of each other, then by the main lemma of the previous section the lower angle made by them is not less than the corresponding angle of the triangle $A_0B_0C_0$. By Theorem 1, the lower angle coincides with the usual angle in this case. Hence, the theorem is proved.

Theorem 3. *The nonoverlapping condition for shortest arcs holds in every manifold with a metric of positive curvature.*

Proof. Let AB and AC be two shortest arcs emanating from a common point A. Assume that these arcs overlap each other on the segment AD and go apart.

Then their segments AD, DB, and DC divide a sufficiently small neighborhood of D into three sectors U, W, and W (Fig. 64).

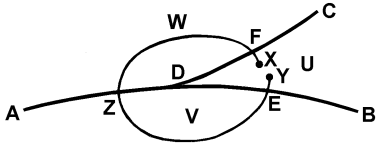


Fig. 64

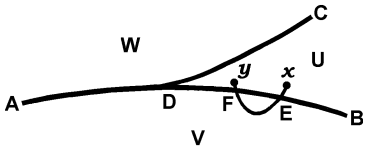


Fig. 65

Let us prove that the sector U made by DB and DC is convex. Let X and Y be two points of this sector (so close to D that the shortest arc XY cannot run apart from the lines DA, DB, and DC, i.e., for example, going from U to V, this shortest arc must intersect DB). Assume that the shortest arc XY does not belong to the sector U; let EF be some segment of this shortest arc lying outside U (of course, it is possible that XY also has other segments outside U). For definiteness, we assume that the point E lies on DB. The following two possibilities are open for the point F: (1) F lies on DC (Fig. 64); (2) F lies on DB (Fig. 65).

If F lies on DC, then the segment EF of the shortest arc XY must intersect AD at a certain point Z. But then the parts EZ and FZ of the segment EF are equal to the segments EZ and FZ of the shortest arcs AB and AC. Hence these segments of the shortest arcs AB and AC together form one shortest arc $\overline{EF}$; this is not possible, since the line formed by these segments can be shortened by excluding the segment ZD. Hence, the point F cannot lie on DC. This also proves that the shortest arc XY cannot intersect AD.

Now let the point F lie on DB. This means that both endpoints of the segment EF lie on the shortest arc DB and so this segment is equal to the segment EF of the shortest arc DB. Therefore, the segment EF of the shortest arc XY can be replaced by the segment EF of the shortest arc DB. Performing such a replacement for all segments of the shortest arc XY that go outside the sector U, we replace this shortest arc by a shortest arc that lies entirely in the sector U. This proves the convexity of the sector U.

Take two points P and Q inside the lines DB and DC such that $DP = DQ$ and draw a shortest arc PQ in the sector U. This is possible by what we have proved

© 2006 by Taylor & Francis Group, LLC

above. We obtain a convex triangle DPQ. Indeed, if X and Y are two points of this triangle, then by what we have proved above, we may connect these points by a shortest arc in the sector U. If this shortest arc does not lie entirely in the triangle DPQ, then this arc intersects the side PQ. But each segment of the shortest arc XY between two intersection points with PQ can be replaced by the corresponding segment of the side PQ. Thus, there always exists a shortest arc XY lying in the triangle DPQ, i.e., the triangle DPQ is convex.

If we adjoin a small segment GD of the shortest arc AD to the triangle DPQ, then we obtain the convex triangle GPQ again; this can easily be verified, since AD is a prolongation of both sides DP and DQ. But the angle made by the sides GP and GQ in the triangle GPQ is equal to zero. Therefore, according to Theorem 2, the corresponding angle of the plane triangle with sides of the same length must be equal to zero, and so

$$GP + PQ = GQ$$

or

$$GQ + PQ = GP.$$

But since $GP = GQ$ by the choice of the points P and Q, this implies $PQ = 0$, i.e., the points P and Q coincide; this contradicts the assumption that the shortest arcs AB and CD go apart. Consequently, the shortest arcs AB and AC, which coincide on the segment AD, cannot then go apart; as required.

Since the nonoverlapping condition for shortest arcs holds in a manifold with a metric of positive curvature, all its consequences obtained in Chapter II are valid.

The following theorem on triangulation of Sec. 5 of Chapter II will be of utmost importance for us.

Theorem 4. *Each convex triangle in a manifold with a metric of positive curvature can be partitioned into arbitrarily small convex triangles.*

In Sec. 4 of Chapter III, we have proved the main theorem on the convergence of angles between shortest arcs on convex surfaces (Theorem 4 of Sec. 4 of Chapter III). Its deduction was based only on the fact that the angles of a triangle on a convex surface are no less than the angles of a plane triangle; this property was stated for convex triangles. Therefore, we can assert the following.

Theorem 5. *If two shortest arcs L and M that are sides of a variable convex triangle in a manifold with a metric of positive curvature converge to shortest arcs L_0 and M_0, then the angle between L_0 and M_0 is no greater than the lower limit of the angles between L and M.*[5]

This theorem itself will be needed only in the next chapter. Here, we also note that the deduction of the theorems on addition of angles given in Sec. 1 of Chapter IV was based only on the existence of the angle, while the deduction of the theorem asserting that the sum of adjacent angles is equal to π given in Sec. 2 of Chapter IV rested in addition on the nonoverlapping condition for shortest arcs and the above theorem on the convergence of angles. Therefore, we can abstract the general

[5]The shortest arcs L_0 and M_0 are the sides of a convex triangle or prolongations of one another, since we can choose a convergent sequence from small triangles having the sides L and M.

© 2006 by Taylor & Francis Group, LLC

theorem on addition of angles on convex surfaces which was proved in Sec. 2 of Chapter IV to a manifold with a metric of positive curvature on formulating this abstraction for three shortest arcs, which is sufficient for us.

Theorem 6. *Let three shortest arcs L_1, L_2, and L_3 emanate from a point O in a manifold with a metric of positive curvature. Assume that there exist angles between them which are denoted by α_{ij}. Then, if for points X_1 and X_3 on L_1 and L_3, which are arbitrarily close to O, the shortest arc X_1X_3 has common points with the shortest arc L_2, we have $\alpha_{13} = \alpha_{12} + \alpha_{23}$.*

In particular, if L_1 and L_2 prolong one the other, then the shortest arc X_1X_3 overlaps them and has the common point O with L_2; at the same time, the angle made by L_1 and L_2 is equal to π in this case, and the theorem reduces to the fact that the sum of adjacent angles is equal to π.

4. The Complete Angle at a Point

Shortest arcs emanating from a common point O in a manifold with a metric of positive curvature divide a neighborhood of this point into sectors since by the nonoverlapping condition no two of them can intersect arbitrarily close to the point O unless one of them is a part of the other. A sector bounded by two shortest arcs L and M is called convex if some segments of the shortest arcs L and M turn out to be the sides of some convex triangle that contains this sector[6] or prolong one the other comprising a single shortest arc. In a convex sector, every two points can be connected by a shortest arc lying in this sector. The sides of a sector are assumed to belong to it. There is an angle between the shortest arcs bounding a convex sector. We will consider an arbitrary number n of shortest arcs $L_1, \dots, L_n$ that emanate from a common point O and are enumerated in the order of their location around the point O, i.e., L_i and L_{i+1} (and also L_n and L_1) bound the sector that contains no other shortest arcs L_j. The angle between L_i and L_j (if this angle exists) will be denoted by α_{ij}.

Theorem 1. *If shortest arcs $L_1, \dots, L_n$ emanating from a point O in a manifold with a metric of positive curvature partition a neighborhood of the point O into convex sectors, then the sum of the angles $\alpha_{12} + \alpha_{23} + \dots + \alpha_{n1}$ is independent of these shortest arcs and can depend only on the point O.*

Proof. Assume that shortest arcs $L_1, \dots, L_n$ and $M_1, \dots, M_k$ are drawn from the point O; let α_{ij} denote the angle between L_i and $L_{i,i+1}$, and let $\beta_{j,j+1}$ denote the angle made by M_j and M_{j+1}. If the shortest arcs L_i and the shortest arcs M_j separately partition a neighborhood of a point into convex sectors, then they also partition this neighborhood into convex sectors jointly. Therefore, if we enumerate all of them in succession and denote them by $N_1, N_2, \dots, N_{k+n}$, then there will appear a definite angle $\gamma_{i,i+1}$ between N_i and N_{i+1}. By the general theorem on addition of angles, which was proved in Sec. 1 of Chapter IV, the angle $\alpha_{i,i+1}$ made by $L_i = N_p$ and $L_{i+1} = N_q$ is equal to the sum of the angles made by N_p and N_{p+1}, N_{p+1} and N_q. Consequently,

$$\alpha_{12} + \dots + \alpha_{n1} = \gamma_{12} + \dots + \gamma_{n+k,1},$$

[6] See the definition of sector in the beginning of Sec. 3 in Chapter IV.

© 2006 by Taylor & Francis Group, LLC

and, in exactly the same way,

$$\beta_{12} + \cdots + \beta_{k1} = \gamma_{12} + \cdots + \gamma_{n+k,1},$$

as required.

Since the sum of the angles of convex sectors that form a neighborhood of a point O depends only on this point itself by what we have proved above, this sum is naturally called the *complete angle at the point* O. This definition differs from that introduced in Sec. 3 of Chapter IV only by the fact that here we consider convex rather than arbitrary sectors.

It is more essential that we do not prove the possibility of partitioning a neighborhood of *any* point into convex sectors, although this can be done. Therefore, we apply the concept of complete angle only to those points for which such a partition is always possible. These points exist, since each convex polygon can be partitioned into arbitrarily small convex triangles. The vertices of these triangles are such points. Moreover, the complete angle always exists at a point inside a shortest arc, since the two branches of the shortest arc emanating from this point divide its neighborhood into two convex sectors. The complete angle at this point is obviously equal to 2π. The concept of complete angle is needed by no other points. When we prove that a neighborhood of each point is isometric to a convex surface, it will become clear that the above-mentioned restrictions are all motivated only by the way of demonstration.

Theorem 2. *The complete angle at a point in a manifold with a metric of positive curvature does not exceed* 2π.

It is more convenient to prove this theorem in the following slightly more general form.

Let shortest arcs $L_1, \ldots, L_n$ *emanating from a common point* O *in a manifold with a metric of positive curvature be enumerated in the order of their location around* O. *If two neighboring arcs* L_i *and* L_{i+1} *(and also* L_n *and* L_1*) make a definite angle* $\alpha_{i,i+1}$ *with each other* (*but not necessarily bound a convex sector*), *then the sum of these angles does not exceed* 2π.

Proof. If there are only two shortest arcs, then this sum reduces to $\alpha_{12}+\alpha_{21} = 2\alpha_{12}$, and since $\alpha_{12} \le \pi$, the theorem holds in this case. In what follows, we can assume that there are at least three shortest arcs L_i.

Take variable points X_1 and X_{n-1} on the shortest arcs L_1 and L_{n-1}. When X_1 and X_{n-1} are arbitrarily close to O, the following three possibilities are open for the shortest arc X_1X_{n-1}:

1. X_1X_{n-1} always has common points with L_n;

2. X_1X_{n-1} always has no common points with L_n;

3. X_1X_{n-1} sometimes has common points with L_n and sometimes does not have them.

© 2006 by Taylor & Francis Group, LLC

When the shortest arc X_nX_{n-1} has no common points with L_n, this arc intersects the shortest arcs $L_2, \ldots, L_{n-2}$ (if these shortest arcs exist, i.e, if $n > 3$, since L_{n-2} is L_2 for $n = 3$).

If, for a certain location of the points X_1 and X_{n-1}, the shortest arc X_1X_{n-1} passes through the point O, then the shortest arcs L_1 and L_{n-1} turn out to be prolongations of each other, and hence, for all X_1 and X_1 that are sufficiently close to O, the shortest arc X_1X_{n-1} has O as a common point with L_n, i.e., in this case, the first of the above-mentioned possibilities holds, and hence, in the other two cases, the shortest arc X_1X_{n-1} never passes through the point O.

First, we consider the third case. Let us use the general Lemma 2 in Sec. 1 of Chapter IV. This lemma asserts that if, for X_1 and X_{n-1} arbitrarily close to O, the shortest arc X_1X_{n-1} intersects L_n at a different point from O and can exist only for a certain location of the point X_1 and X_{n-1}, then

$$\liminf_{x_1,x_n\to 0} \gamma_{n1}(x_1, x_n) + \liminf_{x_{n-1},x_n\to 0} \gamma(x_{n-1}, x_n) \le \liminf_{x_1,x_{n-1}\to 0} \gamma_{1,n-1}(x_1, x_{n-1}), \qquad (1)$$

where $\gamma_{ij}(x_i, x_j)$ has the usual meaning[7] and the limits are taken over those x_1, x_{n-1}, and x_n that correspond to the shortest arcs X_1X_{n-1} intersecting L_n. But since there exist angles made by L_1 and L_n, L_{n-1} and L_n, the lower limits on the left-hand side of (1) can be replaced by these angles; we thus obtain

$$\alpha_{12} + \alpha_{n-1,n} \le \liminf_{x_1,x_{n-1}\to 0} \gamma_{1,n-1}(x_1, x_{n-1}). \qquad (2)$$

In exactly the same way, if X_1X_{n-1} intersects $L_2, \ldots, L_{n-2}$, then

$$\alpha_{12} + \cdots + \alpha_{n-2,n-1} \le \liminf_{x_1,x_{n-1}\to 0} \gamma_{1,n-1}(x_1, x_{n-1}). \qquad (3)$$

But since $\gamma_{1,n-1}(x_1, x_{n-1}) \le \pi$, inequalities (2) and (3) imply

$$\alpha_{12} + \cdots + \alpha_{n-1,n} + \alpha_{n1} \le 2\pi;$$

and so the theorem is proved in the third case.

In the second case, the shortest arc X_1X_{n-1} does not intersect L_n, and thus, always intersects $L_2, \ldots, L_{n-2}$ at different points from O (provided, certainly, that these shortest arcs exist, i.e., $n > 3$). Here, we can use Theorem 3 of Sec. 1 of Chapter IV, which states that under these conditions, there exists an angle made by L_1 and L_{n-1} equal to the sum of the angles between L_1 and L_2,..., L_{n-2} and L_{n-1}. Therefore, without changing the total sum of angles, we can exclude all shortest arcs $L_2, \ldots, L_{n-2}$. Only three shortest arcs remain, and the conditions of the theorem will hold for them.

In the first case, the shortest arc X_iX_{n-1} always intersects L_n. Therefore, if this shortest arc does not pass through the point O, then according again to Theorem 3 in Sec. 1 of Chapter IV, the angle between L_1 and L_{n-1} exists and equals the sum of the angles between L_1 and L_{n-1}, L_n and L_{n-1}. If X_1X_{n-1} passes through the point O, then L_1 and L_{n-1} are prolongations of each other: the angle between

[7] For example, $\gamma_{1n}(x_1, x_n)$ is the angle in the plane triangle with sides equal to $OX_1 = x_1, OX_n = x_n$, and X_1X_n, where X_n is the intersection point of the shortest arc X_1X_{n-1} with L_n.

© 2006 by Taylor & Francis Group, LLC

them is equal to π, and by the theorem on the sum of adjacent angles, this angle is equal to the sum of the angles between L_1 and L_{n-1}, L_n and L_{n-1}. Hence, not changing the total sum of the angles, we can always exclude the shortest arc L_n. The conditions of the theorem will hold for remaining shortest arcs, and we can repeat our argument with three cases for them. This implies that eventually we will arrive at one of the following possibilities:

1. Only two shortest arcs remain; then the theorem is proved.
2. We come to the "third case;" again, the theorem is proved.
3. We come to the "second case;" moreover, there remain only three shortest arcs.

In the "second case," the shortest arc X_1X_2 does not intersect L_3 and hence always lies in the sector between L_1 and L_2 (now $n = 3$, and, thus, $n - 1 = 2$). But if, say, the shortest arc X_2X_3 with endpoints at L_2 and L_3 intersects L_1 for X_2 and X_3 arbitrarily close to O, then using the previous argument (the third case), we obtain that the sum of the angles α_{12}, α_{23}, and α_{31} does not exceed 2π. Therefore, we are left with considering the possibility in which none of the shortest arcs X_1X_2, X_2X_3, and X_3X_1 intersects L_3, L_1, and L_2, respectively; in other words, each of these shortest arcs stays in its own sector; this implies that the sectors bounded by the shortest arcs L_1, L_2, and L_3 are all convex.[8]

Thus we must prove the following assertion.

If shortest arcs L_1, L_2, and L_3 emanate from a point O, partition a neighborhood of this point into convex sectors, and none of every two of them prolongs the other, then the sum of the angles between these shortest arcs does not exceed 2π.

To prove this assertion, we take points X_1, X_2, and X_3 on L_1, L_2, and L_3 that are sufficiently close to O and draw shortest arcs X_iX_J. Each of these shortest arcs lies inside its own sector U_{ij}, since these sectors are convex and none of every two of the shortest arcs L_i prolongs the other. Hence the point O lies inside the triangle $X_1X_2X_3$.

If the points X_1, X_2, and X_3 are taken sufficiently close to the point O, then according to the theorem proved in Sec. 4 of Chapter II, there exists a convex geodesic triangle T that includes the triangle $X_1X_2X_3$. This triangle is characterized by the property that it has minimum perimeter among all geodesic triangles that include $X_1X_2X_3$. The vertices of the triangle T either coincide with X_1, X_2, and X_3, or can be "non-geniune vertices", i.e., two sides meeting at one vertex can form a single geodesic. Since the sectors U_{ij} are convex, their intersections with the triangle T are also convex, i.e., are some convex geodesic triangles T_1, T_2, and T_3 (see Fig. 66). Therefore, if A_1 and A_2 are the intersection points of the perimeter of the triangle T with the shortest arcs L_1 and L_2, then there is a shortest arc $\overline{A_1A_2}$ in the triangle

[8] Let, e.g., points X and Y lie in the sector U_{12} bounded by L_1 and L_2. Assume that the segment EF of the shortest arc XY lies outside U_{12}, and, moreover, the point E lies on L_1. Then the point F cannot lie on L_2 as follows from the nonoverlapping condition for shortest arcs (Theorem 1 of Sec. 3 of Chapter II). But the point F cannot lie on L_2; since, if E lies on L_1 and F lies on L_2, then EF intersects L_3; this contradicts the assumption that no shortest arcs with endpoints at L_1 and L_2 intersect L_3. Consequently, the shortest arc XY cannot have segments outside the sector U_{12}, i.e., this sector is convex.

© 2006 by Taylor & Francis Group, LLC

$T_1 = TU_2$. If the side A_1A_2 of the triangle T_1 is not a shortest arc, then replacing this side by the shortest arc $\overline{A_1A_2}$, we obtain a new geodesic triangle instead of T that contains $X_1X_2X_3$[9] and has a smaller perimeter. However, this contradicts the fact that the triangle T has the shortest perimeter among all triangles that contain $X_1X_2X_2$. Hence, the sides of the triangle T are shortest arcs, its vertices lie on the shortest arcs L_1, L_2, and L_3, and these shortest arcs divide T into three "partial" convex triangles T_1, T_2, and T_3. If the points X_1, X_2, and X_3 are taken sufficiently close to the point O, then the triangle T is sufficiently small, and thus, we have proved that there exist arbitrarily small convex triangles with the indicated properties.

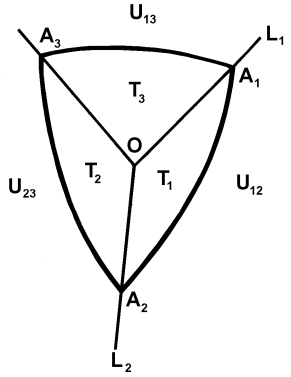


Fig. 66

Let α_{12}, α_{23}, and α_{31} be the angles made by the shortest arcs L_1, L_2, and L_3. Assume by way of contradiction that their sum is greater than 2π and put

$$3\varepsilon = \alpha_{12} + \alpha_{23} + \alpha_{31} - 2\pi > 0. \tag{1}$$

As usual, let $\gamma_{ij}(x_i, x_j)$ denote the angle opposite the side X_iX_j in the plane triangle with sides equal to $OX_i = x_i$, $OX_j = x_j$, and X_iX_j. Since $\lim_{x_i,x_j\to 0} \gamma_{ij}(x_i, x_j) = \alpha_{ij}$, we can take a small convex triangle $A_1A_2A_3$ with vertices at the shortest arcs L_i such that for all points of X_iX_j lying on the segments OA_i and OA_j,

$$|\alpha_{ij} - \gamma_{ij}(x_i, x_j)| < \frac{\varepsilon}{2} \quad (i, j = 1, 2, 3). \tag{2}$$

The triangle $A_1A_2A_3$ is divided by the shortest arcs L_1, L_2, and L_3 into the three triangles OA_iA_j each of which is convex, since the sectors bounded by the shortest arcs L_i are convex.

Let ω and ω_{ij} stand for the curvatures of the triangles $A_1A_2A_3$ and OA_iA_j $(i, j = 1, 2, 3)$, respectively, i.e., the sums of their angles minus π. The angles of the triangle $A_1A_2A_3$ are composed of the angles of the triangles OA_iA_j, and so

$$\omega_{12} + \omega_{23} + \omega_{31} = \omega + \alpha_{12} + \alpha_{23} + \alpha_{31} - 2\pi,$$

or, by (1),

$$\omega_{12} + \omega_{23} + \omega_{31} = \omega + 3\varepsilon.$$

Since $\omega > 0$, at least one of the numbers ω_{ij} should be no less than ε. This is always so, however small triangle $A_1A_2A_3$ would be. Therefore, among the three given triangles OA_iA_j, there is at least one triangle such that for arbitrarily small analogous triangles $AOA_i'A_j'$ contained in this triangle, we have $\omega_{ij}' \geq \varepsilon$.

For definiteness, we assume that the triangle OA_1A_2 has this property.

Take two points X_1 and X_2 on the sides OA_1 and OA_2 of this triangle and draw a shortest arc X_1X_2; we obtain the triangle OX_1X_2 which lies in our triangle. By assumption, this triangle includes the triangles $OA_1'A_2'$ for which $\omega(OA_i'A_j') > \varepsilon$. The triangle $OA_1'A_2'$ lies in the triangle OX_1X_2 and its vertices lie on the sides of

[9] A_1A_2 cannot intersect X_1X_2 as is clear from the nonoverlapping condition for shortest arcs.

© 2006 by Taylor & Francis Group, LLC

OX_1X_2. This easily implies $\omega(OX_1X_2) \geq \omega(OA_1'A_2')$.[10] Hence, for all points X_1 and X_2 on the sides, we have

$$\omega(OX_1X_2) \geq \varepsilon.$$

Denote by ξ_1 and ξ_2 the angles of the triangle OX_1X_2 at the vertices X_1 and X_2, and let ξ_{10} and ξ_{20} be the corresponding angles of the plane triangles with sides of the same length. Then

$$\alpha_{12} + \xi_1 + \xi_2 - \pi = \omega(OX_1X_2) \geq \varepsilon,$$

$$\gamma_{12}(x_1, x_2) + \xi_{10} + \xi_{20} - \pi = 0,$$

and hence

$$(\alpha_{12} - \gamma_{12}) + (\xi_1 - \xi_{10}) + (\xi_2 - \xi_{20}) \geq \varepsilon.$$

At the same time, by inequality (2),

$$\alpha_{12} - \gamma_{12}(x_1, x_2) \leq \frac{\varepsilon}{2},$$

and, therefore,

$$(\xi_1 - \xi_{10}) + (\xi_2 - \xi_{20}) > \frac{\varepsilon}{2}.$$

Hence, for any points X_1 and X_2, i.e., for any x_1 and x_2, at least one of the differences $\xi_1 - \xi_{10}$ and $\xi_2 - \xi_{20}$ is no less than $\varepsilon/4$. But if, say, $\xi_1(x_1, x_2) - \xi_{10}(x_1, x_2) \geq \varepsilon/4$, then, by Lemma 4 of Sec. 2, there exists $x_1' < x_1$ such that

$$\gamma_(x_1', x_2) - \gamma(x_1, x_2) > \sin^2 \frac{\varepsilon}{8} \log \frac{x_1}{x_2}.$$

Therefore, we can proceed almost in the same way as in the proof of the main lemma on convex triangles (Sec. 2). Moving the points X_1 and X_2 from A_1 and A_2 to the point O so that every time, the angle $\gamma(x_1, x_2)$ receives the corresponding increment, we will obtain the values of this angle such that

$$\gamma(x_1, x_2) - \gamma(a_1, a_2) > \sin^2 \frac{\varepsilon}{8} \log \frac{a_1 a_2}{x_1 x_2},$$

where $a_1 = OA_1$ and $a_2 = OA_2$. But for sufficiently small x_1 and x_2, the right-hand side of this inequality becomes arbitrarily large, which is absurd, since $\gamma(x_1, x_2) \leq \pi$. Another possibility of the main lemma is excluded, namely, the case in which the sides OA_1 and OA_2 prolong each other, since, by condition, the shortest arcs L_1 and L_2, which contain them as segments, are not prolongations of each other. We obtain a contradiction, which shows that the assumption that $\alpha_{12} + \alpha_{23} + \alpha_{31} > 2\pi$ is impossible. Hence $\alpha_{12} + \alpha_{23} + \alpha_{31} \leq 2\pi$, as required.

[10] Drawing X_1A, we divide the triangles OX_1X_2 into the following three triangles: $OA_1'A_2'$, $X_1A_1'A_2'$, and $X_2A_2X_1$. These triangles are convex. Therefore, they have angles, and the simple addition of angles yields $\omega(OX_1X_2) = \omega(OA_1'A + 2') + \omega(X_1A_1'A_2') + \omega(X_2A_2'X_1)$, and since $\omega(X_1A_1'A_2')$ and $\omega(X_2A_2'X_1) \geq 0$, we have $\omega(OX_1X_2) \geq \omega(OA_1'A_2')$.

© 2006 by Taylor & Francis Group, LLC

5. Curvature and Two Related Estimates

We now can introduce the concept of the curvature of some simple sets in a manifold with a metric of positive curvature in much the same way as in Sec. 1 of Chapter V. An essential difference consists in the fact that instead of arbitrary triangles, we will consider only sufficiently small convex triangles, and instead of arbitrary points, we will consider only those points at each of which the complete angle is defined. According to the definition of the previous section, the concept of complete angle applies yet only to those points whose neighborhoods may be partitioned into convex sectors. In fact, this is possible for all points, but while it is not proved, we have to specify the fact that we speak only about such points.

We define curvature for the following three types of basis sets: the interiors of convex triangles, shortest arcs without endpoints, and points. *The curvature of the interior of a convex triangle is defined as the sum of its angles minus* π. *The curvature of a shortest arc with ends excluded is equal to zero. The curvature of a point is defined as the difference between* 2π *and the complete angle at this point.* By this definition, curvature is nonnegative, since the angles of a convex triangle are not less than those of a plane triangle and the complete angle at a point is $\leq 2\pi$. After that, we can introduce "elementary" sets and define curvature for them in exactly the same way as in Sec. 1 of Chapter V. However, it is sufficient to restrict exposition to the theorem on the curvature of a polygon.

Theorem. *Let a convex polygon* P *be partitioned into convex triangles* $T_1, \ldots, T_n$ *with vertices* $X_1, \ldots, X_m$ *lying inside* P. *Then the sum of the curvatures of interiors of the triangles* T_i *and the curvatures of the vertices* X_i *does not depend on the form of the partition and is expressed by the formula*

$$\sum_{i=1}^{n} \omega(T_i) + \sum_{i=1}^{m} \omega(X_i) = 2\pi\chi - \sum_{j=1}^{k} (\pi - \alpha_j), \tag{1}$$

where χ *is the Euler characteristic and* α_j *are the angles of the polygon* P.

This theorem repeats almost literally Theorem 1 in Sec. 1 of Chapter V with only one difference that here we speak especially on a convex polygon. But namely due to this fact, the angles of the polygon always exist; these are the angles made by its sides which certainly bound convex sectors. By the same reason, the angle of the polygon P is equal to the sum of the angles of the triangles T_i with the same vertex. Thus, in order to obtain a proof of this theorem, it is necessary only to literarily repeat the proof of Theorem 1 of Sec. 1 of Chapter V. This theorem applies to every convex polygon, since we have proved that each polygon can be partitioned into convex triangles. The magnitude, which is expressed by formula (1), is naturally called the curvature of the interior of the polygon P.

If a polygon P is homeomorphic to a disk then $\chi = 1$, and formula (1) yields

$$\omega(P) = \sum_{j=1}^{k} \alpha_j - (k-2)\pi;$$

© 2006 by Taylor & Francis Group, LLC

in the case of a triangle, this coincides with the definition of its curvature given above. If P is homeomorphic to the sphere then there are no angles α_j, and $\chi = 2$, so that $\omega(P) = 4\pi$.

Formula (1) *contains the following two important properties of curvature*: *nonnegativity and additivity.* The first property is implied by the fact that all curvatures $\omega(T_i)$ and $\omega(X_i)$ are nonnegative. The second is obtained by simple addition: *if a convex polygon P is composed of convex polygons P_i with vertices Y_i lying inside P, then*

$$\omega(P) = \sum \omega(P_i) + \sum \omega(Y_i), \tag{2}$$

where $\omega(P_i)$ denotes the curvature of the interior of the polygon P_i. In order to prove Eq. (2), it is sufficient to decompose every polygon P_i into convex triangles and take the sum of the curvatures of the interiors of all these triangles and their vertices lying inside P.

In what follows, when considering approximation of a given metric of positive curvature by polyhedral metrics, we will need two estimates related with curvature. The first of them is given by the following lemma:

Lemma 1. *Let ABC be a sufficiently small convex triangle in a manifold with a metric of positive curvature, and let X and Y be two points on the sides AB and AC of this triangle. Let x, y, and z denote the distances AX, AY, and XY. Construct the plane triangle $A_0B_0C_0$ with sides of the same length as ABC, i.e., such that $A_0B_0 = AB$, etc. Take two points X_0 and Y_0 on its sides A_0B_0 and A_0C_0 such that $A_0B_0 = x$ and $A_0Y_0 = y$; set $X_0Y_0 = z_0$. If ω is the curvature of the interior of the triangle ABC and d is the length of its maximal side, then we have*

$$|z - z_0| \le \omega d.$$

This lemma repeats almost literally Theorem 2 of Sec. 6 of Chapter V, and its proof is the same since we use nothing but the additivity of curvature and elementary trigonometry in this proof. Hence we may assume that Lemma 1 is proved.

Another estimate, which is of interest in its own right, is expressed by the following lemma.

Lemma 2. *Let two points A and B in a manifold with a polyhedral metric of positive curvature be connected by a shortest arc AB and also by a certain geodesic $\overline{AB}$ so that these lines bound a digon D homeomorphic to a disk. Let a and b stand for the corresponding lengths of the lines AB and $\overline{AB}$, and let ω be the curvature of the interior of the digon D (i.e., the sum of its angles as is clear from the definition of curvature). If $\omega < \pi$, then we have*

$$a \ge b \cos\frac{\omega}{2}.$$

Proof. Cut out the digon D out of our manifold, and taking a second copy D' of this digon, identify their corresponding sides. As a result, we obtain the manifold $D + D'$ which is homeomorphic to a sphere. A polyhedral metric of positive curvature is naturally defined on this manifold, since the angles of our digons are $< \pi$. Hence, the complete angle at every point of this manifold is at most 2π.

© 2006 by Taylor & Francis Group, LLC

By the realization theorem for a polyhedral metric of positive curvature, the manifold $D + D'$ is isometric to a closed convex polyhedron, and the digon D itself is isometric to, so to say, a half of this polyhedron. However, we do not need this remark. The only fact that this digon is convex in the manifold $D + D'$ is important for us. Indeed, if points X and Y of the digon D are connected by a shortest arc in $D + D'$ that lies in D' on the segment $\overline{X_1Y_1}$, then by the complete symmetry of D and D', there is a shortest arc X_1Y_1 in D itself. Replacing the segment $\overline{X_1Y_1}$ by this shortest arc X_1Y_1 and treating all segments in D' likewise, we obtain a shortest arc XY that lies entirely in the digon D. Consequently, we can proceed further under the assumption that the digon D is convex, although this digon can be nonconvex in the initial manifold.

Since the geodesic $\overline{AB}$ is a shortest arc on any sufficiently small segment, we can divide this line into segments each of which is a shortest arc. If two neighboring segments together form a single shortest arc, then we combine these segments. Performing these operations, we finally arrive at the partition of the line $\overline{AB}$ by the points A, C_1, C_2, ..., C_n, B into shortest arcs $\overline{AC_1}$, $\overline{C_1C_2}$, ..., $\overline{C_nB}$ such that no pair of them combines into a single shortest join. Since the line $\overline{AB}$ is a geodesic, then there will be no points C_i at all only if $\overline{AB}$ is also a shortest arc. In this case, $a = b$, and the lemma is trivial. Therefore, we can assume that the points C_i do exist.

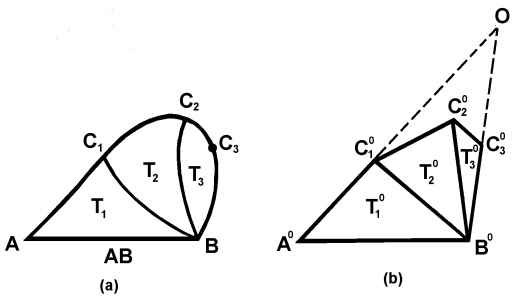


Fig. 67

Draw shortest arcs BC_1, BC_2, ..., BC_n in the digon D (Fig. 67, a). They cannot pass through the vertex A, since the angle at this vertex is less than π. (By the condition of the lemma, even the sum of the angles at the vertices A and B is less than π.) Moreover, these lines do not intersect. All these lines but BC_n lie inside the digon, since by the choice of the points C_i, none of the segments $\overline{BC_i}$ of the line $\overline{AB}$ is a shortest arc, except for the segment $\overline{BC_n}$ which, in contrast, is a shortest arc.[11] We take the line BC_n so that this line coincides with $\overline{BC_n}$. It is clear from what we have said above that the lines BC_1, ..., BC_{n-1} partition the digon into triangles $T_1 = BAC_1$, $T_2 = BC_1C_2$, ..., $T_n = BC_{n-1}C_n$ whose sides are shortest arcs.

Now we construct the triangles $T_1^0 = B^0A^0C_1^0$, $T_2^0 = B^0C_1^0C_2^0$, ..., $T_n^0 = B^0C_{n-1}^0C_n^n$ with sides equal to the sides of the triangles T_1, T_2, ..., T_n. We let the triangles T_i^0 abut on one another in the same way as the triangles T_1, T_2,..., T_n do (Fig. 67, b). As a result, we obtain some polygon Q_n; we assert that this polygon is convex.

To prove this, we imagine that the polygon Q is obtained by sequentially adjoining some triangles to one another. We then first obtain the triangle $Q_1 = T_1^0$, next we obtain the quadrangle $Q_2 = T_1^0 + T_2^0$, and so on. The fact that all these

[11] The fact that the lines BC_i lie inside the digon and do not intersect follows from the general properties of shortest arcs in a polyhedral metric of positive curvature (see Sec. 2 of Chapter III).

© 2006 by Taylor & Francis Group, LLC

figures are convex will be proved by induction. The triangle $Q_1 = T_1^0$ is obviously convex. Assume that we have verified that the polygon Q_k is convex. The polygon Q_{k+1} results from Q_k by adjoining the triangle $T_{k+1}^0 = B^0C_k^0C_{k+1}^0$. Moreover, the angle at the vertex B^0 remains less than π. Indeed, this angle is composed of the angles of the triangles $T_1^0, \ldots, T_{k+1}^0$,[12] while the sum of the angles of the triangles $T_1, \ldots, T_{k+1}$ at the vertex B is no greater than π, since, by the condition of the lemma, even the sum of the angles of the digon D is less than π.

The angle at the vertex C_k^0 of the polygon Q_{k+1} is composed of the angles of the triangles T_k^0 and T_{k+1}^0, which are no greater than the corresponding angles of the triangles T_k and T_{k+1}. And the sum of the latter angles is equal to the angle between the two segments $\overline{C_{k-1}C_k}$ and $\overline{C_kC_{k+1}}$ of the geodesic $\overline{AB}$, i.e., is equal to π. Hence, when the polygon Q_{k+1} results from Q_k by adjoining Q_k to the triangle T_{k+1}^0, the angles at the vertices B^0 and C_k^0 remain no greater than π. This makes it clear that if Q_k is convex, then Q_{k+1} is also convex. This proves the convexity of the polygon Q_n.

The side A^0B^0 of the polygon Q_n corresponds to the shortest arc AB, and the broken line $\overline{A^0B^0} = A^0C_1^0 \ldots C_n^0B^0$ corresponds to the geodesic $\overline{AB}$. Moreover, in the sense of the equality of lengths, we have

$$A^0B^0 = AB = a, \quad \overline{A^0B^0} = \overline{AB} = b.$$

The angles α_0 and β_0 at the vertices A^0 and B^0 in the polygon Q_n are not greater than the angles α and β at the vertices A and B of our digon. This is clear from the fact that the angles do not increase in passage from the triangles T_i to T_i^0. Thus,

$$\alpha_0 \le \alpha, \quad \beta_0 \le \beta,$$

and, therefore,

$$\alpha_0 + \beta_0 \le \alpha + \beta = \omega \le \pi. \tag{3}$$

Prolong the sides of the polygon Q_n, which emanate from the vertices A^0 and B^0, as is shown in Fig. 57, b. Then these sides intersect at some point O, since $\alpha_0 + \beta_0 < \pi$. Let c and d be the lengths of the sides A^0O and B^0O of the triangle A^0B^0O. Obviously, by the convexity of the broken line $\overline{A^0B^0}$, we have $A^0O + B^0O \ge \overline{A^0B^0}$, i.e.,

$$c + d \ge b. \tag{4}$$

If γ_0 is the angle at the vertex O of the triangle A^0B^0O, then

$$a \ge (c + d) \sin \frac{\gamma_0}{2}. \tag{5}$$

Indeed, among all triangles with given angle γ_0 and opposite base a the maximum sum of lateral sides is attained at an equilateral triangle. If $2l$ is the sum of the lateral sides of this equilateral triangle then

$$a = 2l \sin \frac{\gamma_0}{2},$$

[12] The angles of a convex polyhedron are no less than the angles of the plane triangle with sides of the same length.

© 2006 by Taylor & Francis Group, LLC

and since $2l \geq c + d$, we obtain (5).

Since $\alpha_0 + \beta_0 + \gamma_0 = \pi$ and $\alpha_0 + \beta_0 \leq \omega$ by (3), we have

$$\sin \frac{\gamma_0}{2} = \cos \frac{\alpha_0 + \beta_0}{2} \geq \cos \frac{\omega}{2}.$$

Using this inequality together with (4), we obtain immediately from (5) that

$$a \geq b \cos \frac{\omega}{2},$$

as required.

(Note that the estimate given in Lemma 2 can be proved exactly in the same way for digons in every manifold with metric of positive curvature if we use the gluing theorem that will be proved in Sec. 1 of Chapter VIII. This estimate is sharp, i.e., for any $a < b$ and $\omega < \pi$, we can construct a digon on a convex polyhedron such that $a - b \cos(\omega/2)$ is arbitrarily small.)

6. Approximation of a Metric of Positive Curvature

Assume that a convex polyhedron P in a manifold of positive curvature is partitioned into small triangles T_i, and, moreover, the sum of each two sides of the triangles T_i is greater than the third. By the theorem in Sec. 6 of Chapter II, this triangulation is possible. To each triangle T_i, we put in correspondence the plane triangle $\overline{T_i}$ with sides of the same length. The triangles $\overline{T_i}$ do not degenerate into segments, since the sum of every two sides is greater than the third. Therefore, the triangles T_i can be mapped homeomorphically onto the corresponding triangles $\overline{T_i}$. Since the sides of the triangles T_i and $\overline{T_i}$ are equal, this mapping can be chosen so that the segments of the sides of the triangles T_i and $\overline{T_i}$ that correspond to each other under this mapping have equal lengths. The sides of the triangles T_i, which do not lie on the boundary of the polygon P, overlap; if a segment of some side of the triangle T_i coincides with a segment of some side of the triangle T_k, then we identify the corresponding segments of sides of the triangles $\overline{T}_j$ and $\overline{T}_k$. As a result, we obtain the development $\overline{P}$ composed of the triangles $\overline{T}_i$, which is homeomorphic to the polygon P. (In this development, the triangles $\overline{T}_i$ adjoin to each other not along the whole sides in general, since so do the triangles $\overline{T}_i$.)

We define the metric $\overline{\rho}(\overline{XY})$ of the development $\overline{P}$ in a natural way. The length of the shortest arc of two points $\overline{X}$ and $\overline{Y}$ is defined to be the distance between these points. The length of a curve in the development is defined automatically, since in each triangle $\overline{T}_i$, the length of a segment of a curve is defined by the fact that this triangle lies in a plane. The shortest curve exists, since the development $\overline{P}$ is composed of a finite number of triangles and hence $\overline{P}$ is compact. (Here, we slightly generalize the concept of development which was introduced in Chapter I, since the sides of the triangles in the development $\overline{P}$ which comprise its boundary (if this boundary exists) are not glued with any sides of other triangles.)

In passage from the triangles T_i to the plane triangles $\overline{T}_i$ with sides of the same lengths, the angles do not increase by Theorem 2 of Sec. 3. The sum of the angles of the triangles T_i at a point inside the polygon P is equal to the complete angle at

© 2006 by Taylor & Francis Group, LLC

this point and hence is not greater than 2π. Therefore, the sums of the angles at the interior vertices of the development $\overline{P}$ are not greater than 2π, either. Consequently, the polyhedral metric $\overline{\rho}$ in the development $\overline{R}$ is a metric of positive curvature.

(It is worth observing that if a development has no boundary, then points of its boundary have no neighborhood that is homeomorphic to a disk, and, therefore, this development is not a manifold. But when generalizing here the concept of a polyhedral metric, we can assume that the polyhedral metric is defined on the development $\overline{P}$. However, this generalization can be avoided if we restrict ourselves only to the interior of the development $\overline{P}$. Indeed, the angles of the convex polygon P do not exceed π. Therefore, the angles at the exterior vertices of the development $\overline{P}$ do not exceed π either, so that in passage from the triangles T_i to the triangles $\overline{T}_i$, the angles do not increase. Hence the development $\overline{P}$ is a polygon with angles not exceeding π. But in such a polygon, the shortest line between its interior points must lie inside this polygon. Therefore, restricting ourselves to only the interior part of $\overline{P}$, we obtain a manifold with a polyhedral metric of positive curvature in which every two points can be connected by a shortest arc.)

Finally, we make one more remark of use in the sequel. Let $\overline{\alpha}_j$ be the angles on the boundary of the development $\overline{P}$, and let $\chi(\overline{P})$ be its Euler characteristic. Then the curvature of the interior of this development is

$$\omega(\overline{P}) = 2\pi\chi(\overline{P}) - \sum(\pi - \overline{\alpha}_j).$$

Since the development $\overline{P}$ and the polygon P are composed of the triangles $\overline{T}_i$ and T_i in the same way, their Euler characteristics are equal; that is, $\chi(\overline{P}) = \chi(P)$. Further, since in passage to plane triangles the angles do not increase, the angles at the boundary of $\overline{P}$ are no greater than the corresponding angles α_j of the polygon P, i.e., $\overline{\alpha}_j \le \alpha_j$. Therefore, this expression of the curvature implies $\omega(\overline{P}) \le \omega(P)$. All what we said above can briefly be summarized in the form of the following lemma.

Lemma 1. *If a convex polygon in a manifold with a metric of positive curvature is partitioned into small convex triangles, then replacing each of these triangles by a plane triangle with sides of the same length, we obtain the polyhedral metric of positive curvature which is defined by the development whose angles on the boundary* $\le \pi$. *The curvature of this development is no greater than the curvature of the polygon.*

Let X and Y be two points of our polygon P, and let $\overline{X}$ and $\overline{Y}$ be those points in the development $\overline{P}$ to which the points X and Y go under the mapping of the triangles T_i onto the triangles $\overline{T}_i$ defined above. Let $\overline{\rho}(\overline{XY})$ be the distance between the points $\overline{X}$ and $\overline{Y}$ measured in the development $\overline{P}$, and let $\rho(XY)$ be the distance between X and Y in the metric under consideration. Our goal is to estimate the difference $\overline{\rho}(\overline{XY}) - \rho(XY)$. Here, we first prove the following:

Lemma 2. *If the diameters of the triangles T_i into which the polygon P is partitioned do not exceed D and the curvature of the interior of the polygon P is equal to ω, then for every pair of points X and Y in P, we have*

$$\overline{\rho}(\overline{X}\,\overline{Y}) < \rho(XY) + (\omega + 2)d. \tag{1}$$

© 2006 by Taylor & Francis Group, LLC

Proof. Take a shortest arc L in the given metric ρ that connects the points X and Y of the polygon P. Since this polygon is convex, we can assume that L traverses some triangles T_i in this polygon; since all these triangles are convex, this arc has only one segment in each of them. Enumerate these segments and the triangles they traverse in the order of their location on the shortest arc L starting from the point X and ending at the point Y; let them be $L_1, L_2, \ldots, L_n$ and $T_1, T_2, \ldots, T_n$.

With the ends of the segment L_i in the triangle T_i, we associate some points in the triangle $\overline{T}_i$, and to the segment L_i, we put in correspondence the line segment $\overline{L}_i$ connecting these points in the triangle $\overline{T}_i$. As a result, with the whole shortest arc L, we associate the broken line $\overline{L}$ in the development $\overline{P}$ which is composed of the segments $\overline{L}_i$ and connects the points $\overline{X}$ and $\overline{Y}$. Let us estimate the length $\overline{s}(\overline{L})$ of this broken line.

Since the sides of the triangles T_i and $\overline{T}_i$ are equal, the diameters of the triangles $\overline{T}_i$ are also no greater than d. Therefore, the length of every segment $\overline{L}_i$ is not greater than d, and, therefore, we certainly have $\overline{s}(\overline{L}_i) < s(L_i) + d$, where $\overline{s}$ and s are the lengths measured in $\overline{P}$ and P, respectively. In particular, for the first and last segments, we have

$$\overline{s}(\overline{L}_1) < s(L_1) + d, \quad \overline{s}(\overline{L}_n) < s(L_n) + d. \tag{2}$$

Consider now other segments (if they exist). The segment L_i enters the triangle T_i at some point A and then goes away from this triangle at some point B. If these points lie on one side of this triangle then L_i reduces to the segment AB of this side, and the segment $\overline{L}_i$ is the segment $\overline{AB}$ of the corresponding side of the triangle $\overline{T}_i$. By condition, the corresponding segments of the sides of the triangles T_i and $\overline{T}_i$ have equal lengths, and hence, in this case

$$\overline{s}(\overline{L}_i) = s(L_i). \tag{3}$$

If the ends A and B of the segments L_i lie on different sides of the triangle T_i, then the ends $\overline{A}$ and $\overline{B}$ of the segment lie on the corresponding sides of the triangle $\overline{T}_i$ and cut off two segments on them whose lengths are the same as the lengths of the segments that the point A and B cut off on the sides of the triangle T_i. Therefore, according to Lemma 1 of the previous section, we have

$$\overline{s}(\overline{L}_i) \le s(L_i) + \omega_i d, \tag{4}$$

where ω_i is the curvature of the triangle T_i.

The lengths $s(L)$ and $\overline{s}(\overline{L})$ of the lines L and $\overline{L}$ are equal to the sums of the lengths of their segments. Therefore, summing the lengths of these segments and using formulas (2), (3), and (4), we obtain

$$\overline{s}(\overline{L}) < s(L) + \left[2 + \sum_{i=1}^{m} \omega_i\right] d. \tag{5}$$

Since the shortest arc L traverses each triangle at least once, each triangle in the sum in inequality (5), i.e., each ω_i appears not more than once. At the same time, by nonnegativity and additivity of curvature, all $\omega_i \ge 0$, and their sum over all

© 2006 by Taylor & Francis Group, LLC

triangles of the partition is no greater than the curvature ω of the polygon P. Consequently,

$$\sum_{i=1}^{n} \omega_i \le \omega. \tag{6}$$

Since L is a shortest arc, we have

$$s(L) = \rho(XY); \tag{7}$$

since the distance $\overline{\rho}(\overline{X}\,\overline{Y})$ is the greatest lower bound of the lengths of curves that connect the points $\overline{X}$ and $\overline{Y}$,

$$\overline{s}(\overline{L}) \ge \rho(\overline{X}\,\overline{Y}).$$

Taking (6), (7), and (8) into account, we obtain from (5) that

$$\overline{\rho}(\overline{X}\,\overline{Y}) < \rho(XY) + (2+\omega)d,$$

as required.

Lemma 3. *Let a convex polygon P, homeomorphic to a disk and having curvature of the interior $\omega < \pi$, be partitioned into small convex triangles so that the diameters of these triangles do not exceed d. Then for every pair of points X and Y in P, we have*

$$|\rho(XY) - \overline{\rho}(\overline{X}\,\overline{Y})| < Cd,$$

where ρ and $\overline{\rho}$ have the same meaning as above and C is a constant.

We can take $C = 14$ for all P satisfying these conditions. In proving, we show that

$$C = 2\omega + \frac{\omega^2}{2} + 2.$$

Proof. Obviously, by the previous lemma, it is sufficient to prove the existence of the constant C such that

$$\rho(XY) < \overline{\rho}(\overline{X}\,\overline{Y}) + Cd.$$

For the proof, we consider a shortest arc $\overline{L}$ in the development $\overline{P}$ which connects the points $\overline{X}$ and $\overline{Y}$. This shortest arc traverses some triangles $\overline{T}_i$ but *a priori* can have more than one segment in each triangle, since it is unknown whether these triangles are convex in the metric $\overline{\rho}$ in the development $\overline{P}$. Therefore, the proof cannot be performed simply by repeating the proof of the previous lemma with the replacement of ρ by $\overline{\rho}$ and vice versa.

Let us draw the shortest arc $\overline{L}$ from the point $\overline{X}$ to the point $\overline{Y}$.

We replace the first segment $\overline{L}_1 = \overline{X}\,\overline{Z}$ of the shortest arc $\overline{L}$ that lies in the triangle $\overline{T}_i$ containing $\overline{X}$, by a shortest arc XZ of the corresponding points X and Z in the triangle T_i. In exactly the same way, we treat the last segment $\overline{L}_n$ of the shortest arc $\overline{L}$. For the shortest arcs L_1 and L_n in the polygon P which correspond to these segments, we have $s(L_1) \le d$ and $s(L_n) \le d$. Therefore,

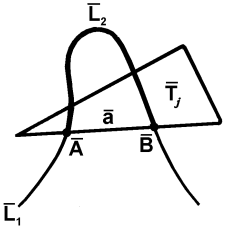
$$s(L_1) < s(\overline{L}_1) + d, \quad s(L_n) < s(\overline{L}_n) + d. \tag{9}$$

© 2006 by Taylor & Francis Group, LLC

Leaving the triangle that contains the point $\overline{X}$, the shortest arc $\overline{L}$ enters into another triangle $\overline{T}_j$. The following three cases are possible here:

(1) $\overline{L}$ goes along a side of the triangle $\overline{T}_j$ and has a common segment $\overline{L}_2$ with this side. Then we take the corresponding segment L_2 of the side of the triangle T_j in L, and for the length, we obtain

$$s(L_2) = \overline{s}(\overline{L}_2). \tag{10}$$



(2) $\overline{L}$ intersects the side $\overline{a}$ of the triangle $\overline{T}_j$ at the point $\overline{A}$, and there are no other intersection points with $\overline{a}$. Then this arc leaves this triangle at some point $\overline{B}$ lying on the other side $\overline{b}$ of this triangle. We replace the segment $\overline{L}_2 = \overline{AB}$ of the shortest arc $\overline{L}$ by the corresponding shortest arc $L_2 = AB$ in the triangle T_j. Then, by Lemma 1 of the preceding section, we have

$$s(L_2) \le \overline{s}(\overline{L}_2) + \omega(T_j)d. \tag{11}$$

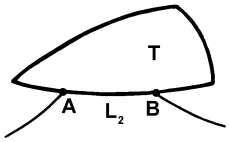


Fig. 68

(3) $\overline{L}$ intersects the side $\overline{a}$ at the point $\overline{A}$ and then again intersects this side at the point $\overline{B}$ (Fig. 68).[13] Then the whole segment $\overline{L}_2$ of the shortest arc $\overline{L}$ between the points $\overline{A}$ and $\overline{B}$ is replaced by the segment $L_2 = AB$ of the corresponding side a in the triangle T_j. Since the sides of the triangles T_j and $\overline{T}_j$ are isometrically mapped onto each other, the length of the segment AB of the side a is equal to the length of the segment $\overline{AB}$ of the side $\overline{a}$; that is,

$$s(L_2) \equiv s(AB) = \overline{s}(\overline{AB}).$$

The segment $\overline{L}_2$ of the shortest arc $\overline{L}$ and the segment $\overline{AB}$ of the side $\overline{a}$ form a simple closed curve, and since the development $\overline{P}$ is homeomorphic to a disk, this curve bounds a domain that is homeomorphic to a disk by the Jordan curve theorem. Hence the segments $\overline{L}_2$ and $\overline{AB}$ bound a geodesic digon; moreover, the segment $\overline{L}_2$ is a shortest arc.

The curvature $\omega(\overline{D})$ of this digon $\overline{D}$ is no greater than the curvature of the whole development $\overline{P}$, and as was shown above, the curvature of the development $\overline{P}$ is no greater than the curvature ω of the polygon P, and since $\omega < \pi$ by assumption, we have $\omega(\overline{D}) < \pi$. Hence we can use Lemma 2 of the preceding section to obtain

$$\overline{s}(\overline{L}_2) \ge \overline{s}(\overline{AB}) \cos \frac{\omega(\overline{D})}{2},$$

or, by (12),

$$\overline{s}(L_2) \le \overline{s}(\overline{L}_2) + \left(1 - \cos \frac{\omega(\overline{D})}{2}\right) s(L_2).$$

[13] Figure 68 is of "topological character," that is, the size is not correct; in fact, $\overline{L}_2$ is shorter than $\overline{AB}$.

© 2006 by Taylor & Francis Group, LLC

Since $s(L_2) = s(AB)$ is obviously no greater than the diameter of the triangle T_j and, thus, no greater than D, we have

$$s(L_2) \le \overline{s}(\overline{L}_2) + \left(1 - \cos\frac{\omega(\overline{d})}{2}\right) d. \tag{13}$$

If the shortest arc $\overline{L}$ intersects the side $\overline{a}$ for the third time at some point $\overline{C}$, then we again obtain the digon with vertices $\overline{B}$ and $\overline{C}$ and proceed in exactly the same way with this digon.

We make these replacements of the segments of the shortest arc $\overline{L}$ by segments $\overline{L}_i$ in the polygon P for all triangles traversed by the shortest arc $\overline{L}$. Moreover, the already replaced segment $\overline{L}_i$ is entirely excluded from consideration under other replacements. As a result, the shortest arc $\overline{L}$ is replaced by a line L in the polygon P which connects the points X and Y. The lengths $\overline{s}(\overline{L})$ and $s(L)$ of the lines $\overline{L}$ and L are equal to the sum of the lengths of their segments. Therefore, summing all these segments and using relations (9), (10), (11), and (13), which are applied to all segments (each in due course), we obtain

$$s(L) < \overline{s}(\overline{L}) + \left[2 + \sum_i \omega(T_i) + \sum_j \left(1 - \cos\frac{\omega(\overline{D}_j)}{2}\right)\right] d. \tag{14}$$

Here, the first sum is taken over all triangles T_i corresponding to those $\overline{T}_i$ inside which we have those segments of the line $\overline{L}$ which do not enter the digons; the second sum is taken over all digons $\overline{D}_j$.

First, we assert that each triangle enters the first sum at most twice. Indeed, let the shortest arc $\overline{L}$ enter the triangle $\overline{T}_i$ at a point $\overline{A}$ on the side $\overline{a}$ (Fig. 69). If this arc intersects this side again then we have a digon. Passing through all such digons with vertices at the side $\overline{a}$, we can enter $\overline{T}_i$ at some point $\overline{B}$ again. But since all digons with vertices at the side $\overline{a}$ are already passed, $\overline{L}$ does not intersect the side $\overline{a}$. When this arc enters the triangle $\overline{T}_i$ at the point $\overline{B}$, it leaves this triangle at some point $\overline{C}$ on the other side $\overline{b}$. The segment $\overline{BC}$ is the first segment lying in the triangle $\overline{T}_i$ and not belonging to any of the digons in question. Further, the shortest arc $\overline{L}$ can form digons with vertices at the side $\overline{b}$. When we have passed all these digons, we can finally intersect the side $\overline{b}$ at some point $\overline{D}$ and enter the triangle $\overline{T}_i$. But we can intersect neither the side $\overline{a}$ nor the side $\overline{b}$, since all digons with vertices at these sides are already considered. Therefore, we have to leave the triangle $\overline{T}_i$ through the third side $\overline{c}$ and intersect this side at some point $\overline{E}$. We cannot ever return to the triangle $\overline{T}_i$, since, traversing this triangle, we have to intersect two of its sides,[14] whereas we do not intersect it between $\overline{a}$ and $\overline{b}$. Consequently, the segment $\overline{D}E$ is the second and last segment of the shortest arc $\overline{L}$ that is not included in any digon.

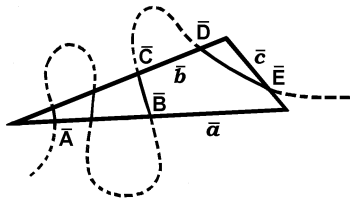


Fig. 69

[14] L_i is a shortest arc, and so its every part T_i is a straight line segment.

© 2006 by Taylor & Francis Group, LLC

This proves that each triangle T_i can enter the first sum in formula (14) no more than two times. Since all curvatures of the triangles T_i are nonnegative and their sum over all triangles is no greater than the curvature ω of the polygon P, we have

$$\sum_i \omega(T_i) \le 2\omega. \tag{15}$$

We now prove that the interiors of the digons $\overline{D}$ are disjoint.

Assume that two digons $\overline{D}_p$ and $\overline{D}_q$ have common interior points. These digons are bounded by the segments $\overline{L}_p$ and $\overline{L}_q$ of the shortest arc $\overline{L}$ and by the segments $\overline{S}_p$ and $\overline{S}_q$ of sides of some triangles $\overline{T}_i$. The segments $\overline{L}_p$ and $\overline{L}_q$ are disjoint, since they are parts of a single shortest arc. The segments $\overline{S}_p$ and $\overline{S}_q$ are disjoint, since the sides of any triangles $\overline{T}_i$ never meet. Since the digons $\overline{D}_p$ and $\overline{D}_q$ have common interior points by assumption, the following two cases are possible: (1) one of these digons lies inside the other, say, $\overline{D}_q$ lies inside $\overline{D}_p$; (2) the boundaries of $\overline{D}_p$ and $\overline{D}_q$ are disjoint. But since the segments $\overline{L}_p$ and $\overline{L}_q$, $\overline{S}_p$ and $\overline{S}_q$ cannot intersect, we see that in the second case, e.g., the segment $\overline{L}_q$ intersects $\overline{S}_p$ and hence enters the digon $\overline{D}_p$.

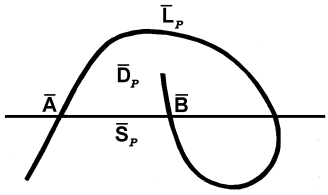


Fig. 70

In the first case when $\overline{D}_q$ lies inside $\overline{D}_p$, the segment $\overline{L}_q$ lies inside $\overline{D}_p$. Hence, in both cases, we can assume that the shortest arc $\overline{L}$ enters the interior of the digon $\overline{D}_p$ (see Fig. 70). Therefore, traveling from the beginning of the segment $\overline{L}_p$ along $\overline{L}$, we will eventually intersect $\overline{S}_p$ and enter the digon $\overline{D}_p$. If $\overline{A}$ is the beginning of $\overline{L}_p$ and $\overline{B}$ is the first intersection point of $\overline{L}$ with $\overline{S}_p$ then we obtain the digon $\overline{D}$ with vertices $\overline{A}$ and $\overline{B}$; one of its sides is the segment $\overline{AB}$ of the shortest arc $\overline{L}$, and the other is the segment $\overline{AB}$ of the line $\overline{S}_p$. Since the shortest arc $\overline{L}$ enters $\overline{D}$ at the point $\overline{B}$, the angle at the vertex $\overline{B}$ in the digon $\overline{D}$ is greater than π. However, this is impossible, since its curvature is not greater than the curvature $\omega(\overline{P})$ of the whole development $\overline{P}$ and $\omega(\overline{P}) \le \omega(P) < \pi$. Hence we have proved that our digons $\overline{D}_j$ have no common interior points.

In this case, the sum of their curvatures does not exceed the curvatures $\omega(\overline{P})$ of the whole development; that is,

$$\sum_i \omega(\overline{D}_j) \le \omega(\overline{P}) \le \omega(P) < \pi.$$

This implies

$$\sum_j \left(1 - \cos\frac{\omega(\overline{D}_j)}{2}\right) = \sum_j 2\sin^2\frac{\omega(\overline{D}_j)}{2}$$

$$\le \sum_j \frac{1}{2}\omega(\overline{D}_j)^2 \le \frac{\omega(\overline{P})}{2}\sum_j \omega(\overline{D}_j) \le \frac{\omega^2}{2}. \tag{16}$$

Now, inserting inequalities (15) and (16) in inequality (14), we obtain

$$s(L) < \overline{s}(\overline{L}) + (2\omega + \frac{\omega^2}{2} + 2)d. \tag{17}$$

© 2006 by Taylor & Francis Group, LLC

Moreover, since $\overline{L}$ is a shortest arc, we have

$$\overline{s}(\overline{L}) = \overline{\rho}(\overline{XY});$$

since the distance $\rho(XY)$ cannot be greater than the length of a curve that connects the points X and Y, we have

$$s(L) \geq \rho(XY).$$

Therefore, inequality (17) implies

$$\rho(XY) < \overline{\rho}(\overline{XY}) + Cd,$$

where

$$C = 2\omega + \frac{\omega^2}{2} + 2.$$

Comparing this with formula (1) of Lemma 1, we see that

$$|\rho(XY) - \overline{\rho}(\overline{XY})| < Cd;$$

as required.

Remark. It seems likely that on refining the above arguments and reasonably choosing the mappings of the triangles T onto the triangles $\overline{T}$, we may obtain the following

$$0 \leq \rho(XY) - \overline{\rho}(\overline{XY}) \leq \omega d$$

which is unimprovable.

7. Realization of a Metric of Positive Curvature Given on a Sphere

Now we prove our main theorem:

Theorem. *Each metric of positive curvature on the sphere can be realized by a closed convex surface.*

Proof. 1. Let a metric $\rho(XY)$ of positive curvature be given on the sphere S. Since the sphere with this metric is a convex polygon in the sense of our general definition, this sphere can be partitioned into arbitrarily small triangles. (Of course, we speak about triangles in the sense of the metric ρ!) We partition these triangles into smaller triangles, and so on. As a result, we obtain a sequence of partitions $R_1, R_2, \ldots$ of the sphere S; moreover, each subsequent partition is a subpartition of the preceding and the maximal diameters $d_1, d_2, \ldots$ of the triangles of these partitions tend to zero.

To each triangle of the partition R_i, we put in correspondence the plane triangle with sides of the same lengths; according to Lemma 1 of the preceding section, we obtain a development $\overline{R}_i$ that is homeomorphic to the sphere and has a polyhedral convex metric. According to the theorem proved in Chapter IV, we can glue a convex closed polyhedron P_i from each development $\overline{R}_i$.

© 2006 by Taylor & Francis Group, LLC

Each triangle of the partition R_i is mapped onto the corresponding triangle of the development $\overline{R}_i$ in such a way that this mapping is isometric on its sides. Therefore, the sphere S turns out to be mapped onto the development $\overline{R}_i$, and, thus, onto the polyhedron P_i glued from this development. We shall denote by $\overline{X}_i$ (or $\overline{A}_i$, etc.) the point on the polyhedron P_i corresponding to a point X (or A, etc.) under this mapping. The overline always signifies the fact that a point lies in the space rather than on the abstract sphere S.

Of course, we can assume that all polyhedra P_i pass through the same point: it suffices to shift each of them appropriately.

2. *Let us prove that we can choose a sequence from polyhedra P_i such that for each point X of the sphere S, the sequence of the corresponding points $\overline{X}_i$ on the polyhedra of the chosen sequence converges.*[15]

We will then prove that the limit points of these sequences $\overline{X}_i$ form a convex surface that realizes our metric.

Since the curvature of the sphere is equal to 4π, according to Lemma 2 of the preceding section, we have the inequality

$$\rho_i(\overline{X}_i\overline{Y}_i) < \rho(XY) + (4\pi + 2)d_i, \tag{1}$$

where ρ_i is the metric on the polyhedron P_i. The spatial distances between points of polyhedron certainly do not exceed these distances on polyhedra. By assumption, the polyhedra P_i pass through one point, and so inequality (1) implies that they all lie in a bounded part of space.

We have countable many vertices $A^1, A^2, \ldots$ of all partitions R_i on the sphere S. With each point A^n on each polyhedron P_i, we associate the point $\overline{A}_i^n$. Since all polyhedra lie in the bounded part of space, we can choose a sequence

$$P_{11}, P_{12}, P_{13}, \ldots$$

from them in which the points corresponding to the point A^1 converge. Further, we can choose another sequence

$$P_{21}, P_{22}, P_{23}, \ldots$$

from this sequence, in which the points corresponding to the point A^2 converge. Proceed with choosing sequences likewise and arrange the diagonal sequence P_{11}, $P_{22}, \ldots$. We thus obtain a sequence of our polyhedra in which the points corresponding to all points A^n converge. We exclude from consideration all polyhedra that do not enter this sequence; to avoid abusing notation, we will denote the polyhedra of this sequence (and together with them, the partitions of the sphere, etc.) in the same way as we have denoted the polyhedra of the initial sequence.

[15] This assertion and its proof are of a very general character; namely, actually we prove the following: assume that a metric $\rho(XY)$ is given on a compact domain S (in this case, on the sphere). Suppose that we have a sequence of surfaces P_i (in our case, polyhedra) that pass through the same point and are homeomorphic to S and, moreover, S is mapped onto each surface P_i. Let the following inequalities hold: $\overline{\rho}_i(\overline{X}_i\overline{Y}_i) < \rho(XY) + \varepsilon_i$, where $\overline{\rho}_i(\overline{X}_i\overline{Y}_i)$ are the distances on P_i between the points corresponding to the points X and Y of the domain S, and $\varepsilon_i \to 0$ as $i \to \infty$. Then we can choose a sequence from the surfaces P_i such that for each point X of the domain S, the sequence of points X_i on the surfaces of the chosen subsequence converges. The limit points $\overline{X}$ form the limit of the chosen sequence of surfaces $\overline{P}_i$.

© 2006 by Taylor & Francis Group, LLC

Therefore, we have the sequence of polyhedra P_i such that the points $\overline{A}_i^n$ corresponding to the points A^n of the sphere S converge to some points $\overline{A}^n$.

Now let X be an arbitrary point on the sphere S, and let $\overline{X}_1, \overline{X}_2, \ldots$ be the points on the polyhedra $P_1, P_2, \ldots$, which correspond to them. Let us prove that these points $\overline{X}_i$ converge to some point $\overline{X}$.

Take an arbitrarily small $\varepsilon > 0$. Since the diameters of triangles of the partition R_i of the sphere S tend to zero, there is a vertex A of some partition such that

$$\rho(XA) < \varepsilon \tag{2}$$

(for simplicity, we omit the superscript at the vertex).

Further, by inequality (1),

$$\rho_i(\overline{X}_i\overline{X}_i) < \rho(XA) + (4\pi + 2)d_i. \tag{3}$$

Since $d_i \to 0$ as $i \to \infty$, there exists N_ε such that

$$(4\pi + 2)d_i\varepsilon$$

for all $i > N_\varepsilon$. Then, by (2) and (3),

$$\rho_i(\overline{X}_i\overline{A}_i) < 2\varepsilon \ (i > N_\varepsilon). \tag{4}$$

We now let $\overline{X}_i, \overline{A}_i$ and $\overline{X}_j, \overline{A}_j$ be the points on the polyhedra P_i and P_j, respectively, which correspond to the points X and A. Let ρ_0 stand for the distance in space. Then we obviously have

$$\rho_0(\overline{X}_i\overline{X}_j) \le \rho_0(\overline{X}_i\overline{A}_i) + \rho_0(\overline{A}_i\overline{A}_j) + \rho_0(\overline{A}_j\overline{X}_j). \tag{5}$$

But the distances in space are no greater than the distances on the polyhedron P_i, i.e., for example $\rho_0(\overline{X}_i\overline{A}_i) \le \rho_i(\overline{X}_i\overline{A}_i)$. Therefore, (5) implies

$$\rho_0(\overline{X}_i\overline{X}_j) \le \rho_i(\overline{X}_i\overline{A}_i) + \rho_j(\overline{X}_j\overline{A}_j) + \rho_0(\overline{A}_i\overline{A}_j). \tag{6}$$

But, by (4),

$$\rho_i(\overline{X}_i\overline{A}_i) + \rho_j(\overline{X}_j\overline{A}_j) < 4\varepsilon \quad (i, j > N). \tag{7}$$

On the other hand, since the points $\overline{A}_k$ converge, $\rho_0(\overline{A}_i\overline{A}_j)$ becomes less than ε whenever i and j are greater than some M_ε. Therefore, inequality (6) implies

$$\rho_0(\overline{X}_i\overline{X}_j) < 5\varepsilon \quad (i, j > \max(M_\varepsilon, N_\varepsilon)).$$

Since ε is arbitrarily small, this means that the points $\overline{X}_i$ converge to some point $\overline{X}$.

3. Thus, all points $\overline{X}_i$ of the polyhedra P_i converge to some points $\overline{X}$. Let F be the set of all these points $\overline{X}$. We now prove that *the polyhedra P_i converge to F*.

Since each point in F is the limit of a sequence of points $\overline{X}_i$ lying on the polyhedra P_i, it suffices to prove only that no point beyond F can be a condensation point of the points lying on distinct polyhedra P_i.

© 2006 by Taylor & Francis Group, LLC

Assume that a point $\overline{B}$ is the limit of points $\overline{C}_i$ lying on polyhedra $\overline{P}_i$. (We take points on all polyhedra for simplicity of notation. If the point $\overline{B}$ is a condensation point but is not the limit of the points $\overline{C}_i$, then we can certainly abstract a convergent sequence from the points $\overline{C}_i$ and restrict consideration only to the corresponding polyhedra.) Let C_i be the points on the sphere S, which correspond to the points $\overline{C}_i$. We can choose a convergent sequence from these points; to keep notation, we will assume that the points C_i converge to some point D. Finally, with the point D we associate the points $\overline{D}_i$ on the polyhedra P_i, and by what we have proved above, $\overline{D}_i$ converge to some point $\overline{D}$ that belongs to F.

Hence we have

$$\overline{C}_i \to \overline{B}, \quad C_i \to D, \quad \overline{D}_i \to \overline{D},$$

where $\overline{C}_i$ correspond to C_i and $\overline{D}_i$ correspond to D.

As above, let ρ_0 stand for the distance in space. Obviously,

$$\rho_0(\overline{BD}) \le \rho(\overline{BC}_i) + \rho_0(\overline{C}_i\overline{D}_i) + \rho_0(\overline{D}_i\overline{D}). \tag{8}$$

Since $\overline{C}_i \to \overline{B}$ and $\overline{D}_i \to \overline{D}$, we have

$$\rho_0(\overline{BC}_i) \to 0, \quad \rho_0(\overline{D}_i\overline{D}) \to 0. \tag{9}$$

Further, since the distance in space is no greater than the distance on the polyhedron P_i, we have

$$\rho_0(\overline{C}_i\overline{D}_i) \le \rho_i(\overline{C}_i\overline{D}_i). \tag{10}$$

Finally, since the points $\overline{D}_i$ and C_i correspond to the points D and C_i on the sphere S, we have

$$\rho_i(\overline{C}_i\overline{D}_i) < \rho(\overline{C}_iD) + (4\pi + 2)d_i$$

by inequality (1), and therefore, inequality (1)) implies

$$\rho_0(\overline{C}_i\overline{D}_i) < \rho(C_iD) + (4\pi + 2)d_i.$$

But $d_i \to 0$ and $C_i \to D$; therefore,

$$\rho_0(\overline{C}_i\overline{D}_i) \to 0. \tag{11}$$

Comparing (8), (9), and (11), we see that $\rho_0(\overline{BD}) = 0$; i.e., the points $\overline{B}$ and $\overline{D}$ coincide. And since the point $\overline{D}$ belongs to F, we infer that $\overline{B}$ belongs to F. Consequently, every condensation point of distinct polyhedra P_i belongs to F, and so we have proved that the polyhedra P_i converge to F.

4. The limit of a bounded convergent sequence of closed convex polyhedra is either a closed convex surface (including doubly-covered plane domains) or a straight line segment or a point (see the Appendix, Sec. 6, Theorem 3). Therefore, F is either a closed convex surface or the segment or a point. For brevity, F will be called a surface.

With each point X of the sphere S, we associate the point $\overline{X}$ on the surface F, i.e., the limit of points $\overline{X}_i$ that correspond to X on the polyhedra P_i. On the other hand, by the definition of the surface F, each point on F is such a limit, and, thus, this point corresponds to some point of the surface S. Consequently, we have the

© 2006 by Taylor & Francis Group, LLC

mapping h of the sphere S onto the surface F; under this mapping, the point X goes to $\overline{X}$.

In the case where F is a doubly-covered domain on the plane, its every interior point is counted twice, i.e., on one and the other side of this domain. This leads to some nonuniqueness in the definition of this mapping of S onto F. Namely, if X goes to $\overline{X}$, then we do not know yet the side on which the point $\overline{X}$ is located. We now show that this nonuniqueness can be eliminated.

Let F lie in some plane Q. Take the positive direction of the normal to the plane Q and orient each line perpendicular to Q in this direction. We call the top or upper side of Q that side from which these lines emanate. The bottom or lower side of Q we call the other side of Q. If a line L that is perpendicular to Q intersects the polyhedron P_i entering this polyhedron at the point $\overline{A}$ and exiting at the point $\overline{B}$, then we assume that $\overline{B}$ lies on the top, while $\overline{A}$ is assumed to lie on the bottom. If L is only tangent to the polyhedron, then we can assume that the tangent points lie either on the top or on the bottom. If a point $\overline{X}$ lies inside F, say, namely on the top (bottom), then a sequence of points $\overline{X}_i$ lying on the polyhedra P_i are assumed to converge to $\overline{X}$ if and only if, for a sufficiently large i, the points $\overline{X}_i$ also lie on the top (bottom) of the polyhedra P_i. If a point $\overline{X}$ lies on the boundary of F, then we impose no additional condition on the convergence of the points $\overline{X}_i$. All these was already specified in Sec. 1 of Chapter III.

As above, let $A^1, A^2, \ldots$ be a sequence of partitions of R_i. If the point $\overline{A}^1$ lies on the boundary of F, then $\overline{A}^1_i$ converge to this point, and we need no additional conditions. If the point $\overline{A}^1$ lies inside F and the points A^1_i lie on various sides for arbitrarily large i, then we choose a sequence

$$P_{11}, P_{12}, P_{13}, \ldots$$

from the polyhedra P_i such that the points $\overline{A}^1_{1k}$ of this sequence lie on the same sides of the polyhedra P_{1k}. We choose a sequence

$$P_{21}, P_{22}, P_{23}, \ldots$$

from this sequence in which the points $\overline{A}^2_{2k}$ lie on the same sides of the polyhedra P_{2k}. If we proceed further in this way and take the diagonal sequence $P_{11}, P_{22}, \ldots$, then all points $\overline{A}^n_{kk}$ in this sequence converge (for each given n and $k \to \infty$) also under the above additional condition. To keep notation, we will assume that the sequence P_{kk} coincides with the initial sequence P_i.

We now prove that for each point X of the sphere S, the points $\overline{X}_i$ on the polyhedra P_i converge to the point $\overline{X}$ on F under our additional condition. This will be proved only in the case where $\overline{X}$ lies inside F.

If $\overline{X}$ lies inside F then we can take a disk of some radius $r > 0$ centered at $\overline{X}$ which lies inside F. We construct the right cylinder C on this disk K (Fig. 71). Take one of the points A^n on the sphere S, which we denote simply by A, such that

$$\rho(XA) < \frac{r}{2},$$

where X is the point on S corresponding to $\overline{X}$.

© 2006 by Taylor & Francis Group, LLC

Since

$$\rho_i(\overline{A}_i\overline{X}_i) < \rho(AX) + (4\pi + 2)d_i$$

by formula (1) and since $d_i \to 0$, we also have

$$\rho_i(\overline{A}_i\overline{X}_i) < \frac{r}{2} \tag{12}$$

for a sufficiently large i. The points $\overline{X}_i$ converge to $\overline{X}$; therefore, for a sufficiently large i, the distance from $\overline{X}_i$ to $\overline{X}$ is at most $r/2$. Then, by (12), the points $\overline{A}_i$ lie at a distance at most r from $\overline{X}$ and, thus, in the cylinder C. Therefore, their limit point $\overline{A}$ lies in the disk K, i.e., inside the plane F.

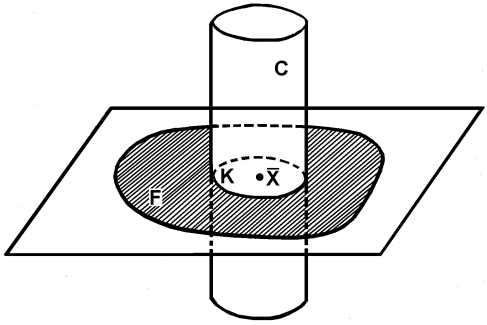


Fig. 71

Assume for definiteness that $\overline{A}$ lies on the top of F. Then the points $\overline{A}_i$ lie on the top of the polyhedra P_i (by our choice of the sequence of polyhedra). Since the cylinder C intersects F along the disk K, which lies inside F, this cylinder intersects the polyhedra P_i for a sufficiently large i in such a way that there are no "extreme" points of the polyhedra P_i, i.e., those points that yield boundary points of the projection of this polyhedron under the projection onto the plane Q.[16]

Therefore, if the points $\overline{X}_i$ lie on the bottom of the polyhedra P_i for arbitrarily large i, then shortest arcs connecting these points with the points $\overline{A}_i$ must pass from bottom to top. Hence they must leave the cylinder C. However, if the point $\overline{X}_i$ is sufficiently close to $\overline{X}$, then each curve that emanates from $\overline{X}_i$ and leaves the cylinder C has length at least $\frac{r}{2}$. Thus, for large i, we have

$$\rho_i(\overline{A}_i\overline{X}_i) > \frac{r}{2}; \tag{13}$$

which contradicts inequality (12). Hence, for sufficiently large i, the points $\overline{X}_i$ stay on the same side of the polyhedra P_i and, thus, converge to $\overline{X}$ and satisfy the above additional condition.

5. Now, when the additional condition holds, our mapping h that sends the points X of the sphere S to the points $\overline{X}$ is defined uniquely. This is a mapping of S onto F. We now prove that this mapping is isometric; i.e., for any X and Y,

$$\rho(XY) = \rho_F(\overline{XY}),$$

[16] If C contains an extreme point $\overline{Z}_i$ of the polyhedron P_i, then there is a supporting plane Q_i of the polyhedron P_i which passes through $\overline{Z}_i$ and is orthogonal to Q. This hyperplane intersects the disk K. Therefore, if the cylinder C contains extreme points of the polyhedra P_i for arbitrarily large i, then the disk K cannot lie inside the limit of the polyhedra P_i, i.e., inside F.

© 2006 by Taylor & Francis Group, LLC

where ρ is the given metric on S and ρ_F is the distance on F. This will demonstrate that *the surface F realizes the metric ρ.*

To prove, we note first that the metrics of the polyhedra P_i converge to the metric of the surface F as follows from Theorem 1 of Sec. 1 of Chapter III. According to this theorem, if the points $\overline{X}_i$ and $\overline{Y}_i$ on the polyhedra P_i converge to the points $\overline{X}$ and $\overline{Y}$ then

$$\rho_F(\overline{XY}) = \lim_{t\to\infty} \rho_i(\overline{X}_i\overline{Y}_i). \tag{14}$$

But, by inequality (1),

$$\rho_i(\overline{X}_i\overline{Y}_i) < \rho(XY) + (4\pi + 2)d_i;$$

and since $d_i \to 0$, using this inequality, we obtain from (14) that

$$\rho_F(\overline{XY}) \le \rho(XY).$$

To prove the reverse inequality, we take a shortest arc $\overline{XY}$ on the surface F and consider one of the partitions R_i of the sphere S. Since the sphere S is mapped onto F, some figure $\overline{T}$ on the surface F corresponds to each triangle T in R_i; this figure is also called a triangle.

Assume that the curvature of the triangle T is less than π. Then we can apply Lemma 3 of the previous section to this triangle. This lemma implies that when we refine the partitions R_i, the metrics ρ_i converge to the given metric ρ in the triangle T. At the same time, the metrics ρ_i are realized on the polyhedra P_i and converge to the metric of the surface F. This implies the following result. Let $\overline{NM}$ be a segment of the shortest arc $\overline{XY}$ that lies in the triangle $\overline{T}$ corresponding to the triangle T. The points N and M on the sphere corresponding to $\overline{N}$ and $\overline{M}$, belong to the triangle T, so that

$$\rho(NM) = \lim_{i\to\infty} \rho_i(\overline{N}_i\overline{M}_i).$$

But, by formula (14),

$$\rho_F(\overline{NM}) = \lim_{i\to\infty} \rho_i(\overline{N}_i\overline{M}_i),$$

and hence

$$\rho(NM) = \rho_F(\overline{NM}). \tag{16}$$

Therefore, when to each segment $\overline{NM}$ of the shortest arc $\overline{XY}$, we put in correspondence the shortest arc NM on the sphere S, we obtain a line whose length is equal to the length of the part of the shortest arc $\overline{XY}$ which lies in the triangle T. (The number of segments $\overline{NM}$ in a single triangle $\overline{T}$ can be infinite. However, this does not change the matter, since the sum of the lengths of infinitely many segments of a curve is equal to the sum of the lengths of these segments.)

However, the shortest arc $\overline{XY}$ can traverse the triangles $\overline{T}$ that correspond to those T whose curvature $\ge \pi$. Let $\overline{XY}$ approach such triangle at the point $\overline{N}$ and leave this triangle at the point $\overline{M}$ (if $\overline{XY}$ intersects $\overline{T}$ many times, then we take the most distant points on this arc which belong to $\overline{T}$). Then to the segment $\overline{NM}$ of our shortest arc, we put in correspondence the shortest arc NM on the sphere

© 2006 by Taylor & Francis Group, LLC

S. Since the diameter of the triangle T in the partition R_i does not exceed d_i, we have

$$\rho(NM) \le d_i. \tag{17}$$

Now, take all shortest arcs NM, which we have in correspondence to the segments of the shortest arc $\overline{XY}$. These arcs form a curve that connects the points X and Y. The length of this curve (in the metric ρ) is $\ge \rho(XY)$.

When combining the segments NM, we meet the segments lying in the triangles T of curvature $\ge \pi$ at most four times. Indeed, the number of such triangles is at most four, since the curvature of the whole sphere is equal to 4π. Hence, summing over all segments NM and using formulas (16) and (17), we obtain

$$\rho(XY) = \sum \rho(NM) \le \sum \rho_F(\overline{NM}) + 4d_i. \tag{18}$$

Relation (18) implies[17]

$$\rho(XY) \le \rho_F(\overline{XY}) + 4d_i.$$

But $d_i \to 0$ as $i \to \infty$; therefore,

$$\rho(XY) \le \rho_F(\overline{XY}). \tag{19}$$

Comparing (19) and (15), we obtain

$$\rho_F(\overline{XY}) = \rho(XY),$$

and, therefore, F realizes the metric ρ.

To complete the proof of the theorem, we have to show also that F is not a point nor a straight line segment. But when we know that F realizes the metric ρ, this is obvious. The isometric correspondence between the sphere with metric ρ and the surface is a homeomorphism, while neither a segment nor a point is homeomorphic to the sphere.

[17] Each segment enters the sum on the right-hand side of inequality (18) at most once, but there are no segments lying in those $\overline{T}$ whose curvature $\ge \pi$. If we add them then we strengthen the inequality, and so $\sum \rho_F(\overline{NM}) = \rho_F(\overline{XY})$.

© 2006 by Taylor & Francis Group, LLC

Chapter VIII

OTHER EXISTENCE THEOREMS

1. The Gluing Theorem

The method for constructing a polyhedral metric from a development composed of plane polygons can be generalized, that is, instead of a development consisting of plane polygons, we can take a set of polygons on manifolds with metric of positive curvature, and identifying their sides, we again obtain a manifold that also has a metric of positive curvature. The gluing theorem, which will be proved here, yields conditions under which this in fact can be done.

By a polygon, we call each closed domain in a manifold with intrinsic metric such that (1) its boundary consists of shortest arcs and (2) no point of the boundary is a condensation point of a set of points lying on some of these shortest arcs. In this definition of polygon, we omit the compactness assumption. Therefore, e.g., an infinite domain on the plane bounded by an infinite broken line or a sector cut out from an open hemisphere are considered now to be polygons. A sector of a hemisphere is open in the direction to the equator but is closed with respect to this hemisphere, i.e., its every limit point belonging to the hemisphere belongs to this sector.

Assume that we cut out polygons $P_1, \dots, P_n$ from manifolds with metric of positive curvature. We shall say that some new intrinsic metric manifold R' originates from the polygons $P_1, \dots, P_n$ by gluing if this manifold can be partitioned into geodesic polygons $P_1' \dots, P_n'$ so that each polygon P_i' is isometric to the polygon P_i.[1] If the manifold R' is a surface in space, then it is natural to say that this surface is obtained by gluing the polygons $P_1, \dots, P_n$.

The sides of the polygons $P_1', \dots, P_n'$ are identified pairwise, and their vertices are also identified somehow. In a natural way, these identifications are extended to the polygons P_i by the isometry of the polygons P_i and P_i'. (It can happen that a whole side of the polygon P_i is divided by vertices of P_i'; then we can assume that the corresponding points are also the vertices of P_i.)

If we perform these identifications for the polygons P_i, then the latter form a manifold that is homeomorphic to R'. In this manifold, the intrinsic metric is defined in a natural way; indeed, we can define the length of a curve in the manifold R as the sum of the lengths of segments of this curve which lie in each of the polygons P_i. After that, the distance between points is defined as the greatest lower bound of the lengths of curves connecting these points. With this definition of metric on the

[1] Its own intrinsic metric is induced inside each polygon P_i, it can differ from the metric that this polygon possesses as a part of the manifold that includes this polygon (see Theorem 5 of Sec. 2 of Chapter II). Here, we speak about the isometry of polygons P_i and P_i' precisely in the sense of their own intrinsic metrics.

© 2006 by Taylor & Francis Group, LLC

manifold R, this manifold is isometric R'; therefore, if R' is an abstract manifold rather than a surface in space, it is possible not to distinguish the manifolds R and R'; also, we say that the manifold R is obtained by gluing the polygons P_i.

A manifold that originates by gluing from given polygons is completely determined by specifying some "law" for identification of its sides and vertices; i.e., by indicating the sides and vertices to be identified and the directions in which these sides will be identified. Recall that these identifications satisfy the following conditions.

1. The angles with a common vertex must meet at this vertex in the same way as sectors dividing a disk at the center of the disk. (This already implies that the sides are identified pairwise.)

2. It is possible to pass from one polygon to the other traversing polygons that have identified sides.

3. The identified sides have equal lengths, and under the identification, their corresponding segments also must have equal lengths.

The first two conditions imply that the polygons form a manifold. Thus, given polygons $P_1, \dots, P_n$ and a law of gluing which satisfies the above three conditions, we obtain a manifold R with an intrinsic metric.

In Sec. 4 of the preceding chapter, we have proved that the complete angle at a point in a manifold with metric of positive curvature cannot exceed 2π. Therefore, for a metric on a manifold R to be of positive curvature, it is necessary that the sum of the angles of the polygons P_i meeting at a common vertex is no greater than 2π. If the polygons P_i are cut out from manifolds with metric of positive curvature or, as we will say, if the polygons P_i have metrics of positive curvature, then this condition turns out to be also sufficient. This is the content of the "gluing theorem". However, before proving this theorem, it is necessary to make an essential stipulation. First, we did not prove yet that *each* polygon in every manifold with metric of positive curvature has definite angles. This was proved only for convex polygons (Sec. 5 of Chapter VIII). Second, we did not prove that the complete angle around *each* point in such a manifold is $\leq 2\pi$. This was proved only for those points about which it is known in advance that their neighborhoods can be divided into convex sectors (Sec. 4 of Chapter VII). Therefore, our assertion on angles of the triangles P_i has a definite sense only for convex polyhedra. But then we will prove that each point in a manifold with metric of positive curvature has a neighborhood isometric to a convex surface. This will imply that all properties of angles which are proved in Chapter IV for the case of a convex surface may be translated to every manifold with metric of positive curvature. Then this will make it clear that there is no necessity to restrict ourselves to convex polygons in the gluing theorem. We prove this theorem in such a way that the convexity of polygons P_i will play no role, except for the fact that *by now* the existence of angles and theorems on their addition are applicable only to convex polygons.

Theorem (Gluing). *If a manifold R is obtained by gluing the polygons $P_1, \dots, P_n$ with metric of positive curvature so that, at each vertex, the sum of angles of these*

© 2006 by Taylor & Francis Group, LLC

polygons meetings at it is no greater than 2π, then the metric on the whole manifold is also a metric of positive curvature.

Proof. A metric of positive curvature is characterized by the fact that this metric is intrinsic and the sum of the lower angles of every sufficiently small convex triangle is $\geq \pi$. The fact that the metric on the manifold R is intrinsic is immediate from its definition.

To prove the second characteristic property, we show that the following properties hold.

1. There always is a definite angle between two sides of a convex triangle in the manifold R.

2. The sum of the angles of a small convex triangle in R is no less than π.

3. There exists an angle in the strong sense between sides of a small convex triangle.

Of course, these three assertions implies the desired result.

If a starting point of two shortest arcs lies inside one of the polyhedra to glue, then there is a definite angle made by these shortest arcs since the metric on this polygons themselves is a metric of positive curvature.

Let a point O lie on the boundary of the polygons, i.e., on a side or at a vertex; assume that the angles of the polygons $P_1, P_2, \ldots, P_m$ meet at this point which are separated by shortest arcs $L_1, L_2, \ldots, L_m$; these shortest arcs are segments of sides of the above polygons. (If O lies on a side, then two "angles" equal to π meet at this point.) Let us show that there is no shortest arc M emanating from the point O and intersecting one of the shortest arcs $L_1, \ldots, L_m$.

Indeed, assume that M intersects L_1 at a point A while traveling from the polygon P_m to the interior of the polygon P_1. Then, replacing the segment OA of the shortest arc M by the segment OA of the shortest arc L_1, we obtain a new shortest arc $\overline{M}$ that lies in the polygon P_1. This shortest arc coincides with L_1 on the part OA and then deviate from L_1. But the metric in the polygon P_1 is a metric of positive curvature, and hence (as was shown in Sec. 3 of Chapter VII) the nonoverlapping condition for shortest arcs holds in this polyhedron which asserts that no overlapping of L_1 and $\overline{M}$ is possible. Hence the shortest arc M cannot intersect L_1. We easily conclude from this that the nonoverlapping condition for shortest arcs holds in the manifold R in general.

Assume that two shortest arcs L and M that are sides of a convex triangle T emanate from the point O. Let $L_1, \ldots, L_k$ be those shortest arcs from the shortest arcs separating the polygons $P_1, \ldots, P_m$ which issue from the point O in the triangle T. For simplicity, we assume that they are enumerated in the order of their location from the shortest arc M to the shortest arc N. There are definite angles $\alpha_0, \alpha_1, \ldots, \alpha_k$ made by pairs of the neighboring shortest arcs M and L_1, L_1 and L_2, $\ldots$, L_k and N, since each of these pairs lies in one polygon P_i.[2] Since M and

[2]If the polygon $P_1(P_m)$ is convex then, cutting off a small convex triangle with vertex O from this polygon, we verify that the shortest arcs $M(N)$ and $L_1(L_k)$ bound a convex sector, and hence there exists an angle between them.

© 2006 by Taylor & Francis Group, LLC

N are sides of the convex triangle T, a shortest arc XY intersects the shortest arcs $L_1, \ldots, L_k$ whenever two points X and Y on them are sufficiently close to O. If this shortest arc passes through the point O, then M and N are continuations of one another, and then, there is an angle made by them, and this angle is equal to π. If the shortest arc XY never passes through the point O, then we can use Theorem 4 in Sec. 1 of Chapter IV, according to which, in this case, the angle made by M and N also exists and is equal to the sum of the angles $\alpha_0, \alpha_1, \ldots, \alpha_k$.

Now let us prove that for each sufficiently small convex triangle, the sum of angles is no less than π. Let T be a small convex triangle. If this triangle is contained in the interior of one of the polygons P_i we glue, then the sum of its angles is $\geq \pi$, since the metric in each P_i is a metric of positive curvature. Assume that the triangle T intersects distinct triangles P_i. Then this triangle is partitioned into polygons $Q_1, \ldots, Q_m$ each of which lies in one polygon P_i, and, therefore, their interiors have nonnegative curvature,[3] i.e.,

$$\omega(Q_k) \geq 0 \quad (k = 1, \ldots, m). \tag{1}$$

Let $A_1, \ldots, A_l$ be the vertices of the polygons Q_k that lie inside the triangle T. Since the complete angle at each of them is $\leq 2\pi$, their curvatures are also nonnegative, i.e.,

$$\omega(A_j) \geq 0 \quad (j = 1, \ldots, l). \tag{2}$$

This is also true for those vertices that are vertices of the polygons P_i, since, by the condition of the theorem, the sum of the angles of these polygons at a single vertex does not exceed 2π.

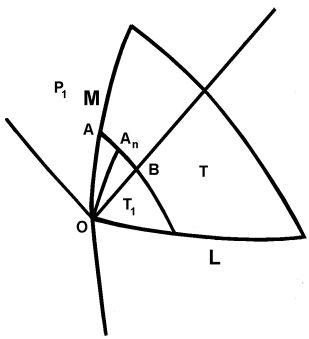


Fig. 72

Applying the theorem on addition of curvatures which was proved in Sec. 1 of Chapter V, we find that the curvature of the interior of the triangle T, i.e., the sum of its angles minus π, is expressed by the formula

$$\omega(T) = \sum_{k=1}^{m} \omega(Q_k) + \sum_{j=1}^{l} \omega(A_j).$$

By inequalities (1) and (2), we have $\omega(T) \geq 0$, as required.

It remains to prove that there is an angle in the strong sense between the sides of a convex triangle in our manifold R. To this end, recall the proof of existence of the angle in the strong sense between two shortest arcs L and M on a convex surface which was presented in Sec. 4 of Chapter III. As was mentioned at the end of Sec. 4 of Chapter III; this proof bases on the following three facts: (1) the existence of an ordinary angle; (2) the nonoverlapping condition for shortest

[3] There are finitely many polygons Q_m, since the nonoverlapping condition for shortest arcs holds in R; by this condition, finitely many of shortest arcs cannot bound infinitely many polygons. Moreover, if the polygons P_i are convex then the polygons Q_m are also convex, since they are intersections of the convex triangle T with the polygons P_i (see Theorem 4 of Sec. 5 of Chapter II).

© 2006 by Taylor & Francis Group, LLC

arcs; (3) the possibility of drawing a shortest arc M' (or L') between two shortest arcs L and M that makes an angle with L (or M) close arbitrarily to the angle made by L and M. But all these three conditions hold in our case. Indeed, these three conditions are satisfied.

1. The existence of an ordinary angle made by sides of a convex triangle is proved.

2. The nonoverlapping condition for shortest arcs is also proved.

3. The possibility of drawing a shortest arc M' that makes an angle with L close to the angle made by L and M is proved as follows. Let L and M emanate from a point O and present sides of a convex triangle T (Fig. 72). Let P_1 be one of the polygons P_i whose part is separated by the shortest arc M participating the triangle T. Cut off from the triangle T a sufficiently small triangle T_1 with vertex O and vertex A inside the shortest arc M. A side of the polygon P_1 cuts off a triangle OAB lying in P_1 from this triangle. Take a sequence of points A_n on the side AB of this triangle which converges to the point A. Then the shortest arcs OA_n converge to OA (see Corollary 4 of Theorem 3 in Chapter II); by Theorem 4 of Sec. 3 of Chapter VII, the angle made by OA_n and OA tends to zero. The additivity of the angle implies that the angle made by OA_n and L tends to the angle made by OA and L, i.e., made by M and L. But this means that it is possible to draw the shortest arc OA_n so that the angle made by OA_n and L is arbitrarily close to the angle made by M and L.

Thus, we can assume that the existence of the angle in the strong sense is proved; which completes the proof of the gluing theorem.

This theorem has many applications. In particular, together with the main realization theorem of the previous chapter, it leads to a number of other theorems on realization of a metric of positive curvature and on the bending of convex surfaces.

2. Application of the Gluing Theorem to the Realization Theorems

Theorem 1. *Each convex polygon, homeomorphic to a disk, on a manifold with metric of positive curvature is isometric to a convex surface.*

Proof. Let P be a convex polygon in a manifold with metric of positive curvature which is homeomorphic to a disk. Take another copy P' of this polygon and identify its sides with the corresponding sides of the polygon P. As a result, we obtain the manifold $P+P'$ that is homeomorphic to the sphere. Since the polygon P is convex, this polygon has definite angles, and all these angles does not exceed π. Therefore, the sum of the angles at each common vertex of the polygons P and P' is no greater than 2π. Hence, by the gluing theorem, the metric on the manifold $P+P'$ is thus a metric of positive curvature; and since this manifold is homeomorphic to the sphere, there exists a convex surface isometric to the manifold $P+P'$ by the main realization theorem of the previous chapter. The part of this surface which corresponds to the polygon P is a convex surface isometric to P.

© 2006 by Taylor & Francis Group, LLC

Theorem 2. *Each point in a manifold with metric of positive curvature has a neighborhood isometric to a convex surface.*

Proof. In Sec. 3 of the previous chapter we have proved that the nonoverlapping condition for shortest arcs holds in a manifold with metric of positive curvature, and in Sec. 4 of Chapter II we have proved that under this condition, each point has a neighborhood that is a convex polygon homeomorphic to the sphere. Applying Theorem 1 to such a neighborhood, we obtain Theorem 2.

Theorem 2 shows that *a manifold with metric of positive curvature is merely a manifold such that* (1) *the metric is intrinsic and* (2) *each point has a neighborhood isometric to a convex surface.* Therefore, *in such a manifold, there hold at least in the small all above results, as well as the results to be obtained in the sequel for intrinsic geometry of convex surfaces.* In particular, this refers to the theorems on angles made by shortest arcs and on angles of sectors. Therefore, it is not necessary to restrict ourselves to convex polygons in the gluing theorem: it the gluing theorem is true for all kinds of polygons to be glued.

The proof of Theorem 1 is a particular application of the following general theorem which can be called the realization theorem of a glued metric:

Theorem 3. *If a manifold R obtained by gluing polygons with metrics of positive curvature is homeomorphic to the sphere and if at each vertex of these polygons, the sum of the contiguous angles does not exceed 2π, then the manifold R is isometric to a closed convex surface. In other words, there exists a closed convex surface that is composed of corresponding parts isometric to the given polygons.*

Indeed, by the gluing theorem, the metric in R is a metric of positive curvature, and since R is homeomorphic to the sphere, R is isometric to a closed convex surface by the realization theorem of the preceding chapter. The theorem on existence of a polyhedron with given development, which is proved in Chapter VI, is obviously a particular case of Theorem 3.

Here is an example of a corollary of Theorem 3.

Theorem 4. *Let a polygon P with metric of positive curvature be homeomorphic to a closed domain on the sphere. This polygon is bounded by closed broken lines $L_1, \dots, L_n$. Assume that on each of these broken lines L_i, all angles of the polygon P, except for one or even two angles that are located so that their vertices divide the broken line L_i in half, do not exceed π. Then there exists a convex surface isometric to the polygon P, more precisely, to its interior.*

Proof. Let $A_1, B_1, \dots, A_n, B_n$ be those vertices of the broken lines $L_1, \dots, L_n$ at which the vertices can be greater than π. By condition, they divide each broken line into two halves. Identifying these halves of each broken line L_i, we transform the polygon P into a manifold R homeomorphic to the sphere. The sums of the angles at all vertices do not exceed 2π, since the angles at all vertices, except for A_i and B_i, do not exceed π.

Consequently, there exists a closed convex surface isometric to R. Performing cuts along the lines corresponding to the broken lines, which are folded in two, in this surface, we obtain a surface that is isometric to the interior of the polygon P.

© 2006 by Taylor & Francis Group, LLC

This surface even has a particular shape of a closed surface with cuts along some broken lines.

We can also obtain surfaces isometric to P but of another shape. If the angles at the complete angles A_i and B_i are equal to α_i and β_i, then we can take a sector S_i that is cut out by meridians from an appropriate surface of revolution and has the angles equal to $2\pi - \alpha_i, 2\pi - \beta_i$ and the perimeter equal to the perimeter of the broken line L_i.[4] Identifying the sides of these sectors with halves of the broken lines L_i, we obtain a manifold that is homeomorphic to the sphere and has a metric of positive curvature again. If we cut out those domains that correspond to the sectors S_i from a closed surface that is isometric to this manifold, then we obtain a surface that is isometric to the polygon P.

Theorem 1 on the existence of a convex surface isometric to a convex polygon is obviously a particular case of Theorem 4.

An analogous application of Theorem 4 allows us to prove the following general realization theorem which was formulated even in Sec. 9 of Chapter I.

Theorem 5. *If a manifold R with metric of positive curvature is homeomorphic to a domain on the sphere and if there is a shortest join of every two points of R then this manifold is isometric to a convex surface.*

The lack of space prevents us from reproducing here the proof of this theorem. We restrict expositions to a sketch of the proof.

Construct a sequence of expanding polygons Q_n in the manifold R whose union covers the whole R. For each polygon Q_n, the following two possibilities are open: (1) Theorem 4 is applicable to this polygon, and hence this polygon can be realized in the form of a closed surface F_n; (2) Theorem 4 is not applicable to the polygon Q_n. But in the second case, we prove that it is possible to glue a plane polygon to each closed broken line lying on the boundary of Q_n so that the conditions of Theorem 2 hold.[5] Then the polygon Q_n is realizable in the form of a domain on a closed surface. Thus, it turns out that for any n, the polygon Q_n is realizable in the form of a certain convex surface F_n. The surface, which is the limit of such surfaces F_n realizes the whole manifold R.

[4] As this sector S_i, we can take a surface that is composed of two right circular cones with complete angles $2\pi - \alpha_i$ and $2\pi - \beta_i$ which is cut along the meridian or, which is the same, the sector S_i can be composed of two plane circular sectors with angles $2\pi - \alpha_i$ and $2\pi - \beta_i$ and with the sum of radii equal to the half-perimeter of the broken line L_i.

[5] Here, the following two theorems on a manifold R that satisfies the conditions of Theorem 5 play an important role:
(A) Every two points in each connected closed domain of R can be connected by a curve shortest in this domain. (If R is not compact then a closed domain of R can fail to be compact either, and, therefore, this theorem in no way issues from the general theorems of Sec. 2 of Chapter II.)
(B) If T is a "triangular domain" in R (i.e., a closed domain bounded by three geodesics that are shortest in the domain R itself), then its angles are no less than the corresponding angles of the plane triangle with sides of the same length. (The domain R can fail to be compact or connected, and hence this assertion is a strong generalization of the familiar theorem on angles of triangles.)
The complement of the interior of the polygon Q_n can consist of several domains. Using Theorem A, it is easy to prove that each of them can be partitioned into triangular domains T. Replacing each domain T by a plane triangle, we obtain a manifold that satisfies the conditions of Theorem 3.

© 2006 by Taylor & Francis Group, LLC

In addition, let us sketch an example of application of Theorem 3 that now refers to the question on the bending of surfaces. *If a convex surface F is obtained from a convex surface by removing a triangle of positive curvature, then there are infinitely many surfaces that are isometric but not equal to F.*

If the triangle ABC, which was removed from our closed convex surface, has a positive curvature, then we can prove that its angles α, β, and γ are strictly less than the angles of the plane triangle with sides of the same length. Therefore, there exist infinitely many distinct convex hexagons $AXBYCZ$ whose sums of sides $AX+XB$, $BY+YC$, and $CZ+ZA$ are equal to the sides of the triangle ABC, and the angles at the vertices A, B, and C are less than α, β, and γ. Identifying sides of such hexagon with the corresponding segments of the boundary of the surface F, we obtain a manifold that satisfies the conditions of Theorem 3. This manifold is realized in the form of a closed convex surface. If we remove a part of this surface that corresponds to the hexagon glued to F, we obtain a surface F' that is isometric to F. Taking various hexagons, we obtain surfaces F' that are isometric to F, but as is easily proved, not equal to each other.

The general principle of applications of the gluing theorem to the problems of realizability and bending becomes to be very clear from the above examples. In what follows, we shall present other applications of this theorem to the problems of another type and indicate (in Sec. 3 of Chapter IX) its wide generalization to the case of gluing domains that are not necessary polygonal.

3. Realizability of a Complete Metric of Positive Curvature

The realizability of a complete metric of positive curvature given on the plane is also proved on the basis of the gluing theorem. The applicability of this theorems becomes to be possible due to the following lemma:

Lemma. *Each polygon on the plane E with a metric of positive curvature can be included into a polygon that is homeomorphic to a disk whose all but possibly one angles do not exceed π.*

Proof. Let Q be a given polygon on the plane E. We can assume that this polygon is homeomorphic to a disk; to this end, it is sufficient to include into this polygons all domains that are surrounded by it. Include the polygon Q into another polygon P in such a way that the distance from the boundary of this polygon to the polygon Q is greater than the perimeter of Q. To prove that this is legitimate, we take the set M of all points whose distance to the polygon Q is no greater than its perimeter q. This set is bounded, and since the metric on the plane E is complete, this set is compact. Each point of the set M can be surrounded by a polygonal neighborhood, and by the Borel lemma, we can refine finitely many these neighborhoods which also cover M. Obviously, these polygonal neighborhoods form a polygon P whose boundary lies from Q at a distance at most the perimeter q of Q. We can assume that this polygon is homeomorphic to a disk; to this end, it suffices to supplement P with the domains it surrounds.

Take a point X on the boundary of the polygon Q and a point Y on the boundary of the polygon P; connect these points by a shortest arc XY. It can happen that along this shortest arc, we encounter other common points with the boundaries of

© 2006 by Taylor & Francis Group, LLC

the polygons Q and P. In any case, we can choose a segment AB from this shortest arc so that the points A and B lie on the boundaries of Q and P and there are no other points of these boundaries on AB. Then the shortest arc AB goes inside the polygon $P - Q$ that is obtained from P by excluding the interior points of Q (see Fig. 73).

If we cut the polygon $P - Q$ along the shortest arc AB, then this polygon transforms into a polygon homeomorphic to a disk. The boundary of this polygon R consists of the boundaries of Q and P that are cut at the points A and B and the two copies of the shortest arc AB; denote by A_1, A_2, and B_1, B_2 the two copies of the points A and B.

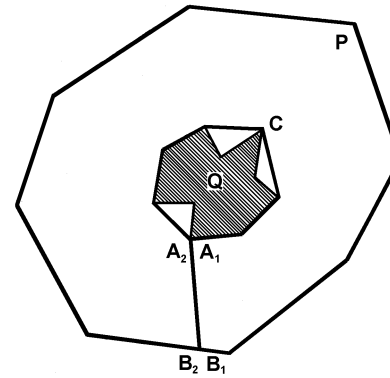


Fig. 73

We have already proved in Chapter II that in each polygon, every two points can be connected by a shortest line in this polygon, and this line is a geodesic broken line with vertices among the vertices of this polygon (Theorem 6 of Sec. 2 of Chapter III). Therefore, there is a geodesic broken line A_1A_2 in the polygon R which is a shortest line in it. The length of this broken line is no greater than the perimeter q of the polygon Q, since the boundary of this polygon also yields a line in R that connects the points A_1 and A_2. By the choice of the polygon P, the distance from its boundary to Q is at least q, and so the broken line A_1A_2 cannot have common points with the boundary of P. Consequently, all vertices of this broken line lies at the vertices of the polygon Q.

Let C be some vertex of the shortest broken line A_1A_2 that is different from A_1 and A_2; the sides of this broken line meeting at this vertex divide the neighborhood under study into two sectors; one of these sectors lies in $P - Q$ (see Fig. 73). If the angle of this sector is less than π, then taking two points X and Y on the sides of this sector which are sufficiently close to C, we have $XY < BX + BY$. Therefore, replacing the part $XB + BY$ of the line A_1A_2 by the shortest arc XY, we shorten it, which is impossible by definition. Consequently, the angles between the sides of the broken line A_1A_2 to the side of the polygon $P - Q$ are no less than π. Other angles opening on the polygon Q are in contract no greater than π.

But the broken line A_1A_2 is a closed broken line on the plane E: its endpoints are the same point A. This closed broken line bounds some polygon Q' that is homeomorphic to a disk and contains Q; all angles of this polygon, except for the angle at the vertex A, are the angles between the sides of the broken line A_1A_2 which open on Q. Therefore, all angles of the polygon Q', but possibly the angle A, does not exceed π. The lemma is proved.

Theorem. *A complete metric of positive curvature on the plane is realizable by an infinite complete convex surface.*

Proof. Assume that a metric of positive curvature is given on the plane E. Construct a sequence of polygons $Q_1, Q_2, \ldots$ on E in such a way that $Q_1 \subset Q_2 \subset \ldots \subset$

© 2006 by Taylor & Francis Group, LLC

$Q_n \subset Q_{n+1} \subset \ldots$ and the family of Q_i covers E. For example, this sequence can be obtained in the following way. The plane E can be mapped onto the Euclidean plane; then with the sequence of the concentric circles of radii 1, 2, ..., n, ..., we associate certain curves on E. Geodesic broken lines with sufficiently small sides, inscribed in these curves, bound the polygons with the required properties.

According to the lemma proved above, each of the polygons Q_n can be included into a polygon P_n homeomorphic to a disk whose all but possibly one angles are at least π. Let A_n be exactly the vertex of the polygon P_n the angle at which is greater than π; if there is no such vertex, then we take an arbitrary vertice of the polygon P_n as A_n. Let B_n be a point on the boundary of P_n such that both segments A_nB_n of the boundary of P_n have the same length. Identify these segments so that the identified points cut out from them some segments that begin at the point A_n and have the same length. As a result, we obtain a manifold R_n that is homeomorphic to the sphere. The sum of the angles at the identified points is $\leq 2\pi$, since the the angles at all vertices, but A_n, are $\leq \pi$. Therefore, we can use the realization theorem for a"glued metric" which was proved in the preceding section. By this theorem, there exists a closed convex surface isometric to the manifold R_n. If we perform a cut along the line corresponding to the boundary of the polygon P_n on this surface, then we obtain a surface that is isometric to the interior of the polygon P_n.[6] Thus, for each polygon P_n, there exists a closed convex surface F_n with a cut which is isometric to this polygon.

By the isometry of the mapping of P_n onto F_n, for each point X (or A, etc.) in P_n, there is the respective point on the surface F_n, which will be denoted by $\overline{X}$ (or $\overline{A}$, etc.). Since the polygons P_n being expanded eventually cover the whole plane E, for each point X of the plane E, there is n_0 such that for all $n > n_0$, we have the corresponding point $\overline{X}_n$. In the sequel, when we speak about points of the plane E and the corresponding points of the surfaces F_n, we do not remind that $\overline{X}_n$ exists for a sufficiently large n.

Take a point A that belongs to all polygons P_n (such a point exists, since $P_n \supset Q_n$ and all Q_n contain Q_1). Then there exist points $\overline{A}_n$ on all surfaces F_n which correspond to A. Translate the surface F_n so that all their points $\overline{A}_n$ coincide. Denote by A the point common to all surfaces F_n; this point corresponds to A on each of these surfaces.

Now let X be an arbitrary point of the plane E, and let $\overline{X}_n$ be the corresponding point on F_n. Since the surface F_n realizes the metric ρ_{F_n} of the polygon P_n, we have $\rho_{P_n}(AX) = \rho_{F_n}(\overline{A}\,\overline{X}_n)$. But the distance in the space is always no greater than the distance on a surface. Therefore, denoting by ρ_0 the distance in the space, for each point X and for all n, we have

$$\rho_0(\overline{A}\,\overline{X}_n) \leq \rho_{F_n}(AX). \tag{1}$$

Now take a countable everywhere dense set of points $A^1, A^2, \ldots$ on the plane E. Formula (1) implies that for any given k, the points $\overline{A}^k_n$ lie in a bounded part of the space; therefore, we can choose a sequence from the surfaces F_n such that the points corresponding to the point A^1 converge. We choose a sequence from

[6] Under the condition that we take the metric in P_n induced by the metric given on the whole plane E.

© 2006 by Taylor & Francis Group, LLC

this sequence such that the points corresponding to A^2 converge, etc. Then in the diagonal sequence, the points corresponding to all A^k converge. The surfaces of this sequence and the corresponding polygons of the plane E will also be denoted by F_n and P_n.

Exactly in the same way as in item 2 of the proof of the realization theorem for a metric on the sphere (Sec. 7 of Chapter VII), we can show that the points $\overline{X}_n$ on the surfaces $\overline{F}_n$ which correspond to the same point X of the plane E converge to a certain point $\overline{X}$. It suffices to take the surfaces F_n instead of the polyhedra in Sec. 7 of Chapter VII, and instead of inequality (1) therein to use inequality (1) of the present section. Let us show that the set F of all these points $\overline{X}$ is a convex surface that realizes a given metric.

Since the limit of convex surfaces is the boundary of a convex set (see the Appendix, Sec. 6, Theorem 3), the set F is a complete convex surface possibly degenerate (a point, a line, etc.). However, if we prove that the distance between every two points X, Y on the plane E is equal to the distance between the corresponding points $\overline{X}, \overline{Y}$ of the set F, then this will certainly imply that F is in fact a convex surface, and this surface realizes a metric on the plane E.

Therefore, let X and Y be arbitrary points of the plane E. Since the polygons P_n expand and cover the whole plane E, there exists n_0 such that for all $n > n_0$, the doubled sum of the distances from points X and Y to the point A chosen above and common to all P_n is less than the distances from the points X and Y to the boundary of the polygon P_n Then the distance between X and Y themselves is less than the distance from them to the boundary of P_n, and the shortest arc XY connecting these points is contained in P_n. Therefore, the distance between X and Y measured in P_n and in the whole plane E are equal; that is,

$$\rho_{F_n}(XY) = \rho_F(XY). \tag{2}$$

In what follows, we restrict consideration to only those n which are greater than n_0.

To the closed surface F_n with a cut which is isometric to the polyhedron P_n, we put in correspondence the closed surface $\overline{F}_n$ without cut. Since the distance $\rho_{P_n}(XY)$ is less than the distances from the points X and Y to the boundary of the polygon P_n and since the surface F_n is isometric to P_n, the distances $\rho_{F_n}(\overline{X}_n\overline{A}_n)$ and $\rho_{F_n}(\overline{Y}_n\overline{A}_n)$ on the surface F_n are less than the distances from the points $\overline{X}_n$, $\overline{Y}_n$, and $\overline{A}_n$ to the cut. This implies that the distance between the points $\overline{X}_n$ and $\overline{Y}_n$ measured on the surface $\overline{F}_n$ is the same as on F_n, and so from Eq. (2), we infer

$$\rho_{\overline{F}_n}(\overline{X}_n\overline{Y}_n) = \rho_E(XY), \tag{3}$$

and, exactly in the same way,

$$\rho_{\overline{F}_n}(\overline{X}_n\overline{A}_n) = \rho_E(AX).$$

Circumscribe the ball S centered at the point $\overline{A}$ of radius

$$R > 2[\rho_F(AX) + \rho_E(AY)].$$

Then all points $\overline{X}_n$ and $\overline{Y}_n$ are contained in this ball and even in the ball of half-radius.

© 2006 by Taylor & Francis Group, LLC

The common part of the ball S and the convex body that is bounded by the surface $\overline{F}_n$ is a convex body whose boundary is denoted by F_n^*. Since the points $\overline{X}_n$ and $\overline{Y}_n$ are contained in the ball of half-radius, the distance between them measured on F_n^* is the same as on $\overline{F}_n$. Therefore, by Eq. (3),

$$\rho_{F_n^*}(\overline{XY}) = \rho_E(XY).$$

The convex surfaces F_n^* converge to a part of the surface F limit for $\overline{F}_n$ which lies in the ball S, and the points $\overline{X}_n$ and $\overline{Y}_n$ converge to the points $\overline{X}$ and $\overline{Y}$ on F. Therefore, by the theorem on the convergence of metrics,

$$\rho_F(\overline{XY}) = \lim_{n\to\infty} \rho_{F_n^*}(\overline{X}_n\overline{Y}_n),$$

and in view of (4),

$$\rho_F(\overline{XY}) = \rho_E(XY), \tag{5}$$

i.e., the distance on the surface F is equal to the distance on the plane E, as required.

If the surface F degenerates into a point or a line then the same result holds, since the theorem on convergence of metrics is also true in this case. But since F realizes the metric on the plane E, it follows that F cannot be a point or a line. *A priori,* the set F, presenting the limit of convex surfaces, can be composed of two parallel planes. However, this is impossible since taking the points $\overline{X}$ and $\overline{Y}$ on distinct planes and taking the corresponding points X and Y on the plane E, we would arrive at a contradiction to Eq. (5).

4. Manifolds on Which a Metric of Positive Curvature Can Be Given

In connection with the theorems on realization of a metric of positive curvature, the question naturally arises of those manifolds that admit a metric of positive curvature in principle i.e., the question about the topological types of manifolds with such a metric.[7] A metric of positive curvature will be called proper if this metric is not reduced to the Euclidean metric at least in a neighborhood of one point. In Sec. 6 of Chapter V, we have proved that for a metric on a domain G on a convex surface to be locally Euclidean, it is necessary and sufficient that the curvature of the domain G vanishes. Since each point in a manifold with metric of positive curvature has a neighborhood isometric to a convex surface by Theorem 2

[7] The question of the topological structure of manifolds admitting a metric that is defined by a line element of some type attracted attention of researchers in geometry for a long time. In the case of closed manifolds, this question was easily solved using the Gauss-Bonnet theorem; here, we merely repeat the classical inference, expressing it in our more general terms. The case of a complete metric of positive curvature, and in particular, the case of a locally Euclidean were considered by Klein and Killing in 1890. The case of a complete metric of positive curvature was considered by Cohn-Vossen; in our inherence, we use the Cohn-Vossen method; see S. Cohn-Vossen, Kürzeste Wege und Total-krümmung auf Flächen. Compositio mathematica, Vol. 2(1935), pp. 69–133. The classical results, stemming from Klein and Killing, are presented, e.g., in the following books: N. V. Efimov, Higher Geometry, Sec. Sec. 87–104 and E. Cartan, Geometry of Riemannian Space, Chap. III.

© 2006 by Taylor & Francis Group, LLC

of Sec. 1, an analogous result holds in each such a manifold R; in particular, *each point of the manifold R has a neighborhood isometric to a part of the Euclidean plane if and only if the curvature of the whole manifold R vanishes.*

The question of topological types of closed, i.e. compact, manifolds with metric of positive curvature is solved by the following theorem.

Theorem 1. *A proper metric of positive curvature can be given on two and only two closed manifolds: on the sphere and the projective plane. A locally Euclidean metric can be given only on the torus and the "Klein bottle" (or a nonorientable torus).*

Proof. Let R be a closed, i.e., compact, manifold with metric of positive curvature, and let $\chi(R)$ be the Euler characteristic of this manifold. According to the result of Sec. 5 of Chapter VII, the curvature of such a manifold is nonnegative and equal to

$$\omega(R) = 2\pi\chi(R).$$

Since $\omega(R) \geq 0$, therefore, $\chi(R) \geq 0$, and hence a metric of positive curvature can be given only on manifolds with nonnegative Euler characteristic. If $\chi(R) = 0$ then $\omega(R) = 0$, and as we have just proved the metric of the manifold R turns out to be locally Euclidean. Consequently, a proper metric of positive curvature can be given only on manifolds of positive Euler characteristic. But it is known that there are only two such manifolds: the sphere and the projective plane.

We can define a metric of positive curvature on each of these manifolds. A metric of the sphere is determined by each closed convex surface. A metric of positive curvature on the projective plane is obtained by identification of antipodes of each closed convex surface with a center of symmetry.

Only two manifolds, the torus and the Klein bottle, have Euler characteristic zero. Therefore, a locally Euclidean metric can be given only on these manifolds. As is known, the torus and the Klein bottle are obtained from a rectangle by pairwise identification of their opposite sides. Moreover, the Klein bottle appears if we orient two opposite sides in opposite directions, so that when we glue these sides, the rectangle is twisted. A rectangle with sides that are pairwise identified is nothing else but a development; it is this development that defines a Euclidean metric on the torus and the Klein bottle, respectively.

The matter is absolutely different for nonclosed manifolds; namely, the following theorem holds:

Theorem 2. *Each nonclosed manifold that is obtained from a closed manifold by removing a certain set of points admits a proper as well as an improper metric of positive curvature.*

For example, the plane and the cylinder are obtained by removing one and two points from the sphere, respectively. Of course, not every manifold can be so-obtained from a closed manifold. It seems to be not improbable however that a metric of positive curvature can be given on an arbitrary nonclosed manifold. However, we are not aware of any proof or disproof of this general assertion.

© 2006 by Taylor & Francis Group, LLC

Proof of Theorem 2. Let a manifold R be obtained from a closed manifold $\overline{R}$ by removing certain points. If $\overline{R}$ is the sphere or projective plane, then we can define a metric of positive curvature on $\overline{R}$ that induces the metric on R, which coincides with it in a small neighborhood of each point (by Theorem 5 in Sec. 2 of Chapter II). This metric is a proper metric of positive curvature. If we remove only one point from the sphere, then from a topological viewpoint, we obtain the plane; of course, we can define the Euclidean metric on the plane. If we remove one point from the projective plane, then we obtain the Möbius band without boundary. We can define the Euclidean metric on the Möbius band, since it is well known that the Möbius band originates by gluing without dilations and contractions from a plane rectangle. Now, if we remove new points from the sphere or projective plane, then the Euclidean metric that results after removing one point induces the metric on the obtained manifold which is also locally Euclidean.

Now let the manifold R be obtained by removing points from a closed manifold $\overline{R}$ that is different from the sphere or projective plane. As is known, such a manifold $\overline{R}$ can be cut by lines emanating from one point so that we obtain a figure that is homeomorphic to a plane polygon.[8] This plane polygon P can in turn be mapped onto $\overline{R}$ so that all its vertices arrive at a single point A. Then a metric that is Euclidean everywhere, but, possibly the point A, is given on $\overline{R}$. But when precisely this point A is removed from $\overline{R}$, we thus obtain the metric on R, which is locally Euclidean. It should be noted here that under the mapping of the polygon P onto $\overline{R}$, the sides of P are glued pairwise. The metric in a neighborhood of every point on the sides glued pairwise is Euclidean exactly as the metric in a neighborhood of a point on an edge of a polyhedron.

The plane polygon P can be replaced by a disk with points on its circle which play the role of vertices. Draw a hemisphere that is based on this circle and map this hemisphere onto the manifold $\overline{R}$ so that the "vertices" arrive at one of the removed points. Then we obtain a proper metric of positive curvature on R, and R is locally isometric to a sphere. Obviously, in a neighborhood of every point on the arcs of the circle we glue, this metric is the same. Indeed, a neighborhood of this point is composed of two halves in the same way as a neighborhood of a point on the equator is composed of pieces of hemispheres. This completes the proof of the theorem.

Applying the same arguments, we can define a rather arbitrary metric on a nonclosed manifold R. However, it is easily verified that our construction leads to incomplete metrics. Each metric is complete on a closed manifold, but for nonclosed manifolds, the question makes sense of manifolds admitting a *complete* metric of positive curvature. An answer to this question is given by the following theorem.

Theorem 3. *The plane is a unique nonclosed manifold admitting a proper complete metric of positive curvature. A complete locally Euclidean metric that can also be given only on the orientable cylinder or the nonorientable cylinder* (*i.e., on the Möbius band without boundary*).

We sketch the proof of this theorem.

[8] See, e.g., H. Seifert and W. Threlfall, Topologie, Chapter VI.

© 2006 by Taylor & Francis Group, LLC

Let R be a nonclosed manifold with a complete metric of positive curvature. Construct an expanding sequences of polygons Q_n on this manifold that cover the whole R, i.e.,

$$Q_1 \subset Q_2 \subset Q_3 \subset \dots \quad \text{and} \quad \sum_{n=1}^{\infty} = R.$$

Take one of the polygons Q_n, and, given a certain positive number r, consider all polygons containing Q_n such that each closed broken line in their boundary lies from Q_n at a distance no greater than r. The completeness of the metric implies that there is a polygon P_n with the minimal perimeter among these polygons. Since it is not known in advance that the manifold R is a plane, the polygon P_n can fail to be simply connected and its boundary can consist of more than one closed broken line.

Let $L_1, L_2, \dots, L_{k_n}$ be closed broken lines that bound the polygon P_n. The distance from each of them to Q_n is at most r. Using elementary arguments that are similar to those of Sec. 4 of Chapter II or in Sec. 3 of the present chapter, and considering the shortest curve that surrounds a given domain, it is easy to prove the following: if a broken line L_i has common points with the boundary of the polygon Q_n, then all its vertices lie at the vertices of Q_n, and the angles of the polygon P_n at these vertices do not exceed π. If the broken line L_i is disjoined from the boundary of Q_n, then this line is a geodesic loop with a unique angle at a point A_i whose distance from Q_n is exactly equal to r. The angle of the polygon P_n at the point A_i is also $\geq \pi$. The only exception is the case in which L_i is a closed geodesic. Thus, on each broken line L_i, all angles but, possibly, that at the point A_i do not exceed π.[9]

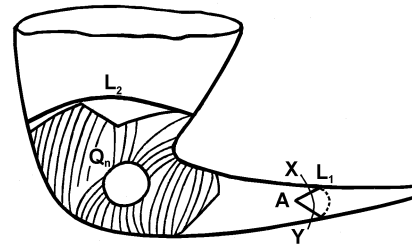


Fig. 74

We now increase the given number r. Then the set of polygons among which we choose the polygon P_n of minimal perimeter is expanded. Therefore, the minimum of their perimeters, i.e, the perimeter of the polygon P_n is a nonincreasing function $p(r)$ of r. Given some $\varepsilon > 0$, we assume that for a given r, the polygon P_n has the angle $\alpha_i > \pi + 2\pi\varepsilon$ at some point A_i (Fig. 74). Take two points X and Y equidistant from A_i on the sides of the broken line L_i which meet at the vertex A_i. Replacing the part $XA_i + A_iY$ of the broken line L_i by the shortest arc XY, we replace the broken line L_i by another broken line L_i^*, and the polygon P_n is replaced by the polygon P_n^*. Since the distance from the point A_i to the polygon Q_n is no greater than r, the distance from the broken line L_i^* to Q_n is no greater than $r + \Delta r = r + A_iX$. Since $p(r + \Delta r)$ is the minimum of the perimeters of the

[9] It is visual in Fig. 74 that the line L_1 tending to shorten must depart from Q_n, but since its distance from Q_n cannot be greater than r, this line tightens and forms the angle $> \pi$ at the point A_1. The line L_2 tending to shorten tightens on the boundary of Q_n.

© 2006 by Taylor & Francis Group, LLC

polygons bounded by broken lines distant from Q_n by no more than $r + \Delta r$, we see that $p(r + \Delta r)$ does not exceed the perimeter $p(P_n^*)$ of the polygon P_n^*, that is,

$$p(r + \Delta r) \leq p(P_n^*) \quad (\Delta r = A_i X). \tag{1}$$

We now estimate the quantity by which the perimeter $p(P^*)$ is less than the perimeter $p(P_n) = p(r)$ of the polygon P_n. Since the complete angle at the point A_i does not exceed 2π, the angle β at the vertex A_i in the triangle A_iXY is not greater than $2\pi - \alpha_i$,[10] and since $\alpha_i > \pi + 2\pi\varepsilon$, we have

$$\beta < \pi - 2\pi\varepsilon. \tag{2}$$

If β_0 is the corresponding angle in the plane triangle with sides of the same length, then $\beta_0 \leq \beta$, and hence, the more so,

$$\beta_0 < \pi - 2\pi\varepsilon. \tag{3}$$

The base XY of the isosceles triangle is related to the length of its lateral side $A_iX = A_iY$ by the equation

$$XY = 2A_i \sin \frac{\beta_0}{2}, \tag{4}$$

and hence

$$XY < 2A_iX \sin \frac{\pi - 2\pi\varepsilon}{2} = 2A \cos \pi\varepsilon.$$

Therefore, the difference of the lengths of the broken lines L_i and L_i^* is

$$A_iX + A_iY - XY = 2A_iX - XY > 2A_i(1 - \cos \pi\varepsilon). \tag{5}$$

But the difference of the lengths of the broken lines L_i and L_i^* is exactly the difference of the perimeters of the polygons P_n and P_n^*, so that

$$p(P_n) - p(P_n^*) > 2A_i(1 - \cos \pi\varepsilon). \tag{6}$$

But since

$$p(P_n) = p(r), \quad p(P_m^*) \leq p(r + \Delta r), \quad \text{and} \quad A_iX = \Delta r,$$

we have

$$p(r + \Delta r) - p(r) < -2\Delta r(1 - \cos \pi\varepsilon). \tag{7}$$

By integration, we may conclude from this that if for all r, there are angles $\alpha_i > \pi + \varepsilon$ on the broken lines L_i, then, for $r = r_1$, we have

$$p(r_1) < p(r) - 2(r_1 - r)(1 - \cos \pi\varepsilon).$$

Taking a sufficiently large r_1, we obtain $p(r_1) < 0$, which is impossible. Hence, there always exists r such that on each of the broken lines L_i bounding the polygon P_n the only angle greater than π is still less than $\pi + 2\pi\varepsilon$.

[10] If X and Y are sufficiently close to A_i, then the angle β is exactly the angle complementing α_i to the complete angle at A_i since $\alpha_i > \pi$.

© 2006 by Taylor & Francis Group, LLC

Let P_n be precisely this polygon. Let k be the number of the broken lines L_i, bounding this polygon, and let χ be its Euler characteristic. Finally, let α_i be those angles of this polygon that are greater than π, and let γ_i be those angles that are no greater π. We have the following formula for the curvature of the polygon P_n:

$$\omega(P_n) = 2\pi\chi - \sum_i(\pi - \alpha_i) - \sum_j(\pi - \gamma_j). \tag{8}$$

Since $\gamma_j \leq \pi$, $\alpha_i < \pi + 2\pi\varepsilon$ and the number of angles α_i does not exceed k, we have

$$\omega(P_n) \leq 2\pi\chi + 2\pi\varepsilon k. \tag{9}$$

If we (topologically) identify each broken line L_i with the circle of a plane disk, then the polygon P_n turns out to be completed up to a closed manifold, and its Euler characteristic increases by k. Therefore, if χ_0 is the Euler characteristic of this closed manifold, then

$$\chi = \chi_0 - k. \tag{10}$$

Substituting this expression for χ on formula (9), we obtain

$$\omega(P_n) \leq 2\pi\chi_0 - 2\pi k(1 - \varepsilon). \tag{11}$$

Since the metric of positive curvature is given in the initial manifold R, we have $\omega(P_n) \geq 0$, and hence

$$\chi_0 \geq k(1 - \varepsilon). \tag{12}$$

And since $k \geq 1$ and ε is arbitrarily small, we have $\chi_0 \geq 1$. There exist only two closed manifold with Euler characteristic $\chi_0 \geq 1$: the sphere for which $\chi_0 = 2$ and the projective plane for which $\gamma_0 = 1$. Therefore formula (12) opens up only the following possibilities: (1) $\chi_0 = 2, k = 1$, and the polygon P_n is homeomorphic to a disk; (2) $\chi_0 = 2, k = 2$, and P_n is homeomorphic to an annulus; (3) $\chi_0 = 1, k = 1$, and P_n is homeomorphic to the Möbius band.

If a proper metric of positive curvature is given in R, then for a sufficiently large n, we have

$$\omega(P_n) \geq \omega(Q_n) > 0. \tag{13}$$

In the last two cases $\chi_0 = k$, and, so (11) implies

$$\omega(P_n) \leq 2\pi k\varepsilon.$$

But $k \leq 2$, and taking ε so small that $4\pi\varepsilon < \omega(Q_n)$, we obtain $\omega(P_n) < \omega(Q_n)$; which contradicts (13). Hence, for sufficiently large n, the polygon P_n is homeomorphic to a disk, and, therefore, the manifold R itself is homeomorphic to a plane.

If the metric in R is locally Euclidean, then all three possibilities are admissible for the polygons P_n, and in accordance with this fact, the manifold R can be homeomorphic to a plane or to a cylinder or to a nonorientable cylinder.

© 2006 by Taylor & Francis Group, LLC

Thus, the question of manifolds admitting a metric of positive curvature is solved.[11]

A more general question consists in revealing those manifolds admitting a metric of positive curvature such that every two points of such a manifold can be connected by a shortest arc. We can answer this question not for arbitrary manifolds but only for manifolds homeomorphic to domains on closed manifolds. However, it is worth supposing that this restriction is not obligatory and the result of the theorem to be formulated remains true without this restriction.

Theorem 4. *Let a metric ρ of positive curvature be given on a manifold R homeomorphic to a domain on some closed manifold so that every two points of the manifold R can be connected by a shortest arc. Then, if the metric ρ is a proper metric of positive curvature, then the manifold R is homeomorphic either to the sphere, or to the projective plane, or to the interior of a disk (i.e., the ordinary plane), or can be completed to a manifold of one of these types by adjoins some of points M; moreover, the completed manifold also has a metric of positive curvature and each point of M has the property that there are no shortest arc passing through this point.*

If the metric ρ is locally Euclidean, then the manifold R is homeomorphic to one of the following manifolds: (1) *the plane;* (2) *the cylinder;* (3) *the nonorientable cylinder;* (4) *the torus;* (5) *the Klein bottle, or it can be completed to a manifold homeomorphic to the sphere, the plane, or the projective plane by adding a certain set of points M; moreover, the completed manifold has a metric of positive curvature and each points of the set M has the property that there are no shortest arc passing through this point.* (An example is given by a convex polyhedron without vertices; the metric is Euclidean everywhere in this polyhedron.) In short, to within the set M, such that no shortest join passes the point of M the manifold can be only one of those admitting on a metric of positive curvature.

The proof of this theorem cannot be given here for lack of space; it proceeds on using the gluing theorem, and, certainly, we can *a priori* assume that the manifold R is not closed. We then construct an expanding sequence of polygons in this manifold whose union covers R. It proves that "holes" in each polygons Q_n can be glued by plane polygons so that we obtain a closed manifold with a metric of positive

[11] Another proof. Let R be a nonclosed manifold with a proper complete metric ρ of positive curvature, i.e., $\omega(R) > 0$. As is known from the topology of manifolds, each nonclosed manifold has the plane as a covering manifold; moreover, the manifold is covered infinitely many times under the covering mapping whenever this manifold is not the plane (see H. Seifert and W. Threlfall, Topologie, Chapter VIII). If the plane E covers R, then the metric ρ_1 is induced on this plane in the following way: as $\rho_1(XY)$ $(X, Y \in E)$, we take the greatest lower bound of the lengths of those curves in L which are the images of curves connecting X and Y under the covering mapping. With this definition of the metric ρ_1, the respective points X and X_0 in E and R have isometric neighborhoods. S. Cohn-Vossen has proved that if the metric ρ in R is complete then the metric ρ_1 on E is also complete (see Doklady Akademii Nauk SSSR, Vol. III, No. 8 (1935), pp. 339–342; the assertions proved therein easily imply this assertion for every complete intrinsic metric). But we have proved that a complete metric of positive curvature on the plane is realizable. This implies that the curvature of the plane E with metric ρ_1 does not exceed 2π. On the other hand, there are many domains on E isometric to parts of R with positive curvature, since $\omega(R) > 0$ and the covering has infinitely many branches. We obtain a contradiction that shows that R is the plane. This idea to use the covering of R by the plane as well as the elegant argument with estimation of $p(r + \Delta r) - p(r)$ is taken from the paper by Cohn-Vossen which is cited in the beginning of this section.

© 2006 by Taylor & Francis Group, LLC

curvature.[12] This immediately implies that R itself is homeomorphic to a domain on one of these closed manifolds; the main assertion of Theorem 4 is proved.

5. The Question of the Uniqueness of a Convex Surface with a Given Metric

After proving the existence of a surface that realizes a given metric, we come to the question of the uniqueness of such a surface. Of course, the uniqueness is understood here to within motion and reflection. In the general case, the answer should be negative, since, as is known, every sufficiently small piece of an analytic convex surface can be bent, i.e., it can be deformed while preserving the lengths of all curves on this piece. However, for "large" pieces of the surface, and the more so, for a complete surface, the situation may be different. Of course, we can bend every convex surface by pressing into this surface. For example, we can intersect the surface by a plane and reflect one of the part of this surface in this plane. Moving continuously the plane starting from some supporting plane and performing these reflections, we obtain a continuous bending of the surface. But the surface is no longer convex under this operation. Therefore, it is natural to restrict consideration only to those bendings that preserve the convexity of the surface. We present the main relevant results without proof.

It was Cauchy who proved that two closed convex polyhedra are congruent if they are equicomposed from congruent faces. However, this does not answer the question of the uniqueness of a convex closed polyhedron with given development. It might happen that, allowing flexes of faces, we would obtain different closed convex polyhedra with the same development. Meanwhile, a simple generalization of the Cauchy method allows us to prove that this is impossible and so to prove the uniqueness of a closed convex polyhedron with a given development.[13] Moreover, using the preservation of the area of the spherical image under isometric mappings, it is easy to prove that a convex surface isometric to a closed convex polyhedron is itself a polyhedron. This result in combination with the just indicated generalization of the Cauchy theorem, leads us to the following theorem: *A convex surface isometric to a closed convex polyhedron is itself a polyhedron congruent to the former.*[14] *In other words, for each polyhedral metric of positive curvature on the sphere, there exists only one convex surface realizing it (to within motion and reflection) which is a polyhedron.*

For every metric of positive curvature on the sphere, the question of the uniqueness of a convex surface realizing this metric seems to be very difficult.[15]

[12] This is proved on using the same arguments as indicated in the footnote of Sec. 2 in regard to the proof of the general realization theorem.

[13] The proof of the Cauchy theorem can be found in the book J. Hadamard, Elementary Geometry, Part II (Stereometry), Appendix, L, p. 534, Uchpedgiz, 1938. For a supplement to the Cauchy method, see the paper by A. D. Aleksandrov, "Existence of a Convex Polyhedron and a Convex Surface with Given Metric", Mat. Sb., Vol. 11, No. 1 (1941), p. 20.

[14] This theorem was firstly proved by S. P. Olovyanishnikov in 1941 without using the invariance of the spherical image.

[15] It is sufficient to mention that as far back as in 1838, Minding conjectured that there are no closed analytic surface of constant curvature except for the sphere; only in 1899 this was proved by Liebman. This question of uniqueness was studied by Liebman, Hilbert, Weyl, and Blaschke; only in 1927 this question was settled by Cohn-Vossen.

© 2006 by Taylor & Francis Group, LLC

Still now, we have only the following two results here. The first result is due to Cohn-Vossen who proved that *if thrice differentiable convex closed surfaces of everywhere positive curvature are isometric then they are congruent.*

The proof of the Cohn-Vossen theorem is given in his article "Bending of Surfaces in the Large," Uspekhi Matematicheskikh Nauk, No. 1.

The Cohn-Vossen theorem solves the question of uniqueness if the curvature is positive and it is *a priori* assumed that the surface is differentiable three times. But it can happen that this metric is realized by a less regular convex surface; for example, only a twice-differentiable metric although the latter seems improbable.

Moreover, it is not known what conditions on the metric ensure the existence of a thrice differentiable realizing surface. Only the result by H. Lewy is known here; this author proved that a metric given on the sphere by an analytic line element with everywhere positive curvature can be realized by an analytic convex surface.

The second result is final, although a particular: We can prove that *the sphere is a unique closed convex surface of constant Gaussian curvature.*

No assumptions on regularity are needed here. This proposition is an obvious consequence of the generalized Gauss theorem and the following theorem:

If two closed convex surfaces have the property that each two domains on these surface with the same spherical image have equal areas, then these surfaces are congruent and translates of each other.[16]

As for nonclosed surfaces, we can prove the following. *If we remove an arbitrarily small domain of positive curvature from a closed convex surface, then we obtain a surface that is no longer a unique convex surface with the same metric.* There exist continuum manysurfaces that are isometric but not congruent to it, and we may assume that it is possible even to bend this surface continuously.

The following theorem holds for infinite convex surfaces. *If an infinite complete convex surface has complete curvature* $< 2\pi$*, then there exist infinitely many surfaces isometric but not congruent to this surface.* Moreover, the same is true for each piece of such surface.

The proof of both above-stated theorems on existence of isometric but not congruent surfaces is based on the available realization theorems and the gluing theorems. They also are existence theorems, and, therefore, our methods apply for them, while the uniqueness theorems are "nonexistence" theorems, i.e., they state that there are no isometric surfaces that are not congruent, and require another methods.

The theorem on infinite surfaces is due to Olovyanishnikov; moreover, this author obtained even a more sharp result. Let F be an infinite complete convex surface of total curvature, or what is the same, a surface with area of the spherical image less than 2π. If we contract a surface F infinitely and homothetically to an arbitrary point O, then this surface converges to a convex cone with vertex O (or to a doubly-covered angle on the plane) which is called the limit cone of this surface. It is easy to verify that the spherical image of this cone coincides with the closure of the spherical image of the surface F. (The spherical image of every cone is obviously closed, while the spherical image of the surface F can be nonclosed.) The Olovyan-

[16] See A. D. Aleksandrov, To the Theory of Mixed Volumes of Convex Bodies, III, Mat. Sbornik, Vol. 3, No. 1 (1938), p. 36.

© 2006 by Taylor & Francis Group, LLC

ishnikov theorem is formulated as follows: *Let F be an infinite complete surface of total curvature $\omega < 2\pi$, and let L be a semi-infinite curve on F each of whose segments is a shortest arc (such curves can be drawn from each point on F). Let K be a convex cone with area of the spherical image equal to the total curvature ω of the surface F, and let L' be some cone generator. There exist two convex surface F_1 and F_2 isometric to F and having the cone K as their limit cones, and when we contract them to the cone K, the lines L_1 and L_2 corresponding to L pass to L'. The surfaces F_1 and F_2 differ from one another by orientation, i.e., if the direction of traversal that is given on F_1 makes the right screw with the outer normal, then the corresponding direction of traversal on F_2 makes the left screw.*[17] (In particular cases, F_1 and F_2 can be congruent).

Each cone can be deformed the complete angle at its apex, and thus, preserving the area of its spherical image. Therefore, we can bent the cone K continuously, and the surface F_1 bends simultaneously with this cone. However, it remains unclear whether this surface bends continuously. We can keep the cone K unchanged and move the generator L' along this cone; then the surface F_i revolves in some sense around this cone. (For surfaces of total curvature 2π, the limit cone degenerates into a half-line, and the theorem, remaining true formally, becomes senseless.)

To conclude this section, we present the following conjectures which seem to be very plausible.

1. A convex surface of total curvature equal 4π is uniquely determined from its metric. (Such a surface can be nonclosed.)

2. An infinite convex surface of curvature 2π is uniquely determined from its metric. (Such a surface can be incomplete.)

   Unique determination is understood in the sense that two isometric convex surface should be congruent.

3. Convex surfaces that are obtained from the surfaces of Conjectures 1 and 2 by excluding some domains of positive curvature and also infinite surfaces of curvature $< 2\pi$ admit continuous bendings. Moreover, it is possible that if two of such surfaces are isometric, then one of them bends continuously into the other or into a surface symmetric to the other. For small pieces of analytic surfaces with everywhere positive curvature this was proved by Lewy (1912).

4. The isometric mappings of infinite convex surfaces which are indicated in the Olovyanishnikov theorem can be realized by a continuous bending with possible reflection, and for given F, K, L, L', and orientation, the surface F_1 is unique. These assertions were proved by Olovyanishnikov for polyhedra.

Complete solution of the posed questions seems to present severe difficulties, and so even partial results that solves a particular question for more or less specific class of surfaces could be of great importance.

[17] S. P. Olovyanishnikov, On Bending Infinite Convex Surfaces, Mat. Sbornik, Vol. 18, No. 3 (1946), p. 434.

© 2006 by Taylor & Francis Group, LLC

6. Various Definitions of a Metric of Positive Curvature

A metric of positive curvature is defined as an intrinsic metric in which the sum of the angles of every sufficiently small triangle is at least π, i.e., choosing various definitions of angle, we obtain different definitions of this metric. Here, of course, we have to choose a definition that allows us to prove that in each manifold with metric of positive curvature, each point has a neighborhood isometric to a convex surface. To avoid the specific condition of existence of the angle, we introduced the concept of lower angle and used it for the definition of metric of positive curvature. However, as we have already seen when introducing the definition of lower angle, it has a shortage consisting in the fact that it involves not only an arbitrarily small neighborhood of a starting point of shortest arcs. Therefore, it is desirable to replace this definition by another one. Other definitions of angle lead to a new definition of metric of positive curvature; and, thus, to the new conditions necessary and sufficient for each point in a manifold with such a metric to have a neighborhood isometric to a convex surface. We present without proofs some relevant results that can be obtained here. They can be regarded as various systems of axioms defining an intrinsic metric of a convex surface.

We thus take the following conditions on a metric space R as a foundation.

1. R is a two-dimensional manifold.
2. The metric on R is intrinsic.
3. The sum of the angles between the sides of each sufficiently small triangle in R is no less than 2π.

First, by the angle we can understand the angle in the sense of our usual definition, i.e., $\lim_{x,y\to\infty}\gamma(x,y)$. In this case, the following condition must be added to conditions 1–3:

4. There is a definite angle between two shortest arcs in R that emanate from a common point and also, say, Condition 5, the nonoverlapping condition for shortest arcs.

The systems of Conditions 1–5 is necessary and sufficient for each point in R to admit a neighborhood isometric to a convex surface. However, the proof of sufficiency turns out to be considerably more difficult than that in the case when the lower angle was introduced. Probably, Condition 5 is superfluous, but we cannot prove this fact by now.

Second, in Condition 3, we can bear in mind the angle that is defined as follows. Let L_1 and L_2 be two shortest arcs emanating from a point O in a manifold with some metric ρ. Let X_1 and X_2 be two variable points in R different from O. Put $z=\rho(X_1X_2), x_i=\rho(OX_i)$, and $\xi_i=\rho(R_iL_i)/\rho(OX_i)$ $(i=1,2)$, where $\rho(X_iL_i)$ is the distance from the point X_i to the shortest arc L_i. Let ϕ be the angle opposite the side z of the plane triangle with sides x_1, x_2, and z. We define the lower limit of the angle ϕ as x_1, x_2 and ξ_1, ξ_2 tend to zero to be the angle $\alpha(L_2L_5)$. (The visual meaning of the condition that x_1, x_2 and ξ_1, ξ_0 tend to zero is quite obvious.) The lower limit always exists, and so the specific condition of existence of the angle turns out superfluous.

© 2006 by Taylor & Francis Group, LLC

The following general assertion is rather easy to prove. The angle just defined $\alpha(L_1L_2)$ is always no greater than the lower angle between L_1 and L_2 in the sense of our definition we have accepted. Therefore, if we take the angle $\alpha(L_1L_2)$ as the angle in Condition 3, than this conditions implies immediately that the sum of the lower angles between the sides of a triangle is always no less than π. But since we have proved the sufficiency of this condition, Condition 3 is thus sufficient also with the angle $\alpha(L_1L_2)$. At the same time, the angle $\alpha(L_1L_2)$ depends, by definition, only on arbitrarily small segments of the shortest arcs L_1 and L_3 near the point O, and so this concept of angle is free of the above-mentioned shortage of the concept of lower angle.

However, the inconvenience of the angle $\alpha(L_1L_2)$ consists in that when introducing this angle, the necessity of Condition 3 is proved in a more difficult way. Namely, using the continuity property of curvature, we can prove that the angle $\alpha(L_1L_2)$ on a convex surface is equal to the angle between L_1 and L_2 in the sense of our usual definition (moreover, the angle ϕ as x_1, x_2, ξ_1, $\xi_2 \to 0$ always has a definite limit). Since the sum of the usual angles between the sides of a triangle on a convex surface is no less than π, we thus have proved the necessity of Condition 3 with the angle $\alpha(L_1L_2)$. However, the continuity of curvature is a very deep property, and in particular, we have succeeded in proving it only in Sec. 4 of Chapter V; therefore, it seems inappropriate to introduce necessary condition that is difficult to prove.

Finally, we can introduce one of the following two definitions of angle:

$$\text{(A)}\ \limsup_{x,y\to 0}\gamma(x,y) \quad \text{and} \quad \text{(B)}\ \liminf_{x,y\to\ 0}\gamma(x,y).$$

Since these upper and lower limits exist, we need not the special condition of the existence of the angle, but other conditions turn out necessary.

When defining the angle by (A), we can introduce Condition 4(A). On each shortest arc L, almost all points X (in the sense of linear measure) are such that the sum of the angles defined by (A) between the branches of L meeting at X and any shortest arc emanating from X does not exceed π. We are familiar with the necessity of this condition, and we can prove that Conditions 1–4 and 4(A) are sufficient. We cannot avoid the use of Condition 4(A) or some other conditions, since it is possible to exhibit simple examples of metrics that satisfy Condition 1–3 (with Definition (A) of angle) but which are not realized locally on any surface. Namely, define the length $|x\overline{x}|$ of a vector $\overline{x}$ on the plane so that the following conditions hold: (1) $|\overline{x}+\overline{y}| \le |\overline{x}|+|\overline{y}|$; (2) $|\overline{x}| = 0$ if and only if $\overline{x} = 0$; (3) $|\lambda\overline{x}| = |\lambda||\overline{x}|$; (4) if $\overline{x} = \overline{y}$ in the sense that $\overline{x}$ and $\overline{y}$ coincide under a parallel translation, then $|\overline{x}| = |\overline{y}|$. As a result, we define a metric on the plane that is intrinsic and satisfies Condition 3 with Definition (A) of the angle as is easily seen. Straight lines are shortest arcs in this metric, but the usual angle made by them ($\lim_{x,y\to 0}\gamma(x,y)$) exists if and only if the introduced metric is equivalent to the Euclidean metric. If,

© 2006 by Taylor & Francis Group, LLC

e.g., we put $|\overline{x}| = |x_1 - x_2| + |y_1 - y_2|$, where X_i and y_i are rectangular coordinates of the endpoints of the segment $\overline{x}$, then all conditions on $|\overline{x}|$ hold, but there are no angle between two straight lines.[18]

If we accept Definition (B) of angle, then the nonoverlapping condition for shortest arcs can be introduced as an additional condition. However, it seems likely that in this case we can avoid using any additional condition. If we succeed in proving this, then Conditions 1–3 with Definition (B) of angle yield the simplest system of conditions necessary and sufficient for a point to admit a neighborhood in the space R isometric to a convex surface.

[18] We can introduce the following condition instead of Condition 4(A): a space has the tangent cone at each point (in the sense of the definition at the end of Sec. 5 of Chapter IV). A metric given by the line element ds with continuous coefficients E, F, and G satisfies this condition *a priori* (in this case, the cone is a plane at each point), and, therefore, if, in advance, we restrict consideration only to these metrics, then Definition (A) of angle turns out to be the most convenient.

© 2006 by Taylor & Francis Group, LLC

Chapter IX

CURVES ON CONVEX SURFACES

1. The Direction of a Curve

Among basic concepts of the theory of curves in intrinsic geometry, we have considered by now only the length of a curve and the angle between curves. We now pass to two other basic concepts, the direction of a curve at a given point and the swerve of the direction of a curve, which is a generalization of the concept of the curvature of a curve. Since shortest arcs can emanate from a given point on a convex surface not in an arbitrary direction, the intrinsic definition of the direction of a curve via a shortest arc that touches this curve seems inappropriate. For example, there are no shortest arc on a right circular cone which touches its base circle and also there are no shortest arcs on a doubly convex lens which touch its edge. Meanwhile, both the base circle of the right circular cone and the edge of the doubly convex lens are circles both in the space and in the sense of intrinsic geometry, so that it would be rather extravagant to avoid considering them as smooth curves. Also, we can give examples of curves that are not edges of a convex surface and are smooth in the spatial sense but which have no shortest arcs tangent to them at some points.[1]

These reasons force us to introduce another definition of the direction of a curve; this was already done in Sec. 7 of Chapter I. Namely, we begin with the concept of the angle made by curves. Let $X_t = X(t)$ and $Y_s = Y(s)$ be two curves that emanate from one point $O = X_0 = Y_0$ on a convex surface, or, in general, on a metric space with metric ρ. We assume here that for t and s sufficiently close to but different from zero, the points X_t and Y_t also differ from O. (In what follows, this condition is assumed to be always fulfilled.[2]) Let $\gamma(t, s)$ be the angle opposite to the side equal to $\rho(X_tY_s)$ in the plane triangle with sides equal to $\rho(OX_t), \rho(OY_s)$, and $\rho(X_tX_s)$. We say that the curve $X(t)$ and $Y(t)$ make the angle α at the point O if the limit $\lim_{t,s\to 0}\gamma(t, s)$ exists and equals α.

Further, we say that a *curve $X(t)$ emanating from the point $O = X(0)$ has some direction at this point if this curve makes some angle with itself.* Applying the definition of angle, we must understand this as follows: we consider the curve $X(t)$ as two coinciding curves with a common parameter; the angle made by these coinciding curves is defined as $\lim_{t,t'\to 0}\gamma(t, t')$, i.e., we take the points $X_t = X(t)$ and $X_{t'} = X(t')$, and the angle $\gamma(t, t')$ is defined in the triangle with the sides

[1] In Sec. 10 of Chapter I, we have exhibited the example of a convex smooth surface on which there are no shortest arcs emanating from a point O in some direction L. A plane section of such a surface which touches L is a smooth curve that has no shortest arc touching it at the point O.

[2] If $X_t = O$ for arbitrarily small t then the angle $\gamma(t, s)$ is not defined for these t, and the above definition has no meaning.

© 2006 by Taylor & Francis Group, LLC

$\rho(OX_t), \rho(OX_{t'})$, and $\rho(X_tX_{t'})$. If we take $t = t'$, then the points X_t and $X_{t'}$ coincide, and, therefore, $\gamma(t,t) = 0$. Consequently, if $\lim_{t,t'\to 0} \gamma(t,t')$ exists, then it is equal to zero; in other words, *if a curve makes a definite angle with itself, then this angle is equal to zero.*

Obviously, each shortest arc has a definite direction at the initial point. The existence of the direction of a curve on the plane in the sense of our definition is equivalent to the existence of the tangent; this fact is a particular corollary of the following general theorem, which allows us to better understand the meaning of the above definitions of angle and direction.

Theorem 1. *Let $X_t = X(t)$ and $Y_s = Y(s)$ be two curves that emanate from one point $O = X_0 = Y_0$ on a convex surface. These curves make some angle at the point O if and only if the angle between the shortest arcs OX_t and OY_s tends to α as t, $s \to \infty$.*

Proof. Let $\alpha(t,s)$ be the angle made by the shortest arcs OX_t and OY_s,[3] and let $\gamma(t,s)$ have the same meaning as above. Since $\alpha(t,s)$ is the angle made by the sides of the triangle OX_tY_s and since $\gamma(t,s)$ is the corresponding angle of the plane triangle with sides of the same lengths, we have

$$|\alpha(t,s) - \gamma(t,s)| \le \omega(t,s), \tag{1}$$

where $\omega(t,s)$ is the curvature of the interior of the triangle OX_tY_s (Theorem 1 in Sec. 6 of Chapter V). But when the points X_t and Y_s tend to O, the interior of the triangle OX_tY_s turns out to be included into smaller and smaller neighborhoods of the point O from which the point O itself is deleted. Such "punctured" neighborhoods form a vanishing sequence of sets, and by the continuity property of curvature, their curvatures tend to zero. Consequently, we also have

$$\lim_{t,s\to 0} \omega(t,s) = 0.$$

In combination with (1), this yields $\lim_{t,s\to 0} \gamma(t,s)$, i.e., the angle made by our curves exists if and only if there exists $\lim_{t,s\to 0}$, which completes the proof.

By our definition of the limit direction, Theorem 1 implies that *a curve $X_t = X(t)$ on a convex surface has a direction at the initial point $O = X_0$ if and only if the angle made by the shortest arcs OX_t and $OX_{t'}$ tends to zero as $t, t' \to 0$.* This makes it clear that the existence of the direction at a point on the plane is equivalent to the existence of the tangent at this point.

Theorem 2. *Two curves on a convex surface emanating from a common point O make some angle with each other if and only if each of them has a definite direction at the point O.*

Proof. Let $X_t = X(t)$ and $Y_s = Y(s)$ be two curves on a convex surface emanating from the point $O = X_0 = Y_0$. Let $\alpha(t,s)$ be the angle made by the shortest arcs

[3] For given X_t and Y_s, the shortest arc OX_t and OY_s can be not unique, and hence the function $\alpha(t,s)$ is not single-valued in general.

© 2006 by Taylor & Francis Group, LLC

OX_t and OY_s, and let $\xi(t,t')$ and $\eta(s,s')$ be the angles made by the shortest arcs OX_t, $OX_{t'}$ and $OY_s, OY_{s'}$, respectively.

According to Theorem 1, the existence of the limit $\lim_{t,s\to 0}\alpha(t,s)$ is necessary and sufficient for the existence of the angle made by our curves, while the relations $\lim_{t,t'\to 0}\xi(t,t')=\lim_{s,s'\to 0}\eta(s,s')=0$ are necessary and sufficient for the existence of directions of these curves.

The general property of angles (Theorem 1 in Sec. 1 of Chapter IV) implies

$$\alpha(t,s)\le\alpha(t',s')+\xi(t,t')+\eta(s,s'),$$

and, similarly,

$$\alpha(t',s')\le\alpha(t,s)+\xi(t,t')+\eta(s,s').$$

It is clear from this that if $\lim_{t,t'\to 0}\xi(t,t')=\lim_{s,s'\to 0}\eta(s,s')=0$, then the limit $\lim_{t,s\to 0}\alpha(t,s)$ exists, i.e., the existence of directions of curves implies the existence of the angle made by them.

Let us prove the converse statement, i.e., that the existence of $\lim_{t,s\to 0}\alpha(t,s)$ implies $\lim_{t,t'\to 0}\xi(t,t')=\lim_{s,s'\to 0}\eta(s,s')=0$. To this end, assume that the limit $\lim_{t,s\to 0}\alpha(t,s)$ exists. Then, for all given positive ε, there exist t_ε and s_ε satisfying the following conditions:

1. For all $t',t<t_\varepsilon$ and $s,s'<s_\varepsilon$, we have

$$|\alpha(t,s)-\alpha(t',s')|<\varepsilon.$$

2. For all $t<t_\varepsilon$ and $s<s_\varepsilon$, the points X_t and Y_s lie in a convex neighborhood U of the point O such that the curvature of this neighborhood without point O is less than ε, i.e.,

$$\omega(U-O)<\varepsilon. \tag{2}$$

This condition may be satisfied in view of the continuity property of the curvature.

Let us show that for $t,t'<t_\varepsilon$ and $s,s'<s_\varepsilon$,

$$\xi(t,t'),\eta(s,s')\le 3\varepsilon.$$

Since ε can be taken arbitrarily small, this will prove that the angles $\xi(t,t')$ and $\eta(s,s')$ tend to zero.

It is sufficient to consider one of these angles, say, ξ.

Assume by way of contradiction that there exists $t_1<t_2<t_\varepsilon$ satisfying[4]

$$\xi(t_1,t_2)>3\varepsilon. \tag{3}$$

Take some $s<s_\varepsilon$ and draw a shortest arc OY_s. Let U_1 and U_2 be those sectors bounded by the pairs of shortest arcs OX_{t_1} and OY_s, OX_{t_2} and OY_s whose angles are equal to the angles made by these shortest arcs themselves. If, e.g., OX_{t_2} lies in the sector U_1 between OX_{t_1} and OY_s, then the angle of this sector is equal to the sum of the angles between OX_{t_1} and OX_{t_2}, OX_{t_2} and OY_s; i.e., $\alpha(t,s)=\alpha(t_2,s)+\xi(t_1,t_2)$, since $|\alpha(t_1,s)-\alpha(t_2,s)|<\varepsilon$ by condition, we have $\xi(t_1,t_2)<\varepsilon$. This contradicts

[4] That is, there exist shortest arcs OX_{t_1} and OX_{t_2} with the angle between them $>3\varepsilon$.

© 2006 by Taylor & Francis Group, LLC

the assumption, and hence OX_{t_2} cannot lie in the sector U_2. Thus, these sectors are located to the opposite sides of the shortest arc OY_s (Fig. 75).

Take some value t between t_1 and t_2 and draw a shortest arc OX_t; by condition, the angle $\alpha(t, s)$ between OX_t and OY_s satisfies the inequalities

$$|\alpha(t, s) - \alpha(t_1, s)| < \varepsilon,$$
$$|\alpha(t, s) - \alpha(t_2, s)| < \varepsilon. \tag{4}$$

Let U be the sector between OX_t and OY_s whose angle is equal to the angle made by these shortest arcs, i.e., to $\alpha(t, s)$. Since the sectors U_1 and U_2 lie to the opposite sides of OY_s, either U includes one of them, or, in the opposite case, U is included in one of them.

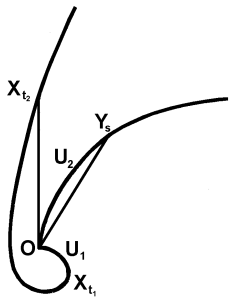


Fig. 75

If, e.g., U is contained in U_1 then the angle made by OX_{t_1} and OY_s is equal to the sum of the angles made by OX_{t_1} and OX_t, OX_t and OY_s, i.e.,

$$\alpha(t_1, s) = \alpha(t, s) + \xi(t_1, t),$$

and, by the first of the inequalities (4), we have $\xi(t_1, t) < \varepsilon$.

In the opposite case, when the sector U contains the sector U_1, we have

$$\alpha(t, s) = \alpha(t_1, s) + \xi(t_1, t);$$

we see again that $\xi(t_1, t) < \varepsilon$.

In exactly the same way, if the sector U is contained in U_2, or, in the opposite case, if U_2 contains U_1, then $\xi(t_2, t) < \varepsilon$. Since one of the above possibilities always holds, we have that for any t ($t_1 \le t \le t_2$), either $\xi(t_1, t) < \varepsilon$ or $\xi(t_2, t) < \varepsilon$, i.e., the shortest arc OX_t cannot simultaneously deviate from OX_{t_1} and OX_{t_2} by more than the angle ε.

Let t_0 be the least upper bound of those t for which the angle $\xi(t_1, t)$ made by OX_t and OX_{t_1} is less than ε. Let OX_{t_0} be the shortest arc connecting the points O and X_{t_0}, which is the limit of the minimal arcs OX_t satisfying $\xi(t_1, t) < \varepsilon$ (Fig. 76; the curve X_t is shown by the bold line).

Since the angle varies continuously with a continuous change of a shortest arc (Theorem 6 in Sec. 4 of Chapter IV), the angle made by OX_{t_0} and X_{t_1} does not exceed ε, i.e.,

$$\xi(t_1, t_0) \le \varepsilon.$$

On the other hand, let $\overline{OX_{t_0}}$ be the shortest arc connecting the points O and X_{t_0}, which is the limit of the shortest arcs OX_t as $t \to t_0 + 0$. Since the shortest arcs OX_t for $t > t_0$ make the angle greater than ε with OX_{t_1}, they make the angle less than ε with OX_{t_1} by what we have proved above. Therefore, for the limit shortest arc $\overline{OX_{t_0}}$, this angle is not greater than ε,[5] i.e.,

$$\overline{\xi}(t_2, t_0) \le \varepsilon.$$

[5]It may happen that $t_0 = t_2$, and then, there are no $t > t_0$; but in this case, we take OX_{t_2} itself as the shortest arc $\overline{OX_{t_2}}$.

© 2006 by Taylor & Francis Group, LLC

Thus, the minimal arcs OX_{t_0} and $\overline{OX_{t_0}}$ make the angles no greater than ε with OX_{t_1} and OX_{t_2}, respectively; since the angle $\xi(t_1, t_2)$ between OX_{t_1} and OX_{t_2} is greater than 3ε by assumption, the angle between OX_{t_0} and $\overline{OX_{t_0}}$ at the point O must be greater than ε.

These minimal arcs OX_{t_0} and $\overline{OX_{t_0}}$ connect the same points O and X_{t_0} and bound the digon D which is contained in the convex neighborhood U by the second condition imposed on the values of t under consideration. Therefore, the curvature $\omega(D)$ of the interior of D is not greater than $\omega(U - O)$, and by (2), we have

$$\omega(D) \leq \omega(U - O) < \varepsilon.$$

The curvature of the interior of a digon is equal to the sum of its angles. Therefore the angle made by OX_{t_0} and $\overline{OX_{t_0}}$ at the point O must be less than ε. Meanwhile, we have just proved the contrary. This contradiction shows the impossibility of the assumption that there exist T_1 and $t_2 < t_\varepsilon$ satisfying $\xi(t_1, t_2) > 3\varepsilon$. Consequently, $\xi(t, t') \leq 3\varepsilon$ for all t and t'. Since ε is arbitrary, the limit $\lim_{t,t' \to 0} \xi(t, t')$ exists, i.e., the curve $X(t)$ has direction at the point O. Since the curves $X(t)$ and $Y(s)$ play the same role, the curve $Y(s)$ also has direction at O; which completes the proof of the theorem.

It is natural to introduce the following definition: *two curves emanating from a common point O have the same direction at this point if the angle between them is equal to zero.* The above theorem implies that in this case, each of them has a definite direction, and, therefore, we can speak about the coincidence of their directions.

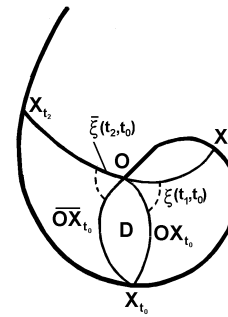


Fig. 76

If curves L_1, L_2, and L_3 make some angles α_{12}, α_{13}, and α_{23} with each other, then by Theorem 1 in Sec. 1 of Chapter IV, $\alpha_{13} \leq \alpha_{12} + \alpha_{23}$. Hence, if $\alpha_{12} = \alpha_{23} = 0$ then also $\alpha_{13} = 0$. Therefore, the coincidence relation just defined is reflexive, symmetric, and transitive. This relation defines a class of curves that have the same direction at the point O. In an abstract sense, a direction itself can be considered as such a class of curves, and the assertion that a curve has a direction d at a point O can be understood so that this curve belongs to the corresponding class. Further, if three curves L_1, L_2, and L_3 make angles α_{12}, α_{13}, and α_{23} and if $\alpha_{23} = 0$, i.e., if the curves L_2 and L_3 have the same direction, then $\alpha_{12} = \alpha_{13}$. This means that the curves with the same direction make equal angles with another curve. Therefore, we can arbitrarily speak about the angles between directions without circumlocutions. An important distinction of angles between any curves, i.e., between their directions, the angles between shortest arcs consists in the fact that for each direction from a point O, there is another direction that makes a given angle with the first (this angle is no greater than half of the complete angle at the point O). This last assertion follows by the way from the following theorem which reveals the extrinsic geometric meaning of the concept of direction.

Theorem 3. *A curve emanating from a point O on a convex surface has direction at this point if and only if this curve has tangent at this point. The angle made by*

© 2006 by Taylor & Francis Group, LLC

two curves emanating from the point O is equal to the angle between the tangents to these curve which is measured on the tangent cone at the point O.

Proof. Let a curve L on a convex surface F have direction at the initial point O. We prove that this curve has tangent at the point O. In Sec. 5 of Chapter V, we have proved that it is possible to draw some shortest arcs from a point on a convex surface that divide its neighborhoods into arbitrarily small sectors. Therefore we can draw shortest arcs M and N from our point that go to the opposite sides of L and make angles with L that are different from zero by less than a given $\varepsilon > 0$.

The curve L does not intersect these shortest arcs in a sufficiently small neighborhood of the point O since this curve has a definite direction at this point.

If we indefinitely dilate the surface F with center at the point O, then the shortest arcs M and N converge to their half-tangents. Since the curve L does not intersect the shortest arcs M and N in a sufficiently small neighborhood of the point O, the limit of this curve as the ratio of the dilation tends to infinity will come to the tangent cone between the half-tangents of M and N. But the angle between these half-tangents is equal to the angle between the shortest arcs themselves (Theorem 2 in Sec. 6 of Chapter IV), and hence this angle is less than ε. Since ε is arbitrarily small, the limit of the curve L is simply a tangent cone generator, which is also the half-tangent of the curve L at the point O.

Assume now that the curve L has the half-tangent T at the point O. We prove that this curve has direction at O. Draw two shortest arcs M and N from the point O so that their half-tangents T_M and T_N bound a sector on the tangent cone whose angle is less than a given ε and, moreover, this sector includes the half-tangent T. This is possible by the theorem of Sec. 5 of Chapter V. The curve L does not intersect M and N in some arbitrary small neighborhood of O, since the half-tangents T_M and T_N are distinct. Therefore, whenever two points X and Y of the curve L are sufficiently close to O, the shortest arcs OX and OY lie in the sector between M and N, and, so, the angle between them is less than ε. Since ε is arbitrarily small, this means exactly that the curve L has direction at the point O by Theorem 1 (by its corollary, to be more exact).

If now L_1 and L_2 are two curves emanating from O, each having definite direction at O, then they have half-tangents T_1 and T_2 by what we have proved above. Applying the above argument to these curves, we include them into the sectors between the shortest arcs M_1, N_1 and M_2, N_2 whose angles are less than ε. Correspondingly, the half-tangents T_1 and T_2 turn out to be included between the half-tangents T_{M_1}, T_{N_1} and T_{M_2}, T_{N_2}.

Under these conditions, the angle α' made by T_1 and T_2 on the tangent cone differs from the angle α_1' made by T_{M_1} and T_{M_2} by less than ε, i.e., $|\alpha' - \alpha_1'| < \varepsilon$; in exactly the same way, the angle α made by L_1 and L_2 differs from the angle α_1 made by M_1 and M_2 also by less than ε, i.e., $|\alpha - \alpha_1| < \varepsilon$. But since the angle made by shortest arcs is equal to the angle made by their half-tangents, which is measured on the tangent cone (Theorem 2 in Sec. 6 of Chapter IV), we have $\alpha_1' = \alpha_1$, and hence $|\alpha - \alpha'| < \varepsilon$; since ε is arbitrarily small, $\alpha = \alpha'$, i.e., the angle made by the curves L_1 and L_2 themselves is equal to the angle made by their half-tangents T_1 and T_2 which is measured on the tangent cone.

© 2006 by Taylor & Francis Group, LLC

Thus, the proof of the theorem is complete. The theorem shows that a generator of the tangent cone corresponds to each direction at a point O. Conversely, some direction corresponds to each cone generator. The curve l whose direction corresponds to a generator T can be obtained by intersecting the surface with the plane passing through T. Therefore, Theorem 3 can also be formulated as follows:

There is a one-to-one angle-preserving correspondence between the generators of the tangent cone at a point O and the directions at this point on a surface. All curves of the same direction have the corresponding generator as half-tangent.

This shows that, for curves on a convex surface, the notion of direction in the sense of the intrinsic definition coincides with the notion of direction in the spatial sense. For example, Theorem 5 of Chapter V reads: given a point of a convex surface, consider the directions in which there are no shortest arcs issuing from this point; then the set of these directions has angular measure zero. When proving this theorem, we use directions on the tangent cone. Now we can endow this theorem with a purely intrinsic meaning and repeat the whole proof using the language of intrinsic geometry not introducing the concept of tangent cone.

By now, we spoke only about the direction of a curve at its initial point.

If a point O lies inside a curve L, then we agree to say that L *has direction at the point O if its branches emanating from O have directions making the angle equal to π with each other.*

If a curve L on a convex surface traverses an edge of this surface at a point O, then the curve L can have no tangent at the point O even if this curve has direction at this point in the sense of the given intrinsic definition. However, the curve L has right and left half-tangents at O, and the angle between these half-tangents, measured on the tangent cone, is equal to π.

We present the following two theorems as examples of applications of the concept of direction; the reader will easily prove these theorems on his or her own (we are talking of curves on convex surfaces).

1. A curve, bounding a convex domain, has directions at its every point O of the branches on both sides of O.

2. We call the geometric locus of points, such that the sum (difference) of their distances to two given points F_1 and F_2 – the focuses – is constant, an ellipse (hyperbola). Take a branch L of an ellipse (hyperbola) which emanates from the point O, and let F_1O and F_2O be the focal radii (i.e., the shortest arcs F_1O and F_2O, that are the limits of the shortest arcs F_1X and F_2X as $X \to O$ along the branch L). The branch L of the ellipse makes equal angles with F_1O and F_2O, while for the hyperbola, the sum of the angles made by L and F_1O, F_2O is equal to π. (We do not specify various possible singularities here; e.g., the ellipse may fail to exist at all on the sphere if its major semi-axis is greater than the length of the great circle; if the focuses lie at the poles and the major semi-axis is equal to the great circle, then the ellipse covers the whole sphere. It is an interesting problem to examine all these and other singularities of the ellipse and hyperbola on convex surfaces. If the focuses F_1 and F_2 coincide then the ellipse reduces to the circle, and in this case, the theorem reduces to the assertion that "the circle is orthogonal to the radius."

© 2006 by Taylor & Francis Group, LLC

The precise formulation and proof of this theorem will be given in Sec. 6 which is especially devoted to the circle.)

2. The Swerve of a Curve

We now pass to a generalization of the curvature of a curve. First we consider a geodesic broken line, i.e., a curve that consists of finitely many shortest arcs. We can distinguish the two sides of this curve, agreeing to call them right and left. If a broken line has no multiple points, then the possibility of so distinguishing the sides is obvious. If a broken line has multiple points, then having defined the right and left sides of some of its segments, we extend this definition to the whole broken line by continuity.

A broken line makes two angles α_i and β_i, one to the right and the other to the left of this broken line at its every vertex. The value

$$\phi = \sum_i (\pi - \alpha_i) \tag{1}$$

is called the right swerve,[6] while the value

$$\psi = \sum_i (\pi - \beta_i) \tag{1'}$$

is called the left swerve.

The sum of the angles $\alpha_i + \beta_i$ is equal to the complete angle at the vertex of the broken line, i.e., $2\pi - \omega_i$, where ω_i is the curvature of this vertex. Therefore,

$$\phi + \psi = \sum_i \omega_i \geq 0 \tag{2}$$

The curvature (the area of the spherical image) of a geodesic at parts is equal to zero, and, therefore, if a broken line has no multiple vertices then formula (2) means that the sum of the right and left swerves of a broken line is equal to the curvature of the broken line without endpoints. On the plane and on any regular surface, the sum of the right and left swerves of a geodesic broken line is equal to zero, so that the distinction between the right and left swerves is immaterial. But, speaking about an arbitrary surface, we must discriminate between these swerves.

If a broken line bounds a polygon that is homeomorphic to a disk and has angles $\alpha_1, \ldots, \alpha_n$, then the curvature of this polygon is equal to

$$\omega = 2\pi - \sum_{i=1}^{n} (\pi - \alpha_i).$$

The sum on the right-hand side here is none other than the swerve of the broken line to the side of the polygon, and, so, we can write

[6] The term "swerve" was minted by H. Busemann. The alternative term "turn" is also in common parlance (Ed.).

© 2006 by Taylor & Francis Group, LLC

$$\omega = 2\pi - \phi.$$

Now let L be an arbitrary curve on a surface with no multiple points; we can also distinguish two sides of the curve L which we will agree to call right and left. Let A and B be the end parts of this curve, and assume that the curve L has directions at the points A and B. Construct a sequence of geodesic broken lines L_i without multiple points which connect the points A and B, travel to the right of the curve L, and converge to this curve. Whenever the broken lines L_i are sufficiently close to the curve L, it makes sense to speak about their traveling to the right of the curve L, and we can define the right side of each of them in correspondence with the right side of the curve L. The definition of these notions can be given as follows.

The broken line L_i, together with the curve L, bounds some domain G_i which however can fall into separate parts if the broken line L has other common points with L_i, except for the endpoints; e.g., the broken line L_i can be inscribed into L. We say that the broken line L_i travels to the right of the curve L if the interior of the domain G lies to the right of the curve L. In contrast, we consider the opposite side to the side of L to which the domain G lies (Fig. 77) as the right side of the broken line L_i.[7]

Since the curve L has directions at the endpoints A and B, it makes angles with the broken line L_i. We take the angles related to the sectors lying in the domain G. Let α_i and β_i be these angles, and let ϕ_i be the right swerve of the broken line L_i. The value $\phi_i + \alpha_i + \beta_i$ represents in a sense the turn from the direction of the curve L at the point A to its direction at the point B, but this turn is taken along the broken line L rather than the curve itself.

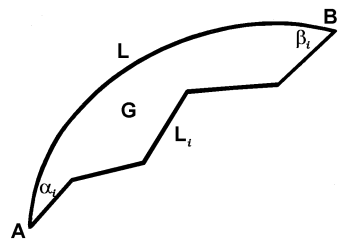


Fig. 77

The magnitude

$$\phi = \lim_{L_i \to l} (\phi_i + \alpha_i + \beta_i) \qquad (3)$$

is called the swerve of the curve L to the right or the right swerve of the curve L.

The left swerve of the curve L is defined in a similar way.

Theorem 1. *For a curve on a convex surface, the limit in formula* (3) *exists and does not depend on the choice of a sequence of broken lines L_i that converge to L traveling to the right of L_0. Consequently, the existence of the swerve of a curve*

[7] If the curve L has finitely many multiple points, then taking its segment without multiple points and defining its left and right sides, we extend this definition to the whole curve L by continuity. The broken lines L_i that converge to L should be taken so that their topological structure is the same as that of the curve L: they must have the same number of multiple points. The notion of a broken line traveling to the right (left) of a curve and the notion of the right (left) side of this broken line can easily be defined. To this end, we have to take a segment of a line, together with some of its neighborhoods on the plane, and map this segment onto the curve L and the neighborhood, onto a neighborhood of the curve L. The notions "to the right" and "to the left" in a neighborhood of the curve should correspond to these notions in a neighborhood of the segment.

© 2006 by Taylor & Francis Group, LLC

without multiple points is ensured only by the existence of directions of this curve at its endpoints.[8]

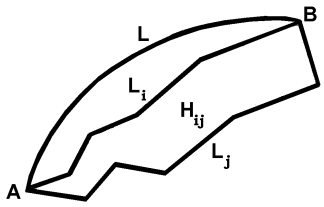


Fig. 78

Proof. Let L_i and L_j be two broken lines traveling to the right of the curve L and connecting the endpoints A and B of this curve. Assume for simplicity that these broken lines are disjoint; then they bound some polygon H_{ij} (Fig. 78). This polygon lies to the right of one of these broken lines and to the left of the other. Assume that this polygon lies to the right of L_i and to the left of L_j. The angles of the polygon H_{ij} are as follows: at the vertices A and B, they are equal to $\alpha_j - \alpha_i$ and $\beta_j - \beta$, respectively; at the vertices of the broken line L_i, they are its angles to the right, and at the vertices of the broken line L_j, they are its angles to the left. Therefore, if ϕ_i stands for the right swerve of the broken line L_i and if ψ_j is the left swerve of the broken line L_i, then the total swerve of the whole broken line $L_i + L_j$ to the side of the polygon H_{ij} is equal to

$$\phi_i + \psi_j + (\pi - \alpha_j + \alpha_i) + (\pi - \beta_j + \beta_i).$$

Therefore, the curvature of the interior of the polygon H_{ij} is equal to

$$\omega(H_{ij}) = 2\pi - \phi_i - \psi_j - (\pi - \alpha_j + \alpha_i) - (\pi - \beta_j + \beta_i),$$

or

$$\omega(H_{ij}) = (-\psi_j + \alpha_j + \beta_j) - (\phi_i + \alpha_i + \beta_i).$$

Let ϕ_i be the right swerve of the broken line, and let $\omega(L_j)$ be the curvature of this broken line without endpoints; then, according to formula (2),

$$\omega(L_j) = \phi_j + \psi_j.$$

Expressing ψ_j from this and inserting the result into the expression for $\omega(H_{ij})$, we obtain

$$\omega(H_{ij}) + \omega(L_j) = (\phi_j + \alpha_j + \beta_j) - (\phi_i + \alpha_i + \beta + i).$$

[8] If we speak about a curve on the plane then this assertion turns out to be almost obvious, since in this case, the value of $\phi_i + \alpha_i + \beta_i$ is the same for all broken lines sufficiently close to the curve. This value is equal, up to some multiplier of 2π, to the angle made by the direction of the curve at its endpoints. If a curve on the plane has right and left half-tangents at each point, then its swerve can be defined as follows. Choosing some direction for counting the angles and some direction of traveling along a curve, we consider, e.g., the right tangent vector. In passage along the curve, this vector is rotated, and its complete turn yields the swerve of the curve. We can say that the swerve of a curve is equal to $\sum \Delta\phi$, where $\Delta\phi$ is the angle of the turn of the tangent vector in passage from one point of the curve to the next one.

This definition of the swerve is based on the fact that we can define the angle between two vectors emanating from different points of the plane. However, on a curved surface the notion of the angle between two directions that emanate from different points is indeterminate. In particular, the concept of parallel directions is indeterminate. The concept of parallel translation along a given path on a curved surface was introduced by Levi-Civita (see, e.g., P. K. Rashevskii, A Course of Differential Geometry, Secs. 59–61).

© 2006 by Taylor & Francis Group, LLC

If we add the broken line L_j to the polygon H_{ij} and denote the result by $\overline{H_{ij}}$, we can write

$$\omega(\overline{H}_{ij}) = (\alpha_j + \alpha_j + \beta_j) - (\phi_i + \alpha_i + \beta_i). \tag{4}$$

The figure $\overline{H}_{ij}$ lies in a neighborhood of the curve L and has no common points with this curve, since the endpoints of the broken line L_j are excluded. If U is a neighborhood of the curve L which contracts to L then, by the continuity property of the curvature, we have

$$\dot{\omega}(U - L) \to 0.$$

Therefore, as the broken lines L_i and L_j converge to L, we have

$$\omega(\overline{H}_{ij}) \to 0,$$

and formula (4) implies

$$|(\phi_j + \alpha_j + \beta_j) - (\phi_i + \alpha_i + \beta_i)| \to 0,$$

i.e., there exists $\lim_{L_i \to L}(\phi_i + \alpha_i + \beta_i)$.

We have assumed that the broken lines L_i and L_j are disjoint. If they meet then they bound several polygons rather than one. Computing the curvatures of all these polygons[9] and using the continuity property of curvature, we prove in a similar way the existence of the limit $\lim_{l_i \to l}(\phi_i + \alpha_i + \beta_i)$ in the general case.

If a curve has finitely many multiple points and directions at the endpoints, then the existence of its swerve is proved by analogy.

Now let us reveal in what way the swerves of separate arcs of a curve are added. Assume that a point C lies inside a curve $\overset{\frown}{AB}$ and that both arcs $\overset{\frown}{AC}$ and $\overset{\frown}{BC}$ have directions at the point C. Let the angle between these directions to the right of the curve be equal to γ. The value $\phi(C) = \pi - \gamma$ will naturally be called the swerve of the curve at the point C. If we take broken lines that travel to the right of the curve $\overset{\frown}{AB}$ so that these lines have a common point C together with the points A and B, then, summing the swerves of the sides AC and BC of these broken lines and passing to the limit, we easily obtain

$$\phi(\overset{\frown}{AB}) = \phi(\overset{\frown}{AC}) + \phi(\overset{\frown}{BC}) + \phi(C), \tag{5}$$

where ϕ stands for the right swerves of the corresponding arcs.

It seems natural to relate the swerve to an arc without endpoints, and when adjoining the endpoints to an arc, to include in the sum the swerves at the endpoints. For example, a curve $\overset{\frown}{AB}$ can be divided into two parts: the arc $\overset{\frown}{AC}$ without the

[9] Let $\overline{H}^1_{ij}, \overline{H}^2_{ij}, \ldots, \overline{H}^n_{ij}$ be these polygons with segments of the broken lines L_i and L_j being added; we add to H^k_{ij} the segment of the broken line for which H^k_{ij} lies to the left. The computation of curvatures yields the formula

$$|(\phi_j + \alpha_j + \beta_j) - (\phi_i + \alpha_i + \beta_i)| = \left| \sum_{k=1}^{n} (-1)^k \omega(\overline{H}^k_{ij}) \right|.$$

© 2006 by Taylor & Francis Group, LLC

points C and the arc $\widehat{BC}$ with the adjoint point C; the swerve of the first is $\phi(\widehat{AC})$, and the swerve of the second is $\phi(\widehat{CB}) + \phi(C)$.

Under this condition, formula (5) shows that the *swerve is an additive function of an arc*, i.e., if a curve is partitioned into finitely many arcs disjoint and each of them having a swerve, then the swerve of the whole curve is equal to the sum of the swerves of separate arcs.

As we have shown, a curve has swerve it only this curve has directions at its endpoints, and hence the existence of the swerve of the whole curve does not imply the existence of the swerves of its arcs. If both branches of each point of a curve have directions, then the swerve is defined for all arcs, and as is just proved, presents an additive function of an arc.

If the curve L has length and if each of its arcs has swerve, then we can introduce the concept of geodesic curvature of a curve. The right (left) geodesic curvature of a curve L at the point X is the limit of the ratio of the right (left) swerve of an arc of the curve L to the length of this arc as this arc contracts to the point X. Of course, this limit may fail to exist.

For sufficiently smooth curves on regular surfaces, this definition of geodesic curvature coincides with the usual definition,[10] and the left and right curvatures differ only by the sign on such surfaces (the sign of curvature shows the side of concavity of this curve). In this case, the swerve turns out to be the integral of the geodesic curvature with respect to the arc length. Thus, the concept of swerve is a generalization of the familiar concept of geodesic curvature, or, more precisely, of its integral. In the general studies concerning curves on arbitrary convex surfaces, the concept of swerve rather than the much more particular concept of geodesic curvature turns out to be useful due to the simple fact that the latter is not applicable to geodesic broken lines.

The right and left swerves can differ not only by the sign on nonregular surfaces. For example, the circle of the base of a right circular cylinder has zero geodesic curvature to the side of the lateral surface at each point, while this curvature is equal to $1/r$ to the side of the base, where r is the radius of the base. In what follows, we will show that the sum of the right and left swerves of a curve on a convex surface is equal to the spherical image of this curve without endpoints.

Let L be a closed curve without multiple points on a convex surface. We can distinguish the two sides of this curve which we agree to call right and left. The total swerve of this curve to the right (left) can merely be defined as the limit of swerves of broken lines that converge to L to the right (left). If there is a point A on the curve L such that both branches emanating from it have directions, the swerve of the curve L can be defined in another way. Namely, the curve L can be considered as nonclosed but such that its endpoints lie at one point A. Then the total swerve of the curve L to the right (left) is equal to the sum of the right (left) swerves of this nonclosed curve and the right (left) swerve at the point A. Additivity of swerve

[10] In what follows, by a regular surface we mean a surface that satisfies all usual propositions of differential geometry. (In this case, it is sufficient that a surface and a curve on it are twice differentiable.) In regard to the geodesic curvature of curves on regular surfaces see, e.g., W. Blaschke, Differential Geometry, Sec. 68 or P. K. Rashevskii, A Course of Differential Geometry, Sec. 61.

© 2006 by Taylor & Francis Group, LLC

implies easily that this sum is the same for each choice of the point A whenever both branches of the curve L have directions at this point. This second definition of swerve of a closed curve coincides with the first; this is immediate considering broken lines that converge to L from the right (left) and have a vertex at the point A. Such broken lines exist by existence of the directions of the branches of the curve L which emanate from the point A.

Obviously, the swerve of a closed curve without multiple points on the plane is equal to 2π. The following assertion holds on a convex surface.

Theorem 2. *If a curve L bounds a domain G homeomorphic to a disk and has curvature ω and the swerve of this curve L to the side of the domain G is equal to ϕ, then*

$$\omega = 2\pi - \phi.$$

In the case where L is a broken line, this assertion was already proved. If L is an arbitrary curve then we construct a sequence of broken lines L_n converging to L from the interior of the domain G. Let ϕ_n and ω_n denote the swerve of the broken line L_n and the curvature of the polygon bounded by this line, respectively. Then

$$\omega_n = 2\pi - \phi_n.$$

The definition of the swerve of a closed curve implies $\phi_n \to \phi$, while the complete additivity of the curvature yields $\omega_n \to \omega$; hence

$$\omega = 2\pi - \phi.$$

In exactly the same way, we prove the following more general assertion:

Theorem 2^***.** *If the interior of a compact domain bounded by curves whose swerves from the side of the domain are equal to $\phi_1, \phi_2, \ldots, \phi_n$ has curvature ω and Euler characteristic χ, then*

$$\omega = 2\pi\chi - \sum_{i=1}^{n} \phi_i.$$

Theorem 3. *The sum of the right and left swerves of a curve without multiple points is equal to the curvature (the area of the spherical image) of this curve without endpoints.*

Proof. Let a curve with endpoints A and B have right and left swerves ϕ_1 and ϕ_2. Include the curve L into a geodesic polygon P so that its ends lie on the boundary of this polygon. The curve L divides the polygon P into two domains so that P_1 lies to the right of L and P_2 lies to the left of L; the curvature of the interior of the whole polygon is

$$\omega(P) = \omega(P_1) + \omega(P_2) + \omega(L), \tag{6}$$

where $\omega(L)$ is the curvature of the curve L without endpoints.

The points A and B divide the boundary of the polygon P into two broken lines L_1 and L_2. Let ψ_1 and ψ_2 be the swerves of these broken lines counting inside the corresponding domains P_1 and P_2. Finally, let α_1 and α_2 be the angles, which are made by the curve L with the broken lines L_1 and L_2 at the point A, and let β_1

© 2006 by Taylor & Francis Group, LLC

and β_2 be analogous angles at the point B (we take interior angles of the domains P_1 and P_2). Then the swerves of the curves $L_1 + L_2$, $L_1 + L$, and $L_2 + L$ (i.e., the swerves of the boundaries of the polygon P and the domains P_1 and P_2 counting inside these domains) are equal to

$$\begin{aligned}\phi(L_1 + l_2) &= \psi_1 + \psi_2 + 2\pi - (\alpha_1 + \alpha_2 + \beta_1 + \beta_2),\\ \phi(L_1 + L) &= \psi_1 + \phi_1 + 2\pi - (\alpha_1 + \beta_1),\\ \phi(L_2 + L) &= \psi_2 + \phi_2 + 2\pi - (\alpha_2 + \beta_2).\end{aligned} \tag{7}$$

By Theorem 2,

$$\omega(P) = 2\pi - \phi(L_1 + L_2),$$

and, similarly,

$$\omega(P_1) = 2\pi - \phi(L_1 + L), \quad \omega(P_2) = 2\pi - \phi(L_2 + L).$$

Therefore, formula (6) implies

$$\omega(P) = \phi(L_1 + L) + \phi(L_2 + L) - \phi(L_1 + L_2) - 2\pi;$$

using formulas (7), we obtain

$$\omega(P) = \phi_1 + \phi_2;$$

as claimed.

We have presented only the basics of the theory of the swerve of a curve; many principal questions of this theory still remain unsolved. Before formulating some of them, we introduce one more definition (which will also be needed in the next section) where some applications of the concept of the swerve of a curve will be given.

Let a curve L have directions at each point to both its sides; then each arc of the curve has swerve; the latter, as we have shown, is an additive function of an arc. In accordance with the terminology of the theory of additive functions, we shall say that the swerve has bounded variation if the sum of absolute values of swerves of each finite family of disjoint arcs never exceeds some number. If t is a parameter varying along the curve from 0 to 1, then the swerve ϕ of the arc corresponding to the interval $(0, t)$ is a function of t. The requirement that the swerve has bounded variation is obviously equivalent to the requirement that this function $\phi(t)$ is a function of bounded variation. If, in particular, the swerves of all arcs of the curve have the same sign, then the function $\phi(t)$ is monotone, and the swerve certainly has bounded variation.

It is possible to prove that *if a curve L on a convex surface has swerve of bounded variation then this curve is rectifiable* (*i.e., has length*) *and its swerve is a completely additive function of an arc.* In particular, this holds for the curves all of whose arcs have the swerve to one side of the same sign; such curves will be considered in the next section.

When posing the following questions about swerve, we must imply the curves all of whose arcs have swerves and each swerve is a function of bounded variation; without this assumption, our questions will hardly be tractable.

© 2006 by Taylor & Francis Group, LLC

The first question consists in revealing the extrinsic-geometric sense of the swerve of a curve on any convex surface. If a curve L on a regular surface F has continuous geodesic curvature, then its swerve can be defined in the following way. Take consecutive points $X_0, X_1, \ldots, X_n$ on the curve L, where X_0 and X_n are the endpoints of the curve, and draw the tangent planes $P_0, P_1, \ldots, P_n$ to the surface F at these points. Let $T_0, T_1, \ldots, T_n$ be the tangents to the curve L at the points $X_0, X_1, \ldots, X_n$; they lie in tangent planes. Revolving the plane P_0 around the intersection line with the plane P_1, we can make it coincident with the plane P_2. Proceeding likewise, we unfold all planes $P_0, P_1, \ldots, P_n$ onto one plane P_n. The tangents $T_0, T_1, \ldots, T_n$ are unfolded together with the planes, and this unfolding defines a broken line on the plane P_n with vertices at the intersection points T_0 with T_1, T_1 with T_2, etc. Let ϕ be the swerve of this broken line. We shall take points on the curve X_i more and more densely and define the swerves ϕ of the broken lines, which are obtained by the method just described. The limit of these swerves ϕ yields the swerve of the curve L.

If the curve L lies on a nonregular convex surface F, then the above construction can become impossible simply by the fact that this surface can have no tangent plane at the points of the curve L. However, we can proceed as follows. At each point of the curve L, the surface F has tangent cone K, while the curve L has half-tangents, i.e., the half-lines that are tangent to its branches emanating from this point. These half-tangents divide the cone K into two sectors, and if we define the left and right sides of the curve, then one of these sectors lies to the right and the other sector lies to the left. Taking points $X_0, X_1, \ldots, X_n$ on the curve L and the right (left) sectors of the tangent cones at these points, we can sequentially unfold these sectors onto the plane similar to the above unfolding of tangent planes. Of course, this process of unfolding should be clarified, since each sector itself can fail to be flat. After unfolding the sectors, we can define the swerve of the broken line, formed by the unfolding half-tangents to the curve L. Take the limit of these swerves as the sequence of points $X_0, X_1, \ldots, X_n$ on the curve L indefinitely condenses. The following two questions are in order: (1) Does this limit exist if the swerve of the curve L has bounded variation? (2) Whenever this limit exists, does it equal to the right (left) swerve of the curve L?

If the surface F is a polyhedron then we can obtain a positive answer to these questions without much effort. In the general case, the answer remains unknown.

After the above posed problem of the extrinsic-geometric meaning of swerve, there arises another closely related problem. If a curve L on a regular surface F has continuous geodesic curvature, then it is easily seen that the above method for finding its swerve described above is equivalent to the following. Take the envelope of the tangent planes to the surface F at the points of the curve L. In a neighborhood of the curve L, this envelope is a developable surface without singularities. In this case, the curve L passes to some plane curve whose swerve is equal to the swerve of the curve L.

If the curve L lies on a nonregular convex surface F, then we draw tangent cones of the surface F through the points of the curve L. These cones jointly bound some convex body with surface F^*. It can be shown that the right and left swerves of the curve L, measured on the surface F, are equal to the corresponding swerves of this curve which are measured on the surface F^*. However, in general, the surface F^* is

© 2006 by Taylor & Francis Group, LLC

not developable. If the curve L has no continuous geodesic curvature but only has swerve of bounded variation, then even arbitrarily narrow strips of the surface F^* to the right and to the left of the curve L may fail to be developable. Therefore, the measurement of the swerve of a curve by unfolding these strips onto the plane turns out impossible in general.

However, we can address this question from another point of view. A developable surface tangent to a regular surface along the curve L is isometric to a narrow neighborhood of the curve L on the surface to within infinitesimals of higher order. Therefore, a neighborhood of the curve L can be mapped isometrically onto the plane to within infinitesimals of higher order; moreover, the curve L passes to a curve whose curvature is equal to the geodesic curvature of the curve L.

If the surface F is nonregular then instead of a neighborhood of the curve L, we have to consider right and left half-neighborhoods separately. The following question arises: is it possible to map a right (left) half-neighborhood of the curve L onto the plane in such a way that the following conditions hold: (1) the curve L passes to some curve L' each of whose arcs has the same length and the same swerve as the corresponding arc of the curve L and (2) near the curve L, this map is isometric to within infinitesimals of higher order?

Roughly speaking, the question is about the coincidence of the geometry of an infinitely narrow half-neighborhood of the curve L with the geometry of a half-neighborhood of the curve L'. For example, it is known that the variation of the length of a plane curve under infinitesimal displacements of its points depends on the curvature of this curve. It is natural to reveal to what extent we can describe the connection between the variation of the length and the swerve of the curve. In the general case, this connection can hardly be expressed by some equations; nevertheless, this question is still interesting in the general case, since it is probably related to variational problems. For example, one of these problems follows by way of example.

Let F be a closed convex surface. Consider closed curves lying on the surface F and dividing its area in a given proportion. There is a curve of minimal length among such curves; this is easy from the general theorems on convergence of rectifiable curves. If the surface F is regular then the usual methods of the calculus of variations yield the well-known result that this curve has constant geodesic curvature. The problem consists now in studying this curve on an arbitrary closed convex surface. (We are unaware of the answer.) A similar but not completely equivalent problem consists in searching for a curve of given length which bounds maximal area, or better, which divides the area of the whole surface in proportion that is as close to one half as possible.

3. The General Gluing Theorem

The concept of the swerve of a curve allows us to generalize the gluing theorem proved in Sec. 1 of Chapter VIII for the case when not only polygons but also arbitrary domains, bounded by curves having the swerve of bounded variation, are glued. Assume that closed domains $G_1, G_2, \ldots, G_n$ each of which is bounded by finitely many curves are cut out from manifolds with metric of positive curvature. These domains are not necessarily compact or bounded in the sense that the dis-

© 2006 by Taylor & Francis Group, LLC

tances between points of every domain do not exceed some number. For example, these domains can be parts of infinite convex surfaces. Therefore, the curves bounding these domains G_i can be of the following two types: closed curves homeomorphic to a circle and open curves homeomorphic to a line. However, we assume that each arc of such a curve has a swerve, and the swerve of the whole curve is of bounded variation in the sense of the preceding section. (By an arc, we mean each segment of a curve which is homeomorphic to a line segment.) Then we can prove that each arc of these curves has finite length; the curve itself can "go to infinity" and, thus, can have infinite length. (However, in order to avoid appealing to the theorems that are not proved yet, we can merely claim that each bounded arc constituting the boundary of the domain G_i has finite length.)

Each domain G_i has its own intrinsic metric induced from the metric of the manifold from which this domain is cut out, and this metric coincides with the latter in small neighborhoods (see Theorem 5 of Sec. 2 of Chapter II). We must say that a manifold with intrinsic metric R' originates by gluing from the domains $G_1, \ldots, G_n$ if this manifold can be partitioned into domains $G'_1, \ldots, G'_n$ isometric to the domains $G_1, \ldots, G_n$, respectively. If the manifold R' is a surface in space, it is natural to say that this surface is obtained by gluing the domains $G_1, \ldots, G_n$.

Some segments of the boundaries of the domains $G'_1, \ldots, G'_n$ are pairwise identified; these segments can be called the sides of the domains $G'_1, \ldots, G'_n$. Those points at which more than two domains meet can be called vertices. We assume that each domain has finitely many sides and vertices. The identification of sides and vertices of the domains G'_i can be transferred to the initial domains G_i by isometry, and then these domains themselves form a manifold R homeomorphic to R'. An intrinsic metric is defined in the manifold R in a natural way. The length of a curve on the manifold R can be defined as the sum of the lengths of its segments each of which lies in its own domain G_i. After that, the distance between points in R is defined as the greatest lower bound of the lengths of the curves connecting these points. With this definition of the metric on the manifold R, this manifold is isometric to the manifold R'; therefore, if R is an abstract manifold rather than a surface in space, we can make no distinction between the manifolds R and R' and also say that the manifold R is obtained by gluing the domains $G_1, \ldots, G_n$.

A manifold originating by gluing from some domains is completely determined from indication of those parts of the boundaries of the domains which must be identified. This "gluing law" must satisfy the following usual conditions.

1. The domains meeting at one vertex must approach it in the same way as the sectors partitioning a disk meet at the center of this disk. This immediately implies that the sides are pairwise identified.

2. We can pass from one domain to another traversing some domains that have identified sides.

3. The identified sides as well as identified segments must have equal lengths.

Under the first two conditions, the domains we glue form a manifold. Thus, given the domains $G_1, \ldots, G_n$ and indicating the "gluing law" which satisfies the above three conditions, we obtain a manifold with intrinsic metric R. The question

© 2006 by Taylor & Francis Group, LLC

is to find conditions under which this metric is a metric of positive curvature. An answer is given by the following "general gluing theorem."

Theorem. *The metric of the manifold R, obtained by gluing by domains $G_1, \dots, G_n$ with metrics of positive curvature, is itself a metric of positive curvature if and only if the following two conditions hold:* (1) *the sum of swerves of every two identified segments of the boundaries of domains is nonnegative;* (2) *the sum of angles of the domains G_i meeting at one vertex does not exceed 2π for each vertex.* (Of course, we speak here about those domains that satisfy the above conditions relative the swerves of their boundaries; we imply the swerves of the segments of the boundaries to the side of the corresponding domains.)

In essence, we are already aware of the necessity of these conditions. Indeed, according to Theorem 3 of the preceding section, the sum of swerves to both sides of each curve in a manifold with a metric of positive curvature is equal to the curvature of this curve without endpoints. Hence this sum is always nonnegative, which proves the necessity of the first condition. The necessity of the second condition follows from the fact that the complete angle at a point in a manifold of positive curvature cannot exceed 2π.

A complete proof of the sufficiency of the above conditions entails some difficulties in setting details; in essence, overcoming them requires mostly obvious but rather tedious arguments. Therefore, we do not consider these details here and only sketch the idea of the proof which is completely similar to the idea of proof of the gluing theorem for polyhedra in Sec. 1 of Chapter VIII. The proof reduces to the consecutive verification of the following properties of the manifold R, obtained by gluing the domains G_i: (1) the metric of R is intrinsic; (2) there are angles between the sides of each convex triangle in R; (3) the sum of the angles of each small convex triangle is no less than π; (4) there is an angle in the strong sense between the sides of a convex triangle in R. Of course, these properties imply that the metric of R has positive curvature.

The fact that the metric on R is intrinsic is clear from its definition. If a point O lies inside one of the domains G_i, then each pair of shortest arcs emanating from this point makes an angle, since the metric on G_i is a metric of positive curvature. Let the point O lie on the boundary of the domains G_i, and let L and M be the sides of a convex triangle T which meet at this point. Some segments $N_1, \dots, N_m$ of the boundaries of the domain G_i can go from the point O in the triangle T. Each pair of neighboring segments belongs to one domain and, therefore, makes an angle. If each of the shortest arcs L and M also belongs to one of the domains G_i near the point O, then they make angles with the neighboring segments N_1 and N_m of the boundaries. Hence, the convex sector between L and M is partitioned into sectors with some angles (between L and N_1, N_1 and $N_2, \dots, N_m$ and M), and then, by Theorem 3 in Sec. 1 of Chapter IV, there is an angle between L and M themselves. A singular case is, e.g., when the shortest arc L intersects a segment of the boundary arbitrarily close to the point O and passes infinitely many times from one domain to the other. In this case, the angle between L and N_1 is simply equal to zero, but this fact requires special proof, since L does not lie in one domain here, and, therefore, we cannot simply refer to the fact that in each domain a shortest

© 2006 by Taylor & Francis Group, LLC

arc makes an angle with a segment of the boundary. However, we will not elaborate this detail.

Now let T be a convex triangle on R, and let $\omega(T)$ be the curvature of its interior, i.e., the sum of the angles minus π. If T lies in one of the domains G_i, then certainly $\omega(T) \geq 0$. Assume that the triangle intersects several domains G_i. Then this triangle is partitioned into domains $Q_1, Q_2, \ldots$ each of which lies in one domain G_i. Since the sides of the triangle T can intersect the boundaries of the domains G_i infinitely many times, the number of the domains Q_j can be infinite. Since each domain Q_j lies in one domain G_i, its curvature $\omega(Q_j) \geq 0$. Calculation of curvatures, similar to that in the deduction of formula (1) for the curvature of a polygon in Sec. 1 of Chapter 5, leads to the following result: the curvature of the interior of the triangle T is equal to the sum of the curvatures of the domains Q_j plus the sum of the curvatures of the segments of their boundaries lying inside T plus the sum of the curvatures of the vertices of the domains Q_j that also lie inside T. The curvature $\omega(N_p)$ of the segment N_p of the common boundary of the domains Q_j and Q_k is equal to the sum of its swerves, and since this sum is nonnegative by the condition of the theorem, we have $\omega(N_p) \geq 0$. The curvature $\omega(X_q)$ of the common vertex of several domains Q_i is equal to 2π minus the sum of the angles meeting at this vertex; by the condition of the theorem, this sum is $\leq 2\pi$, and hence $\omega(X_q) \geq 0$. Thus,

$$\omega(T) = \sum_j \omega(Q_j) + \sum_p \omega(N_p) + \sum_q \omega(X_q);$$

herewith, all summands are nonnegative here. Hence $\omega(T) \geq 0$, i.e., the sum of the angles of the triangle T is no less than π. This argument, as well as the proof of the existence of an angle, which was presented above, needs refining if the number of the domains Q_j is infinite. In this case, calculation of the sum of curvatures requires certain caution, and, in particular, here the assumption that the swerve of the boundaries of the domains G_i has bounded variation plays an essential role.

It remains to prove that there is an angle in the strong sense between the sides of a convex triangle on R. If each of the two sides L and M of the triangle emanating from a common point O travels only in one of the domains G_i at least near the point O, then the proof of this assertion is literarily the same as in the gluing theorem of polygons in Sec. 1 of Chapter VIII. If at least one of these shortest arcs L and M passes from one domain to the other arbitrarily close to the point O, then some additional arguments are needed; however, we will not expatiate on them.

If the manifold R turns out to be such that we may realize this manifold as a convex surface, e.g., if it is homeomorphic to the sphere, the gluing theorem transforms to the existence theorem of a convex surface obtained by "gluing" some pieces of the given manifold or, in particular, some pieces of given convex surfaces. Gluing a surface from pieces is omnipresent. The simplest case that does not reduce to the gluing of polyhedra is the gluing of a convex surface from pieces of the plane. The well-known methods of gluing cylinders and cones or an approximate construction of the sphere from digons bounded by equal arcs of circles can serve as examples.

If the sum of the swerves of two glued segments of the boundary of domains is positive, then an edge is obtained on the glued surface, since the swerve sum

© 2006 by Taylor & Francis Group, LLC

of a curve to both sides is equal to the area of its spherical image (Theorem 3 of Sec. 2). If the sum of the swerves of two glued segments of the boundaries is equal to zero, then either no edge results or we obtain a rectilinear edge as in the case of a polyhedron.

In the case where the curves that bound the glued domains have piecewise-continuous geodesic curvature, the condition of the sum of swerves reduces to the fact that the sum of geodesic curvatures of the glued curves is nonnegative everywhere. For example, we can glue a closed surface from a spherical segment by bending this segment so that both halves of its boundary circle coincide.

However, we have to keep in mind that in the general case we do not prove at all that the gluing of a convex surface from pieces can be performed in reality in the sense that these pieces can continuously be bent into the corresponding pieces of the surface we glue. The proof of the possibility of such a bending with possible addition of a reflection is a very interesting but, possibly, very difficult problem.

In connection with the gluing theorem, we can pose the problem of characterization of the metric of an arbitrary convex surface. Let F be a bounded convex surface; the boundary of its convex hull is a closed convex surface $\overline{F}$ on which the "holes" possibly present in the initial surface were glued by developable surfaces. (This follows from the fact that the areas of the spherical images of these surfaces without boundary are equal to zero.) This remark, together with the realization theorem of a metric on the sphere, leads to the following result: a compact domain G on a manifold with metric of positive curvature, homeomorphic to a domain on the sphere, is isometric to a convex surface if and only if we can glue plane domains[11] to it so that the conditions of the gluing theorem are fulfilled, i.e., so that the swerves of the segments of the boundary of these domains are no less than the swerves of the corresponding segments of the boundary of the domain G taken with the opposite sign. Thus, the question of the existence of a surface isometric to the domain G reduces to the question of the existence of plane domains that satisfy this condition. But this is a problem of plane geometry which seems to involve no insurmountable difficulties. It is interesting to study this problem at least in the case where the domains we glue to G are polygons.

4. Convex Domains

As an application of the concept of the swerve of a curve, consider curves that bound convex domains of convex surfaces.

Theorem 1. *Each arc lying on the boundary of a convex domain has a nonnegative swerve to the side of this domain.*

Proof. Let O be a point on the boundary of a closed convex domain G, and let L be an arc of the boundary of the domain G that emanates from the point O (Fig. 79). Take a point Y on L and draw a shortest arc OX in the domain G; this is possible by the very definition of a convex domain. If OX does not go along the arc L, then L has direction at the point O. If OX does not go along L and the point X lies sufficiently close to O, then OX passes inside the domain G and separates

[11] Strictly speaking, these domains can be not plane domains in the rigorous sense failing to unfold onto the plane univalently.

© 2006 by Taylor & Francis Group, LLC

a convex domain G_1 from G that is bounded by the shortest arc OX itself and by the segment $\overset{\frown}{OX}$ of the arc L. If we take a point X_1 on $\overset{\frown}{OX}$, then a shortest arc OX_1 goes in G_1 and separates a new convex domain G_2 from G_1. Further, we can take a point X_2 on the segment $\overset{\frown}{OX_1}$ of the arc $\overset{\frown}{OX}$, and so on. The angle between the shortest arcs OX_n and the shortest arc OX_1 increases (or at least does not decrease) and, therefore, tends to a limit. As we have proved in Sec. 1, this means that the arc L has direction at the point O. This proves that each arc of the boundary of a convex domain has directions at the endpoints and, thus, has swerve.

We can inscribe into the arc L a broken line L' that lies in the domain G. The angles of this broken line to the side of G are no greater than π, since the domain G is convex. Hence the swerve of the broken line L' to the side of G is nonnegative. If we increase the number of vertices of the broken line L', then this line tends to the arc L, and the angles between L and L' tend to zero at the endpoints of L as is clear from the above proof of existence of the directions at the endpoints of the arc L. Therefore, the swerve point of the arc L to the side of the domain G is equal to the limit of the swerve points of the broken line L' and hence is also nonnegative.

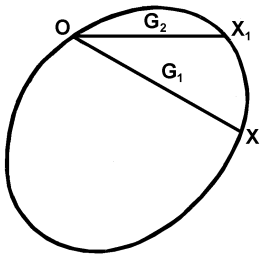


Fig. 79

Theorem 1 is essentially supplemented by the following assertions: (1) each arc on the boundary of a convex domain has length; (2) the ratio of the length of such an arc $\overset{\frown}{OX}$ to the distance OX tends to 1 as $X \to O$ and the point O is fixed; (3) the swerve point of $\overset{\frown}{OX}$ tends to zero as $X \to O$ and the point O is fixed. The proof of these assertions is left to the reader.

Theorem 2. *A compact (i.e., closed and bounded) convex domain on a closed surface can be only one of the following types:* (1) *the whole closed surface;* (2) *a domain homeomorphic to a disk;* (3) *a domain homeomorphic to the lateral surface of a right circular cylinder.*

Proof. Each bounded domain G on a convex surface can be considered as a part of a closed convex surface F. If the domain G is the whole surface F, then we have the first case. Assume that the convex domain G occupies only a part of the surface F. *A priori,* the surface G can be bounded by infinitely many curves; but first, we consider the domain G bounded by finitely many curves. Let ω be the curvature of the interior of the domain G, and let τ_i be the swerves to the side of G of the curves bounding this domain. Then by Theorem 2^* in Sec. 2,

$$\omega = 2\pi\chi - \sum_i \tau_i, \tag{1}$$

where χ is the Euler characteristic of the domain G. We know that $\omega \geq 0$, and as we just have proved, all swerve points $\tau_i \geq 0$. Therefore, we see that $\chi \geq 0$. As is known, the domains on the sphere with Euler characteristic $\chi \geq 0$ can be of the following three types only: (1) $\chi = 2$; the domain is homeomorphic to the

© 2006 by Taylor & Francis Group, LLC

sphere; (2) $\chi = 1$; the domain is homeomorphic to a disk; (3) $\chi = 0$; the domain is homeomorphic to an annulus or, which is the same, to the lateral surface of a cylinder.[12] The first two types yield the first two cases of the theorem. Assume that $\chi = 0$. Then we have only two curves L_1 and L_2 bounding the domain G. Since their swerves τ_1 and τ_2 cannot be negative, formula (1) implies $\tau_1 = \tau_2 = 0$ and $\omega = 0$. We have proved in Sec. 6 of Chapter V that a domain of zero curvature is locally isometric to the plane, i.e., the metric is Euclidean in small parts of this domain. Since the swerves of the curves L_1 and L_2 to the side of the domain G are equal to zero, the arcs of these curves go to line segments under the isometric mapping onto the plane. These properties are characteristic of the lateral surface of a right cylinder, and, hence, the domain G is isometric in the large to such a surface.

We have assumed that the domain G is bounded only by finitely many curves. Assume now that this domain is bounded by infinitely many curves. Each curve surrounds some domain on the closed surface F on which G lies. If we add an arbitrary number of such domains to G, then it is easily seen that we again obtain a convex domain. Thus, we can replace G by another convex domain G' that is bounded by a finite but arbitrarily large number of curves. But we have just proved that if the number of the curves bounding a convex domain is finite, then it is no greater than 2. Consequently, it is not possible that a convex domain is bounded by infinitely many curves, and so the proof of the theorem is complete.[13]

A convex surface F, bounded by a closed plane curve L, such that the projection of F onto the plane of this curve coincides with the domain bounded by L is called a *cap*. A spherical segment is an example of a cap.

Theorem 3. *Each convex domain on a convex surface homeomorphic to a disk is isometric to a cap.* In other words, each convex domain on a convex surface homeomorphic to a disk can be isometrically mapped so that its boundary transforms to a plane curve and its projection to the plane of this curve coincides with the convex domain bounded by the curve.

[12] The Euler characteristic of the sphere is equal to 2. When we exclude a domain homeomorphic to a disk from the sphere, the Euler characteristic decreases by 1; this becomes completely obvious if we take a partition of the sphere one of whose domains is exactly the excluded domain.

[13] In much the same way as the above study of the construction of compact convex domains, we can determine all types of infinite convex domains on infinite complete convex surfaces. For this, it is necessary to use the connection between the swerve of the boundary with the curvature of a domain for infinite domains. The swerve of an infinite curve is defined as the limit of the swerve of its indefinitely enlarging arcs (under the assumption that the swerve points of these arcs exist and there is a limit of this swerve independent of the choice of a sequence of enlarging arcs). Accepting this definition, we can prove the following: Let G be a domain on an infinite complete convex surface homeomorphic to a half-plane, let ω be the curvature of this domain G, and let τ be the swerve of the curve, bounding G (the swerve to the side of G); then $\omega \leq \pi - \tau$. We leave it to the reader to prove this assertion as an interesting problem, to obtain a more general result for an infinite domain of arbitrary structure, and then to apply this more general result for seeking all types of infinite convex domains on complete convex surfaces. See the paper by Cohn-Vossen in Matematicheskii Sbornik for 1936 which was already cited in Sec. 10 of Chapter I, for the interplay of curvature and swerve in the case of infinite domains. Cohn-Vossen considers a metric given by a line element, but his methods may be abstracted in general to an arbitrary metric of positive curvature.

© 2006 by Taylor & Francis Group, LLC

Proof. The proof is based on the gluing theorem proved in Sec. 1 of Chapter VIII.

We first consider a convex polygon P on a convex surface which is homeomorphic to a disk. Take an infinite circular cylinder with the length of some cross-section equal to the perimeter of the polygon P. Cut up this cylinder along some cross-section, and let C be one of the right "half-cylinders" obtained in such a way.

Identify the boundary of the polygon P with the boundary of the half-cylinder C and consider the resulting manifold $P+C$. Since all angles of the polygon P are no greater than π and the boundary of the half-cylinder C is a closed geodesic, the metric of the manifold $P + C$ has positive curvature in accordance with the gluing theorem. At the same time, this manifold is homeomorphic to the plane and its metric is complete, since the half-cylinder C is infinite. Therefore, by the theorem of Sec. 3 of Chapter VIII, the manifold $P+C$ is isometric to some infinite convex surface. This surface consists of two surfaces P' and C' that are isometric to the polygon P and the half-cylinder C, respectively (Fig. 80). But it is easy to prove that every surface that is isometric to an infinite right cylinder is itself a right half-cylinder.[14] Hence the surface C' is a right half-cylinder, and its boundary and, thus, the boundary of the surface P' is some plane curve L. Since the generators of the half-cylinder C are orthogonal to the plane of the curve L, the projection of the surface P' on this plane coincides with the domain bounded by the curve L; otherwise, the surface $P' + C'$ is not convex. Consequently, P' is a cap, and, thus, our theorem is proved for polygons.

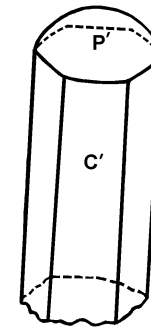


Fig. 80

Now let G be an arbitrary convex domain on a convex surface which is homeomorphic to a disk. Inscribe a sequence of polygons P_n, which converge to G, into G. Each polygon P_n is convex and, therefore, is isometric to some cap P_n'. We can choose a convergent sequence from the caps P_n', and the limit of this sequence is obviously a cap isometric to the domain G by the theorem on the convergence of metrics. (These conclusions are easily justified, e.g., by the method of the proof of the realization theorems in Sec. 7 of Chapter VII and in Sec. 3 of Chapter VIII.)

If we use the general gluing theorem, then we may immediately glue the domain G with a suitable half-cylinder and, in this way, obtain the required result.

Now it is natural to pose the question on the possibility of conversing Theorems 1 and 2, i.e., to ask (1) if it is true that each domain on a convex surface bounded by curves whose every arc has nonnegative swerve is a convex domain and (2) if it is true that each cap cut out from a convex surface is a convex domain on this surface. It is easy to see that both questions have a negative answer. For example, take a closed surface F consisting of a spherical segment S and a plane disk C glued to this segment. The segment S is a cap, and each arc of its boundary has a positive swerve. However, this segment is not a convex domain on the surface F,

[14] The surface C' isometric to a half-cylinder is developable. Let L be one of its rectilinear generators; this generator has the property that each of its arbitrarily long segment is a shortest arc. With the line L, we associate a curve on the half-cylinder with the same property. But only the generator on the half-cylinder has this property. Consequently, the generator of the cylinder C and the surface C' correspond to each other, which makes it clear that C' is also a semicylinder.

© 2006 by Taylor & Francis Group, LLC

since a shortest arc connecting points on its boundary certainly travels in the disk C. Despite the fact that the question of conversing Theorems 1 and 2 has a negative answer as we see, it is possible, however, to generalize the concept of convex domain so that for the domains convex in this generalized sense Theorems 1 and 2, as well as the theorems converse to them, are valid. This generalization of the concept of convex domain consists in the following.

A closed domain on a convex surface is called *convex in itself* if every two its interior points can be connected by a curve that passes inside the domain and is the shortest among all these curves. In other words, every two interior points can be connected by a line that is shortest in the domain but, probably, not on the whole surface and that lies inside the domain.[15]

All general properties of convex domains, proved in Sec. 4 of Chapter II, turn out true for the domains that are convex in themselves if by shortest arcs we mean the shortest arcs in the domain itself. For the domains convex in themselves, the following theorems also hold; they contain Theorems 1, 2, and 3 and the theorems converse to Theorems 1 and 3.

Theorem 1^***.** *A closed domain G bounded by a curve L on a convex surface is convex in itself if and only if each arc of the curve L has nonnegative swerve to the side of G.*

Theorem 2^***.** *A closed bounded domain on a convex surface convex in itself can be one of the following three types:* (1) *a complete closed surface;* (2) *a domain homeomorphic to a disk;* (3) *a domain isometric to the lateral surface of a right circular cylinder.*

Theorem 3^***.** *Each domain on a convex surface which is homeomorphic to a disk and convex in itself is isometric to every cap, and any cap is convex in itself.*

We do not provide the proofs of these theorems;[16] we will show only that *every cap is convex in itself and each arc of a curve bounding a cap has nonnegative swerve to the side of the cap.*

To prove this, we take some cap G; let L be the curve bounding this cap, and let P be the plane of this curve. We draw a half-line from each point of the curve L which is orthogonal to the plane P and aims at the half-space that has no points of the cap G. These half-lines form a half-cylinder with base curve L; we obtain the infinite complete convex surface F consisting of the cap G and this semicylinder C. The cap G is a convex domain on the surface F. Indeed, if a shortest arc AB connecting two points A and B of the cap G passes not in the cap itself, then

[15] Above, e.g., in Sec. 2 of Chapter VIII, we spoke about a surface whose every two points can be connected by a shortest arc. Here, we speak about a closed domain on a surface. But each domain on the surface is itself a surface by definition, and, therefore, this distinction is immaterial. An open domain convex in itself differs from a closed domain by the fact that its boundary is excluded and there can be some set of points such that the complete angles at them can be $< 2\pi$.

[16] In the part where they merely generalize Theorems 1, 2, and 3, these theorems are proved in a similar way. Of course, before proving Theorem 1^*, it is necessary to prove the properties of convex domains which were used in the proof of Theorem 1 for the domains convex in themselves. It is not absolutely easy, but when this is done, the further arguments repeat those that prove Theorems 1, 2, and 3.

© 2006 by Taylor & Francis Group, LLC

its segment lying on the half-cylinder C is a shortest arc connecting two points of the base line L of this half-cylinder. But a shortest arc on the half-cylinder that connects two points of its plane directrix, which is orthogonal to its ruling, obviously overlaps the directrix. Therefore, the shortest arc AB travels in the cap, and hence the cap is a convex domain on the surface F. This cap is also convex in itself, since, according to Theorem 1 in Sec. 5 of Chapter II, a shortest arc between interior points of a convex domain travels inside this domain. Moreover, according to Theorem 1, each segment of the boundary of a convex domain has nonnegative swerve to the side of this domain. Consequently, the curve bounding the cap G has this property.

From the general gluing theorem it follows easily that conversely, each domain on a convex surface which is homeomorphic to a disk and bounded by a curve such that the swerve of its every arc to the side of the domain is nonnegative is isometric to a cap. Indeed, if we glue an infinite semicylinder C to such a domain G, then, by the gluing theorem, we obtain a manifold with complete metric of positive curvature which is homeomorphic to the plane. The surface, isometric to this manifold, consists of the semicylinder corresponding to C and the cap isometric to the domain G; this is proved in exactly the same way as in Theorem 3 for convex domains.

This argument is also applicable in the case where G is a domain not on a surface but on an arbitrary manifold with the metric of positive curvature. This remark, together with Theorem 1^* generalized to domains in abstract manifolds, leads to the following theorem.

Theorem 4. *Let G be a domain on a manifold with metric of positive curvature which is homeomorphic to a disk and bounded by a curve L. The following three conditions turn out to be equivalent:* (1) *the domain G is convex in itself:* (2) *each arc of the curve L has nonnegative swerve to the side of G;* (3) *the domain G is isometric to a cap.*

This theorem contains a complete characterization of the intrinsic metric of caps and is a refinement of the general realization theorem which was discussed in Sec. 2 of Chapter VIII (Theorem 5 in Sec. 2 of Chapter VIII) for the case of domains convex in themselves and homeomorphic to a disk. In the general theorem in Sec. 2 of Chapter VIII, we spoke about the realization of a metric of positive curvature in each domain convex in itself and homeomorphic to a domain on the sphere. Each nonclosed domain of this sort is obtained from a closed domain by deleting the boundary curves and, probably, some set of the points lying in no shortest arc. Theorem 3^* reveals the structure of all closed domains convex in themselves, and hence it reveals the structure of all nonclosed domains of this sort to the extent of what is known about the points lying in no shortest arcs.

Theorem 4 opens a way for studying the curves that are not necessarily closed and such that the swerve of each of their arcs to one side is nonnegative. For example, we sketch the proof of the following theorem.

Theorem 5. *Let convex surfaces F_n converge to F, and let curves L_n lying on the surfaces F_n converge to a curve L. Assume that the curves L_n have no multiple points and the swerves of all their arcs to one side are nonnegative. Then, for all*

© 2006 by Taylor & Francis Group, LLC

points A and B lying inside the curve L, the arc AB of this curve has nonnegative swerve to one side, and the length of AB is equal to the limit of the lengths of the arcs of the curves L_n that converge to AB.[17]

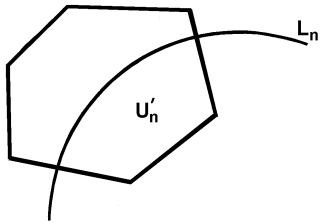


Fig. 81

To prove this theorem, we take some point O inside the curve L and the points O_n on the curves L_n that converge to O; then we surround them by small convex neighborhoods U_n so that the latter converge to some neighborhood U of the point O. The curves L_n divide the neighborhoods U_n, and since their swerve to one side is nonnegative and the neighborhoods U_n themselves are convex, one of the parts U'_n in each of the neighborhoods U_n is bounded by a curve such that all its arcs have nonnegative swerve to the side of U'_n (Fig. 81). Therefore, by Theorem 3, each domain U'_n can be mapped isometrically onto some cap U''_n. The limit of these caps is obviously a cap; and, at the same time, it is isometric to one of the pieces U' of the neighborhood U. By what we have proved above, each arc of the boundary of the cap has nonnegative swerve. Therefore, in particular, the arc of the curve L, which cuts out the piece U' from the neighborhood U, has nonnegative swerve. But O is an arbitrary point inside the curve L, so that the same is true for each sufficiently small arc L lying inside L_0. Summing the swerves of separate arcs, we find that each arc lying inside L has nonnegative swerve.

The boundaries of the caps are plane convex curves. Therefore, if the caps U''_n converge, then the length of each arc l of the boundary of the limit cap is equal to the limit of the lengths of the arcs of the boundaries of U''_n which converge to l. This implies that the length of each sufficiently small arc of L is equal to the limit of the arcs of L_n which converge to it. Summing the lengths of small arcs, we obtain the same result for every arc lying inside the curve.

5. Quasigeodesics

Among curves on convex surface, having the swerve of their every subarc, nonnegative to one side, of special interest are the curves having the swerve of their every arc nonnegative to both sides. These curves are called *quasigeodesics* since it will become clear in the sequel that these curves are a natural generalization of geodesics. The swerves to both sides of all arcs of a geodesic are equal to zero, so the geodesics prove a particular case of quasigeodesics. Conversely, on regular surfaces each quasigeodesic is a geodesic. Indeed, by Theorem 3 of Sec. 2, the sum of the right and left swerves of a curve is equal to the area of its spherical image, and on a regular surface, this sum is thus equal to zero. Therefore, the swerves of all arcs of a quasigeodesic on a regular surface is equal to zero, i.e., the geodesic curvature of a quasigeodesic is equal to zero. As is known, on a regular surface a

[17] The curve L_n itself can have no definite swerve, since it can turn around its endpoints in a spiral way surrounding each of them infinitely many times. If the curves L_n have multiple points, then the theorem is no longer true. For example, we can take the curves of positive curvature on the plane which have loops contracting to a point. At this point, the swerve of the limit curve can be negative.

© 2006 by Taylor & Francis Group, LLC

curve of zero geodesic curvature is a shortest arc on each sufficiently small segment, i.e., is a geodesic.

There can exist quasigeodesics on nonregular surfaces which are not geodesics. It is easy to see that on a convex polyhedron, a curve is a quasigeodesic if and only if it is a geodesic broken line whose vertices lie at the vertices of the polyhedron; herewith, both angles between segments of this broken line meeting at a vertex should be $\leq \pi$. Therefore, e.g., the sides of one face of a cube form a closed quasigeodesic on it.

The fact that each arc of the boundary of a cap has a nonnegative swerve to the side of the cap implies that a curve that divides a closed convex surface into two caps is a quasigeodesic. The edge of a doubly-convex lens or the base circle of a circular cone are examples of such quasigeodesics (in the latter case, one of the caps proves to be flat.)

It is worthy of note that as demonstrated by examples even curves of zero geodesic curvature on nonregular surfaces are not always geodesic, i.e., shortest arcs on each sufficiently small part.

Theorem 5 of the previous section yields immediately that the limit of a convergent sequence of quasigeodesics without multiple points is again a quasigeodesic. This result can easily be generalized to quasigeodesics with multiple points, and then we obtain the following assertion.

Theorem 1. *The limit of quasigeodesics is a quasigeodesic. In more detail, let convex surfaces F_n converge to a surface F, and let quasigeodesics L_n on the surfaces F_n converge to a line L. Then L is a quasigeodesic on the surface F and its length is equal to the limit of lengths of the quasigeodesics L_n. (This limit exists by the very fact of the convergence of L_n).*

The last part of the theorem ensues from the corresponding assertion of Theorem 5 of the previous section.

Since geodesics are quasigeodesics; therefore, in particular, *the limit of geodesics is a quasigeodesic.*

To some extent, this result can be converted. To this end, consider some quasigeodesic on a convex surface F. Take an arc L of this quasigeodesic which has no multiple points and cuts out a domain G from the surface F so that the arc L divides this domain into two parts. Cut the domain G along the arc L and "glue" a plane rectangle P to both sides of this cut by the opposite sides of P. The lengths of the bases of the rectangle must be equal to the length of the arc L. Since the swerves of all segments of the arc L are nonnegative, the so obtained manifold has a metric of positive curvature by the general gluing theorem. The median of the rectangle P is obviously a geodesic in this convex metric. If we tend the height of the rectangle P to zero, then the metric in the manifold R tends to the metric of the domain G and the median of the rectangle P converges to the arc L. This result can somewhat loosely be expressed in the form of the following theorem.

Theorem 2. *A quasigeodesic without multiple points is the limit of geodesics in the sense of the intrinsic metric.*

It is likely that each sufficiently small arc of a quasigeodesic on a convex surface is the limit of geodesics on surfaces which converge to this surface not only in the

© 2006 by Taylor & Francis Group, LLC

sense of the intrinsic metric. At the same time, it is worth noting that not any, even arbitrarily small, arc of a quasigeodesic is the limit of geodesics on the same surface. For example, an arbitrarily short quasigeodesic on a convex polyhedron which passes through the vertex and makes the angles less than π at this vertex is not the limit of geodesics of the same polyhedron.

The limit of geodesics can fail to be a geodesic. For example, the limit of geodesics on a convex polyhedron whose midpoints tend to a vertex of the polyhedron is not a geodesic. Hence the class of geodesics is not closed, and it is not possible in general to pass to the limit in this class. According to Theorem 1, the class of quasigeodesics is closed, and by Theorem 2, this class consists, in a sense, of the limits of geodesics. We can say that the class of quasigeodesics is the closure of the class of geodesics on all convex surfaces and is the shortest class of lines that contains all geodesics and at the same time admits passage to the limit.[18]

As for geodesics, we have proved early that (1) there are points on convex surfaces from which geodesics emanate not in all directions (Sec. 5 of Chapter V) and (2) if a geodesic emanates from a given point in a given direction, then it is unique, since the angle made by nonoverlapping shortest arcs and, hence, by geodesics cannot be equal to zero. It turns out that for quasigeodesics, the converse statement holds, namely, we can prove the following theorem.

Theorem 3. *A quasigeodesic can be drawn in each direction from each point on an arbitrary convex surface. But there may exist directions in each of which it is possible to draw several quasigeodesics rather than one.*

We sketch a construction that leads to a quasigeodesic emanating from a given point O in a given direction d. Take two small positive numbers r and ε. If a shortest arc emanates from the point O in the direction d, then this shortest arc is a desired quasigeodesic. If there is no shortest arc emanating from O in the direction d, then there exists a digon D_1 that includes the direction d whose angle at the vertex O is less than ε (this was proved in Sec. 5 of Chapter V). Let X_1 be the other vertex of this digon, and let L_1 and L_1' be its sides (Fig. 82). The angle between L_1 and L_1' at the point X_1 is greater than zero, and so we can draw a shortest arc L_2 from the point X_1 that lies outside of the digon D_1 and makes the angles $< \pi$ with L_1 and L_1'. (This follows for the fact that shortest arcs emanate from X_1 in almost all directions as was proved in Sec. 5 of Chapter V.) Let X_2 be the endpoint of the shortest arc L_2. If it is impossible to extend the shortest arc L_2 beyond the point X_2, then there exists a digon D_2 with sides L_3 and L_3' whose angles with L_2 are

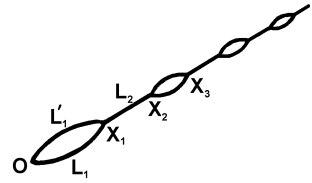


Fig. 82

[18] The precise formulation of this assertion is as follows. A class of curves is called closed if it contains the limit of every convergent sequence of curves in this class. Consider the class K of lines on all convex surfaces which satisfies the following conditions: (1) K contains all shortest arcs; (2) the class K is closed; (3) if a curve L on a convex surface F belongs to K then the corresponding line on each surface isometric to F also belongs to K; (4) if each point of a line L can be placed in a segment of this line which belongs to K, then L itself belongs to K. Among these classes K, there is one class included in each of them; this is exactly the class of quasigeodesics; a single point is also viewed as a quasigeodesic here.

© 2006 by Taylor & Francis Group, LLC

both less than $< \pi$ and the angle between L_3 and L_3' is less than ε. Further, we again draw a shortest arc L_4 from the vertex X_3 of the digon D_2, and so on. As a result, we obtain two broken lines $L_1 + L_2 + L_3 + L_4 + \ldots$ and $L_1' + L_2' + L_3' + L_4' + \ldots$ that have nonnegative swerves each to its own side everywhere. If such broken lines do not attain the given length r, then we continue the construction infinitely and obtain broken lines with infinitely many sides which have a common direction d' at their endpoint Y. Further, we continue the constriction beyond the point Y by drawing a digon whose sides make the angles less than π with the direction d' and such that the angle between these sides is less than ε, and so on. As a result, we arrive at two broken lines L and L' of length r with infinitely many sides which have nonnegative swerves each to its own sides everywhere. The proof of this fact can certainly be performed without transfinite arguments but rather on considering the least upper bound of the lengths of such broken lines as usual.

If we now let ε tend to zero, the broken lines L and L' converge to the same curve, which, by Theorem 5 of the previous section, has a nonnegative swerve to both sides, i.e., it is a quasigeodesic. It seems quite obvious and we can prove it that this geodesic emanates from the point O exactly in the given direction d.

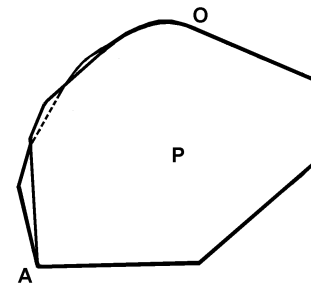


Fig. 83

To prove the second assertion of Theorem 3, we construct an example of a surface on which three quasigeodesics emanate from some given point in some direction. Consider a doubly-covered convex polygon P with infinitely many sides converging to a point O (Fig. 83). The boundary of the polygon P is a quasigeodesic emanating from the point O. At the same time, among all curves that connect the point O with some vertex A of the polygon P and twist from one side to the other, surrounding each vertex between A and O, there is a shortest arc as can be easily seen. It is easy to verify that this curve is a geodesic on its every segment that does not contain the point O. Hence, this curve is a quasigeodesic. Moreover, it obviously goes from the point O in the same direction as the boundary of the polygon P. There are two such lines, symmetric with respect to the boundary of the polygon P. Thus, three quasigeodesics emanate from the point O in the same direction.

Note one more important property of quasigeodesics which was proved by Pogorelov.

Theorem 4. *Let L be a quasigeodesic on a closed convex surface F. Let C be a cylinder with base line L and generator entering inside the body bounded by the surface F. When this cylinder unfolds onto the plane, the line L passes to a convex curve whose convexity is directed to the side corresponding to the exterior part of the cylinder C.*[19]

In Sec. 5 of Chapter IV, we have proved the Liberman theorem which asserts that every shortest arc has this property. Thus, all corollaries for shortest arc

[19] Pogorelov gave examples of curves that have the property indicated in Theorem 4 but not in the class of quasigeodesics.

© 2006 by Taylor & Francis Group, LLC

deduced from this property also hold for quasigeodesics. There is no necessity to repeat their deduction.

Let us also especially consider the question of the existence of closed quasigeodesics on a closed convex surface. It is known that at least three closed geodesics without multiple points exist on each regular closed convex surface. This famous theorem was conjectured by Poincaré in 1905, but its complete proof was given about 25 years later by Lyusternik and Shnirelman.[20] However, it is not difficult to see that there can be no closed geodesic without multiple points on a nonregular closed convex surface.

Indeed, such a geodesic must divide the surface into two domains homeomorphic to a disk, and, moreover, the boundaries of these domains have swerve zero as geodesics. Therefore, according to Theorem 2 in Sec. 2, the curvature of each of these domains equals to 2π. Hence for the existence of at least one closed geodesic on a closed convex surface, it is necessary anyway that this surface can be divided into two domains that have curvatures equal to 2π. But this turns out impossible sometimes. For example, we can construct a convex polyhedron with an arbitrary number of vertices whose curvatures are chosen so that the sum of curvatures of no part of vertices is equal to 2π. For a tetrahedron and all the more so for a doubly-covered triangle this can be done straightforwardly. Since the curvature of a polyhedron is supported only at its vertices, even on a convex polyhedron there can thus be no closed geodesics without multiple points.[21]

However, a generalization of the theorem on three closed geodesics proves possible for any closed convex surface if we consider quasigeodesics instead of geodesics.

To reveal the expected result, we take some closed convex surface F and construct a sequence of regular closed convex surfaces F_n that converge to F. On the surfaces F_n, there exist closed geodesics without multiple points; moreover, the well-known properties of these geodesics imply easily that there are those geodesics among them whose lengths are uniformly bounded. Therefore, we can choose some convergent sequence $L_{n_1}, L_{n_2}, \ldots$ from them. Each of these geodesics L_{n_i} divides the corresponding surface F_{n_i} into two domains $F^1_{n_i}$ and $F^2_{n_i}$ isomorphic to some caps as follows from Theorem 3 of Sec. 4. The limit of caps is either a cap or a

[20] L. A. Lyusternik and L. G. Shnirelman, Topological methods in variational problems, Uspekhi Mat. Nauk, Vol. II, No. 1 (17), 1947. Some remarks on closed geodesics can be found in W. Blaschke, Differential geometry, Sec. 97.

[21] Moreover, the polyhedra admitting closed geodesics without multiple points are rare exceptions. Indeed, let $\omega_1, \ldots, \omega_n$ be the curvature of the vertices of a polyhedron with n vertices. The numbers ω_i satisfy the condition

$$\omega_1 + \ldots + \omega_n = 4\pi; \qquad (*)$$

it is possible to prove that in any case, the numbers ω_i can vary within some bounds in an arbitrary way but so that condition (*) holds. Hence the set of the tuples of the possible values of the curvatures of vertices is of dimension $(n-1)$. Meanwhile, the condition that the sum of certain ω_i is equal to 2π is an equation that distinguishes an $(n-2)$-dimensional set in the set of all tuples $(\omega_1, \ldots, \omega_n)$. Starting from this fact, we can prove that the set of polyhedra with n vertices such that the sum of the curvatures of some of its vertices is equal to 2π has dimension less by 1 than the dimension of the set of all polyhedra with n vertices. (As we know, the latter is equal to $3n-6$.) About the determination of a polyhedron from curvatures of its vertices, see A. D. Aleksandrov, Application of a domain invariance theorem on to the proofs of existence, Izv. Akad. Nauk SSSR, Ser. Mat., 1939, No. 3, p. 249. Along with condition (*), all ω_i satisfy the obvious condition $0 < \omega_i < 2\pi$. It is interesting to find out whether these numbers satisfy some other conditions or not.

© 2006 by Taylor & Francis Group, LLC

line segment or a point. However, it is easy to show that the domains $F^1_{n_i}$ and $F^2_{n_i}$ cannot yield a point in the limit, and, therefore, only the following two possibilities are open: either the limits of the domains $F^1_{n_i}$ and $F^2_{n_i}$ are isometric to caps or one of these limits degenerates into a line while the second is isometric to a cap whose boundary is glued with itself.

In the first case, the limit of the lines L_{n_i} is a closed quasigeodesic on the surface F without multiple points. In the second case, this limit is a closed quasigeodesic that overlaps itself in much the same way as a segment which is first traversed in one direction and then, in the opposite direction.

An edge of a polyhedron connects two vertices whose complete angles $\leq \pi$ yield an example of such self-overlapping closed quasigeodesic. The complete angle at each vertex of a right tetrahedron is equal to π. Therefore, each of its edges turns out to be a closed quasigeodesic if we consider each edge as doubly covered.

We have proved that on every closed surface *there is at least one closed quasigeodesic that either has no multiple points or is self-overlapping.*

For proving the existence of three quasigeodesics of this sort, our simple argument is not sufficient any longer, since three closed geodesics on regular surfaces that converge to a given surface can *a priori* converge to the same curve. However, by a reasonable choice of regular surfaces that converge to a given surface, we can avoid this as was shown by Pogorelov.[22] He has thus proved that *there are three closed quasigeodesics on every convex surface which have no multiple points or overlap each other.*

All what was said above about quasigeodesics makes it clear that a detailed study of quasigeodesics is possible and promises to yield interesting and important results. We must mention the most important question concerning quasigeodesics.

It is known that geodesics on regular surfaces can be defined as the curves that yield the extremum of length, e.g., among all curves connecting given points. In other words, the variation of length is equal to zero at them. The concept of variation of length can lose sense for very simple curves on nonregular surfaces, and, therefore, the usual methods of the calculus of variations become inapplicable. For example, even the length of a broken line on a polyhedron is not a differentiable function of the coordinates of its vertices if at least one of its vertices coincides with the vertex of the polyhedron (we speak about the coordinates on the polyhedron but not in space.) This makes it clear that the definition of an extremal itself as a curve for which the variation of length is equal to zero turns out inappropriate. It is necessary to use other definitions of an extremum and an extremal. Such a definition was already available in the calculus of variations in connection with topological methods; roughly speaking, this definition reduces to the fact that a point in the space R on which a function $f(X)$ is defined is called extremal or critical if the family of surfaces $f(X) = \text{const}$ has singularities at some neighboorhood about this point. In this case, the role of points is played by rectifiable curves, while the space is the set of all rectifiable curves under consideration on which the concept of

[22] This result was communicated to me by A. V. Pogorelov in a letter in 1945.

© 2006 by Taylor & Francis Group, LLC

distance is introduced in an appropriate way. We formulate our question, referring the reader to special literature for rigorous definitions.[23]

Is it possible to prove that a curve yields an extremum of length among all curves connecting two given points on a convex surface if and only if this curve is a quasigeodesic? A similar question arises in the case of other limit conditions, e.g., when considering closed curves, etc.

It is known that the trajectory of a mass point moving along a surface in the absence of external forces is a geodesic provided a surface is regular. This follows by the way from the Hamilton principle which reduces to the condition that each arc of the trajectory must yield an extremum length in this case. It seems likely that all quasigeodesics themselves are feasible trajectories of a point moving in the absence of external forces. Since nonregular convex surfaces can be constructed with any available accuracy from an arbitrarily resistant material, the question of the motion of a mass point along such a surface is meaningful in reality. If we slightly smooth the surface so that it becomes regular, then the trajectories of a point will be geodesics. Taking this smoothing infinitesimal and falling beyond the limits of precision with which the surface is manufactured, we can arrive at the conclusion that motion happens along the limits of geodesics, i.e., along quasigeodesics.

If the complete angle θ at a point B on a convex surface is $< 2\pi$, then a quasigeodesic that approaches the point O in the direction d can be extended beyond this point remaining a quasigeodesic in each direction that makes the angle less than π with d, i.e., in each direction that lies in the angle, equal to $\theta - \pi$, and, in general, in every direction if $\theta \leq \pi$. This involves a curious uncertainty of the motion of a mass point along a quasigeodesic, which is not surprising since the velocity and reaction forces at such point O are completely uncertain. These remarks on the mechanical essence of quasigeodesics do not pretend to be rigorous; also their goal is to show that the study of quasigeodesics might be of interest for mechanics.

6. A Circle

A circle on a convex surface, i.e., the locus of points equidistant from a given point, can be not a curve in the sense of its possible continuous parametric representation. Recall the example of the circle on a cone with complete angle at the vertex $< \pi$ which was considered in Sec. 10 of Chapter I (Fig. 11). If we take a point A on this cone different from its apex, then a sufficiently small circle centered at the point A is certainly identified with a circle on the plane. But with the increase of the radius, this circle will touch itself at some instant of time and then fall into two curves one of which contracts to the apex with the further increase of the radius and finally disappears. Whenever we take the point A sufficiently close to the apex of the cone, we obtain arbitrarily small circles, which fall into two curves or have an isolated point.

We can exhibit some examples of circles that resemble curves in the usual sense to a lesser degree. For example, take a plane circular sector of radius 1 centered at O with angle $< \pi/2$ and distinguish an arbitrary closed set M containing two

[23] H. Seifert und W. Threlfall, Variationsrechnung in Grossen; M. Morse. The calculus of variations in the large; L. A. Lyusternik, Topology and the calculus of variations, Uspekhi Mat. Nauk, Vol. I, No. 1 (11), 1946.

© 2006 by Taylor & Francis Group, LLC

points A and B on the arc AB bounding this sector (Fig. 84). The convex hull of this set and the center O is some convex domain bounded by the radii OA and OB to the sides. Gluing a part of the plane which is bounded by the prolongation of the radius OA and the tangent to the plane at the point B to this domain along the radius OB, we obtain the convex domain G, that is marked with dashes in the figure. Assuming that this domain is doubly-covered, we transform it into a complete convex surface F. It is easy to see that the circle of radius 1 centered at the point O on this surface consists of two arcs BC and the chosen set M. We can bend this surface F in such a way that it becomes nondegenerate (this is possible by the Olovyanishnikov theorem cited in Sec. 4 of Chapter VIII). This example shows that *a circle on a convex surface can be homeomorphic to an arbitrary closed set on a plane circle that contains at least one arc* (corresponding to the arc CB).

The following theorem expresses an important general property of a circle and is completely analogous to the well-known theorem of plane geometry which asserts that the circle is orthogonal to its radius.

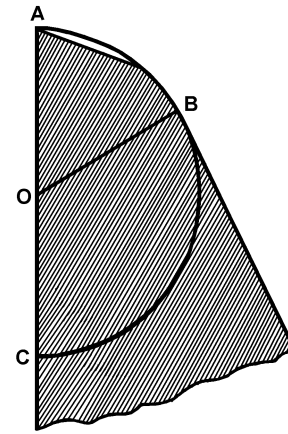


Fig. 84

Theorem 1. *Let A be a point on a circle C centered at a point O on a closed surface. Let a sequence of points X_n of the circle C converge to the point A in such a way that the shortest arcs OX_n, the radii of this circle, converge to some shortest arc OA. Then the angles between the shortest arcs AX_n and OA tend to the right angle.*

Proof. Consider the triangles OAX_n and plane triangles T_n with sides of the same lengths. Let α_n be the angle made by OA and AX_n, α_n^0 be the corresponding angle of the triangle T_n, and let ω_n be the curvature of the interior of the triangle OAX_n. Then

$$|\alpha_n - \alpha_n^0| \le \omega_n.$$

Since $OA = OX_n$, the triangle T_n is isosceles, and hence its angle α_n^0 tends to $\pi/2$ as $X_n \to A$. At the same time, the shortest arcs OX_n converge to OA, and, therefore, the interiors of the triangles OAX_n can be included into a vanishing sequence of domains.[24] By the continuity property of curvature, this implies that ω_n tends to zero. Consequently, the limit of the angle α_n coincides with the limit of α_n^0, i.e., is equal to $\pi/2$.

If two continuous branches C_1 and C_2 of the circle C emanate from the point A, then, taking points X_n on one of these branches, we obtain that the radius OA, which is the limit of the radii OX_n, is orthogonal to this branch. Consequently, both branches C_1 and C_2 have directions at the point O which are orthogonal to the corresponding radii $(OA)_1$ and $(OA)_2$. Of course, these radii may fail to coincide, and then the point A is a corner point of the circle.[25] If there is only one radius,

[24] Namely, into the sequence of neighborhoods of the radius OA in each of which OA is deleted.
[25] Two radii go to the point B in Fig. 84. In Fig. 11, two radii also go to each of the points C and D.

© 2006 by Taylor & Francis Group, LLC

i.e., only one shortest arc OA then both branches are orthogonal to this radius. In this case, we can merely say that the circle is orthogonal to the radius.

The following conclusion can be drawn from Theorem 1.

A point such that the complete angle at it is $< \pi$ can be only an isolated point of any circle.

Indeed, let the complete angle θ at the point A be less than π. Assume that this point is not an isolated point of a circle C centered at a point O, i.e., assume that there are points on C that are arbitrarily close to A. We choose a sequence X_n of these points such that the shortest arcs OX_n converge to some shortest arc OA. Then by Theorem 1, the angles α_n between OA and AX converge to $\pi/2$. But the angle made by shortest arcs emanating from the point A cannot be greater than half of the complete angle θ at the point A. Hence $\alpha_n \leq \theta/2$ and $\lim_{n\to\infty} \alpha_n \leq \theta/2 < \pi/2$. We obtain a contradiction, which shows that there are no points arbitrarily close to the point A on the circle C.

After the remarks on the possible singularities of a circle made in the beginning of this section, the following theorem will not seem to be trivial.

Theorem 2. *For each point O on a convex surface, there exists a positive r_0 such that each circle of radius less than r_0 centered at O turns out a simple closed curve, i.e., homeomorphic to a circle on the plane.*

Proof. Take a small disk U_r of radius r around the point O; its boundary is a circle. Show that for a sufficiently small r, the interior of the disk U_r is a simply connected domain. The fact that the interior of the disk is a connected domain is obvious. If the disk is sufficiently small, it lies in a neighborhood of the point O, which is homeomorphic to a plane disk, and hence if its interior is not simply connected, then it has "holes." Since the disk U'_r lies in the disk U_r for $r' < r$, its exterior boundary shrinks and the "holes" enlarge when the radius of the disk U_r decreases; also, some new "holes" may appear.[26] Eventually, one of the holes V in the process of its enlargement touches the exterior boundary of the disk U_r at least at one point A. Two branches C_1 and C_2 emanate from the point A which are the boundaries of the hole V under consideration. We can take points converging to A on these branches; and then the radii dropped to these points yield in the limit two radii $(OA)_1$ and $(OA)_2$ that approach the point A from the different sides of the hole V (Fig. 85). Hence these radii bound some digon D that includes the hole V. Since each radius is orthogonal to its own branch, C_1 or C_2, the angle of the digon D at the vertex A is no less than π. Consequently, the sum of both angles of the digon D, i.e., the curvature of its interior is greater than π.

At the same time, the interior of the digon D lies in the domain $W - O$ obtained from the neighborhood W of the point O by deleting the point O. By the continuity property of curvature, the curvature of the domain $W - O$ becomes arbitrarily small as the neighborhood W decreases. Therefore, there are no digons of curvature greater than π in a sufficiently small domain W, and hence a sufficiently small disk U_r cannot have holes. Thus, the interior of a sufficiently small disk centered at a

[26] In Fig. 11, the arc CC is the exterior boundary of the disk, and the arc DO is a "hole" in it. When the radius decreases, the boundary of the "hole" touches the exterior boundary at the point B.

© 2006 by Taylor & Francis Group, LLC

given point is simply connected; this easily implies that its boundary, i.e. the circle is a closed simple curve.

At the same time it might be worth noting that even a circle of an arbitrarily small radius can have corner points. Indeed, we have seen that a circle has no corner points only under the condition that only one radius goes to its every point. As is easily seen, this means that shortest arcs emanate from its center in all directions; they are the radii of this circle. Meanwhile, we know that shortest arcs do not emanate from each point in all directions in general. Thus, even an arbitrarily small circle with a given center can have singularities that essentially differ it from a circle on the plane.

We now pass to the consideration of the dependence of the length of a circle on its radius. To this end, we use the following simple remark.

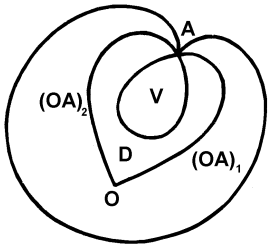


Fig. 85

Lemma. *Let X and Y be points equidistant from O on two shortest arcs emanating from a point O, that is $OX = OY = r$; and let $z(r)$ be the distance between them. Then the ratio $z(r)/r$ is a nonincreasing function of r, and if φ is the angle between these shortest arcs and $\omega(r)$ is the curvature of the triangle OXY, then*

$$2\sin\frac{\varphi}{2} \geq \frac{z(r)}{r} \geq 2\sin\frac{\varphi-\omega(r)}{2}. \qquad (1)$$

Indeed, if $\varphi_0(r)$ is the corresponding angle in the plane triangle with the same sides as the triangle OXY, then

$$\frac{z(r)}{r} = 2\sin\frac{\varphi_0(r)}{2}.$$

But by the convexity condition, $\varphi_0(r)$ does not increase as r increases, and hence the same is true for $z(r)/r$. Moreover, we know that the angles $\varphi_0(r)$ and φ satisfy the relation $|\phi_0(r) - \phi| \leq \omega(r)$. This implies (1).

The following theorem on the length of a circle turns out to be an almost obvious corollary of the previous lemma.

Theorem 3. *The ratio of the length of a circle $l(r)$ to its radius r does not increase with the increase of the radius and*

$$\theta \geq \frac{l(r)}{r} \geq \theta - \omega(r),$$

where θ is the complete angle at the center and $\omega(r)$ is the curvature of the disk with the deleted center which is bounded by this circle. Instead of the whole circle, we can take its every arc subtending the sector with a given angle φ; then, in case $l(r)$ is the length of this arc and $\omega(r)$ is the curvature of this sector with the deleted center, we have $\phi \geq l(r)/r \geq \phi - \omega(r)$; moreover, $l(r)/r$ again does not increase with the increase of r.[27]

[27] If we cut out this sector and glue the radii that bound it together with each other, then we obtain a disk whose complete angle at the center is equal to φ. Thus, the operation of gluing reduces the case of an arc to the case of the whole circle.

© 2006 by Taylor & Francis Group, LLC

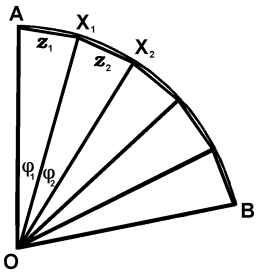


Fig. 86

We assume that an arc of a circle is a curve in the usual sense.[28] We say that an arc subtends a sector if it is possible to draw a radius to its every point which travels in the sector U.

Proof. Assume that the arc AB of a circle of radius r subtends the sector with angle ϕ. Take points $X_0, X_1, \ldots, X_n$ on this arc which are located sequentially between the ends A and B of this arc; and, moreover, $X_0 = A, X_n = B$. Drawing radii to all points $X_0, X_1, \ldots, X_n$ from the center O, we obtain the isosceles triangles $OX_{i-1}X_i$. Let φ be the angle between the radii OX_{i-1} and OX_i, and let z_i be the distance $X_{i-1}X_i$ (Fig. 86). According to our lemma, for a given ϕ_i, the ratio z_i/r does not increase with the increase of the radius r. At the same time, the length of the arc AB is by definition the least upper bound of the sums $\sum_{i=1}^{n} z_i$. We conclude from this easily that the length of the arc also does not increase with the increase of the radius.[29] This proves the first assertion of the theorem.

Further, according to formula (1),

$$z_i \geq 2r \sin \frac{\varphi_i}{2};$$

this implies

$$\sum_{i=1}^{n} z_i \leq 2r \sum_{i=1}^{n} \sin \frac{\varphi_i}{2} \leq r \sum_{i=1}^{n} \varphi_i = r\varphi.$$

[28] In the general case, an arc of the circle can be not a curve in the usual sense. However, its length can always be defined in the following way. Draw the radii OA and OB that bound the sector U. A point X of the arc AB is considered as a preceding point of a point Y if the sector OAX lies in the sector OAY. This defines the order of points on the arc AB, and the length of the arc AB can be defined as the limit of the sums of distances between its sequential points under the condition that we take only the distances between points that are distant from each other by no more than some $\varepsilon \to 0$.

[29] Take two arcs AB and $A'B'$ of radii r and r' with $r' < r$. The radius OX_i that goes to the point X_i of the arc AB intersects the arc $A'B'$ at the point X_i'. If $X_{i-1}'X_i + Z_i'$ and $l(r')$ is the length of the arc $A'B'$, then

$$\frac{1}{r'} l(r') \geq \lim \frac{1}{r'} \sum_i z_i'. \tag{a}$$

Here the strict inequality is possible, since the radii going to points on the arc AB can form digons that contain part of the arc $A'B'$. Further,

$$\frac{z_i'}{r'} \geq \frac{z_i}{r}; \tag{b}$$

moreover, the points X_i can in advance be chosen so that

$$\sum_i z_i > l(r) - \varepsilon. \tag{c}$$

Since ε is arbitrary, it follows from (a), (b), and (c) that

$$\frac{l(r')}{r'} \geq \frac{l(r)}{r}.$$

© 2006 by Taylor & Francis Group, LLC

Therefore, the length $l(r)$ of the arc AB is also no greater than $r\varphi$; this proves the first part of the inequality of the theorem.

To prove the second part of this inequality, we consider all those points of the arc AB each of which may be joined with the center by more than one radius. Two radii that go to one point bound a digon, and among all radii going to a given point, there are two radii bounding the maximal digon. Thus, with each point Y admitting more than one radius, we associate some angle ψ between the radii OY and $\overline{OY}$, bounding the maximal digon. Since the sum of all these angles ψ cannot exceed the angle of the whole sector subtended by the arc AB, the number of such digons with angles ψ greater than a given angle is finite and so their total number is at most countable.

Let $Y_1, \ldots, L_m$ be all endpoints to which the radii OY_j and $\overline{OY_j}$ forming the angles greater than some $\varepsilon > 0$ go. We number the points $Y_1, \ldots, L_m$ in the order of their location from one endpoint A to another endpoint B of the arc AB; the radius OY_j is assumed to be closer to the radius OA than the radius $\overline{PY_j}$ (Fig. 87). The sector between the neighboring radii $\overline{OY_j}$ and OT_{j+1} can be divided into sectors with the angles that are less than 2ε. Indeed, if this is impossible then there exist radii OX and OY with arbitrarily close endpoints X and Y on the arc Y_jY_{j+1} which make the angle no less than 2ε with each other. In the limit, when the points X and Y coincide, these radii will bound a digon with angle not less than 2ε. But this is impossible, since the points Y_j have already exhausted all vertices of digons of such a type.

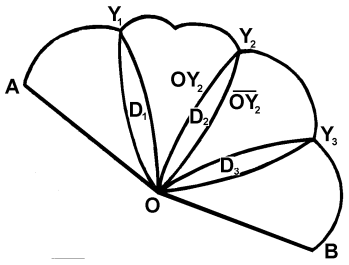


Fig. 87

Thus, we take those points X_i on each of the arcs AY_1, $Y_1Y_2, \ldots, Y_mB$ for which the angles φ_i between the radii OX_{i-1} and OX_i are less than 2ε. We add the points A, Y_1, $\ldots$, Y_m, B to the set of the points X_i so that if, e.g., the point X_i is Y_j, then the angle ϕ_i is the angle made by the radii OX_{i-1} and OY_j, while the angle ϕ_{i+1} is the angle between the radii OX_{i+1} and $\overline{OY_j}$.

We thus come to the isosceles triangles $OX_{i-1}X_i = T_i$ with angles $\phi_i \le 2\varepsilon$. The sum of all angles ϕ_i is less than the angle ϕ of the whole sector subtended by the arc AB, since the angles of the digons D_j with vertices Y_j are excluded, and in order to obtain the whole angle ϕ, we have to add the angles ψ_j of these digons. Thus,

$$\phi = \sum_i \phi_i + \sum_j \psi_j. \tag{2}$$

The digons D_j, together with the triangles T_i, form some polygon P; the curvature of the interior of this polygon is equal to the sum of the curvatures of their interiors; that is,

$$\omega(P) = \sum_i \omega(T_i) + \sum_j \omega(D_j).$$

© 2006 by Taylor & Francis Group, LLC

The curvature of a digon is equal to the sum of its angles, and so $\omega(D_i) > \psi_j$; consequently,

$$\omega(P) \geq \sum_i \omega(T_i) + \sum_j \psi_j. \tag{3}$$

Subtracting (3) from (2), we obtain

$$\sum_i [\phi_i - \omega(T_i)] \geq \phi - \omega(P). \tag{4}$$

If $X_{i-1}X_i = z_i$ as above, then just by the definition of length, for the length l of the arc AB, we have the inequality

$$l \geq \sum_i z_i.$$

According to (1),

$$z_i \geq 2r \sin \frac{\phi_i - \omega(T_i)}{2} \geq 2r \sin \frac{\phi_i}{2} - r\omega(T_i),$$

and hence

$$l \geq 2r \sum_i \sin \frac{\phi_i}{2} - r \sum_i \omega(T_i). \tag{5}$$

Since $2\sin(\phi_i/2) > \phi_i - (\phi_i^3/24)$ and $\phi_i \leq 2\varepsilon$ and $\sum_i \phi_i \leq \phi$, we have

$$2 \sum_i \sin \frac{\phi_i}{2} > \sum_i \phi_i - \delta,$$

where $\delta = (1/3)\varepsilon^2\phi$. Due to this fact, (5) implies

$$l > r \sum_i \phi_i - r \sum_i \omega(T_i) - r\delta. \tag{6}$$

Now, using inequality (4), we deduce from this that

$$l > (\phi - \omega(P))r - r\delta.$$

When we insert some newer and newer points X_i in the arc AB, the resulting polygons P converge to the sector U subtended by this arc. By the continuity of curvature, this implies that the limit of the curvatures of the interiors of the polygons P does not exceed the curvature ω of the sector U with the deleted center O; therefore, (7) implies

$$l > (\phi - \omega)r - r\delta.$$

Since δ, together with ε, is arbitrarily small, we have

$$l \geq (\phi - \omega)r. \tag{8}$$

But this is the second part of the inequality whose validity is claimed in the theorem.

© 2006 by Taylor & Francis Group, LLC

A finer analysis leads to the fact that, as ω in inequality (8) we can take the curvature of the interior of the sector U, i.e., with deletion of the arc AB itself rather than just the center. This makes a difference since the arc AB can have curvature (the area of the spherical image) other than zero. Each parallel of a surface of revolution which is an edge of this surface yields an example of this.

Further, inspection of the above proof leads to the fact that $l = \phi r$ if and only if the curvature of the interior of the sector U is zero, i.e., if and only if the sector U can be developed onto the plane. In exactly the same way, we can show that $l = (\phi - \omega)r$ only in the case where the curvature ω of the interior of the sector U is zero. Thus, whenever $\omega \neq 0$, we have $\phi r > l > (\phi - \omega)r$.

If only one radius goes to each point of an arc of a circle, then the length of this arc can be represented as an integral. Let AB be an arc of a circle having this property, let ϕ be the angle of the sector subtended by this arc, and let r be the radius of this arc. Let $\omega(x)$ be the curvature of the interior of the circular sector of radius $x \leq r$ between the radii OA and OB. The length l of the arc AB is expressed by the formula

$$l = \phi r - \int_0^r \omega(x)\, dx. \tag{9}$$

If only one radius goes to each point of the circle of radius r centered at O, then the length of this circle is expressed by the same formula

$$l = \theta r - \int_0^r \omega(x)\, dx, \tag{10}$$

where θ is the complete angle at the point O and $\omega(x)$ is the curvature of the concentric disk of radius r with the point O deleted.

For the derivative of the arc length l with respect to the radius r, we have the formula

$$\frac{dl}{dr} = \varphi - \omega(r), \tag{11}$$

where dl/dr is the left derivative if $\omega(r)$ is the curvature of the sector with the deleted arc, and dl/dr is the right derivative if $\omega(r)$ is the curvature of the sector with the arc ajoined. This may differ as is shown by the example of parallels of a surface of revolution which are edges of this surface. Since $\omega(r)$ does not decrease with the increase of r, dl/dr does not increase, and hence the arc length turns out to be a convex function of the radius. However, we speak of a particular case in which only one radius goes to each point of the arc. Without this assumption, formula (9), and, respectively, (10) and (11) prove false and must be replaced by more intricate expressions.

Namely, if more than one radius goes to some points of the arc AB, then we cut out from the surface all digons that are bounded by the radii going to the points of the arc AB and pairwise glue together the sides of the resulting holes. By the gluing theorem, we obtain a manifold with metric of positive curvature (which can certainly be realized as a convex surface). In this manifold, the only radius goes to each point of the arc AB, so we can apply the same formula (9) to this arc.

© 2006 by Taylor & Francis Group, LLC

Consequently, formula (9) also holds in the general case whenever by φ we mean the angle of the subtended sector without angles of all digons and $\omega(x)$ stands for the curvature of the interior of the sector of radius x without the curvature of those parts of this interior that lie in our digons. Since the cut-off digons change with the radius r, formula (11) for dl/dr does not hold, and the arc length can be not a convex function of the radius as can be seen by examining a circle on a cone whose center does not lie at the vertex of this cone.[30]

To conclude, we mention one more theorem: *If a sequence of convex surface F_n converges to a surface F and a sequence of circles C_n on the surfaces F_n converges to a circle C on F, then the lengths of the circles C_n converge to the length of the circle C.*

If the circles C_n and C are closed curves, then this theorem insues from the following two theorems: 1. The swerve of a circle has variation less than 6π (the variation of the swerve of a circle is easily estimated via the curvature of the disk and the complete angle at its center). 2. If the variations of the swerve of curves L_n are uniformly bounded and the curves L_n converge to the curve L, then the length of L is equal to the limit of the lengths of the curves L_n provided necessarily that the curves L_n lie on convex surfaces converging to the surface over which the curve L travels. In the case where the circles C and C_n are not curves in the usual sense, we are aware of no proof nor disproof of this theorem.

[30] Let F be a doubly covered polygon with infinitely many vertices accumulating symmetrically at the point O. On such a surface F, the length of a circle centered at O of radius r is not a convex function of r even for arbitrarily small r.

© 2006 by Taylor & Francis Group, LLC

Chapter X

AREA

1. The Intrinsic Definition of Area

Let P be a polygon on a convex surface. This polygon can be divided into arbitrarily small triangles, and, therefore, there exists a sequence Z_n of its partitions such that the maximal diameter of triangles in the nth partition tends to zero as $n \to \infty$. To each triangle of the nth partition, we put in correspondence the plane triangle with sides of the same length and take the sum S_{Z_n} of the areas of these plane triangles. It turns out that these sums S_{Z_n} converge to some limit S as $n \to \infty$ independently of the choice of the partitions Z_n whenever these partitions are infinitely refined. This limit we define to be the area of the polygon P.

Justification of this definition of area should consist in the proof of the following three facts: (1) The limit of the sums S_{Z_n} always exists, i.e., each polygon on a convex surface has area. (2) The area in the sense of this definition of every polygon on a convex surface coincides with its usual area defined by polyhedral approximation. (3) Area is completely additive, i.e., if a polyhedron P is partitioned into a finitely or countably many polygons, then its area is the sum of the areas of these polygons. After that, using the familiar technique of measure theory, we can define the area of each closed or open set and then that of each Borel set on a convex surface. The first of the three facts will be proved in this section, while the other two will be established in the next section. Here, we obtain even more than the existence of the area of each polygon; namely, we prove that if a polygon P on a convex surface is partitioned into triangles of diameter $\leq d$, then the difference of the sum S_Z of the areas of the corresponding plane triangles and the sum S of the polygon P itself satisfies the inequality

$$0 \leq S_Z \leq \frac{1}{2}\omega(P)d^2, \tag{1}$$

where $\omega(P)$ is the curvature of the interior of the polygon P. This inequality estimates quite sharply how close the sum S_Z is to the limit value S depending on the maximal diameter d of the triangles of the partition.[1]

Without loss of generality, we can assume that a given polygon P belongs to some closed convex surface F. Each partition Z_n of the polygon P can be completed to $\overline{Z}_n$ of the whole surface F. Replacing each triangle of $\overline{Z}_n$ by a plane triangle, we obtain the development R_n. The more refined partition $\overline{Z}_n$, the more close is the metric defined by this development to the metric of the surface F. The

[1] Inequality (1) can be improved only in the sense that the factor $1/2$ can be replaced by a smaller one.

© 2006 by Taylor & Francis Group, LLC

surface F has metric ρ, and by the realization theorem of Chapter VII, we can choose a sequence from polyhedra obtained by gluing from the developments R_n which converges to a surface having this metric ρ, i.e., to a surface isometric to F. Therefore, our intrinsic definition of area is similar to its usual definition as the limit of the areas of polyhedra converging to a given surface. We only carry out polyhedral approximation in intrinsic rather than in extrinsic manner; that is, the polyhedral metrics defined by the developments R_n converge to the metric of the surface F.

In order to derive estimate (1) and, at the same time, to establish the existence of area, we first prove the following two lemmas.

Lemma 1. *Assume that T is a triangle on a convex polyhedron or, in general, on a manifold with polyhedral metric of positive curvature, and T_0 is the plane triangle with sides of the same length, d is the diameter of the triangle T, and ω is the curvature of its interior. Let S and S_0 be the areas of the triangles T and T_0, respectively. Then the following inequality holds*:

$$0 \le S - S_0 \le \frac{1}{2}\omega d^2. \tag{2}$$

Proof. If the curvature of the triangle T is equal to zero, then this triangle is isometric to the triangle T_0, and inequality (2) is trivial. Therefore, we can assume that $\omega > 0$.

Let A, B, and C be the vertices of the triangle T. Assume, e.g., that the vertices A and B can be connected by a shortest arc $\overline{AB}$ in T different from the side AB. Then this shortest arc divides the triangle T into a digon D and a triangle T_1 with sides of the same length as T. The curvature of the interior of the triangle T consists of the curvatures of the interiors of the triangle T_1 and the digon D, i.e.,

$$\omega(T) = \omega(T_1) + \omega(D). \tag{3}$$

The diameters of the digon D and the triangle T_1 are obviously not greater than the diameter of the triangle T.

If the angle at the vertex B of the digon D is greater than π, then there are points X in D such that the shortest arc AX passes through B. Then the shortest arc BX does not pass through A for certain. Thus, we can draw a shortest arc of each point X of the digon and each of its vertices which does not pass through other vertices.

Draw shortest arcs AX from the vertex A of the digon D to all those points X of this digon for which the shortest arc AX does not pass through the vertex B. Take a point A_0 on the plane and draw the segments A_0X_0 from this point which are equal to the shortest arcs AX and make the same angles with each other. As a result, we obtain a figure D_A on the plane whose area S_A is equal to the area of the part of the digon D which is formed by all points X under consideration. (This construction can be also visualized as follows: we draw all geodesics from the vertex A in D by the length at which they all remain shortest arcs and then draw the segments on the plane which are equal to these arcs and make the same angles with each other. Two shortest arcs will arrive at some points X; these points will be the endpoints of the geodesics in question.)

© 2006 by Taylor & Francis Group, LLC

If α is the angle at the vertex A of the digon D, then the figure D_A lies in the sector with angle α and radius equal to the diameter of D. But the diameter of D is no greater than the diameter d of the triangle T, and so

$$S(D_A) < \frac{1}{2}\alpha d^2.$$

If, starting from the vertex B, we construct the corresponding figure D_b, then we have

$$S(D_b) < \frac{1}{2}\beta d^2$$

for its area $S(D_B)$.

Since each point of D has its image at least in one of the figures D_A and D_B, then we obtain

$$S(D) \le S(D_A) + S(D_B) < \frac{1}{2}(\alpha + \beta)d^2$$

for the area of the digon. But the sum of the angles of a digon is the curvature of its interior, and hence

$$S(D) < \frac{1}{2}\omega(D)d^2. \tag{4}$$

After deleting D from T, there remains the triangle T_1 for which

$$\begin{aligned} S(T_1) &= S(T) - S(D) > S(T) - \frac{1}{2}\omega(D)d^2, \\ \omega(T_1) &= \omega(T) - \omega(D). \end{aligned} \tag{5}$$

If it is still possible to draw shortest arcs in T_1 which connect the vertices of T_1 and thus cut out new digons from T_1, then we also delete these digons. As a result, there remains a triangle T_2 without these shortest arcs. If ω' is the curvature of all deleted digons, inclusive the first digon D then, according to formulas (5), we have

$$\begin{aligned} S(T) > S(T_2) &> S(T) - \frac{1}{2}\omega' d^2, \\ \omega(T_2) &= \omega(T) - \omega'. \end{aligned} \tag{6}$$

If the curvature of T_2 turns out to be zero, then this triangle is isometric to the plane triangle T_0 with sides of the same length as the above triangle T. In this case, formulas (6) yield

$$S(T) > S(T_0) > S(T) - \frac{1}{2}\omega(T)d^2;$$

which is the required inequality (1).

It remains to assume that the curvature of T_2 is away from zero. Of course, it could happen that there are no shortest arcs between the vertices of the initial triangles at all except for the sides. Then the triangle T itself would play the role of the triangle T_2.

If the curvature of T_2 is other than zero then there are "interior vertices" in this triangle, i.e., the points such that the complete angle at each of them is less than 2π. Take such a point O and connect this point with the vertex A of the triangle T by

© 2006 by Taylor & Francis Group, LLC

a shortest arc AO. If this shortest arc passes through another vertex, then we take the latter for A, and so we can always assume that the shortest arc AO goes inside the triangle. If there are two shortest arcs AO, then these arcs bound a digon whose interior must contain other interior vertices, and instead of O, we can take one of these vertices. Therefore, we can always assume that AO is a unique shortest arc that connects the "interior" vertex O with the "exterior" vertex A.

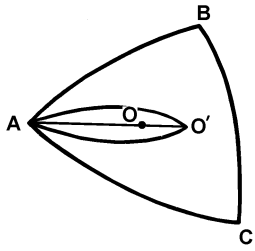


Fig. 88

Draw a shortest arc OO' from O that makes equal angles to both sides with AO (these angles are equal to half of the complete angle at O) (see Fig. 88, which is of "topological character"). If the point O' is sufficiently close to O then there are exactly two shortest arcs AO' traveling symmetrically to both sides of AO. These shortest arcs bound a certain digon D. Let $\Delta\alpha$ be the angle at the vertex A of this digon. The digon D is composed of two equal triangles AOO', and so the area of D is

$$S(D) = AO \cdot AO' \cdot \sin\frac{1}{2}\Delta\alpha d^2.$$

If we cut out the digon D and identify the shortest arcs AO' then the triangle T_2 is replaced by a new one in which the vertex O' appears instead of the vertex O. The angle at the vertex A diminishes by $\Delta\alpha$, and hence the curvature decreases by

$$\Delta\omega = \Delta\alpha.$$

In turn, the area decreases by $\Delta S = S(D) < (1/2)\Delta\alpha d^2$, and so the deficit of area is

$$\Delta S < \frac{1}{2}\Delta\omega d^2. \tag{7}$$

When cutting out the digon D and gluing together the shortest arcs AO', the lines intersecting these shortest arcs are replaced by shorter lines in general. Therefore, it might happen that there appears, for example, a line $\overline{BC}$ in the triangle T_2 which connects two vertices B and C and is shorter than the corresponding side. However, this is impossible if the point O' is sufficiently close to the point O; otherwise, as $O' \to O$, the limit of such lines $\overline{BC}$ yields a line in triangle whose length is not greater than the side BC; which is impossible simply by the definition of T_2. In any case, we must ensure that the lines $\overline{BC}$, shorter than the side BC, do not appear. If there appears a line $\overline{BC}$ equal to the side BC, then so does a digon, which we cut out in the same way as in the beginning of the proof. In this case, the subtracted area is estimated by formula (6), i.e., the deficit of the area and that of the curvature are related by the same inequality (7).

The operation of cutting digons can be repeated with various interior vertices O; moreover, the number of the interior vertices will not increase, the diameter of the triangle will not increase either, and the deficit of area ΔS is constantly related to the deficit of curvature $\Delta\omega$ by inequality (7). If there appears a shortest arc that connects two vertices and is different from the side, then we cut out the resulting digon; moreover, the number of interior vertices decreases, while ΔS and $\Delta\omega$ again are connected by inequality (7). In this way, we arrive at the case in which there

© 2006 by Taylor & Francis Group, LLC

are no interior vertices, i.e., we obtain a triangle whose curvature is equal to zero; this is the triangle T_0.

The total deficit of area is $S(T) - S(T_0)$, the deficit of curvature is $\omega(T)$; and, since inequality (7) holds as usual, on summing up we obtain

$$S(T) - S(T_0) < \frac{1}{2}\omega(T)d^2.$$

One exception is the case in which the triangle T_0 is isometric to the triangle T: then $S(T) = S(T_0)$. Hence, in the general case,

$$0 \leq S(T) - S(T_0) \leq \frac{1}{2}\omega(T)d^2,$$

as required.

Of course, this inference is not perfectly rigorous; but it can be carried out impeccably by using a routine trick. Indeed, we have proved that by cutting out a digon it is possible to pass from the initial triangle T to a triangle T' such that

$$0 \leq S(T) - S(T') \leq \frac{1}{2}[\omega(T) - \omega(T')]d^2. \tag{8}$$

This is inequality (7). Consider all triangles T' with sides of the same length, having diameter at most d, and the number of interior vertices at most that of T such that inequality (8) holds for T'. Let T_0 be a triangle whose curvature has a minimum value. This triangle exists since, by the boundedness of the number of vertices and the diameters of the triangles T', we can choose a convergent sequence from each sequence of these triangles.[2] Then

$$0 \leq S(T) - S(T^0) \leq \frac{1}{2}[\omega(T) - \omega(T^0)]d^2. \tag{9}$$

But if $\omega(T^0) \neq 0$ then, by the above, we can cut out a digon from the triangle T^0; moreover, ΔS and $\Delta\omega$ are connected by inequality (8) and the resulting triangle T^1 satisfies all the requirements and also inequality (8), since

$$\Delta S = S(T^0) - S(T^1) < \frac{1}{2}\Delta\omega d^2 = \frac{1}{2}[\omega(T^0) - \omega(T^1)]d^2.$$

Hence, if $\omega(T^0) \neq 0$, then the curvature of the triangle T^0 could decrease in contradiction to the definition of this triangle. Therefore, $\omega(T^0) = 0$, i.e., $T^0 = T_0$, and (9) is nothing else but the required inequality

$$0 \leq S(T) - S(T^0) \leq \omega(T)d^2.$$

The lemma is proved.

[2]Since the number of interior vertices is bounded above, the triangle T' can be partitioned into triangles isometric to plane triangles whose number is also bounded above. At the same time, the sides of these triangles do not exceed d. Consequently, each triangle T' is given by finitely many parameters each ranging over a bounded interval, and so, the set of these triangles is, indeed, compact.

© 2006 by Taylor & Francis Group, LLC

Lemma 2. *Assume that T is a triangle on a convex surface, d is the diameter of T, ω is the curvature of the interior of T, and S_0 is the area of the plane triangle T^0 with sides of the same length. Then for each $\varepsilon > 0$, there exists a partition Z of the triangle T into "smaller" triangles T_i such that the sum S_Z of the areas of the plane triangles T_i^0 with sides of the same length satisfies the inequality*

$$-\varepsilon < S_Z - S_0 < \frac{1}{2}\omega(T)d^2 + \varepsilon. \tag{10}$$

Proof. Consider some partition Z of the triangle T into small triangles T_i. Replacing each triangle T_i by the plane triangle T_i^0, we obtain some development. It is convenient to consider this development to be glued in such a way that we obtain some "polyhedral" polygon Q that replaces the triangle T. If we consider the triangle T lying on some closed convex surface F and complete its partition to a partition $\overline{Z}$ of this whole surface, then, according to the above remark, a certain development and the polyhedron glued from it will correspond to the partition $\overline{Z}$. Therefore, the polygon Q will be a polygon on this polyhedron. However, it is sufficient to consider the polygon Q itself from the viewpoint of its intrinsic metric only, whereas it is entirely irrelevant whether this polyhedron is realized in space or not. The only important thing is that, for a sufficiently fine partition Z, the metric of the polygon Q is arbitrarily close to the metric of the triangle T.

There are the following three types of vertices of the polyhedral polygon Q: (1) the three vertices A, B, and C that correspond to the vertices of the triangle T; (2) the vertices of the "second kind" that correspond to the vertices of the partition lying on the sides of the triangle T; (3) the "interior" vertices that correspond to the vertices of the partition lying inside the triangle T.

Let α, β, and γ be the angles at the vertices A, B, and C, and let $\delta_1, \ldots, \delta_m$ be angles at the vertices of the second kind. The curvature $\omega(Q)$ of the interior of the polygon Q is expressed by the formula

$$\omega(Q) = 2\pi - \sum_{j=1}^{m}(\pi - \delta_j) - [(\pi - \alpha) + (\pi - \beta) + (\pi - \gamma)],$$

i.e.,

$$\omega(Q) = (\alpha + \beta + \gamma - \pi) - \sum_{j=1}^{m}(\pi - \delta_j). \tag{11}$$

Since the angles do not increase when replacing the triangles T_i of our partition by the plane triangles, the angles α, β, and γ are no greater than the corresponding angles of the initial triangle; thus,

$$\alpha + \beta + \gamma - \pi \le \omega(T), \tag{12}$$

where $\omega(T)$ is the curvature of the interior of the triangle T.

The sum of the angles of the triangle T_i at the vertex lying on some side of the triangle T is equal to π, and since the angles never increase when passing to plane triangles, the angles at all vertices of the second kind do not exceed π, that is,

$$\delta_j \le \pi \ (j = 1, \ldots, m). \tag{13}$$

© 2006 by Taylor & Francis Group, LLC

By inequalities (12) and (13), Eq. (11) yields

$$\omega(Q) \le \omega(T), \tag{14}$$

that is, the curvature of the polygon Q is no greater than the curvature of the triangle T.

Drawing lines $\overline{AB}$, $\overline{AC}$, and $\overline{BC}$ that connect the vertices A, B, and C and are shortest in the triangle Q, we divide this triangle into the triangle $\overline{T} = ABC$ and three polygons Q_{AB}, Q_{AC}, and Q_{BC} (Fig. 89). These lines $\overline{AB}$, $\overline{AC}$, and $\overline{BC}$ do not pass through the vertices of the second kind since the angles at these vertices are $< \pi$. It might happen that, for example, the line $\overline{AB}$ passes through the vertex C if the angle $\gamma \ge \pi$. In this case, $\overline{AB} = \overline{AC} + \overline{CB}$. If this is true for an arbitrarily small partition into triangles T, then we obtain in the limit that $AB = AC + CB$, i.e., the sum of the two sides AC and CB of the triangle T is equal to the third and the angle at the vertex C is equal to π. But the angle γ in the triangle Q is always no greater than this angle, and hence, γ would be equal to π. In this case, the line $\overline{AB}$ can pass through the vertex C only if this line travels over the boundary of the polygon Q i.e., if the segments $\overset{\frown}{AC}$ and $\overset{\frown}{CB}$ of its boundary form a single shortest arc. Then the triangle ABC degenerates into a segment, and there remains only one of the polygons Q_{AB}, Q_{AC}, and Q_{BC} which coincides with Q itself. Therefore, without loss of generality, we can assume that the triangle ABC and the polygons Q_{AB}, Q_{AC}, and Q_{BC} exist, although some of them may degenerate into lines.

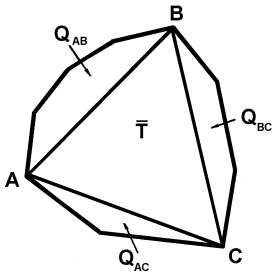


Fig. 89

We partition each of the polygons Q_{AB}, Q_{AC}, and Q_{BC} into triangles $\overline{T}^j$ by the diagonals drawn from the vertices A, B, and C (this is possible, since the angles at all vertices of the second kind are $\le \pi$). Thus, the polygon Q is represented as the union of triangles; i.e.,

$$Q = \overline{T} + \sum_{j=1}^{n} \overline{T}^j \quad (\overline{T} = ABC).$$

If $\overline{T}_0$ and $\overline{T}_0^j$ are two triangles with sides of the same length, then, by Lemma 1, we have

$$0 \le S(\overline{T}^j) - S(\overline{T}_0^j) \le \frac{1}{2}\omega(\overline{T}^j)d^2,$$

where d is the diameter of the polygon Q which is certainly not less than the diameters of the triangles T^j. Since the curvature of the polygon Q is equal to the sum of the curvatures of the triangles $\overline{T}^j$, adding these inequalities, we obtain

$$0 \le S(Q) - S(\overline{T}_0) - \sum_{j=1}^{n} S(\overline{T}_0^j) \ \le \frac{1}{2}\omega(Q)d^2. \tag{15}$$

© 2006 by Taylor & Francis Group, LLC

Consider the plane triangles $\overline{T}_0^j$ corresponding to the triangles $\overline{T}^j$ lying, say, in the polygon Q_{AB}. Adjoining them to each other in the same way as the triangles $\overline{T}^j$ do, come to the polygon Q_{AB}^0 that is bounded by the segment $\overline{A_0B_0}$, equal to the shortest line $\overline{AB}$, and the broken line A_0B_0, which is equal to the segment AB of the boundary of the polygon Q, i.e., to the side AB of the initial triangle T. But when refining the partition Z, the metric on Q converges to the metric of the triangle T; and, therefore, for a sufficiently fine partition, the difference $A_0B_0 - \overline{A_0B_0}$ becomes arbitrarily small. Together with this difference, the area of the polygon Q_{AB}^0 is also arbitrarily small. Consequently, for a sufficiently fine partition, the sum of the areas of all triangles $\overline{T}_0^j$ $(j > 0)$ is less than $\varepsilon/2$, where ε is a given positive number. In this case, (15) implies

$$0 \le S(Q) - S(\overline{T}_0) \le \frac{1}{2}\omega(Q)d^2 + \frac{\varepsilon}{2}. \tag{16}$$

The triangle $\overline{T}_0$ has the sides $\overline{AB}$, $\overline{BC}$, and $\overline{CA}$ which differ arbitrarily little from the sides AB, BC, and CA of the triangle T itself for a sufficiently fine partition Z. Therefore, there exists an arbitrarily fine partition such that modulus of the difference of the areas of the plane triangles $\overline{T}_0$ and T_0 is less than $\varepsilon/2$. In this case, substituting $S(T_0)$ for $S(\overline{T}_0)$ in (16), we obtain

$$-\varepsilon \le S(Q) - S(T_0) \le \frac{1}{2}\omega(Q)d^2 + \varepsilon;$$

since $\omega(Q) \le \omega(T)$ by (14), we finally infer that

$$-\varepsilon \le S(Q) - S(T_0) \le \frac{1}{2}\omega(T)d^2 + \varepsilon,$$

as required.

Now it is easy to prove our main assertion.

Theorem. *Each polygon on a convex surface has area. Let S be the area of a polygon P, and let ω be the curvature of its interior. If the polygon P is partitioned into triangles of diameter $\le d$, then the sum S_Z of the areas of the plane triangles with sides of the same length satisfies the inequality*

$$0 \le S - S_Z \le \frac{1}{2}\omega d^2.$$

Proof. Let Z_1 and Z_2 be two given partitions of the polygon P into triangles whose diameters are not greater than d_1 and d_2, respectively. Let S_{Z_1} and S_{Z_2} be the sums of the areas of the plane triangles corresponding to the triangles of these partitions. Show that

$$-\frac{1}{2}\omega d_1^2 \le S_{Z_1} - S_{Z_2} \le \frac{1}{2}\omega d_2^2.$$

To this end, we consider both partitions Z_1 and Z_2 simultaneously. Each triangle of one partition meeting with the triangles of the other yields finitely many polygons. Each of these polygons can be triangulated, and hence there exists an arbitrarily

© 2006 by Taylor & Francis Group, LLC

fine partition Z of the polygon P that is simultaneously a subpartition of each of the given partitions Z_1 and Z_2.

Let N be the number of triangles of the partition Z, and let ε be a given positive number. Take a certain triangle T of the partition Z_1; in the partition Z, this triangle falls into small triangles t_i. Replacing each of these triangles by a plane triangle, we obtain a "polyhedral" polygon Q. According to Lemma 2, the partition Z can be taken so fine that

$$-\frac{\varepsilon}{N_1} \leq S(Q) - S(T_0) \leq \frac{1}{2}\omega(T)d_1^2 + \frac{\varepsilon}{N_1},$$

where $S(T_0)$ is the area of the plane triangle with sides of the same length as those of the triangle T. Sum up all these inequalities for all triangles T of the partition Z_1. Then the sums of the areas $S(Q)$ and $S(T_0)$ yield the sums S_Z and S_{Z_1} of the areas of the plane triangles corresponding to the partitions Z and Z_1, while the sum of the curvatures $\omega(T)$ yields a value that is no greater than the curvature of the polygon P. Consequently, we have

$$-\varepsilon \leq S_Z - S_{Z_1} \leq \frac{1}{2}\omega d_1^2 + \varepsilon. \tag{17}$$

But if the partition Z is sufficiently fine, then, in exactly the same way, we infer

$$-\varepsilon \leq S_Z - S_{Z_2} \leq \frac{1}{2}\omega d_2^2 + \varepsilon. \tag{18}$$

Subtracting (17) from (18) and considering that ε can be taken arbitrarily small, we obtain

$$\frac{1}{2}\omega d_1^2 \leq S_{Z_1} - S_{Z_2} \leq \frac{1}{2}\omega d_2^2. \tag{19}$$

Now, if we have a sequence of partitions Z_n of the polyhedron P into triangles of diameter $\leq d_n$, then

$$\frac{1}{2}\omega d_m^2 \leq S_{Z_m} - S_{Z_n} \leq \frac{\omega}{d_n}^2.$$

Consequently, if $d_n \to 0$ as $n \to \infty$ then the numbers S_{Z_n} form a convergent sequence. But this means that the area of P exists. On the other hand, if we refine the partition Z_1 in inequality (19) on letting $d_1 \to 0$, then, by what we have just proved, the areas S_{Z_1} converge to the area S of the polygon P. Therefore, passage to the limit as $d_1 \to 0$, in inequality (19) yields

$$0 \leq S - S_Z \leq \frac{1}{2}\omega d^2 \tag{20}$$

where the subscript 2 is omitted in the notation of Z and d. This proves the theorem. It is hardly necessary to note that our arguments are of purely intrinsic character and, therefore, the same result is true in an arbitrary manifold with a metric of positive curvature.

In particular, if the polygon P is a triangle T and the partition Z consists only of this triangle, then formula (20) shows that the area of the triangle T is no less than the area of the plane triangle with the sides of the same length.

© 2006 by Taylor & Francis Group, LLC

Now it is easy to prove that *if a polygon P is partitioned into polygons $P_1, \ldots, P_k$, then the area of the former is equal to the sum of the areas of the latter.* To prove this, it is sufficient to consider those of the polygon P which are simultaneously triangulations of the polygons $P_1, \ldots, P_k$.

After that, we can define the area of a figure composed of finitely many disjoint polygons as the sum of the areas of these polygons. The definition of area of sets of a more general shape can be carried out as follows. We agree that the area of an open set G is equal to the least upper bound of the areas of the sets composed of finitely many polygons and lying in G. The area of a closed set M is defined to be the greatest lower bound of the areas of the sets including M and also composed of finitely many polygons. Finally, for an arbitrary set M, we can define outer and inner measures; the former is the greatest lower bound of the areas of the open sets including M, and the latter is the least upper bound of the areas of the closed sets lying in M. A set at which both measures coincide is said to be measurable, and the common value of the outer and inner measures is the area or measure of this set. All Borel sets turn out to be measurable, and it is possible to prove that the so-defined measure is completely additive. The proof of these assertions is a repetition of the familiar inferences of Lebesgue measure theory.

In this way, we can obtain an instance of intrinsic-geometric measure theory on convex surfaces. However, this way is in fact very long, and, therefore, we proceed otherwise. Namely, in the next section, we will derive an expression for area in the shape of a usual double integral; after that, the required properties of area (measure) will become simple corollaries to the well-known properties of the integral. This way, although being not intrinsic-geometric, will let us attain the goal much faster, not mentioning the fact that the expression of area in the shape of the integral is of interest in its own right.

2. The Extrinsic–Geometric Meaning of Area

Theorem 1. *Let convex surfaces F_n converge to a convex surface F, and let polygons P_n on the surfaces F_n converge to a polygon P on F; moreover, assume that the number of sides of each of the polygons P_n is bounded above. Then the areas of P_n converge to the area of P.*

Proof. Consider partitions Z_n of the polygons P_n into triangles of diameter less than a given d. Since the curvature of the polygon P_n cannot exceed 4π, we have

$$|S(P_n) - S_{Z_n}| \le 2\pi d^2 \tag{1}$$

by the theorem of the previous section. Since the number of the sides of the polygons P_n is bounded, we can take the partitions Z_n so that the number of the triangles in every P_n is less than some possibly very large number. Therefore, we can chose a sequence P_{n_i} from the sequence of polygons P_n such that all triangles of the partitions Z_{n_i} converge, and these triangles comprise in the limit some partition Z of the polygon P into triangles of diameter no greater than d. For this partition Z, the theorem of the previous section yields

$$|S(P) - S_Z| \le 2\pi d^2. \tag{2}$$

© 2006 by Taylor & Francis Group, LLC

But when the triangles of the partitions Z_{n_i} converge to the triangles of the partition Z, their sides also converge, and hence, the areas of the plane triangles corresponding to them also converge. Therefore, we have

$$S_Z = \lim_{i\to\infty} S_{Z_{n_i}}$$

for the sum of the areas of these triangles. So, inequalities (1) and (4) imply

$$\limsup_{i\to\infty} S(P_n) - 2\pi d^2 \le S(P) \le \liminf_{i\to\infty} S(P_{n_i}) + 2\pi d^2.$$

Since d can be taken arbitrarily small, $\lim_{i\to\infty} S(P_{n_i})$ exists and equals $S(P)$. Consequently, each sequence P_{n_i} chosen from the polygons P_n contains a subsequence in which the areas converge to the area of the polygon P. This implies that the areas of the polygons P_n converge to the area of the polygon P, as required.

If we consider polyhedra as the surface F_n in Theorem 1, then this theorem reduces to the following.

Theorem 2. *The area of a polygon P on a convex surface F is equal to the limit of the areas of polygons P_n on convex surfaces that converge to P provided that the number of sides of the polygons P_n remains bounded.* (However, we will show that this condition is superfluous.)

This theorem shows that our intrinsic definition of area is equivalent to its extrinsic definition as the limit of the areas of polyhedra converging to a given surface, since the polygon P itself can be considered as a surface while the polygons P_n, as polyhedra converging to this surface.

In order to obtain the expression of area in the shape of a double integral starting from this observation, we need the following theorem.

Theorem 3. *The projection to a plane of the set of points of a convex surface at which this surface has no tangent plane is a set of plane measure zero.*[3]

Proof. It suffices to prove the theorem for a surface that has no tangent planes orthogonal to the projection plane E. Indeed, each complete convex surface can be decomposed into three parts F_1, F_2, and F_3 so that F_3 consists of the points at which there are supporting planes orthogonal to E, and F_1 and F_2 have no such supporting planes and are in a one-to-one fashion projected to E. The projection of the set F_3 on the plane is a convex curve bounding the projection of the complete surface F and hence has measure zero.

Let the convex surface F have no support planes that are orthogonal to a given plane E. Assume that this plane is $z = 0$ in the rectangular coordinate system x, y, z; then the surface is represented by the equation

$$z = z(x, y).$$

If the function $z(x, y)$ has both partial derivatives z_x and z_y at a given point (x, y) then the surface F has two tangent lines at the corresponding point X. The

[3]This theorem was proved by Reidemeister in 1921; see K. Reidemeister, Über die singulären Randpunkte eines konvexen Körpers. Math. Ann., Vol. 83 (1921), pp. 116–118.

© 2006 by Taylor & Francis Group, LLC

tangent cone at X includes these lines, and since it is convex, this cone is a plane. Consequently, the surface F has tangent plane at X if and only if the function $z(x, y)$ has both partial derivatives z_x and z_y.

Since the partial derivatives z_x and z_y are the limits of the continuous functions $(z(x + h, y) - z(x, y))/h$ and $z(x, y + h) - z(x, y)/h$, the sets M_x and M_y of the points (x, y) at which they do not exist are measurable.[4]

Intersect the surface F by the plane $y = c$ to obtain a convex curve $z = z(x, c)$ in the cross-section. The set of points at which it has no tangent, i.e., at which there is no derivative $z_x(x, c)$ is at most countable.[5] Hence, each line $y = c$ intersects the set M_x at the points of an at most countable set. As is known, if the set M_x is countable then this set is of measure zero. In exactly the same way, the set M_y has zero measure. The set of the points at which at least one of the derivatives z_x and z_y does not exist is $M_x + M_y$ and hence this set is also of measure zero. But we have proved that this set is exactly the projection of the set of the points of the surface F at which there is no tangent plane. Thus, the theorem is proved.

Theorem 4. *Let a convex surface F be represented by the equation $z = z(x, y)$ in the rectangular coordinate system and have no supporting planes parallel to the axis z. Then the area of each polygon P on the surface F is expressed by the integral*

$$S(P) = \iint_{P'} \sqrt{1 + z_x^2 + z_y^2}\, dx\, dy, \tag{3}$$

which is taken over the projection P' of this polygon to the plane $z = 0$.

Proof. Construct a sequence of convex polyhedra F_n converging to F and take polygons P_n on these polyhedra which converge to P.

Let $\phi(x, y)$ be the angle between the supporting plane of the surface F at the point $(x, y, z(x, y))$ with the plane $z = 0$, and let $\phi_n(x, y)$ be a similar angle for the polyhedron F_n. The functions $\phi(x, y)$ and $\phi_n(x, y)$ are not single-valued in general, since there can be points on F and F_n at more than one supporting plane. However, by Theorem 3, the set of the points at which these functions are not single-valued has measure zero, so it can be neglected if the functions $\phi(x, y)$ and $\phi_n(x, y)$ stand under the integral sign.

Since the limit of supporting planes to the polyhedra F_n is a supporting plane to the surface F (Lemma 2 of Sec. 2 of Chapter V), we have $\lim_{n\to\infty} \phi_n(x, y) = \phi(x, y)$

[4] See, e.g., P. S. Aleksandrov and A. N. Kolmogorov, Introduction to Real Function Theory, Chap. 8, Sec. 1.

For brevity, we put

$$\frac{1}{h}[z(x + h, y) - z(x, y)] = z_h(x, y).$$

Let M_{nm} be the set of points (x, y) for which the following condition holds: if $|h| \le 1/m$ and $|k| \le 1/m$ then

$$|z_h(x, y) - z_k(x, y)| < \frac{1}{n}.$$

It is easy to verify that the set M_{nm} is closed. The set $\prod_{nm} \sum M_{nm}$ is the set of the points at which the derivative z_x exists, and this set is measurable; this is an $F_{\sigma\delta}$-set.

[5] The tangent to a convex curve revolves monotonically, and its total swerve is not greater than 2π. Consequently, the number of jumps of the tangent is at most countable.

© 2006 by Taylor & Francis Group, LLC

at all points (x, y) where $\phi(x, y)$ is single-valued, i.e., almost everywhere. But since the surface F has no vertical supporting planes, $\phi_n(x, y)$ do not approach $\pi/2$ in any closed set lying in the projection of the surface F.

We have the following obvious formula for the surface area of the polygon P_n on the polyhedron:

$$S(P_n) = \iint\limits_{P_n'} \frac{dx\,dy}{\cos\phi_n(x, y)}, \tag{4}$$

where P_n' is the projection of P_n to the plane $z = 0$. When P_n converge to the polygon P, their projections P_n' converge to its projection P'. Therefore, if we take integral (4) over the domain P_n', then the error tends to zero as n increases.[6]

On the other hand, $\lim_{n\to\infty} \phi_n(x, y) = \phi(x, y)$ almost everywhere, and since $\phi_n(x, y)$ approaches $\pi/2$ nowhere in P', the function $1/\cos\phi_n$ is bounded. Therefore, on the basis of the well-known theorem on passage to the limit under the integral sign, we have

$$\lim_{n\to\infty} S(P_n) = \lim_{n\to\infty} \iint\limits_{P_n'} \frac{dx\,dy}{\cos\phi_n} = \iint\limits_{P'} \frac{dx\,dy}{\cos\phi}. \tag{5}$$

But, by Theorem 1, the limit of the areas of the polygons P_n is the area of the polygon P. Consequently,

$$S(P) = \iint\limits_{P'} \frac{dx\,dy}{\cos\phi(x, y)}.$$

This is formula (9), since[7]

$$\cos\phi(x, y) = \frac{1}{\sqrt{1 + z_x^2 + z_y^2}}$$

Now, let M be a set on the surface F such that its projection M' to the plane $z = 0$ is measurable, and M has no points at which their supporting planes make arbitrarily small angles with the axis z, so that the function $\sqrt{1 + z_x^2 + z_y^2}$ is bounded on M'. Then we can define the area of M by the formula

$$S(M) = \iint\limits_{M'} \sqrt{1 + z_x^2 + z_y^2}\,dx\,dy. \tag{6}$$

We can show that this definition coincides with that sketched at the end of the previous section. However, there is a certain inconvenience here that consists in the fact that formula (6) has a meaning only in the case where the set M does not

[6] Of course, it can happen that the projection of F_n does not cover P' and, thus, the function $\phi_n(x, y)$ is not defined on P', but this is impossible for large n.

[7] The representation of the area by formula (3) is well known not only for convex surfaces but for a wider class of surfaces. See, e.g., H. Lebesgue, Intègrale, longue, aire. Annali di Mat. (3), Vol. VII (1922), p. 315.

© 2006 by Taylor & Francis Group, LLC

contain points at which its supporting planes make arbitrarily small angles with the axis z; and moreover, this formula is not applicable if we consider a surface that is not uniquely projected to the plane $z = 0$. Therefore, it is more convenient in some cases to use the representation of a surface not in the rectangular but rather in the spherical coordinates. Let F be a convex surface not degenerating into a plane domain. Choose a point O inside the convex body, bounded by this surface, and take this point as the origin of the spherical coordinates r, ϑ, and ψ. With each point of the surface F, we associate the point (ϑ, ψ) on the unit sphere E centered at the point O, and the surface F is represented by the equation $r = r(\vartheta, \psi)$. In full analogy with Theorem 3, we can prove the following theorem.

Theorem 3^***.** *The set of the points on the sphere S which correspond to those points of the surface F at which there is no tangent plane has measure zero.*

Let $\phi(\vartheta, \psi)$ be the angle between the normal to the supporting plane of the surface F at the point $(\vartheta, \psi, r(\vartheta, \psi)$ and the ray going to this point from the origin O. If the surface F is a polyhedron then we have the following obvious formula for the area of a polygon P on this surface:

$$S(P) = \iint\limits_{P'} \frac{r^2}{\cos\phi}\, d\sigma,$$

where $d\sigma = \sin\vartheta d\vartheta d\psi$ is the surface element of the sphere E and P' is the projection of P to the sphere E. This formula can be extended to arbitrary convex surfaces in exactly the same way as above for formula (4). In this way, we obtain the following theorem which is similar to Theorem 4:

Theorem 4^***.** *If a convex surface is represented by the equation $r = r(\vartheta, \psi)$ in the spherical coordinates, then the area of a polygon P on this surface is expressed by the formula*

$$S(P) = \iint\limits_{P'} \frac{r(\vartheta, \psi)^2}{\cos(\phi(\vartheta, \psi)}\, d\sigma; \tag{7}$$

moreover, $d\sigma = \sin\vartheta d\vartheta d\psi$ where $\phi(\vartheta, \psi)$ has a unique value, i.e., almost everywhere, and $\phi(\vartheta, \psi)$ is the angle between the normal and the radius.

With each set M on a surface F, we associate its central projection M' to the sphere E. If the set M' is measurable, then the integral

$$\iint\limits_{M'} \frac{r^2}{\cos\phi}\, d\sigma \tag{8}$$

exists, and we can take this integral as the area $S(M)$ of M. This definition coincides with that at the end of the previous section. Indeed, if M is an open set then we can represent this set as the union of an increasing sequence of "elementary" sets $M_1 \subset M_2 \subset \ldots$, i.e., those sets each of which is composed of finitely many polyhedra. In this case, the measure of the set $M' - M'_n$ tends to zero and, therefore,

$$\iint\limits_{M'} \frac{r^2}{\cos\phi}\, d\sigma = \lim_{n\to\infty} \iint\limits_{M'_n} \frac{r^2}{\cos\phi}\, d\sigma.$$

© 2006 by Taylor & Francis Group, LLC

Consequently,

$$S(M) = \lim_{n\to\infty} S(M_n),$$

i.e., the area of an open set is the least upper bound of the areas of "elementary" sets lying in it.

If M is closed then we can represent M as the intersection of a nested sequence of elementary sets $M_1 \supset M_2 \supset \ldots$. Therefore, in full analogy with the above, we obtain $S(M) = \lim_{n\to\infty} S(M_n)$ and, therefore, the area of a closed set is the greatest lower bound of the areas of the elementary sets lying in it.

Now let M be some set on a surface F whose projection is measurable. As is known, for each measurable set M' there exists a sequence of open sets M'_n including M'_n such that the measure of M' is equal to the limit of the measures of M'_n. In this case,

$$\iint\limits_{M'} \frac{r^2}{\cos\phi}\, d\sigma = \lim_{n\to\infty} \iint\limits_{M'_n} \frac{r^2}{\cos\phi}\, d\sigma. \tag{9}$$

But the open set M'_n on the sphere E is the projection of some measurable set to the surface F. Therefore, formula (9) implies that the area of M is the greatest upper bound of the areas of the open sets that include M. In a similar way, we prove that the area of M is the least upper bound of the areas of closed sets lying in M.

If $M = \sum_{i=1}^{\infty} M_i$ where the sets M_i are pairwise disjoint and have definite area, i.e., there exist the corresponding integrals (8), then

$$\iint\limits_{M'} \frac{r^2}{\cos\phi}\, d\sigma = \iint\limits_{M'_n} \sum_{i=1}^{\infty} \frac{r^2}{\cos\phi}\, d\sigma,$$

i.e.,

$$S(M) = \sum_{i=1}^{\infty} S(M_i).$$

This means that area is a completely additive function.

Finally, the projection of a closed (open) set of the surface F to sphere E is a closed (open) set on the sphere; the projection of a union (intersection) of sets is the union (intersection) of their projections to the sphere E. Consequently, the projection of a Borel set on the surface F is a Borel set to the sphere E. Therefore, if M is a Borel set then integral (8) exists, and so this set has area.

Therefore, all assertions formulated at the end of the previous section are proved. The so-obtained result can be summarized as the following theorem.

Theorem 5. *Measure theory on the sphere can be translated to each* (*nondegenerate*) *closed convex surface as a result of the central projection of this surface to the sphere.*

In turn, measure theory on a surface degenerating into a doubly-covered domain is constructed in an obvious way.

© 2006 by Taylor & Francis Group, LLC

3. Extremal Properties of Pyramids and Cones

Here, we deal with the problems of the following type: among all convex surfaces homeomorphic to a disk, enjoying given properties and bounded by a curve, find a surface of maximal area. For example, we will prove that among all surfaces that have a given perimeter, i.e., a given length of the bounding curve, and that also have a given curvature $\omega < 2\pi$, the maximal area is attained at the lateral surface of the right circular cone. If the curvature $\omega = 0$ then this result transforms into the well-known theorem asserting that the disk has maximal area among all plane domains with a given perimeter. The curvature of a convex surface bounded by a closed curve is expressed through the swerve τ of this curve by the formula $\omega = 2\pi - \tau$, derived in Sec. 2 of Chapter IX; therefore, the assignment of the curvature of a surface is equivalent to the assignment of the swerve of the bounding curve, and the condition $\omega < 2\pi$ is equivalent to the requirement that the swerve τ is positive. The requirement that the curvature is less than 2π is necessary, since otherwise the possible maximum of area turns out infinite. For example, if the curvature of a cone with a given perimeter tends to 2π then its height, together with its area, increases indefinitely.

Our arguments will be of intrinsic-geometric character, and in order to avoid using the realization theorem without necessity, *by a convex surface we mean a domain that is homeomorphic to a disk and cut out from a manifold with a metric of positive curvature.* Similarly, *by a convex polyhedron we mean a polygon that is homeomorphic to a disk and cut out from a manifold with polyhedral metric of positive curvature. The sides and angles of this polygon are the sides and angles of this polyhedron.*

The "polyhedron" under consideration has vertices of the following two types: "exterior" vertices lying on the boundary and "interior" one. From the viewpoint of the intrinsic metric, the exterior vertices are the vertices of the corresponding polygon, while the interior vertices are those points at which the complete angle is less than 2π.

Theorem 1. *Among all convex polyhedra with given sides and angles that are less than π, a polyhedron containing one interior vertex has maximal area. This polyhedron is isometric to the lateral surface of a pyramid.*

Proof. Let P be a polyhedron with given sides whose angles are less than π. Assume that this polyhedron contains at least two vertices A and B with curvatures ω_A and ω_B. Draw the shortest line AB on P; since all the angles of the polyhedron P are less than π, this is a geodesic traveling inside P.

Take two equal plane triangles $A_1B_1C_1$ and $A_2B_2C_2$ with bases $A_1B_1 = A_2B_2 = AB$ and with angles $\angle A_1 = \angle A_2 = \omega_A/2, \angle B_1 = \angle B_2 = \omega_B/2$. These triangles exist, since the curvature of the polyhedron P is less than 2π, and hence, $\omega_A/2 + \omega_B/2 < \pi$. Cut the polyhedron P along the line AB and glue our triangles to both sides of the cut in such a way that their vertices A_1 and A_2 coincide with the point A, while the vertices B_1 and B_2 coincide with the point B. After that, we identify the side A_1C_1 with the side A_2C_2 and identify the side B_1C_1 with the side B_2C_2. As a result, we glue the cut away. Two angles equal to $\omega_A/2$ and $\omega_B/2$ are added at the point A and B, i.e., exactly those angles by which the complete angles at

© 2006 by Taylor & Francis Group, LLC

the points A and B differ from 2π. Consequently, the points A and B are no longer the vertices, and instead of them, there appears only one vertex C. At the same time, the sides and angles of the polyhedron P are not changed, while the area of P increases.

If there are other interior vertices of the polyhedron P, then we can perform the same operation on them and continue until we arrive at a polyhedron P_0 that has only one interior vertex. This polyhedron has the same sides and angles, but its area is greater than that of the initial polyhedron P. Now, if we show that there is only one convex polyhedron (up to an isometry) with given sides and angles and with a unique interior vertex, then we will thus prove that the obtained polyhedron is exactly the polyhedron with a maximal area.

Let P_1 be another polyhedron with a single interior vertex and the same sides and angles as those of the polyhedron P_0. Since the angles of these polyhedra are equal, their curvatures are equal; but this means that the complete angles at their vertices are equal.

Take respective vertices A_0 and A_1 on the boundaries of the polyhedra P_0 and P_1 and connect them by shortest arcs O_0A_0 and O_1A_1 with the interior vertices O_1 and O_2 of these polyhedra; then cut our polyhedra along these shortest arcs and develop them onto the plane. We obtain two polygons P_0' and P_1' (Fig. 90). The polygon P_0' has the following vertices: O_0' which corresponds to O_0, A_0' and A_0'' which correspond to the vertex A_0, and also two vertices that correspond to other exterior vertices of the polyhedron P_0. The broken line L_0 bounding P_0 transforms into the broken line L_0' which, together with two segments $O'A_0 = O_0'A_0'' = O_0A_0$, bounds the polygon P_0'. The polygon P_1' has a similar structure. Moreover, the angles at the vertices O_0' and O_1' of our polygons are equal since they are equal to the complete angles at the interior vertices of the polyhedra P_0 and P_1. Thus, the isosceles triangles $O_0'A_0'A_0''$ and $O_1'A_1'A_1''$ have equal bases and opposite angles; therefore, their lateral sides and angles are also equal. Consequently, the polygons P_0' and P_1' have equal sides and angles so that these polygons themselves are equal; but this means that the polygons P_0 and P_1 are isometric.

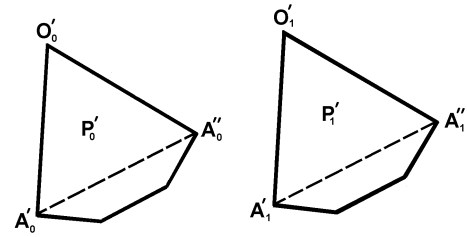


Fig. 90

Thus, the polyhedron P_0 is uniquely determined and, therefore, has maximal area. Since all its angles are less than π, according to what was proved in Sec. 4 of Chapter IX, this polyhedron is isometric to a "cap", i.e., in this case, to a polyhedron with plane boundary. This polyhedron is obviously the lateral surface of a pyramid.

The result obtained can be generalized as follows.

Theorem 2. *Consider convex surfaces bounded by curves with the following properties:* (1) *the lengths of these curve are equal;* (2) *the swerves of the arcs of these curves are always nonnegative and the total swerves of the curves themselves are*

© 2006 by Taylor & Francis Group, LLC

positive; (3) *every two curves admit a mapping from one to the other which preserves length and under which the respective arcs have equal swerves. Among all such surfaces, a cone isomorphic to the lateral surface of a solid convex cone has maximal area.*

Since the swerve of a broken line is the sum of the complements of its angles to π, then the condition of Theorem 1 which says that the angles of the polyhedron are less than π corresponds exactly to condition (2) of this theorem, while the condition of equality of angles corresponds to condition (3).

We have observed in Sec. 4 of Chapter IX that a surface bounded by a curve each of whose arcs has nonnegative swerve is "convex in itself", i.e., every two of its points can be connected by a shortest arc that does not travel over the boundary whenever the points do not lie on the boundary. However, this assertion was given without proof, and, therefore, we can simply require in Theorem 2 that the surfaces under consideration are convex in themselves. We sketch the proof of Theorem 2 exactly under this assumption.

Let F be a surface of the type considered, and let L be the curve bounding this surface. Inscribe a closed broken line L' into L and partition each polygon bounded by it into small triangles; then replace each of these triangles by the plane triangle with sides of the same length. As a result, we obtain a polyhedron P given by its development. This polyhedron satisfies the conditions of Theorem 1, and, therefore, its area $S(P)$ is no greater than the area $S(P_0)$ of the corresponding polyhedron P_0 with a single interior vertex. If we take the broken line L' closer and closer to the curve L and take the partition finer and finer, then the area of the polyhedron P tends to the area of our surface F; that is,

$$S(F) = \lim S(P).$$

Simultaneously, the polyhedra P_0 converge to a certain cone F_0 whose boundary curve has the same length and the same swerves of the arcs as the boundary curve L of the surface F. (The fact that such cone is unique up to isometry is proved almost in the same way as the corresponding assertion in Theorem 1.)

Therefore, $S(F_0) = \lim S(P_0)$; and since $S(P_0) \geq S(P)$, (1) implies $S(F_0) \geq S(F)$, i.e., the cone F_0 has maximal area in fact.

Show that at the same time, no other surface of the type under consideration can have the same area. Let F_1 be such a surface. If this surface itself is not a cone, then we can find two disjoint convex polygons Q_1 and Q_2 on this surface, each of positive curvature. As was just proved, to each of the polygons Q_1 and Q_2, there corresponds a cone, i.e., in this case, the lateral surface of a pyramid with the same sides and angles and at least the same area. By the gluing theorem, we can cut out the polygons Q_1 and Q_2, and glue the corresponding cones instead of them. After that, we make a cut along the shortest arc connecting the vertices of these cones and glue the cut by two equal triangles in the same way as in the proof of Theorem 1. The area of the surface increases. This proves that only some cone can have maximal area, but, as was already mentioned, such a cone is unique.

Of course, it is very essential for our arguments that the swerve of each of the arcs of the boundary curve is nonnegative; without this assumption, the existence of shortest lines connecting interior vertices and having no common points with

© 2006 by Taylor & Francis Group, LLC

the boundary of a surface cannot be guaranteed. The following question arises: what result can be obtained if we omit this condition, i.e., if we search for the maximal surface area given not necessarily nonnegative swerves of the arcs of the boundary curve? Of course, the swerve of the whole curve should be positive, since otherwise the curvature of the surface is $\geq 2\pi$; and so, as was shown, the existence of a maximum cannot be guaranteed. If the curvature is $< 2\pi$, then the maximum exists, but the existence of a surface which realizes this maximum is not obvious. It is interesting to study this question even in the case of polyhedra.

Now we turn to the following problem: among all convex surfaces with a given perimeter and curvature $\omega < 2\pi$, find a surface of maximal area. Here, no conditions are imposed on the swerves of the arcs of the boundary curve. As above, we first solve this problem for polyhedra. Namely, we will consider convex polyhedra that satisfy the following conditions: (1) the number of all vertices of each of these polyhedra does not exceed a given number n; (2) their perimeters do not exceed a given l; (3) their curvatures do not exceed a given $\omega < 2\pi$. The angles of a polyhedron are always less than 2π, but we add those for which there are angles equal to 2π to the polyhedra considered; this corresponds to the case in which some neighboring sides can overlap. The consideration of these limit cases makes the set of the polyhedra under study closed, and thus simplifies the proof of the existence of a polyhedron of maximal area. In this case, the existence of such a polyhedron cannot be proved as directly as in Theorem 1; however, it ensues from the following lemma.

Lemma 1. *The diameter of a convex polyhedron of perimeter l and curvature ω is at most $2l/(2\pi - \omega) + l/2$.*

Proof. Let P be a convex polyhedron of perimeter l and curvature ω, and let O be a point that lies inside this polyhedron and is not an interior vertex. Draw shortest lines to all points of the boundary of P from O. Since we add to P its boundary, these shortest lines exist but can have common points with the boundary of P. Each of these shortest lines is a geodesic broken line with vertices at those exterior vertices of P at which the complete angles $> \pi$ (Theorem 6 of Sec. 2 of Chapter II). Since a geodesic cannot pass through vertices of a polyhedral metric, all interior vertices turn out to be contained in domains D_i each of which is bounded by two shortest lines that go to the same point of the boundary. Cutting out these domains D_i and identifying the pairs of the shortest lines bounding these domains, we obtain a new polyhedron P_1 with a single interior vertex O. Under this procedure, the complete angle at the point O decreases by the sum of the angles α_i we have cut out. Each angle α we have cut out belongs to one of the domains D_i. The curvature of the domain D_i is expressed by the formula

$$\omega(D_i) = \alpha_i + \beta_i + \sum_{p=1}^{k-2} \gamma_i^{(p)} - k\pi, \tag{1}$$

where α_i is the angle at O, while β_i is the angle at the endpoint on the boundary, and $\gamma_i^{(p)}$ are other angles of the domain D_i which can exist if the shortest line bounding the domain D_i passes through exterior vertices of the polygon; finally,

© 2006 by Taylor & Francis Group, LLC

k is the number of all vertices of the domain D_i. All angles $\gamma_i^{(p)} \geq \pi$, and so (1) implies $\omega(D_i) > \alpha_i$.

Since the domains D_i contain all interior vertices of the polyhedron P, the sum of their curvatures is equal to the curvature ω of this polyhedron, and so

$$\omega(P) = \sum_i \omega(D_i) > \sum_i \alpha_i. \tag{2}$$

But the complete angle θ at the vertex O of the polyhedron P is

$$\theta = 2\pi - \sum_i \alpha_i, \tag{3}$$

and hence this angle satisfies the inequality

$$\theta > 2\pi - \omega(P). \tag{4}$$

Since the polyhedron P_1 contains only one vertex O, cutting this polyhedron along the shortest line traveling from O to the boundary, we can develop P onto the plane. The resulting plane polygon has angle θ at the vertex O and contains the circular sector centered at O whose radius r is equal to the minimal distance from O to the boundary of the polyhedron P_1. The arc length θr of this sector is not greater than the length of the broken line into which the boundary of P_1 develops, i.e.,

$$\theta r < l;$$

together with (4), this yields

$$r < \frac{l}{2\pi - \omega}.$$

Since the point O in the polyhedron P was chosen arbitrarily, the distance from its every point to the boundary is thus less than $l/(2\pi - \omega)$. Now, if X and Y are two points of P and X' and Y' are the points of the boundary which are nearest to them, then XX' and $YY' < l/(2\pi - \omega)$, and obviously, $X'Y' \leq l/2$. Consequently,

$$XY \leq XX' + YY' + X'Y' < \frac{2l}{2\pi - \omega} + \frac{l}{2};$$

as claimed.[8]

Lemma 2. *There exists a polyhedron of maximal area among the polyhedra with a bounded number of vertices n of perimeter l and curvature at least a given $\omega < 2\pi$.*

Proof. Let P be a given polyhedron. Connect one of its interior vertices A_1 to other interior vertices A_i by shortest lines in this polyhedron. Cutting P along these lines A_1A_i, we transform P into a polygon that does not contain interior vertices and has at most $3n$ exterior vertices (the vertex A_1 counted as many times as the number of the lines A_i1A_i meeting at this vertex). This polygon can be partitioned by diagonals into at most $3n$ triangles (the proof of this fact is the same as Lemma

[8]The obtained estimate is far from being sharp. The sharp estimate is as follows: the diameter of P is $< 1/2$ if $\omega \leq \pi$, and the diameter of P is $\leq l/(2\sin\frac{2\pi-\omega}{2}) = l/2\sin(\omega/2))$.

© 2006 by Taylor & Francis Group, LLC

1 of Sec. 1 of Chapter IV.) We obtain a development that defines the polygon P. Such a development with a given structure is determined by the length of its edges. But the number of the structures of developments constructed from a given set of triangles is finite, and according to Lemma 1, the lengths of all edges are bounded. Consequently, each development, together with the polygon P, is given by finitely many variables that assume only bounded values. Since the set of the polygons under consideration is closed,[9] there is a polygon of maximal area among them by the Weierstrass principle.

Lemma 3. *If a polyhedron P of a given perimeter and curvature has maximal area, its every angle is at most π.*

Proof. Assume that the polygon P of maximal area has angles equal to 2π. Then, gluing the segments of the sides meeting at some vertex with such an angle, we diminish the perimeter of P keeping the number of vertices, curvature, and area. But then the area of P may be increased keeping the perimeter of P the same. To this end, it suffices to glue an appropriate triangle to some side a of the polygon P whose one side is a prolongation of the side of the polygon P which is adjacent to a. Therefore, the polygon P cannot have angles equal to 2π.

Assume that the angle at the vertex A of the polygon P is greater than π but less than 2π. Take a plane quadrangle $A'B'C'D'$ with sides $A'B' = AB$, $A'C' = AC$ and with the angle $\alpha' = 2\pi - \alpha$ at the vertex A' and angles at B' and C' so small that their sums with the angles β and γ at the vertices B and C are less than 7π. Of course, such a quadrangle can be chosen so that $D'C' + D'B' < A'C' + A'B' = AC + AB$ (Fig. 91).

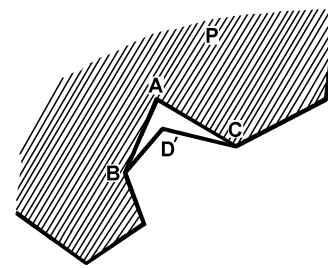


Fig. 91

Identify the sides AW and AC of the polygon P with the sides $A'B'$ and $A'C'$ of our quadrangle. As a result, we obtain a new polygon P_1 of a smaller perimeter, since $D'C' + D'B' < AC + AB$. Its curvature is the same, since the complete angle at the point A turns out to be equal to 2π, so that no new vertex appears. But the area of P_1 is greater than the area of P, which contradicts the assumption that P has maximal area.

Theorem 3. *The lateral surface of a pyramid has maximal area among all convex polyhedra with a given number of sides and curvature $\omega < 2\pi$.*

[9] If a sequence of polyhedra P_m of given perimeter and number of vertices bounded above converges, then the limit polyhedron obviously has the same perimeter and the number of its vertices is no greater than the above bound. However, the curvature of this polyhedron can be less than the limit of the curvatures of the polyhedra P_m, since in the limit, an interior vertex can enter the boundary. The curvature of the polyhedron P_m is

$$\omega(P_m) = 2\pi - \sum(\pi - \alpha_m^{(i)}),$$

where $\alpha_m^{(i)}$ are its angles. If $\alpha^{(i)}$ are the angles of the limit polyhedron, then by the theorem on convergence of angles (Theorem 4 of Sec. 4 of Chapter III), $\alpha^{(i)} \le \liminf_{m\to\infty} \alpha_m^{(i)}$, and hence $\omega(P) \le \liminf_{m\to\infty} \omega(P_m)$. That is why we consider polyhedra of curvature at most a given ω rather than equal to ω precisely.

© 2006 by Taylor & Francis Group, LLC

Proof. Consider convex polyhedra with the number of vertices at most n of perimeter e and curvature at most $\omega < 2\pi$. By Lemma 2, there exists a polyhedron P of maximal area among them; by Lemma 3, this polyhedron has no angles $> \pi$. Hence, according to Theorem 1, this polyhedron can have only one interior vertex. We have to show that all the sides of the polyhedron P are equal and all its exterior vertices are equidistant from the interior vertex. To this end, we cut the polyhedron P along the shortest line connecting its interior vertex O with some exterior vertex A and develop this polyhedron onto the plane. We obtain a polygon P' such that the angle at the vertex O', which corresponds to O, is equal to the complete angle at P, the sides meeting at the vertex O' are equal, and the other sides and vertices correspond to the sides and exterior vertices of the polyhedron P. We have to show that the polygon with these data has maximal area if and only if all its sides opposite to the vertex O' are equal and their endpoints are at equal distances from O'. This is a planimetry problem, which can easily be solved by using the well-known maximal properties of isosceles triangles; we leave this to the reader.

Thus, the polyhedron P is isometric to the lateral surface of a right pyramid. Since we search for the maximum in the set of polyhedra with the number of vertices at most a given integer and with curvature at most a given ω, it is not known *a priori* if the curvature and the number of sides of the polyhedron P are less than these given values. However, it is easy to see that for a given perimeter, the area of the lateral surface as well as the number of sides of a pyramid increases with the increase of curvature. Consequently, the polyhedron must have the given curvature and number of sides.

The following theorem on arbitrary convex surfaces corresponds to this theorem on polyhedra.

Theorem 4. *The lateral surface of a right circular cone has maximal area among all convex surfaces of given perimeter and curvature $\omega < 2\pi$.*

The surface area of a right circular cone of curvature ω and perimeter l is equal to $l^2/(2(2\pi - \omega))$. Therefore, this result can also be expressed as follows.

The area S of each convex surface of perimeter l and curvature $\omega < 2\pi$ satisfies the inequality

$$S \leq \frac{l^2}{2(2\pi - \omega)},$$

with equality holding if and only if the surface is isometric to the lateral surface of a right circular cone.

Since $2\pi - \omega = \tau$ is the swerve of the boundary of a surface, this inequality can also be rewritten as

$$2\tau S \leq l^2.$$

The proof of this theorem is performed on using Theorem 3 in exactly the same way as Theorem 2 was proved on using Theorem 1.

Applying the same methods, we can also solve, for example, the following problem: among all convex surfaces bounded by shortest arcs of given length and having given curvature $\omega < 2\pi$, find a surface of maximal area. The answer reads: take a polygon P on the plane inscribed into a disk and whose sides are equal to the

© 2006 by Taylor & Francis Group, LLC

shortest arcs bounding the surfaces under consideration. If the center of the disk lies in P, then the desired surface is the lateral surface of the pyramid with base P and apex that projects to the center of the disk. If the center of the disk lies outside of P, then the following two cases are possible: (1) the central angle subtended by the side a facing the center is greater than $(2\pi - \omega)/2$; (2) this angle is no greater than $(2\pi - \omega)/2$. In the first case, the desired surface is also the lateral surface of the pyramid with base P and apex that projects to the center of the disk. In the second case, the desired surface is obtained by gluing the figure composed of two isosceles triangles with the base a and the angle $(2\pi - \omega)/2$ opposite the side a.

It is important here that the figure is bounded by shortest arcs; if we consider a surface that is bounded by geodesics of a given length, then the answer will change, since, first of all, the limit of geodesics can be not a geodesic but a quasigeodesic. For a surface of maximal area, one of the bounding geodesics can transform to a quasigeodesic.

Our gluing method also allows us to solve the extremal problems concerning not only area but other quantities related to the convex surfaces specified by some data. For example, in order to refine Lemma 1, we can easily prove the following theorem.

The diameter d of a convex surface of given perimeter l and curvature $\omega \leq \pi$ does not exceed $l/2$. The value $l/2$ is attained only in the limit when the surface degenerates into a segment of length $l/2$. If $\pi < \omega < 2\pi$, then the maximum of the diameter is $l/(2\sin(\omega/2))$, and this maximum is attained only at the surface composed of two right triangles with leg $l/2$ and opposite angle $\pi - (\omega/2)$; the gluing of the triangles is carried out along the hypotenuse and leg contiguous to this angle.

In closing, we state one more problem: among all closed convex surfaces of given diameter, find a surface of maximal area. Here, we speak about the intrinsic diameter, i.e., the maximal distance between points in the sense of the intrinsic metric of a surface. No solution to this problem seems to be known. There are some reasons to think that a solution is given by a surface which is a doubly-covered disk of given diameter.

© 2006 by Taylor & Francis Group, LLC

Chapter XI

THE ROLE OF SPECIFIC CURVATURE

1. Intrinsic Geometry of a Surface

Any domain G on a convex surface has some curvature $\omega(G)$ and area $S(G)$. The ratio $\omega(G)/S(G)$ is naturally called the specific curvature of the domain G; we will denote this ratio by $\kappa(G)$. The specific curvature at a point X or the *Gaussian curvature* at this point is the limit of the specific curvatures of the domains that shrink to the point X. Of course, this limit does not necessarily exist for each point of the surface in general. Moreover, we can impose various additional conditions on the domains under consideration; under some of these conditions, this limit can exist for a given point X, whereas there is no limit under other conditions. For example, we can assume that the ratio of the area of a domain to the area of some geodesic disk centered at the point X which includes this domain is always greater than some arbitrarily small positive number. Under this condition the limit of the specific curvature of a domain that shrinks to a point exists almost everywhere on each convex surface, i.e., the set of points of the surface at which there is no limit has measure (area) zero.[1] If there are no restrictions imposed on the admissible sequences of domains shrinking to a point, then this limit may fail to exist for all points of the surface. A surface with an everywhere dense set of conical points can serve as an example. Indeed, let X be a given point of such a surface, and let $X_1, X_2, \ldots$ be a sequence of conical points converging to X; the curvatures $\omega_1, \omega_2, \ldots$ of these points are positive. Surround each point X_n by a domain G_n of area $S_n < \omega_n/n$; then the ratio ω_n/S_n is greater than n and tends to infinity as $n \to \infty$. We can achieve the situation in which all domains G_n include the point X; to this end, it suffices to add a narrow band containing the point X with area less than ω_n/n to each G_n. Consequently, it is perfectly immaterial that we restrict exposition only to the domains that contain a given point. We can present examples of smooth convex surfaces for which the limit of the specific curvature of a domain shrinking to a point does not exist either for any point unless some extra conditions are imposed on admissible sequences of domains.

We will say that *a surface F has Gaussian curvature equal to K at a point X which is equal to K if the limit of the specific curvature of a domain tends to the*

[1] This is an immediate corollary of a general theorem on set functions. For each completely additive set function $f(M)$ defined, e.g., on a plane domain D for all Borel sets M, the limit $\lim_{M \to X}(f(M)/\mathrm{meas}M)$ exists almost everywhere whenever sets M shrink to a point X so that the ratio of meas$M$ to the area of the minimal disk centered at X that includes M remains greater than some positive number. See Ch.-J. de la Vallee-Poussin, Course of Analysis of Infinitesimals, Vol. II, Chap. III, Sec. 4.

© 2006 by Taylor & Francis Group, LLC

limit K whenever this domain shrinks to the point X.[2] We will reveal the consequences of the existence of Gaussian curvature at each point of a convex surface. The role of the Gaussian curvature in the geometry of regular surfaces is well known, and so we will address the question of what follows from the existence of Gaussian curvature without any additional regularity assumptions.

First of all, we mention the following assertion.

Lemma 1. *If the Gaussian curvature exists at each point of a domain U on a convex surface, then this curvature is a continuous function of the point on the domain U.*

Proof. Let the Gaussian curvature at the point X be equal to K, and let this curvature be equal to K_n at some points X_n converging to X. For each point X_n, there exists a domain G_n of specific curvature $\kappa(G_n)$ that differs arbitrarily little from K_n. Therefore, we can take domains G_n such that (1) $|\kappa(G_n) - K_n| < 1/n$ and (2) G_n lies in the disk of radius $1/n$ centered at X_n. Then the domains G_n shrink to the point X, and hence we have

$$\lim_{n\to\infty} \kappa(G_n) = K.$$

And since $|\kappa(G_n) - K_n| < 1/n$, we also have $\lim_{n\to\infty} K_n = K$; as required.[3]

Theorem 1. *If the specific curvature of every domain lying in some neighborhood of a point O on a convex surface does not exceed some positive number K then there exists $r_0 > 0$ such that it is possible to draw a shortest arc of length at least r_0 from the point O in each direction.*

Proof. Given some $r > 0$, draw the circle of radius r centered at the point O. We assume that r is so small that this circle is a closed curve (cf. Theorem 2 of Sec. 6 of Chapter IX). If we move a point X along our circle then the shortest arc OX will jump over those directions in which there is no shortest arc of length r. There arises a digon D at the place of a jump. This digon can be considered as a triangle such that the sum of its two sides is equal to the third. Therefore, the area of the plane triangle with sides of the same length is equal to zero. Hence, by the theorem which was proved in Sec. 1 of Chapter X, we have the following inequality for the area of the digon D:

$$S(D) < \frac{1}{2}\omega(D)d^2,$$

[2] It is easy to prove that for the existence of such a limit for any domains, it is sufficient that it exists for sequences of triangles.

[3] The following assertion holds; it is, in a sense, converse to the previous one. If the Gaussian curvature defined as the limit of specific curvatures of domains that satisfy special conditions (e.g., imposed on curvatures of disks centered at a given point) exists everywhere in the domain U and is continuous, then at each point of the domain U there is the Gaussian curvature in the sense of our definition.

Indeed, let $K(X)$ be the Gaussian curvature at the point X defined under additional conditions. The curvature of the domain G is equal to $\omega(G) = \iint_G K(X)dS$, and if the domain G shrinks to a point A, then the ratio $\omega(G)/S(G)$ tends to $K(A)$, since $K(X)$ is continuous. Consequently, the existence of the Gaussian curvature without conditions imposed on domains is equivalent to the continuity of the Gaussian curvature or to some conditions imposed on the domain.

© 2006 by Taylor & Francis Group, LLC

where d is the diameter of the digon D. Therefore,

$$\kappa(D) = \frac{\omega(D)}{S(D)} > \frac{2}{d^2},$$

and if the specific curvature $\kappa \le K$ for all domains, then also $\kappa(D) \le K$, and hence the digon D may appear only for $d > 2/\sqrt{K}$. But the diameters of the digons D also tend to zero as $r \to 0$. Therefore, there exists r_0 such that only one shortest arc goes from the center to each point of the circle of radius r_0; therefore, shortest arcs of length r_0 emanate from O in all directions.

The above theorem implies that in a neighborhood of the point O, in which the specific curvature is bounded, we can introduce geodesic polar coordinates. Namely, with each point X of the disk U of radius r_0 centered at the point O, we associate two numbers r and ϕ, the distance from O and the angle made by OX and some fixed shortest arc emanating from O which is measured in a given direction. Conversely, with every two numbers r and ϕ ($r < r_0, 0 \le \phi < 2\pi$), we associate the point X with these coordinates.

Let the coordinates u, v be introduced in a domain V on a surface; we say that in the domain V, *the metric can be given by the line element*

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2$$

if the length of each curve $u = u(t), v = v(t)$ $(0 \le t \le 1)$ traveling in V can be expressed by the integral

$$\int_0^1 \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\, dt$$

(provided that the derivatives $\dot{u}$ and $\dot{v}$ are continuous), while the distance between two points X and Y in V is equal to the greatest lower bound of these integrals that are taken over all curves connecting the points X and Y. We can prove that, in the disk U, the metric of a surface of bounded specific curvature can be given by the line element $ds^2 = dr^2 + Gd\phi^2$ in the coordinates r, ϕ. However, the coefficient G is not even a continuous function of r and ϕ in the general case. Therefore, this general case is not interesting, and we shall study the case in which not only the specific curvature of each domain is bounded, but also the Gaussian curvature exists at each point of the disk U. To this end, we first prove one lemma on isosceles triangles.

Let OAB be an isosceles triangle on a convex surface ($OA = OB$), and let X and Y be two points on its lateral sides which are equidistant from the vertex O. Assume that these points can always be connected by a shortest arc traveling in the triangle OAB. If these two points X and Y can be connected by several shortest arcs, then there is a "leftmost" arc among them, i.e., an arc such that in the triangle OXY bounded by this arc there is no shortest arc connecting the points X and Y except for this arc itself. The fact that such shortest arc exists is easily proved using the nonoverlapping condition for shortest arcs and from the fact that the limit of shortest arcs is a shortest arc. However, the existence of a leftmost shortest arc was

© 2006 by Taylor & Francis Group, LLC

already proved in Sec. 3 of Chapter VII. Also, it was proved there that if points X' and Y' converge to X and Y from the left, i.e., from the side of the point O, then the leftmost shortest arcs $X'Y'$ converge to the leftmost shortest arc XY. In what follows, we shall consider only the leftmost shortest arcs XY.

Lemma 2. *Let X and Y be two points inside the lateral sides of the triangle OAB which are equidistant from O. Put $OX = OR = r$ and $XY = z = z(r)$; let ϕ be the angle of the triangle OAB at the vertex O, and let $\omega(r)$ be the curvature of the interior of the triangle OXY, which is distinguished in OAB by the leftmost shortest arc XY. Then the left derivative dz/dr of the function $z(r)$ with respect to r exists and can be represented by the formula*

$$\frac{dz}{dr} = 2(1-\varepsilon)\sin\frac{\phi-\omega(r)}{2}, \quad \textit{where} \quad 0 \le \varepsilon \le \sin^2\frac{\omega}{4}.$$

Proof. Let two points X' and Y' on the sides OA and OB converge to the given points X and Y from the left. Then the shortest arcs $X'Y'$ converge to XY. Therefore, if ξ and η, ξ' and η' are angles at the vertices X and Y, X' and Y' in the triangles OXY and $OX'Y'$, then, by Theorem 5 of Sec. 4 of Chapter IV on convergence of angles, we have

$$\lim\xi' = \xi, \quad \lim\eta' = \eta. \tag{1}$$

On the other hand, if ξ_0 is the angle corresponding to ξ in the plane triangle with the same sides as those of the triangle XYX' then

$$\lim_{X'\to X}\xi_0 = \xi$$

(e.g., this follows from Theorem 5 of Sec. 4 of Chapter III on the existence of an angle in the strong sense)

If we put

$$XX' = \Delta x, \quad XY = z = z(x,y),$$
$$XY' = z' = z(x-\Delta x, y),$$

then

$$z'^2 = z^2 + \Delta x^2 - 2z'\Delta x\cos\xi_0$$

or

$$\cos\xi_0 = \frac{z-z'}{\Delta x}\cdot\frac{z+z'}{2z'} + \frac{\Delta x}{2x'}.$$

Passing to the limit as $\Delta x \to 0$ and taking into account that $\lim\xi_0 = \xi$, we infer

$$\cos\xi_0 = \lim_{\Delta x\to+0}\frac{z(x,y)-x(x-\Delta x,y)}{\Delta x} = \frac{\partial z}{\partial x}, \tag{2}$$

where $\partial z/\partial x$ is the left partial derivative. In exactly the same way, we obtain the following expression for the left partial derivative $\partial z/\partial y$ with respect to y:

$$\cos\eta = \frac{\partial z}{\partial y}.$$

© 2006 by Taylor & Francis Group, LLC

Formulas (1) imply that the partial derivatives $\partial z/\partial x$ and $\partial z/\partial y$ are left-continuous and so there exists the left derivative of $z(r,r)$ with respect to r,[4] which is equal to

$$\frac{dz(r,r)}{dr} = \left(\frac{\partial z}{\partial x}\right)_{x=y=z} + \left(\frac{\partial z}{\partial y}\right)_{x=y=z} = \cos\xi + \cos\eta.$$

But $z(r,r)$ is exactly the length $z(r)$ of the shortest arc of the points X and Y equidistant from O. Therefore,

$$\frac{dz(r)}{dr} = \cos\xi + \cos\eta = 2\cos\frac{\xi+\eta}{2}\cos\frac{xi-\eta}{2}. \tag{3}$$

The curvature of the interior of the triangle OXY is equal to

$$\omega(r) = \xi + \eta + \phi - \pi;$$

this implies

$$\frac{\xi+\eta}{2} = \frac{\pi}{2} - \frac{\phi-\omega(r)}{2},$$

so that

$$\cos\frac{\xi+\eta}{2} = \sin\frac{\phi-\omega(r)}{2}.$$

Therefore, (3) implies

$$\frac{dz(r)}{dr} = 2\cos\frac{\xi-\eta}{2}\sin\frac{\phi-\omega(r)}{2} = 2(1-a\sin^2\frac{\xi-\eta}{2})\sin\frac{\phi-\omega(r)}{2}. \tag{4}$$

The angles ξ and η of the triangle OXY differ from the angles ξ_0 and η_0 of the plane triangle with the same sides by less than $\omega(r)$; that is,

$$0 \le \xi - \xi_0 \le \omega(r), \quad 0 \le \eta - \eta_0 \le \omega(r). \tag{5}$$

Since the triangle OXY is isosceles, we have

$$|\xi - \eta| \le \omega(r),$$

so that

$$|\sin\frac{\xi-\eta}{4}| \le \sin\frac{\omega(r)}{4}. \tag{6}$$

Consequently, formula (4) just yields the expression for $dz(r)/dr$ which was claimed in the lemma.

Now let the Gaussian curvature exist at each point of a certain closed neighborhood V of a point O on a convex surface. By Lemma 1, this curvature is a continuous function of a point and hence it is bounded in the closed neighborhood V; therefore, by Theorem 1, we can circumscribe a disk U around the point O in which the polar coordinates r, ϕ with the origin at the point O can be introduced.

[4] $z(r,r) - z(r-\Delta r, r-\Delta r) = z(r,r) - z(r-\Delta r) + z(r-\Delta r, r) - z(r-\Delta r, r-\Delta r) = \int_{r-\Delta r}^{r} \frac{\partial z(x,r)}{\partial x}\,dx + \int_{r-\Delta r}^{r} \frac{\partial z(r-\Delta r, y)}{\partial \Delta y}\,dy$; since the derivatives are left continuous, this implies our assertion.

© 2006 by Taylor & Francis Group, LLC

Theorem 2. *Let the polar coordinates r, ϕ with the origin at the point O be introduced in the disk O on a convex surface, and let the Gaussian curvature $K(r, \phi)$ exist everywhere on the disk U. Then the metric on the disk U can be given by the line element*

$$ds^2 = dr^2 + G(r, \phi)d\phi^2; \tag{7}$$

moreover, the coefficient G is a continuous function of r and ϕ and has two continuous partial derivatives with respect to r, while the Gaussian curvature $K(r, \phi)$ is expressed through the coefficient G by the Gauss formula

$$K(r, \phi) = -\frac{1}{\sqrt{G(r, \phi)}} \frac{\partial^2 \sqrt{G(r, \phi)}}{\partial r^2}. \tag{8}$$

Proof. Let $z = z(r, \phi, \Delta\phi)$ stand for the distance between the points X and Y with the coordinates r, ϕ and $r, \phi + \Delta\phi$, and let $\omega = \omega(r, \phi, \Delta\phi)$ denote the curvature of the triangle OXY. If dz/dr is the left derivative of z with respect to r, then, according to Lemma 2,

$$\frac{dz}{dr} = 2(1 - \varepsilon) \sin \frac{\Delta\phi - \omega}{2},$$

where $|\varepsilon| \leq 2 \sin^2(\omega/4)$. Integrating this equation, we obtain

$$z(r, \phi, \Delta\phi) = 2 \int_0^r (1 - \varepsilon) \sin \frac{\Delta\phi - \omega}{2}\, dx. \tag{9}$$

Since $\omega(x, \phi, \Delta\phi)$ does not decrease as x increases, we have $|\varepsilon| \leq 2 \sin(\omega(r)/2)$ under the integral sign; since $\omega(r, \phi, \Delta\phi) \to 0$ as $\Delta\phi \to 0$, we see that ε tends uniformly to zero. Dividing Eq. (9) by $\Delta\phi$, we pass to the limit as $\Delta\phi \to 0$. Since $\varepsilon \to 0$ and $\lim_{\xi \to 0} \sin \xi/\xi = 1$, we obtain the following expressions for the upper and lower limits of the ratio $z/\delta\phi$:

$$\overline{B}(r, \phi) = \limsup_{\Delta\phi \to 0} \frac{z(r, \phi, \Delta\phi)}{\Delta\phi} = r - \liminf_{\Delta\phi \to 0} \int_0^r \frac{\omega(x, \phi, \Delta\phi)}{\Delta\phi}\, dx, \tag{10a}$$

$$\underline{B}(r, \phi) = \liminf_{\Delta\phi \to 0} \frac{z(r, \phi, \Delta\phi)}{\Delta\phi} = r - \limsup_{\Delta\phi \to 0} \int_0^r \frac{\omega(r, \phi, \Delta\phi)}{\Delta\phi}\, dx. \tag{10b}$$

These limits $\overline{B}(r, \phi)$ and $\underline{B}(r, \phi)$ are uniformly bounded, since, first, we have $\overline{B} \geq \underline{B} \leq 0$, and second, (9) implies

$$z(r, \phi, \Delta\phi) \leq 2r \sin \frac{\Delta\phi}{2} < r\Delta\phi;$$

therefore, $\underline{B} \leq \overline{B} \leq r$.

We now calculate the value $\omega(x, \phi, \Delta\phi)$. This is the curvature of the isosceles triangle OXY with lateral sides equal to x and angle $\Delta\phi$ at the vertex O. Take a series of points on the lateral sides of the triangle OXY which are equidistant from O, connect them pairwise, and partition the triangle OXY into small quadrangles (Fig. 92). The sides of these quadrangles are Δx and $z(x, \phi, \Delta\phi)$. Consider one quadrangle q in the set of these quadrangles. If the angle $\Delta\phi$ is sufficiently small

© 2006 by Taylor & Francis Group, LLC

then the angles of the quadrangle q are close to the right angle, and hence this triangle is partitioned into two "almost right" triangles t_1 and t_2. Since the curvatures of these triangles are small, the corresponding plane triangles t_1^0 and t_2^0 with the same sides are also "almost right" and, for their areas, we have

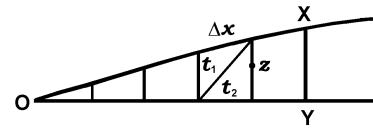


Fig. 92

$$S(t_1^0) = \frac{1}{2} z(x, \phi, \Delta\phi) \cdot \Delta x \sin\alpha_1,$$
$$S(t_2^0) = \frac{1}{2} z(x + \Delta x, \phi, \Delta\phi) \cdot \Delta x \sin\alpha_2, \quad (11)$$

where α_1 and α_2 tend to the right angle as $\Delta\phi \to 0$.

By the theorem of Sec. 1 of Chapter X the area of the quadrangle q is related to the sum of the areas of the triangles t_1^0 and t_2^0 by the inequality

$$0 \le S(q) - [S(t_1^0) + S(t_2^0)] \le \frac{1}{2}\omega(q)d^2,$$

where d is the diameter of the triangles t_1 and t_2. Introducing the specific curvature $\kappa(q) = \omega(q)/S(q)$ instead of $\omega(q)$, we can rewrite this inequality as

$$|S(q) - [S(t_1^0) + S(t_2^0]| \le \frac{1}{2}\kappa(q)S(q)d^2. \quad (12)$$

Since the Gaussian curvature exists everywhere in the domain under consideration, $\kappa(q)$ is bounded; moreover, $d \to 0$ as $\Delta\phi \to 0$ and $\Delta x \to 0$. Therefore, combining formulas (11) and (12), we readily see that

$$S(q) = \frac{1}{2}[z(x, \phi, \Delta\phi) + z(x + \Delta x, \phi, \Delta\phi)]\Delta x(1 + \varepsilon), \quad (13)$$

where ε is an infinitesimal relative to the diameter of the quadrangle q, i.e., relative to $\Delta\phi$ and Δx,

The curvature of the triangle OXY is equal to the sum of the curvatures of all quadrangles q, and hence

$$\omega(OXY) = \omega(x, \phi, \Delta\phi) = \sum \kappa(q)S(q). \quad (14)$$

The specific curvature $\kappa(q)$ is exactly the mean value of the Gaussian curvature in the quadrangle g; since the Gaussian curvature is continuous, its mean value in the quadrangle q is equal to its value K_q at one of the points of this quadrangle. If we make the quadrangle q narrower and narrower while keeping $\Delta\phi$ constant, then the values K_q converge to the values of the curvature at certain points of the context shortest arcs $X'Y'$ of the points equidistant from the center O on the given radii OX and OY. Denoting by $\overline{K}(y, \phi, \Delta\phi)$ this value of the curvature for the shortest arc $X'Y'$ such that $OX' = OY' = y$, we come in the limit l_o to the following formula instead of formula (14):

$$\omega(x, \phi, \Delta\phi) = \int_0^x \overline{K}(y, \phi, \Delta\phi) z(y, \phi, \Delta\phi)(1 + \varepsilon)\, dy, \quad (15)$$

© 2006 by Taylor & Francis Group, LLC

where ε is an infinitesimal relative to $\Delta\phi$.

If $\Delta\phi \to 0$ then the shortest arc $X'Y'$ shrinks to the point X', and the value of the Gaussian curvature $\overline{K}(y, \phi, \Delta\phi)$ at some point on $X'Y'$ converges to the Gaussian curvature at the point X', since the Gaussian curvature is continuous. Therefore, (15) implies

$$\limsup_{\Delta\phi\to 0} \frac{\omega(x,\phi,\Delta\phi)}{\Delta\phi} = \limsup_{\Delta\phi\to 0} \frac{1}{\Delta\phi} \int_0^x \overline{K} z(1+\varepsilon)\, dy$$

$$\leq \int_0^x K(y,\phi) \limsup_{\Delta\phi\to 0} \frac{z(y,\phi,\Delta\phi)}{\delta y}\, dy, \tag{16}$$

where $K(y, \phi)$ is the value of the Gaussian curvature at the point on the given radius OX whose distance from O is y. A similar formula holds for the lower limit.

Using formula (16) and a similar formula for the lower limit, we obtain the following relations instead of (10a) and (10b):

$$\overline{B}(r,\phi) \leq r - \int_0^r \left(\int_0^x K(x,y) \underline{B}(y,\phi)\, dy \right) dx, \tag{17a}$$

$$\underline{B}(r,\phi) \geq r - \int_0^r \left(\int_0^x K(x,y) \overline{B}(y,\phi)\, dy \right) dx. \tag{17b}$$

Therefore,

$$\overline{B} - \underline{B} \leq r - \int_0^r \left(\int_0^x K(\overline{B} - \underline{B})\, dy \right) dx. \tag{18}$$

This is true for all r; moreover, since always $\overline{B} - \underline{B} \geq 0$, this implies $\overline{B} = \underline{B}$.[5] By the definition of $\overline{B}$ and $\underline{B}$, this means that, for all r, there exists

$$B(r,\phi) = \lim_{\Delta\phi\to 0} \frac{z(r,\phi,\Delta\phi)}{\Delta\phi}. \tag{19}$$

Therefore, inequalities (17a) and (17b) reduce to the following integral equation for $B(r, \phi)$:

$$B(r,\phi) = r - \int_0^r \left(\int_0^x K(x,y) B(y,\phi)\, dy \right) dx. \tag{20}$$

[5] Let M and N be the maxima of $|\overline{B} - \underline{B}|$ in the interval $(0, r)$. Then for all $r' \leq r$, we have

$$\overline{B}(r') - \underline{B}(r') \leq MN\frac{r^2}{2}.$$

In particular, if $\overline{B}(r') - \underline{B}(r') = M$ then $M \leq MN(r^2/2)$; if $r^2 < 2/N$ then $M = 0$. Consequently, $\underline{B} = \overline{B}$ in the interval from $r = 0$ to $r = 2/\sqrt{\max K}$. But the equality $\underline{B} = \overline{B}$ extends further beyond this interval by the same argument.

© 2006 by Taylor & Francis Group, LLC

Since the right-hand side of this equation has the second derivative with respect to r, this is also true for $B(r,\phi)$, and hence Eq. (20) leads to the differential equation

$$\frac{\partial^2 B(r,\phi)}{\partial r^2} + K(r,\phi)B(r,\phi) = 0. \tag{21}$$

Since $K(r,\phi)$ depends continuously not only on r but also on ϕ, $B(r,\phi)$ is a continuous function of r and ϕ as well.

It remains only to prove that the metric on the disk U can be given by the line element

$$ds^2 = dr^2 + B^2(r,\phi)d\phi^2. \tag{22}$$

Then this makes it clear that Eq. (21) is the Gaussian expression (8) for the curvature $K(r,\phi)$.

To prove this, we consider some smooth curve $L : r = r(t), \phi = \phi(t)$ $(0 \le t \le 1)$. Take a point $A(r,\phi)$ on this curve and a point $X(r+\Delta r, \phi+\Delta\phi)$ that is infinitely close to it; then calculate the length of the shortest arc AX. To this end, we take one more point Y with coordinates $(r, \phi+\Delta\phi)$ and consider the triangle AXY (Fig. 93).

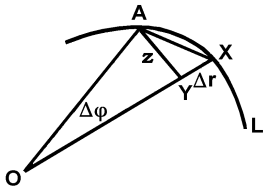


Fig. 93

As $\Delta\phi \to 0$, the angle at Y in this triangle tends to the right angle and the curvature of the triangle tends to zero. Therefore, the plane triangle with sides of the same length is close to the rectangular triangle, so that we have

$$AX^2 \cong AY^2 + XY^2$$

to within infinitesimals of higher order.[6] But $XY = \Delta r$, and AY is exactly $z(r,\phi,\Delta\phi)$; moreover, formula (19) implies

$$AY = z(r,\phi,\Delta\phi) \cong B(r,\phi)\Delta\phi$$

to within infinitesimals of higher order. Therefore, if we denote the length of the chord AX by Δs then, instead of (23), it is possible to write

$$\Delta s^2 \cong \Delta r^2 + B(r,\phi)\Delta\phi^2. \tag{24}$$

Since the length of the curve L is the limit of the sums of the length of the chords comprising an inscribed broken line, this obviously implies that the length of the

[6] Here and in what follows, care must be exercised in dealing with infinitesimals, since otherwise we run the risk of obtaining a ghost of the proof. However, it is easy to give a rigorous estimate for the neglected magnitudes. For example, denote by ξ the angle at the vertex A of the triangle AXY. Since the curve has direction at the point A, there exists $\lim \xi = \alpha$. Therefore, the limit of the corresponding angle ξ_0 in the plane triangle with the same sides is also equal to α, and hence

$$\lim \frac{AY}{AX} = \cos\alpha, \quad \lim \frac{XY}{AX} = \sin\alpha.$$

Squaring and adding, we have

$$\lim \frac{AY^2 + XY^2}{AX^2} = 1,$$

yielding (23).

© 2006 by Taylor & Francis Group, LLC

curve L is represented by the formula

$$s(L) = \int_L \sqrt{dr^2 + B^2 d\phi^2} = \int_0^1 \sqrt{\dot{r}^2 + B^2 \dot{\phi}^2}\, dt.$$

Consequently, the metric on the disk is in fact given by the line element $ds^2 = dr^2 + B^2 d\phi^2$, and so the proof of our theorem is complete.

This theorem shows that the only fact of existence of the Gaussian curvature ensures to a certain extent the possibility of applying the methods of classical differential geometry; however, this is true only to some extent since, for example, we did not prove at all that the coefficient $B(r, \phi)$ has at least the first derivative with respect to ϕ; in fact, this derivative may fail to exist anywhere.[7] The usual expression for the geodesic curvature of a curve involves this derivative. Therefore, it seems that this expression becomes inapplicable, and, for example, it is not possible to write out the differential equation for geodesics which is obtained by equating the geodesic curvature to zero. However, we can write the expression for the geodesic curvature in the shape that does not contain $\partial B/\partial \phi$ explicitly. Namely, we can prove the following theorem.

Theorem 3. *Assume that the conditions of Theorem* 2 *hold. Let* $\phi = \phi(r)$ *be the equation of a curve in the polar geodesic coordinates. If this curve has geodesic curvature* k *(the derivative of the swerve with respect to the arc length) at the point* A*, then*

$$k = \frac{(B\phi')' + [1 + (B\phi)'^2](1 + (\partial B)/\partial r)\phi'}{[1 + (B\phi')^2]^{3/2}}; \tag{25}$$

moreover, the derivative $(B\phi')' = d[(B(r, \phi(r))\phi(r)]/dr$ *exists.*

Proof. Let A be a given point, and let X be a variable point on the curve. Since shortest arcs emanate from each point in all directions by hypotheses, we can draw shortest arcs at the points A and X which touch our curve; if the point X is close to A then these tangents intersect each other (Fig. 94). If ψ is the angle between them and τ is the swerve of the arc AX, then the curvature ω of the domain bounded by this arc and the tangents is equal to $\psi - \tau$; since ω tends to zero in the same way as the area of this domain (i.e., faster than the square of the arc s), then $\lim_{X \to A} \tau/s = \lim_{X \to A} \tau/s$. Consequently, in this case, the geodesic curvature can be defined in the same way as for curves on the plane. If ξ and $\xi + \Delta\xi$ are the angles made by the radii OA and OX and the tangents at the points A and X and if $\Delta\phi$ is the angle made by OA and OB then, for the curvature of the quadrangle $OXZA$, we have

$$\Delta\omega = \xi + \Delta\xi + \Delta\phi + (\pi - \psi) + (\pi - \xi) - 2\pi = \Delta\phi + \Delta\xi - \psi.$$

[7] For example, let the curvature $K(r, \phi) = f(\phi)$ in the annulus $r_1 < r < r_2$, where $f(\phi)$ is a continuous positive function that has no derivative anywhere. Then, solving Eq. (21), we obtain $B(r, \phi) = C_1 \cos \sqrt{f(\phi)} \cdot r + C_2 \sqrt{f(\phi)} \cdot r$ in this annulus, so that $B(r, \phi)$ is now differentiable with respect to ϕ everywhere. It is easy to prove that the line element $ds^2 = dr^2 + B^2(r, \phi) d\phi^2$ with this B defines a metric of positive curvature on the annulus $r_1 < r < r_2$ (it suffices to consider $f_n(\phi) \to f(\phi)$) analytically.

© 2006 by Taylor & Francis Group, LLC

This implies

$$k = \lim_{X \to A} \frac{\psi}{s} = \frac{d\phi}{ds} + \frac{d\xi}{ds} - \frac{d\omega}{ds}; \tag{26}$$

moreover, if k, $d\phi/ds$, and $d\omega/ds$ exists, then so does $d\xi/ds$.

If K is the Gaussian curvature then, for $\Delta\omega$, we have

$$\Delta\omega = \iint\limits_{(OXZA)} K\, dS \cong \Delta\phi \int_0^r KB\, dr,$$

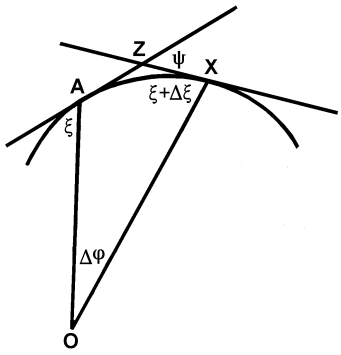


Fig. 94

where the last equation holds certainly to within infinitesimals of higher order. Also, from the Gauss equation $B'_{rr} + KB = 0$, we obtain

$$\Delta\omega \cong -\frac{\partial B}{\partial r}\Delta\phi.$$

Consequently, the derivative $d\omega/ds$ exists and is equal to

$$\frac{d\omega}{ds} = \frac{d\omega}{d\phi}\frac{d\phi}{dr}\frac{dr}{ds} = -\frac{\partial B}{\partial r}\phi' \frac{1}{[1 + (B\phi')^2]^{1/2}}. \tag{27}$$

For the angle ξ, it is easy to deduce the usual expression

$$\tan\xi = \frac{Bd\phi}{dr} = B\phi';$$

this implies

$$\frac{d\xi}{ds} = \frac{d\xi}{dr}\frac{dr}{ds} = \frac{(B\phi')'}{[1 + (B\phi')^2]^{3/2}}; \tag{28}$$

moreover $(B\phi')'$ exists since so does $d\xi/ds$.

Inserting (27) and (28) into (26), we obtain formula (25).

We can prove that *on a convex surface with the Gaussian curvature defined everywhere each line of zero geodesic curvature is a geodesic, i.e., a shortest arc on small segments.* As we know, the converse is always true. Therefore, equating relation (25) for the geodesic curvature to zero, we obtain a differential equation for geodesics. If we put $B\phi' = \vartheta$, this equation reduces to the system

$$\phi' = \frac{\vartheta}{B(r,\phi)}, \quad \vartheta' = -(1 + \vartheta^2)\left(1 + \frac{\partial B(r,\phi)}{dr}\right)\frac{\vartheta}{B(r,\phi)}.$$

In our arguments, we use the polar coordinates, but nothing prevents us from introducing other, so-called semigeodesic coordinates u, v in which tu is the arc length of a geodesic. Namely, we can take a shortest arc L in the disc U and draw all shortest arcs through the points of L orthogonal to L. It is easy to prove that this is possible on assuming the boundedness of the specific curvature; moreover, the shortest arcs drawn in such a way never meet in some neighborhood of the

© 2006 by Taylor & Francis Group, LLC

shortest arc L. In this neighborhood, we can introduce the distance from a point X to L as the coordinate u of the point X and take the arc length of the shortest arc L from some initial point to the foot of the geodesic perpendicular dropped from X to L as the coordinate v. In these coordinates, the metric on the neighborhood under consideration is given by the line element $ds^2 = du^2 + B^2dv^2$; moreover, $B(u, v)$ is continuous and $B_{uu} + KB = 0$, where K is the Gaussian curvature. The proof repeats the proof of Theorem 2. Instead of a shortest arc, we can take each curve with continuous geodesic curvature as the line L ($u = 0$), and the result will be the same.

2. Intrinsic Geometry of a Surface of Bounded Specific Curvature

We shall say that the specific curvature is $> K$ ($< K$, or $= K$) on a surface F, if this curvature is $> K$ ($< K$, or $= K$) for each domain on the surface F.

Theorem 1. *If the specific curvature is equal to K on a surface F, then each point of the surface F has a neighborhood isometric to a part of the sphere S_K of radius $1/\sqrt{K}$. If the surface F is closed, then F is isometric (and even equal) to this sphere in the large.*

Proof. If all specific curvatures are equal to K on the surface F, then, obviously, the Gaussian curvature exists everywhere and is equal to K. Therefore, according to what was proved in the previous section, we can introduce polar coordinates r, ϕ in a neighborhood of each point and define the metric of the surface by the linear element $ds^2 = dr^2 + B^2d\phi^2$, where

$$\frac{\partial^2 B}{\partial r^2} + KB = 0.$$

Since K is constant and $B(0, \phi) = 0, B'_r(0, \phi) = 1$ (this is obvious geometrically by itself and also is implied by formula (20) of the previous section), we have

$$B(r, \phi) = \frac{1}{\sqrt{K}} \sin \sqrt{K} r,$$

and the element ds turns out to be the line element of the sphere of radius $1/\sqrt{K}$. This proves the first part of the theorem.

Let the surface F be closed. Take a point O on this surface, and map isometrically a small disk centered at O onto the sphere S_K of radius $1/\sqrt{K}$. It is easy to show that this disk can always be extended so that the extended disk is any time isometrical to the sphere S_K, and we may proceed likewise until this disc covers the whole surface F and its image covers the whole sphere S_K. This proves that the sphere F is isometric to the sphere S_K. (The proof of the fact that the surface F itself is the sphere was sketched in Sec. 5 of Chapter VIII. However, this remark is immaterial if we speak only about intrinsic geometry.)

If the range of the specific curvature of a certain convex surface lies within some bounds K_1 and K_2, then the intrinsic geometry of this surface turns out to be, in a

© 2006 by Taylor & Francis Group, LLC

sense, intermediate between the intrinsic geometries of the spheres of radii $1/\sqrt{K_1}$ and $1/\sqrt{K_2}$. We present here a number of theorems revealing a more rigorous meaning of this assertion.

Theorem 2. *If the specific curvature of a convex surface F is bounded above then, as was proved in the previous section, in a neighborhood of its every point, we can introduce polar geodesic coordinates r, ϕ. Assume that the specific curvature is $\geq K_1$ and $\leq K_2$ on such a neighborhood U. Introduce the polar geodesic coordinates on the sphere S_{K_i} $(i = 1, 2)$, and to each point of the neighborhood U with coordinates r, ϕ, put in correspondence the point of the sphere S_{K_i} with the same coordinates. If, to a curve L on the surface F, this mapping puts in correspondence the curves L_1 and L_2 on the spheres S_{K_1} and S_{K_2}, then the lengths of these curves are connected by the inequalities*

$$s(L_1) \geq s(L) \geq s(L_2).$$

On the sphere of radius $1/\sqrt{K}$ each geodesic is a shortest arc on the segment of length $\leq \pi/\sqrt{K}$ and, in contrast, is not a shortest arc on the segment of length $> \pi/\sqrt{K}$. We have similar estimates for a geodesic on a convex surface, which depends on the boundaries of the specific curvature; these estimates were found by Bonnet for geodesics on regular surfaces;[8] the following theorem is an abstraction of these Bonnet estimates to the convex surfaces not subjected to any additional regularity conditions:

Theorem 3. *If the specific curvature is $\leq K$ in a neighborhood of a geodesic L on a convex surface, then each arc of the geodesic L of length at most $\pi/\sqrt{K}$ is the shortest path among all lines sufficiently close to it.*[9] *If in a neighborhood of L, the specific curvature is $\geq K$, then each arc of this geodesic of length at least $\pi/\sqrt{K}$ is not a shortest path as compared with arbitrarily close lines.*

We present two more theorems on triangles. In these theorems, T stands for a triangle on some closed surface, and T_K denotes the triangle on the sphere S_K of radius $1/\sqrt{K}$ which has the sides of the same length and is entirely included in the half-sphere. Moreover, we assume that the lengths of the sides of the triangle T are less than $\pi/\sqrt{K}$, so that this triangle always exists.

Theorem 4. *If the specific curvature is $\geq K$ $(\leq K)$ in the triangle T, then the angles of T are no less (greater) than the corresponding angles of the triangle T_K.*

Moreover, if the angle at least at one vertex A of the triangle $T = ABC$ is equal to the corresponding angle of the triangle T_K, then either the triangle T is isometric to the triangle T_K or we can draw a shortest arc in the triangle T between two other vertices B and C of this triangle so that it cuts out a triangle isomorphic to T_K from T. The latter case is possible only if the specific curvature is $\geq K$ on the triangle.

[8] See, e.g., W. Blaschke, Differential geometry, Sec. 100.

[9] It should be noted that if the specific curvature is $\leq K$ on a complete convex surface, then each segment of a geodesic of length at most $\pi/\sqrt{K}$ is a shortest path in general but not as compared with close lines.

© 2006 by Taylor & Francis Group, LLC

Theorem 5. *If the specific curvature is* $\geq K$ $(\leq K)$ *on the triangle* T*, then the area of the triangle* T *is no less (greater) than the area of the triangle* T_K*.*

We can give some estimates for the difference $S(T) - S(T_K)$ of the areas of the triangles T and T_K that are similar to the estimate obtained in Sec. 1 of Chapter X for the difference $S(T) - S(T_0)$:

If the specific curvature is $\geq K$ on T, then

$$0 \leq S(T) - S(T_K) \leq \frac{1}{2}[\omega(T) - \omega(T_K)]d^2, \tag{1}$$

if the specific curvature is $\leq K$ on T, then

$$0 \geq S(T) - S(T_K) \geq \frac{1}{2}[\omega(T) - \omega(T_K)]d^2. \tag{2}$$

Here, d is the diameter of the triangle T, while ω is its curvature, and, certainly, $\omega(T_K) = KS(T_K)$. Moreover, it turns out that if equality holds at least in one of the parts of the formulas (1) and (2), then the triangles T and T_K are isometric.

We mention one more theorem that can easily be deduced from Theorem 4 and, at the same time, obviously covers Theorem 4.

Let L and M be two shortest arcs emanating from a point O of some manifold R with metric ρ. Let X and Y be variable points on L and M, $OX = x$, $OY = y$, $XY = z$, and let $\gamma_K(x, y)$ be the angle opposite the side z in the triangles on the sphere S_K of radius $1/\sqrt{K}$ with sides x, y, and z. We will say that the metric ρ satisfies the K-convexity condition (convexity with respect to the metric of constant curvature K) or is K-convex if for all L and M, $\gamma_K(x, y)$ is a nonincreasing function in each interval $0 < x < x_0, 0 < y < y_0$ for which there exists a shortest arc XY. Also, we shall say that the metric satisfies the K-convexity condition or is K-concave if $\gamma_K(x, y)$ turns out to be a nondecreasing function of x and y on the same interval. For $K = 0$, the K-convexity condition reduces to the convexity condition fulfilled by the metric of each convex surface, as was proved in Chapter III. As regards K-convexity and K-concavity, we have the following theorem.

Theorem 6. *If the specific curvature is* $\geq K$ $(\leq K)$ *on a convex surface* F*, then the* K*-convexity (*K*-concavity) condition holds on* F*.*

Also, we note that Theorems 2, 4, and 6 yield conditions that are not only necessary but sufficient for the specific curvature of the surface F to satisfy the corresponding restrictions. Of course, in this case, the result of Theorem 2 must hold for a neighborhood of each point of the surface F, and the result of Theorem 4 must hold for all triangles on F. Probably, Theorem 5 also yields not only the necessary but also a sufficient condition for the corresponding restrictions on the specific curvature. However, we do not know how to prove this.

The lack of space does not allows us to present a complete proof of Theorems 2–6 which are formulated above. We give here only their fundamentals, and only Theorem 2 and 3 will be proved completely. We assume that the bounds of the specific curvature are positive since otherwise the sphere S_K reduces to the plane and our theorems reduce to the comparison of the metric of a convex surface with the metric of the plane; this was treated in the most part of our previous presentation.

© 2006 by Taylor & Francis Group, LLC

Lemma 1. *Let the specific curvature be $\geq K$ on a convex surface F. Then for each $K' < K$ $(K' > 0)$, there exists d such that for each triangle T on the surface F whose diameter is less than d, the sum of the angles of T is greater than the sum of angles of the triangle $T_{K'}$ with sides of the same length on the sphere $S_{K'}$ of radius $1/\sqrt{K'}$.*

Since the specific curvature is $\geq K$ on the surface F, we have

$$\omega(T) = \alpha + \beta + \gamma - \pi \geq KS(T) \tag{3}$$

for every triangle T with angles α, β, and γ whose area is $S(T)$.

Let T_0 and T_K be the triangles with sides of the same length on the plane and the sphere S_K, respectively. According to Lemma 1 in Sec. 1 of Chapter X,

$$S(T_0) \leq S(T);$$

and, at the same time,

$$S(T_0) \geq S(T_K) - \frac{1}{2}\omega(T_K)d^2 = S(T_k)(1 - \frac{1}{2}Kd^2).$$

By these two inequalities, (3) implies

$$\omega(T) \geq K(1 - \frac{1}{2}Kd^2)S(T_K). \tag{4}$$

Given $K' < K$, we choose d_0 such that

$$K' = K(1 - \frac{1}{2}Kd_0).$$

Then by (4), for each triangle of diameter $d \leq d_0$,

$$\omega(T) \geq KS(T_K). \tag{5}$$

It is easy to prove that when the curvature of the sphere decreases, the area of the spherical image of the triangle with sides of a given length decreases.[10] Therefore, $S(T_K) > S(T_{K'})$, and hence (5) implies

$$\omega(T) > K'S(T_{K'}) = \omega(T_{K'}),$$

i.e.,

$$\alpha + \beta + \gamma > \alpha_{K'} + \beta_{K'} + \gamma_{K'}. \tag{6}$$

where α, β, γ and $\alpha_{K'}, \beta_{K'}, \gamma_{K'}$ are the angles of the triangles T and $T_{K'}$. This proves the lemma.

A similar lemma for a surface on which the specific curvature $\leq K$ is obtained by the replacement of the "greater" sign by the "less" sign.

[10] Of course, we can omit the condition $S(T_K) > S(T_{K'})$ for $K > K'$. Indeed, since $|S(T_0) - S(T_K)| \leq \omega(T_K)d^2/2$, the areas $S(T_{K_1})$ and $S(T_{K_2})$ are close to each other for close values of K_1 and K_2; therefore, we can choose K'' so that $K > K'' > K'$ and $K'S(T_k) \geq K''S(T_{K''})$; then (5) implies

$$\omega(T) \geq K''S(T_{K''}) = \omega(T_{K''}).$$

© 2006 by Taylor & Francis Group, LLC

Lemma 2. *If inequality* (6) *holds for any triangle of diameter* $< d_0$ *on a convex surface, then each angle of a convex triangle* T *of diameter* $< d_0$ *is greater than the corresponding angle of the triangle* $T_{K'}$*; i.e.,* (6) *implies* $\alpha > \alpha_{K'}$ $\beta > \beta_{K'}$*, and* $\gamma > \gamma_{K'}$.

Let ABC be a convex triangle of diameter $< d_0$, and let $A_0B_0C_0$ be the triangle on the sphere $S_{K'}$ with sides of the same length. Assume the contrary, i.e., suppose that $\alpha \le \alpha_{K'}$ and we may prove that this is impossible.

If we may connect the vertices B and C of the triangle ABC by a shortest arc different from the side BC, then we take the leftmost of these shortest arcs, i.e., the arc cutting out a triangle that is free of such shortest arcs. Instead of the initial triangle ABC, we shall consider exactly this triangle ABC.

Since inequality (6) holds for the triangle ABC and $\alpha \le \alpha_{K'}$ by assumption, we have that either $\beta > \beta_{K'}$ or $\gamma > \gamma_{K'}$. Assume that $\beta > \beta_{K'}$.

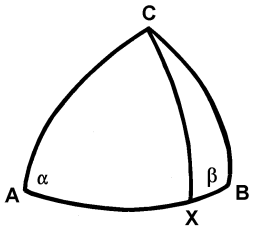


Fig. 95

Take a variable point X on the side AB (Fig. 95). Then by the theorem on the existence of the angle in the strong sense,[11]

$$\cos\beta = \lim_{X\to B} \frac{CB - CX}{BX}. \tag{7}$$

If X_0 is a variable point on the side A_0B_0 of the triangle $A_0B_0C_0$, then

$$\cos\beta_{K'} = \lim_{X_0\to B_0} \frac{C_0B_0 - C_0X_0}{B_0X_0}. \tag{8}$$

Since $\beta > \beta_{K'}$, we have $\cos\beta < \cos\beta_{K'}$; and so (7) and (8) imply that for sufficiently small BX and B_0X_0, we have

$$\frac{CB - CX}{BX} < \frac{C_0B_0 - C_0X_0}{B_0X_0}.$$

But, by the construction of the triangle $A_0B_0C_0$, we have $C_0B_0 = CB$, so that if

$$BX = B_0X_0, \tag{9}$$

then

$$CX > C_0X_0. \tag{10}$$

Thus, we can choose the point X so that relations (9) and (10) hold.

Consider the triangles ACX and $A_0C_0X_0$. The triangle ACX is convex since this triangle is cut out from the convex triangle ABC by the shortest arc CX; its diameter is also $< d_0$, and the angle at its vertex A is also α. The sides of the triangle $A_0C_0X_0$ are related to the sides of the triangle ACX by the relations

$$A_0C_0 = AC, \quad A_0X_0 = AX, \quad C_0X_0 < CX. \tag{11}$$

[11] If β_0 is the angle in the plane triangle with sides CB, CX, and BX, then $CX^2 = CB^2 + BX^2 - 2CB\cdot BX\cdot\cos\beta_0$; this implies $\cos\beta_0 = \frac{CB-CX}{BX}\cdot\frac{CB+CX}{2CB} + \frac{BX}{2CB}$; but by the theorem on the existence of the angle in the strong sense, $\lim_{X\to A}\cos\beta_0 = \cos\beta$. Since obviously $\frac{CB+CX}{2CB}\to 1$ and $\frac{BX}{2CB}\to 0$, we obtain (7).

© 2006 by Taylor & Francis Group, LLC

The angle at the vertex A_0 in the triangle $A_0C_0X_0$ is equal to $\alpha_{K'}$. If we construct the triangle $A_0C_0X_0^1$ with sides equal to AC, AX, and CX on the sphere $S_{K'}$ then, by (11), the angle $\alpha_{K'}^{(1)}$ at its vertex A_0 is greater than $\alpha_{K'}$; that is,

$$\alpha_{K'}^{(1)} > \alpha_{K'}. \tag{12}$$

Since $\alpha_{K'} \geq \alpha$ by assumption, we have $\alpha_{K'}^{(1)} > \alpha$. We put

$$\alpha_{K'}^{(1)} - \alpha = \varepsilon. \tag{13}$$

Since the triangle ACX is small, relation (6) holds for it, and, moreover, $\alpha_{K'}^{(1)} > \alpha$. Consequently, *we have the same situation for the triangle ACX as that for the triangle ABC, and the same argument can be applied to the former*, i.e., at least one of the angles $\angle C$ and $\angle X$ is greater than the corresponding angle $\angle C_0$ or $\angle X_0^1$ of the triangle $A_0C_0X_0^1$. If, for example, $\angle C > \angle C_0$ then, applying the above argument, we find a point Y on AC such that, in analogy with (12), $\alpha_{K'}^{(2)} > \alpha_{K'}^{(1)}$ for the triangle $A_0Y_0^1X_0^1$ (on the sphere $S_{K'}$) with sides equal to the sides of the triangle AYX. Therefore, (13) implies

$$\alpha_{K'}^{(2)} - \alpha > \varepsilon. \tag{14}$$

This argument can be continued by taking newer and newer points X and Y on the sides AB and AC which approach A closer and closer.

We now consider the set M of all triangles AXY that have the following properties.

1. The vertices X and Y lie on the sides AB and AC of the triangle ABC, and the triangle AXY is a convex triangle lying in ABC.

2. If $AX = x$, $AY = y$, and $\alpha_{K'}(x, y)$ is the angle at the vertex A of the triangle $A_0X_0Y_0$ on the sphere $S_{K'}$ which has the same sides as AXY, then $\alpha_{K'}(x, y) \geq \alpha + \varepsilon$.

The above argument shows that the set M is not empty and has the following properties.

(I) For each triangle AXY in M, there exists a triangle $AX'Y'$ in M such that $AX' \leq AX$, $AY' \leq AY$, and the "less" sign takes place in at least one of these relations.

Let us prove that the set M also has the following property.

(II) If a sequence of triangles AX^nY^n in M converges and, moreover, if the points X^n and Y^n do not converge to A, then the limit triangle AXY also belongs to the set M.

Indeed, the triangle AXY is convex as the limit of convex triangles;[12] the vertices X and Y lie on AB and AC. At the same time, the lengths of the sides of the triangles AX^nY^n converge to the lengths of the sides of the triangle AXY,

[12]The fact that the limit of convex figures is a convex figure immediately follows from the fact that the limit of shortest arcs is a shortest arc.

© 2006 by Taylor & Francis Group, LLC

and so the angles $\alpha_{K'}(x^n, y^n)$ of the corresponding spherical triangles converge to $\alpha_{K'}(x, y)$. Since $\alpha_{K'}(x^n, y^n) \geq \alpha+\varepsilon$, we also have $\alpha_{K'}(x, y) \geq \alpha+\varepsilon$. Consequently, the triangle AXY belongs to M.

Properties (I) and (II) of the set M imply the following property.

(III) There are triangles in the set M such that at least one of the vertices X and Y is arbitrarily close to A. In other words, the set of the numbers $x = AX$ and $y = AY$ contains numbers that are arbitrarily close to zero.

Indeed, (I) implies that for each pair x, y there is a pair x', y' such that $x'y' < xy$; (II) implies that if $x^n \to x$ and $y^n \to y$ and if the pairs $x^n y^n$ belong to the set under consideration then the pair x, y also belongs to this set whenever $xy \neq 0$. This makes it clear that the greatest lower bound of the product xy equals zero, i.e., there are numbers among the numbers x and y that are arbitrarily close to zero.

Recall now the definition of angle in the strong sense and apply it to the angle α. According to this definition, α is the limit of the angles $\gamma(x, y)$ in the plane triangles $A_0X_0Y_0$ provided that $xy \to 0$, where $x = AX = A_0X_0$ and $y = AY = A_0Y_0$. But the angles of this plane triangle slightly differ from the angles of the spherical triangle with the same sides, and so

$$\alpha = \lim_{xy\to 0} \alpha_{K'}(x, y). \tag{15}$$

But, for the triangles AXY in the set M, we have $\alpha_{K'}(x, y) \geq \alpha + \varepsilon$, and so

$$\lim_{xy\to 0} \alpha_{K'}(x, y) \geq \alpha + \varepsilon,$$

which contradicts (15). This contradiction shows that the above-assumed inequality $\alpha_{K'} \geq \alpha$ is impossible; consequently, $\alpha_{K'} < \alpha$, as required.

An analogous lemma holds for a surface on which the specific curvature $\leq K < K'$; this lemma is obtained from the above lemma by substituting the "less" sign for the "greater" sign.

Lemma 3. *Let the specific curvature is $\geq K > K' (\leq K < K')$ on a convex surface F. Let OAB be an isosceles convex geodesic triangle on F whose lateral sides OA and OB are sufficiently close to each other. Then the base AB of this triangle is less (greater) than the base of the isosceles triangle on the sphere $S_{K'}$ which has the same sides making the same angle.*

Proof. Let the specific curvature $\geq K > K'$ on the surface F. Then, by Lemma 2, the angles of each sufficiently small convex triangle are greater than the angles of the corresponding triangle on the sphere $S_{K'}$. Let the sides OA and OB of the triangle OAB be so close to each other that this triangle can be partitioned into some triangles T^i with vertices on OA and OB that are sufficiently small for Lemma 2 to be applicable to them (Fig. 96). Construct triangles $T^i_{K'}$ on the sphere $S_{K'}$ whose sides are equal to the sides of the triangles T^i. Adjoining these triangles to each other in the same way as the triangles T^i adjoint to each other, we obtain a

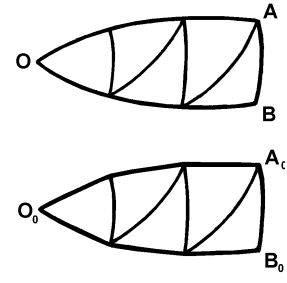


Fig. 96

© 2006 by Taylor & Francis Group, LLC

polygon Q. Lemma 2 implies that all angles of this polygon do not exceed π; therefore, this polygon is convex. Straightening the broken lines O_0A_0 and O_0B_0 that correspond to the sides OA and OB, we transform the polygon Q into a triangle $T_{K'}$ whose sides are equal to the sides of the triangle OAB. It is obvious and easy to prove that the angle at the vertex O_0 decreases in straightening. But this angle is already less than the angle $\angle O$ in the triangle OAB, and hence the corresponding angle in the triangle $T_{K'}$ is also less than the angle $\angle O$.

Therefore, if we construct an isosceles triangle on the sphere $S_{K'}$ whose lateral sides are equal to OA and OB and the angle between them is equal to the angle $\angle O$, then the base of this triangle is greater than AB, as claimed.

In the case where the specific curvature is $< K'$ on the surface F, we proceed in a similar way; but here the polygon Q is not convex; and when we transform this polygon into the triangle $T_{K'}$, the angle at the vertex $O_)$ increases.

We now prove Theorem 3 which was formulated above:

If the specific curvature is $\leq K$ in a neighborhood of a geodesic L on a convex surface, then each arc of the geodesic L of length at most $\pi/\sqrt{K}$ is the shortest line as compared with all sufficiently close lines. If the specific curvature is $\geq K$ in a neighborhood of L, then each arc of this geodesic of length at least $\pi/\sqrt{K}$ is not shortest.

Let the specific curvature be $\leq K$ in a neighborhood of the geodesic L. Then, as was proved in Sec. 1, we can draw shortest arcs in each direction from each point of this neighborhood. Therefore, we can draw a geodesic of each length (at least up to the boundary of the domain under consideration) from each point of the geodesic L, say, from its endpoint O. In some sector, these geodesics do not meet, and each of them is a shortest line in this sector.

Take some $K > K'$. Draw two geodesics M_1 and M_2 from O which are so close to each other that Lemma 3 holds for each isosceles triangle OXY with lateral sides lying on these geodesics. (It is easy to see that if M_1 and M_2 are sufficiently close and if two points X and Y lie inside them then the triangle OXY is convex.) We put $OX = OY = r$; let ϕ be the angle between M_1 and M_2, and let $z = z(r, \phi) = XY$. Let $z_{K'}(r, \phi)$ be the base of the isosceles triangle on the sphere $S_{K'}$ with the lateral side r and the same angle ϕ. Then, by Lemma 3,

$$z(r, \phi) > z_{K'}(r, \phi).$$

The value $z_{K'}(r, \phi)$ remains greater than zero on the sphere $S_{K'}$ while the inequality $r < \pi/\sqrt{K'}$ holds, and so $z(r, \phi)$ is surely greater than zero for $r < \pi/\sqrt{K'}$. Consequently, any two close geodesics M_1 and M_2 do not intersect on segments of length $< \pi/\sqrt{K'}$. If such a segment AB of the geodesic L is included into a narrow geodesic triangle then AB is a shortest line in this triangle, since the shortest line in this triangle should be a geodesic while a geodesic that connects the points A and B is unique by what we have proved.

Consequently, each segment of the geodesic L of length $< \pi/\sqrt{K'}$ is the shortest as compared with sufficiently close curves. Since K' can be taken arbitrarily close to K, the same is true for segments of length $< \pi/\sqrt{K}$.

Now let the specific curvature be $\geq K$ in a neighborhood U of the geodesic L. Let the segment OA of this geodesic be a shortest line as compared with all sufficiently

© 2006 by Taylor & Francis Group, LLC

close curves. Take a point B inside this segment and construct a sequence of points B_n that lie outside L and converge to the point B. For sufficiently large n, there exist lines OB_n shortest in U. Otherwise, we can construct a sequence of lines $\widehat{OB}_n$ whose lengths converge to the length of OB such that the limit line does not coincide with the segment OA of the geodesic L in the closed neighborhood U, which would clearly contradict the nonoverlapping condition for shortest arcs.

Thus, we can draw two close shortest lines OB and OB_n in U. Let ϕ be the angle between them, and let $z(r,\phi)$ be the distance between their points X and Y that cut out equal segments $OX = OY - r$. Then, according to Lemma 3,

$$z(r,\phi) \le z_{K'}(r,\phi),$$

where $K' < K$ and $z_{K'}$ have the same meaning as above. But $z_{K'}(r,\phi)$ assumes zero value at $r = \pi/\sqrt{K'}$, and hence $z(r,\phi)$ also equals zero. This means that the shortest arcs OB and OB_n are of length $\le \pi/\sqrt{K'}$; otherwise, they would intersect and could not be shortest arcs. Consequently, on a segment of length $> \pi/\sqrt{K'}$, the geodesic L cannot be a shortest line. Since K' can be taken arbitrarily close to K, the same is true for segments of length $\pi/\sqrt{K}$; which completes the proof of the theorem.

We now prove Theorem 2:

Assume that the polar coordinates are introduced in a domain U on a convex surface where the specific curvature is $\ge K_1$ and $\le K_2$. On the sphere S_{K_i} $(i = 1, 2)$, we introduce the polar coordinates, and to each point of the domain U, we put in correspondence the point of the sphere S_{K_i} with the same coordinates. If to a curve L in the domain U, this mapping puts in correspondence the curves L_1 and L_2 on the spheres S_{K_1} and S_{K_2}, then the lengths of these curves are related by the inequalities

$$s(L_1) \ge s(L) \ge s(L_2).$$

Let O be the origin of the polar coordinate system in the domain U. Let X and Y be two points of the domain U with coordinates r, ϕ and $r+\Delta r, \phi+\Delta\phi$, respectively, and let $\Delta r \ge 0$. Take a point Z on the radius OA so that $OZ = OX = r$. Put

$$XY^2 = (ZY^2 + XZ^2)(1 + \xi). \tag{16}$$

Then ξ becomes infinitesimal together with XY, since when the triangle XYZ is small, its curvature is small and this triangle is close to a right triangle on the plane. Since the specific curvature is bounded, ξ is uniformly infinitesimal independently of the coordinates r, ϕ of the point X, depending only on Δr and $\Delta\phi$.[13] Take $K_1' < K_1$, and let X_0, Y_0, and Z_0 be the points on the sphere $S_{K_1'}$ with the same coordinates as X, Y, and Z. Then, in exactly the same way, we have

$$X_0Y_0^2 = (Z_0Y_0^2 + X_0Z_0^2)(1 + \eta). \tag{17}$$

But

$$ZY = \Delta r = Z_0Y_0$$

[13]The sharp estimate of ξ is easy in terms of the curvature of the triangles XYZ and OXZ and, eventually, in terms of XY and a bound for the specific curvature K.

© 2006 by Taylor & Francis Group, LLC

and, by Lemma 3,

$$XZ < X_0Z_0.$$

Therefore, formulas (16) and (17) imply

$$XY < X_0Y_0(1+\varepsilon), \tag{18}$$

where ε becomes infinitesimal together with Δr and $\Delta\phi$.

Now, if we take a broken line L in the domain U partitioning this line into small segments and using formula (18) when, passing to indefinitely fine partitions, we obtain

$$s(L) \le s(L_{K_1'}).$$

Since K_1' can be taken to be arbitrarily close to K_1, in much the same way as above, we have

$$s(L) \le s(L_{K_1}).$$

Finally, approximating each curve L by broken lines, we obtain exactly the same result for L.

Of course, the inequality $s(L) \ge s(L_{K_2})$ is proved in exactly the same way.

Using the so-obtained result, it is easy to prove Theorem 4 in the following weakened form.

If the specific curvature is $\ge K_1$ and $\le K_2$ on a surface F, then the angles of any sufficiently small triangle T on F are bounded by the corresponding angles of the triangles T_{K_1} and T_{K_2} on the spheres S_{K_1} and S_{K_2}.

Let the triangle $T = ABC$ be so small that each of its vertices can be taken as the origin of the geodesic polar coordinates defined on the disk of radius $< \pi/\sqrt{T}_2$ which includes the triangle T. Then, taking, for example, the vertex A as the center of such coordinates, we can map the triangle T onto the sphere S_{K_2} by sending points with equal polar coordinates to each other. Under this mapping (Fig. 97), the sides AB and AC of the triangle T transform into the arcs A_0B_0 and A_0C_0 of great circles which make the same angle

$$\angle A_0 = \angle A,$$

and the side BC passes to some curve B_0C_0 whose length is no greater than the length of BC according to Theorem 2. Therefore, the shortest arc B_0C_0 is also not longer than BC; that is,

$$B_0C_0 \le BC.$$

Thus, the elements of the spherical triangle $A_0B_0C_0$ are related to the elements of the triangle ABC as follows

$$A_0B_0 = AB, \quad A_0C_0 = AC, \quad B_0C_0 \le BC, \quad \angle A_0 = \angle A. \tag{19}$$

Now, if we construct the spherical triangle $T_{K_2} = A_1B_1C_1$ with sides equal to the sides of the triangle ABC then, by the first three relations (19), we have

$$A_1B_1 = A_0B_0, \quad A_1C_1 = A_0C_o, \quad B_1C_1 \ge B_0C_0.$$

This implies

$$\angle A_1 \ge \angle A_0 = \angle A,$$

i.e., the angle of the triangle T is no greater than the corresponding angle of the triangle T_{K_2}, as required.

© 2006 by Taylor & Francis Group, LLC

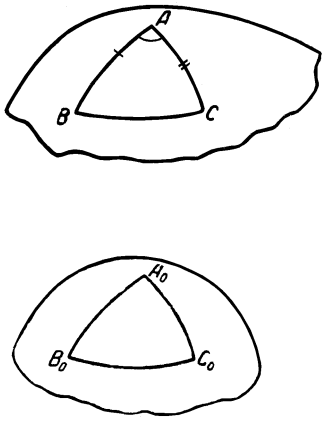


Fig. 97

The reverse inequality between the angles of the triangles T and T_{K_1} is proved in a similar way. We take the triangle $A_0B_0C_0$ on the sphere S_{K_1} in such a way that $A_0B_0 = AB, A_0C_0 = AC$, and $\angle A_0 = \angle A$; then we map this triangle onto a given surface in such a way that A_0B_0 and A_0C_0 become AB and AC. Then the side B_0C_0 transforms to the curve $\overset{\frown}{BC}$ which, according to the just proved theorem, is not longer than B_0C_0. Since obviously $BC \leq \overset{\frown}{BC}$, we also have that $BC \leq B_0C_0$. Comparing now the triangle $A_0B_0C_0$ with the triangle T_{K_1} having the same sides as the triangle $T = ABC$, we arrive at the fact that $\angle A_1 \leq \angle A_0 = \angle A$, i.e., the angle of the triangle T is no greater than the corresponding angle of the triangle T_{K_1}.

This result is weaker than Theorem 2 formulated in the beginning of this section in the sense that we do not assume in this theorem that the triangle T can be covered by a domain with polar coordinates. There can be no such domains if the specific curvature is only bounded below but is not bounded above. Therefore, the complete proof of Theorem 2 requires essentially different arguments.

Let the specific curvature be $\geq K$ on the triangle T; take $K' < K$ and partition the triangle T into so small triangles T^i that Lemma 2 is applicable to them, i.e., the angles of these triangles are greater than the angles of the corresponding triangles $T^i_{K'}$ on the sphere $S_{K'}$. Replacing each triangle T^i by the triangle $T^i_{K'}$, we obtain a development that is composed of spherical triangles. If we indefinitely refine the partition of the triangle T then the metric of this development converges to the metric of the triangle T. In essence, all arguments we have used in considering "planar" polyhedral metrics can be applied to such developments composed of spherical triangles or, say, to "spherical" polyhedral metrics. Approximating a given metric by such spherical metrics, we can prove Theorems 4, 5, and 6 by the same methods as the theorems on the convexity condition of Secs. 2–4 of Chapter III and the theorems on area of Sec. 1 of Chapter X, respectively. We leave these proofs to the reader.

3. Shape of a Convex Surface Depending on Its Curvature

The points of a convex surface can be of the following three types: (1) conical points at which the tangent cone has complete angle $< 2\pi$; (2) "edge" points at which the tangent cone is a dihedral angle (but not a plane); (3) "smooth" points at which the tangent cone turns out to be a plane. We have proved in Sec. 2 of Chapter X that almost all points of a convex surface are smooth, i.e., the set of those points at which there is no tangent plane is of measure zero. The fact that the curvature of a conical point is > 0 and the curvature of the whole surface is finite implies that there are no more than countably many conical points.

© 2006 by Taylor & Francis Group, LLC

Busemann and Feller have proved the following remarkable theorem: *every convex surface is twice differentiable almost everywhere*,[14] so that at all points, except possibly a set of measure zero, there exist curvatures of normal cross-sections which satisfy the well-known Euler theorem. These points admit the Dupin indicatrix of the ordinary shape, i.e., it is an ellipse, or a pair of parallel lines, or it "lies at infinity," the case in which the curvatures of normal cross-sections are equal to zero. At these points, the Meusnier, Rodrigues, etc. theorems hold along with the Euler theorem.

The problem consists in studying the extent to which the singularities of a convex surface depend on its intrinsic metric and what conditions on the intrinsic metric of a convex surface guarantee smoothness or twice differentiability. In this direction, we now have only the following results.

Theorem 1. *If the specific curvature of each sufficiently small domain containing a point A on a convex surface does not exceed some constant number, then A is a smooth point or a rectilinear edge of the surface passes through this point* (*the point A is not an endpoint of this edge*).[15]

A rectilinear edge on a convex surface always has an endpoint, and it can have no endpoints on an infinite complete convex surface only in the case where this surface is a closed or open cylinder; in the latter case, the surface is isometric to the plane. Therefore, Theorem 1 implies the following assertion.

Theorem 2. *If the specific curvature is bounded on a closed or infinite complete convex surface whose metric is not Euclidean everywhere, then this surface is smooth.*

If a rectilinear edge of a convex surface has endpoints on the boundary of this surface or at infinity, the surface can be unbended along this edge, so that this edge obliterates. Therefore, Theorem 1 implies the following theorem.

Theorem 3. *If the specific curvature is bounded on a convex surface F then there exists a smooth convex surface isometric to F. In other words, a metric of positive curvature such that its specific curvature is bounded for all domains is realizable by a smooth convex surface* (*whenever this metric is realizable by a convex surface in principle*).

It is possible in Theorem 1 that some rectilinear edge passes through the point A. A plane that is flexed along a line yields such an example. The following theorem holds for a rectilinear edge.

[14] That is, if we take rectangular coordinates x, y, and z such that the surface (or its part) is represented by the equation $z = f(x, y)$, then the function $f(x, y)$ has second differential at all (x, y), possibly except for a set of measure zero. See H. Busemann und W. Feller, Krümmungseigenschaften konvexer Flächen, Acta Mat., Vol. 66 (1935), pp. 1–47.

Some improvement of the results by Busemann and Feller is given in the paper by A. D. Alexandrov "Almost everywhere existence of the second differential of a convex function and some related properties of convex surfaces," Uchenye Zapiski Leningradskogo Gos. Universiteta, Seriya Matem. Nauk, No. 6 (1939).

[15] See A. D. Alexandrov, Smoothness of a convex surface of bounded Gaussian curvature, Doklady Akademii Nauk SSSR, Vol. 36, No. 7 (1942), pp. 211–216.

© 2006 by Taylor & Francis Group, LLC

Theorem 4. *If a line segment (whether an edge or not) passes through a point A of a convex surface F then there are arbitrarily small domains on F that have arbitrarily small specific curvature.*

Theorems 1 and 4 imply the following assertion.

Theorem 5. *If the specific curvature changes within positive bounds common for all domains on a convex surface F, then the surface F is smooth and contains no line segments.*

These results allow us to translate the proved general theorems into the usual language of differential geometry. The realization problem in differential geometry is formulated as follows. We introduce coordinates u, v in a neighborhood of each point of a manifold R and define the quadratic form[16]

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2. \tag{1}$$

The problem is to find a surface Φ and a homeomorphic mapping of the manifold R onto this surface so that the line element of the surface coincides with the line element ds^2 given on R. That is, if a surface is represented by the equation $\underline{x} = \underline{x}(u, v)$ in a neighborhood of an arbitrary point,[17] then

$$\underline{x}_u^2 = E, \quad \underline{x}_u\underline{x}_v = F, \quad \underline{x}_v^2 = G.$$

The statement of the problem implies immediately that this problem makes sense if and only if the function $\underline{x}(u, v)$ is differentiable, i.e., if the surface Φ is smooth; we can say that in differential geometry, the talk is about the realization of a given line element rather than the realization of a metric.

Since a metric given by a line element with a continuous nonnegative Gaussian curvature is a particular case of the general metric of positive curvature, the above Theorem 3 implies that such a metric is realizable by a smooth convex surface (if it is realizable by a convex surface in principle; for example, it is in any case realizable locally).

Consequently, this metric is realizable in the sense of the differential-geometric statement of the realization problem. Thus, our general realization theorems in the language of differential geometry yield the following results:

1. A line element with nonnegative curvature is locally realizable by a convex surface;

[16] If two neighborhoods U and U' meet, then the coordinates u', v' in their common part are expressed through u, v so that the Jacobian $D(u', v')/D(u, v)$ and u', v' are thrice differentiable with respect to u, v. When passing from the coordinates u, v to u', v', the coefficients must be transformed so that the values of the quadratic form ds^2 do not change. (Thrice differentiability is necessary since the second derivatives of E, F, and G enter the expression of the Gaussian curvature, while E, F, and G themselves are expressed through E', F', and G' and the first derivatives of some coordinates with respect to the others.) The coefficients E, F, and G are assumed to be twice continuously differentiable functions of the coordinates u, v.

[17] The coordinates u, v are translated into the surface by the mapping of the manifold R onto this surface.

© 2006 by Taylor & Francis Group, LLC

2. The same line element on the sphere is realizable by a closed convex surface and, finally,

3. the same line element on the plane, defining a complete metric on this plane, is realizable by an infinite complete convex surface.

By the generalized Gauss theorem, on the surface that realizes the line element under consideration, the Gaussian curvature determined from it coincides with the Gaussian curvature defined as the limit of the ratio of the area of the spherical image of a domain to the area of the domain. However, the Gaussian curvature can differ from the product of principal curvatures because of a simple reason that its existence does not necessarily imply the existence of the principal curvatures. At all points where the surface is twice differentiable (according to the Busemann–Feller theorem, this holds almost everywhere), the Gaussian curvature is equal to the product of the principal curvatures. But there can be points at which the Gaussian curvature exists, but no normal section has curvature. The surface given by the equation

$$2z = ax^2 + \frac{1}{a}y^2,$$

where

$$a = 1 + \frac{1}{2}\sin\log\log z,$$

provides such an example.

On this surface, the Gaussian curvature exists and is continuous in a neighborhood of the point $x = y = 0$, while, at the point $x = y = 0$, it is equal to 1 but no normal sections have curvature. The surface with the equation $z^2 = |x|^7 + |y|^3$ provides another example. At the point $x = y = 0$, one of the principal curvatures of this surface vanishes, while the other assumes the infinite value; the Gaussian curvature, in turn, is continuous and is equal to zero.[18]

Consequently, the existence and continuity of the Gaussian curvature do not ensure the twice differentiability of a surface at all.[19]

In closing, we pose a few more problems.

Conical points are precisely those points of a surface at which the curvature is > 0. Is it possible to give an intrinsic characterization of other types of points of a convex surface, e.g., of isolated "edge" points or points at which the surface has an elliptic Dupin indicatrix?

[18] We can exhibit examples in which the Dupin indicatrix is a line (i.e., one principal curvature vanishes, while the other is infinite) and the Gaussian curvature is positive and continuous.

[19] This question arises not only in connection with the realization theorem but also in connection with a number of other existence theorems for convex surfaces, for example, in connection with the celebrated Minkowski theorem on existence of a closed convex surface with Gaussian curvature given as a function of the normal or in connection with another theorem on existence of a convex surface with a given curvature which was stated at the end of Sec. 4 of Chapter V. For these theorems, see the following papers by A.D. Alexandrov: (1) Almost everywhere existence of the second differential of a convex function and some related properties of convex surfaces, Uchenye Zapiski LGU, Seriya Matem. Nauk, No. 6 (1939), pp. 1–35; (2) To the theory of mixed volumes of convex bodies, III, Matem. Sbornik, Vol. 3, No. 1 (1938), pp. 27–44; (3) Existence and uniqueness of a convex surface with given integral curvature, Doklady Akademii Nauk SSSR, Vol. 35, No. 5 (1942), pp. 143–147.

© 2006 by Taylor & Francis Group, LLC

In differential geometry, the relation of the intrinsic metric and exterior shape of a surface is expressed by the well-known Gauss–Codazzi formulas.[20] These formulas yield the necessary and sufficient conditions on a quadratic differential form which ensure the existence of a surface whose second quadratic form is the given form provided that the line element is also given. These conditions are necessary and sufficient and make sense only for a form whose coefficients are differentiable since the first derivatives of the coefficients of the second form enter the Gauss–Codazzi formulas.

It is well known that from the first and second forms given, a surface is determined uniquely up to motion and reflection. This is true for surfaces that are at least thrice differentiable, i.e., for those surface that can be represented by the equation $z = f(x, y)$ in a neighborhood of its every point (for an appropriate choice of the Cartesian coordinates x, y, and z) where the function $f(x, y)$ has continuous derivatives up to the third order. The Codazzi formulas make no sense for the surfaces that do not satisfy this condition; for the surfaces that are not twice differentiable, even the second quadratic form makes no sense at all points. However, we can pose the problem of searching for some replacement of the second quadratic form and the Gauss–Codazzi formulas for all convex surfaces. Is it possible to replace the second form by some integral concept that makes sense for all convex surfaces in much the same way as we have replaced the first form by the integral concept of distance on a surface?

The corresponding replacement of the Gauss formula is already given by the generalized Gauss theorem which is the integral expression of this formula. In this connection, we mention the following approach to the solution of the posed problem which was suggested by L. Pikus.

The third quadratic form of a surface is the expression for the square of the length element of the spherical image of the surface. Namely, if $\underline{n}$ is the unit normal vector to the surface with some polar coordinates u, v, then the third quadratic form is

$$d\underline{n}^2 = \underline{n}_u^2 + 2\underline{n}_u\underline{n}_v du dv + \underline{n}_v^2 dv^2.$$

It is known that the second form of a surface can be expressed through the first and third. Therefore, instead of the second form, we can consider the third form of the surface and deduce formulas for this form which replace the Gauss–Codazzi formulas. It turns out that the first and third forms define a surface uniquely up to motion and reflection, except for the case in which the mean curvature vanishes (i.e., except for minimal surfaces).

The spherical image is defined for each convex surface, and instead of the length element of the spherical image, we can consider distances between its points, i.e., merely the angles between normals to supporting planes of the surface. Since infinitely many supporting planes may pass through some points of a surface, there is no one-to-one correspondence between the points of the surface and normals. Therefore, it might be reasonable to consider a surface as a set of "elements" consisting of a point and a supporting plane at this point. There are two definite distances between these "elements;" these are the distance between points as measured on the surface and the angle between the normals to supporting planes; the first replaces

[20] See any course of differential geometry, e.g., W. Blaschke. Sec. 58.

© 2006 by Taylor & Francis Group, LLC

the first form while the second replaces the third form of the surface. It is almost obvious for polyhedra that each polyhedron is determined uniquely from these distances between all "elements." Pogorelov has proved a similar fact for an arbitrary convex surface; namely, he proved that if two convex surfaces admit an isometric mapping onto each other which preserves the angles between the normals at the corresponding points, then these surfaces are equal.[21]

The question of connections between the intrinsic metric of a surface and its spherical image which may replace the Codazzi formulas remains open.

[21]This result was communicated to me by A. V. Pogorelov in 1946 and has not been published yet.

© 2006 by Taylor & Francis Group, LLC

Chapter XII

GENERALIZATION

1. Convex Surfaces in Spaces of Constant Curvature

1. Along with Euclidean space, there exist other spaces in which the free movability of a solid body takes place; these are spaces of constant curvature. From the point of view of topological structure, the spherical spaces and Lobachevskii space are the simplest among them. Each of them is characterized by the value of its curvature K, i.e., by the ratio of the excess of the triangle $\alpha + \beta + \gamma - \pi$ to its area; we have $K > 0$ for spherical spaces and $K < 0$ for Lobachevskii space. Euclidean space can be considered as the limit case of Lobachevskii space with $K = 0$. The spherical space of curvature K can be realized by the three-dimensional sphere of radius $1/\sqrt{K}$ in four-dimensional Euclidean space.

In the sequel, R_K stands for the spherical space of curvature K in the case $K > 0$; in the case $K < 0$, it stands for Lobachevskii space of curvature K, and in the case $K = 0$, R_K denotes Euclidean space.

There exist lines and planes in these spaces; therefore, a line on the three-dimensional sphere is a great circle, and a plane is a two-dimensional sphere that is obtained in the diametrical section of the three-dimensional sphere.

If by a segment AB we mean a shortest line in space, which connects the point A and B, i.e., a segment of a line, then it is possible to introduce the concept of a convex body as a closed set that has interior points and enjoys a property that a segment connecting its every two points is contained in it. Now there appears the concept of a convex surface as a domain on the boundary of a convex body; then it is possible to pose the problem of studying the intrinsic geometry of the convex surfaces in spherical spaces and Lobachevskii space. It turns out that this study can be performed exactly by the same methods that we have applied to convex surfaces of Euclidean space; moreover, the results turn out to be a natural generalization of the results obtained in the case of Euclidean space. The matter is that in our arguments it is possible to avoid the parallel axiom; this guarantees that these arguments can be extended to Lobachevskii space. Passing to a spherical space certainly requires changing other axioms, since lines are closed in this space and each two of them intersect; but this turns out to be perfectly immaterial, since it is easy to prove that each convex surface on the sphere is a convex body in spherical space (except for the whole sphere) and always lies in the hemisphere; therefore, when studying convex surfaces, it is sufficient to restrict consideration to the limits of one hemisphere.

2. The study of the intrinsic geometry of convex surfaces in the space R_K of constant curvature K must start with proving the theorem on the convergence of

© 2006 by Taylor & Francis Group, LLC

metrics for convergent convex surfaces in order to make it possible to apply the method of polyhedral approximation.

The proof of this theorem in the case of Euclidean space R_0, which was given in Sec. 1 of Chapter III, uses essentially the homothety transformation and, therefore, cannot be extended to R_K with $K \neq 0$. However, it is possible to give the proof using no homothety transformation; such a proof in the case of Lobachevskii space was given in the author's paper "Complete convex surfaces in Lobachevskii space", Izvestiya Akademii Nauk SSSR, Seriya Matem., Vol. 9, No. 2 (1945), pp. 113–120.

When the theorem on the convergence of metrics is proved, we pass to the proof of the fact that *the metric of a convex surface satisfies the convexity condition*; i.e., if X and Y are variable points on shortest arcs L and M that emanate from a common point O, and $OX = x$, $OY = y$, $XY = z$, then the angle $\gamma(x, y)$ opposite to the side z in the plane triangle with sides x, y, and z, is a nonincreasing function of x and y. Here, we speak about the plane in the given space whose convex surfaces are under study. From the point of view of an intrinsic metric, this plane is a surface of constant curvature K, and, following the terminology introduced in Sec. 2 of Chapter XI, we can speak about the K-convexity condition.

The proof of the fact that the K-convexity condition holds on a convex surface is performed in exactly the same way as it was done in the case $K = 0$ in Secs. 2 and 3 of Chapter III; of course, in this case, we shall consider polyhedra (or the polyhedral metric) that are composed of plane triangles in R_K, i.e., by triangles cut out from a surface of constant curvature K; in other aspects, the whole proof is repeated word by word.

3. The corollaries of the convexity condition in Sec. 4 of Chapter III also hold, and their proof remains the same. In particular, it proves that *the angles between sides of a triangle on a convex surface in R_K are no less than the corresponding angles of the triangle with sides of the same length on the plane in R_K, i.e., on a surface of constant curvature K.* Hence, in a spherical space, the sum of the angles between the sides of a triangle on a convex surface is always greater than π. Further, as in Sec. 4 of Chapter III, the K-convexity condition implies that there exists the angle in the strong sense between two shortest arcs on a convex surface in R_K that emanate from a common point. Therefore, each convex surface in spherical space has a metric of positive curvature. At the same time, it is obvious that in spherical space, each convex body is bounded, and, therefore, any convex surface in it is a part of a closed convex surface. We have already proved in Chapter VII that each metric space that is homeomorphic to the sphere and has the metric of positive curvature is isometric to the convex surface in Euclidean space. Consequently, *each convex surface in the spherical space R_K is isometric to a convex surface in Euclidean space*; *namely, as it can be shown, to a surface of specific curvature $\geq K$.* Thus, from the viewpoint of an intrinsic metric, convex surfaces of spherical spaces give nothing new as compared with convex surfaces of Euclidean space, and, therefore, there is no necessity to study them separately. Also, it can be proved that every convex surface of Euclidean space on which the specific curvature is $\geq K$ and whose each two points can be connected by a shortest arc is isometric to a convex surface in the spherical space of curvature K. However, if not every two points can be connected by a shortest arc on a convex surface of Euclidean space, then this surface cannot be isometric to a convex surface in the spherical space R_K, although the specific

© 2006 by Taylor & Francis Group, LLC

curvature is no less than K on it. (A convex surface of revolution of constant positive curvature K and the length greater than $2\pi/\sqrt{K}$ is an example.)

4. The matter is quite different for convex surfaces in Lobachevskii space. For example, a plane of Lobachevskii space is certainly a convex surface, but it has constant negative curvature and, hence, is never isometric to any convex surface of Euclidean space. Therefore, polyhedra in Lobachevskii space are also not isometric to convex surfaces of Euclidean space. This shows that from the viewpoint of the intrinsic metric, the convex surfaces of Lobachevskii space yield something essentially new. Moreover, it turns out that the number of topologically different types of complete convex surfaces in Lobachevskii space is infinite rather than finite in Euclidean space; namely, the following theorem holds.

A convex surface in Lobachevskii space is homeomorphic to a domain on the sphere, and for each domain on the sphere, there exists a complete convex surface homeomorphic to it in Lobachevskii space.

To prove this assertion, we use the interpretation of Lobachevskii space; in the latter it is represented as the interior of a certain ball E in Euclidean space and the lines of this space are represented by segments of Euclidean lines.[1] Then each segment in the Lobachevskii sense is depicted by a segment in the ball E; and, conversely, any segment of the ball E depicts a segment of the Lobachevskii line. Therefore, a convex body of Lobachevskii space is depicted by an Euclidean convex body such that points not belonging to the interior of the ball E are deleted from it. Conversely, each of these Euclidean convex bodies is a convex body in Lobachevskii space. Thus, we obtain a clear visual idea of a convex body and, hence, that of a convex surface in Lobachevskii space.

Let H be a Euclidean convex body lying in the ball E; moreover, assume that points common with the surface S of the ball E are deleted from it. The projection of the boundary of the ball H to the sphere S from any point lying inside H is a homeomorphism. Therefore, the boundary of the body H is homeomorphic to an open set on the sphere. This implies that each convex surface of Lobachevskii space is homeomorphic to a domain on the sphere.

We now take an open set G on the sphere S such that its complement $S-G$ does not lie in one plane (this can always be attained by a topological transformation of the set G whenever $G \neq S$). The convex hull of the set $S - G$ becomes a convex body in Lobachevskii space if we exclude this set itself from the convex hull. Projecting its boundary on the sphere S, we see that it is homeomorphic to the set G. Consequently, for each open set G on the sphere, there exists a convex body in Lobachevskii space whose boundary is homeomorphic to this set. (For the case $G = S$, this is self-evident; each convex body has boundary homeomorphic to the sphere.) The set G is not necessary connected, but the connectedness requirement includes the definition of surface, and if G is connected, i.e., is a domain, then our construction yields a complete convex surface that is homeomorphic to it.

5. After we have proved that the metric of convex surfaces in the space of constant curvature K satisfies the K-convexity condition and derived corollaries analogous to Theorems 1–5 in Sec. 4 of Chapter III from this condition, further elaboration of the intrinsic geometry of convex surfaces in R_K proceeds along the

[1] This is the well-known projective interpretation of Lobachevskii space.

© 2006 by Taylor & Francis Group, LLC

same lines as for the convex surfaces in Euclidean space R_0. First of all, all the results concerning angles between shortest arcs and angles of sectors in Chapter IV are extended to a convex surface in R_K without changes. Some distinction appears in the item (Secs. 5 and 6 of Chapter IV) where we have obtained the spatial interpretation of the angle between shortest arcs by reducing it to the angle between tangents to them, which is measured on the tangent cone. The matter is that we have obtained the tangent cone as a result of the infinite homothetic dilation of a surface; however, in the space R_K, for $K \neq 0$, homothety transformations are impossible; therefore, it is necessary to use another definition of a tangent cone as a cone formed by limits of rays going from a given point O to the variable point X of the surface on condition that X tends to O.

Herewith, the proof of the fact that the metric of the tangent cone approximates the metric of a surface in an infinitely small neighborhood of the point O must be carried out in some other fashion. Such a proof is available, but we do not present it here. As for the proof of the existence of the half-tangent to a shortest arc on a convex surface, then the Liberman method used for this in Sec. 6 of Chapter IV is extended to any R_K almost literally.[2] Thus, all the results concerning angles between the shortest arcs may be extended to the case of every R_K literally keeping the same formulations.

6. Next comes the theory of curvature developed in Chapter V.

The definition of the curvature of an open triangle, an open shortest arc, and a point given in Sec. 1 of Chapter V, and also all the conclusions of this section are literally repeated with only one essential distinction. Namely, in the case $K < 0$, i.e., in Lobachevskii space, the sum of the angles of a triangle on a convex surface can be less than π, and, therefore, the curvature of a triangle (the sum of its angles minus π) can be negative. If in Chapter V we would have proceeded in a purely intrinsic way when expounding the theory of curvature, then, when passing to any R_K, nothing would have changed, and we would have arrived at the fact that the curvature is not only an additive but completely additive set function on a convex surface in each R_K. In exactly the same way, we would have arrived at the extension of the curvature to all Borel sets on R_K. Finally, the role of curvature as the measure of "non-Euclidicity" is certainly preserved in the case of convex surfaces in each R_K.

However, the generalized Gauss theorem, according to which the curvature is equal to the area of the spherical image, cannot be literally extended to each R_K by the simple reason that for $K \neq 0$, in R_K, we cannot use the usual parallel translation of a vector, and, therefore, the concept of the spherical image itself obliterates. In order to generalize the Gauss theorem to each R_K, first, it is necessary to introduce the concept of the area of a domain on a convex surface.

However, the theory of the area of a domain on a convex surface developed in Secs. 1 and 2 of Chapter X can be extended to any R_K without change. The distinction consists in the fact that it is more convenient to use not Euclidean plane triangles but the triangles that are flat in a given R_K (i.e., the triangles that are cut out from a surface of constant curvature K). However, this does not change the result, since the ratio of the areas of triangles from R_K and R_0 that have the same

[2] Applying this method, we construct a cylinder with a shortest arc as the directrix and with parallel generators. However, instead of a cylinder, one can consider a cone; this leads to the same result.

© 2006 by Taylor & Francis Group, LLC

sides tends to 1 as the maximal side of these triangles tends to zero. Therefore, the concept of the area of a domain on convex surfaces in an arbitrary R_K can be studied in the same way as in the case of R_0.[3] In particular, e.g., the areas $S(T)$ and $S(T_K)$ of some triangles with equal sides on a convex surface and in the plane are connected by the relation

$$0 \le S(T) - S(T_K) \le \frac{1}{2}[\omega(T) - KS(T_K)]d^2.$$

We now define the "extrinsic" curvature of a set on a convex surface in R_K. Let M be some set on a convex surface F in R, and let O be some point in R_K. At each point X of the set M, we draw all supporting planes to the surface F; let n be the outer normals to these planes. To each normal n, we put in correspondence the ray n' that emanates from the point O and makes the angle with the segment OX which complements the angle between n and OX up to 180°. (We can say that we translate the normal n along the segment OX in the parallel way in the sense of Levi–Civita.) All rays n' form some solid cone. The value of the solid angle at the vertex of this cone is denoted by $\psi_F(M, O)$.[4] In the case of Euclidean space, this value does not depend on the choice of the point O and is nothing else but the area of the spherical image of the set M; however, in the case of an arbitrary R_K, the value $\psi_F(M, O)$ depends not only on the set M but also on the choice of the point O; therefore, it cannot be taken as the measure of the "extrinsic" curvature of M. We define the "extrinsic" curvature in the following way. Partition the set M into sets M_i and take a point O_i in each of them. Define the sum $\sum_i \psi_F(M_i, O_i)$. It turns out that if the sets M_i are taken smaller and smaller in such a way that the maximal diameter of them tends to zero, then these sums converge to a limit. We accept this limit as the extrinsic curvature $\psi_F(M)$ of the set M on the surface F. If the surface F is regular, then $\psi_F(M)$ is the integral of the product of the principal curvatures over the area of the set M, i.e.,

$$\psi_F = \iint\limits_M \frac{1}{R_1 R_2} dS.$$

The generalized Gauss theorem in R_K is formulated as follows: *let* M *be a set on a convex surface* F *in* R_K *having the (intrinsic) curvature* $\omega(M)$*, extrinsic curvature* $\psi(M)$*, and area* $\sigma(M)$*. Then the following equation holds:*

$$\psi(M) = \omega(M) - K\sigma(M).$$

Consequently, the magnitude $\psi(M)$ has an intrinsic-geometric meaning; it plays the role of the measure of deviation of the metric of a given surface from the metric of the plane in R_K, i.e., from the metric of the surface of constant curvature K.

[3] Isoperimetric problems of Sec. 3 of Chapter X can also be posed in every R_K. But some essential changes are in order here; e.g., in Lobachevskii space, the requirement that the curvature of a surface is less than 2π turns out to be insufficient for the surface of maximum area for a given perimeter to exist.

[4] This $\psi_F(M, O)$ is defined if the set M is Borel. We restrict ourselves to the consideration of Borel sets.

© 2006 by Taylor & Francis Group, LLC

7. After the theories of angle, curvature, and area are available, we are left with the theory of direction and band of a curve among the general questions of intrinsic geometry. All definitions and results of Secs. 1 and 2 of Chapter IX remain valid for surfaces in any R_K. In exactly the same way, the theorem remaining valid is the one on the assertion that shortest arcs emanate from each point in almost all directions (proved in Sec. 5 of Chapter V).

Thus, *all* basic concepts and facts of intrinsic geometry of convex surfaces of Euclidean space are mostly word by word generalized to convex surfaces in every R_K. More specific questions concerning, e.g., circles, curves bounding convex domains, etc. cam also be generalized *mutatis mutandis* in any R_K. The conditions that characterize the metric of convex surfaces in R_K and the related realizability theorems connected with them will be considered in the next section.

2. Realization Theorems in Spaces of Constant Curvature

The theorem on the existence of a polyhedron with a given development proved in Chapter VI is literally extended to any R_K if by plane polygons we mean the polygons that are planar in R_K or, in other words, polygons that are cut out from a surface of constant curvature K.

We can glue a closed convex polyhedron from each development composed of plane polygons in R_K if this development is homeomorphic to the sphere and the sums of the angles meeting at each of its vertices do not exceed 2π. Here we include the doubly covered convex polygons plane (in R_K) in the set of convex polyhedra.

The proof of this theorem can be almost literally performed as was done in Chapter VI for the Euclidean case. In our proof, we use the parallel axiom but in implicit form when proving the possibility of a continuous passage from a given metric (development) to a realizable item.[5] However, it is easy to prove this in a manner that will suit any R_K.

The conditions characterizing the intrinsic metric of an arbitrary convex surface in R_K can be formulated by analogy with the conditions characterizing the intrinsic metric of convex surfaces in R_0. We will say that some *manifold has a metric of curvature $\geq K$ if this metric is intrinsic, and for each sufficiently small triangle the sum of lower angles between its sides is no less than the sum of the angles of the triangle with sides of the same length on a surface of constant curvature K; that is,*

$$\alpha + \beta + \gamma \geq \alpha_K + \beta_K + \gamma_K.$$

Here, the lower angle is understood in the sense we have defined in Sec. 9 of Chapter I and understood it as in Chapter VII. If $K = 0$, then $\alpha_K + \beta_K + \gamma_K = \pi$, and the "metric of curvature $\geq K$" turns out to be a metric of "positive" curvature in the sense of Sec. 9 of Chapter I. In the previous section, we mentioned that on convex surfaces in R_K, the angles between shortest arcs do exist in the strong sense and the angles between the sides of each triangle are no less than the angles of the plane (in R_K) triangle with sides of the same length, i.e., $\alpha \geq \alpha_K, \beta \geq \beta_K, \gamma \geq \gamma_K$.

[5] Namely, we use the construction of a plane triangle with a given base and adjacent angles. In the case $K \geq 0$, the condition that the sum of these angles is $< \pi$ is sufficient for the existence of such a triangle, but for $K < 0$ it is no longer sufficient.

© 2006 by Taylor & Francis Group, LLC

Therefore, we obviously have $\alpha + \beta + \gamma \geq \alpha_K + \beta_K + \gamma_K$, so that the *metric of every convex surface in R_K is of curvature $\geq K$.*

Similar to the theorem of Chapter VII, the following general theorem holds for any R_K.

A metric of curvature $\geq K$ on the sphere is realizable in R_K by a closed convex surface. (Here and in what follows, we again include doubly covered convex domains on the plane in R_K to convex surfaces.)

The proof of this theorem is obtained by literally repeating its proof in the case $K = 0$, which was given in Chapter VII; it suffices only to compare triangles from a given manifold with triangles that are plane in R_K and, in closing, to appeal the existence theorem of a polyhedron with a given development in the space R_K under consideration.

Further, we have the following local realization theorem.

In a manifold with a metric of curvature $\geq K$, each point has a neighborhood isometric to a convex surface in R_K.

The proof of this theorem literally repeats its proof in the case $K = 0$, which was given in Sec. 1 of Chapter VIII.

The gluing theorem which was proved in Sec. 1 of Chapter VIII (as also the gluing theorem whose proof was sketched in Sec. 13 of Chapter IX) is literally extended to the metric of curvature $\geq K$.

The application of this theorem in Sec. 3 of Chapter VIII leads to the realization theorem of a complete metric of positive curvature given on the plane. In the case $K > 0$, there is no analog of this theorem, since in the spherical space, any complete convex surface is closed; but first of all, since a complete metric of curvature $\geq K > 0$ can be given only on the sphere and the projective plane.[6]

The matter is quite different in the case $K < 0$, i.e., in the case of Lobachevskii space. As we have shown in Sec. 1, in this space there are complete convex surfaces that are homeomorphic to any domain on the sphere. All them have a complete metric of curvature $\geq K$. The following theorem serves as the realization theorem yielding the characterization of the metric of all such surfaces.

In the case $K < 0$, the complete metric of curvature $\geq K$ given on an arbitrary domain on the sphere is realizable by a complete convex surface in Lobachevskii space of curvature K.

The proof of this theorem is performed by using the gluing theorem in much the same way as in the proof of the realizability of a complete metric of curvature ≥ 0 given on the plane in Sec. 3 of Chapter VIII. The distinction consists only in the fact that if a domain G of the definition of the complete metric is not simply connected, then the enlarging polygons P_n that eventually cover the whole domain G can be not simply connected. Each polygon P_n bounds in G several domains that are infinite in the sense of the metric on G. Given a certain number $l > 0$, consider all those polygons containing P_n whose every connected component of the boundary is distant from P_n by no more than l. The completeness of the metric implies that there is a polygon Q_n of minimal perimeter among these polygons. It

[6] A metric of curvature $\geq K > 0$ is obviously a metric of positive curvature with specific curvature $\geq K > 0$ everywhere, but in the plane, such a metric would yield the infinite complete curvature; this contradicts the fact that the complete curvature of a metric of positive curvature on the plane is always $\leq 2\pi$.

© 2006 by Taylor & Francis Group, LLC

is easy to prove that on each connected component of the boundary of the polygon Q_n all angles but, possibly, one do not exceed π. (This exceptional angle can be at the point such that the distance from this point to P_n is exactly equal to l). Therefore, we can glue a manifold homeomorphic to the sphere and obtain the metric of curvature $\geq K$ from the polygon Q_n. In other respects, the arguments of Sec. 3 in Chapter VIII remain the same. By the way, we see that in Lobachevskii space the realization theorem of a complete metric becomes reacher in content, and, according to this, the gluing theorem has reacher applications.

Finally, we can prove the following general realization theorem.

If a metric of curvature $\geq K$ is given on a domain on the sphere in such a way that every two points can be connected by a shortest arc, then this metric can be realized by a convex surface in R_K.

For the case $K = 0$, this theorem was addressed though not proved in Sec. 2 of Chapter VIII. Then in Sec. 4 of Chapter IX we obtained additional results regarding the convex surfaces "convex in themselves," i.e., those whose each two points are connected by a shortest arc. If we ignore the closed surfaces and lateral surfaces of cylinders, then finite convex surfaces "convex in themselves" are isometric to "caps" in R_0. An analogous theorem holds in spherical space. But in Lobachevskii space the matter is more complicated, since here convex surfaces "convex in themselves" can be homeomorphic to any domain on the sphere. We do not know if it is possible to shape each of these surfaces canonically, e.g., in some particular sense so that, similar to the case of a "cap," all bounding lines are planar. Probably, using the gluing theorem, this interesting question can be solved rather easily.

Now consider the connection of the general metric of curvature $\geq K$ with a metric given by the line element of a certain Gaussian curvature. It is known that if the Gaussian curvature is $\geq K$ everywhere on a regular surface, then the angles of each triangle on such a surface are no less than the angles of the triangle with sides of the same length on a surface of constant curvature K. For small triangles, this result was in essence proved by Gauss. It implies that a regular surface with Gaussian curvature $\geq K$ has a metric of curvature $\geq K$. Since the Gaussian curvature is undoubtedly bounded from below on every regular finite surface, i.e., is no less than some K, the intrinsic geometry of a surface with the metric of curvature $\geq K$ thus embraces in particular the intrinsic geometry of all bounded regular surfaces. Each of these metrics, at least locally, is realizable by a convex surface in corresponding Lobachevskii space; therefore, we can say that the *intrinsic geometry of convex surfaces in Lobachevskii space embraces the intrinsic geometry of small domains of all regular surfaces in general. Consequently, passage to convex surfaces in Lobachevskii space allows us to overview the whole of the intrinsic differential geometry "in the small."*

Further, the Gaussian curvature is bounded from below on a regular surface homeomorphic to the sphere. Therefore, each of these surfaces is isometric to a closed convex surface in Lobachevskii space (whose curvature is no greater than the minimum of the Gaussian curvature on this surface). Consequently, from the viewpoint of intrinsic geometry, all regular surfaces homeomorphic to the sphere belong to the set of convex surfaces of Lobachevskii space.

© 2006 by Taylor & Francis Group, LLC

In Sec. 1 of Chapter XI, we have shown that if the Gaussian curvature exists at each point on a convex surface of Euclidean space, then the metric of this surface can be given by a line element in the geodesic polar coordinates. Moreover, all basic formulas of intrinsic differential geometry are valid.

Exactly the same result holds for convex surfaces in any R_K, and it is deduced in the same way. This allows us to give an abstract definition of a metric given by a line element with arbitrary Gaussian curvature. Namely, we can prove the following theorem.

Let an intrinsic metric be given in a two-dimensional manifold R that satisfies the following condition.

Let α, β, and γ be lower angles of triangle T in R, and let σ_0 be the area of the Euclidean triangle with sides of the same length. Put

$$\frac{\alpha+\beta+\gamma-\pi}{\sigma_0}=\overline{\kappa}(T).$$

The condition reads: whenever the triangle T_n shrinks to some point X, the values $\overline{\kappa}(T_n)$ converge to some limit $K(X)$.

Under this condition, in a neighborhood of each point of the manifold R, we can introduce the geodesic polar coordinates r,ϕ and define the metric of the manifold by the line element

$$ds^2=dr^2+B^2d\phi^2$$

in these coordinates. The coefficient B has continuous derivatives $\partial B/\partial r$ and $\partial^2 B/\partial r^2$. Also, all classical formulas of intrinsic differential geometry hold: the Gauss formula for the curvature $K(X)\equiv K(r,\phi)$, i.e.,

$$\frac{\partial B(r,\phi)}{\partial r^2}+K(r,\phi)B(r,\phi)=0,$$

the formula for the geodesic curvature of the curve $\phi=\phi(r)$ in the form

$$k=\frac{(B\phi')'+[1+(B\phi')^2]\phi'}{[1+(B\phi')^2]^{\frac{3}{2}}},$$

and also the formulas for the length, area, and angle between the curves.

Thus, this theorem yields a purely metric definition of "Gaussian metric" to which basically all tools of classical differential geometry can be applied.

This theorem is a simple consequence of results formulated earlier. Indeed, since the magnitude

$$\overline{\kappa}(T)=\frac{\alpha+\beta+\gamma-\pi}{\sigma_0}$$

has limit as T shrinks to a point, it is bounded in a small neighborhood of each point, i.e., for each small domain U, there exist K_1 and K_2 such that

$$K_1\sigma_0\geq\alpha+\beta+\gamma-\pi\geq K_2\sigma_0.$$

If σ_K is the area of the triangle on a surface of constant curvature K which has sides of the same length as those of a given triangle T, then the ratio σ_K/σ_0 tends

© 2006 by Taylor & Francis Group, LLC

to 1 when the length of the maximal side of the triangle T tends to zero. Therefore, we can find K such that for all sufficiently small triangles T in the domain U, we have

$$\alpha + \beta + \gamma - \pi \geq K\sigma_K;$$

since

$$K\sigma_K = \alpha_K + \beta_K + \gamma_K - \pi,$$

we thus have

$$\alpha + \beta + \gamma \geq \alpha_K + \beta_K + \gamma_K,$$

i.e., the metric on the domain U has curvature $\geq K$. Consequently, we can use all properties of this metric.

If σ is the area of the triangle T itself, ω is its curvature, and d is its diameter, then, as was mentioned in Sec. 1, we have

$$0 \leq \sigma - \sigma_K \leq \frac{1}{2}[\omega - K\sigma_K]d^2.$$

Since

$$\omega = \alpha + \beta + \gamma - \pi \leq K_1\sigma_0$$

and the ratio $\sigma_K/\sigma_0 \to 1$ as $d \to 0$, we have that in the same way $\sigma/\sigma_0 \to 1$. Therefore, the existence of the limit of the magnitudes

$$\overline{\kappa}(T) = \frac{\alpha + \beta + \gamma - \pi}{\sigma_0}$$

implies the existence of a similar limit of the specific curvatures

$$\kappa(T) = \frac{\alpha + \beta + \gamma - \pi}{\sigma}.$$

Thus, the metric satisfying all conditions of the theorem turns out to be a metric of curvature $\geq K$ with the Gaussian curvature defined everywhere. We can use the arguments of Sec. 1 in Chapter XI for this metric.

3. Surfaces of Indefinite Curvature

Although intrinsic geometry of surfaces in Lobachevskii space embraces the intrinsic geometry of small domains of all regular surfaces, this geometry, e.g., undoubtedly does not embrace the intrinsic geometry of all polyhedra. A polyhedron such that its complete angle at some vertices exceeds 2π cannot be isometric to a convex surface, since the complete angle at each point on a convex surface is $\leq 2\pi$. Therefore, the problem is to extend further the class of surfaces under study in such a way that this class will contain all possible polyhedra and also all those surfaces that admit a sufficiently good polyhedral approximation. Of course, the concept of a good polyhedral approximation is slightly uncertain, but, anyway, we must involve those surfaces whose intrinsic metric admits a polyhedral approximation.

Along with the possibility of approximation by polyhedral metrics, the intrinsic metric of a surface under study must also have other properties, so that one can set

© 2006 by Taylor & Francis Group, LLC

forth their intrinsic geometry with the content as rich as that of the intrinsic geometry of convex surfaces. Of course, first of all, it is necessary that there should exist a definite angle made by any two shortest arcs emanating from one point. Further, in our intrinsic geometry of convex surfaces, as in Gaussian intrinsic geometry, the concept of curvature, and, in particular, its complete additivity property plays an extremely important role. Therefore, it is natural to search for the class of those surfaces on which it is possible to define curvature as an additive set function.

These general reasons lead us to the following two series of conditions which can naturally be imposed on those surfaces that we want and can study by our methods (for simplicity, we shall keep in mind the surfaces in Euclidean space):

I. For a surface F, there exists a sequence of polyhedra P_n converging to it that satisfies the following conditions: (1) if points X_n and Y_n on P_n converge to points X and Y on F, then the distances $\rho_{P_n}(X_nY_n)$ converge to the distance $\rho_F(XY)$; (2) the sums of the absolute values of the curvatures of the vertices of the polyhedra P_n are bounded uniformly (the curvature ω of a vertex is equal to 2π minus the complete angle θ at this vertex, and if $\theta > 2\pi$, then $\omega < 0$).

II. The intrinsic metric of a surface F must satisfy the following two conditions: (1) there is an angle between every two pairs of shortest arcs that emanate from a common point; (2) the sum of absolute values of the excesses of the triangles on the surface F are bounded uniformly for the triangles with disjoint sides (the excess of a triangle is the sum of its angle minus π, i.e., $\alpha + \beta + \gamma - \pi$, and if $\alpha + \beta + \gamma < \pi$, then $\alpha + \beta + \gamma - \pi < 0$).

Conditions I(2) and II(2) are intended to ensure the existence of the curvature on the surface F. The curvature must be defined through the excesses of triangles, and without condition II(2) it will certainly fail to be an additive set function taking no infinite values.

It turns out that both these series of conditions I and II are equivalent from the viewpoint of intrinsic geometry. Namely, the following theorem holds: *Let R be the manifold ρ. The following two series of conditions I^* and II^* are equivalent*:

I^*. *In every polyhedron P, the metric ρ admits a uniform approximation by polyhedral metrics ρ_n such that the sums of the absolute values of their vertices are uniformly bounded.*

II^*. (1) *There exists some angle between every pair of shortest arcs emanating from a common point, and* (2) *for any polygon P, the sum of the absolute values of the excesses of the triangles in P with disjoint sides has an upper bound depending only on P.*

One can prove that *each surface represented in the rectangular coordinates by the equation $z = f(x, y)$, where $f(x, y)$ is a difference of two convex functions, satisfies condition* I *and, therefore, condition* II*; thus, the same holds for a surface coverable by finitely many of these surfaces.*

However, the latter do not exhaust all surfaces that satisfy condition I, though, in a sense, they seem to represent the majority of them. Anyway, we have here a broad class of surfaces the intrinsic geometry of which should admit as contensive development as the intrinsic geometry of convex surfaces does. But the theory of these surfaces has not been elaborated yet, so the fact that we speak about them below is an indication of arising problems to no less extent than the presentation

© 2006 by Taylor & Francis Group, LLC

of results obtained. The following can be proved for the surfaces F or manifolds R that satisfy condition I (or I*) and, therefore, conditions II (or II*):

1. There exists an angle in the strong sense made by two shortest arcs emanating from a common point on each surface F.

2. The general theorems on addition of angles and angles of sectors obtained in Secs. 1 and 3 of Chapter IV also hold on a surface F.

3. Shortest arcs emanate from each point of a surface F in almost all directions (in the sense of angular measure).

4. For each point O of the surface F, there exists a cone such that a neighborhood of its vertex is isometric to a neighborhood of the point O to within the infinitesimals of higher order. The complete angle at the vertex of this cone is equal to the complete angle at the point O in the sense of the definition given in Sec. 3 of Chapter IV.

5. Each polygon on a surface F has area in the sense of our intrinsic definition based on triangulation.

6. Let T be a triangle on F, let T_0 be the plane triangle with sides of the same length, and let ω^+ and ω^- be the least upper bound and the greatest lower bound of the excesses of triangles that are contained in T. For each pair of the corresponding angles α and α_0 of the triangles T and T_0, the following inequality holds:
$$\omega^- \leq \alpha - \alpha_0 \leq \omega^+.$$

7. Given two shortest arcs L and M emanating from a common point O on the surface F, we define the angle $\gamma(x, y)$ in the same way as done in the convexity condition. If x and y are given as increasing functions of the parameter t, then the angle $\gamma(x(t), y(t))$ has bounded variation as a function of the parameter t, i.e., a difference of two monotone functions. This is a generalization of the convexity condition.

We can list a number of other results similar to those for convex surfaces. However, many questions remain unsolved. First of all, it is necessary to pose the question of conditions under which a metric on abstract manifold is realizable, at least locally, by a surface of the class under consideration. The spatial shape of these surfaces has not been studied. It is not known what geometric properties of a surface ensure the possibility of defining it by the equation $z = f(x, y)$, where $f(x, y)$ is a difference of two convex functions. Finally, it is necessary to extend the Gauss theorem to these surfaces while it remains unknown even how to define the spherical image. For surfaces that have a tangent plane at each point, this is done in the usual way: for a convex surface, we used support planes; but for nonconvex nonsmooth surfaces, we have to proceed in a different fashion.

Here, the intrinsic curvature cannot be defined as easily as in the case of convex surfaces. If the complete angle at a point X is equal to θ, then the curvature of this point is defined by the formula $\omega(X) = 2\pi - \theta$. Here, θ can be greater than 2π, and a

© 2006 by Taylor & Francis Group, LLC

shortest arc can pass through such a point. For example, shortest arcs pass through the vertex of a cone with complete angle $> 2\pi$. The points with complete angle $< 2\pi$ push away the shortest arcs, in a sense, whereas points with complete angle $> 2\pi$ attract them. If we draw three generators on a cone with complete angle $\theta > 2\pi$ separating this cone into three sectors U_1, U_2, and V with angles π,π, $\theta - 2\pi$, then all points of the sector V are connected with points of the generator, which separate the sectors U_1 and U_2 by shortest arcs passing through the vertex of the cone. Since points with curvature < 0 can lie on a shortest arc, the curvature of the shortest arc itself cannot be considered to be equal to zero. Moreover, a shortest arc that is a side of some triangle can flex when passing through points with curvature < 0 and form a sector with angle $> \pi$. Therefore, even on a polyhedron, the curvature of the interior of a triangle can fail to be equal to the sum of angles at the vertex minus π.

For example, take a cone with complete angle $\theta > 2\pi$ and construct a triangle ABC on this cone so that its side BC passes through the vertex of the cone O and forms a sector that lies in the triangle and has the angle $\theta - \pi$; this angle is greater than π. The triangle ABC can be considered as the quadrangle $ABOC$ with angles $\alpha, \beta, \theta - \pi$, and γ. This quadrangle can be unfolded onto the plane. Therefore, first, $\alpha + \beta + (\theta - \pi) + \gamma = 2\pi$, i.e., $\alpha + \beta + \gamma = 3\pi - \theta < \pi$, and second, the curvature of the interior of this quadrangle, i.e., the triangle ABC must certainly be considered to be equal to zero. Meanwhile, the sum of the angles of the triangle ABC minus π is $\alpha + \beta + \gamma - \pi < 0$.

In the case of a nonpolyhedral surface, the matter becomes more complicated, since here a shortest arc can have not only flex, which can be taken into account easily, but the curvature on it can be, in a sense, continuously distributed. For example, take two copies of a plane domain, which is the exterior of some disc, and glue together both these copies along the bounding circles. We obtain a circle which doubly covers the exterior domain of the disk and has the circle L of this disk as an edge. Of course, this surface admits a good polyhedral approximation. Obviously an arc of the circle that is less than the half-circle is a shortest arc on this surface. The tangent cone has the complete angle 2π at all points of this circle, but its curvature cannot be considered to be equal to zero and must be taken as equal to -4π.

Take a square surrounding the circle L on one of the leaves of our surface. Along with the circle L, this square bounds a geodesic polygon P with Euler characteristic $\chi = 0$ and four angles equal to 4π. If the curvature of the interior of the polygon is defined as in the case of a convex surface, i.e., by the formula

$$\omega = 2\pi\chi - \sum_i (\pi - \alpha_i),$$

then we obtain $\omega(P) = -2\pi$. Meanwhile, the curvature of the interior of the polygon P must be equal to zero, since P is a piece of the plane. All these show that we need a new definition of curvature in the case of nonconvex surfaces. However, we do not present this definition here and leave the question of the most successful choice open.

Among still unsolved problems, we can list the question of the existence of a half-tangent to a shortest arc on a surface that is given by a difference of two

© 2006 by Taylor & Francis Group, LLC

convex functions or admit a good polyhedral approximation. The related issues are the question on the extrinsic-geometrical meaning of the direction of a curve and the angle between the curves, and further, the question of the swerve of a curve and its extrinsic-geometric meaning. The latter question still remains unstudied even in the case of convex surfaces, as was already stressed in Sec. 2 of Chapter IX. Finally, we see a wide field of the problems of intrinsic geometry of the surface under consideration which concerns the properties of shortest arcs, geodesics, triangles, circles and convex domains, extremal problems similar to those solved in Sec. 3 of Chapter X, etc. We can hope that the progress of the theory sketched here would lead not only to a wide generalization of the classical intrinsic geometry, but will also give new interesting visual geometric results.

Two viewpoints prevail in this theory, which make it different from the classical differential geometry. First, in the latter theory, curves and surfaces are given by functions from calculus, whereas in the new theory, we start from the constructive principle, i.e., a surface is considered as the limit of polyhedra, and its properties are determined by how and by which polyhedra it can be approached. From this point of view, convex surfaces are defined as the limits of convex polyhedra, while, e.g., the smooth surfaces are defined as those that can be approximated by polygons in such a way that, if points X_n of these polygons converge, then the planes of the faces passing through these points X_n also converge. Second, instead of the differential concepts of line element, Gaussian curvature, etc., as the base, we take integral concepts of distance, curvature as a set function, etc. The same viewpoint in especially simple form can be realized in the theory of curves, which should be treated as the limits of broken lines by defining the length, integral curvature, and integral torsion of a curve by approximation by broken lines.

In conclusion, let us dwell on one more class of surfaces whose properties are, in a sense, opposite to those of convex surfaces. These are surfaces of negative curvature. A polyhedron of negative curvature is characterized by the fact that the complete angle at each of its vertices is greater than 2π. A metric of negative curvature is defined in a corresponding way.

Consider a surface whose intrinsic metric can be obtained as the limit of polyhedral metrics of negative curvature. On such a surface, the sum of angles of a triangle is always less than π, and the complete angle at each point is $\geq 2\pi$. The metric of such a surface satisfies the "concavity" condition, i.e., the angle $\gamma(x, y)$, which is defined in the same way as in the convexity condition, turns out to be a nondecreasing function of x and y. Every regular surface of negative curvature has this property. If we glue a new surface from pieces of surfaces of negative curvature in such a way that the sum of swerves of segments of boundaries is always nonpositive, then we obtain a surface with the same properties. Here, we come across the gluing operation, which in the case of general nonconvex surfaces also yields a simple and visual method of proof. We mention one theorem whose proof is convenient when using this method. On a surface of negative curvature, the area of a geodesic polygon homeomorphic to a disk is no greater than the area of the plane polygon with sides of the same length inscribed into this polygon (in particular, the area of a triangle in the concave metric is no greater than the area of the plane triangle with sides of the same length). The area of each domain homeomorphic to a disk is no greater than the area of a plane disk such that the length of its circle

© 2006 by Taylor & Francis Group, LLC

is equal to the length of the curve bounding the domain considered. Hence, if S is the area of the domain and l is the length of the curve, which bounds this domain, then $S \leq l^2/(4\pi)$. The sign of equality appears here if and only if the domain is isometric to a bounding plane disk.

© 2006 by Taylor & Francis Group, LLC

Appendix

BASICS OF CONVEX BODIES

In Appendix, we present all basic facts on convex bodies used in this book. Restricting consideration to only necessary facts, we do not pretend to provide a complete presentation of the fundamentals of the theory of convex bodies.[1]

All results proved here are known to everybody who dealt with the preliminaries of the theory, and, therefore, the experienced reader may skip this Appendix – this will have no harmful effect on understanding the basic content of the book.

1. Convex Domains and Curves

A *convex* set is a set of points which includes the whole segment between every two points together with these points. A *convex domain* is a closed convex set on the plane which has interior points. Examples of convex domains are given by a disk, the domain bounded by a parabola, a half-plane, a band between two parallel lines, etc. A segment, a disk with the deleted circle, and a single point are examples of convex sets that are not convex domains in the sense of our definition. A convex domain is said to be *bounded* if this domain can be included in some disk.

The boundary of a bounded closed domain is called a *closed convex curve*. As is known, the boundary of a set is the set of points such that in each neighborhood of each of which there are points that belongs to this set as well as points that do not belong to this set. It is useful to keep in mind the easily proved fact that the boundary is always a closed set. Of course, the term "closed" curve as well as the term "closed" surface has a completely different origin.

Theorem 1. *A closed convex curve is homeomorphic to a circle.*

Let L be a closed convex curve, and let G be the convex domain whose boundary is the curve L. Take a point O inside the domain G (Fig. 98). Each half-line that emanates from O intersects the curve L, since otherwise, the domain G would not be bounded. At the same time, each such half-line intersects the curve L only at one point. Indeed, assume that some half-line emanating from the point O intersects L at two points A and B; moreover, assume that, e.g., the point A lies farther from

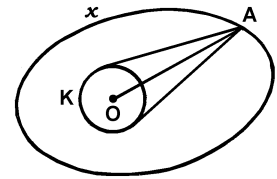


Fig. 98

[1] It was created by papers of Brunn and Minkowski at the end of XIX century and then developed by a number of researchers in geometry. See, e.g., Minkowski. Gesammelte Abhandlungen, Bd. II; W. Blaschke, Kreis und Kugel; T. Bonnesen und Fenchel, Theorie der konvexen Körper.

© 2006 by Taylor & Francis Group, LLC

O than B. Since the point O lies inside the domain G, we can circumscribe a disk K about this points which lies entirely in G. Connecting the point A with all points of the disk K, we obtain a figure that is bounded by the tangents to the disk K which are drawn from A, and by an arc of the disk K. This figure lies in the domain G, since, by the convexity of the domain G, each segment connecting A with any point of the disk K lies in the domain G. But our figure contains the whole segment AO but for the point A in its interior; hence this figure includes the point B in its interior. Thus, the point B also lies inside the domain G but not on its boundary.

It follows from what we have proved that projecting the curve L to the circle C of the disk K from the point O, we obtain a one-to-one correspondence between the points of the circle and those of the curve L. Let us prove that this correspondence is also bicontinuous. When the point X moves continuously along the curve L, the segment OX revolves continuously, and, therefore, the corresponding point Y on the circle C moves continuously as well. To prove the continuity of the inverse dependence of the point X on the point Y, it suffices to prove that if points Y_i on the circle converge to a point Y, then the corresponding points X_i converge to the point X corresponding to the point Y. Since the curve L is bounded, we can choose a convergent subsequence from the points X_i; the limit point $\overline{X}$ of this subsequence also belongs to L (L is the boundary of the domain G and, as any boundary, is a closed set). At the same time, the point $\overline{X}$ lies on the half-line, which is drawn from the point O through the point Y, since Y_i converge to Y. Therefore, it follows from what was proved above that the point $\overline{X}$ coincides with X. Thus, every convergent subsequence of points X_i converges to the point X, and, therefore, the whole sequence of these points converges to X as required.

We have thus proved that each closed convex curve can bicontinuously and bijectively be projected to the circle, i.e., every closed convex curve L is homeomorphic to the circle. Therefore, in particular, every two points of a closed curve divide it into two arcs.

A *convex line* is a subarc of a closed convex curve.

A *support line* to a set M is a line such that it has at least one common point with M and the whole set M lies in one of the half-planes bounded by this line; moreover, the line itself belongs certainly to this half-line.

Theorem 2. *At least one support line passes through each point of a convex curve.*

Proof. Obviously, it suffices to prove that a support plane to any convex domain passes through each its boundary point. Let G be a convex domain, and let A be a point of its boundary. Drawing half-lines from the point A through all points of the domain G, we obtain some angle V. If points X and Y lies on the so-drawn half-lines, then there are points X_0 and Y_0 of the domain G on these half-lines. The segment X_0Y_0 belongs to G, and, therefore, the segment XY belongs to the angle V. Consequently, the angle V is convex. However, this angle may fail to include its sides; for example, this is the case if the domain G is a disk. However, we include the sides into the angle V. Since the angle V is convex, every its side, together with a prolongation of this side, forms a support line to V at its vertex. Since the domain G lies in the angle V and the point A lies on the boundary of G, this line is also a support line to the domain G at the point A.

© 2006 by Taylor & Francis Group, LLC

Theorem 3. *Every convex curve is rectifiable, i.e., has a definite length; if a closed curve L_1 encloses a convex curve L, then the length of L is no greater than the length of L_1.*

Proof. If a closed broken line L_1 encloses a closed convex broken line L, then the length of L is less than the length of L_1. Indeed, prolonging one of the sides of the broken line L to both sides up to the intersection of the broken line L_1, we cut out a broken line L_1' from L_1 which also encloses L but has length less than the length of L_1. Repeating this operation, we cut out the broken line L from the domain, which is bounded by the broken line L_1. But since at each step, the length of the enclosing broken line decreases, the length of L is less than the length of L_1.

Let L be a closed convex curve, and let L_1 be some broken line that encloses this curve. It is easy to verify that each broken line inscribed into a convex curve is convex (we leave the proof to the reader).[2] Therefore, it follows from what we just have proved that the length of some broken line inscribed into the curve L is less than the length of the broken line L_1. Consequently, there exists a least upper bound of the lengths of the inscribed broken lines, i.e., the curve L has length.

Each arc of a rectifiable curve is also rectifiable, and since every closed convex curve is rectifiable, every convex curve is also rectifiable.

Now the second part of our theorem follows from a similar assertion for broken lines, which was just proved; it suffices to make an obvious passage to the limit.

2. Convex Bodies. A Supporting Plane

A *convex body* is a closed convex set in the three-dimensional space that has interior points; in the plane, the role of convex bodies is played by convex domains. A *closed convex surface* is the boundary of a finite convex body.

Theorem 1. *If a convex set does not lie in a single plane, then it has bounded interior points.*

Proof. Assume that a convex set M do not lie in a single plane. Then there are four points A, B, C, and D in this set that do not lie in the same plane. The segment AB lies in M since M is convex. By the same reason, the segment connecting the point C with an arbitrary point of the segment AB also lies in M, and hence the whole triangle ABC lies in M. Therefore, by convexity of M, the segment connecting the point D with any point of the triangle ABC also lies in M. This means that the whole tetrahedron $ABCD$ lies in M, so that M contains interior points.

This theorem makes it clear that closed convex sets can be of the following four types (1) convex bodies; (2) plane convex domains; (4) parts of a line, i.e., a segment, a half-line, and a line itself; (4) a point.

[2]For example, we can use the concept of convex hull, which is defined below in § 5. Taking the convex hull of a finite set of points lying on a convex curve L, we obtain a convex polygon whose boundary is a convex broken line inscribed into L. By an inscribed broken line, we always mean a broken line that is obtained when we connect by segments some points of the curve in the order of their location on the curve. The order of points on a convex curve is naturally defined since we have proved that a closed convex curve is homeomorphic to a circle.

© 2006 by Taylor & Francis Group, LLC

In turn, convex bodies can be of the following five different topological types:

1. Bounded convex bodies; they are homeomorphic to a ball; for example, an ellipse, a cube, etc.;
2. Infinite or unbounded convex bodies homeomorphic to a half-space; for example, the body with boundary a paraboloid of revolution;
3. Infinite convex cylinders; they are homeomorphic to a circular cylinder;
4. Layers or slabs between pairs of parallel planes;
5. The whole space.

It is quite obvious that all these types are topologically different, i.e. a body of one type cannot bijectively and bicontinuously be transformed into a body of another type. The fact that these types exhaust all possible topological types of convex bodies is proved here in Sec. 4.

Theorem 2. *The intersection of every family of convex sets is a convex set.*

Indeed, if two points X and Y belong to convex sets M_i, i.e., belong to their intersection, then the segment XY lies in each of these sets, i.e., is contained in their intersection as claimed.

Since a plane is a convex set, this theorem implies that a plane passing through interior points of a convex set intersects this set along a convex domain and intersects its boundary along a convex curve.

A plane P is called a supporting plane to a set M at a point A of P if this plane passes through the point A and the whole set M lies to one side of this plane, i.e., M lies entirely in one of the half-spaces defined by the plane P; moreover, the plane P is included in this half-plane.

In particular, if a set lies in a single plane, then this plane is a supporting plane to this set at its every point. The plane of a face of a cube is a supporting plane to this cube at each point of the face. We can draw infinitely many supporting planes through a vertex of the cube. A tangent plane to the surface of a convex body is always a supporting plane.

Theorem 3. *A convex body has at least one supporting plane at each point of its boundary.*

Proof. Let A be a point on the boundary of a convex body H (Fig. 99). Draw a plane P through the point A and some interior point of the body H. This plane intersects H along a certain convex domain G. By what we have proved in Sec. 1, the domain G has a support line l at the point A. Take the plane Q orthogonal to the line l, and project the body H to this plane. We obtain a convex domain G_1 in the projection (the proof of the fact that the projection of a convex body is a convex domain is obvious.)

Let A_1 be the projection of the point A on the plane Q. The line l, serving as a support line to the domain G at the point A, divides the plane P into two parts in one of which there are no points of the domain G and, hence, no points of the body H.

© 2006 by Taylor & Francis Group, LLC

Therefore, the intersection line l_1 of the planes P and Q is divided by the point A into two parts one of which does not include the points of the projection of the body H at all. Consequently, the point A_1 lies on the boundary of the projection of the body H, i.e., on the boundary of the domain G. In this case, a support line l_2 to the domain G passes through the point A.

The plane R passing through the line l_1 and l_2, is a supporting plane to the body H at the point A. Indeed, since the projections of the body H lie to one side of the line l_2, the body H itself lies to one side of the plane R. Moreover, the plane R passes through the point A.

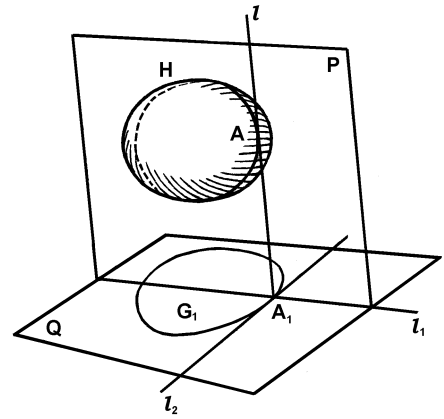


Fig. 99

Note that the theorem is trivial for convex sets without interior points: by Theorem 1, such a set lies in a single plane and, hence, this plane is a supporting plane at each point of this set.

The outer normal to a supporting plane of a set M is a unit vector orthogonal to this plane and is pointing to the half-space without the points of the set M. If the set M lies in a supporting plane, then both normals are outer. If $\mathbf{n}$ is the unit vector orthogonal to the plane, and h is the distance from the plane to the origin, which is assumed to be positive in direction from the origin to $\mathbf{n}$ and negative in the opposite direction, then the equation of the plane can be written in the following well-known normal form:

$$\mathbf{n}x = h,$$

where x is the varying vector drawn from the origin.

If $\mathbf{n}$ is an outer normal to a supporting plane of the set M, then the set M lies in the half-space defined by the inequality

$$\mathbf{n}x \leq h.$$

Theorem 4. *If closed convex sets H_1 and H_2 are disjoint and one of them is bounded, then there is a pair of nearest points A_1 and A_2 in them. There are supporting planes P_1 and P_2 to A_1 and A_2 at these points that are orthogonal to the segment A_1A_2. Each of these planes separates one of the sets H_1 and H_2 from the other.*

Proof. If H_2 is bounded, then this set contains a point A_2 that is nearest to the set H_1.[3] Then the set H_1 contains a point A_1 that is nearest to A_2. Draw the plane P_1

[3]Let two points X_1 and X_2 belong to H_1 and H_2, respectively. Let $\rho(X_1X_2)$ be the distance between them. The distance from the point X_2 to the set H_1 is the greatest lower bound of the distances $\rho(X_1X_2)$ when X_1 runs over all H_1. Let $r(X_2)$ be this greatest lower bound. A nearest point to H_1 is a point at which $r(X_2)$ has the minimal value. It is easy to prove that $r(X_2)$ is a continuous function of the point X_2; then the closedness and the boundedness of the set H_2 imply that there is a point in H_2 at which $r(X_2)$ attains the minimal value.

© 2006 by Taylor & Francis Group, LLC

that is orthogonal to the segment A_1A_2 and passes through the point A_1 (Fig. 100). Assume that the set H_1 contains a certain point X that lies to the side of the plane P_1 to which the point A_2 does. Then the convexity of H_1 implies that the whole segment A_1X lies in the set H_1. Since the points X and A_2 lies to one side of the plane P_1, the perpendicular dropped from the point A_2 to the line A_1X intersects this line at a certain point Y lying on the segment AX. Since the segment A_1X lies in the set H_1, the point Y belongs to H_1. The perpendicular A_2Y is shorter than the slanting line A_2A_1, and hence the point Y is closer to A_2 than A_1. However, this contradicts the fact that the point A_1 is the nearest point to A_2 among all points of the set H_1 Consequently, the set H_1 has no points in the half-space, which the point A_2 belongs to. This means that the plane P_1 is a supporting plane to H_1.

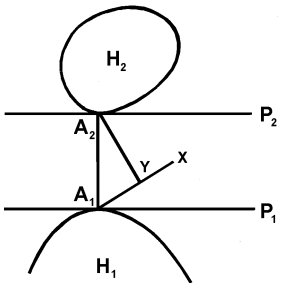


Fig. 100

In exactly the same way, we prove that the plane P_2 drawn through the point A_2 and orthogonal to the segment A_1A_2 is a supporting plane to the set H_2. The planes P_1 and P_2 are parallel. The plane P_1 divides the space into two half-spaces, and we have shown that there are no points of the set H_1 in the half-space containing the point A_2. The same is true for the plane P_2. This makes it obvious that the layer between the planes P_1 and P_2 does not contain points of H_1 and H_2 and separates these sets.

This theorem is formulated in a very simple manner if the set H_2 reduces to a single point A_2. Then H_1 contains a point A_1 that is nearest to the point A_2 and the plane orthogonal to the segment A_1A_2 at the point A_1 is a support plane of H_1 separating the point A_2 from the set H_1.

3. A Convex Cone

A *convex cone* is a convex body shaped by half-lines emanating from a common point, the vertex of this cone; moreover, we exclude the case where a cone is the whole space. A dihedral angle and a half-space are particular cases of a cone. However, in the main part of our book, a cone is the *surface* of such a solid cone. Let K be a convex cone with vertex O. If we circumscribe the unit sphere around the point O, i.e., the sphere of radius 1, then the cone cuts out a certain domain G from S. This domain lies on a hemisphere, namely, on the one of those hemispheres which are cut out by each plane that is a support plane to the cone K at its vertex. The domain G is spherically convex, i.e., the shorter the arc of the great circle which connects every two points of the domain G belongs to the domain G. The proof of this assertion is easy from the convexity of K.

Taking a point A inside the domain G and circumscribing a small disk around this point, we can draw arcs of great circles from A up to their intersection with the boundary of the domain G. Then, repeating the arguments of the proof of Theorem 1 of Sec. 1, we show that the boundary of the domain G is a curve homeomorphic to a circle. This curve is called a *spherically convex curve*. For spherical broken lines (i.e., lines that are composed of arcs of great circles) that lie on one hemisphere, we

© 2006 by Taylor & Francis Group, LLC

can prove that if a broken line L_1 encloses a spherically convex broken line L, then the length of L is no greater than the length of L_1. The proof literally repeats the proof of the corresponding assertion for broken lines on the plane, which is given in Sec. 1. Therefore, repeating the arguments of the proof of Theorem 3 of Sec. 1, we show that the length of a spherically convex curve is no greater than the length of the curve that encloses it and lies on the same hemisphere. In particular, this length is no greater than the length of the equator of the hemisphere, i.e., no greater than 2π, since the radius of the sphere is equal to 1. If we cut the surface of a cone along some its generator and develop the result onto the plane, then the curve L along which the surface of the cone K intersects the sphere S passes into an arc of the unit circle. The center of this circle is the image of the vertex of the cone passes. Since the length of the curve L is no greater than 2π, the angle unfolding on the surface of the cone under the development is also no greater than 2π. This angle is called the *complete angle at the vertex of the cone* K. In particular, if the cone K is a polyhedral angle, then its complete angle is equal to the sums of its plane angles. Consequently, the sum of the plane angles of a convex polyhedral angle does not exceed 2π. It is easy to see that this sum is always less than 2π whenever the polyhedral angle is not reduced to a dihedral angle or a half-space.

4. Topological Types of Convex Bodies

Theorem. *Convex bodies can be only of the following five topologically different types*: (1) *bounded convex bodies that are homeomorphic to a ball*; (2) *infinite convex bodies homeomorphic to a half-space*; (3) *cylinders homeomorphic to an infinite circular cylinder*; (4) *layers between parallel planes*; (5) *the whole space.*

If a closed component of the boundary of a convex body is called a complete convex surface, then this theorem implies that complete convex surfaces can be of the following three types: (1) closed surfaces homeomorphic to the sphere; (2) infinite surfaces homeomorphic to a plane; (3) cylindrical surfaces homeomorphic to the surface of infinite circular cylinder.

The proof of this theorem rests on the following two lemmas.

Lemma 1. *If a point O lies inside a convex body H, then a ray issuing from O either has no common points with the boundary of H or has a unique common point with it.*

Assume that a ray L issuing from O has a common point A with the boundary of the body H. Circumscribe a ball S around O that lies in the interior of H. Then, by the convexity condition, all segments connecting the point A with points of the ball S lies in the body H, and hence the body H include the whole cone composed of them.[4] The segment OA lies in the body H and hence in the interior of the body H itself. Obviously, this implies that the ray OA can contain at most one point of the boundary of the body H. Indeed, assuming that there are two intersection points A and B such that B lies closer to O than A, we arrive at a contradiction to the fact that the whole segment OA lies inside H.

[4] A similar construction was performed in the proof of Theorem 1 in Sec. 1 (Fig. 98).

© 2006 by Taylor & Francis Group, LLC

Lemma 2. *If a convex body H contains whole rays emanating from its interior point O, then these rays form one of the following four figures*: (1) *the whole space;* (2) *a line*; (3) *a plane*; (4) *a convex cone included also a closed plane angle and a single ray inclusively.*

Let L_1 and L_2 be two rays lying in the body H and emanating from the point O. Take two points A_1 and A_2 on them. Since the points A_1 and A_2 belongs to the body H, the whole triangle OA_1A_2, together with this segment, lies in H. Since the point A_1 and A_2 can be taken to be however far from O, the whole angle between L_1 and L_2 lies in H whenever the rays L_1 and L_2 are not the prolongations of each other.

This makes it clear that the set of all rays issuing from O and are lying in H forms a convex figure. Therefore, we have only the following possibilities for this figure: (1) this set includes only one ray; (2) there are only two such rays and they form a single straight line; (3) all such rays lie in one plane, but not on one straight line, and, thus, fill a convex cone or the whole space.

If rays L_n that are lying in H converge to a ray L, then the ray L is also contained in H, since H is a closed set. Consequently, the figure composed of the rays under consideration is a closed convex set.

We now prove the theorem formulated above.

Let H is a convex body. Take a point O inside this body and circumscribe the unit sphere S around this point. The following three cases are possible:

1. None of the rays issuing from O intersects the boundary of the body H.
2. Each ray issuing from O intersects the boundary of the body H.
3. Not all rays issuing from O intersect the boundary of the body H.

In the first case, the body H is obviously the whole space, and we may omit this case in what follows. Therefore, we shall assume that the body H has the boundary, which is denoted by F.

According to Lemma 1, a ray L going from O intersects the boundary F of the body H only at one point. Therefore, if, to each point X, we put in correspondence the point Y of the sphere S lying on the same ray L, then we obtain a one-to-one mapping h of the surface F onto the sphere S (or onto on a part of S). This mapping h is also continuous, since this mapping is the projection to the sphere by rays emanating from its center.

We show that the mapping h is also bicontinuous. Let the points $Y_1, Y_2, \ldots$ of the sphere S corresponding to points $X_1, X_2, \ldots$ of the surface F converge to a point Y. Assume that Y corresponds to the point X of the surface F. We have to prove that the points X_n converge to X. Assume however that the points X_n do not converge to X. The we have the following two possibilities: (1) it is possible to choose a subsequence X_{n_i} converging to a point $\overline{X}$ that is different from X in this sequence; (2) it is possible to choose a sequence of points X_{n_i} that go away from the point O from the points X_n to infinity.

Consider the first possibility. The point $\overline{X}$ belongs to the boundary of the body H as the limit point of the boundary points of the body H.[5] At the same time, the

[5] Recall that the boundary of each set is a closed set.

© 2006 by Taylor & Francis Group, LLC

rays OX_{n_i} converge to the ray OX since they pass through the points Y_{n_i} of the sphere S that converge to the point Y corresponding to the point X. Consequently, the point $\overline{X}$ must lie on the same ray OX. But, according to Lemma 1, a ray issuing from O can intersect the boundary of the body H only at a single point. Therefore, the point $\overline{X}$ must coincide with X.

We now consider the second possibility where the points X_{n_i} go away from O to infinity. By convexity, each segment OX_{n_i} is entirely included in the body H. At the same time, these segments converge to the ray OX since the points Y_{n_i} on the sphere S converge to the point Y corresponding to X. But since the body H is a closed set, the whole ray OX lies in H. But in this case Lemma 1 implies that this ray cannot contain any boundary points of the body H at all; this contradicts the fact that such a point X lies on H. Consequently, the second possibility is excluded, and we have proved the bicontinuity of the mapping h.

We now easily prove that *each ray emanating from O intersects the boundary of the body H, then the body H is homeomorphic to the ball, and its boundary is homeomorphic to the sphere.*

Indeed, if each ray issuing from O intersects the boundary F of the body H, then the mapping h defined above is bijective and bicontinuous mapping of the surface F onto the sphere S. Consequently, F is homeomorphic to the sphere. If X is a point of the body H and Y is its boundary point that lies on the ray going from O through X, then, to the point X, we put in correspondence the point X' in the ball S such that $OX' = OX/OY$ (this is possible since the ball S has radius 1). This defines a homeomorphic mapping of the whole body H onto the ball S.

Assume now that not every ray emanating from O intersects the boundary F of the body H. Then, according to Lemma 2, these rays which not intersecting F form one of the following three figures: (1) a line; (2) a plane; (3) a convex cone, a convex plane angle and a single ray inclusively.

Consider the first case.

The straight line L that does not intersect F meets the sphere S at two points A and B. All rays going from O to other points of the sphere S intersect F, and, therefore, the mapping h is a homeomorphic mapping of the surface F onto the sphere with two deleted points. Consequently, in this case, F is homeomorphic to an infinite cylindrical surface.

Let X be some point of the body H, and let Y be a point on the line L. The segment XY lies in H. Since the whole line L lies in H, we can take the point Y arbitrarily far from the point X. The set H is a closed set, which implies that this body includes the whole line that is parallel to L and passes through the point X. But X is an arbitrary point of the body H; hence the whole body H consists of these lines, i.e., is a cylinder. The cross-section of this cylinder by the plane orthogonal to the line L is a convex domain; this domain is bounded since its boundary is bijectively and bicontinuously projected onto a great circle of the sphere S.

Consider now the second case where all rays going from O that do not intersect the boundary of the body H fill some plane E. Then, arguing almost literally as in the previous case, we prove that the body H includes each line that is parallel to the plane E and passes through some point X of the body H. Therefore, the body H also includes the whole plane parallel to E and passing through X. Consequently,

© 2006 by Taylor & Francis Group, LLC

this body consists of all these planes; this obviously implies that this body is a layer between two planes parallel to E.

We now consider the third case where all rays going from O and not meeting the boundary F of the body H fill a cone K (or a convex plane angle or a single ray). The cone K intersects the sphere S along a spherically convex domain lying in a hemisphere. This domain is homeomorphic to a disk which is proved almost literally as in the case of convex domains on the plane. A domain G homeomorphic to a disk without boundary, or, which is the same, to the plane is the complement of this domain. If K reduces to a plane angle or to a ray, then the domain G is certainly homeomorphic to the plane. Therefore, the mapping h (the projection from the point O to the sphere S) defined above yields a homeomorphic mapping of the surface F onto the domain G, and thus, the surface F is homeomorphic to the plane.

Let Q be a body obtained from the ball bounded by the sphere S on deleting all points of the sphere S that do not belong to G. Define a mapping f of the body H onto the body Q as follows. Let Y be a point of the body H. If the ray going from O through the point Y intersects the boundary of H at the point X then, to the point Y, we put in correspondence the point Y' such that $OY' = OY/(1 + OY)$. It is easy to see that the mapping f so-defined is a homeomorphism, and hence the body H is homeomorphic to a half-space.

Thus, the proof of the theorem is complete.

5. A Convex Polyhedron and the Convex Hull

A *convex polyhedron* is a convex body whose surface consists of finitely many plane polygons. In the main text of our book, however, the *surface* of such convex body is called a convex polyhedron. Faces of a convex polyhedron are convex polygons. Indeed, let a point A lie inside a face Q of a convex polyhedron P. A supporting plane R at the point A does not intersect P and, therefore, R coincides with the plane of the face of Q. The common part of the plane R and the polyhedron P is exactly the face of Q; at the same time, it is convex as the intersection of two convex sets, the polyhedron P and the plane R.

Vertices of a convex polyhedron are those points of its boundary that belong to supporting planes containing no other points of the polyhedron.

The *convex hull* of a set M is the intersection of all convex sets that includes M. The fact that the intersection of convex sets is a convex set implies that the convex hull of each set is a convex set. Obviously, the convex hull of a convex body is this body itself.

Lemma. *The convex hull of a finitely many points is the geometric locus of the centers of gravity of various masses located at these points.*

By definition, the center of gravity of masses $m_1, m_2, \ldots, m_n$ at points $X_1, X_2, \ldots, X_n$ is the point obtained in the following way.

We draw a vectors x_i from some origin to each of the points X_i. Then the end of the vector

$$x = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i} \tag{1}$$

© 2006 by Taylor & Francis Group, LLC

that is drawn from the same origin is exactly the center of gravity of the given masses. Of course, it is assumed that all masses are nonnegative and $\sum_{i=1}^{n} m_i \geq 0$. Elementary properties of the center of gravity are assumed to be known. Putting $m_i = t_i \sum_{i=1}^{r} m_i$, instead of inequality (1), we obtain the relation

$$x = \sum_{i=1}^{n} t_i x_i, \tag{2}$$

where all $t_i \geq 0$ and $\sum_{i=1}^{n} t_i = 1$.

In order to prove our lemma, we have to prove the following two assertions: (1) The center of gravity of arbitrary masses m_i located at points X_i belongs to the convex hull of these points X_i. (2) The convex hull of the points X_i is lies in the set of centers of gravity of masses at these points.

Let us prove the first assertion. Let X be the center of gravity of masses $m_1, m_2, \ldots, m_n$ at the points $X_1, X_2, \ldots, X_n$. The center of gravity can be obtained in the following way. First, we find the center of gravity Y_1 of masses at the points X_1 and X_2 (if the masses at the points X_1 and X_2 are equal to zero, then we merely do not consider these points.) Further, putting the mass $m_1 + m_2$ at Y_1, we find the center of gravity Y_2 of this mass and the mass m_3 lying at the point X_3. At the same time, the point Y_2 is the center of gravity of the masses m_1, m_2, and m_3 that are located at the points X_1, X_2, and X_3. Continuing this construction, we arrive at the center of gravity of all given masses $m_1, m_2, \ldots, m_n$. At the same time, the point Y_1 lies on the segment X_1X_2 and, therefore, belongs to the convex hull of the points $X_1, X_2, \ldots, X_n$. The point Y_2 lies on the segment Y_1X_3; since Y_1 and X_3 belong to the convex hull of the points $X_1, X_2, \ldots, X_n$, we have that the points Y_2 also belongs to this hull. Repeating this argument, we finally arrive at the center of gravity of all given masses and verify that this center lies in the convex hull of the points $X_1, X_2, \ldots, X_n$.

We now prove the second assertion. Take a mass $m > 0$ at a certain point X_i and zero masses at all other points. Then the point X_i is the center of gravity. Consequently, the set of all centers of gravity of masses at the points X_i contains all points X_i. Prove that this set is convex. Let A and B be the centers of gravity which are obtained for certain distributions of masses at the points X_i; let these masses be a_i in the first case and b_i in the second; moreover, without loss of generality, we can assume that $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i = 1$. Then according to (2), the vectors from the origin to the points A and B are as following:

$$a = \sum_{i=1}^{n} a_i x_i, \quad b = \sum_{i=1}^{n} b_i x_i. \tag{3}$$

We now put the masses

$$m_i = (1 - t)a_i + tb_i, \tag{4}$$

where $0 \leq t \leq 1$, at the points X_i. Then

$$\sum_{i=1}^{n} m_i = (1 - t) \sum_{i=1}^{n} a_i + t \sum_{i=1}^{n} b_i = 1,$$

© 2006 by Taylor & Francis Group, LLC

and the location of the center of gravity of these masses is determined from the equation

$$x = \sum_{i=1}^{n} m_i x_i.$$

Using (3) and (4), we obtain

$$x = (1-t)\sum_{i=1}^{n} a_i x_i + t\sum_{i=1}^{n} b_i x_i = (1-t)\mathbf{a} + t\mathbf{b}.$$

It is clear from this formula that when t varies from zero to 1, the end of the vector x runs over the segment between the vectors $\mathbf{a}$ and $\mathbf{b}$. This means that this whole segment consists of the centers of gravity, i.e., if the points A and B belong to the set of the centers of gravity, then the segment AB is also lies in this set. Consequently, this set is convex; since it contains all points X_i, we see that this set contains their convex hull. Thus, the lemma is proved.

Theorem 1. *The convex hull of finitely many points is a convex polyhedron provided that these points do not lie in a single plane*; *otherwise, this set is either a convex polygon, or a segment, or a single point.*

Proof. Let $X_1, \ldots, X_n$ be given points, and let V be their convex hull. Let us prove that this hull is a closed set. Let a sequence of points A_i in V converge to a point A. Each point A_i is the center of gravity of some masses $m_1^i, \ldots, m_n^i$ at the points $X_1, \ldots, X_n$. We can assume that the sum of these masses is equal to 1 for each i. Then we can choose a sequence of superscripts i such that the masses $m_1^i, \ldots, m_n^i$ converge for each of the points $X_1, \ldots, X_n$. The center of gravity of the limit masses is the point A. This proves the closedness of the convex hull.

The convex hull of a single point is this point itself. The convex hull of points lying on one line is the segment of this line, which connects the leftmost and rightmost points of these points. Let points $X_1, \ldots, X_n$ lie in one plane but not in one line. Let V be the convex hull of these points, and let A be a point on the boundary. There is a support line l support to V which passes through the point A; since V lies in one of the half-planes determined by the line l, we see that all points $X_1, \ldots, X_n$ lie in the same half-planes. Then it is clear that the point A is the center of gravity of masses at the points $X_1, \ldots, X_n$ if and only if there is at least one of the points X_i on the line l itself. If this point is unique then it coincide with A, and then, A is a vertex of V. All other points of the boundary of V lie on lines passing through pairs of the points X_i. The set such lines is finite, and hence the boundary of V consists of finitely many line segments. This proves that V is a convex polygon.

Assume now that the points $X_1, \ldots, X_n$ do not lie in a single plane. Then the convex hull V of them is a convex body since, by what we have proved, is closed and has interior points by Theorem 1 of Sec. 2.

To prove that V is a convex polyhedron, it remains to verify that the boundary of V consists of finitely many polygons. Let A be a point on the boundary of V, and let P be a support plane to V at this point. Then V lies in one of the half-spaces determined by the plane P. This makes it obvious that some of the given points X_i lie in the plane P itself since otherwise the point A cannot be the center of gravity

© 2006 by Taylor & Francis Group, LLC

of masses at these points. At the same time, if a certain point X_i does not lie in the plane P and the mass at this point is > 0, then the center of gravity cannot lie in the plane P and belongs to the side of this plane where points X_i that are not lying in the plane itself. Thus, the point A is the center of gravity of masses at points on the plane P, i.e., this point belongs to the convex hull of P. This implies that the whole boundary of V consists of the convex hull of those sets of points X_i each of which consists of the points X_i lying in a single plane. But there are finitely many such sets of points X_i, and the convex hull of each of them is either (1) a single point, or (2) a segment, or (3) a convex polygon. The theorem is proved.

Theorem 2. *Each convex polyhedron is the convex hull of its vertices and hence is completely determined when the vertices are given.*

Proof. It is clear from the definition of convex hull that the convex hull of the vertices of a convex polyhedron lies in this polyhedron. Therefore, it suffices to prove that, conversely, a polyhedron lies in the convex hull of its vertices.

Let A be a certain point of a convex polyhedron P. If this point lies on an edge of P, then A obviously belongs to the convex hull of the vertices of this edge. If the point A lies inside a face Q, we draw a segment through this point to its intersection point with the boundary of the face Q. Then the endpoint s of this segments lie on edges and, hence, belong to the convex hull of the vertices. But in this case, the segment itself, as well as the point A, belongs to this convex hull. Finally, if the point A lies inside the polyhedron, then we draw a segment through this point to its intersection point with the boundary of the polyhedron. Then, by what we have proved, the endpoints of this segment belong to the convex hull of the vertices, and, therefore, the segment itself, as well as the point A, belongs to this convex hull.

6. On Convergence of Convex Surfaces

We say that a sequence of sets M_n converges to a set M or the set M is the limit of the sets M_n if (1) for each point X of the set M, there exists a sequence of points X_n converging to this point and belonging to the corresponding set M_n; (2) any point that does not belong to the set M is not a point of condensation of any sequence of points that belong to distinct sets M_n (not necessarily the point of each set, but of least a point in some of the sets M_n with distinct indices); in other words, all condensation points of sequences of points in distinct M_n belong to the set M. This definition of convergence of sets is constantly used is not formulated it in explicit form. For example, a tangent is the limit of secant lines precisely in this sense.

It should be specified that if the set M is empty, then the sequence of sets M_n is not considered as convergent, although, in this case, both requirements of the definition of the limit, can be satisfied. The first requirement holds undoubtedly for empty M despite making no sense. For example, a sequence of points that go to infinity has now points of condensation, and, therefore, we have to assume that its limit is empty set on formally applying the above definition. However, we repeat that the sequence is not considered as convergent in this case.

Lemma. *Let sets M_i lie in half-spaces that are determined by planes P_i passing through points X_i. The outer normal is a normal to the plane P_i which aims at*

© 2006 by Taylor & Francis Group, LLC

the half-space not including the set M_i. Then the following holds: if the points X_i converge to a point X and the outer normals $\mathbf{n}_i$ to the planes P_i converge to a vector $\mathbf{n}$, then (1) *the planes P_i converge to the plane P passing through the point X and having the normal $\mathbf{n}$;* (2) *every limit point of a sequence of points belonging to distinct sets M_i lies in the half-space determined by the plane P from which the normal $\mathbf{n}$ emanates.*

The first assertion of the lemma is obvious. Let us prove the second assertion. Let points Y_i lying in distinct sets M_i converge to a point Y. Let $\mathbf{y}$ and $\mathbf{y}_i$ be two vectors that go from some origin to the points Y and Y_i, respectively.

Moreover, assume that h_i and h are the distances from this origin to the planes P_i and P, which are assumed to be positive from the origin in the direction of the outer normal. Since the points Y_i lie in those half-spaces from which the normals $\mathbf{n}_i$ emanate, we have

$$\mathbf{n}_i y_i \le h_i.$$

Since the normals $\mathbf{n}_i$ converge to $\mathbf{n}$ and the planes P_i converge to P, we have that h_i converge to h. Since the points Y_i converge to H, the points y_i converge to y. Consequently, passing to the limit in (1), we obtain

$$\mathbf{n}y \le h,$$

which means that the point Y lies in the half-space determined by the plane P from which the normal $\mathbf{n}$ emanates.

Theorem 1. *Let sets M_i converge to M, and let points X_i of the sets M_i converge to a point X. If the sets M_i have support planes P_i at the points X_i, then the limit of each converging sequence of these planes is a supporting plane to the set M at the point X.*

Proof. If a sequence of planes P_i converges, then its limit is some plane P that passes through the point X. Of course, we can choose a subsequence from this sequence of the plane P_i such that the outer normals to them converge. Then the previous lemma implies that the limit of the sets M_i lies to one side of the plane P. Moreover, the plane P passes through the point X which belongs to the limit M of these sets. Consequently, the plane P is a supporting plane to the set M at the point X.

The just proved theorem can somewhat imprecisely be formulated as follows: *the limit of supporting planes of converging sets is a supporting plane of the limit of these sets*. However, this formulation is true only in the case where these sets are bounded uniformly; otherwise, the points X_i at which the planes P_i are touch the sets M_i can go to infinity, and the limit of the planes P_i can have no common points with the limit of the sets M_i at all. It is easy to give examples in which all happens precisely so. If the sets M_i are bounded uniformly, the points X_i have a condensation point, and this point is a point at which the limit of the plane P_i touches the limit of the sets M_i. Consequently, the above formulation of the theorem is admissible in this case.

A closed convex polyhedron is the boundary of a convex polyhedron, i.e., the boundary of a convex body bounded by finitely many polygons. We shall say that

© 2006 by Taylor & Francis Group, LLC

a convex polyhedron is inscribed into a convex surface F if its vertices lie on the surface F.

Theorem 2. *For each closed convex surface, there exists a sequence of closed convex polyhedra inscribed into this surface and converging to it.*

Proof. Let F be a closed convex surface, and let ε be a given positive number. Partition the whole space into cubes with sides equal to $\varepsilon/\sqrt{3}$ and take all those cubes that have common points with the surface F. Take one point of the surface F in each of the chosen cubes and denote these points by $A_1, \ldots, A_n$. If these points lie in a single plane, then we can add other points of the surface F to them so that these points, together with those chosen earlier, no longer lie in a single plane; otherwise, the surface F lies in a single plane, which is impossible; therefore, we can assume that the points $A_1, A_2, \ldots, A_n$ do not lie in a single plane.

It is easy to see that the cube with side $\varepsilon/\sqrt{3}$ containing the point A_i lies in the ball of radius ε centered at the point A_i. Since the chosen cubes cover the surface F, all balls of radius ε circumscribed around the points A_i also cover this surface. In other words, each point of the surface F is distant from one of the points A_i at most by ε.

Let P be the boundary of the convex hull of the points $A_1, \ldots, A_n$. By Theorem 1 in Sec. 5, P is a closed convex polyhedron. This polyhedron is inscribed into the surface F.[6]

We now take a sequence of numbers ε_i converging to zero; for each ε_i, we take points $A_1^i, A_2^i, \ldots, A_{n_i}^i$ as indicated. Let P_i be the convex polyhedra inscribed into the surface F via the above construction just indicated. Let us prove that the polyhedra P_i converge to the surface F.

Let X be a point on the surface F. By the choice of the points $A_1^i, A_2^i, \ldots, A_{n_i}^i$, there is a point A_{ki}^i among them that is distant from K by at most ε_i. Since ε_i tends to zero, these points A_{ki}^i converge to the point X. The points A_{ki}^i lie on the polyhedra P_i, and hence each point of the surface F is the limit of points lying on the polyhedra P_i.

Now if we prove that each convergent sequence of points lying on distinct polyhedra P_i converges to a point of the surface F, this will prove that F is the limit of P_i.

Let a sequence of points X_j lying on distinct polyhedra P_j converge to some point X. The point X cannot lie outside the body H bounded by the surface F. Indeed, the polyhedra P_i are the boundaries of the convex hulls of some points of H; since the body H is a closed set, the limit of the points X_j lies in the body H.

Thus, the point X lies in the body H. Support planes R_j of the polyhedra P_j pass through the points X_j. We can choose a sequence R_{jk} from these planes such that the outer normals to the plane R_{jk} converge. Then the planes R_{jk} themselves converge to a certain plane R passing through the point X. By the lemma proved above, the plane R is a support plane to the set of points that are limit to points lying on the polyhedra P_{jk}. We have already proved that each point of the surface F is such a limit point. Therefore, the surface F lies to one side of the plane R. But the point X lies in the body H bounded by the surface F, and at the same

[6] Only the points A_i can be vertices of P as is implied by the properties of the convex hull. Therefore, all vertices of P lie on the surface F.

© 2006 by Taylor & Francis Group, LLC

time, lies on the plane R. Consequently, this point lies on the surface of the body H, i.e., on F; as required.

Using similar reasons, we can prove that each closed convex curve is the limit of inscribed closed convex broken lines.

Theorem 3. *The limit of the boundaries of convex sets is the boundary of a convex set.*

Proof. Let the boundaries F_n of certain convex sets converge to a certain limit F. For brevity, F_n and F will be called surfaces. Let X be a point on the surface F, and let X_n be points on surfaces F_n that converge to X. The surface F_n has a supporting plane at the point X_n. We can choose a subsequence P_{n_i} from the planes P_n whose exterior normals converge. Then the plane P_{n_i} converge to a certain plane P that is a support plane to the surface F at the point X by Theorem 1.

Taking all possible points X on the surface F and all possible converging sequences of support planes P_{n_i}, we obtain all possible planes P that are limit for the planes P_n. Let H be the intersection of those half-spaces that are defined by those planes in which the surface F lies. By what we have proved above, all planes P are support to F, and therefore, the set H contains the surface F, and hence, is not empty. At the same time, this set is convex as an intersection of half-spaces. We have seen that at least one of the planes P passes through each point of the surface F. Therefore, the surface F lies on the boundary of the set H. Now, if we prove that no point of F lies on the boundary of the set H, then the theorem will be proved.

Let a point X not belong to F. Since F is the limit of the sequences F_n, the point X cannot be a condensation point of any sequences of points lying on distinct surfaces F_n. Therefore, we can circumscribe a ball S around X such that for sufficiently large n, none of the surfaces F_n intersects this ball. Assume that for arbitrarily large n, the ball S lies in those convex sets whose boundaries are the surfaces F_n. Then this ball lies in all half-spaces including the surfaces F_n and, hence, also lies in the intersection of these half-spaces. Recalling the definition of the set H, we see that the ball S lies in this set. Consequently, X is an interior point of this set and does not lie on its boundary.

Assume now that the ball S lies outside the bodies bounded by the surfaces F_n with arbitrarily large numbers n. Then by Theorem 8 of Sec. 3, for each of these surfaces, there is a support hyperplane separating this surface from the ball S. The limit of a convergent sequence of such planes is one of the planes P, which are considered above, and this limit separates the set H from the ball S. Therefore, the center of the ball S, i.e., the point X does not belong to the boundary of H.

Consequently, if a point X does not belong to F, then this point does not belong to the boundary of H. Since we have proved that F lies on the boundary of H, the set F is thus the boundary of the set H. The set H is convex, and, therefore, F is the boundary of a convex set. In particular, if F_n are closed convex surfaces that are bounded uniformly, then their limit is either a closed convex surface, or a plane convex domain, or a segment, or a point.

© 2006 by Taylor & Francis Group, LLC