© 2007 European Mathematical Society, FIZ Karlsruhe & Springer-Verlag

---

0960.01035

Alexandrov, A.D. (Reshetnyak, Yu.G.(ed.); Kutateladze, S.S.(ed.); Aleksandrov, A.D.)

Selected works. Part 1: Selected scientific papers. Ed. by Yu. G. Reshetnyak and S. S. Kutateladze, transl. from the Russian by P. S. V. Naidu. (English)

Classics of Soviet Mathematics. 4(pt.1). Amsterdam: Gordon and Breach Publishers. x, 322 p. $ 160.00; £96.00; ecu 123.00 (1996). [ISBN 2-88124-984-1/hbk; ISSN 0743-9199]

The present edition is the first part of the selected works of Alexandr Danilovich Alexandrov one of the most brilliant geometers, who exerted such a powerful influence on the development of modern mathematics, "...a giant of 20th century mathematics..." (Professor D. E. Edmunds, University of Sussex). His scientific contributions address a wide range of problems of modern mathematics and its applications.

The 16 papers included in this volume represent some of his most seminal work. Topics treated include convex polyhedrons and closed surfaces, an elementary proof and extension of Minkowski's theorem, a chapter on a generalization of Riemannian geometry, a general method for majorizing the solutions of Dirichlet problems and principles of relativity theory.

At the beginning of this century, geometry was in a position to deal with objects "in the large." However, approaches to them were not prompted by the available methods of differential geometry or by the methods for studying the solvability of partial differential equations that were propounded in the 19th century. The efforts of such outstanding mathematicians as H. Minkowski, D. Hilbert, H. Weyl, et al. led only to isolated results. At the same time, their works contained statements of many important unresolved problems that anticipated the further development of geometry "in the large" in this century. Major discoveries in the study of these problems were due to Alexandrov, who solved many of the discipline's hard and enduring problems.

The work of Alexandrov on the theory of generalized Riemannian spaces is closely related to the area of mathematical analysis that studies the concept of weak (generalized) solution to a differential equation (see for instance, the famous articles by S. L. Sobolev and L. Schwartz). Remind, that just Alexandrov has ensured customary now the use of weak convergence of functions of sets by geometrical theory of weak convergence, that he has advanced the topological methods in proofs of existence theorems, forestalling the appropriate process in theory of differential equations.

Alexandrov was the first to apply many of the tools and methods of the theory of real functions and functional analysis that are now current in geometry. One of his first geometric papers using such pioneering method is the article [On infinitesimal bendings of nonregular surfaces: Mat. Sb. (Rec. Math. Moscou, N. Ser.) 1, 307-322 (1936; Zbl Zbl 0014.41304)]. Here, the concept of infinitesimal bending is proposed and necessary and sufficient conditions are given for a vector field on a surface to be a bending field on an arbitrary surface rectifiable in the Lebesgue sense.

The paper "An elementary proof of the Minkowski and some other theorems on convex

© 2007 European Mathematical Society, FIZ Karlsruhe & Springer-Verlag

polyhedra" [Izv. Akad. Nauk SSSR, Ser. Mat. 4, 597-606 (1937)] contains a proof of a uniqueness theorem for convex polyhedra in three-dimensional space. The author derives from it a proof of the Minkowski theorem stating that two convex polyhedra coincide or are translates of one another provided that their parallel faces have the same area.

The results obtained in the papers [To the theory of mixed volumes of convex bodies. Part I: Extension of certain concepts of the theory of convex bodies, Mat. Sb. (Rec. Math. Moscou, N. Ser.) 2(44), No. 5, 947-972 (1937; Zbl 0017.42603); Part II: New inequalities for mixed volumes and their applications, Mat. Sb. (Rec. Math. Moscou, N. Ser.) 2(44), No. 6, 1205-1238 (1937; Zbl 0018.27601); Part III: Extension of two Minkowski theorems on convex polyhedra to all convex bodies, Mat. Sb. (Rec. Math. Moscou, N. Ser.) 3, No. 1, 27-46 (1938; Zbl 0018.42402); Part IV: Mixed discriminants and mixed volumes, Mat. Sb. (Rec. Math. Moscou, N. Ser.) 3, No. 2, 227-251 (1938; Zbl 0019.32804); On the area function of a convex body, Mat. Sb. (Rec. Math. Moscou, N. Ser.) 6, No. 1, 167-174 (1939; Zbl 0022.40203)] are now considered fundamental in the theory of convex bodies and the present collection offers an opportunity to become better acquainted with the theory by consulting its definitive source. In the papers [A general uniqueness theorem for closed surfaces, Dokl. Akad. Nauk SSSR 19, No. 4, 227-229 (1938; Zbl 0019.08104); Uniqueness theorems for closed surfaces, Dokl. Akad. Nauk SSSR 22, No. 3, 99-102 (1939; Zbl 0020.40202)] a general uniqueness theorem is proved for regular closed surfaces in three-dimensional space. Alexandrov indicated a new direction of research in geometry, namely the theory of nonregular Riemannian spaces. Central to Riemannian geometry is the theory of curvature of a space. Alexandrov constructed a theory of nonregular Riemannian spaces satisfying a specific generalized boundedness condition on curvature. His articles [The inner geometry of an arbitrary convex surface, Dokl. Akad. Nauk SSSR 32, No. 7, 467-470 (1941; Zbl 0061.37602); Existence of a convex polyhedron and of a convex surface with a given metric, Dokl. Akad. Nauk SSSR 30, No. 2, 103-106 (1941; Zbl 0024.35802)] devoted to the topic further develop the geometric concept of space along the lines of the tradition stemming from Lobachevskij, Riemann and E. Cartan. Thereby mathematics was enriched with new ideas. These papers deal with the intrinsic geometry of arbitrary convex surfaces in the space $\mathbb{R}^3$. A surface is considered as a metric space in which the distance between points is defined as the greatest lower bound of the lengths of curves on the surface joining the points. The challenge here is to find conditions for an arbitrary metric space to be isometric with a convex surface. Alexandrov introduces the concept of a metric of positive curvature. It is proved that a two-dimensional manifold with metric of positive curvature and homeomorphic to the sphere is isometric to a closed convex surface. This theorem provides a new solution to the familiar Weyl problem, but is now stated in an essentially more general form than that considered by Weyl himself. The solution given by Alexandrov is purely geometric in nature. The crux of the argument is the existence theorem for a closed convex polyhedron with a given development. The work [On tiling a space with polyhedra, Vestnik LGU 2, 33-43 (1954)] relates to Alexandrov's research into mathematical problems of crystallography. Here he considers the problem of tiling a space with polyhedra of which only a finite number are geometrically distinct. The chapter "On a generalization of Riemannian geometry" (pp. 187-249) is a revised and enlarged version of the lecture Alexandrov read in March 1955 at the Humbold

© 2007 European Mathematical Society, FIZ Karlsruhe & Springer-Verlag

University of Berlin. A part of the chapter was published in the paper [Trudy Mat. Inst. Steklova 38, 5-23 (1951; Zbl 0049.39501)] but a greater part is being published for the first time. In connection with particular research of intrinsic geometry of convex surfaces in two-dimensional case Alexandrov has allocated and studied the space of curvature $\leqslant K$, denoted by R_K, which is a particular case of the space bounded curvature. For metric spaces the author defines a shortest arc, a triangle, an upper angle between two curves, and introduces the relative excess $\delta_K(T)$ of a triangle T (excess relative to the curvature K) as the difference between the sum of the upper angles of a triangle T and the sum of the angles of the corresponding triangle T^K on K-plane (on a surface of constant curvature K). Observe that in the theory of manifolds of bounded curvature the excess is defined to be $\delta(T)=\alpha+\beta+\gamma-\pi$, where α,β and γ are the upper angles of T. A space R_K is such that for every "sufficiently small" triangle T its relative excess $\delta_K(T)$ is less than or equal to zero. For example, a polyhedron consisting of triangles in a K-plane is a space R_K, if the sum of the angles at each vertex not lying on a boundary is $\geqslant 2\pi$. In some sense one can assert that a space with a metric presenting the limit of metrics of spaces R_K is itself a space R_K. In particular, the limit of Riemannian metrics of curvature $\leqslant K$ is a metric of curvature $\leqslant K$. General propositions concerning upper angles, main properties of the space R_K, the direction of a curve and the angle of the direction cone, the area of a surface and the isoperimetric inequality in R_K are expounded, and some results concerning an angle in the strong sense, a space with curvature $\geqslant K$, a ruled surface in R_K, a cone in R_K and deviation of a curve from a shortest arc are stated. A wide range of articles by Alexandrov are devoted to the theory of partial differential equations. The starting point for this research is within geometry. The present volume contains only two articles reflecting this side of his activities. In [The Dirichlet problem for the equation $\det\|z_{ij}\|=\varphi(z_1,\dots,z_n,z,x_1,\dots,x_n)$ Part I: Vest. Leningr. Univ. 13, No. 1 (Ser. Mat. Mekh. Astron. No. 1), 5-24 (1958; Zbl 0114.30202)] where z is an unknown function, z_i are its first derivatives and z_{ij} are second derivatives, the existence and uniqueness of the solutions are studied. The concept of solution is treated here in a generalized sense, and is determined by geometric means. The paper [A general method for dominating solutions of the Dirichlet problem, Sib. Mat. Zh. 7, No. 3, 486-498 (1966; Zbl 0146.34701)] outlines a method for estimating solutions to second order partial differential equations that relies upon geometric considerations connected with the convex envelope of a solution. The final paper [On the principles of relativity theory, Vestnik LGU 19, 5-28 (1976)] gives an overview of Alexandrov's research related to the mathematical foundations of relativity theory. The question addressed and solved here is to find minimal conditions characterizing Lorentz transformations in relativity theory. This monograph offers any student or professional working in the field a unique access to a selection of Alexandrov's best quality work, published here for the first time in English translation.

S.M.Pokas (Odessa)

Keywords : convex polyhedra; mixed volumes of convex bodies; uniqueness theorem for closed surfaces; convex surface with given metric; tilings in n dimensions; generalization of Riemannian geometry; existence of generalized solutions of PDE; principles of relativity

Classification :

$*$01A75 Collected or selected works

© 2007 European Mathematical Society, FIZ Karlsruhe & Springer-Verlag

Cited in ...

FOREWORD

The present edition is the first part of the selected works by Alexandr Danilovich Alexandrov, one of our most brilliant geometricians. His scientific contributions address a wide range of the problems of modern mathematics and its applications.

At the turn of this century, geometry was in a position to deal with the objects "in the large." However, approaches to them were prompted neither by the available methods of differential geometry nor by the methods for studying solvability of the Cauchy problem and boundary-value problems for partial differential equations which were propounded in the 19th century. The efforts of such outstanding mathematicians as H. Minkowski, D. Hilbert, H. Weyl, et al. led only to isolated results. At the same time, their works contained the statements of many important unresolved problems which anticipated the further development of geometry "in the large" in this century. The major discoveries in studying the problems belong to Alexandrov who alone solved many hard and enduring problems of the discipline.

The works by Alexandrov on the theory of generalized Riemannian spaces are closely related to the background ideas of the section of mathematical analysis which studies the concept of weak (generalized) solution to a differential equation (see, for instance, the famous articles by S. L. Sobolev and L. Schwartz).

Alexandrov is also responsible for some important results on existence of weak solutions to some nonlinear partial differential equations (the equations of Monge-Ampère type). In the case considered by Alexandrov, the concept of weak solution is not reducible to that of the theory of partial differential equations.

Alexandrov was the first to apply many of the tools and methods of the theory of real functions and functional analysis which are current today in geometry. One of his first geometric papers using such pioneering methods is the article of Chapter 1 which concerns the infinitesimal bendings of a surface of revolution. Here the concept of infinitesimal bending is proposed and necessary and sufficient conditions are given for a vector field on a surface to be a bending field on an arbitrary surface rectifiable in the sense of Lebesque. Then Alexandrov proves a rigidity theorem for a surface of revolution which asserts absence of infinitesimal bendings other than motions. The subtlety of this result is that in studying the equations of infinitesimal bendings with classical tools there appears the second derivative of the function describing a meridian of the surface under consideration. If

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

the function is an arbitrary convex function then the methods requiring its second derivative cannot be applied directly.

Chapter 2 contains a proof of some uniqueness theorem for convex polyhedra in three-dimensional space. The author derives from it a proof of the Minkowski theorem claiming that two convex polyhedra coincide or are translates of one another provided that their parallel faces have the same area. The proof of the uniqueness theorem is of pure synthetic nature. It is based on the idea underlying the proof given by Cauchy to his remarkable theorem claiming that convex polyhedra composed of congruent similarly-placed faces are congruent.

Chapters 3–6 present parts of a single treatise published in 1937 and 1938 in Matematicheskiĭ Sbornik. Chapter 9 also relates to the treatise. The obtained results are now considered fundamental to the theory of convex bodies and the present collection offers an opportunity to become better acquainted with the theory by consulting its definitive source. A short survey of the contents of the treatise is given in Chapter 3. The main topic of the treatise is the theory of mixed volumes of systems of convex bodies. The concept of mixed volume was suggested by H. Minkowski. In the works by Brunn and Minkowski some inequalities were obtained between various mixed volumes. The purpose of the treatise by Alexandrov which is reprinted in Chapters 3–6 is the establishment of some new inequalities between mixed volumes, derivation of various geometric corollaries of the inequalities and proof of numerous uniqueness theorems of the theory of convex bodies and the theorems on extremal properties of a ball and bodies circumscribed around a ball. (Alexandrov's inequalities were independently obtained by Fenchel and Jessen.) This treatise is characterized by the prolific use of some tools from functional analysis and the theory of real functions. Among the ingenious gadgets of research there are some countably additive set functions on the unit sphere, the area functions, curvature functions and some other functions introduced by Alexandrov. In the series of results of these chapters of the collection we also mention the inequality between mixed discriminants of positive definite quadratic forms which are established in Chapter 6 and are analogous to those for mixed volumes. These results belong principally to algebra. Some of their geometric applications are given in Chapter 6. This chapter also contains the proofs of Alexandrov's main inequalities between mixed volumes which are obtained by means of functional analysis (the Hilbert method, as the author calls it).

In Chapters 7 and 8 one general uniqueness theorem is proven for regular closed convex surfaces in three-dimensional space. In Chapter 7 the validity of the theorem is established under certain hypotheses which are modified in Chapter 8. The uniqueness theorem of Chapter 2 is an analog of the

general theorem applied to polyhedra.

Alexandrov indicated a new direction of research in geometry, the theory of nonregular Riemannian spaces. Central to Riemannian geometry is the theory of curvature of a space. Alexandrov constructed a theory of nonregular Riemannian spaces satisfying a specific generalized boundedness condition on curvature. His articles devoted to the topic further develops the geometric concept of space along the lines of the tradition stemming from Lobachevskiĭ, Riemann and E. Cartan. Thereby mathematics was enriched with new beneficial ideas. Chapters 10, 11 and 12 contain the articles by Alexandrov which reflect this side of his activities.

Chapters 10 and 11 deal with the intrinsic geometry of arbitrary convex surfaces in the space $\mathbb{R}^3$. A surface is considered as a metric space whose distance between points is defined as the greatest lower bound of the lengths of curves on the surface which join the points. The challenge here is to find conditions for an arbitrary metric space to be isometric with a convex surface. Alexandrov introduces the concept of a metric of positive curvature. The sought conditions are as follows: the metric space under test is to be a two-dimensional manifold with metric of positive curvature. In particular, it is proven that a two-dimensional manifold with metric of positive curvature and homeomorphic with the sphere is isometric to a closed convex surface. This theorem provides a new solution to the familiar Weyl problem but is stated now in an essentially more general form than that considered by Weyl himself. The solution given by Alexandrov is purely geometric in nature. The crux of the argument is the existence theorem for a closed convex polyhedron with a given development.

Chapter 12 relates to the Alexandrov research into mathematical problems of crystallography. Here he considers the problem of tiling a space with polyhedra among which only finitely many are geometrically distinct.

Chapter 13 is devoted to study of metric spaces of curvature not greater than K, with K a constant. A particular instance of such a space is a Riemannian space whose sectional curvature at every point is at most K. The definition is given of a space with curvature not greater than K and various properties of such spaces are established.

A wide circle of articles by Alexandrov is devoted to the theory of partial differential equations. A starting point for this research is within geometry. The present Part 1 of the collection contains only two of Alexandrov's papers reflecting this side of his activities. In Chapter 14 an existence theorem is proven for a solution to the Dirichlet problem for the equation of the form $\det\|z_{ij}\| = \varphi(z_1, \dots, z_n, z, x_1, \dots, x_n)$, where z is an unknown function, z_i are its first derivatives and z_{ij} its second derivatives. The concept of solution is treated here in some generalized sense and is determined by geometric means. The proof of the main result is based on making use of the methods

developed in the theory of convex bodies. Chapter 15 outlines a method for estimating solutions to second order partial differential equations which is indebted to some geometric considerations connected with using the convex envelope of a solution.

The final Chapter 16 gives an overview of Alexandrov's research which relates to mathematical foundations of relativity theory. The question addressed and solved here is to find minimal conditions characterizing Lorentz transformations in relativity theory.

Alexandrov has many students and followers. Virtually all of his articles, including those reproduced here, are continued and developed in the research of other scientists. Analysis and comparison would require so ample room that we were impelled to decline the challenge (with a natural relief). We hope that the reader will find the book of relevance, the paucity of comments notwithstanding.

Yu. G. Reshetnyak
S. S. Kutateladze

CHAPTER I

ON INFINITESIMAL BENDINGS OF NONREGULAR SURFACES

Matematicheskiĭ Sbornik, vol. 1 (43), No. 3, 307–321 (1936)

Study of continuous length-preserving deformations of a surface is sensible not only for regular surfaces: it suffices that on a surface there be enough curves for which the concept of length is meaningful. Lebesgue[1] defined the corresponding class of the so-called rectifiable surfaces (la famille des surfaces rectifiables) as follows.

A surface S,

$$x = f(u, v), \quad y = \varphi(u, v), \quad z = \psi(u, v),$$

given over a domain D in the plane (u, v) is said to be *rectifiable*, if to each rectifiable curve on D there corresponds a rectifiable curve on S.

For a surface homeomorphic to a sphere, we naturally take a sphere of unit radius as the domain of parameters. Then, every convex surface is rectifiable, as seen if a sphere in the interior of the surface is mapped onto it in such a way that a point of the sphere is sent to the point on the surface lying on the same radius.

At time $t = 0$ let a vector $\bar{x}(u, v)$ describe a sphere. On gradually contracting the sphere, we obtain a continuous family of surfaces $\bar{x}(u, v, t)$ isometric to the sphere such that $\bar{x}(u, v, 0) = \bar{x}(u, v)$, where every point on the surface has a definite velocity $\frac{\partial \bar{x}}{\partial t}$ at each time instant. This example demonstrates that every surface admits some bending, if bending is defined in rather general terms. Therefore I confine myself only to an infinitesimal bending defined as follows.

[1]Lebesgue, Intégrale, longueur, aire, Annali de Math., (3), **VII**, 315 (1902).

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

On a rectifiable surface $\bar{x}(u, v)$ let a continuous velocity field $\bar{z}(u, v)$ be given such that every rectifiable curve on the initial surface is transformed to a rectifiable curve on the deformed surface $\bar{x}(u, v) + t\bar{z}(u, v)$. In § 1 we prove that under this condition there always exists a one-sided time derivative $\frac{\partial s}{\partial t}$ of the length of a rectifiable curve. At $t = 0$, if $\frac{\partial s}{\partial t} = 0$ for every rectifiable curve on the surface $\bar{x}(u, v)$, the surface is said to admit an *infinitesimal bending* at rate $\bar{z}(u, v)$, where $\bar{z}(u, v)$ is the bending field.

A surface admitting no infinitesimal bendings other than motions is said to be *rigid*. This definition is apparently in line with the mechanical rigidity encountered in deforming a surface made of a flexible but virtually non-tensile material. In fact, even nonregular surfaces can be produced in practice. However, it is very likely that there is no straightforwardly perceptible meaning in the concept of a surface lacking tangent planes at the points of an everywhere dense set and other similar ideas that are invoked in the theory of functions of a real variable. Regrettably, at our disposal there is no sufficiently general method for studying all functions, surfaces, etc. that are arbitrarily nonregular within the limits dictated by reality. Therefore, we are impelled to follow the path that was paved by Lebesgue.

In the first two sections we derive the equation

$$dx\, d\bar{z} = 0$$

that is satisfied by a bending field almost everywhere. We then prove that a convex surface of revolution having no plane parts and whose spherical image covers the whole sphere, i.e., a closed surface or a surface having a pole and one or two parallel disks[2] lying in the tangent plane is rigid. Here only the convexity of any meridian is assumed.

§ 1. Infinitesimal Bending of a Curve

Let the end of a vector $\bar{x}(s)$ describe a continuous rectifiable curve, where s is the *arc length*. We study a bending of this curve at rate $\bar{z}(s)$ so that $\bar{x}(s) + t\bar{z}(s)$ is a continuous rectifiable curve. In other words, the function $\bar{z}(s)$ is continuous and has bounded variation. For the sake of brevity, I say that a vector function of several parameters satisfies some condition (continuity, bounded variation, absolute continuity, or Lipschitz's condition), if each of its three components, as a function of these parameters, satisfies this condition. By $s(t)$ I denote the arc length of a deformed curve. We derive an expression for $\frac{\partial s(t)}{\partial t}$ and thereby prove its existence for $t \geqslant 0$. Thus, we

[2] By a *parallel disk* is meant a disk lying in a plane orthogonal to the axis of rotation and centered at a point of the axis.—Eds.

have

$$s(t) = \int\limits_0^S |\bar{x}' + t\bar{z}'|\, ds + \sigma(s,t), \tag{1}$$

where the prime stands for differentiation with respect to s and the modulus sign for the length of a vector. The second term $\sigma(s,t)$ is the lower limit of the sum of the lengths of arcs of the deformed curve containing the points where $|\bar{x}' + t\bar{z}'|$ is infinite or, since $|\bar{x}'| \leqslant 1$, the points where $|\bar{z}'|$ is infinite.[3]

Choosing a sequence of sets of arcs $\triangle s(t)$ such that

$$\lim \sum \Delta s(t) = \sigma(s,t),$$

and recalling that the arc length is by definition

$$\triangle(s,t) = \lim \sum |\delta\bar{x} + t\delta\bar{z}|,$$

where $\delta\bar{x} + t\delta\bar{z}$ is a chord, we can write

$$\sigma(s,t) = \lim \sum \lim \sum |\delta\bar{x} + t\delta\bar{z}|. \tag{2}$$

Since, for $t > 0$,

$$t|\delta\bar{z}| - |\delta\bar{x}| \leqslant |\delta\bar{x} + t\delta\bar{z}| \leqslant t|\delta\bar{z}| + |\delta\bar{x}|;$$

clearly, we have

$$\begin{gathered} t \lim \sum \lim \sum |\delta\bar{z}| - \lim \sum \lim \sum |\delta\bar{x}| \leqslant \sigma(s,t) \\ \leqslant t \lim \sum \lim \sum |\delta\bar{z}| + \lim \sum \lim \sum |\delta\bar{x}|. \end{gathered}$$

However, $\bar{x}(s)$ is an absolutely continuous function of s and the set of the points s where $|\bar{z}'(s)|$ is infinite has zero measure because $\bar{z}(s)$ is of bounded variation. Therefore the second terms on the right and left sides vanish. Consequently, we obtain

$$\sigma(s,t) = t \lim \sum \lim \sum |\delta\bar{z}| = tv(s). \tag{3}$$

Hence, it follows that

$$\frac{\partial s(t)}{\partial t} = \lim_{\tau=0} \int\limits_0^S \frac{|\bar{x}' + (t+\tau)\bar{z}'| - |\bar{x}' + t\bar{z}'|}{\tau} ds + v(s).$$

[3]See Lebesgue: Lecons sur l'integration et la recherche des fonctions primitives (1928).

Applying the triangle axiom, we readily find that the integrand is not greater in absolute value than $|\bar{z}'(s)|$ which is summable because $\bar{z}(s)$ is of bounded variation. So we can take the limit under the integral sign and obtain[4]

$$\frac{\partial s(t)}{\partial t} = \int^{S} \frac{\bar{x}'\bar{z}' + t\bar{z}'^2}{|\bar{x}' + t\bar{z}|'} ds + v(s). \tag{4}$$

In particular, for $t = 0$ since $|\bar{x}'| = 1$ almost everywhere, we have

$$\frac{\partial s(0)}{\partial t} = \int_0^S \bar{x}'\bar{z}'\, ds + v(s). \tag{5}$$

We now require that $\bar{z}(s)$ be a bending field, i.e., $\frac{\partial s(0)}{\partial t} = 0$ for every arc. Then

$$\int_0^S \bar{x}'\bar{z}'\, ds = -v(s).$$

This integral is certainly an absolutely continuous function of s, whereas $v(s)$ suffers, as is evident from its definition (3), if it does not vanish identically, a total change on a set of zero measure where $|\bar{z}'|$ is infinite. Hence, $v(s) = 0$, and

$$\int_0^S \bar{x}'\bar{z}'\, ds = 0$$

for all s. Therefore $\bar{x}'\bar{z}'$ vanishes almost everywhere on s.

If $v(s) = 0$, then $\bar{z}(s)$ is absolutely continuous because only those points where $|\bar{z}'(s)|$ is infinite contribute to the change of $\bar{z}(s)$ on a set of arbitrarily small measure.

From the foregoing reasoning it is obvious that for $\bar{z}(s)$ to be a bending field, it is sufficient that $\bar{z}(s)$ be absolutely continuous and $|\bar{x}'\bar{z}'|$ be zero almost everywhere. Hence we have the

Theorem. *For a velocity field $\bar{z}(s)$ to be a bending field on a rectifiable curve $\bar{x}(s)$, it is necessary and sufficient that $\bar{z}(s)$ be an absolutely continuous function of the arc length s and that $\bar{x}(s)\bar{z}'(s)$ vanish almost everywhere on s.*

The conditions that a bending field obeys are linear, i.e., if $\bar{z}_1(s)$ and $\bar{z}_2(s)$ are two bending fields, $a\bar{z}_1(s) + \bar{z}_2(s)$ is also a bending field. On a given curve let there be defined a continuous family of bending fields

[4]This is the right derivative. The left derivative contains $-v(s)$.

$\bar{z}(s,v)$ dependent on a parameter $v\,(0 \leqslant v \leqslant 1)$. Let the length of the deformed curve $\bar{x}(s) + t\bar{z}(s,v)$ be uniformly bounded for every t. Then

$$\bar{z}(s) = \int\limits_0^1 \bar{z}(s,v)\,dv$$

is also a bending field on our curve. For sufficiently regular bendings of a smooth curve, the above assertion is rather trivial: everything is reduced to differentiation with respect to s under the integral sign because

$$\bar{x}'(s)\bar{z}'(s) = \bar{x}'(s)\frac{d}{ds}\int\limits_0^1 \bar{z}(s,v)dv = \int\limits_0^1 \bar{x}'(s)\bar{z}'(s,v)dv = 0,$$

as $\bar{x}'(s)\bar{z}'(s,v) = 0$ for every v. Under our assumption, this simple assertion requires certain explanation. The integral

$$\int\limits_0^S \left|\frac{\partial}{\partial s}\bar{z}(s,v)\right|\,dv,$$

being a function of v, is bounded because, by assumption, the length of the deformed curve $\bar{x}(s) + t\bar{z}(s,v)$ is bounded, being a function of v. Hence, from a well-known theorem,[5] $\left|\frac{\partial \bar{z}(s,v)}{\partial s}\right|$ is surface summable, i.e., $\frac{\partial}{\partial s}z(s,v)$ is summable in the domain of s and v. Consequently, we can assert that[6]

$$\bar{z}'(s) = \int\limits_0^1 \frac{\partial}{\partial s}\bar{z}(s,v)\,dv$$

for almost all s. Hence it is readily seen that $\bar{z}(x)$ is absolutely continuous (as the integral of its derivative) and that

$$\bar{x}'(s)\bar{z}'(s) = 0,$$

for almost all s because $\bar{z}(s)$ is indeed a bending field. This note is helpful in studying infinitesimal bendings of a surface of revolution.

[5]See, e.g., La Vallée Poussin, A Course in Analysis, vol. 2, p. 108, § 94.
[6]ibid, pp. 109–110, § 97.

§ 2. Infinitesimal Bending of a Surface

Let the end of a vector $\bar{x}(u,v)$ describe a rectifiable surface when the parameters u and v vary in some domain D. In his work cited above, Lebesgue proved that for $\bar{x}(u,v)$ to be a rectifiable surface, it is necessary and sufficient that $\bar{x}(u,v)$ satisfy the *Lipschitz condition*:

$$|\bar{x}(u_1,v_1)-\bar{x}(u_2,v_2)|<M\sqrt{(u_1-u_2)^2+(v_1-v_2)^2}.$$

Under these conditions, as shown by Rademacher,[7] the total differential

$$d\bar{x}=\bar{x}_u du+\bar{x}_v dv$$

exists almost everywhere in D.

Let a bending field $\bar{z}(u,v)$ be given on our surface. Then, as is clear from the definition of bending field, the surface $\bar{x}(u,v)+t\bar{z}(u,v)$ is rectifiable and, consequently, $\bar{z}(u,v)$ also satisfies the Lipschitz condition, and the total differential

$$d\bar{z}=\bar{z}_u du+\bar{z}_v dv$$

exists almost everywhere in D. Let us take in D three families of lines $u=\text{const}$, $v=\text{const}$ and $u-v=\text{const}$. Their images on our surface are rectifiable and are under a bending at rate of $\bar{z}$. Therefore the equation

$$\frac{d\bar{x}}{ds}\frac{d\bar{z}}{ds}=0$$

is satisfied almost everywhere on each of these lines, where s is arc. If $\frac{du}{ds}$ is finite, then

$$\frac{d\bar{x}}{ds}\frac{d\bar{z}}{ds}=\bar{x}_u\frac{du}{ds}\,\bar{z}_u\frac{du}{ds}=0,$$

and, as the Lipschitz condition asserts

$$\frac{du}{ds}\geqslant\frac{1}{M}>0,$$

$\bar{x}_u\bar{z}_u$ vanishes. However, if $\frac{ds}{du}=0$, then $\bar{x}_u$ vanishes.

Obviously, $\bar{x}_u$ vanishes almost everywhere on every set of positive measure on the line $v=\text{const}$ which is transformed to a set of zero measure on the arc s on a surface. And since $\bar{z}(u,v)$ satisfies the Lipschitz condition, $\bar{z}_u$ is bounded. Hence $\bar{x}_u\bar{z}_u$ vanishes almost everywhere on every line $v=\text{const}$.

---

[7] Rademacher, Über partielle und totale Differenzierbarkeit, Math. Ann., **81**, 52 (1920).

Introducing the parameter $w = \sqrt{u^2+v^2}$, we can also apply the above reasoning to the lines $u =$ const and $u - v =$ const. Thus we find $\bar{x}_v\bar{z}_v$ and $\bar{x}_w\bar{z}_w$ vanish almost everywhere on these lines, respectively. The derivatives $\bar{x}_u, \ldots, \bar{x}_w$ all are surface measurable in D and as our equations are satisfied almost everywhere on straight lines, they are also satisfied almost everywhere in D (in the sense of surface measure). If $\bar{x}(u,v)$ and $\bar{z}(u,v)$ have total differentials almost everywhere in D, then

$$\bar{x}_w\bar{z}_w = \bar{x}_u\bar{z}_u\left(\frac{du}{dw}\right)^2 + (\bar{x}_u\bar{z}_v + \bar{x}_v\bar{z}_u)\frac{du}{dw}\frac{dv}{dw} + \bar{x}_v\bar{z}_v\left(\frac{dv}{dw}\right)^2$$

holds almost everywhere and since

$$\bar{x}_u\bar{z}_u = \bar{x}_v\bar{z}_v = \bar{x}_w\bar{z}_w = 0$$

almost everywhere, we find that

$$\bar{x}_u\bar{z}_v + \bar{x}_v\bar{z}_u = 0$$

almost everywhere. Thus we obtain the

Assertion. *If a bending field $\bar{z}(u,v)$ is given on a rectifiable surface $\bar{x}(u,v)$, then*

(1) *$\bar{z}(u,v)$ satisfies the Lipschitz condition*

$$|\bar{z}(u_1,v_1) - \bar{z}(u_2,v_2)| < M\sqrt{(u_1-u_2)^2 + (v_1-v_2)^2}, \quad \textit{and}$$

(2) *$dxdz = 0$ almost everywhere in the domain of u and v or in fully expanded form*

$$\bar{x}_u\bar{z}_u = \bar{x}_v\bar{z}_v = \bar{x}_u\bar{z}_v + \bar{x}_v\bar{z}_u = 0$$

almost everywhere in the domain of u and v.

These conditions are, as can be easily verified by simple examples, not sufficient for $\bar{z}(u,v)$ to be a bending field. In order to obtain more refined results, I confine myself to a more particular class of surfaces than rectifiable ones. If a sequence of polygons $P_1, P_2, \ldots$ in a domain D converges to a curve L such that the length of the curve L is the limit of the lengths of the polygons P_n, then I assume that the limit of the lengths of the images of these polygons on a surface is equal to the length of the image of the curve L. A surface having this property I call *continuously rectifiable*, e.g., every surface consisting of smooth pieces on which the curves E, F and G are continuous or the surfaces of revolution (studied in the sequel) are continuously rectifiable.

For $\bar{z}(u,v)$ to be a bending field on a continuously rectifiable surface traced by the endpoint of a vector $\bar{x}(u,v)$, it is necessary and sufficient that

(1) *$\bar{z}(u,v)$ satisfy the condition*

$$|\bar{z}(u_1,v_1)-\bar{z}(u_2,v_2)|<M\rho,$$

where (u_1,v_1) and (u_2,v_2) are two arbitrary points and ρ is the geodesic distance between them on the surface, and

(2) *the equation*

$$d\bar{x}\,d\bar{z}=0$$

hold almost everywhere in D.

That the second condition is necessary has already been proved. To prove that the first condition is necessary, assume the contrary, i.e., there exists a sequence of pairs of points $A_1,B_1,\ A_2,B_2,\dots$ such that

$$\frac{|\bar{z}(A_n)-\bar{z}(B_n)|}{\rho(A_n,B_n)}$$

increases indefinitely as n tends to infinity. Clearly, we could suppose that the points A_n converge to a point A and the B_n converge to a point B, where A and B should coincide, else $|\bar{z}(u,v)|$ would be infinite. Take a sequence of positive numbers $\varepsilon_1,\varepsilon_2,\dots$ such that $\sum_{n=1}^{\infty}\varepsilon_n$ is finite. From our sequence choose a subsequence of pairs of points $A_{\alpha_1},B_{\alpha_1},\ A_{\alpha_2},B_{\alpha_2},\dots$ satisfying the following three conditions

$$\left.\begin{aligned}\rho(A,A_{\alpha_n}) &< \varepsilon_n,\\ \rho(A_{\alpha_n},B_{\alpha_n}) &< \varepsilon_n,\\ \frac{|\bar{z}(A_n)-\bar{z}(B_n)|}{\rho(A_n,B_n)} &> \frac{1}{\varepsilon_n}.\end{aligned}\right\}\tag{1}$$

Since the surface is continuously rectifiable, every pair $(A_{\alpha_n},B_{\alpha_n})$ of points can clearly be joined by an infinite number of arcs having no common point other than the endpoints and of lengths arbitrarily close to $\rho((A_{\alpha_n},B_{\alpha_n})$. Joining the points A_{α_n} and B_{α_n} by a sufficiently large number of small arcs, we can obtain a continuous curve of a length greater than ε_n but less that $2\varepsilon_n$. Our arc S_n on the deformed surface traced by the endpoint of the vector $\bar{x}(u,v)+t\bar{z}(u,v)$ has length

$$\begin{aligned}s_n &\geqslant |\bar{x}(A_{\alpha_n})+t\bar{z}(A_{\alpha_n})-\bar{x}(B_{\alpha_n})-t\bar{z}(B_{\alpha_n})|\\ &\geqslant t\big|\bar{z}\,|(A_{\alpha_n})-\bar{z}(B_{\alpha_n})|-\rho(A_{\alpha_n},B_{\alpha_n})\big|.\end{aligned}$$

By virtue of the last inequality in (1), we find

$$s_n > \left| t \frac{1}{\varepsilon_n} - 1 \right| \rho(A_{\alpha_n}, B_{\alpha_n})|.$$

Therefore the length of our curve on the deformed surface is greater than $|t\frac{1}{\varepsilon_n} - 1|\varepsilon_n$. Joining in this way all pairs of points $(A_{\alpha_n}, B_{\alpha_n})$ and then joining all A_{α_n} to A, we obtain a continuous curve on the initial surface. Its length is less than $3\sum_{n=1}^{\infty} \varepsilon_n$ because by (1) $\rho(A, A_{\alpha_n}) < \varepsilon_n$. On the other hand, the length of the corresponding curve on the deformed surface is $> \sum_{n=1}^{\infty} |t - \varepsilon_n|$, i.e., is infinite. So $\bar{z}(u, v)$ is not a bending field.

To prove that these conditions are sufficient, let us take in D a continuously rectifiable curve L. Let $s(L)$ and $s(L, t)$ denote the length of the image of the curve on the initial surface and on the deformed surface, respectively. In D construct a continuous sequence, of polygons P_r' covering a certain neighborhood of L, such that the polygons P_r converge to L and their lengths to the length of L as r tends to infinity.[8] Then, by virtue of continuous rectifiability, we have

$$\lim_{r=\infty} s(P_r) = s(L),$$

and on the deformed surface

$$\lim_{r=\infty} s(P_r, t) \geqslant s(L, t).$$

Hence, we obtain

$$\frac{s(L,t) - s(L)}{t} \leqslant \lim_{r=\infty} \frac{s(P_r, t) - s(P_r)}{t}. \tag{2}$$

Since $d\bar{x} d\bar{z}$ vanishes almost everywhere in D and the polygons P_r cover a certain set, there exists a subsequence of polygons $P_{r_1}, P_{r_2}, \dots$ such that they converge to L and that $\bar{x}'(s)\,\bar{z}'(s)$ vanishes almost everywhere on the image of these polygons on the surface. Since

$$|\bar{z}(u_1, v_1) - \bar{z}(u_2, v_2)| < M\rho, \tag{3}$$

the field $\bar{z}(s)$ is absolutely continuous and

$$|\bar{z}'(s)| \leqslant M.$$

[8] Of course, r runs through the continuum of natural numbers.

Therefore, using formula (4) of § 1, we find that

$$\frac{\partial s(P_r,t)}{\partial t}=t\int\limits_0^s\frac{\bar z'^2}{|\bar x'+t\bar z'|}\,ds\leqslant tM^2\frac{s(P_r)}{1-tM}\tag{4}$$

for $t<\frac{1}{M}$ if the denominator is positive. By the finite increment formula, we have

$$\frac{s(P_r,t)-s(P_r)}{t}=\frac{\partial s(P_r,\tau)}{\partial t}\leqslant\frac{\tau M^2 s(P_r)}{1-\tau M}.\tag{5}$$

Hence, from (2), we obtain

$$\frac{s(L,t)-s(L)}{t}\leqslant\frac{\tau M^2 s(P_r)}{1-\tau M},\tag{6}$$

and taking the limit as $t\to 0$, we obtain

$$\frac{\partial s(L,0)}{\partial t}\leqslant 0.\tag{7}$$

This inequality holds for every curve. However, by formula (5) of § 1, we have

$$\frac{\partial s(L,0)}{\partial t}=\int\limits_0^s\bar x'\bar z'\,ds.$$

Substituting $-\bar z(u,v)$ for $+z(u,v)$, we once again obtain a bending field satisfying the conditions of the theorem, but $\frac{\partial}{\partial t}s(L,0)$ changes its sign, although inequality (7) admits only like signs or zero value for this derivative. Hence, $\frac{\partial}{\partial t}s(L,0)$ vanishes definitely and our surface admits a bending on every curve L. This result leads to a surprising application. In the domain D let us take three lines $u=a$, $v=b$, and $u-v=c$, where the constants a,b and c take the values in an everywhere dense set in an interval admissible for each of them. The corresponding figure on a continuously rectifiable surface defined for D is called a *densely triangulated grid*. Let $\rho(A,B)$ and $\rho'(A,B)$ denote the geodesic distances between the points A and B on the surface and on the densely triangulated grid, respectively. For every two points A and B on the grid we require that

$$\frac{\rho'(A,B)}{\rho(A,B)}<N.\tag{8}$$

Assume that a bending field $\bar z(u,v)$ is given on a densely triangulated grid. Applying the reasoning we used to prove the above theorem, we can show that

$$|\bar z(A)-\bar z(B)|<M\rho'(A,B)$$

on the grid and, by virtue of (8)

$$|\bar{z}(A) - \bar{z}(B)| < MN\rho(A, B).$$

Hence, the bending field $\bar{z}(u, v)$ defined on the grid can clearly be extended onto the whole of the surface so that the field satisfies the first condition of the theorem. Take, for instance, a line $u = \alpha$ (not equal to any of the values chosen for a). Taking a sequence of lines $u = a_1$, $u = a_2, \ldots, \lim_{n\to\infty} a_n = \alpha$ and applying the arguments used above, we find that for every arc of the image of the line $u = \alpha$ on the surface

$$\frac{\partial s(0)}{\partial t} = 0.$$

Consequently,

$$\bar{x}_v \bar{z}_v = 0$$

almost everywhere on the line.

Similarly, $\bar{x}_u \bar{z}_u$ vanishes almost everywhere on the line $v = \beta$. Now introducing the parameter $w = \sqrt{u^2 + v^2}$, we find $\bar{x}_w \bar{z}_w$ vanishes almost everywhere on every line $u - v = \text{const}$. The equalities $\bar{x}_u \bar{z}_u = \bar{x}_v \bar{z} v = \bar{x}_w \bar{z}_w = 0$, as before, yield

$$d\bar{x}\, d\bar{z} = 0$$

almost everywhere in the domain D.

Thus the field extended onto the surface also satisfies the conditions of our theorem; consequently, it is a bending field on the surface.

Clearly, we would also arrive at the same result, if we take any densely triangulated grid other than our special grid. Thus we obtain the

Theorem. *If a bending field is defined on a densely triangulated grid satisfying condition (8) on a continuously rectifiable surface, then the field can be extended onto the entire surface. And, if the surface is rigid, then the densely triangulated grid is also rigid.*

Note. If condition (8) is not satisfied then, as is shown by particular examples, the theorem fails, i.e., the bending field on the grid may fail to induce a bending field on the surface.

A densely triangulated grid can be constructed on a surface of revolution by taking a dense set of meridians and parallels, i.e., the lines diagonal to meridians. The proof, given below to demonstrate the rigidity of closed (and certain other) convex surfaces of revolution, implies the rigidity of a densely triangulated grid as well.

§ 3. Infinitesimal Bending of a Surface of Revolution

In his paper[9] demonstrating the existence of nonrigid closed surfaces of revolution, Cohn-Vossen expanded a bending field in a Fourier series and thus reduced the problem to studying certain individual coefficients of the expansion. Incidentally, he proved that a closed convex surface of revolution is rigid, if a meridian of it has a definite curvature everywhere. I demonstrate that the Cohn-Vossen method is easily extended to every convex surface of revolution and that the following theorem holds.

Theorem. *A convex surface of revolution that has no plane pieces and whose spherical image covers the entire sphere is rigid.*

The term "spherical image" has been introduced here for the sake of brevity. For a convex surface with edges, conic points and other singularities, the concept of spherical image can be attributed the following meaning. At the interior points of the surface we draw all possible supporting planes, whereas at the points lying on its boundary we draw only those that are limits of supporting planes at interior points. The unit vectors directed along the outward normals of these supporting planes generate the spherical image of the surface, if the vectors are drawn from one point. Hence, it is clear that three types of convex surfaces of revolution are involved in our theorem, namely,

(1) closed convex surfaces,
(2) convex surfaces having a pole and one parallel disk lying in the tangent plane of the surface,
(3) convex surfaces having two parallel disks (like the outer parts of a torus).

Such a disk I call a *parabolic disk.*

Every convex surface of revolution that is a closed set and does not belong to any of the three types listed above is nonrigid.[10]

Let the endpoint of a vector $\bar{x}(u,v)$ describe a surface of revolution, where u is the height on the axis of rotation and v the angle on a parallel disk. Introduce three mutually perpendicular unit vectors $\bar{a}$, $\bar{b}$ and $\bar{c}$ at every point on the surface such that $\bar{a}$ is parallel to the axis of rotation, $\bar{c}$ is tangential to the parallel disk and is related to $\bar{b}$ by the expression:

$$\frac{\partial \bar{b}}{\partial v} = \bar{c}, \quad \frac{\partial \bar{c}}{\partial v} = -\bar{b}. \tag{1}$$

[9] Cohn-Vossen, Math. Ann., **102**, 10 (1929).

[10] A nonrigid surface is a surface on which a nontrivial bending field can be defined. The surface need not necessarily admit continuous regular bendings. For instance, an analytic surface with a parabolic disk, although not rigid, does not admit such bendings. See, Rembs, Math. Z., **36**, 10 (1932).

We consider the vectors in these rotating axes and assume that the vector $\bar{z}(u)$ depends only on $\bar{u}$, if its components on the axes $\bar{a}$, $\bar{b}$ and $\bar{c}$ do not depend on v, i.e., if the field of the vectors $\bar{z}$ on the surface is transformed into itself when rotated about the axis.

Let a bending field $\bar{z}(u,v)$ be given on a surface of revolution. Take a rectifiable curve $\bar{x}(s)$ on this surface and let a bending field $\bar{z}(s)$ be induced on the curve $\bar{x}(s)$. Rotating the bending field about the axis of rotation, we obtain a bending field $\bar{z}(u, u-v')$, where v' is the angle of rotation. This is obvious because a surface of revolution is transformed into itself when revolved about the axis of rotation. Thus we obtain a family that definitely satisfies the condition formulated at the end of § 1. Therefore,

$$\bar{z}_k(u,v) = \frac{1}{2\pi}\int_0^{2\pi} \bar{z}(u, v-v')\cos kv'\,dv$$

is a bending field. Substituting v'' for $v-v'$, we easily find

$$\bar{z}_k(u,v) = \bar{\zeta}_k(u)\cos kv + \bar{\eta}_k(u)\sin kv, \tag{2}$$

where $\bar{\zeta}_k(u)$ and $\bar{\eta}_k(u)$ are the Fourier coefficients of the function $\bar{z}(u,v)$. Similarly,

$$\bar{z}_k^*(u,v) = \frac{1}{2\pi}\int_0^{2\pi} \bar{z}(u, v-v')\sin kv'\,dv'$$

$$\bar{z}_k^*(u,v) = \bar{\zeta}_k(u)\sin kv - \bar{\eta}_k(u)\cos kv \tag{3}$$

is also a bending field.

Thus, fields of the type (2) and (3) can be obtained from a bending field defined on a surface of revolution. Since every function satisfying the Lipschitz condition can be expanded in a unique Fourier series, we are justified in confining ourselves to fields of the type (2) or (3). Calculation is simplified, on putting as usual

$$\bar{z}_k(u,v) \pm i\bar{z}_k^*(u,v) = (\bar{\zeta}_k(u) \mp i\bar{\eta}_k(u))e^{\pm ikv} = \bar{\xi}_{\pm k}(u)e^{\pm ikv}. \tag{4}$$

The terms on the right side are the kth and the $-k$th terms of the Fourier series:

$$\bar{z}(u,v) = \sum_{k=-\infty}^{+\infty} \bar{\xi}_k(u)e^{ikv},$$

whose terms each represents a bending field. Assuming that the surface

$$\bar{x}(u) = \bar{a}u + \bar{b}r(u) \tag{5}$$

is described by the vector

$$\bar{\xi}_k(u) = \bar{a}\varphi_k(u) + \bar{b}\psi_k(u) + \bar{c}\chi_k(u). \tag{6}$$

In the bending equation

$$\bar{x}_u\bar{z}_u = \bar{x}_v\bar{z}_v = \bar{x}_u\bar{z}_v + \bar{x}_v\bar{z}_u = 0, \tag{7}$$

substituting expression (5) for $\bar{x}$, and $\bar{\xi}_k(u)e^{ikv}$ for $\bar{z}$, and by virtue of (1), we obtain the relations:

$$\varphi_k'(u) + r'(u)\psi_k'(u) = 0, \tag{8}$$
$$\psi_k(u) + ik\chi_k(u) = 0, \tag{9}$$
$$ik\varphi_k(u) + r'(u)[ik\psi_k(u) - \chi_k(u)] = 0. \tag{10}$$

Now, using (9), we can eliminate $\psi_k(u)$, and thus obtain

$$\varphi_k'(u) = ikr'(u)\chi_k'(u), \tag{11}$$
$$ik\varphi_k(u) + (k^2-1)r'(u)\chi_k(u) + r(u)\chi_k'(u) = 0. \tag{12}$$

For $k = 0, +1$ and -1, these equations are easily integrated, if

$$r(u) = \int r'(u)\,du.$$

And it is a simple matter to verify that the corresponding bending fields represent translations of the surface, so that any other translation of the surface is a combination of these translations. Therefore, strictly speaking, bending is possible only if $|k| \geqslant 2$.[11] The velocity has a definite value at a pole; therefore, φ_k, ψ_k and χ_k must vanish at the poles for $|k| \geqslant 2$.

In order to study the behavior of φ_k and χ_k on a parabolic disk, let us take r as the independent variable. Then equations (11) and (12) are rewritten as follows:

$$u'(r)\varphi_k'(r) = ik\chi_k'(r), \tag{11$'$}$$
$$iku'(r)\varphi_k(r) + (k^2-1)\chi_k(r) + r\chi_k'(r) = 0. \tag{12$'$}$$

For a bending field satisfying the Lipschitz condition, $\varphi_k'(r)$ should be finite in the neighborhood of the parabolic disk, and on a convex meridian the derivative $u'(r)$ steadily tends to zero on approaching the parabolic disk.

[11] See Cohn-Vossen, loc. cit.

Hence, by virtue of (11), it is clear that ${\chi_k}'(r)$ tends to zero on approaching this parabolic disk. Therefore, from (12), we find $\chi_k = 0$ on the parabolic disk, if $k^2 - 1 \neq 0$.

Thus, for $|k| \geqslant 2$, the component $\chi_k(u)$ vanishes also at the endpoints of the interval in which u varies. Equation (12) yields

$$\chi_k(u) = -\int\limits_0^u \frac{ik\varphi_k(u) + (k^2-1)r'(u)\chi_k(u)}{r(u)}\,du.$$

And, as $r(u)$, $\varphi_k(u)$ and $\chi_k(u)$ are continuous and $r'(u)$ is monotone, the integrand has definite right and left limits for every u (possible exceptions are the endpoints of the interval at which it may happen that $r(u) = 0$ and $r'(u) = \infty$). Hence, a well-known theorem yields that the function $\chi_k(u)$ has for, every u, definite one-sided derivatives satisfying equation (12) everywhere, if $r'(u)$ in (12) is interpreted as the respective one-sided derivative. Therefore, we may confine ourselves, say, to the right derivatives, and equation (12) will hold everywhere for them. For this reason, the same is also true of equation (11).

In (12) using the substitution

$$f(u + \triangle u) - f(u) = \triangle f(u),$$

for all u and $\triangle u > 0$, we obtain

$$\begin{aligned} ik\triangle\varphi_k(u) + (k^2-1)r'(u)\triangle\chi_k(u) + (k^2-1)\chi_k(u+\triangle u)\triangle r'(u) \\ +\chi_k'(u)\triangle r(u) + r(u+\triangle u)\triangle\chi_k'(u) = 0. \end{aligned} \tag{13}$$

And equation (11) yields

$$\triangle\varphi_k(u) = \varphi_k'(u)\triangle u + \varepsilon\triangle u = ikr'(u)\chi_k'(u)\triangle u + \varepsilon\triangle u. \tag{14}$$

However,

$$\triangle\chi_k(u) = \chi_k'(u)\triangle u + \delta\triangle u \tag{15}$$

and

$$\triangle r(u) = r'(u)\triangle u + \gamma\triangle u. \tag{16}$$

Substituting (14), (15) and (16) into (13), and collecting the like terms, we obtain

$$(k^2-1)\chi_k(u+\triangle u)\triangle r'(u) + r(u+\triangle u)\triangle\chi_k'(u) = \alpha\triangle u, \tag{17}$$

where $\alpha = -ik\varepsilon - (k^2-1)r'(u)\delta - \chi_k'(u)\gamma$ tends to zero as $\triangle u$ tends to zero.

Since each meridian is convex, we have $\triangle r'(u) \leqslant 0$ and $r(u + \triangle u) > 0$. For the sake of definiteness, putting $\chi_k(u) > 0$, we obtain[12]

$$\chi_k(u + \triangle u) > 0$$

for sufficiently small $\triangle u$.

If $\frac{\triangle r'(u)}{\triangle u}$ is less than some negative number, then $\frac{\chi_k'(u)}{\triangle u} > 0$. If $\frac{\triangle r'(u)}{\triangle u}$ tends to zero as $\triangle u$ tends to zero by some law, then $\frac{\triangle \chi_k(u)}{\triangle u}$ also tends to zero. Consequently,

$$\varliminf_{\triangle u = +0} \frac{\triangle \chi_k'(u)}{\triangle u} \geqslant 0. \tag{18}$$

Since $\triangle u > 0$, and we are studying the right derivative, we do not know anything about the jumps of $\chi_k'(u)$. Let $r'(u)$ be discontinuous at the point u_0. [From (12) it is clear that $\chi_k'(u)$ is continuous at the points where $r'(u)$ is continuous.] Using (12), let us calculate the right and left derivatives at the point u_0 and then subtract the latter from the former. Thus, we obtain an equation for the jumps $\triangle r'(u_0)$ and $\triangle \chi_k'(u)$:

$$(k^2 - 1)\chi_k(u_0)\triangle r'(u_0) + r(u_0)\triangle \chi_k'(u_0) = 0. \tag{19}$$

Since $\triangle r'(u_0) < 0$, assuming that $\chi_k'(u_0) > 0$, we obtain

$$\triangle \chi_k'(u_0) > 0. \tag{20}$$

As the left derivative for this case is the left limit of the right derivative, expression (20) implies that at every point where $\chi_k'(u)$ suffers a jump, the jump is positive for positive $\chi_k(u)$. Hence, by virtue of (18), we find that $\chi_k'(u)$ is a nondecreasing function, if $\chi_k(u) > 0$.

The function $\chi_k(u)$ vanishes at the left endpoint of the variation interval of u. Let $\chi_k(u) > 0$ at some interior point, then there exists a point u_0 where $\chi_k(u_0) > 0$ and $\chi_k'(u_0) > 0$. Since $\chi_k'(u)$ is nondecreasing, we have $\chi_k'(u) > \chi_k'(u_0)$, if $u > u_0$. Consequently, $\chi_k(u)$ cannot vanish at the right endpoint of the variation interval of u. This is contrary to what we proved above. Hence, for $|k| \geqslant 2$, the function $\chi_k(u)$ vanishes everywhere. And by virtue of (9) and (10), we have $\varphi(u) = \psi_k(u) = 0$, i.e., bending is virtually impossible.

Without demonstrating that a convex surface of revolution is nonrigid if it is a closed set not belonging to any of the three types of surfaces

[12] We take either the real or the imaginary part.

listed above, I only mention that the proof is based on the following two observations:

(1) *For a system of explicit linear differential equations that are solved with respect to derivatives to have a solution with absolutely integrable derivatives for any initial value, it is necessary and sufficient that the coefficients of the system be absolutely integrable. This lemma must be applied to the equations for* $\varphi_k(u)$ *and* $\psi_k(u)$ *derived from (11) and (12).*

(2) *The existence of solutions that vanish at poles or on a parabolic disk can be demonstrated by taking the solutions of* $\varphi_k(u)$ *and* $\psi_k(u)$ *that do not vanish at poles or a parabolic disk and then searching for the required solution in the form in which any other linearly independent solution is expressed by certain known formula. For instance, if* $y' = py + qz$, $z' = ry + sz$ *and* (y_1, z_1) *is a system of solutions, another solution is expressed as follows*:

$$y_2 = y_1 \int_0^x \frac{q}{y_1^2} e^{\int (p+s)\,dx} dx, \quad z_2 = z_1 \int_0^x \frac{q}{y_1^2} e^{\int (p+s)\,dx} dx + \frac{e^{\int (p+s)\,dx}}{y_1}.$$

Interestingly, using these two propositions, we can construct a bounded velocity field satisfying the condition $d\bar{x}d\bar{z} = 0$ on a surface having a pole and one or two parabolic disks, provided $r(u)^2$ is finitely integrable over the interval containing the value of u to which one of the parabolic disks corresponds. This is in general possible, although $r'(u)^2$ is infinite over there; for instance, if the curvature of a meridian increases rapidly to infinity. This does not however imply that such a surface is necessarily deformable.

An example of this situation is a cuspidal surface generated by a convex arc rotating about the tangent at one of its endpoints. A continuous bounded velocity field satisfying the condition $d\bar{x}d\bar{z} = 0$ can be defined on such a surface. However, by the condition stated in § 2, if s is an arc of a parallel disk, then

$$\left|\frac{\partial \bar{z}}{\partial s}\right| = \frac{1}{r}\left|\frac{\partial \bar{z}}{\partial v}\right| \leqslant M.$$

For $\bar{z}(u,v) = [\bar{a}\varphi_k(u) + \bar{b}\psi_k(u) + \bar{c}\chi_k(u)]e^{ikv}$, the above inequality yields

$$|\chi_k(u)| \leqslant Mr(u).$$

If $r(u)$ is an analytic function, then taking the limit $\Delta u \to 0$ in (17), we obtain

$$\chi_k''(u) + (k^2 - 1)\frac{r''(u)}{r(u)}\chi_k(u) = 0.$$

Assuming u vanishes at the cusp point, we obtain

$$\frac{r''(u)}{r(u)} = \frac{a}{u^2}(1 + f(u)),$$

where $f(u)$ is an analytic function vanishing for $u = 0$ and $a > 1$. Putting, as usual,

$$\chi_k(u) = u^\rho \sum_{n=0}^{\infty} a_n u^n,$$

we derive an equation for ρ:

$$\rho(\rho - 1) + (k^2 - 1)a = 0.$$

From this equation it follows that for $|k| \geqslant 2$, the function ρ is complex valued and its real part is equal to $\frac{1}{2}$. So $\chi_k(u)$ increases when u tends to zero as $\sqrt{u}$. However,

$$r(u) = u^2 \sum_{n=0}^{\infty} c_n u^n.$$

Therefore for sufficiently small u,

$$|\chi_k(u)| > Mr(u),$$

which contradicts the inequality derived above. Thus, a cuspidal surface is an exceptional example of a rigid, although arbitrarily small surface.

Institute of Mathematics,
Leningrad State University

Received December 1, 1935

CHAPTER II

AN ELEMENTARY PROOF OF THE MINKOWSKI AND SOME OTHER THEOREMS ON CONVEX POLYHEDRA

(*Communicated by Academician I. M. Vinogradov*)

Izvestiya Akademii Nauk SSSR, No. 4, 597–606 (1937).

An elementary proof is given to the Minkowski theorem stating that a convex polyhedron is completely determined by the areas and the directions of its faces. Additionally, a convex polyhedron has been proved to be completely determined also by the directions and the perimeters of its faces.

In the introduction to his paper *Allgemeine Lehrsätze über die konvexen Polyeder* [1], Minkowski wrote:

> Der vorliegende Aufsatz entstand bei Gelegenheit von Versuchen den folgenden Satz zu beweisen, den ich seit längerer Zeit vermutete und dessen elementare Fassung nicht auf die Schwierigkeiten seiner Verifizierung schliessen lässt: Wenn aus einer endlichen Anzahl von lauter Körpern mit Mittelpunkt, die unter einander nur in den Begrenzungen zusammenstossen, sich ein konvexen Körper aufbaut, so hat dieser stets ebenfalls einer Mittelpunkt.

He derived this theorem as a corollary to a more general theorem, viz., the main result of the mentioned paper [1] on polyhedra:

A convex polyhedron is uniquely determined, up to translations, by the directions and the areas of its faces. By the *direction of a face*, we mean the direction of the outward normal to the face.

The proof of this theorem, as given by Minkowski, is based on the Brunn inequality which needs by far nonelementary and intricate reasoning for its demonstration. This fact, noted by Minkowski in the excerpt quoted above, distinguishes his theorem from the other results of the theory of polyhedra.

Therefore, a need remains for a proof of this theorem which would be as elementary as the statement of the theorem itself.

In the present article I resolve the problem in a rather unexpected way. Basing on the long-known results in the theory of polyhedra, I prove in an elementary manner even a more general theorem from which the Minkowski theorem follows as a particular case, but only for three-dimensional polyhedra.

The relationship between the Minkowski theorem and the Brunn inequality is reversible: using the Minkowski theorem, I demonstrate that the Brunn inequality holds for convex polyhedra.

§ 1. Basic Elements of the Method

The well-known *Euler theorem* is fundamental for the theory of polyhedra. Basing on an extension of this theorem, Cauchy proved another elegant topological theorem concerning polyhedra which he applied to demonstrating that convex polyhedra composed of the same number of equal and similarly-placed faces are congruent.

The *Cauchy theorem*, which we are dealing with, asserts that it is impossible for every edge of a convex polyhedron be assigned a positive or negative sign or zero in such a way that at least one edge has a definite sign (but not zero) and not less than four sign changes occur in passage around each face not all of whose edges are assigned zero.[1]

It is this theorem that serves a principal idea behind my method. Thus, the Cauchy idea has reaffirmed its potentiality as a powerful proving tool: I therefore refer to it as the *Cauchy principle*.

Another important technique of my method is the *mixing* of two convex polyhedra or polygons. This operation, introduced by Brunn, consists in the following. Let two convex polyhedra H_1 and H_2 (or two convex polygons in parallel planes) be given. Join each point of one of them (including its interior points) with every point of the other by a straight line segment and take the locus of the midpoints of these segments. The locus is also a convex polyhedron (or a polygon lying in a plane parallel to the planes of the polygons under mixing). This polyhedron is denoted by $\frac{1}{2}(H_1 + H_2)$ [4, 5].

The main property of mixing which we need is that every boundary element (a face, edge, or vertex) of the polyhedron $\frac{1}{2}(H_1 + H_2)$ results from mixing some boundary elements of H_1 and H_2 lying in parallel supporting planes. (By *parallel supporting* planes, supporting lines or faces, we mean

[1]See [2]. As is known, Cauchy himself did not give any proof of this theorem. For a complete proof see, e.g., [3]. Usually, the Cauchy theorem is formulated differently: in place of passage around a face, in the Cauchy theorem we pass around the vertices. This does not however make any difference, as is clear on taking the dual of a polyhedron.

such planes, lines or faces with parallel outward normals.) A face of the polyhedron $\frac{1}{2}(H_1 + H_2)$ is obtained by mixing a pair of parallel faces, or by mixing a face of one polyhedron with an edge or vertex of the other, or by mixing a pair of nonparallel edges of H_1 and H_2 lying in parallel supporting planes. An edge of the polyhedron $\frac{1}{2}(H_1 + H_2)$ is obtained by mixing either a pair of parallel edges or an edge of one polyhedron with a vertex of the other lying in parallel supporting planes.

On mixing two polygons P_1 and P_2, we obtain a polygon $\frac{1}{2}(P_1 + P_2)$ whose sides are equal to the half-sums of parallel sides of P_1 and P_2. If a side of one of the polygons has no corresponding parallel side, then we assume that it does exist but its length is zero. In the sequel we presume this without any special mention.

§ 2. A Theorem on Convex Polygons

A polygon P_1 is said to *be embedded in* a polygon P_2, if all the points of P_1 lie in P_2. If, in addition, at least one side of P_1 lies in the interior of P_2, then the polygon P_1 is said to *be embedded inside* P_2. In studying the sides of two polygons, we only compare those sides whose outward normals are parallel (tacitly assuming the condition that governs a side of zero length). Finally, the sides $l_1, \dots, l_n$ of a polygon are said to be greater than the sides $l'_1, \dots, l'_n$ of another polygon, if $l_1 \geqslant l'_1, \dots, l_n \geqslant l'_n$ and for at least one pair $l_i > l'_i$.

Theorem. *If two convex polygons P_1 and P_2 are such that one is not embeddable in the other by translation, then the signs of the differences of the lengths of their sides change at least four times in passage around each of these polygons.*

Lemma A. *If all sides of a polygon P_1, except one side l_0, are less than the sides of a polygon P_2, the polygon P_1 can be embedded inside the polygon P_2 by translation.*

Take a pair of corresponding vertices A_1 and A_2 in P_1 and P_2 such that the supporting lines with outward normal antiparallel to the normal to the side l_0 pass through A_1 and A_2. Now superpose the polygons so that the vertex A_1 coincides with the vertex A_2. The vertex A_1 and the side l_0 then subdivide the polygon P_1 into two subpolygons P'_1 and P''_1. Similarly the polygon P_2 is subdivided into two subpolygons P'_2 and P''_2. The polygon P'_1 does not protrude from the polygon P_2 through P'_2, for, as can be easily verified, the polygon P'_1 would have a side greater than a side of the polygon P'_2. Similarly the polygon P''_1 does not protrude from the polygon P_2 through P''_2. Therefore the polygon P_1 is embedded inside P_2.

Lemma B. *Let polygons P_1 and P_2 have a pair of common supporting lines intersecting at a point O. Let P_1' and P_2' be those parts of the boundaries of P_1 and P_2 whose convexity looks out onto the point O. If a ray from O intersects the polygon P_1' before intersecting the polygon P_2', then the polygon P_1' has a side smaller than the corresponding side of the polygon P_2'.*

Proof. To prove this lemma, let us contract the polygon P_2' to the point O so that it always remains similar to itself. At the instant when P_2' lies within the angular region having vertex at O and bounded by the polygon P_1' and is still in contact with the polygon P_1', each common side of P_2' is not smaller than the corresponding side of P_1'. This completes the proof of our lemma.

Let P_1 and P_2 be two polygons that satisfy the conditions of our theorem. If every side of the polygon P_1 is smaller than the side of P_2, then the polygon P_1 can be embedded inside P_2. Therefore the signs of the differences of the lengths of their sides change at least twice. Assume that there are only two sign changes. Of course, we could as well assume that, for a given direction of reckoning the angles, the angle between the normals to the extreme sides, which are greater on the polygon P_1 than on the polygon P_2, is less than π (if not, let us interchange the numberings of the polygons). Let us subdivide the boundaries of the polygons P_1 and P_2 into P_1', P_1'' and P_2', P_2'' such that the sides of P_1' are less than the sides of P_2' and vice versa for the polygons P_1'' and P_2''.

By Lemma A, the polygon P_1' can be embedded inside P_2. By the condition of the theorem, however, the polygon P_1'' protrudes from the polygon P_2. Take a point on P_1'' where there is a supporting line that does not intersect the polygon P_2. Moving this point over the polygon P_1'' and thereby rotating the supporting line first in one direction and then in the opposite direction, we obtain a pair of common supporting lines a and b of the polygons P_1 and P_2 that touch the polygon P_1' at points belonging to P_1''.

Since the angle between the normals to the extreme sides of the polygon P_1'' is less than π, the polygon P_1'' is convex relative to the point O, the point of intersection of the lines a and b. The points where the lines a and b touch the polygon P_2 are farther away from the point O than the points where lines touch the polygon P_1. These points of contact however belong to the polygon P_2''. (This is evident, because the sides of P_1' and P_2' are parallel and the polygon P_1' lies inside the polygon P_2.) Hence, by lemma B, the sides of the polygon P_2'' are greater than the sides of P_1''. This completes the proof.

Remark. This theorem and its proof can be readily extended to any closed convex curves. Taking a closed convex curve P and an arc φ on

the unit circle E, let us denote by $l(\varphi)$ the length of the arc on P that consists of those points through which pass the supporting planes with outward normals directed into φ when drawn from the center of the circle E. Our theorem, when generalized, reads as follows.

If there are two closed convex curves P_1 and P_2, of which one cannot be embedded in the other by translation, then the unit circle is subdivided into at least four arcs φ_k such that the sign of the difference $l_1(\varphi_k) - l_2(\varphi_k)$ changes in passing from one arc to its adjacent arc (where $l_1(\varphi_k)$ and $l_2(\varphi_k)$ are the lengths of the arcs of P_1 and P_2, respectively).

For twice differentiable curves, the length $l(\varphi)$ is the integral of the radius of curvature over the arc. Therefore the "theorem of four vertices" (Vierscheitelsatz) follows directly from the theorem stated above.

§ 3. Proof of the Minkowski Theorem

Theorem. *If to each face of a convex polyhedron there corresponds a face of another convex polyhedron with parallel outward normal and vice versa and if their corresponding faces are such that one cannot be embedded inside the other by translation, then the polyhedra are translates of one another.*

(If one of the polyhedra has no side with the same outward normal as the outward normal of the corresponding side of the other, then this side is supposed to exist but degenerate into an edge lying in the corresponding supporting plane.)

Let H_1 and H_2 be two polyhedra satisfying the conditions of our theorem. Construct the polyhedron

$$H = \frac{1}{2}(H_1 + H_2).$$

Each edge of the polyhedron H is generated either by mixing a pair of parallel edges lying in parallel supporting planes of H_1 and H_2, or by mixing an edge of one polyhedron with the vertex of the other, where the edge and the vertex under mixing lie in parallel supporting planes. In the latter case, we can also suppose that an edge of the polyhedron H is generated by mixing the edges of H_1 and H_2, the length of only one edge being zero. Under this assumption, we can assert that to every edge of the polyhedron H there corresponds one edge of H_1 and H_2 each, which when mixed generate the edges of H.

To each edge of the polyhedron H let us now assign either plus or minus sign, depending on whether the corresponding edge of H_1 is longer or shorter than the edge of H_2. In case the edges of H_1 and H_2 are equal, the

corresponding edge of H is assigned zero. (Recall the condition concerning an edge of zero length.)

We now prove that, in passing around each face of the polyhedron H with at least one edge labelled, we meet at least four sign alterations.

There are two kinds of faces of the polyhedron H. A face of the first kind is obtained by mixing parallel faces of the polyhedra H_1 and H_2. A face of the second kind is obtained by mixing a pair of nonparallel edges lying in parallel supporting planes of the polyhedra H_1 and H_2. Other cases are excluded because the faces of the polyhedra H_1 and H_2 are supposed to be pairwise parallel.

Since parallel faces of H_1 and H_2 cannot be embedded one into another by translation; therefore, by the theorem proved in § 2, not less than four sign alterations occur in passage around each face of the first kind, if at least one edge is assigned a definite sign.

Each face P of the second kind, being the result of mixing two nonparallel edges, is a parallelogram. Let the respective edges L_1 and L_2 of the polyhedra H_1 and H_2 generate a face P of the second kind. There is no edge parallel to L_1 in the supporting plane of H_2 that is parallel to the plane of the face P, otherwise the supporting plane would include a whole face, because the edge L_2 lies in this plane. The edge L_1 is mixed with two vertices that are the endpoints of the edge L_2 and generates two opposite sides of the parallelogram P. The same is true of the edge L_2. Hence, exactly four alterations of sign occur in passage around each face of the second kind.

Now, by the Cauchy principle, we find that zero should be assigned to all the edges of the polyhedron H. This implies that the parallel faces of the polyhedra H_1 and H_2 are congruent. Therefore the polyhedra themselves are translates of one another.

If the areas of the corresponding parallel faces of the polyhedra H_1 and H_2 are equal, then these faces cannot be embedded one inside the other. Therefore the Minkowski theorem is a particular case of our theorem. In the same way we obtain the following theorem.

If to each face of a convex polyhedron there corresponds a face of another with parallel outward normal and with the same perimeter, and vice versa, then the polyhedra are translates of one another.

§ 4. Rigidity of a Convex Polyhedron with Stationary Directions and Areas of Faces

Let us choose an origin in a plane and draw n rays so that they are not directed into one half-plane. The straight lines perpendicular to the rays and intersecting them generate a convex polygon. We call these lines the *boundary lines* of the polygon.

Lemma. *If the boundary lines of a polygon are subject to infinitesimal displacements such that the area of the polygon is stationary (i.e., does not vary up to infinitesimals of second order), then the variations of the lengths of the sides of the polygon suffer at least four sign changes in passage around the polygon, provided the resulting deformation of the polygon is not an infinitesimal translation. Moreover, if a new side appears, it is supposed to increase in length. (A new side may appear, if one of the boundary lines touches the initial polygon at a vertex but not on an edge.)*

Let dl_i $(i = 1, \ldots, n)$ denote the variations of lengths of the sides of the polygon. Assume that dl_i changes its sign only twice in passage around the polygon. Let the vertices A_1 and A_2 separate the elongating sides from the contracting ones. Take the origin at the intersection of two supporting lines passing through the vertices A_1 and A_2. Then the support numbers h_i of the polygon change their signs in passage from the vertex A_1 to the vertex A_2, whereas the product $l_i dl_i$ does not. Therefore

$$\sum_{i=1}^{n} h_i \, dl_i \tag{1}$$

vanishes if and only if all dl_i vanish. However, the sum (1) is the differential of the area of the polygon and, consequently, is zero[2] by the condition of the theorem.

Therefore, either all dl_i vanish and the polygon is not deformed but is only displaced, or the differentials dl_i change their sign more than twice, i.e., not less than four times.

Theorem. *If the planes of the faces of a convex polyhedron are subject to infinitesimal displacements such that the areas of the faces are stationary, then the polyhedron itself suffers only an infinitesimal translation.*

The lines formed by the intersection of the plane of a given face P with the planes of other faces are the boundary lines of the face P. Under infinitesimal displacements of the faces, the boundary lines also suffer displacement.

[2]The area of a polygon is $S = \frac{1}{2}\sum_{i=1}^{n} h_i l_i$; $dS = \frac{1}{2}\sum_{i=1}^{n} h_i dl_i + \frac{1}{2}\sum_{i=1}^{n} l_i dh_i$. However, $dS = \sum_{i=1}^{n} l_i dh_i$ because $l_i dh_i$ is the area of the figure covered by the ith side upon displacement dh_i. Hence, $dS = \sum_{i=1}^{n} h_i dl_i$.

So, if the face P does not undergo simple translation and its area remains stationary, then, by the previous lemma, the variations in the lengths of its edges change their sign at least four times in passage around the face. (The condition concerning a newly generated edge is the same as in the lemma.)

If a plus is assigned to an elongating side, a minus to a contracting side and zero to a side of invariable length, then by the Cauchy principle, we find that zero should be assigned to all the edges. This implies that the lengths of all edges are stationary and, consequently, the polyhedron suffers only an infinitesimal translation.

The theorem just proved relates to the Minkowski theorem in the same way as the rigidity theorem for a convex polyhedron with stationary faces relates to the Cauchy congruence theorem for convex polyhedra composed of equal and similarly-placed faces.[3]

§ 5. The Brunn-Minkowski Inequality for Convex Polyhedra

Let H and L be two convex polyhedra. Joining each point of one of them with every point of the other by a straight line segment and taking the locus of the points dividing these segments in the ratio $\vartheta : 1-\vartheta$ $(0 \leqslant \vartheta \leqslant 1)$, we obtain a polyhedron $K_\vartheta = (1-\vartheta)H + \vartheta L$. As shown by Minkowski, its volume is

$$\begin{aligned} V(K_\vartheta) = (1-\vartheta)^3\, V(H,H,H) + 3(1-\vartheta)^2\,\vartheta V(H,H,L) \\ +3(1-\vartheta)\,\vartheta^2\, V(H,L,L) + \vartheta^3\, V(L,L,L), \end{aligned} \tag{1}$$

where $V(H,H,H)$ and $V(L,L,L)$ are the volumes of the polyhedra H and L, respectively; $V(H,H,L)$ and $V(H,L,L)$ are their mixed volumes:

$$V(H,H,L) = \frac{1}{3}\sum L_i\, F_i(H), \tag{2}$$

$$V(H,L,L) = \frac{1}{3}\sum H_i\, F_i(L). \tag{3}$$

Here L_i stands for the distance of the origin from the supporting plane of L parallel to the plane of the ith face of the polyhedron H, and $F_i(H)$ the area of the ith face of the polyhedron H. The symbols H_i and $F_i(L)$ have similar meaning.

If we could prove the following two *Minkowski inequalities*

[3]The corresponding rigidity theorem for a polyhedron with stationary perimeters and directions of its faces does not differ from the theorem formulated at the end of Section 3 because the variations in the perimeters are linearly dependant on the displacements of the faces. For this reason, the condition of their stationarity is equivalent to exact constancy.

$$\left.\begin{aligned} V(H,H,L)^3 &\geqslant V(H,H,H)^2\, V(L,L,L), \\ V(H,L,L)^3 &\geqslant V(H,H,H)\, V(L,L,L)^2, \end{aligned}\right\} \tag{4}$$

with equality holding only if H and L are homothetic. Thereby we would automatically prove the following assertion:

If the volumes of two polyhedra H and L are equal, then the volume of the polyhedron K_ϑ is not less than the volume of either. And the volume of the polyhedron K_ϑ is equal to these volumes, if and only if the polyhedra H and L are translates of one another.

This is precisely the *Brunn inequality* combined with Minkowski's supplement.

Let H be a given convex polyhedron and L a variable polyhedron whose faces are parallel to the faces of H. Let $L_1, \dots, L_i, \dots, L_n$ be the support numbers of L, i.e., the distances of the planes of the faces from the origin. These numbers, of course, completely determine the polyhedron.

Let us consider the following function of the support numbers L_i:

$$\Phi(L) = \frac{V(H,H,L)^3}{V(L,L,L)}. \tag{5}$$

By virtue of (2), it is a homogeneous function of degree zero. Therefore it is the same for all polyhedra L homothetic to one another. So it is sufficient to study this function for polyhedra of equal volumes.

If the ith support number L increases unboundedly, then the mixed volume $V(H,H,L)$, by virtue of formula (2), also increases unboundedly; but by the condition of the theorem, the volume $V(L,L,L)$ remains constant. Hence, $\Phi(L)$ attains its minimum at a bounded polyhedron.

The function $\Phi(L)$ cannot attain a minimum at a polyhedron, if at least one of its faces degenerates into an edge or vertex. Let the ith face degenerate into an edge or vertex. Move its plane inside the polyhedron L through a small distance δL_i. Then the change in its volume $\delta V(L,L,L)$ is of the order of $\delta L_i F_i(L)$, where $F_i(L)$ is the area of the face generated as a result of displacement. Therefore the volume decreases by an infinitesimal of an order not less than two. However, the reduction in the volume $V(H,H,L)$ is equal to $\delta L_i F_i(H)$ (this is clear from (2)) and, consequently, the reduction is equal to an infinitesimal of order one. Hence, the function $\Phi(L)$ decreases when the plane of a vanishing face is displaced.

Consequently, the function $\Phi(L)$ attains a minimum at a polyhedron L, all whose faces are nondegenerate and are parallel to the faces of the polyhedron H. Therefore the function $\Phi(L)$ has partial derivatives with respect to all L_i at the minimum point. And these partial derivatives must

equal zero. Clearly,

$$\frac{\partial V(L,L,L)}{\partial L_i} = F_i(L), \tag{6}$$

and, by formula (2),

$$\frac{\partial V(H,H,L)}{\partial L_i} = \frac{1}{3}F_i(H). \tag{7}$$

Therefore, calculating the derivatives $\frac{\partial \Phi(L)}{\partial L_i}$ and equating them to zero, we obtain

$$F_i(L) = \lambda F_i(H) \qquad (i = 1,\dots,n), \tag{8}$$

where

$$\lambda = \frac{V(L,L,L)}{V(H,H,L)}. \tag{9}$$

Equality (8) shows that the areas of parallel faces of the polyhedra H and L are proportional to one another. Therefore, by virtue of the Minkowski theorem, the polyhedra H and L are homothetic. In this case,

$$\Phi(L) = \frac{V(H,H,L)^3}{V(L,L,L)} = V(H,H,H)^2,$$

and since this is the minimum value of $\Phi(L)$, we have

$$V(H,H,L)^3 \geqslant V(H,H,H)^2\, V(L,L,L).$$

We proved this inequality only for a polyhedron L having the same faces as the polyhedron H. For two arbitrary polyhedra H and L and for any ϑ $(0 < \vartheta < 1)$, the polyhedra $(1-\vartheta)H + \vartheta L$ and $(1-\vartheta)L + \vartheta H$ have parallel faces. Therefore inequalities (4) hold for them. When ϑ tends to zero, these inequalities are transformed into the inequalities that hold for the polyhedra H and L.

If the equality holds in (4), then, as it follows from (1), the volumes of the polyhedra K_ϑ are equal, provided the volumes of H and L are equal. For $\vartheta_1 < \vartheta < \vartheta_2$, the polyhedron K_ϑ is

$$K_\vartheta = \frac{\vartheta_2 - \vartheta}{\vartheta_2 - \vartheta_1} K_{\vartheta_1} + \frac{\vartheta - \vartheta_1}{\vartheta_2 - \vartheta_1} K_{\vartheta_2},$$

and its volume is equal to the volume of the polyhedron K_{ϑ_1} or K_{ϑ_2}. The latter have parallel faces; consequently, they are homothetic. Hence the polyhedra H and L are also homothetic.

Institute of Mathematics and Mechanics,
Leningrad State University

Received April 28, 1937

References

[1] Minkowski H., *Allgemeine Lehrsätze über die konvexen Polyeder*, Ges. Abh.; Bd. II.
[2] Cauchy A., *Sur les polygones et les polyedres. Second memoire*, J. Ecole Polyt.; **9** (1813), 87.
[3] Cohn-Vossen S., *Bendability of a surface in the large*, Uspekhi Mat. Nauk, **I** (1936), 39–40; in Russian.
[4] Minkowski H., *Theorie der konvexen Körper*, Ges. Abh.; Bd. II, Teubner (1911).
[5] Bonnesen T. and Fenchel W., *Theorie der konvexen Körper*, Springer (1934).

CHAPTER III

TO THE THEORY OF MIXED VOLUMES OF CONVEX BODIES

Matematicheskiĭ Sbornik, vol. 2 (44), No. 5, 947–972 (1937).

In this paper we study convex bodies in n-dimensional Euclidean space, always denoting by n the dimension of the space.

By a *convex body* we mean a closed bounded set that includes, together with every pair of its points, also the entire segment joining them. By this definition, a convex body may have (in certain cases) no interior points; in this case it is said to be *degenerate*.

Brunn, and after him, Minkowski introduced the operation of *mixing* in the theory of convex bodies, which is defined as follows.

Let $H_1, \dots, H_m$ be convex bodies and let $\lambda_1, \dots, \lambda_m$ be nonnegative numbers. Choose an origin and draw vectors $\bar{x}_1, \dots, \bar{x}_m$ from it to the points of the bodies $H_1, \dots, H_m$. When the endpoints of these vectors vary independently of each other on the corresponding convex bodies, the endpoint of the vector $\lambda_1\bar{x}_1 + \cdots + \lambda_m\bar{x}_m$ describes some set of points. It is a simple matter to prove that this set is a convex body expressed as

$$H = \lambda_1 H_1 + \cdots + \lambda_m H_m.$$

When the numbers $\lambda_1, \dots, \lambda_m$ vary, always remaining nonnegative, the body H also varies and, as shown by Minkowski, its volume is a homogeneous polynomial of degree n in $\lambda_1, \dots, \lambda_m$:

$$V(H) = \sum_{k_1,\dots,k_n} \lambda_{k_1} \dots \lambda_{k_n} V_{k_1,\dots,k_n},$$

where summation is taken over all the subscripts $k_1, \dots, k_n$ running independently of one another from 1 to m. The quantities $V_{k_1,\dots,k_n}$ are defined such

that they do not depend on the order of occurrence of subscripts. Assuming that all the λ's except those involved in a given product $\lambda_{k_1} \dots \lambda_{k_n}$ are zero, we easily find that $V_{k_1,\dots,k_n}$ depends only on the bodies $H_{k_1}, \dots, H_{k_n}$. It is therefore natural to write it as $V(H_{k_1}, \dots, H_{k_n})$. It is called the *mixed volume*. Hence, the mixed volume of convex bodies in the n-dimensional space generally depends on n bodies. Thus, we defined some functional that depends on n convex bodies, namely, their mixed volume $V(H_1, \dots, H_n)$.

In the sequel we assume that the properties of convex bodies and their mixed volumes are known and also we do not always mention where their proofs were given by Minkowski.[1]

Using the Brunn inequality, Minkowski derived a number of inequalities between mixed volumes and demonstrated that they give the solutions to certain extremal problems concerning convex bodies. He also proved, using the same Brunn inequality, the uniqueness of a convex polyhedron with given normals and areas of faces as well as the uniqueness of a convex body with the Gaussian curvature of its surface[2] given as a function of its outward normal.[3] In this way Minkowski discovered a new powerful tool for studying convex bodies. For instance, it is worth noting that Weyl's proofs of the rigidity of a convex polyhedron and of a regular closed convex surface are indeed based on the use of the results concerning mixed volumes. It is therefore natural to call the problems associated with Minkowski's concepts and methods as the *theory of mixed volumes of convex bodies*.

The aim of this paper is to elaborate this theory further. It is subdivided into four parts: Part I is presented below and the remaining three parts will be published in the coming issues of MATEMATICHESKIĬ SBORNIK. Their contents are briefly outlined below.

Part I. Extension of Certain Concepts of the Theory of Convex Bodies

So far many theorems on convex bodies were established either for a polyhedron or for a body with a sufficiently regular surface. However, other convex bodies, which are "nasty," also deserve attention. Therefore it is reasonable to modify and extend the concept of Gaussian curvature in particular and that of curvature function in general (*Krümmungsfunktion*, an elementary symmetric function of the principal radii of curvature) so that

[1]Minkowski, *Theorie der konvexen Körper. Volumen und Oberfläche,* Ges. Abh., Bd. II. See also Bonnesen and Fenchel, *Theorie der konvexen Körper,* Ergebnisse der Math. (1934).

[2]Uspekhi Matematicheskikh Nauk, vol. II (1936). Delaunay, *A proof of the Brunn-Minkowski inequality*. Lyusternik, *Application of the Brunn-Minkowski inequality to extremal problems*. Minkowski, *Allgemeine Lehrsätze über die konvexen Polyeder*.

[3]A *normal* is tacitly assumed to be of length 1.—Eds.

they could be applied not only to regular bodies but also to arbitrary convex bodies.

This is just the main objective of Part I. The results of the next two parts are formulated and proved in terms of the generalized concepts thus introduced, so they can be applied to every convex body.

Part II. New Inequalities for Mixed Volumes and Their Applications

The inequality between mixed volumes

$$V(H_1, \dots, H_n)^2 \geqslant V(H_1, \dots, H_{n-1}, H_{n-1})\, V(H_1, \dots, H_{n-2}, H_n, H_n)$$

is proved. Several other inequalities are derived from it. For the mixed volumes

$$V(\underbrace{H, \dots, H,}_{m} E, \dots, E) = V_m(H),$$

where E is the unit ball, a theorem similar to the Brunn-Minkowski theorem has been derived.

If $H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1$ $(0 \leqslant \vartheta \leqslant 1)$, *then*

$$\sqrt[m]{V_m(H_\vartheta)} \geqslant (1-\vartheta)\sqrt[m]{V_m(H_0)} + \vartheta \sqrt[m]{V_m(H_1)}. \qquad (*)$$

For the bodies H_0 *and* H_1 *having interior points, equality holds in* $(*)$, *if and only if the bodies are homothetic.*

Up to a factor, $V_m(H)$ is the mean of (the areas of) the projections[4] of the body H to m-dimensional planes; $V_{n-1}(H)$ is the area of H divided by n.

Uniqueness of a convex body with given curvature function is proved and, finally, new extremal properties of the ball following from inequality $(*)$ are derived.

Part III. Extension of Two Minkowski Theorems on Convex Polyhedra to Arbitrary Convex Bodies

This part is closely related to Minkowski's paper *Allgemeine Lehrsätze über die konvexen Polyeder*.

Using the results of Part I, we generalize the theorems concerning existence of a polyhedron with given normals and areas of faces by directly applying Minkowski's arguments to the space of continuous functions. Thus, the possibility appears of realizing an arbitrary countably additive nonnegative set function that is defined on the surface of the ball by means of a convex surface. Moreover, some other results are also obtained.

[4] A *projection* is an orthoprojection.—Eds.

Part IV. On Mixed Discriminants and Mixed Volumes

Certain properties of the invariants of a pencil of quadratic forms, called *mixed discriminants*, are established; for positive-definite forms, in particular, inequalities similar to the inequalities between mixed volumes are derived. These results are applied to problems in the theory of convex bodies.

PART I

EXTENSION OF CERTAIN CONCEPTS OF THE THEORY OF CONVEX BODIES

§ 1. The Area of a Convex Surface

A *convex surface* is a part of the boundary or, as we usually say, of the surface of a convex body. If a convex set does not have interior points, then it degenerates into a plane figure. So, the area of each piece of its surface is determined in the same way as for a plane figure. Therefore we ignore this case as trivial and confine ourselves to the surface of a convex body having interior points.

Let H be a convex body containing interior points. Take the origin O in the interior of H and describe the unit ball E about O. Every ray starting from O pierces the surface of E and the surface of H at the only point each. Thus, some one-to-one bicontinuous correspondence is established between the points on E and the points on H.[5] We call this correspondence the *radial with the origin* O. (In the same way, we can define the radial correspondence between arbitrary convex surfaces containing the origin as an interior point. Thus, the radial correspondence is defined between the sets of points on the surfaces of bodies or, as we say, each set σ_1 on H_1 corresponds radially to some set σ_2 on H_2.)

Now let us construct convex polyhedra $H_1, H_2, \dots$, that include H (and consequently, the origin) and converge to H.

Let σ be a set on the surface of H, let ω be a set on E corresponding radially to σ; and let the sets $\sigma_1, \sigma_2, \dots$ on $H_1, H_2, \dots$, correspond radially to σ, and consequently, to ω. We define the area or, which is the same thing, the measure of the set σ as the limit of the areas of $\sigma_1, \sigma_2, \dots$ which make sense if the sets $\sigma_1, \sigma_2, \dots$ are not "too nasty," i.e., in development of the polyhedra $H_1, H_2, \dots$ on a plane they yield measurable sets.

Let $R, R_1, R_2, \dots$, denote the respective distances of points on the convex polyhedra $H, H_1, H_2, \dots$ from the origin; $\varphi, \varphi_1, \varphi_2, \dots$, the angles between the radius vectors joining the origin to the points on $H, H_1, H_2, \dots$ and the normals to the supporting planes of $H, H_1, H_2, \dots$ at these points. By virtue

[5]Here and in what follows by the phrase "on H" we mean "on the surface of H."

of the radial correspondence between E and $H, H_1, H_2, \ldots$, the distance and the angle mentioned above are functions of the points on the unit ball E. It is a simple matter to prove by any conventional method that the area $F(\sigma_k)$ of the polyhedral surface σ_k corresponding radially to a domain ω on E is equal to

$$F(\sigma_k) = \int_{\omega} \frac{R_k^{n-1}}{\cos\varphi_k}\, d\omega \tag{1}$$

where $d\omega$, here and in the sequel, stands for the area element on the surface of the unit ball.

Taking the integral in (1) as measure for the set σ_k, also in the case when if ω is an arbitrary measurable set, we find the measure thus defined has all the properties required of that concept. Moreover, this measure certainly agrees with the measure of the development of the set σ_k on a plane. Now we have to extend the concept of measure, or area $F(\sigma)$, to a set on the surface of an arbitrary convex body by a limit process.

Here, as everywhere else in this work, we denote a point as well as the vector joining it to the origin by the same symbol $\bar{x}$. We say there is a normal $\bar{n}$ to a body H at a point $\bar{x}$ on H, if some supporting plane of H with the outward normal $\bar{n}$ passes through $\bar{x}$.

Lemma I. *Let the sequence of convex bodies $H_1, H_2, \ldots$ converge to a body H and $\bar{x}_1, \bar{x}_2, \ldots$ be a convergent sequence of points on $H_1, H_2, \ldots$. The limit point $\bar{x}$ of this sequence lies on H. Let $\bar{n}_1, \bar{n}_2, \ldots$ be normals to $H_1, H_2, \ldots$ at the points $\bar{x}_1, \bar{x}_2, \ldots$. Every convergent subsequence of these normals converges to a normal to H at the point $\bar{x}$.*

Suppose that a subsequence of normals $\bar{n}_{k_1}, \bar{n}_{k_2}, \ldots$ converges to $\bar{n}$. The supporting planes $P_1, P_2, \ldots$ to the bodies $H_{k_1}, H_{k_2}, \ldots$ at $\bar{x}_{k_1}, \bar{x}_{k_2}, \ldots$ with the normals $\bar{n}_{k_1}, \bar{n}_{k_2}, \ldots$ converge to the plane P passing through $\bar{x}$ and having $\bar{n}$ as a normal. Since H_{k_ν} lies on one side of P_ν, the limit body H too lies on the same side of P. Consequently, the limit plane P is a supporting plane to the body H and passes through $\bar{x}$. This completes the proof.

Lemma II. *To the set of the points on H at which there is more than one normal (i.e., the points at which there is no tangent plane), there corresponds radially a set of measure zero on E.*

We omit the proof of this well-known lemma.[6]

These two lemmas imply that the angles $\varphi_1, \varphi_2, \ldots$ introduced above converge to φ almost everywhere on E. If the bodies $H_1, H_2, \ldots$ are all

[6]See Reidemeister, *Über die singularen Randpunkte eines konvexen Körpers,* Math. Ann., **83:** 116–118 (1921).

included in a ball of radius A and include a ball of radius a, both centered at the origin, then

$$\frac{R_k^{n-1}}{\cos \varphi_k} \leqslant \frac{A^n}{a}.$$

Consequently, the integrands in (1) are uniformly bounded and converge almost everywhere on E to $\frac{R^{n-1}}{\cos \varphi}$. Therefore for every measurable ω there exists a limit of the integrals in (1); hence there exists a limit of the sequence of the areas $F(\sigma_k)$. This limit is taken to be the area $F(\sigma)$ of the set σ on H that corresponds radially to ω.

Since to each closed (open) set on H there corresponds radially a closed (open) set on E, it follows in particular that every closed or open set on a convex body has area (measure). Grounding on these results, we assert the following

Lemma III. *The entire theory of measure and measurable functions on the sphere can be extended to every convex body by radially mapping the surface of the unit ball onto the surface of the convex body.*

By virtue of this lemma, instead of the unit ball, any convex body can be taken as the *fundamental body* (Eichkörper).

Let H_1 and H_2 be two convex bodies and let σ_1 and σ_2 be radially corresponding measurable sets on their surfaces. Then

$$F(\sigma_k) = \int\limits_{\omega} \frac{R_k^{n-1}}{\cos \varphi_k}\, d\omega \quad (k = 1, 2). \tag{1}$$

Since $\frac{R_k^{n-1}}{\cos \varphi_k}$ is the derivative of its indefinite integral almost everywhere, we have

$$F(\sigma_2) = \int\limits_{\sigma_1} \frac{R_2^{n-1}}{R_1^{n-1}} \frac{\cos \varphi_1}{\cos \varphi_2}\, dF(\sigma_1). \tag{2}$$

Lemma IV. *If the sequence of convex bodies $H_1, H_2, \dots$ converges to a body H, and $\sigma_1, \sigma_2, \dots$ are measurable sets on $H_1, H_2, \dots$ which correspond radially to a set σ on H, then*

$$F(\sigma) = \lim_{k=\infty} F(\sigma_k).$$

It is a direct corollary of the fact that the almost everywhere convergence of the functions $\frac{R_k^{n-1}}{\cos \varphi_k}$ implies the convergence of their integrals.

§ 2. The Area Function of a Convex Body

Suppose that a set ω is given on the surface of the unit ball E. Let $\sigma(\omega)$ denote the set of all those points on the surface of a given convex body H such that the normals $\bar{n}$ at these points when drawn from the center of the ball E meet ω. I call the set $\sigma(\omega)$ thus defined a *set mapped spherically onto* ω. To each set ω for which $\sigma(\omega)$ is measurable, we assign the number $F[\sigma(\omega)]$ that equals the area of the set $\sigma(\omega)$. Thus we defined a set function on the unit sphere. It is called the *area function* of the body H and is denoted by $F(H, \omega)$. Our problem here is to pick out the class of those sets ω for which the area function of an arbitrary convex body is well defined and to determine the properties of this set function.[7]

Lemma I. *Let $\bar{x}_1, \bar{x}_2, \dots$ be a sequence of points on the surface of a convex body H that converges to a point $\bar{x}$ on this surface. Let $\bar{n}_1, \bar{n}_2, \dots$ be a sequence of normals to the body H at the points $\bar{x}_1, \bar{x}_2, \dots$. Then every convergent subsequence of the sequence $\bar{n}_1, \bar{n}_2, \dots$ converges to a normal $\bar{n}$ at the point $\bar{x}$.*

We can imagine the body H as the limit of a sequence of identical bodies $H_1, H_2, \dots$ which converges to H. Take the normal $\bar{n}_1$ to H_1 at $\bar{x}_1$, the normal $\bar{n}_2$ to H_2 at $\bar{x}_2$ and so on. Thus, this lemma is reduced to Lemma I proved in the previous section.

Lemma II. *If ω is closed, then $\sigma(\omega)$ is also closed.*

Let $\bar{x}_1, \bar{x}_2, \dots$ be a sequence of points belonging to $\sigma(\omega)$ and converging to the point $\bar{x}$. We now have to demonstrate that $\bar{x}$ belongs to $\sigma(\omega)$ too.

Take normals at the points $\bar{x}_1, \bar{x}_2, \dots$ so that the points on E corresponding to them belong to ω. From the sequence of normals thus obtained, take a convergent subsequence. Its limit is, by Lemma I, a normal at the point $\bar{x}$. And since ω is closed, to this normal there corresponds a point on E lying in ω. Hence the point $\bar{x}$ belongs to $\sigma(\omega)$.

In § 1 every closed set on the surface of a convex body was shown to be measurable. Therefore, the area function $F(H, \omega)$ of each convex body H is defined for every closed set ω.

Lemma III. *If ω_1 and ω_2 have no common points, then the intersection $\sigma(\omega_1)\sigma(\omega_2)$ has zero measure.*

Suppose that the sets $\sigma(\omega_1)$ and $\sigma(\omega_2)$ have common points and let $\bar{x}$ be one such common point. (If the sets have no common points, then the lemma is trivial.) Then at the point $\bar{x}$ there is a normal $\bar{n}_1$ falling into[8]

[7] For set functions, see, for instance, Glivenko, *Stieltjes Integrals*, Chapter VIII, ONTI (1936) (in Russian).

[8] By "falling into ω," we mean the fact that the end of $\bar{n}$ lies in the set ω if $\bar{n}$ is drawn from the center of E.

ω_1 and a normal $\bar{n}_2$ falling into ω_2. Consequently, the normals at $\bar{x}$ all are distinct. However, by Lemma II of § 1, the set of all such points on H has measure zero.[9] Therefore the measure of the intersection $\sigma(\omega_1)\sigma(\omega_2)$ is zero.[10]

Lemma IV.

$$\sigma(\omega_1 + \omega_2) = \sigma(\omega_1) + \sigma(\omega_2).$$

Indeed, if $\bar{x}$ belongs to $\sigma(\omega_1)$, then there is a normal at $\bar{x}$ falling into ω_1, and consequently into $\omega_1 + \omega_2$. The same is true of the points belonging to $\sigma(\omega_2)$. If $\bar{x}$ belongs to the set $\sigma(\omega_1 + \omega_2)$, then the normal at $\bar{x}$ falling into $(\omega_1 + \omega_2)$ points either into ω_1 or into ω_2. Hence $\bar{x}$ belongs either to $\sigma(\omega_1)$ or to $\sigma(\omega_2)$ or to both.

Let ω be an open set and Ω be the whole surface of the unit ball E. The complement $\Omega - \omega$ is closed and has no points in common with the set ω. Therefore the intersection $\sigma(\omega)\sigma(\Omega - \omega)$ has measure zero. Furthermore, $\sigma(\Omega - \omega) + \sigma(\omega) = \sigma(\Omega)$. The sets $\sigma(\omega)$ and $\sigma(\Omega)$, being closed, are measurable. ($\sigma(\Omega)$ is simply the whole surface of H.) Therefore the set $\sigma(\Omega - \omega)$ too is measurable and

$$F(\sigma(\omega)) = F(\sigma(\Omega)) - F(\sigma(\Omega - \omega)).$$

Hence the function $F(H, \omega)$ is well defined for all open sets ω.

Let ω_1 and ω_2 be two disjoint sets for which the function $F(H, \omega)$ is meaningful. We now prove that $F(H, \omega_1 + \omega_2)$ too makes sense and

$$F(H, \omega_1 + \omega_2) = F(H, \omega_1) + F(H, \omega_2). \tag{1}$$

Since $\sigma(\omega_1)$ and $\sigma(\omega_2)$ are measurable sets, their sum $\sigma(\omega_1) + \sigma(\omega_2) = \sigma(\omega_1 + \omega_2)$ is also measurable and, consequently, the function $F(H, \omega_1 + \omega_2)$ has a meaning. Moreover, as the intersection $\sigma(\omega_1)\sigma(\omega_2)$ has measure zero (because ω_1 and ω_2 are disjoint), the measure of their sum is the sum of their measures. This completes the proof of (1).

Lemma V. *If sets $\omega_1, \omega_2, \dots$ at which some area function $F(H, \omega)$ is defined form a vanishing sequence (i.e., $\omega_k \supset \omega_{k+1}$ and $\prod_{k=1}^{\infty} \omega_k = 0$), then*

$$\lim_{k=\infty} F(H, \omega_k) = 0.$$

Since $\omega_k \supset \omega_{k+1}$, it follows that $\sigma(\omega_k) \supset \sigma(\omega_{k+1})$. If the intersection of all $\sigma(\omega_k)$ is empty, then its measure is clearly zero. Suppose that

[9] Formula (1) of § 1 says that a set of measure zero on E corresponds radially to a set of measure zero on H.

[10] Observe that the old (now obsolete) notation is retained here for the set-theoretic operations.—Eds.

$\prod_{k=1}^{\infty} \sigma(\omega_k) = \sigma$ is not empty and a point $\bar{x}$ belongs to the set σ. At $\bar{x}$ there are normals $\bar{n}_1, \bar{n}_2, \dots$ (not necessarily all distinct) falling into the sets $\omega_1, \omega_2, \dots$. Not all of these normals coincide, for if this happened each normal thus obtained would point into all ω_k and their intersection would not therefore be empty. So, the normals at $\bar{x}$ all are distinct. The set of all such points has measure zero and therefore $F(\sigma) = 0$. As is known, if $\sigma(\omega_k) \supset \sigma(\omega_{k+1})$, then

$$\lim_{k=\infty} F(\sigma(\omega_k)) = F\left(\prod_{k=1}^{\infty} \sigma(\omega_k)\right),$$

and since $F[\sigma(\omega_k)] = F(H, \omega)$, the limit $\lim_{k\to\infty} F(H, \omega_k)$ tends to zero as k tends to infinity. This completes the proof.

Let ω_k be disjoint sets and suppose that $F(H, \omega)$ is well defined for all ω_k and that

$$\omega_0 = \sum_{k=1}^{\infty} \omega_k.$$

The sets

$$\omega^{(m)} = \omega_0 - \sum_{k=1}^{m} \omega_k$$

form a vanishing sequence and, consequently, $\lim_{m=\infty} F(H, \omega^{(m)}) = 0$. Then we have

$$F(H, \omega^{(m)}) = F(H, \omega_0) - \sum_{k=1}^{m} F(H, \omega_k),$$

so

$$F(H, \omega_0) = \sum_{k=1}^{\infty} F(H, \omega_k).$$

This result can be formulated as a

Theorem. *The area function of a convex body is a nonnegative countably additive set function on the surface of the unit ball which is defined for the unions of closed and open sets.*

If H is a polyhedron, then $F(H, \omega) = 0$ only if none of the normals to its faces falls into the set ω. However, if the normals to the faces of areas $F_1, F_2, \dots, F_m$ fall into the set ω, then $F(H, \omega) = F_1 + F_2 + \dots + F_m$. Such a function $F(H, \omega)$ may be said to be “discrete.” Suppose that H has positive Gaussian curvature K at each point of its surface. Then the

function $F(H,\omega)$ is absolutely continuous, and its derivative at every point $\bar{n}$ of the sphere is

$$\lim_{\omega\to\bar{n}} \frac{F(H,\omega)}{\operatorname{mes}(\omega)} = \frac{1}{K(\bar{n})},$$

where $K(\bar{n})$ is the *Gaussian curvature* at the point on H where $\bar{n}$ is a normal.

§ 3. Representation of Mixed Volume by Means of Area Functions

Minkowski proved that the mixed volume of a convex body L and a convex polyhedron H can be expressed as

$$V(L,H,\dots,H) = \frac{1}{n}\sum_{\nu=1}^{N} L(\bar{n}_\nu)\,F_\nu, \tag{1}$$

where $\bar{n}_\nu$ is the normal to a face of H, F_ν is the area of the face and $L(\bar{n})$ is the *support function* of the body L.[11]

We generalize this result by proving

Lemma 1. *The mixed volume $V(L,H,\dots,H)$ of any two convex bodies L and H is*

$$V(L,H,\dots,H) = \frac{1}{n}\int\limits_{\Omega} L(\bar{n})\,F(H,d\omega). \tag{2}$$

The integral on the right is in fact a Lebesgue-Stieltjes integral. It makes sense because $L(n)$ is continuous and $F(H,\omega)$ is countably additive.

We prove this lemma by passage to a limit from polyhedra to every convex body as is usually done in such situations. Naturally, a finite sum transforms to an integral. However, a rigorous proof of this kind would be rather tedious and dry.

Suppose that a body H degenerates and has no interior points. In this case, either $F(H,\omega)$ vanishes if H is of dimension less than $(n-1)$, so $V(L,H,\dots,H)=0$, or H is $(n-1)$-dimensional and

$$V(L,H,\dots,H) = \frac{1}{n}(L(\bar{n}) + L(-\bar{n}))\,F, \tag{3}$$

[11] The symbols $L(\bar{n})$ and $H(\bar{n})$ denote the support functions of the bodies L and H "on the unit sphere." The support function is defined in the sense of Minkowski for all vectors. Here it is taken only for unit vectors $\bar{n}$, i.e., for the corresponding points on the unit ball. The function $L(\bar{n})$ stands for the distance from the origin to the supporting plane of the body L with outward normal $\bar{n}$. It is taken positive in the direction of $\bar{n}$, and negative in the opposite direction.

where $\bar{n}$ and $-\bar{n}$ are the normals to the plane in which H lies and F is its $(n-1)$-dimensional volume. This formula holds if H is an $(n-1)$-dimensional polyhedron.

Let us construct a sequence of polyhedra $H_1, H_2, \dots$ lying in the same plane as the polyhedron H and converging to H. Since the mixed volume is a continuous function of the bodies involved, the volumes $V(L, H_k, \dots, H_k)$ converge to $V(L, H, \dots, H)$. On the other hand, the $(n-1)$-dimensional volumes of the polyhedra H_k converge to the $(n-1)$-dimensional volume of H. Hence we obtain formula (3). That (3) is a particular case of (2) is evident, because if $\bar{n}$ falls into ω but $-\bar{n}$ does not, or vice versa, then $F(H, \omega) = F$, but if neither $\bar{n}$ nor $-\bar{n}$ falls into ω, then $F(H, \omega) = 0$.

Now assuming that H has interior points, construct a sequence of convex polyhedra including H and converging to H. Take an $\varepsilon > 0$ and choose a subsequence $H_1, H_2, \dots$ from our sequence such that for all k

$$|V(L, H, \dots, H) - V(L, H_k, \dots, H_k)| < \frac{\varepsilon}{2}. \tag{4}$$

This is possible because of the continuity of mixed volumes.[12]

Furthermore, take a $\rho > 0$ small enough so that for $|\bar{n}_1 - \bar{n}_2| < \rho$

$$|L(\bar{n}_1) - L(\bar{n}_2)| < \frac{\varepsilon}{8A}, \tag{5}$$

with A standing for an upper bound for the areas of the surfaces of the polyhedra $H_1, H_2, H_3, \dots, H_k, \dots$. Now let us subdivide the surface of the unit ball into subdomains $\omega_1, \dots, \omega_N$ such that, if $\bar{n}_1$ and $\bar{n}_2$ fall into one of these subdomains, then

$$|\bar{n}_1 - \bar{n}_2| < \frac{\rho}{2} \quad (\bar{n}_1, \bar{n}_2 \in \omega_\nu). \tag{6}$$

There exist sets $\sigma_1, \dots, \sigma_N$, on the surface of H, spherically corresponding to the sets $\omega_1, \dots, \omega_N$.

Take the origin in the interior of the polyhedron H and, consequently, in the interior of all polyhedra $H_1, \dots, H_k, \dots$. Consider a radial correspondence between these bodies. We denote by σ^k a set on H_k that corresponds radially to some set σ on H. Let σ^k_μ denote a face of the polyhedron H_k. A set on H_k which corresponds radially to a set σ_ν from the sets $\sigma_1, \dots, \sigma_N$ is denoted by σ^k_ν. Finally, let $\sigma^k_{\mu\nu}$ denote the intersection of σ^k_μ and σ^k_ν.

Recall the reasoning in § 1. To every direction of the radius vector $\bar{r}$ starting from the origin there corresponds exactly one point on each of

---

[12]By "continuity of mixed volumes" we mean the fact that if bodies $H_1^{(k)}, \dots, H_n^{(k)}$ converge to bodies $H_1, \dots, H_n$, their mixed volume $V(H_1^{(k)}, \dots, H_n^{(k)})$ converges to the mixed volume of the limit bodies.

bodies $H, H_1, H_2, \dots$ The normals $\bar{n}, \bar{n}^1, \bar{n}^2, \dots$ at these points are functions of $\bar{r}$, i.e., functions of the point $\bar{x}$ on H. These normals converge almost everywhere to $\bar{n}$ on H. Therefore, by the Egorov theorem, we can choose a perfect set Σ that differs in measure from the whole surface of H by a quantity not greater than a given $\delta > 0$ and such that $\bar{n}^1, \bar{n}^2, \dots$ converge uniformly to $\bar{n}$ on Σ.

Let C be an upper bound for all the quantities $\frac{R_k^{n-1}}{R^{n-1}} \frac{\cos\varphi}{\cos\varphi_k}$ (for notation see § 1). Then, by formula (2) of § 1, the measure of the set that corresponds radially to Σ on H_k differs from the area of H_k by a quantity less than $C\delta$ for all k. We take

$$\delta = \frac{\varepsilon}{8C \max |L(\bar{n})|}.$$

The intersections of all above-introduced sets σ with Σ and those with the set on H_k corresponding radially to it is denoted by an overbar on the same symbol. By the choice of δ, a set with an overbar differs in measure from the corresponding set with no overbar by a quantity not greater than

$$C\delta = \frac{\varepsilon}{8C \max\, |L(n)|}.$$

Now we proceed to certain estimates. By the choice of the size of the domains $\omega_1, \dots, \omega_N$, we have

$$\left| \frac{1}{n} \int\limits_{\Omega} L(\bar{n})\, F(H, d\omega) - \frac{1}{n} \sum_{\nu} L(\bar{n}_\nu)\, F(\sigma_\nu) \right| < \frac{\varepsilon}{8} \tag{7}$$

where $\bar{n}$ falls into ω_ν and $F(\sigma_\nu)$, the measure of the set σ_ν, is equal to $F(H, \omega)$.

By the choice of the set Σ,

$$\sum_{\nu} |F(\sigma_\nu) - F(\bar{\sigma}_\nu)| < \frac{\varepsilon}{8 \max |L(\bar{n})|}.$$

Therefore

$$\left| \frac{1}{n} \sum_{\nu} L(\bar{n}_\nu)\, F(\sigma_\nu) - \frac{1}{n} \sum_{\nu} L(\bar{n}_\nu)\, F(\bar{\sigma}_\nu) \right| < \frac{\varepsilon}{8} \tag{8}$$

and

$$\left| \frac{1}{n} \int\limits_{\Omega} L(\bar{n})\, F(H, d\omega) - \frac{1}{n} \sum_{\nu} L(\bar{n}_\nu)\, F(\bar{\sigma}_\nu) \right| < \frac{\varepsilon}{4}. \tag{9}$$

Since, by formula (1),

$$V(L, H_k, \ldots, H_k) = \frac{1}{n} \sum_{\mu} L\left(n_\mu^k\right) F\left(\sigma_\mu^k\right), \tag{10}$$

and, by the choice of Σ,

$$\sum_{\mu} \left|F\left(\sigma_\mu^k\right) - F\left(\bar{\sigma}_\mu^k\right)\right| < \frac{\varepsilon}{8 \max|L(\bar{n})|},$$

we obtain

$$\left|V(L, H_k, \ldots, H_k) - \frac{1}{n} \sum_{\mu} L\left(\bar{n}_\mu^k\right) F\left(\bar{\sigma}_\mu^k\right)\right| < \frac{\varepsilon}{8}, \tag{11}$$

or, since

$$\bar{\sigma}_\mu^k = \sum_{\nu} \bar{\sigma}_{\mu\nu}^k,$$

we have

$$\left|V(L, H_k, \ldots, H_k) - \frac{1}{n} \sum_{\mu,\nu} L\left(\bar{n}_\mu^k\right) F\left(\bar{\sigma}_\mu^k\right)\right| < \frac{\varepsilon}{8}. \tag{12}$$

Since the normals $\bar{n}^k(\bar{x})$ converge to $\bar{n}(\bar{x})$ uniformly on Σ, there exists an m such that, for every $k > m$ and for all $\bar{x}$ belonging to Σ, we have

$$\left|\bar{n}^k(\bar{x}) - \bar{n}(\bar{x})\right| < \frac{\rho}{2} \tag{13}$$

(for the meaning of ρ, see formula (5)). In particular, the normal $\bar{n}_\mu^k$ to a face of H_k is not farther than $\frac{1}{2}\rho$ from the normal $\bar{n}$ at the corresponding point on H. If this point belongs to $\bar{\sigma}_{\mu\nu}$, then, by the choice of the domains ω_ν,[13] we have $|\bar{n} - \bar{n}_\nu| < \frac{1}{2}\rho$, and since

$$\left|\bar{n} - \bar{n}_\mu^k\right| < \frac{1}{2}\rho,$$

then

$$\left|\bar{n}_\mu^k - \bar{n}_\nu\right| < \rho \tag{14}$$

if the point with the normal $\bar{n}_\mu^k$ belongs to $\bar{\sigma}_{\mu\nu}^k$ and $\bar{n}_\nu$ falls into ω. Consequently, under this condition, we obtain

$$\left|L\left(\bar{n}_\mu^k\right) - L\left(\bar{n}_\nu\right)\right| < \frac{\varepsilon}{8A}. \tag{15}$$

[13]The set Σ does not contain those points where there is more than one normal and therefore, if $\bar{x} \in \bar{\sigma}_\nu$, there is only one normal at $\bar{x}$ and it falls into the set ω.

So, for $k > m$, we obtain

$$\left| \frac{1}{n} \sum_{\mu,\nu} L\left(\bar{n}_{\mu}^{k}\right) F\left(\bar{\sigma}_{\mu\nu}^{k}\right) - \frac{1}{n} \sum_{\mu,\nu} L\left(\bar{n}_{\nu}\right) F\left(\bar{\sigma}_{\mu\nu}^{k}\right) \right| < \frac{\varepsilon}{8}. \tag{16}$$

In the second term, summing over μ and involving (12), we infer

$$\left| V(L, H_k, \dots, H_k) - \frac{1}{n} \sum_{\nu} L(\bar{n}_{\nu}) F(\bar{\sigma}_{\nu}) \right| < \frac{\varepsilon}{4} \tag{17}$$

(because $\bar{\sigma}_{\nu}^{k} = \sum_{\mu} \bar{\sigma}_{\mu\nu}^{k}$). Now, combining (9) and (17), we obtain

$$\left| V(L, H_k, \dots, H_k) - \frac{1}{n} \int\limits_{\Omega} L(\bar{n}) F(H, d\omega) \right| < \frac{\varepsilon}{2}. \tag{18}$$

Finally, applying (4), we derive from (18)

$$\left| V(L, H, \dots, H) - \frac{1}{n} \int\limits_{\Omega} L(\bar{n}) F(H, d\omega) \right| < \varepsilon, \tag{19}$$

and since ε is arbitrary, we arrive at (2). Putting $L = H$ in (2), find $V(L, H, \dots, H)$ is the volume of the body H, so the volume of a convex body can be expressed as

$$V(H) = \frac{1}{n} \int\limits_{\Omega} H(\bar{n}) F(H, d\omega). \tag{20}$$

Lemma II. *The area function of a convex body satisfies the condition*

$$\int\limits_{\Omega} \bar{n} F(H, d\omega) = 0. \tag{21}$$

(The integral of a vector x is always the vector whose components are the integrals of the components of x.)

The visual meaning of formula (21) is as follows. Imagine the surface of the unit ball loaded with masses such that the mass of a set ω is $F(H, \omega)$. Then formula (21) implies that the centroid of this system is the center of the ball.

If the body H is translated by a vector $\bar{a}$, then the summand $(\bar{a}, \bar{n})$ is added to its support function on the unit ball; but the volume is constant under translation; therefore, for every $\bar{a}$

$$\int\limits_{\Omega} (\bar{a}\bar{n}) F(H, d\omega) = 0; \tag{22}$$

hence we obtain (21).

Lemma III. *The area of the projection of a convex body H in the direction of $\bar{n}_0$*[14] *is equal to*

$$\frac{1}{2}\int_{\Omega} |\bar{n}_0\bar{n}|\, F(H, d\omega), \tag{23}$$

where $|\bar{n}_0\bar{n}|$ is the modulus of the cosine of the angle between $\bar{n}_0$ and $\bar{n}$.

It is a well-known formula and its proof is simple if H is a polyhedron. In this case the integral reduces to the sum of the terms each of which represents the area of the projection of a face of H in the direction $\bar{n}_0$.

It is readily seen that $\frac{1}{2}|\bar{n}_0\bar{n}|$ is the support function (on the unit ball) of a unit segment L parallel to $\bar{n}_0$ and having its midpoint at the origin. Therefore the integral in (23) is n times the mixed volume $V(L, H, \dots, H)$. Now if we construct a sequence of polyhedra convergent to H, then, on the one hand, the areas of their projections converge to the area of the projection of H to the same plane and, on the other hand, their mixed volumes with L converge to $V(L, H, \dots, H)$. In the limit we obtain (23).

If H has interior points, then there exists a ball included in H. Therefore the areas of all projections of H are bounded below by the area of the projection of this sphere. Hence we obtain

Lemma IV. *The area function of a convex body with interior points satisfies the condition*

$$\int_{\Omega} |\bar{n}_0\bar{n}|\, F(H, d\omega) > a > 0 \tag{24}$$

for every $\bar{n}_0$, where a is a constant which is the same for all $\bar{n}_0$.

§ 4. Mixed Area Functions

Let $H_1, H_2, \dots, H_{n-1}$ be convex bodies and let L be a variable convex body. The mixed volume $V(L, H_1, \dots, H_{n-1})$ is a continuous additive functional on the set of convex bodies or, which is the same thing, on the set of positively homogeneous convex functions $L(\bar{n})$.[15] Indeed, for $\lambda, \lambda \geqslant 0$,

$$V(\lambda_1 L_1 + \lambda_2 L_2, H_1, \dots, H_{n-1}) = \lambda_1 V(L_1, H_1, \dots, H_{n-1}) \\ + \lambda_2 V(L_2, H_1, \dots, H_{n-1}), \tag{1}$$

[14] The *projection in the direction* of $\bar{n}$ means the orthoprojection to the plane with normal $\bar{n}$.—Eds.

[15] Strictly speaking, as was already shown by Minkowski, for a vector function $L(\bar{u})$ to be the support function of a convex body, it is necessary and sufficient that it be a positively homogeneous function of degree one, i.e., for $\lambda \geqslant 0$, $L(\lambda\bar{n}) = \lambda L(\bar{n})$, and be convex, i.e., $L((1-\vartheta)\bar{u} + \vartheta\bar{v}) \leqslant (1-\vartheta)L(\bar{u}) + \vartheta L(\bar{v})$, or, by virtue of the homogeneity condition, $L(\bar{u} + \bar{v}) \leqslant L(\bar{u}) + L(\bar{v})$. [See, Bonnesen and Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Math., 1934, § 2.]

and $V(L, H_1, \dots, H_{n-1})$ depends continuously on L.

We consider the functions defined on the unit sphere. I call any such function *convex*, if it is a support function on the unit sphere. We now extend the functional $V(L, H_1, \dots, H_{n-1})$ defined on the set of convex functions $L(\bar{n})$ to the set of functions $Z(\bar{n})$ that are the differences of convex functions. Let $Z(\bar{n}) = L_1(\bar{n}) - L_2(\bar{n})$, where $L_1(\bar{n})$ and $L_2(\bar{n})$ are convex. Put

$$V(Z, H_1, \dots, H_{n-1}) = V(L_1, H_1, \dots, H_{n-1}) - V(L_2, H_1, \dots, H_{n-1}). \quad (2)$$

This definition is sound. Indeed, let $Z(\bar{n}) = L_1'(\bar{n}) - L_2'(\bar{n})$; then

$$L_1'(\bar{n}) + L_2(\bar{n}) = L_1(\bar{n}) + L_2'(\bar{n}),$$

but

$$\begin{aligned} &V(L_1', H_1, \dots, H_{n-1}) + V(L_2, H_1, \dots, H_{n-1}) \\ &= V(L_1, H_1, \dots, H_{n-1}) + V(L_2', H_1, \dots, H_{n-1}), \end{aligned}$$

i.e.,

$$\begin{aligned} &V(L_1', H_1, \dots, H_{n-1}) + V(L_2', H_1, \dots, H_{n-1}) \\ &= V(L_1, H_1, \dots, H_{n-1}) + V(L_2, H_1, \dots, H_{n-1}). \end{aligned}$$

This completes the proof.

Lemma I. *The functional $V(Z, H_1, \dots, H_{n-1})$ is linear.*

Let $Z = Z' + Z''$ and $Z' = L_1' - L_2'$, $Z'' = L_1'' - L_2''$; then

$$Z = (L_1' + L_1'') - (L_2' + L_2'').$$

The functions $L_1'(\bar{n}) + L_1''(\bar{n})$ and $L_2'(\bar{n}) + L_2''(\bar{n})$ are convex. Applying formula (2) to them, by virtue of the additivity of the mixed volumes $V(L, H_1, \dots, H_2)$, we find that

$$\begin{aligned} V(Z' + Z'', H_1, \dots, H_{n-1}) &= V(Z', H_1, \dots, H_{n-1}) \\ &\quad + V(Z'', H_1, \dots, H_{n-1}). \end{aligned} \quad (3)$$

Let $Z = L_1 - L_2$ and $\lambda > 0$. Now applying formula (2) to λZ, by virtue of the homogeneity of $V(H, H_1, \dots, H_{n-1})$, we derive

$$V(\lambda Z, H, \dots, H_{n-1}) = \lambda V(Z, H_1, \dots, H_{n-1}). \quad (4)$$

However, if $\lambda < 0$, then $\lambda Z = |\lambda| L_2 - |\lambda| L_1$. Therefore the result is valid for negative λ as well.

Minkowski proved that if L is an arbitrary convex body and $H_1, \ldots, H_{n-1}$ are convex polyhedra, then

$$V(L, H_1, \ldots, H_{n-1}) = \frac{1}{n} \sum_{\nu=1}^{N} L(\bar{n}_\nu)\, F_\nu(H_1, \ldots, H_{n-1}), \tag{5}$$

where $F_\nu(H_1, \ldots, H_{n-1})$ is the mixed volumes (i.e., the mixed area) of the faces of the polyhedra $H_1, \ldots, H_{n-1}$ lying in parallel planes with normals $\bar{n}_\nu$. The faces are of dimension less than or equal to $(n-1)$.

If $Z(\bar{n}) = L_1(\bar{n}) - L_2(\bar{n})$ and $H_1, \ldots, H_{n-1}$ are polyhedra, then (5) and (2) yield

$$V(Z, H_1, \ldots, H_{n-1}) = \frac{1}{n} \sum_{\nu=1}^{N} Z(\bar{n}_\nu)\, F_\nu(H_1, \ldots, H_{n-1}). \tag{6}$$

Hence, it follows, first, that for $Z(\bar{n}) \geqslant 0$

$$V(Z, H_1, \ldots, H_{n-1}) \geqslant 0, \tag{7}$$

and, second, that

$$|V(Z, H_1, \ldots, H_{n-1})| \leqslant \frac{1}{n} \sum_{\nu=1}^{n} F_\nu(H_1, \ldots, H_{n-1}) \max |Z(\bar{n})|. \tag{8}$$

If E denotes the unit ball, then we can write

$$\frac{1}{n} \sum_{\nu=1}^{N} F_\nu(H_1, \ldots, F_{n-1}) = V(E, H_1, \ldots, H_{n-1}), \tag{9}$$

hence

$$|V(Z, H_1, \ldots, H_{n-1})| \leqslant V(E, H_1, \ldots, H_{n-1}) \max |Z(\bar{n})|. \tag{10}$$

If our polyhedra converge to arbitrary convex bodies, then, for a given $Z(\bar{n}) = L_1(\bar{n}) - L_2(\bar{n})$, the values of the functionals $V(Z, H_1, \ldots, H_{n-1})$ converge to the values of the functional at the limit bodies since the mixed volumes

$$V(L_1, H_1, \ldots, H_{n-1}) \text{ and } V(L_2, H_1, \ldots, H_{n-1})$$

enjoy this property. Therefore, (7) and (10) can be applied to arbitrary convex bodies $H_1, \ldots, H_{n-1}$.

Formulas (3), (4), and (10) show that the functional $V(Z, H_1, \ldots, H_{n-1})$ is linear.

Lemma II. *Every continuous function on the unit sphere can be uniformly approximated with any accuracy by the difference of convex functions.*

The Weierstrass approximation theorem asserts that every continuous function on the unit sphere can be uniformly approximated with any accuracy by a twice continuously differentiable function. We now show that every such function is the difference of convex functions.

Let $Z(\bar{n})$ be a twice continuously differentiable function on the unit sphere. Extend it to the whole space by putting

$$Z(\bar{u}) = |\bar{u}|\, Z\left(\frac{\bar{u}}{|\bar{u}|}\right),$$

where $|\bar{u}|$ is the length of the vector $\bar{u}$, so $\frac{\bar{u}}{|\bar{u}|}$ is a unit vector.

The second differential $d^2Z(\bar{u})$ of the function $Z(\bar{u})$ is a quadratic form whose eigenvalues all are functions continuous everywhere except at the point $\bar{u} = 0$ and positively homogeneous of degree minus one (because $Z(\bar{u})$ is a positively homogeneous function of degree one). So we can find a number C such that

$$Cd^2|\bar{u}| - d^2Z(\bar{u}) = d^2(C|\bar{u}| - Z(\bar{u}))$$

is a positive form for all $\bar{u} \neq 0$. However, we know that the function $H(\bar{n}) = C|\bar{n}| - Z(\bar{n})$, which is a positively homogeneous function of degree one, is the support function of some convex body.[16] Now on returning from the whole space back to the unit sphere, we obtain

$$Z(\bar{n}) = C - H(\bar{n}),$$

where $H(\bar{n})$ is a convex function on the unit sphere.

Now assume that $Z(\bar{b})$ is an arbitrary continuous function on the unit sphere and $Z_1(\bar{n}), \dots, Z_m(\bar{n}), \dots$ is a sequence of the differences of convex functions converging uniformly to $Z(\bar{n})$. Then, for an arbitrary $\varepsilon > 0$, there exists an M such that, for $m \geqslant M$ and for any k,

$$|Z_{m+k}(\bar{n}) - Z_m(\bar{n})| < \varepsilon.$$

Therefore, for $n > M$, from (10) we derive

$$|V(Z_{m+k}, H_1, \dots, H_{n-1}) - V(Z_m, H_1, \dots, H_1)| < \varepsilon\, V(E, H_1, \dots, H_{n-1}),$$

[16]From the condition $d^2H(\bar{u}) \geqslant 0$, we readily find that the function $H(\bar{u})$ is convex. The function $C|\bar{u}|$ here is the support function of the ball of radius C.

where $H_1, \dots, H_{n-1}$ are the given convex bodies. Hence, the extended mixed volumes $V(Z_m, H_1, \dots, H_{n-1})$ converge to some limit which we take equal to the value of the functional $V(Z, H_1, \dots, H_{n-1})$ for the given continuous function $Z(\bar{n})$. This definition is sound because, if $Z_1(\bar{n}), Z_2(\bar{n}), \dots$ and $Z_1'(\bar{n}), Z_2'(\bar{n}), \dots$ are two sequences of the differences of convex functions converging to $Z(\bar{n})$, then the sequence $Z_1(\bar{n}), Z_1'(\bar{n}), Z_2(\bar{n}), Z_2'(\bar{n}), \dots$ also converges to $Z(\bar{n})$. That the functional $V(Z, H_1, \dots, H_{n-1}$ so defined for all continuous functions is linear is easily proved. On taking the limit inequality (10) still holds, indicating thereby that our functional is continuous. Its norm, as shown by (10), is

$$V(E, H_1, \dots, H_{n-1}).$$

By the Riesz representation theorem, a linear functional defined for all continuous functions $Z(\bar{n})$ can be represented as the Lebesgue-Stieltjes integral of the continuous function $Z(\bar{n})$ with respect to a uniquely defined countably additive set function on the unit sphere. Our reasoning shows the functional $V(Z, H_1, \dots, H_{n-1})$ is completely determined by the convex bodies $H_1, \dots, H_{n-1}$ So we can assert

$$V(Z, H_1, \dots, H_{n-1}) = \frac{1}{n} \int\limits_{\Omega} Z(\bar{n})\, F(H_1, \dots, F_{n-1}; d\omega), \tag{11}$$

where $F(H_1, \dots, F_{n-1}; \omega)$ is a function, of the sets ω on the unit sphere, which is uniquely defined by the given convex bodies $H_1, \dots, H_{n-1}$. I call this function the *mixed area function* of the convex bodies $H_1, \dots, H_{n-1}$.

Since $V(Z, H_1, \dots, H_{n-1}) \geqslant 0$ for $Z(\bar{n}) \geqslant 0$ (see (7)), it follows that[17]

$$F(H_1, \dots, H_{n-1}; \omega) \geqslant 0. \tag{12}$$

Let $Z(\bar{n}) = \bar{a}\ \bar{n}$ be the support function of the endpoint of the vector $\bar{a}$. Then, obviously,

$$V(Z, H_1, \dots, H_{n-1}) = 0.$$

Therefore, substituting $Z(\bar{n}) = \bar{a}\bar{n}$ into (11) and observing that the vector $\bar{a}$ is arbitrary, we obtain

$$\int\limits_{\Omega} \bar{n} F(H_1, \dots, H_{n-1}; d\omega) = 0. \tag{13}$$

[17] Let $Z_1, Z_2, \dots$ be a sequence of nonnegative functions converging to one at the point $\bar{n} \in \omega$ and vanish at all other points on the unit sphere. For all m, we have $V(Z_m, H_1, \dots, H_{n-1}) \geqslant 0$. By definition,

$$F(H_1, \dots, H_{n-1}; \omega) = \lim_{m=\infty} V(Z_m, H_1, \dots, H_{n-1})$$

and, consequently, $F(H_1, \dots, H_{n-1}; \omega) \geqslant 0$.

§ 5. Geometric Meaning of Mixed Area Functions

1. In § 3 we expressed the mixed volume $V(L, H, \dots, H)$ through the area function of a body H as

$$V(L, H, \dots, H) = \frac{1}{n} \int_{\Omega} L(\bar{n})\, F(H, d\omega). \tag{1}$$

Hence it follows that if $Z(\bar{n})$ is the difference of convex functions, then

$$V(Z, H, \dots, H) = \frac{1}{n} \int_{\Omega} Z(\bar{n})\, F(H, d\omega). \tag{2}$$

Since every continuous function can be uniformly approximated by the difference of convex functions, formula (3) holds also for every continuous function $Z(\bar{n})$, provided $V(Z, H, \dots, H)$ is interpreted as the functional defined in § 4. Comparing (2) with the formula (11) of § 4, for $H_1 = H_2 = \dots = H_{n-1} = H_n$, we obtain

$$F(H, \dots, H; d\omega) = F(H, \omega), \tag{3}$$

i.e., the function $F(H, \dots, H; d\omega)$ that we defined abstractly is simply the area function of the body H.

2. Substituting H_n for Z in formula (11) of § 4, we obtain an expression for the mixed volume of the bodies $H_1, \dots, H_n$:

$$V(H_1, \dots, H_n) = \frac{1}{n} \int_{\Omega} H_n(\bar{n})\, F(H_1, \dots, H_{n-1}; d\omega). \tag{4}$$

Substituting in formula (11) of § 4 the body H_1 on H_n and $Z(\bar{n})$ on $H_1(\bar{n})$, we obtain

$$V(H_1, \dots, H_n) = \frac{1}{n} \int_{\Omega} H_1(\bar{n})\, F(H_2, \dots, H_n; d\omega). \tag{5}$$

Moreover, since $V(Z, H_1, \dots, H_{n-1})$ is invariant under the permutations of $H_1, \dots, H_{n-1}$, we can assert that the bodies $H_1, \dots, H_n$ can be arbitrarily permuted in expression (4) for the mixed volume. By simple analogy, this important property may be called the *self-conjugacy* of mixed area functions.

3. Let $H_1 = \lambda' H' + \lambda'' H''$; then

$$V(H_1, \dots, H_n) = \lambda' V(H', H_2, \dots, H_n) + \lambda'' V(H'', H_2, \dots, H_n) \tag{6}$$

and by (6), we have

$$V(H_1,\dots,H_n) = \lambda' \frac{1}{n} \int\limits_{\Omega} H_n(\bar{n})\, F(H',\dots,H_{n-1}; d\omega)$$
$$+ \lambda'' \frac{1}{n} \int\limits_{\Omega} H_n(\bar{n})\, F(H'',\dots,H_{n-1}; d\omega). \tag{7}$$

From the reasoning of the previous section it follows that $H_n(\bar{n})$ in (7) can be replaced by an arbitrary continuous function, and since

$$V(Z, H_1,\dots,H_{n-1}) = \frac{1}{n} \int\limits_{\Omega} Z(\bar{n})\, F(\lambda' H' + \lambda'' H'',\, H_2,\dots,H_{n-1}; d\omega), \tag{8}$$

we have

$$F(\lambda' H' + \lambda'' H'', H_2,\dots,H_{n-1}; \omega)$$
$$= \lambda' F(H', H_2,\dots,H_{n-1}; \omega) + \lambda''(H'', H_2,\dots,H_{n-1}; \omega). \tag{9}$$

Putting

$$H = \sum_{k=1}^{m} \lambda_k\, H_k,$$

and then repeatedly applying formula (9) to the function $F(H,\dots,H;\omega) = F(H,\omega)$ we find it to be a homogeneous polynomial of degree $(n-1)$ in λ_k:

$$F\left(\sum_{k=1}^{m} \lambda_k H_k, \omega\right) = \sum_{k_1,\dots,k_{n-1}} \lambda_{k_1} \dots \lambda_{k_{n-1}}\, F(H_{k_1},\dots,H_{k_{n-1}}; \omega), \tag{10}$$

where $k_1,\dots,k_{n-1}$ take the values from 1 to m independently of each other.

Just this formula gives the geometric meaning of mixed area functions and justifies the name in analogy with mixed volumes. However, since the area function of H is not altered when the body H is translated (as is clear from its definition), the mixed area function is not altered either when the bodies determining it are translated.

We now sum up these results.

Theorem. *The area function $F\left(\sum_{k=1}^{m} \lambda_k H_k, \omega\right)$ of a linear combination of convex bodies is a homogeneous polynomial of degree $n-1$ in λ_k:*

$$F\left(\sum_{k=1}^{m} \lambda_k H_k, \omega\right) = \sum_{k_1,\dots,k_{n-1}} \lambda_{k_1} \dots \lambda_{k_{n-1}}\, F(H_{k_1},\dots,H_{k_{n-1}}; \omega),$$

where the subscripts $k_1, \dots, k_{n-1}$ run independently of one another through all values from 1 to m. The factor $F(H_{k_1}, \dots, H_{k_{n-1}}; \omega)$ depends solely on the bodies $H_{k_1}, \dots, H_{k_{n-1}}$ and does not depend on the other bodies involved in the linear combination. It is defined in such a way that it is invariant under the permutations of the bodies $H_{k_1}, \dots, H_{k_{n-1}}$. It is called a mixed area function.

A mixed area function has the following properties.

(1) *It is nonnegative and countably additive set function on the unit sphere and is well defined for the sums of closed and open sets.*
(2) *It is invariant under translations of the bodies determining it.*
(3) *If the bodies $H_1, \dots, H_{n-1}$ are all translates of a body H, then $F(H_1, \dots, H_{n-1}, \omega)$ is the area function of the body H.*
(4) *The mixed volume is expressed via the mixed area function as*

$$V(H_1, \dots, H_n) = \frac{1}{n} \int\limits_{\Omega} H_n(\bar{n})\, F(H_1, \dots, H_{n-1}; d\omega),$$

where the bodies $H_1, \dots, H_n$ can be arbitrarily permuted.
(5) *It always satisfies the condition*

$$\int\limits_{\Omega} \bar{n} F(H_1, \dots, H_{n-1}; d\omega) = 0.$$

(6) *If one of the bodies determining some given mixed area function is a linear combination of some other bodies, i.e., if $H_1 = \sum_{k=1}^{m} \lambda_k H^{(k)}$, then $F(H_1, \dots, H_{n-1}; \omega) = \sum_{k=1}^{m} \lambda_k F(H^{(k)}, H_2, \dots, H_{n-1}; \omega)$, i.e., it is a linear combination of the corresponding mixed area functions. Consequently, in general for $H = \sum_{k=1}^{m} \lambda_k H^{(k)}$, we have*

$$\begin{aligned} &F(H, \dots, H,\, H_1, \dots, H_{n-1};\, \omega) \\ &= \sum_{k_1, \dots, k_p} \lambda_{k_1} \dots \lambda_{k_p} F\left(H^{(k_1)}, \dots, H^{(k_p)}, H_1, \dots, H_{n-p-1};\, \omega\right), \end{aligned}$$

where the subscripts $k_1, \dots, k_p$ run independently through the values from 1 to m.

If the bodies $H_1, \dots, H_{n-1}$ are polyhedra, then the function $F(H_1, \dots, H_{n-1}; \omega)$ is discrete and has the following meaning. Let $\bar{n}_\nu$ be the normal to an $(n-1)$-dimensional face of the polyhedron $H_1 + H_2 + \dots + H_{n-1}$. This face is a linear combination of those faces (of any dimension) of the polyhedra $H_1, \dots, H_{n-1}$ that lie in the supporting planes of these polyhedra

with the normal $\bar{n}_\nu$. Denote these faces by $H_1, \dots, H_{n-1}$, and denote their $(n-1)$-dimensional mixed volumes by $F(H_1^\nu, \dots, H_{n-1}^\nu)$.

If ω contains a point that corresponds to such a normal $\bar{n}_\nu$, and does not contain any other such points, then

$$F(H_1, \dots, H_{n-1};\, \omega) = F\left(H_1^\nu, \dots, H_{n-1}^\nu\right). \tag{11}$$

However, if ω does not contain points corresponding to the normals $\bar{n}_\nu$ at all, then

$$F(H_1, \dots, H_{n-1};\, \omega) = 0.$$

A particular case for which our theorem is known is the case of regular convex bodies.[18] I call the mixed area function $F(\underbrace{H, \dots, H}_{m}, E, \dots, E; \omega)$ the mth curvature function of the body H and denote it by $F_m(H;\omega)$. In particular, the $(n-1)$th curvature function is the area function.

Given a polyhedron H, take a domain ω on the unit sphere and the domain $\sigma(\omega)$ on H mapped spherically onto ω. Construct the body $H+\lambda E$ parallel to H and consider the domain $\sigma_\lambda(\omega)$ on $H+\lambda E$ mapped spherically onto ω and derived from $\sigma(\omega)$ when all possible normals at every point of $\sigma(\omega)$ that fall into ω are extended to a distance λ. Over the pieces of the $(n-1)$-dimensional faces belonging to $\sigma(\omega)$ we obtain just plane pieces and the sum of the areas of these pieces is $F_{n-1}(H;\omega)$. Over the pieces of the $(n-2)$-dimensional faces belonging to $\sigma(\omega)$ we obtain cylindrical surfaces and the sum of the areas of these cylindrical surfaces is clearly proportional to λ and is equal to $\lambda F_{n-2}(H;\omega)(n-1)$. Over the pieces of the $(n-3)$-dimensional faces we obtain cylindrical surfaces with $(n-3)$-dimensional generators and so on. Finally, over each vertex we obtain a spherical sector and the sum of the areas of these sectors is obviously equal to $\lambda^{n-1}\omega$. Thus, in this simple case every curvature function has a perfectly visual meaning and the formula

$$F_{n-1}(H+\lambda E,\, \omega) = \sum_{k=0}^{n-1} \lambda^k\, C_{n-1}^k\, F_{n-k-1}(H,\omega), \tag{12}$$

which is a particular case of the general formula indicated in property (6), is readily derived in an elementary way.

Suppose that the body H is regular, i.e., at every point on its surface it has well defined principal radii of curvature $R_1, R_2, \dots, R_{n-1}$ which are nowhere zero and are continuous functions of the normal $\bar{n}$. For such a body,

[18] Bonnesen and Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Math., Items 37 and 38 of § 8 (1934).

the area of the domain $\sigma(\omega)$ mapped spherically onto ω is given by the formula

$$F_{n-1}(H,\omega) = \int_{\omega} R_1 \dots R_{n-1}\, d\omega. \tag{13}$$

The principal radii of curvature on the surface of the outer parallel body $H + \lambda E$ are, as is well known, equal to $R_i + \lambda$. Therefore the area of the domain on $H + \lambda E$ that correspond to $\sigma(\omega)$ is

$$F_{n-1}(H + \lambda E,\, \omega) = \int_{\omega} (R_1 + \lambda) \dots (R_{n-1} + \lambda)\, d\omega. \tag{14}$$

Expanding in powers of λ, we find the coefficient of λ^k is an elementary symmetric function $S_{n-k-1}(R_1, \dots, R_{n-1})$ of degree $(n-k-1)$ in the variables $R_1, \dots, R_{n-1}$. This elementary symmetric function is generally called the $(n-k-1)$th *curvature function*.[19] The meaning of this term as modified by us is justified because the usual curvature function makes sense only for a regular body, whereas our curvature function is well defined for any convex body.

From (14) we obtain

$$F_{n-1}(H + \lambda E,\, \omega) = \sum_{k=0}^{n-1} \lambda^k \int_{\omega} S_{n-k-1}(R_1, \dots, R_{n-1})\, d\omega, \tag{15}$$

and since

$$F_{n-1}(H + \lambda E,\, \omega) = \sum_{k=0}^{n-1} \lambda^k\, C_{n-1}^k\, F_{n-k-1}(H, \omega), \tag{16}$$

we have

$$F_m(H,\omega) = \frac{1}{C_{n-1}^m} \int_{\omega} S_m(R_1, \dots, R_{n-1})\, d\omega. \tag{17}$$

The last relation establishes a connection between our general concepts and the concepts well known in differential geometry.

§ 6. Certain Generalizations of the Previous Concepts

The concepts of mixed volume and mixed area function defined for convex bodies are easily extended to the differences of convex functions on the unit

[19] See Bonnesen and Fenchel, Section 38, § 8.

sphere.[20] For this purpose it is sufficient to repeat the reasoning we applied in Section 1 to define the functional $V(Z, H_1, \ldots, H_{n-1})$ for the differences of convex functions $Z(\bar{n})$.

Assuming that the "mixed volume," $V(Z_1, \ldots, Z_m, H_1, \ldots, H_{n-m})$, has already been defined, where $Z(\bar{n}), \ldots, Z_m(\bar{n})$ are the differences of convex functions, we let

$$Z_{m+1}(\bar{n}) = H'_{n-m}(\bar{n}) - H''_{n-m}(\bar{n}).$$

Put

$$\begin{aligned} &V(Z_1, \ldots, Z_{m+1}, H_1, \ldots, H_{n-m-1}) \\ &= V(Z_1, \ldots, Z_m, H_1, \ldots, H'_{n-m}) - V(Z_1, \ldots, Z_m, H_1, \ldots, H''_{n-m}). \quad (1) \end{aligned}$$

Soundness of the definition is demonstrated in the same way as we did earlier.

Thus, the functional $V(Z_1, \ldots, Z_n)$ is defined, where $Z_1(\bar{n}), Z_2(\bar{n}), \ldots, Z_n(\bar{n})$ are the differences of convex functions. It is symmetric in the functions $Z_1, \ldots, Z_n$. As before, it is a simple matter to prove that this functional is homogeneous and additive in each of the functions $Z_1(\bar{n}), \ldots, Z_n(\bar{n})$ separately; in other words,

$$\begin{aligned} &V(\lambda' Z' + \lambda'' Z'', Z_1, \ldots, Z_{n-1}) \\ &= \lambda' V(Z', Z_1, \ldots, Z_{n-1}) + \lambda'' V(Z'', Z_1, \ldots, Z_{n-1}). \quad (2) \end{aligned}$$

The continuity of $V(Z, Z_1, \ldots, Z_{n-1})$ in $Z(\bar{n})$, for given $Z_1(\bar{n}), \ldots, Z_{n-1}(\bar{n})$, is easily proved as follows. Let

$$Z_1(\bar{n}) = H_1^0(\bar{n}) - H_1^1(\bar{n}), \ldots, Z_{n-1}(\bar{n}) = H_{n-1}^0(\bar{n}) - H_{n-1}^1(\bar{n}).$$

Expanding $V(Z, Z_1, \ldots, Z_{n-1})$, we obtain

$$V(Z, Z_1, \ldots, Z_{n-1}) = \sum (-1)^{\sum i_k} V(Z, H_1^{i_1}, \ldots, H_{n-1}^{i_{n-1}}), \quad (3)$$

where summation is taken over all combinations of $i_1, \ldots, i_{n-1}$ which take the values 0 and 1. Taking the absolute value of each term in this sum and applying the inequality

$$|V(Z, H_1, \ldots, H_{n-1})| \leqslant V(E, H_1, \ldots, H_{n-1}) \max |Z(\bar{n})| \quad (4)$$

proved earlier, we obtain

$$|V(Z, Z_1, \ldots, Z_{n-1})| \leqslant \max |Z(\bar{n})| \sum V(E, H_1^{i_1}, \ldots, H_{n-1}^{i_{n-1}}),$$

or in a brief form

$$|V(Z, Z_1, \ldots, Z_{n-1})| \leqslant A \cdot \max |Z(\bar{n})|. \quad (5)$$

This completes the proof. Thus we arrive at the following assertion.

[20] Defined in § 5.

There exists a unique functional $V(Z_1,\dots,Z_n)$ defined for the differences of convex functions that is linear in each of the functions $Z_1(\bar n),\dots,Z_n(\bar n)$ and is equal to the mixed volume of the convex bodies $H_1,\dots,H_n$, provided $Z_1(\bar n)=H_1(\bar n),\dots,Z_n(\bar n)=H_n(\bar n)$ are their support functions.

Along the same lines as before, the functional $V(Z,Z_1,\dots,Z_{n-1})$ defined for given $Z_1,\dots,Z_{n-1}$ can be extended to arbitrary continuous functions $Z(\bar n)$ on the unit sphere. Then, clearly the functional can be uniquely represented as

$$V(Z,Z_1,\dots,Z_{n-1})=\frac{1}{n}\int\limits_{\Omega} Z(\bar n)\,F(Z_1,\dots,Z_{n-1};d\omega), \tag{6}$$

and if

$$Z_1(\bar n)=H_1^0(\bar n)-H_1^1(\bar n),\dots,Z_{n-1}(\bar n)=H_{n-1}^0(\bar n)-H_{n-1}^1(\bar n),$$

then

$$F(Z_1,\dots,Z_{n-1};\,\omega)=\sum(-1)^{\sum i_k}F(H_1^{i_1},\dots,H_{n-1}^{i_{n-1}};\,\omega), \tag{7}$$

where summation is taken over all combinations of $i_1,\dots,i_{n-1}$, which take the values 0 and 1.

In this way the “mixed area functions” $F(Z_1,\dots,Z_{n-1};\omega)$ are uniquely defined for the differences of convex functions. There is no need here to prove that they have all the properties of the mixed area functions of convex bodies, except, of course, nonnegativity.

The concepts of mixed volumes and mixed area functions were formally extended to the differences of convex functions for the following purpose. Minkowski as well as Hilbert had proved that several problems in the theory of convex bodies could be resolved by reducing them to variational problems. For instance, the foremost examples are the famous Minkowski theorem on the existence of a convex polyhedron with given areas of faces and the Hilbert method of proving inequalities between mixed volumes. Suppose that we have reduced in some way a certain problem to search for extrema of some functional $f(H)$ defined for convex bodies. Also, assume that the existence of extrema has been demonstrated, as this can be done without extra effort for many problems relating to convex bodies. However, the formal rules of variational calculus are, generally speaking, inadequate for investigating these extrema. For these rules to be applicable, it is necessary that the functional makes sense both for $H(\bar n)+\delta H(\bar n)$ and for $H(\bar n)-\delta H(\bar n)$. At the same time, the difference $H(\bar n)-\delta H(\bar n)$ is in general not convex, while $H(\bar n)$ and $H(\bar n)+\delta H(\bar n)$ are convex. If, however, our functional is defined for the differences of convex functions and has a first variation, then this obstacle falls away. In Part II of this paper we make use of

the possibility of applying the rules of variational calculus to the extended mixed volume $V(Z, Z, H_1, \dots, H_{n-1})$ in proving the uniqueness of a convex body with given curvature functions. Of course, we could as well use linear combinations of mixed volumes and area functions of convex bodies defined by formulas (3) and (7), but this would make the calculation and the results rather tedious.

These general considerations also require that we extend the concept of mixed volume to positive continuous functions on the unit sphere. We do this in Part III. This extension allows us to generalize the Minkowski theorem concerning polyhedra without digressing much from the simple ideas of the proof.

Let $Z(\bar{n}), Z_1(\bar{n}), \dots, Z_{n-m}(\bar{n})$ be given differences of convex functions. By definition,

$$\delta V(Z, \dots, Z,\, Z_1, \dots, Z_{n-m}) = \lim_{t=0} \frac{V(Z + t\delta Z, \dots) - V(Z, \dots)}{t}, \tag{8}$$

where $\delta Z(\bar{n})$ is a difference of convex functions and

$$V(Z + t\delta Z, \dots,\, Z + t\delta Z,\, Z_1, \dots, Z_{n-m}) = \sum_{k=0}^{m} t^k C_m^k\, V(\underbrace{\delta Z, \dots, \delta Z}_{k}, Z, \dots, Z,\, Z_1, \dots, Z_{n-m}). \tag{9}$$

Therefore $\delta V(Z, \dots, Z,\, Z_1, \dots, Z_{n-m})$ exists and is equal to

$$m\, V(\delta Z,\, Z, \dots, Z,\, Z_1, \dots, Z_{n-m}). \tag{10}$$

Up to a factor, the variation thus found is an extended mixed volume. Therefore, applying the method of § 1, it can be extended, and in a unique way, to arbitrary continuous functions $\delta Z(\bar{n})$ on the unit sphere. From (10) it follows that

$$\delta V(Z, \dots, Z,\, Z_1, \dots, Z_{n-m}) = \frac{1}{n} \int_{\Omega} \delta\, Z(\bar{n})\, F(Z, \dots, Z,\, Z_1, \dots, Z_{n-m};\, d\omega). \tag{11}$$

If this variation is defined, then $F(Z, \dots, Z,\, Z_1, \dots, Z_{n-m}; \omega)$ is also defined because, by the Riesz theorem, the linear functional $\delta Z(\bar{n})$ defined for all continuous functions can be uniquely represented as an integral of $\delta Z(\bar{n})$ over a countably additive set function.

We now prove a lemma needed in the sequel.

Lemma. *If the sequences of convex bodies*

$$H_1^{(1)}, H_1^{(2)}, \dots, H_1^{(m)}, \ \dots \ , H_{n-1}^{(1)}, H_{n-1}^{(2)}, \dots, H_{n-1}^{(m)}, \ \dots$$

converge to bodies $H_1, H_2, \dots, H_{n-1}$, *then for every continuous function* $Z(\bar{n})$

$$V(Z,\, H_1, \dots, H_{n-1}) = \lim_{m=\infty} V\Big(Z,\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big). \tag{12}$$

In other words, the convergence of these sequences of convex bodies implies the weak convergence of the functional $V\Big(Z,\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big)$.

Approximating the continuous function $Z(\bar{n})$ by the difference of convex functions $Z'(\bar{n})$ such that

$$|V(Z,\, H_1, \dots, H_{n-1}) - V(Z',\, H_1, \dots, H_{n-1})| < \frac{\varepsilon}{3} \tag{13}$$

(where, as usual, ε is an arbitrary positive real number) and, for all m,

$$\Big| V\Big(Z,\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big) - V\Big(Z',\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big) \Big| < \frac{\varepsilon}{3}. \tag{14}$$

This can be done because, in general, for arbitrary $H_1, \dots, H_{n-1}$, we have

$$\begin{aligned} &|V(Z,\, H_1, \dots, H_{n-1}) - V(Z',\, H_1, \dots, H_{n-1})| \\ &\qquad \leqslant \max |Z - Z'|\ V(E,\, H_1, \dots, H_{n-1}). \end{aligned} \tag{15}$$

Therefore it suffices to choose $Z'(\bar{n})$ so that

$$|Z(\bar{n}) - Z'(\bar{n})| < \frac{\varepsilon}{3B},$$

where B is an upper bound for $V\Big(Z,\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big)$.

Let $Z'(\bar{n}) = H'(\bar{n}) - H''(\bar{n})$. By the continuity of mixed volumes, there exists an M such that, for $m > M$, the volumes $V\big(H', H_1^{(m)},\, \dots, H_{n-1}^{(m)}\big)$ and $V\big(H'', H_1^{(m)},\, \dots, H_{n-1}^{(m)}\big)$ differ slightly from their limits so that for their difference, i.e., for $V\big(Z',\, H_1^{(m)},\, \dots, H_{n-1}^{(m)}\big)$, we have

$$\Big| V\Big(Z',\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big) - V(Z',\, H_1, \dots, H_{n-1}) \Big| < \frac{\varepsilon}{3}. \tag{16}$$

From (13), (14) and (16) it is immediate that for $m > M$

$$\Big| V(Z,\, H_1, \dots, H_{n-1}) - V\Big(Z,\, H_1^{(m)}, \dots, H_{n-1}^{(m)}\Big) \Big| < \varepsilon.$$

This completes the proof of our lemma.

In conclusion we note that the results derived here can be carried over to the case in which, as in Minkowski geometry, the fundamental body is not the ball but rather an arbitrary convex body having interior points. In this case, the n-dimensional space in which we consider convex bodies is no longer the Euclidean space, but rather a linear normed space whose metric is not necessarily symmetric.

Institute of Mathematics and Mechanics,
Leningrad State University

Received March 31, 1937

CHAPTER IV

TO THE THEORY OF MIXED VOLUMES OF CONVEX BODIES PART II[1]

Matematicheskiĭ Sbornik, vol. 2 (44), No. 6, 1205–1238 (1937).

NEW INEQUALITIES FOR MIXED VOLUMES AND THEIR APPLICATIONS

Using the Brunn theorem, Minkowski demonstrated that the inequality

$$V(H_1,\dots,H_1,H_2)^2 \geqslant V(H_1,\dots,H_1)\,V(H_1,\dots,H_1,H_2,H_2)$$

holds between the mixed volumes of two convex bodies. Basing on it, we prove a

Theorem. *Let $H_1, H_2, \dots, H_{n-1}$ be convex bodies in Euclidean space of dimension n and let Z be the difference of some convex functions on the unit sphere,[2] then*

$$V(H_1,\dots,Z)^2 \geqslant V(H_1,\dots,H_{n-1},H_{n-1})\,V(H_1,\dots,H_{n-2},Z,Z).$$

In particular, if $Z = H_n$ is the support function (on the unit sphere) of a convex body H_n, then

$$V(H_1,\dots,H_n)^2 \geqslant V(H_1,\dots,H_{n-1},\,H_{n-1})\,V(H_1,\dots,H_{n-2},\,H_n,H_n).$$

[1]For Part I "Extension of certain concepts of the theory of convex bodies," see Chapter III. In the sequel the reader is supposed to be familiar with the results of Part I.

[2]By the definition given in Part I, a convex function on the sphere is the support function of a convex body which is restricted to unit vectors.

If the convex bodies $H_1, \ldots, H_{n-1}$ have interior points, then

$$V(H_1, \ldots, H_{n-1}, H_{n-1}) > 0.$$

Therefore if Z satisfies the condition $V(Z_1, \ldots, H_{n-1}, Z) = 0$, then

$$V(H_1, \ldots, H_{n-2}, Z, Z) \leqslant 0.$$

§ 1. Analogous Polyhedra

We study convex polyhedra in n-dimensional space which are not degenerate (i.e., have interior points). As no other types of polyhedra are considered, we omit the adjectives "convex" and "nondegenerate" in our discussion.

Two polygons are said to be *analogous*, if to each side of one there corresponds a parallel side of the other and conversely.[3] Two polyhedra are said to be *analogous*, if to each $(n-1)$-dimensional face of one there corresponds a parallel and analogous face of the other and conversely. Several polyhedra are said to be analogous to one another, if each pair of them is analogous.

Lemma I. *If two polyhedra H_1 and H_2 are analogous, then they have identical* **habitus**, *i.e.,*

(1) *To each face of H_1 there corresponds a parallel and analogous face of H_2 and vice versa,*
(2) *Corresponding faces meet at corresponding faces (of lower dimension),*
(3) *To each vertex of the polyhedron H_1 there corresponds a vertex of the polyhedron H_2 and vice versa; the faces of either polyhedron at the corresponding vertices form homothetic polyhedral angles (i.e., one is a translate of the other).*

We distinguish the corresponding faces and vertices of the polyhedra H_1 and H_2 by affixing the superscripts 1 and 2. For the sake of brevity, we call a polyhedral angle simply an angle.

1. By the definition of analogy, $(n-1)$-dimensional faces of H_1 are analogous to their parallel faces of H_2. Again by analogy, (lower-dimensional) faces of these faces are analogous, and so on.

2. Let $P_i^{(1)}$ and $P_k^{(1)}$ be $(n-1)$-dimensional faces of H_1 adjacent at an $(n-2)$-dimensional face $P_{ik}^{(1)}$. If the faces $P_i^{(2)}$ and $P_k^{(2)}$ of H_2 are not adjacent at $P_{ik}^{(2)}$, then $P_j^{(2)}$ is adjacent to $P_i^{(2)}$ at $P_{ik}^{(2)}$. Let $\bar{n}_i, \bar{n}_k$ and $\bar{n}_j$ be the normals to $P_i^{(2)}, P_k^{(2)}$ and $P_j^{(2)}$ respectively. They all are parallel to

[3] By parallelism we mean the parallelism of outward normals.

the same two-dimensional plane perpendicular to $P_{ik}^{(2)}$ and, consequently, to $P_{ik}^{(1)}$. Project H_1 and H_2 to this plane. The projections of $P_i^{(2)}$ and $P_j^{(2)}$ meet at the vertex that is the projection of $P_{ik}^{(2)}$. On the contrary, when projected, the corresponding sides of H_1 do not meet at one vertex but the sides with normals $\bar{n}_i$ and $\bar{n}_k$ meet at one vertex. Under such a situation, the projection of at least one of the polyhedra H_1 and H_2 is not convex. Consequently, our assumption is false, so $P_i^{(2)}$ meets $P_k^{(2)}$ at $P_{ik}^{(2)}$. This reasoning can also be applied to faces of lower dimension.

3. The third part of the lemma is obvious for polygons. Suppose that it holds for $(n-1)$-dimensional polyhedra, i.e., for the faces of an n-dimensional polyhedron. Let $A^{(1)}$ be a vertex of the polyhedron H_1 where its $(n-1)$-dimensional faces $P_1^{(1)}, P_2^{(1)}, \ldots$ meet. Let $A^{(2)}$ be the corresponding vertex of H_2. Now translate H_2 so that $A^{(2)}$ coincides with $A^{(1)}$. In this case the angle at the vertex $A^{(2)}$ of $P_1^{(2)}$ coincides with the angle at the vertex $A^{(1)}$ of $P_1^{(1)}$ because, by assumption, the lemma holds for faces. If the face $P_2^{(2)}$ is adjacent to $P_1^{(2)}$ at $P_{12}^{(2)}$, then, since the angle of $P_{12}^{(2)}$ coincides with the angle of $P_{12}^{(1)}$, the angles $P_2^{(2)}$ and $P_2^{(1)}$ at the vertices $A^{(2)}$ and $A^{(1)}$ also coincide. Continuing this reasoning, we conclude that the angle on all the faces $P_1^{(2)}, P_2^{(2)}, \ldots$ coincide with the angles on the faces $P_1^{(1)}, P_2^{(1)}, \ldots$. Consequently, the angles at the vertices $A^{(2)}$ and $A^{(1)}$ coincide on the polyhedra themselves.

Lemma II. *If the vertices of the polyhedra H_1 and H_2 can be put in correspondence in pairs so that the polyhedral angles at the corresponding vertices are homothetic, then the polyhedra H_1 and H_2 are analogous.*

The lemma is obvious for polygons because pairwise parallelism of sides is implied in the homothety of angles. Now suppose that the lemma valid for $(n-1)$-dimensional polyhedra.

Let $A^{(1)}$ and $A^{(2)}$ be the corresponding vertices of the polyhedra H_1 and H_2. Since the angles at $A^{(1)}$ and $A^{(2)}$ are homothetic, first, a face of H_1 and the corresponding face of H_2 which is parallel to it meet at these vertices so that to each face of H_1 there corresponds a parallel face of H_2 and conversely. Second, the angles at the vertices $A^{(1)}$ and $A^{(2)}$ on these faces are also homothetic. Consequently, parallel faces of H_1 and H_2 are analogous. This means the polyhedra themselves are analogous.

Lemma III. *Let $H_1, \ldots, H_m$ be certain given polyhedra (possibly degenerate) and $\lambda_1, \ldots, \lambda_m$ variable positive real numbers. Then the polyhedra $H = \lambda_1 H_1 + \cdots + \lambda_m H_m$ are analogous to one another if they are not*

degenerate.[4]

An $(n-1)$-dimensional face of the polyhedron $H = \lambda_1 H_1 + \cdots + \lambda_m H_m$ is obtained as a linear combination of the faces of $H_1, \ldots, H_m$ lying in parallel supporting planes. Let $P = \lambda_1 P^{(1)} + \cdots + \lambda_m P^{(m)}$ be an $(n-1)$-dimensional face of H. It exists if and only if it is possible to draw $(n-1)$ segments on $P^{(1)}, \ldots, P^{(m)}$ none of which is parallel to a plane[5] of dimension less than $(n-1)$. If this is not possible, then the faces $P^{(1)}, \ldots, P^{(m)}$ can be translated to one such plane including a face P, the result of translation, and consequently of dimension less than $(n-1)$. However, if this is possible and if $a_1, \ldots, a_{n-1}$ are the segments we need, then, clearly by the definition of translation, P modules an $(n-1)$-dimensional parallelepiped with sides $\lambda_{i_1} a_1, \lambda_{i_2} a_2, \ldots, \lambda_{i_{n-1}} a_{n-1}$, provided a_1 lies in $P^{(i_1)}$, a_2 in $P^{(i_2)}$ and so on. Therefore, if $\lambda_1, \ldots, \lambda_m$ all are greater than zero, then P is a face of every polyhedron H, if it is the face of at least one such polyhedron.

Hence our lemma holds for polygons. Suppose that it holds for a polyhedron of dimension less than $(n-1)$. Since $(n-1)$-dimensional faces on the polyhedron H are obtained by mixing the faces of the polyhedra $H_1, \ldots, H_m$, they are, by the assumption of induction, analogous. Therefore the polyhedra themselves are analogous.

Lemma IV. *Given finitely many analogous polyhedra, it is possible to made them primitive by arbitrarily small displacements of their* $(n-1)$*-dimensional faces, keeping the polyhedra analogous.*

A polyhedron is said to be *primitive*, if not more than n faces of dimension $(n-1)$ meet at each of its vertices. Every face of a primitive polyhedron is also primitive. The number of $(n-1)$-dimensional faces meeting at a vertex is said to be the *multiplicity* of the vertex.

It suffices to demonstrate the lemma for a pair of analogous polyhedra which we denote by H_1 and H_2. Let $A^{(1)}$ and $A^{(2)}$ be the corresponding vertices of H_1 and H_2. By Lemma I, their multiplicities are equal. Suppose that their multiplicity is greater than n. Let us move the planes of two corresponding faces meeting at $A^{(1)}$ and $A^{(2)}$ over their outward normal. If the displacement is sufficiently small, then the multiplicities of the vertices $A^{(1)}$ and $A^{(2)}$ of the polyhedra H_1 and H_2 do not increase: they are decreased. Furthermore, new vertices and new faces appear. Clearly, no new $(n-1)$-dimensional faces appear and, if the displacement is sufficiently small, none of them vanishes. New vertices appear as a result of "splitting" of the vertices belonging to the displaced faces. Therefore, after a sufficiently small displacement, the new vertices lie near the old ones. Since the

[4]This lemma was first proved by Minkowski *(Theorie der konvexen Körper*, § *19)*.

[5]An $(n-1)$-dimensional plane we call a plane, without mentioning its dimension. In all other cases the dimension is mentioned.

angles at the old vertices were homothetic, the angles at the new vertices are also be homothetic. Consequently, after this operation, the polyhedra H_1 and H_2 remain analogous. However, repeating this operation with different faces a large number of times, we finally arrive at a situation where the multiplicity of each vertex is minimal, i.e., equal to n.

Lemma V. *Finitely many convex bodies can be approximated with any accuracy by primitive polyhedra analogous to one another.*

Suppose that we are given m convex bodies. Let us approximate them by nondegenerate polyhedra $H_1, \dots, H_m$. This, we know, is possible. Choosing a sufficiently small $\varepsilon > 0$, construct $H_1 + \varepsilon H_2 + \dots + \varepsilon H_m$, $\varepsilon H_1 + H_2 + \dots + \varepsilon H_m$, $\varepsilon H_1 + \dots + \varepsilon H_{m-1} + H_m$ which, by Lemma III, are analogous polyhedra. We make them primitive through Lemma IV.

§ 2. Mixing of Analogous Primitive Polyhedra

We now study primitive polyhedra which are analogous to one another. The system of outward normals $\bar{n}_1, \dots, \bar{n}_N$ to their $(n-1)$-dimensional faces is common to them, i.e., each polyhedron has the same system of outward normals. The distance between the plane of an $(n-1)$-dimensional face of H and the origin (positive in the direction of the outward normal and negative in the opposite direction) is called the *support number* of the polyhedron. A system of arbitrary numbers $Z_1, \dots, Z_N$ assigned to the normals $\bar{n}_1, \dots, \bar{n}_N$ is denoted by Z. Here λZ stands for the system of numbers $\lambda Z_1, \dots, \lambda Z_N$, and $Z + Y$ for the system of numbers $Z_1 + Y_1, \dots, Z_N + Y_N$. If H denotes a polyhedron, then H is also used to denote its system of support numbers. Obviously, a polyhedron is completely determined when its system of support numbers is defined for the given system of its normals.

Lemma I. *If H is a primitive polyhedron and Z an arbitrary set of numbers assigned to its normals, then for a sufficiently small ε, the polyhedron $H + \varepsilon Z$ is analogous to H.*

The polyhedron $H + \varepsilon Z$ is obtained from the polyhedron H by displacing the plane of each $(n-1)$-dimensional face to a distance εZ_i along the normal. If ε is sufficiently small, then the multiplicities of vertices do not increase and no new vertices appear because each vertex has the least possible multiplicity. Therefore the polyhedron $H + \varepsilon Z$ has just as many vertices as the polyhedron H and, accordingly, faces which are parallel and correspond to each other meet at the corresponding vertices. Hence, by Lemma II of § 1, the polyhedron $H + \varepsilon Z$ is analogous to the polyhedron H. Furthermore, a polyhedron analogous to a primitive polyhedron is itself primitive as is implied in the homothety of the angles of analogous polyhedra.

Lemma II. *If $H^{(1)}, \dots, H^{(m)}$ are analogous polyhedra, and $\lambda_1, \dots, \lambda_m$ are arbitrary nonnegative numbers, not all zero, then the polyhedron $H = \lambda_1 H^{(1)} + \dots + \lambda_m H^{(m)}$ is analogous to the polyhedra $H^{(1)}, \dots, H^{(m)}$.*

Parallel faces analogous to one another[6] lie on parallel supporting planes to the polyhedra $H^{(1)}, \dots, H^{(m)}$. Hence the $(n-1)$-dimensional faces of H are obtained exclusively by mixing the corresponding $(n-1)$-dimensional faces of polyhedra $H^{(1)}, \dots, H^{(m)}$. The $(n-1)$-dimensional faces of H are therefore parallel to the $(n-1)$-dimensional faces of the polyhedra $H^{(1)}, \dots, H^{(m)}$. Second, our lemma holds for polygons; so, we assume that it also holds for the faces of a polyhedron (i.e., for a polyhedron of dimension less than n). And, since the faces of H are obtained by mixing the analogous faces of $H^{(1)}, \dots, H^{(m)}$, the polyhedron H is analogous to the polyhedra $H^{(1)}, \dots, H^{(m)}$. The support number H_i is simply the value of the support function of the polyhedron H at the corresponding normal n_i, i.e.,

$$H_i = H(\bar{n}_i).$$

The support function of the polyhedron $H = \lambda_1 H^{(1)} + \dots + \lambda_m H^{(m)}$ is a like linear combination of the support functions of its constituent polyhedra $H^{(1)} \dots H^{(m)}$. Therefore the support numbers of the polyhedron H are

$$H_i = \lambda_1 H_i^{(1)} + \dots + \lambda_m H_i^{(m)}.$$

Since every face of H is parallel to some face, one on each of the polyhedra $H^{(1)} \dots H^{(m)}$, there are no more support numbers than what are given above. Thus, the mixing of analogous polyhedra gives rise to the same linear combination of their support numbers.

Lemma III. *Let $H^{(1)} \dots H^{(n)}$ be variable polyhedra always remaining analogous to a given polyhedron. Then their mixed volume $V(H^{(1)} \dots H^{(n)})$ is a homogeneous polynomial of degree n in their support numbers.* (The coefficients of the polynomial are naturally constants. They vary only when $H^{(1)} \dots H^{(n)}$ cease to be analogous to the given polyhedron.)

We know that

$$V(H^{(1)} \dots H^{(n)}) = \frac{1}{n} \sum_i H^{(1)}(\bar{n}_i) F_i(H^{(2)} \dots H^{(n)}), \tag{1}$$

where $F_i(H^{(2)} \dots H^{(n)})$ is the $(n-1)$-dimensional mixed volume of the faces of the polyhedra $H^{(2)} \dots H^{(n)}$ that lie on the supporting planes with

[6] This follows from Lemma 1, § 1. Homothety of angles implies that parallel supporting planes contain only the corresponding vertices and, consequently, the corresponding faces.

normal $\bar{n}_i$. If $H^{(2)}, \ldots, H^{(n)}$ are analogous, then the $(n-1)$-dimensional faces of the polyhedron obtained by mixing them are obtained by mixing the corresponding faces of the polyhedra $H^{(2)}, \ldots, H^{(n)}$. Therefore $F_i(H^{(2)}, \ldots, H^{(n)}) \neq 0$, if and only if $\bar{n}_i$ is the normal to the ith $(n-1)$-dimensional face of these polyhedra. In this case $H^{(1)}(\bar{n}_i) = H_i^{(1)}$ is the ith support number of the polyhedron $H^{(1)}$. Thus, for analogous polyhedra, we have

$$V(H^{(1)}, \ldots, H^{(n)}) = \frac{1}{n} \sum_{i=1}^{N} H_i^{(1)} F_i(H^{(2)}, \ldots, H^{(n)}). \tag{2}$$

Let θ denote the angle between the normals[7] $\bar{n}_i$ and $\bar{n}_k$. Take the projection of the spatial origin to the plane of the ith face of the polyhedron H as the origin O on this plane. The distance between O and the intersection of the ith and kth faces is

$$H_{ik} = \frac{H_k - H_i \cos\theta_{ik}}{\sin\theta_{ik}}, \tag{3}$$

as can be shown by simple calculation. The number H_{ik} is a support number of the ith face, only if the corresponding intersection of these planes actually yields some $(n-2)$-dimensional face of the ith face.

Applying formula (2) to the mixed volume $F_i(H^{(2)}, \ldots, H^{(n)})$ of the faces of the polyhedra $H^{(2)}, \ldots H^{(n)}$, we obtain

$$F_i(H^{(2)}, \ldots, H^{(n)}) = \frac{1}{n-1} \sum_k H_{ik}^{(2)} F_{ik}(H^{(3)}, \ldots, H^{(n)}), \tag{4}$$

where $F_{ik}(H^{(3)}, \ldots, H^{(n)})$ is the $(n-2)$-dimensional mixed volume of the faces of the polyhedra $H^{(3)}, \ldots, H^{(n)}$. Since the faces of analogous polyhedra are analogous, the note to formula (1) applies equally to formula (4). Therefore in (4) $F_{ik}(H^{(3)}, \ldots, H^{(n)}) \neq 0$, if and only if each of the polyhedra $H^{(3)}, \ldots, H^{(n)}$ has an ikth $(n-2)$-dimensional face, i.e., if and only if

$$H_{ik}^{(2)} = \frac{H_k^{(2)} - H_i^{(2)} \cos\theta_{ik}}{\sin\theta_{ik}}$$

are the support numbers for the ith face of the polyhedron $H^{(2)}$.

Repeating our reasoning for the mixed volumes of faces of lower and lower dimensions, we eventually arrive at edges. Thus, as is clear from (3), each time we find that the support numbers of the $(k-1)$-dimensional faces are linear expressions of the support numbers of the k-dimensional faces, that is to say, of the support numbers of the given polyhedra. The coefficients in these linear expressions depend only on the directions of the normals

[7]The system of normals is always supposed to be given.

$\bar{n}_1, \ldots, \bar{n}_N$. The coefficients $F_{ik\ldots p}$ in the expressions [of the type (4)] for the mixed volumes of k-dimensional faces vanish for the same combinations of the subscripts $i, k, \ldots, p$ only if the variable polyhedra $H^{(1)}, \ldots, H^{(n)}$ remain analogous to the given polyhedron. Finally, as it follows from (3), the length of the qth edge belonging to some two-dimensional face with support numbers $h_1, \ldots, h_r$ and with angles $\varphi_{12}, \ldots, \varphi_{rl}$ between the normals to adjacent edges is

$$l_q = \frac{h_{q-1} - h_q \cos\varphi_{q-1,q}}{\sin\varphi_{q-1,q}} + \frac{h_{q+1} - h_q \cos\varphi_{q,q+1}}{\sin\varphi_{q,q+1}}. \tag{5}$$

Substituting these expressions into (2), we obtain $V(H^{(1)}, \ldots, H^{(n)})$ in the form of a homogeneous polynomial of degree n in the support numbers of the polyhedra $H^{(1)}, \ldots, H^{(n)}$.

Indeed, as it follows from our reasoning, the coefficients of this polynomial are constants if the polyhedra $H^{(1)}, \ldots, H^{(n)}$ are analogous to a given polyhedron. In the polynomial thus determined, we can substitute an arbitrary system of numbers $Z^{(1)}, \ldots, Z^{(n)}$ for the support numbers $H^{(1)}, \ldots, H^{(n)}$.

The whole set of arguments can be applied with equal success to corresponding faces of our polyhedra as they themselves are analogous polyhedra. Therefore $F_i(H^{(2)}, \ldots, H^{(n)})$ is a homogeneous polynomial of degree $n-1$ in the support numbers of the ith faces. And, by virtue of formula (3), it is a homogeneous polynomial of degree $n-1$ in the support numbers of the polyhedra $H^{(2)}, \ldots, H^{(n)}$. Hence, what we mean by $F_i(Z^{(2)}, \ldots, Z^{(n)})$ is quite clear.

Lemma IV. *The polynomials $V(Z^{(1)}, \ldots, Z^{(n)})$ and $F_i(Z^{(2)}, \ldots, Z^{(n)})$ satisfy the following conditions:*

(1) *They are invariant under the permutations of the numbers $Z^{(k)}$,*
(2) $V((Z^{(1)}, \ldots, Z^{(n)}) = \frac{1}{n}\sum_{i=1}^{N} Z_i^{(1)} F_i(Z^{(2)}, \ldots, Z^{(n)})$,
(3) $V(\lambda Z + \mu Y, Z^{(2)}, \ldots, Z^{(n)}) = K_1 + K_2$,

where $K_1 = \lambda V(Z, Z^{(2)}, \ldots, Z^{(n)})$, $K_2 = \mu V(Y, Z^{(2)}, \ldots, Z^{(n)})$, and the same is true of $F_i(Z^{(2)}, \ldots, Z^{(n)})$.

Conditions (1) and (3) follow from similar properties of mixed volumes, and condition (2) from formula (2).

§ 3. Proof of the Main Inequality

Theorem. [8] *If $H^{(1)}, \dots, H^{(n-1)}$ are primitive analogous polyhedra, and Z an arbitrary system of numbers assigned to their normals, then*

$$V(H^{(1)}, \dots, H^{(n-1)}, Z)^2 \geqslant V(H^{(1)}, \dots, H^{(n-1)}, H^{(n-1)}) \\ \times V(H^{(1)}, \dots, H^{(n-2)}, Z, Z), \tag{1}$$

where equality holds, if and only if Z is a system of support numbers for a polyhedron homothetic to $H^{(n-1)}$.

First, we demonstrate that our theorem is equivalent to the following

Theorem. *Let $H^{(1)}, \dots, H^{(n-1)}$ have the same meaning [as in the previous theorem] and, in addition, let Z satisfy the condition*

$$V(H^{(1)}, \dots, H^{(n-1)}, Z) = 0, \tag{2}$$

then

$$V(H^{(1)}, \dots, H^{(n-2)}, Z, Z) \leqslant 0 \tag{3}$$

and vanishes if and only if Z is the system of support numbers of a single point.

We now prove just this assertion. First, recall that a polyhedron homothetic to H can be represented as $\lambda H + \bar{a}$, where λ is the coefficient of similitude transformation to which H is subjected, and $\bar{a}$ is the translation vector. Therefore the system of support numbers of a polyhedron homothetic to H is expressed as $\lambda H_i + \bar{a}\bar{n}_i$, where $\bar{a}\bar{n}_i$ are simply the support numbers of the point $\bar{a}$. The converse is equally true, namely, a *polyhedron having the same system of normals as the polyhedron H and whose support numbers are $\lambda H_i + \bar{a}\bar{n}_i$ is homothetic to H.* For the sake of brevity, by A we denote the system of support numbers of a point $\bar{a}$ that correspond to the normals $\bar{n}_i$. Using this notation, we can assert that

$$V(H^{(1)}, \dots, H^{(n-1)}, A) = 0. \tag{4}$$

Since $V(H^{(1)}, \dots, H^{(n-1)}, H^{(n-1)}) > 0$, from (1), under condition (2), we actually obtain inequality (3).

---

[8] The idea underlying the proof of this theorem was published in Dokl. Akad. Nauk SSSR, **14**, 4 (1937), where the polyhedra were supposed to be only analogous without being primitive. The numbers Z therein are therefore not arbitrary but are the differences of support numbers of polyhedra analogous to the given ones.

Now let Z be arbitrary. Choose a λ such that

$$V(H^{(1)},\dots,H^{(n-1)},Z)=\lambda V(H^{(1)},\dots,H^{(n-1)},H^{(n-1)}). \tag{5}$$

Such a λ does exist because $V(H^{(1)},\dots,H^{(n-1)},H^{(n-1)})>0$. Put

$$Z-\lambda H^{(n-1)}=Z', \tag{6}$$

then

$$V(H^{(1)},\dots,H^{(n-1)},Z')=0, \tag{7}$$

and, if our second theorem is true, then

$$V(H^{(1)},\dots,H^{(n-2)},Z',Z')\leqslant 0. \tag{8}$$

Substituting $Z'=Z-\lambda H^{(n-1)}$ and expanding in powers of λ [see item 3 of Lemma IV, § 2], after simple transformations, we obtain inequality (1). Equality holds in (1), if and only if equality holds in (8), i.e., by the assertion, only if $Z'=A$, which is to say, only if $Z=\lambda H^{(n-1)}+A$.

Thus to prove our theorem it suffices to demonstrate just the second theorem, its equivalent.

Lemma I. *If $H^{(1)}=\cdots=H^{(n-1)}=H$, then our theorem is equivalent to the assertion that in the Minkowski inequality*

$$V(H,\dots,H,H')^2\geqslant V(H,\dots,H)V(H,\dots,H,H',H'), \tag{9}$$

where H and H' are analogous primitive polyhedra, with equality holding if and only if H and H' are homothetic.

That this assertion is implied in our theorem can be verified by substituting in (1) H' for Z, and H for $H^{(1)},\dots,H^{(n-1)}$. Suppose that our assertion also holds for the case of equality in the Minkowski inequality. Let Z satisfy the condition

$$V(H,\dots,H,Z)=0. \tag{10}$$

Take a small $\varepsilon\neq 0$ so that $H'=H+\varepsilon H$ is a polyhedron analogous to H. Since $\varepsilon Z=H'-H$, condition (10) yields

$$V(H,\dots,H,H')=V(H,\dots,H). \tag{11}$$

Under this condition, the Minkowski inequality (9) yields

$$V(H,\dots,H,H')\geqslant V(H,\dots,H,H',H'). \tag{12}$$

Subtracting the left side from the right side, and since $H' - H = \varepsilon Z$, we obtain

$$V(H, \ldots, H, H', \varepsilon Z) \leqslant 0. \tag{13}$$

Subtracting ε times equality (10) from (13), we obtain

$$V(H, \ldots, H, H', \varepsilon Z, \varepsilon Z) \leqslant 0. \tag{14}$$

Taking out the positive factor ε^2 and then omitting it, we finally infer

$$V(H, \ldots, H, Z, Z) \leqslant 0, \tag{15}$$

where equality holds if and only if it holds in (12), i.e., in the Minkowski inequality.

Lemma II. *Our theorem holds for polygons.*

Indeed, equality in the Minkowski inequality

$$V(H, H')^2 \geqslant V(H, H) \;\; V(H', H')$$

holds only if H and H' are homothetic. Therefore Lemma II follows from Lemma I.

On the basis of this lemma we can prove our theorem by induction, i.e., we assume that our theorem holds for the faces of a polyhedron.

The mixed volume $V(H^{(1)}, \ldots, H^{(n-2)}, Z, Z)$, as is clear from Lemma IV of § 2, is a quadratic form in the variables $Z_1, \ldots, Z_N$. We prove our theorem through a study of the eigenvalues of this form.

Lemma III. *For arbitrary primitive analogous polyhedra $H^{(1)}, \ldots, H^{(n-2)}$, the quadratic form $V(H^{(1)}, \ldots, H^{(n-2}, Z, Z)$ has zero eigenvalue of multiplicity n to which correspond the "eigenvectors" Z representing the support numbers of a point.*

For a given system of normals $\bar{n}_1, \ldots, \bar{n}_N$, there always exist exactly n linearly independent systems of support numbers of a point, where n denotes, as usual, the dimension of the space. Indeed, if A represents the system of support numbers of the point $\bar{a}$, then $A_i = \bar{a}\bar{n}_i$. And, since in each n-dimensional space there are at most n linearly independent vectors $\bar{a}$, there always exist exactly n linearly independent systems of numbers A_i.

An eigenvector Z of the form $V(H^{(1)}, \ldots, H^{(n-2)}, Z, Z)$ that corresponds to the zero eigenvalue of this form is a solution of the system of equations

$$\frac{\partial}{\partial Z_i} V(H^{(1)}, \ldots, H^{(n-2)}, Z, Z) = 0. \tag{16}$$

Turning to item 2 of Lemma IV of § 2, we can rewrite these equations as

$$F_i(H^{(1)}, \dots, H^{(n-2)}, Z) = 0 \qquad (i = 1, \dots, N), \tag{17}$$

where $F_i(H^{(1)}, \dots, H^{(n-2)}, Z)$ is a linear form which has the same meaning for the ith face as $V(H^{(1)}, \dots, H^{(n-1)}, Z)$ has for the polyhedra themselves. Therefore, assuming that our theorem holds for faces, we find, by virtue of (17), that

$$F_i(H^{(1)}, \dots, H^{(n-2)}, Z, Z) \leqslant 0. \tag{18}$$

We always assume that all $H_i^{(k)} > 0$. This can be achieved by translating the polyhedra so that the origin always lies in their interior. Since translation does not alter the mixed volume of polyhedra, our assumption does not restrict the generality of arguments.

Since $H_i^{(n-2)} > 0$, multiplying the corresponding inequalities (18) by H_i^{n-2} and then adding up [see item 2 of Lemma IV, § 2], we obtain

$$V(H^{(1)}, \dots, H^{(n-2)}, Z, Z) \leqslant 0. \tag{19}$$

Here equality holds if and only if equality holds in each of inequalities (18).

Similarly, multiplying (15) by Z_i and then summing over i, we obtain

$$V(H^{(1)}, \dots, H^{(n-2)}, Z, Z) = 0. \tag{20}$$

Consequently, equality holds in each inequality of (18). By the assumption of induction, this is possible only when Z gives a system of support numbers of a point in the plane of every $(n-1)$-dimensional face. For the ith face, we denote this system by $Z_{(i)}$. Since the polyhedron $H^{(n-2)}$ is primitive, for sufficiently small ε, we find that $H^{(n-2)} + \varepsilon Z$ is also a polyhedron analogous to $H^{(n-2)}$. When support numbers of polyhedra are summed up, the support numbers of their faces also become added as is clear from the linear relation between the former and the latter [see formula (3) of § 2]. The same is also true of the addition of arbitrary systems of numbers $Z_{(i)}$ in the planes of the faces in the same way. The numbers of the system $Z_{(i)}$ correspond to the normals of the $(n-2)$-dimensional faces of the ith $(n-1)$-dimensional face. Therefore, in our notation,

$$(H^{(n-2)} + \varepsilon Z)_{(i)} = H_{(i)}^{(n-2)} + \varepsilon Z_{(i)}. \tag{21}$$

Since $Z_{(i)}$ are systems of support numbers of points, each face of the polyhedron $H^{(n-2)} + \varepsilon Z$ is a translate of a face of the polyhedron $H^{(n-2)}$. Therefore these two polyhedra themselves are translates of one another. Hence Z is a system of support numbers of a point (the end of the vector, which moves $H^{(n-2)}$ to $H^{(n-2)} + \varepsilon Z$, multiplied by $\frac{1}{\varepsilon}$ and drawn from the origin).

Lemma IV. *Our theorem is equivalent to the assertion that the quadratic form $V(H^{(1)}, \ldots, H^{(n-2)}, Z, Z)$ has only one positive eigenvalue and a zero eigenvalue of multiplicity n stated in Lemma III.*

We reduce the form $V(H^{(1)}, \ldots, H^{(n-2)}, Z, Z)$ to a sum of squares by a transformation that preserve the form

$$\frac{1}{n} \sum_{i=1}^{N} \frac{F_i(H^{(1)}, \ldots, H^{(n-1)})}{H_i^{(n-1)}} Z_i^2, \tag{22}$$

which is positive-definite because

$$H_i^{(n-1)} > 0 \quad \text{and} \quad F_i(H^{(1)}, \ldots, H^{(n-1)}) > 0.$$

So, our problem is reduced to the system of equations

$$F_i(H^{(1)}, \ldots, H^{(n-2)}, Z) = \lambda \frac{F_i(H^{(1)}, \ldots, H^{(n-1)})}{H_i^{(n-1)}} Z_i, \tag{23}$$

[see item 2 of Lemma IV, § 2].

For $\lambda = 1$, this system has an obvious solution $Z = H^{(n-1)}$. Therefore the condition

$$V(H^{(1)}, \ldots, H^{(n-1)}, Z) = \frac{1}{n} \sum_{i=1}^{N} \frac{F_i(H^{(1)}, \ldots, H^{(n-1)})}{H_i^{(n-1)}} H_i^{(n-1)} Z_i = 0 \tag{24}$$

is simply a condition for the admissible Z's of $H^{(n-1)}$ to be "weighted" orthogonal, i.e., to the eigenvector corresponding to the eigenvalue $\lambda = 1$. If $V(H^{(1)}, \ldots, H^{(n-1)}, Z, Z)$ has no other positive eigenvalue, then, by condition (24), it is nonpositive and is equal to zero only when Z corresponds to zero eigenvalue. Conversely, if $V(H^{(1)}, \ldots, H^{(n-1)}, Z, Z)$ is nonpositive under condition (24) and is zero only when Z is the system of support numbers of a point, then $\lambda = 1$ is the only positive eigenvalue since if Z were to correspond to such an eigenvalue, then the orthogonality condition (24) would be satisfied and form itself would be positive.

Thus, to complete the proof of our theorem it only remains to demonstrate

Lemma V. *The quadratic form $V(H^{(1)}, \ldots, H^{(n-1)}, Z, Z)$ has only one positive eigenvalue.*

In Lemma I, starting from the Minkowski inequality, we showed that, if

$$V(H, \ldots, H, Z) = 0, \tag{25}$$

then

$$V(H, \dots, H, Z, Z) \leqslant 0. \tag{26}$$

Since, under the orthogonality condition, the quadratic form (26) is non-positive under the orthogonality condition (25), it has only one positive eigenvalue $\lambda = 1$ found in the course of proving the previous lemma.

Let us construct a family of polyhedra $H^{(1)}, (1-\vartheta)H^{(1)} + \vartheta H^{(2)}, \dots, (1-\vartheta)H^{(1)} + \vartheta H^{(n-2)}$. As ϑ increases from zero to one, these polyhedra vary continuously from $H^{(1)}$ to $H^{(1)}, H^{(2)}, \dots, H^{(n-2)}$. In this process they always remain analogous to $H^{(1)}$ as implied in Lemma II of § 2. And the coefficients of the form

$$V(H^{(1)}, (1-\vartheta)H^{(1)} + \vartheta H^{(2)}, \dots, (1-\vartheta)H^{(1)} + \vartheta H^{(n-2)}, Z, Z)$$

also change continuously. So its eigenvalues change continuously as well. However, by Lemma III, none of them can pass through zero (except the case, if any at all, where the eigenvalue remains equal to zero). Therefore the number of positive eigenvalues is invariable. For $\vartheta = 0$, it is equal to one; hence, for $\vartheta = 1$, it is also equal to one, which is what was required to be proven.

Thus we finally showed the main inequality between the mixed volumes for analogous, primitive polyhedra. It now remains to extend it to the general case. For this purpose, naturally, Lemma IV of § 2 is of use to us, which asserts that finitely many convex bodies can be approximated with any accuracy by primitive polyhedra analogous to one another.

Let $H_1, \dots, H_{n-1}$ be arbitrary convex bodies and Z the difference of convex functions H_n and H_{n+1}: $Z = H_n - H_{n+1}$. We now approximate the bodies $H_1, \dots, H_{n+1}$ by analogous primitive polyhedra $H_m^{(1)}, \dots, H_m^{(n+1)}$ which tend to H_m as n tends to infinity. For these polyhedra, by the proven theorem, we have

$$\begin{gathered} V(H^{(1)}, \dots, H^{(n-1)}, H^{(n)} - H^{(n+1)})^2 \\ \geqslant V(H^{(1)}, \dots, H^{(n-1)}, H^{(n-1)}) \\ \times V(H^{(1)}, \dots, H^{(n)} - H^{(n+1)}, H^{(n)} - H^{(n+1)}). \end{gathered}$$

As m tends to infinity, the approximating polyhedra $H_m^{(1)}, \dots, H_m^{(n-1)}$ tend to $H_1, \dots, H_{n-1}$ and the difference $H_m^{(n)}, \dots, H_m^{(n+1)}$ tends to Z. Hence the mixed volumes in the inequality tend to the corresponding mixed volumes and, in the limit, we obtain

$$\begin{aligned} V(H_1, \dots, H_{(n-1)}, Z)^2 \geqslant{} & V(H_1, \dots, H_{(n-1)}, H_{(n-1)}) \\ & \times V(H_1, \dots, H_{(n-2)}, Z, Z). \end{aligned}$$

§ 4. Inequalities Between Mixed Volumes of Convex Bodies

From the main inequality between mixed volumes

$$V(H_1, \dots, H_{n-1})^2 \geqslant V(H_1, \dots, H_{n-1}, H_{n-1})\, V(H_1, \dots, H_{n-2}, Z, Z) \quad (1)$$

proven in the previous section, we can derive a whole series of interesting inequalities between mixed volumes of convex bodies. Their derivation is a rather formal procedure, so we study only those inequalities that are important in application.

In the main inequality (1) putting $Z = H_n$ we obtain an inequality

$$V(H_1, \dots, H_n)^2 \geqslant V(H_1, \dots, H_{n-1}, H_{n-1})\, V(H_1, \dots, H_{n-2}, H_n, H_n) \quad (2)$$

between mixed volumes of convex bodies.

Let us fix the bodies $H_1, H_2, \dots, H_{n-m}$, and construct a linear family of bodies $H^\vartheta = (1-\vartheta)H^0 + \vartheta H^1$ $(0 \leqslant \vartheta \leqslant 1)$ joining the given convex bodies H^0 and H^1. Consider the mixed volume

$$V(\vartheta) = V(\underbrace{H^\vartheta, \dots, H^\vartheta}_{m}, H_1, \dots, H_{n-m}). \quad (3)$$

The Brunn inequality can be generalized as:

$$\sqrt[m]{V(\vartheta)} \geqslant (1-\vartheta)\sqrt[m]{V(0)} + \vartheta\sqrt[m]{V(1)}. \quad (4)$$

To prove this, consider the function

$$\Phi(\vartheta) = \sqrt[m]{V(\vartheta)} - (1-\vartheta)\sqrt[m]{V(0)} - \vartheta\sqrt[m]{V(1)}. \quad (5)$$

It vanishes on the endpoints of the interval $0 \leqslant \vartheta \leqslant 1$. So, to prove it is nonnegative, or which is to say, to prove (4), it suffices to demonstrate, first, its first derivative is nonnegative and, then, decreasing. In other words, it suffices to prove $\Phi''(0) \leqslant 0$. Indeed, let ϑ_1 be a given value of ϑ for which we wish to find the second derivative $\Phi''(\vartheta)$. Some members of the family of bodies H^ϑ, for $\vartheta \geqslant \vartheta_1$, can be represented again as a linear family joining the bodies H^{ϑ_1} and H^1. For this, put

$$\vartheta' = \frac{\vartheta - \vartheta_1}{1 - \vartheta_1}\ (0 \leqslant \vartheta \leq 1), \quad (6)$$

and, for $\vartheta \geqslant \vartheta_1$ and $H^\vartheta = H^{\vartheta'}$, we have

$$H^{\vartheta'} = (1-\vartheta')H^{\vartheta_1} + \vartheta' H^1. \quad (7)$$

To this family there corresponds a function $\bar{\Phi}(\vartheta')$ differing from $\Phi(\vartheta)$ by a linear term,[9] and

$$\frac{d^2\bar{\Phi}(0)}{d\vartheta'^2} = (1-\vartheta_1)^2 \frac{d^2\Phi(\vartheta_1)}{d\vartheta^2}, \tag{8}$$

so if $\bar{\Phi}''(0) \leqslant 0$, then $\bar{\Phi}''(\vartheta_1) \leqslant 0$. Recall that

$$\begin{aligned} &V(H^\vartheta, \dots, H^\vartheta, H_1, \dots, H_{n-m}) \\ &= \sum_{k=0}^{m} (1-\vartheta)^{m-k}\vartheta^k C_m^k V(\underbrace{H^0, \dots, H^0}_{m-k}, \underbrace{H^1, \dots, H^1}_{k}, H_1, \dots, H^{n-m}), \end{aligned} \tag{9}$$

or in obvious and more abbreviated notation,

$$V(\vartheta) = \sum_{k=0}^{m} (1-\vartheta)^{m-k}\vartheta^k C_m^k V_k. \tag{10}$$

Here $V(0) = V_0$ and $V(1) = V_m$. Now calculating $\Phi''(0)$, we find

$$\Phi''(0) = (m-1)\frac{V_0V_2 - V_1^2}{V_0^{\frac{2m-1}{m}}}. \tag{11}$$

Here the denominator contains a positive quantity. The volume V_0 could vanish, were the convex body in question degenerate. We however disregard this case because the generalized Brunn inequality (4) for this case is obtained by a trivial passage to the limit from nondegenerate bodies (having interior points) to degenerate bodies. The numerator of (11) contains the factor $V_0V_2 - V_1^2$, which is nonpositive because putting $H_n = H^1$, $H_{n-m+1} = \cdots = H_{n-1} = H^0$ in the main inequality (2), we obtain

$$V_1^2 \geqslant V_0V_2. \tag{12}$$

Consequently,

$$\Phi''(0) \leqslant 0.$$

Now assume that the first k bodies in (12) are H^1 and the remaining $(m-k)$ are H^0. Then we obtain an inequality that in our abbreviated notation takes the form

$$V_k^2 \geqslant V_{k-1}\, V_{k+1}. \tag{13}$$

---

[9] $\bar{\Phi}(\vartheta') = \Phi(\vartheta) + \Big[(1-\vartheta)\sqrt[m]{V(0)} + \vartheta\sqrt[m]{V(1)} - (1-\vartheta')\sqrt[m]{V(\vartheta_1)} + \vartheta'\sqrt[m]{V(1)}\Big]$.

Here k varies from 1 to $m-1$. They are called *intermediate quadratic inequalities*. Suppose that equality holds in the generalized Brunn inequality. Then

$$V(\vartheta) = [(1-\vartheta)\sqrt[m]{V(0)} + \vartheta\sqrt[m]{V(1)}\,]^m. \tag{14}$$

Expanding the right side by the Newton binomial formula and comparing it with $V(\vartheta)$ expressed in terms of the mixed volumes V_k [formula (10)], we obtain

$$V_k^m = V_0^{m-k}\,V_m^k. \tag{15}$$

Therefore, in this case, equality stands in all intermediate quadratic inequalities (13). We now sum up the results.

Theorem. *Let $H_1, \dots, H_{n-m}, H^0, H^1$ be given convex bodies and let $H^\vartheta = (1-\vartheta)H^0 + \vartheta H^1$ $(0 \leqslant \vartheta \leqslant 1)$. Then the following generalized Brunn inequality*

$$\sqrt[m]{V(\vartheta)} \geqslant (1-\vartheta)\sqrt[m]{V(0)} + \vartheta\sqrt[m]{V(1)}$$

holds, where, for brevity, $V(\vartheta)$ stands for $V(H^\vartheta, \dots, H^\vartheta, H_1, \dots, H_{n-m})$ and

$$V(\vartheta) = \sum_{k=0}^{m}(1-\vartheta)^{m-k}\vartheta^k C_m^k V_k.$$

The intermediate quadratic inequalities

$$V_k^2 \geqslant V_{k-1}V_{k+1} \quad (k = 1, \dots, m-1),$$

hold, and if equality is attained in the generalized Brunn inequality, then equality is attained in all the intermediate quadratic inequalities.

We now mention a particular case of the generalized Brunn inequality which plays a key role in applications. Put $H_1 = \cdots = H_{n-m} = E$, where E is the unit ball (in sequel we always use E to denote the unit ball). Recalling the results of § 2, Part I, we can write

$$V(\underbrace{H, \dots, H}_{m}, E, \dots, E) = V_m(H) = \frac{1}{m}\int\limits_{\Omega} F_m(H, d\omega), \tag{16}$$

because $E(\bar{n}) = 1$ if the origin is taken at the center of the unit ball E. Briefly,

$$V_m(H) = \frac{1}{n}F_m(H, \Omega)$$

(Ω is the whole surface of the unit ball), and

$$F_m(H, \omega) = F(\underbrace{H, \dots, H}_{m}, E, \dots, E; \omega)$$

is the mth curvature function of the body H. For a regular body it represents the integral of the mth curvature function in the usual sense (i.e., the elementary symmetric function of the principal radii of curvature) taken over the set ω. Thus, $V_m(H)$ represents, up to a factor, the integral of the curvature of H. In particular, for $m = n - 1$, we find that $V_{n-1}(H)$ is the area of the surface of H divided by n.

The mixed volume $V_m(H)$ has another geometric meaning. The m-dimensional volume of the projection of a body H to some m-dimensional plane is called the *m-dimensional cross-sectional measure* of the body H. Then $V_m(H)$ represents up to a constant factor the mean m-dimensional cross-sectional measure of the body H or the mth integral of cross-sectional measures (mth quermassintegral).[10])

Putting $H_1 = \cdots = H_{n-m} = E$ in the generalized Brunn inequality, we derive an assertion[11]: *if the bodies H^0 and H^1 have equal mth quermassintegrals (or, which is the same thing, equal mean m-dimensional cross-sectional measures, then the body $H^\vartheta = (1-\vartheta)H^0 + \vartheta H^1$ always has the integral of the curvature integral not less than that of H^0 or H^1.*

Later we show that the mth quermassintegral of H^ϑ is always greater, except in the case where the bodies H^0 and H^1 are translates of one another (excluding, of course, the case $m = 1$, for which we always have $V_1(H^\vartheta) = (1-\vartheta)V_1(H^0) + \vartheta V_1(H^1)$ and the case of degenerate bodies). (We note in parentheses that the volume of the body H can also be interpreted as the nth curvature integral. Then the theorem formulated above includes the Brunn theorem as well.)

We turn again to the function $\Phi(\vartheta) = \sqrt[m]{V(\vartheta)} - (1-\vartheta)\sqrt[m]{V(0)} - \sqrt[m]{V(1)}$. We proved that it is nonnegative and since $\Phi(0) = 0$, its first derivative must be nonnegative when $\vartheta = 0$ and vanish if and only if $\Phi(\vartheta) = 0$, i.e., if equality holds in the generalized Brunn inequality. Expressing $V(\vartheta)$ in terms of the mixed volumes V_k, we obtain

$$\Phi'(0) = \frac{V_1 - V_0}{V_0^{\frac{m-1}{m}}} + V_0^{\frac{1}{m}} - V_m^{\frac{1}{m}} \geqslant 0, \tag{17}$$

hence

$$V_1^m \geqslant V_0^{m-1} V_m, \tag{18}$$

[10]See Bonnesen and Fenchel, Theorie der konvexen Köper, § 7, Section 32 and § 8, Section 38.

[11]It was first proved by Minkowski for $m = n - 1$ and $m = 2$ (Volumen under Oberfläche, § 7). However, the question regarding equality remained unresolved.

or in explicit form

$$V(\underbrace{H^0,\ldots,H^0}_{m-1},H^1,H_1,\ldots,H_{n-m})^m \geqslant V(\underbrace{H^0,\ldots,H^0}_{m},H_1,\ldots,H_{n-m})^{m-1} \times V(\underbrace{H^1,\ldots,H^1}_{m},H_1,\ldots,H_{n-m}).$$

Here m is any number between 2 and n. We now prove the following general inequality which includes (2) and (18) as particular cases:

$$\boxed{V(H_1,\ldots,H_n)^m \geqslant \prod_{k=1}^{m} V(\underbrace{H^k,\ldots,H^k}_{m},H_{m+1},\ldots,H_n).} \tag{19}$$

For $m = 2$, it is reduced to inequality (2). Supposing it holds for an arbitrary m, we show that it also holds for $(m+1)$. To this end, note that (18) can be rewritten, substituting $m+1$ for m, as follows:

$$\begin{gathered} V(H_k,\ldots,H_k,H_{m+1},\ldots,H_n)^{m+1} \\ \geqslant V(H_k,\ldots,H_k,H_{m+2},\ldots,H_n)^m V(H_{m+1},\ldots,H_{m+1},\ldots,H_n) \end{gathered} \tag{20}$$

for $1 \leqslant k \leqslant m$.

Raise inequality (19) to the $(m+1)$th power. Then, using (20) in an obvious manner, and taking the mth root, we find

$$V(H_1,..,H_n)^{m+1} \geqslant \prod_{k=1}^{m+1} V(H_k,\ldots,H_k,H_{m+2},\ldots,H_n). \tag{21}$$

This completes the proof of our general inequality.

From the reasoning which leads to the proof of (21), it is clear that equality holds in (21), if and only if equality stands in each of the inequalities of the type (20) or, which is the same thing, in the corresponding generalized Brunn inequalities of the type (4).

For the particular case of $m = n$, inequality (19) yields

$$V(H_1,\ldots,H_n)^n \geqslant \prod_{k=1}^{n} V(H_k,\ldots,H_k), \tag{22}$$

i.e., the nth power of the mixed volume is greater than or equal to the product of the volumes of the bodies composing the mixed volume.

In application we often need an inequality between curvature integrals derived from (19). In (19) putting $H_1 = \cdots = H_k = H$ $(k < m)$, and $H_{k+1} = \cdots = H_n = E$, we obtain

$$V_k(H)^m \geqslant V_m(H)^k V(E)^{m-k}, \tag{23}$$

where $1 \leqslant k < m \leqslant n$, V is the volume of the unit ball and $V_m(H), V_k(H)$ have the meaning as defined in formula (16).

The equality holds in this inequality if and only if equality holds in all inequalities of the type:

$$\sqrt[l]{V_l(H_\vartheta)} \geqslant (1-\vartheta)\sqrt[l]{V_l(H)} + \vartheta\sqrt[l]{V_l(E)},$$

where $H_\vartheta = (1-\vartheta)H + \vartheta H + \vartheta E$ and $l \leqslant m$. This directly follows from the remark just made regarding equality in the general inequality (19).

Serious difficulties occur in determining the conditions for equality to hold in the general inequalities just derived. This question has not been solved even for the Minkowski quadratic inequality in three dimensions space. Clearly, this question is reduced to investigating the condition for equality to hold in the main inequality. For this we prove a lemma needed in the sequel which may pave the way for establishing the conditions for equality to hold in the general inequalities between mixed volumes.

Lemma. *For the equality to hold in the main inequality (1)*

$$V(H_1, \ldots, H_{n-1}, Z)^2 \geqslant V(H_1, \ldots, H_{n-1}, H_{n-1}) V(H_1, \ldots, H_{n-2}, Z, Z),$$

it is necessary and sufficient that the function Z satisfy the equation

$$\lambda F(H_1, \ldots, H_{n-1}; \omega) = \mu F(H_1, \ldots, H_{n-2}, Z;\ \omega), \tag{24}$$

where

$$\lambda = V(H_1, \ldots, H_{n-1}, Z), \quad \mu = V(H_1, \ldots, H_{n-1}, H_{n-1}).^{12}$$

Sufficiency. Integrating $Z(\bar n)$ over the set function in (24), obtain

$$\lambda \int\limits_\Omega \bar Z(\bar n) F(H_1, \ldots, H_{n-1}; d\omega) = \mu \int\limits_\Omega \bar Z(\bar n) F(H_1, \ldots, H_{n-2}, Z; d\omega),$$

and, after dividing by n,

$$V(H_1, \ldots, H_{n-1}, Z)^2 = V(H_1, \ldots, H_{n-1}, H_{n-1}) V(H_1, \ldots, H_{n-2}, Z, Z). \tag{25}$$

[12] For a description of the generalized mixed volume $V(H_1, \ldots, H_{n-2}, Z, Z)$ and the mixed area function $F(H_1, \ldots, H_{n-2}, Z,\ Z)$ used here, see § 6 of Part I.

Necessity. Suppose that equality holds in (25). Since the main inequality (1) holds for an arbitrary difference of convex functions on the unit sphere, the function $Z(\bar{n})$ satisfying condition (25) yields the minimal difference

$$V(H_1, \dots, H_{n-1}, Z)^2 - V(H_1, \dots, H_{n-1}, H_{n-1})V(H_1, \dots, H_{n-2}, Z, Z). \tag{26}$$

By their nature, generalized mixed volumes are differentiable functionals, i.e., admit a first variation. For a Z satisfying condition (25), the variation of the difference (26) must vanish. Taking the variation, we obtain

$$\lambda \int\limits_{\Omega} \delta Z(\bar{n}) F(H_1, \dots, H_{n-1}; d\omega) = \lambda \int\limits_{\Omega} \delta Z(\bar{n}) F(H_1, \dots, H_{n-2}, Z; d\omega). \tag{27}$$

According to the remark made in § 6 of Part I, $\delta Z(\bar{n})$ may be regarded an arbitrary continuous function. Hence from (27) it follows that

$$\lambda F(H_1, .., H_{n-1};\ \omega) = \mu F(H_1, .., H_{n-2}, Z;\ \omega).$$

Prior to taking up the applications of these inequalities, we prove two auxiliary lemmas needed in the sequel.

§ 5. Two Lemmas on the Projections of Convex Bodies

Lemma I. *If the projections of two convex bodies to every plane are translates of one another, then the bodies themselves are translates of one another.*[13]

Here we are dealing with convex bodies in a space of dimension not less than three because the lemma is a *a fortiori* false for convex figures on the plane.

Let H_1 and H_2 be two convex bodies whose projections to every plane are translates of one another (in n-dimensional space). Take two mutually perpendicular normals to the planes of projection $\bar{n}_1$ and $\bar{n}_2$. Translate the body H_2 so that its projection in the direction of $\bar{n}_1$ coincides with the corresponding projection of H_1. Then the body H_2 can be moved only along (the line spanned by) $\bar{n}_1$. Move H_2 so until its projection in the direction of $\bar{n}_2$ coincides with corresponding projection of H_1. This can be done because the projection of H_2 on the plane with normal $\bar{n}_2$ has just the same projecting cylinder as that of the corresponding projection

[13] This lemma was first proved by Süss, Zusammensetzung von Eikörpern und homotetische Eiflächen, Tôhoku Math. J., **35**: 47-50 (1932). The elegant proof given here is due to I. Liberman, a student of Leningrad State University.

of H_1 with generators parallel to $\bar{n}_1$. Since the superimposed projections are translates of one another, the motion of one of these projections in this cylinder can bring them into coincidence. We show that, after this, all the projections of H_1 and H_2 coincide with one another. Take some vector $\bar{n}$ not parallel to the plane spanned on the normals $\bar{n}_1$ and $\bar{n}_2$. Then, H_1 and H_2 each have $(n-1)$ supporting planes which are parallel either to $\bar{n}_1$ and $\bar{n}$ or to $\bar{n}_2$ and $\bar{n}$ and which are in general position.[14] These supporting planes are also the supporting planes to the projections of H_1 and H_2 in the directions of $\bar{n}_1$ or $\bar{n}_2$. Since the projections of both bodies coincide, the supporting planes coincide too. However, these supporting planes are parallel to $\bar{n}$. Thus, our two projections are translates of one another such that $(n-1)$ supporting planes in general position coincide. That means the projections themselves coincide. They differ only in their location in the plane. The vector translating one of them to the position of the other has zero projections on the normals to the $(n-1)$ supporting planes in general position. So the vector is zero.

Now , having shown that the projections coincide which are taken in the directions not parallel to the two-dimensional plane spanned by the normals $\bar{n}_1$ and $\bar{n}_2$, we can demonstrate the projections also coincide in the direction of each vector in this plane. For this purpose, it suffices to take as $\bar{n}_1$ and $\bar{n}_2$ two normals not parallel to that plane. Every supporting plane of a convex body is also a supporting plane for some one of its projections. Therefore, if all projections of two convex bodies coincide, then all their supporting planes coincide too, i.e., the bodies themselves coincide.

Lemma II. *Let $H_1, \dots, H_{n-1}$ be given convex bodies and $H'_1, \dots, H'_{n-1}$ their projections in the direction of $\bar{n}_0$. The mixed volume of the projection is representable in terms of the mixed area functions of the bodies being projected as*

$$V(H'_1, \dots, H'_{n-1}) = \frac{1}{2} \int\limits_{\Omega} |\bar{n}_0 \bar{n}| \; F(H_1, \dots, H_{n-1}; d\omega), \tag{1}$$

where $|\bar{n}_0 \bar{n}|$ is the absolute value of the cosine of the angle between $\bar{n}_0$ and $\bar{n}$.

We proved that the area of the projection of a body H in the direction of $\bar{n}_0$ is

$$V(H', \dots, H') = \frac{1}{2} \int\limits_{\Omega} |\bar{n}_0 \bar{n}| \; F(H, \dots, H;\; d\omega) \tag{2}$$

[14] $(n-1)$ planes are said to be in general position, if they are not parallel to any two-dimensional plane. Then their normals do not lie in any plane of dimension less than $(n-1)$.

[see formula (23), § 3 of Part I]. Put

$$H = \sum_{i=1}^{n-1} \lambda_i H_i.$$

Since the projection of a linear combination of convex bodies is just the same linear combination of their projections in the same direction, we have

$$H' = \sum_{i=1}^{n-1} \lambda_i H_i'.$$

Thus, from formula (2), we find

$$V\left(\sum_{i=1}^{n-1} \lambda_i H_i', \ldots, \sum_{i=1}^{n-1} \lambda_i H_i'\right) = \frac{1}{2} \int_{\Omega} |\bar{n}_0 \bar{n}|\, F\left(\sum_{i=1}^{n-1} \lambda_i H_i, \ldots; d\omega\right). \quad (3)$$

Expanding the right and left sides of this inequality and then equating the coefficients of $\lambda_1 \lambda_2 \ldots \lambda_{n-1}$, we obtain formula (1).

We already noted that the mixed volume $V_m(H)$ of a convex body H and the unit ball E is proportional to the mth quermassintegral of the body H, i.e.,

$$V_m(H) = \frac{1}{n} F_m(H, \Omega), \quad (4)$$

where Ω as usual is the whole surface of the unit ball. So there is no confusion if we call $V_m(H)$ the mth quermassintegral of the body H, as we do in the sequel. The lemma just proven implies a result needed later: *if the m-curvature functions $F_m(H; \omega)$ of two bodies are equal, then the mth curvature integrals of their projections are also equal.*

§ 6. Preliminary Remarks on Degenerate Bodies

We set before us two problems: (1) to find the necessary and sufficient conditions for the equality to hold in the generalized Brunn inequality for quermassintegrals, and (2) to determine to what extent can a convex body be defined by specifying some curvature function. Even from the proof of the Minkowski theorem on the uniqueness of a polyhedron with given normals and areas of faces, we can trace a close connection between these two problems. We solve them simultaneously, by reducing one to the other. The first problem has been solved in full to date only for the Brunn inequality governing the volumes (nth quermassintegrals). The second has been solved for the $(n-1)$th and the first curvature functions but only for curvature functions in the usual sense, i.e., only for regular bodies, as well as for

the $(n-1)$th curvature function of polyhedra.[15] Thanks to the general concept of mixed area function, particularly, the curvature function, we can solve these two problems in full generality, without any assumption about regularity whatever. First, we consider strongly degenerate bodies for which both our problems turn out to be essentially trivial. The first quermassintegrals have to be discarded from our study because we always have

$$V_1((1-\vartheta)H_0+\vartheta H_1)=(1-\vartheta)V_1(H_0)+\vartheta V_1(H_1).$$

For the mixed volume $V(H_1,\dots,H_n)$ not to vanish, it is necessary and sufficient that segments $a_1,\dots,a_n$ not parallel to the same plane could be drawn in the bodies $H_1,\dots,H_n$. This fact, which can be easily proved, was first demonstrated by Minkowski.[16]

We consider, as usual, convex bodies in the n-dimensional space. A body is said to be m-dimensional, if it lies in a plane of at least m dimensions. From the condition just stated for a mixed volume not to vanish, we find that $V_m(H)>0$ if and only if the dimension of H is not less than m. Hence, if the dimension of one of the bodies H_0 or H_1 in the polyhedron $H_\vartheta=(1-\vartheta)H_0+\vartheta H_1$ is less than m, then the generalized Brunn inequality

$$\sqrt[m]{V_m(H_\vartheta)}\geqslant(1-\vartheta)\sqrt[m]{m}V_m(H_0)+\vartheta\sqrt[m]{V_m(H_1)} \tag{1}$$

is trivial. Suppose that the dimension of H_1 is less than m, then

$$V_m(H_1)=0. \tag{2}$$

Moreover,

$$V_m(H_m)=\sum_{k=0}^{m}(1-\vartheta)^k\vartheta^{m-k}C_m^kV(\underbrace{H_0,\dots,H_0}_{k},\underbrace{H_1,\dots,H_1}_{m-k},E,\dots,E), \tag{3}$$

and, consequently, we always have

$$V_m(H_\vartheta)\geqslant(1-\vartheta)^mV_m(H_0) \tag{4}$$

and this proves our assertion. Moreover, from (3) it follows that equality stands in (4) if and only if

$$V(\underbrace{H_0,\dots,H_0}_{k},\underbrace{H_1,\dots,H_1}_{m-k},E,\dots,E)=0 \qquad (k=0,1,\dots,m-1). \tag{5}$$

Hence, from the condition for the mixed volume to vanish, we easily conclude the following.

---

[15]See Bonnesen and Fenchel, Theorie der konvexen Körper, § 13; Minkowski, Vulumen und Oberfläche, Ges. Abh. II; Christoffel, Über die Bestimmung einer Krummen Oberfläche usw., Ges. Abh., I.

[16]Theorie der konvexen Körper, 21–22, Ges. Abh. II. See also Bonnesen and Fenchel, § 7, item 29, p. 41.

If, in the generalized Brunn inequality (1), the body H_1 is of dimension less than m, then equality holds if and only if H_1 is a point and H_0 is an arbitrary body, or if H_1 is not a point and H_0 and H_1 can be placed (by translation) in one plane of dimension less than m.[17]

Suppose that H is a body of dimension less than m. Then, it is clear from the condition for the mixed volumes to vanish, that for an arbitrary convex body L and for all $k \geqslant m$

$$V(L, \underbrace{H, \dots, H}_{k}, E, \dots, E) = 0. \tag{6}$$

From the definition of mixed volume for continuous functions it follows that

$$V(Z, \underbrace{H, \dots, H}_{k}, E, \dots, E) = \frac{1}{n} \int\limits_{\Omega} Z(\bar{n}) F_k(H, d\omega) = 0. \tag{7}$$

Hence, $Z(\bar{n})$ being arbitrary, we obtain

$$F_k(H, \omega) = 0. \tag{8}$$

Consequently, *if a convex body is of dimension less than m, all of its curvature functions of order greater than or equal to m vanish identically.*

Now suppose that H is an m-dimensional body. Recall the intuitive meaning we attributed to the curvature function of a polyhedron in § 5, Part I. Let us construct the body $H + \lambda E$ parallel to H. That part of its surface which is a cylinder with generator H and directrix the surface of the $(n-m)$-dimensional ball is in magnitude proportional to λ^{n-m}. The area of this part of the cylinder, which can be mapped spherically onto the set ω, is the value of the mth curvature function. This area is proportional to the m-dimensional volume of H. The supporting planes of this part are parallel to that m-dimensional plane in which H lies and the normals of these supporting planes determine that set on the unit sphere on which the mth curvature function is not zero, i.e., if ω has no common points with this set, then $F_m(H, \omega) = 0$. This set on the unit sphere is the cross-section of that surface made by the $(n-m)$-dimensional plane perpendicular to the plane in which H lies. If ω' is that part of the set ω which belongs to this cross-section, then the area of the part of our cylindrical surface mapped spherically onto ω is proportional to ω'. Consequently, the mth curvature function of an m-dimensional convex body is completely defined when its

[17] Otherwise, in H_0 and H_1 we can choose m segments not parallel to any plane of dimension less than m, then at least one of inequalities (5) is impossible.

m-dimensional volume and the m-dimensional plane in which the body lies are given.

For $m = n - 1$ all these considerations are trivial. In this case we have an $(n-1)$-dimensional body, and its $(n-1)$th curvature function, i.e., the area function, is discrete. If $\bar{n}$ is a normal to the plane in which our body lies, its area function is reduced to two point-loadings at the points $\bar{n}$ and $-\bar{n}$ on the surface of the unit ball. This result can of course be formulated as follows.

The mth curvature functions of two m-dimensional convex bodies are equal, if and only if both the bodies have equal m-dimensional volume and lie in parallel m-dimensional planes.

§ 7. Uniqueness of a Convex Body with Given Curvature Function

Now, having essentially studied the trivial solution to the question concerning the sign of equality in the generalized Brunn inequality and the definability of a convex body by its curvature functions, we turn our attention to the cases in which the question has a fuller content. We prove two theorems.

Theorem I. *If two convex bodies of dimension not less than m have the same $(m-1)$th curvature function, then they are translates of one another.*

Theorem II. *Let H_0 and H_1 be two convex bodies of dimension not less than m and $H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1$. In the generalized Brunn inequality for the mth quermassintegrals*

$$\sqrt[m]{V_m(H_\vartheta)} \geqslant (1-\vartheta)\sqrt[m]{V_m(H_0)} + \vartheta\sqrt[m]{V_m(H_1)}, \tag{1}$$

equality holds if and only if H_0 and H_1 are homothetic.

Lemma I. *For equality to hold in the generalized Brunn inequality (1) it is necessary and sufficient that the bodies H_0 and H_1 have the $(m-1)$th curvature functions proportional.*

Necessity. In § 4 we proved that if equality holds in the generalized Brunn inequality, then it holds in all the intermediate quadratic inequalities. In other words, if equality holds in (1), then the following equalities hold:

$$V(\underbrace{H_0,\ldots,H_0}_{k},\underbrace{H_1,\ldots,H_1}_{m-k},E,\ldots E)^2 = V(\underbrace{H_0,\ldots,H_0}_{k-1},\underbrace{H_1,\ldots,H_1}_{m-k-1},E,\ldots E)$$
$$\times V(\underbrace{H_0,\ldots,H_0}_{k-1},\underbrace{H_1,\ldots,H_1}_{m-k+1},E,\ldots E), \tag{2}$$

where k runs from 1 to $(m-1)$.

In § 4 a lemma was proved which states the conditions for equality to hold in the main inequality between mixed volumes. We now apply the result obtained there to (2), a particular case of the main inequality. Thus, we obtain the series of equalities[18]:

$$\begin{aligned}&\lambda_k F(\underbrace{H_0,\dots,H_0}_{k},\underbrace{H_1,\dots,H_1}_{m-k-1},E,\dots,E;\omega)\\ &=\lambda_{k+1}F(\underbrace{H_0,\dots,H_0}_{k-1},\underbrace{H_1,\dots,H_1}_{m-k},E,\dots,E;\omega),\end{aligned}\tag{3}$$

where

$$\lambda_k=V(\underbrace{H_0,\dots,H_0}_{k},\underbrace{H_1,\dots,H_1}_{m-k},E,\dots,E).$$

These equalities generate a whole chain of equalities as k runs from one to $(m-1)$. Each of them shows that the mixed area functions are proportional. Hence the extreme members in this chain are proportional to each other. They are obtained for $k=0$ and $k=m-1$, and they are the $(m-1)$th curvature functions of H_0 and H_1, respectively.

Sufficiency. Suppose that the $(m-1)$th curvature functions of the bodies H_0 and H_1 are proportional

$$F_{m-1}(H_0;\ \omega)=\lambda F_{m-1}(H_1,\ \omega).\tag{4}$$

Dilating the body H_0 similar to itself $\sqrt[m-1]{\lambda}$ times, we obtain two bodies with equal curvature functions. We denote these bodies also by H_0 and H_1. "Differentiating" (4), and then multiplying by H_0 and finally integrating over the unit sphere, we obtain

$$\begin{aligned}&V_m(H_0)=V(\underbrace{H_0,\dots,H_0}_{m},E,\dots,E)\\ &=V_m(H_0)=V(H_0,\underbrace{H_1,\dots,H_1}_{m-1},E,\dots,E),\end{aligned}\tag{5}$$

as λ is now equal to one. In § 4, from the generalized Brunn inequality we derived the generalized Brunn-Minkowski inequality [formula (18), § 4], which for our particular case takes the form

$$V(H_0,\underbrace{H_1,\dots,H_1}_{m-1},E,\dots,E)^m\geqslant V_m(H_1)^{m-1}V_m(H_0).\tag{6}$$

[18]Note that the dimensions of H_0 and H_1 are not less than m, else all λ_k vanish and equalities (3), though true, lose their meaning.

Hence, by virtue of (5), we obtain

$$V_m(H_0) \geqslant V_m(H_1). \tag{7}$$

Since H_1 is no worse than H_0, the opposite inequality must also hold, i.e.,

$$V_m(H_0) = V_m(H_1), \tag{8}$$

and equality holds in the generalized Brunn-Minkowski inequality. In § 4 we proved that equality holds in the generalized Brunn inequality (1). That we replaced the initial body H_0 by a homothetic body makes no difference because when H_0 is multiplied by an arbitrary $\lambda > 0$ equality still holds in (6).

We actually showed that if the $(m-1)$th curvature functions of two convex bodies H_0 and H_1 are equal, then their mth curvature integrals are also equal [formula (8)], and we also showed that equality holds in the generalized Brunn inequality. Hence follows

Lemma II. *If the bodies H_0 and H_1 have equal $(m-1)$th curvature functions, then the bodies $H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1$ have equal mth quermassintegrals.*

Choosing from the family $(1-\vartheta)H_0+\vartheta H_1$ a subfamily for which $\vartheta > \vartheta_1$, we find that equality holds in the generalized Brunn inequality for the terminal members H_{ϑ_1} and H_1 of this subfamily. So their $(m-1)$th curvature functions are proportional:

$$F_{m-1}(H_\vartheta, \omega) = \lambda F_{m-1}(H_1, \omega). \tag{9}$$

Hence, as before,

$$V_m(H_{\vartheta_1}) = \lambda V(H_{\vartheta_1}, H_1, \ldots, H_1, E, \ldots, E). \tag{10}$$

And substituting H_{ϑ_1} for H_0 in the generalized Brunn-Minkowski inequality, we obtain

$$V_m(H_{\vartheta_1})^{m-1} = \lambda^m V_m(H_1)^{m-1}. \tag{11}$$

Since H_{ϑ_1} and H_1 have the same mth quermassintegrals, we find that $\lambda = 1$. Therefore, formula (9) yields

Lemma III. *If the bodies H_0 and H_1 have the same $(m-1)$th curvature functions, then each of the bodies $H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1$ has this same $(m-1)$th curvature function.*

Lemma IV. *If two convex bodies have equal first curvature functions, then these bodies are translates of one another.*

Suppose that H is a convex body with twice continuously differentiable support function $H(\bar{u})$. Draw the vectors $\bar{u}$ all issuing from the origin. Let r be a radius drawn from the origin, i.e., $r = |\bar{u}|$ and let $\bar{n}$ be a variable unit vector or, which is the same thing, a point on the unit sphere:

$$\bar{u} = \bar{n}r. \tag{12}$$

Since each support function is positively homogeneous, we have

$$H(\bar{u}) = rH(\bar{n}). \tag{13}$$

The sum of the principal radii of curvature of the body H, i.e., its first curvature function in the usual sense, is obtained by applying to $H(\bar{n})$ the Laplace operator

$$\Delta = \sum_{k=1}^{n} \frac{\partial^2}{\partial u_k^2} \tag{14}$$

and putting $r = |\bar{u}| = 1$ in the result.[19] In spherical coordinates, a surface element is

$$ds^2 = dr^2 + r^2 d\sigma^2, \tag{15}$$

where $d\sigma$ is the line element on the unit sphere. And the Laplace operator takes the form

$$\Delta = \frac{1}{r^{n-1}} \frac{\partial}{\partial r}\left(r^{n-1}\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\Delta^*, \tag{16}$$

where Δ^* is the "Laplace operator on the unit sphere" not containing r-derivatives. Hence the sum of the principal radii of curvature of the body H is [$\Delta rH(\bar{n})$ for $r = 1$]

$$\triangle^* H(\bar{n}) + (n-1)H(\bar{n}). \tag{17}$$

Consequently, if the body H has twice continuously differentiable support function, its curvature function[20] is

$$F_1(H, \omega) = \int\limits_{\Omega} [\Delta^* H(\bar{n}) + (n-1)H(\bar{n}) d\omega. \tag{18}$$

[19] See Bonnesen and Fenchel, § 8 or Blaschke, Differentialgeometrie I, § 94.

[20] According to formula (17), § 5 of Part I, the integral in (18) ought to be divided by $(n-1)$. However, for the sake of brevity, we omit this unimportant factor.

Now if $Z(\bar{n})$ is an arbitrary twice continuously differentiable function on the unit sphere, as shown in § 4 of Part I, it can be represented as the difference of some support functions on the unit sphere

$$Z(\bar{n}) = H'(\bar{n}) - H''(\bar{n}),$$

and by the definition of mixed area function, for such function $Z(\bar{n})$ we have

$$F_1(Z,\omega) = F_1(H',\omega) - F_1(H'',\omega) = \int\limits_{\Omega} [\Delta^* Z(\bar{n}) + (n-1)Z(\bar{n})]\, d\omega. \quad (19)$$

Now let H_1 and H_2 be two convex bodies with equal first curvature functions

$$F_1(H_1,\ \omega) = F_1(H_2,\ \omega), \quad (20)$$

and let $Y_1(\bar{n})$ be a spherical harmonic function (on the n-dimensional unit sphere). From (20) we find

$$\int\limits_{\Omega} Y_l(\bar{n}) F_1(H_1,\ d\omega) = \int\limits_{\Omega} Y_l(\bar{n}) F_1(H_2,\ d\omega), \quad (21)$$

or, using self-conjugacy of mixed area functions, we obtain

$$\int\limits_{\Omega} H_1(\bar{n}) F_1(Y_l,\ d\omega) = \int\limits_{\Omega} H_2(\bar{n}) F_1(Y_l,\ d\omega). \quad (22)$$

Since spherical harmonic functions are twice continuously differentiable, we can use (19) and thus obtain

$$\int\limits_{\Omega} H_1(\bar{n})[\Delta^* Y_l(\bar{n}) + (n-1)Y_l(\bar{n})]\, d\omega = \int\limits_{\Omega} H_2(\bar{n})[\Delta^* Y_l(\bar{n}) + (n-1)Y_l(\bar{n})]\, d\omega. \quad (23)$$

Spherical harmonic functions on the n-dimensional unit sphere satisfy the condition

$$\Delta^* Y_l(\bar{n}) + l(l+n-2)Y_l(\bar{n}) = 0. \quad (24)$$

Or putting $-l(l+n-2) + (n-1) = \lambda_l$, we find

$$\Delta^* Y_l(\bar{n}) + (n-1)Y_l(\bar{n}) = \lambda_l Y_l(\bar{n}). \quad (25)$$

For $\lambda_l = 0$, i.e., for $l = 1$, we obtain the first order spherical harmonic functions

$$Y_1(\bar{n}) = (\bar{a}\bar{n}), \quad (26)$$

where $\bar{a}$ is an arbitrary vector. This represents the factor which the support function $H(\bar{n})$ on the unit sphere is to be multiplied with when the body H is translated in the direction of the vector[21] $\bar{a}$.

By virtue of (25), we can rewrite (23) as follows:

$$\lambda_l \int\limits_{\Omega} H_1(\bar{n}) Y_l(\bar{n})\, d\omega = \lambda_l \int\limits_{\Omega} H_2(\bar{n}) Y_l(\bar{n})\, d\omega. \tag{27}$$

Consequently, for $\lambda_l \neq 0$ all the corresponding Fourier coefficients of the functions $H_1(\bar{n})$ and $H_2(\bar{n})$ are pairwise equal. Since the system of spherical harmonic functions is closed, this is possible only if the functions $H_1(\bar{n})$ and $H_2(\bar{n})$ differ in the eigenfunction that corresponds to the eigenvalue $\lambda_l = 0$. And this eigenfunction represents the translation from one body to the other.

Thus our two theorems are demonstrated for the first curvature function and for the second quermassintegral. It remains to prove them for higher order curvature functions and integrals. Note that only first order curvature functions exist for plane convex figures; therefore, both our theorems are proven to the full extent for such figures. Therefore, assuming that both the theorems hold for a body in the $(n-1)$-dimensional space, we prove that they hold for a body in n-dimensional space.

Suppose that the bodies H_0 and H_1 have the mth curvature functions equal, with $m > 1$. Then by Lemma III, the same is true of all the bodies of the family

$$H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1.$$

In § 5 we proved that in this case the mth quermassintegrals of the projections of these bodies in the same direction are equal. The projections of the bodies of the linear family $(1-\vartheta)H_0 + \vartheta H_1$ likewise form linear families for which equality holds in the generalized Brunn inequalities for the mth quermassintegrals. We assume that the bodies H_0 and H_1 are of dimension

[21] All properties of a spherical harmonic function on the n-dimensional sphere follow from its definition: if $P_l(\bar{u})$ is a harmonic polynomial of degree l, then $P_l(\bar{u}) = r^l Y_l(\bar{n})$. The completeness of the system of spherical harmonic functions is easily proved if we note that every homogeneous polynomial of degree l can be represented as

$$U_1(\bar{u}) = P_l(\bar{u}) - r^2 P_{l-2}(\bar{u}) - r^4 P_{l-4}(\bar{u}) + \dots$$

which is proved by calculating the number of arbitrary coefficients on either side of the above equality.

The idea of applying spherical harmonic functions to proving the uniqueness of a convex body given the sum of principal radii of curvature belongs to Hurwitz (Sur quelques applications géométriques des séries de Fourier, Ann. de l'École Normale, (3), 19 (1902).

greater than m (since the question here is about their mth curvature functions). Therefore their projections have dimension not less than m. So, by the assumption of induction, equality holds in the Brunn inequality if and only if these bodies H_0 and H_1 are homothetic (recall that $m > 1$) and, moreover, their mth quermassintegrals are equal. So, the projections of H_0 and H_1 in the same direction are translates of one another. Hence, by virtue of Lemma I of § 1, the bodies H_0 and H_1 themselves are translates of one another.

Thus we completed the proof of Theorem I. Since, by Lemma I, both the theorems are equivalent, we also proved Theorem II.

§ 8. Uniqueness of a Centrally Symmetric Convex Body with Given Cross-Sectional Measures

In this section we give an application of the general uniqueness theorem of a convex body with given curvature function.

The projection of a convex body H to an arbitrary plane P of dimension less than $(n-1)$ is also the projection of its projection to an arbitrary $(n-1)$-dimensional plane parallel to P. So, if all m-dimensional cross-sectional measures of a convex body are known, the mean m-dimensional cross-sectional measures, the mth quermassintegrals, of all of its plane projections are also known. By the mean $(n-1)$-dimensional cross-sectional measure of a plane projection we understand the area of the projection.

In what follows, we call a plane projection, i.e., an $(n-1)$-dimensional projection of a convex body simply a *projection*. The projection of a body H in the direction of $\bar{n}$ is denoted by $H_{\bar{n}}$, and the mth quermassintegral and mth curvature function of a body H by $V_m(H)$ and $F_m(H, \omega)$, respectively. Here $V_m(H_{\bar{n}})$ stands for the mth curvature integral of the projection $H_{\bar{n}}$ and, for $m = n-1$, its area. Since the mean cross-sectional measures are proportional to the quermassintegrals, it makes no difference which of these quantities we are dealing with in our study.

In § 5 it was shown that

$$V_m(H_{\bar{n}}) = \frac{1}{2} \int\limits_{\Omega'} |\bar{n}\bar{n}'| F_m(H, d\omega'). \tag{1}$$

The prime pertains to the integration variable and the domains over which integration is carried out. Our immediate problem is to clarify to what extent a knowledge of the mth quermassintegrals of the projections of a convex body determines its mth curvature function.[22]

[22]See Blaschke, Kreis und Kugel, Anhang (1916).

Lemma I. *The eigenvalues of the integral equation*

$$Y(\bar{n}) = \lambda \int\limits_{\Omega'} |\bar{n}\bar{n}'| Y(\bar{n}') d\omega \tag{2}$$

are spherical harmonic functions of even order.

Let $Y_{2l}(\bar{n}')$ be a spherical function of even order. Take a point $\bar{n}$ on the unit sphere and let it be a pole. Then $Y_{2l}(\bar{n}')$ can be represented as a linear combination of spherical harmonic functions of the same order for which $\bar{n}$ is the pole of the coordinate system on the unit sphere

$$Y_{2l}(\bar{n}') = \sum_m a_m Y_{2l,m}^{(\bar{n})}(\bar{n}'). \tag{3}$$

If θ' is the polar distance to the point $\bar{n}$ and ν' is a point on an $(n-1)$-dimensional sphere obtained as an equatorial section of the n-dimensional sphere, then

$$Y_{2l,m}^{(\bar{n})}(\bar{n}') = P_{2l,m}(\theta') Y_m(\nu'). \tag{4}$$

(This result is readily obtained, if the Laplace operator Δ^* on the sphere is expressed as usual through the variables θ' and ν'.) Here $Y_m(\bar{\nu}')$ is a spherical function on the $(n-1)$-dimensional sphere. Thus, for $m \neq 0$, we have

$$Y_{2l,m}^{(\bar{n})} = P_{2l,m}(0) Y_m(\bar{\nu}') = 0, \tag{5}$$

otherwise it would be many-valued at the pole and

$$Y_{2l}(\bar{n}) = a_0 Y_{2l,0}^{(\bar{n})}(\bar{n}). \tag{6}$$

At the same time, since $|\bar{n}\bar{n}'| = |\cos\theta'|$, for $m \neq 0$, we have

$$\int\limits_{\Omega'} |\bar{n}\bar{n}'| Y_{2l,m}^{(\bar{n})}(\bar{n}')\, d\omega' = 0, \tag{7}$$

therefore

$$\int\limits_{\Omega'} |\bar{n}\bar{n}'| Y_{2l}(\bar{n}')\, d\omega' = a_0 \int\limits_{\Omega'} |\bar{n}\bar{n}'| Y_{2l,0}^{(\bar{n})}(\bar{n}')\, d\omega'. \tag{8}$$

The integral on the right side does not depend on the choice of $\bar{n}$. Put

$$\frac{1}{Y_{2l,0}^{(\bar{n})}(\bar{n})} \int\limits_{\Omega'} |\bar{n}\bar{n}'| Y_{2l,0}^{(\bar{n})}(\bar{n}')\, d\omega' = \frac{1}{\lambda_l}. \tag{9}$$

Now, comparing (6) and (8), we obtain

$$Y_{2l}(\bar{n}) = \lambda_l \int_{\Omega'} |\bar{n}\bar{n}'| Y_{2l}(\bar{n}')\, d\omega'. \tag{10}$$

This completes the proof of our assertion.

Equation (2) has no other eigenfunctions as the parity of its kernel ($|\bar{n}\bar{n}'| = |-\bar{n}\bar{n}'|$) implies that all its eigenfunctions are even; and the spherical functions of even order form a complete system of even functions.

Now multiplying (1) by $Y_{2l}(\bar{n})$ and then integrating, we deduce

$$\int_{\Omega} V_m(H_{(\bar{n})}) Y_{2l}(\bar{n})\, d\omega = \frac{1}{2} \int_{\Omega} Y_{2l}(\bar{n})\, d\omega \int_{\Omega} |\bar{n}\bar{u}'| F_m(H,\ d\omega'). \tag{11}$$

Changing the order of integration on the right side and noting that $Y_{2l}(\bar{n})$ satisfies equation (2), we obtain

$$\int_{\Omega} V_m(H_{\bar{n}}) Y_{2l}\, d\omega = \frac{2}{\lambda} \int_{\Omega} Y_{2l}(\bar{n}) F_m(H,\ d\omega). \tag{12}$$

Consequently, when all $V_m(H^{\bar{n}})$ are given, the integrals on the right side are known.

Now let $Z(\bar{n})$ be a function equal to one on some set ω and on its symmetric set $-\omega$ on the surface of the unit ball, and zero on the remaining points of the surface. This function is even, so we can construct a sequence which converges to it and consists of linear combinations of spherical functions $Z_k(\bar{n})$ of even order:

$$Z(\bar{n}) = \lim_{k=\infty} Z_k(\bar{n}). \tag{13}$$

The integrals

$$\int_{\Omega} Z_k(\bar{n}) F_m(H,\ d\omega)$$

are known, so is their limit (ω and $-\omega$ are supposed to have no common points)[23]:

$$\int_{\Omega} Z(\bar{n}) F_m(H, d\omega) = F_m(H, \omega) - F_m(H, -\omega). \tag{14}$$

[23]Let ω be closed. The function $Z(\bar{n})$ can be uniformly approximated by bounded continuous functions so that they converge uniformly to $Z(\bar{n})$ on ω and on $-\omega$ and vanish outside some set $(\omega + \omega^*)$ and on $-(\omega + \omega^*)$. These continuous functions, in turn, can be uniformly approximated by linear combinations of the spherical functions $Z_k(\bar{n})$ of even order. Then, it is a simple matter to verify by trivial evaluation that

$$\lim_{k=\infty} \int Z_k'(\bar{n}) F_m(H,\ d\omega) = \int Z(\bar{n}) F_m(H,\ d\omega).$$

When the mth mean cross-sectional measures of a convex body are given, the even part

$$F_m(H,\omega) + F_m(H,-\omega)$$

of its curvature function is defined. In other words, we have

Lemma II. *If all the m-dimensional cross-sectional measures of the projections of two convex bodies H_1 and H_2 are equal, the even parts of their mth curvature functions are also equal, i.e.,*

$$F_m(H_1,\omega) + F_m(H_1,-\omega) = F_m(H_2,\omega) + F_m(H_2,-\omega). \tag{15}$$

If the bodies H_1 and H_2 are centrally symmetric, then their curvature functions are even and

$$F_m(H_1,\omega) = F_m(H_2,\omega). \tag{16}$$

Hence, from the uniqueness theorem proved earlier, we obtain the following

Theorem. *If the mean cross-sectional measures of given dimension of all projections of a centrally symmetric convex body are given, then the body, up to translation, is unique.*

For example, if for two centrally symmetric convex bodies, the areas of projections in the direction of every vector are equal, then the bodies themselves are translates of one another.

When the cross-sectional measures are given, we know the cross-sectional measures of the bodies for all projections. Whether the converse holds for all convex bodies still remains unclear, except for one case [excluding the trivial case of $(n-1)$-dimensional cross-sectional measures because the mean in this case simply coincides with the cross-sectional measure itself], namely, the case of the one-dimensional cross-sectional measure. From the equality of the mean one-dimensional cross-sectional measures of all projections of two convex bodies, it follows that

$$F_1(H_1,\omega) + F_1(H_1,-\omega) = F_1(H_2,\omega) + F_1(H_2,-\omega), \tag{17}$$

or putting $B(\bar{n}) = H(\bar{n}) + H(-\bar{n})$, we obtain

$$F_1(B_1,\omega) = F_1(B_2,\omega). \tag{18}$$

Hence it follows that $B_1(\bar{n}) = B_2(\bar{n})$.

§ 9. Extremal Properties of a Ball

At the end of § 4 we derived the inequality

$$V_k(H)^n \geqslant V_m(H)^k V(E)^{m-k} \tag{1}$$

for the quermassintegrals, where $1 \leqslant k < m \leqslant n$. The quermassintegral $V_m(H)$ vanishes, if and only if the dimension of H is not less than m. Only under this condition inequality (1) is not trivial. We also showed there that equality holds in (1) if and only if equality holds in the generalized Brunn inequality:

$$\sqrt[l]{V_l(H_\vartheta)} \geqslant (1-\vartheta)\sqrt[l]{V_l(H)} + \vartheta\,\sqrt[l]{V_l(E)}, \tag{2}$$

where $l \leqslant m$ and $H_\vartheta = (1-\vartheta)H + \vartheta E$.

By Theorem II proved in § 6, equality can hold in the generalized Brunn inequalities if and only if the body H is homothetic to the unit ball E, provided the dimension of H is not less than m since l is less than or equal to m.

Consequently, equality holds in (1) if and only if H is a ball, provided its dimension is not less than m. However, in this case the equality is reduced to the assertion that zero equals zero.

Thus, we obtain a generalization of the well-known property: for given volume a ball has least area.

Theorem. *Amongst convex bodies with given nonzero mth quermassintegrals, a ball and only a ball has the least quermassintegral of each order less than m.*

We have already mentioned that the mth quermassintegral of a convex body is equal, up to a constant factor, to its mean m-dimensional cross-sectional measure. So in the extremal property of a ball proved above, we may as well speak of the mean cross-sectional measure. In view of this remark, from the theorem proved above, we immediately deduce the following corollary.

Amongst all convex bodies with given minimal m-dimensional cross-sectional measure, a ball and only a ball has the least mean cross-sectional measure of every dimension less than m.

This corollary, just like the theorem itself, can of course be reformulated as follows: *Amongst all convex bodies with given maximal m-dimensional cross-sectional measure, a ball and only a ball has the greatest mean cross-sectional measure of every dimension greater than m.*

For example, among all convex bodies with given diameter, a ball and only a ball has greatest area.

In all these statements, by the nth quermassintegral and n-dimensional cross-sectional measure of a convex body, we mean its volume.

The theorem just proved is equivalent to the following assertion: the ratio

$$\Phi_{km}(H) = \frac{V_k(H)^m}{V_m(H)^k} \qquad (k < m) \tag{3}$$

is a minimum only when H is a ball.

This result can be refined by the Minkowski method as follows:

The ratio of the quermassintegrals $\Phi_{km}(H)$ decreases in passing from a convex body H to the outer parallel body $H + \lambda E$ and is constant if and only if H is a ball.

It suffices to demonstrate this for the ratio

$$\Phi_{m-1,m}(H) = \frac{V_{m-1}(H)^m}{V_m(H)^{m-1}}, \tag{4}$$

because the ratio $\Phi_{km}(H)$ for an arbitrary $k < m$ is obtained on multiplying together the ratios $\Phi_{l-1,l(H)}$ raised to some positive power, l being varied from $(k+1)$ to m.[24] Consider the family of bodies $H_\vartheta = H + \vartheta E$ $(0 \leqslant \vartheta \leqslant 1)$ which join the body H with one of its outer parallel bodies. From the generalized Brunn inequality it follows that $\sqrt[m]{V_m(H_\vartheta)}$ is a concave function of ϑ and is linear if and only if H and $H + E$ are homothetic, i.e., only if H is a ball. Therefore the derivative $\sqrt[m]{V_m(H_\vartheta)}$ is a monotone decreasing function of ϑ and is constant if and only if H is a ball:

$$\frac{d}{d\vartheta}\sqrt[m]{V_m(H_\vartheta)} = \frac{1}{mV_m(H_\vartheta)^{\frac{m-1}{m}}}\frac{d}{d\vartheta}V_m(H_\vartheta). \tag{5}$$

Since $H_\vartheta = H + \vartheta E$, we have

$$V_m(H_\vartheta) = \sum_{k=0}^{m} \vartheta^k C_m^k V_{m-k}(H). \tag{6}$$

Differentiating with respect to ϑ, and noting that $kC_m^k = mC_{m-1}^{k-1}$, we obtain

$$\frac{d}{d\vartheta}V_m(H_\vartheta) = m\sum_{k=0}^{m-1} \vartheta^k C_{m-1}^k V_{m-k-1}(H). \tag{7}$$

---

[24]Thus, we have

$$\Phi_{l,l+1}^{l+2} \cdot \Phi_{l+1,l+2}^{l} = \Phi_{l,l+2}^{l+1}.$$

Using this relation, which can be easily verified, and varying l from k to $m-2$, we readily find the truth of our assertion. The corresponding analytical expression is rather unduly complicated, so we omit it.

The sum on the right side is nothing else than $V_{m-1}(H_\vartheta)$. So

$$\frac{d}{d\vartheta}\sqrt[m]{V_m(H_\vartheta)} = \frac{V_{m-1}(H_\vartheta)}{V_m(H_\vartheta)^{\frac{m-1}{m}}} = \sqrt[m]{\Phi_{m-1,m}(H_\vartheta)}. \tag{8}$$

Since the derivative $\sqrt[m]{V_m(H_\vartheta)}$ does not increase and remains constant when and only when H is a ball, the same can be asserted about the ratio $\Phi_{m-1,m}(H+\vartheta E)$.

Institute of Mathematics and Mechanics,
Leningrad State University

Received April 17, 1937

CHAPTER V

TO THE THEORY OF MIXED VOLUMES OF CONVEX BODIES PART III[1]

Matematicheskiĭ Sbornik, vol. 3(45), No. 1, 27–46 (1938).

EXTENSION OF TWO MINKOWSKI THEOREMS ON CONVEX POLYHEDRA TO ALL CONVEX BODIES

In this paper we extend two results that were reported by Minkowski in his work *Allgemeine Lehrsätze über die konvexen Polyeder*,[2] namely, the theorem on the minimal property of polyhedra circumscribed about the ball and the theorem on the existence and uniqueness of a convex polyhedron with given directions and areas of faces. In the proofs we literally follow Minkowski, and this, I hope, underlines the elegance and power of his method.

§ 1. On a Convex Body with a Given Domain of Definition of the Support Function

As in Parts I and II, we consider a convex body in n-dimensional space, where n always stands for the dimension of the space. Once for all we choose the origin of the space and use the symbol x to denote a point as well as the vector joining the origin with the point. We assume that the unit ball E is always given in the space. To each unit vector $\bar{n}$ there corresponds on E a unique point $\bar{n}$, the end of the vector equal to $\bar{n}$ and drawn from the center of E. By the *support function* $H(\bar{n})$ of a convex body H we understand the function of unit vectors $\bar{n}$ that specifies the distance from the origin to the

[1]For Parts I and II of this work, see Chapters III and IV in this monograph.
[2]Uspekhi Matematicheskikh Nauk, **II** (1936).

supporting plane of H with outward normal $\bar{n}$. This distance is reckoned positive along $\bar{n}$ and negative in the opposite direction. We say that there is a normal $\bar{n}$ to the body H at its point $\bar{x}$, if there is a supporting plane to H with outward normal $\bar{n}$ which passes through $\bar{x}$.

On the surface of the unit ball E, take a closed set Ω' not lying in one hemisphere. (The boundary of a hemisphere is supposed to belong to the hemisphere.) Let us define a continuous positive function $H^*(\bar{n})$ on the set Ω'. Referring to the origin, construct a family of planes such that $H^*(\bar{n})$ represents the distance from the origin to the plane with normal $\bar{n}$ drawn from the origin. In the set Ω' we can choose $(n+1)$ points $\bar{n}_1, \dots, \bar{n}_{n+1}$ so that the vectors $\bar{n}_1, \dots, \bar{n}_{n+1}$ are not directed into one half-space. The planes of our family with such normals bound a certain simplex. Therefore the half-spaces defined by the inequalities

$$\bar{n}\bar{x} \leqslant H^*(\bar{n}) \quad (\bar{n} \in \Omega) \tag{1}$$

give, as their intersection, some convex body H. Since $H^*(\bar{n}) > 0$, the origin lies in the interior of the body H. We call it the *body with* Ω' *as the domain of definition of the support function*. If Ω' consists of a finite number of points $\bar{n}_1, \dots, \bar{n}_N$, then the convex body with this domain of definition for its support function is a polyhedron with $\bar{n}_1, \dots, \bar{n}_N$ as normals to its boundary planes.

Lemma I. *Through each point on the surface of the body H there passes at least one of the planes*

$$\bar{n}\bar{x} = H^*(\bar{n}) \quad (\bar{n} \in \Omega). \tag{2}$$

The body H is defined as the locus of a point $\bar{x}$ that satisfies all inequalities (1). If, for a point $\bar{x}_0$ and for every normal $\bar{n}$ belonging to Ω',

$$\bar{n}\bar{x} < H^*(\bar{n}) \quad (\bar{n} \in \Omega),$$

then, by virtue of the continuity of $H^*(\bar{n})$, there exists an $a > 0$ such that the point $\bar{x}_0$ is at a distance not less than a from any of the planes (2). Then the ball of radius a and with center $\bar{x}_0$ lies in the interior of every half-space (1), so $\bar{x}_0$ is an interior point of either of the half-spaces. That means $\bar{x}_0$ is an interior point of the body H.

Let $H(\bar{n})$ denote the support function of the body H. Since for every $\bar{n}$, there exists on H a point $\bar{x}$ for which $\bar{n}\bar{x} = H(\bar{n})$, and if, in addition, for an $\bar{n}$ belonging to Ω', we have $\bar{n}\bar{x} \leqslant H^*(\bar{n})$ for every point on H, then

$$H(\bar{n}) \leqslant H^*(\bar{n}) \quad (\bar{n} \in \Omega'). \tag{3}$$

Lemma II. *If $\bar{n}_0$ does not fall into Ω', then at the point x_0 on the surface of H with normal $\bar{x}_0$ there are other normals, i.e., such a point is singular.*

By "$\bar{n}_0$ falls into Ω'" we understand that the end of the vector $\bar{n}_0$, when drawn from the center of E, belongs to the set Ω'.

The supporting plane with normal $\bar{n}_0$ passes through the point $\bar{x}_0$, and, by Lemma I, one of the planes (2) passes through that point. The normal of this plane falls into Ω' and, consequently, differs from $\bar{n}_0$.

Lemma III. *If a normal $\bar{n}_0$ falling into Ω' satisfies the condition $H(\bar{n}_0) < H^*(\bar{n}_0)$, then the point $\bar{x}_0$ on the surface of H with normal $\bar{n}_0$ is singular.*

A supporting plane with normal $\bar{n}_0$ passes through the point $\bar{x}_0$. By Lemma I, one of the planes (2) also passes through that point. The normal to such a plane is different from $\bar{n}_0$ as $H(\bar{n}_0) < H^*(\bar{n}_0)$. So, the supporting plane with normal $\bar{n}_0$ passes closer to the origin than its parallel plane in the family (2).

The set of all singular points on the surface of a convex body has measure zero. The measure of the set of points with normals falling into a set ω on the unit sphere is equal to the value of the area function $F(H,\, \omega)$ at the set ω. Hence, by Lemma II, the value of the area function of the body H at the complement of Ω' is zero,[3] i.e.,

$$F(H,\, \Omega - \Omega') = 0. \tag{4}$$

Similarly, by Lemma III, for a set ω_0' where

$$H(\bar{n}) < H^*(\bar{n}) \quad (\omega_0' \subset \Omega'),$$

we have

$$F(H,\ \omega_0) = 0. \tag{5}$$

However, at all other points $\bar{n}$ belonging to Ω', we have

$$H(\bar{n}) = H^*(\bar{n}).$$

Therefore

$$\int\limits_{\Omega'} H(\bar{n})\, F(H,\, d\omega) = \int\limits_{\Omega'} H^*(\bar{n})\, F(H,\, d\omega), \tag{6}$$

and, since in addition $F(H, \Omega - \Omega') = 0$, we have

$$\int\limits_{\Omega} H(\bar{n})\, F(H,\, d\omega) = \int\limits_{\Omega'} H^*(\bar{n})\, F(H,\, d\omega). \tag{7}$$

Hence we immediately obtain

[3]Since $F(H;\, \omega)$ is nonnegative, $F(H; \omega) = 0$ only if $\omega \subset \Omega - \Omega'$.

Lemma IV. *The volume of the body H defined by a function $H^*(\bar{n})$ given on Ω' is*

$$V(H^*) = \frac{1}{n} \int\limits_{\Omega'} H^*(\bar{n})\, F(H, d\omega). \tag{8}$$

Lemma V. *If a sequence of continuous functions $H_1^*(\bar{n}), H_2^*(\bar{n})\dots$, defined on Ω' converges uniformly to a positive function $H^*(\bar{n})$, then the convex bodies $H_1, H_2, \dots$ defined by these functions converge to the body H defined by the limit function.*

For the convex bodies $H_1, H_2, \dots$ to converge to a body H (having interior points), it is necessary and sufficient that, first, every interior point of H be interior to all H_m for sufficiently large m and, second, any point not belonging to H should not also belong to any of the bodies H_m for sufficiently large m.[4]

Suppose that a point $\bar{x}_0$ lies in the interior of H and let a be the minimum distance from x_0 to the supporting planes of H. Then for all $\bar{n}$ belonging to Ω', we have

$$\bar{n}\bar{x}_0 \leqslant H(\bar{n}) - a \leqslant H^*(\bar{n}) - a \quad (\bar{n} \in \Omega'). \tag{9}$$

Choose an M large enough so that, for $m > M$,

$$|H^*(\bar{n}) - H_m^*(\bar{n})| < \frac{a}{2} \quad (\bar{n} \in \Omega'). \tag{10}$$

Then for every $m > M$,

$$\bar{n}\bar{x}_0 < H_m^*(\bar{n}) - \frac{a}{2}. \tag{11}$$

That means, $\bar{x}_0$ lies in the interior of the body H_m for all m greater than M, as is clear from the proof of Lemma I.

Suppose that $\bar{x}_0$ lies outside the body H. Then there exists an $\bar{n}_0$ belonging to Ω' such that

$$\bar{n}_0\bar{x}_0 = H^*(\bar{n}_0) + a, \tag{12}$$

where $a > 0$ is the distance of $\bar{x}_0$ from the plane $\bar{n}_0\bar{x} = H^*(\bar{n}_0)$. For sufficiently large m,

$$|H^*(\bar{n}_0) - H_m^*(\bar{n}_0)| < \frac{a}{2}. \tag{13}$$

[4]See Blaschke, Kreis und Kugel, § 18.

So

$$\bar{n}_0\bar{x}_0 > H^*_m(\bar{n}_0) + \frac{a}{2}. \tag{14}$$

This means the point $\bar{x}_0$ lies outside the body H_m.

Let us define, as is usually done, the variation of the volume of the body given by the function $H^*(\bar{n})$ as

$$\delta V(H^*) = \lim_{t=0} \frac{V(H^* + t\delta H^*) - V(H^*)}{t}. \tag{15}$$

Lemma VI. *The volume of the body defined by a continuous positive function $H^*(\bar{n})$ given on a set Ω' has the first derivative*

$$\delta V(H^*) = \int\limits_{\Omega'} \delta H^*(\bar{n}) F(H, d\,\omega), \tag{16}$$

where $F(H, \omega)$ is the area function of the body defined by the function $H^*(\bar{n})$.

Let $H^*(\bar{n})$ be a continuous function defined on Ω. For sufficiently small t, the function $H^*(\bar{n}) + t\delta H^*(\bar{n})$ is positive and therefore defines the convex body H_t. By Lemma V, as t tends to zero, the bodies H_t converge to the body H_0 defined by the function $H^*(\bar{n})$. For the sake of definiteness, assume that t tends to zero through positive values. If, on the contrary, t does not take positive values, then the inequalities we derive below are to be reversed and the final result would still remain valid. And if the limit in the formula for the variation is the same regardless of whether t tends to zero only through positive values or only through negative values, the limit is the same when t tends to zero by any law. We can therefore readily assume t positive.

Let $H_t(\bar{n})$ be the support function of the body H_t:

$$H_t(\bar{n}) \leqslant H^*(\bar{n}) + t\delta H^*(\bar{n}) \quad (\bar{n} \in \Omega'). \tag{17}$$

Therefore, since $F(H_0,\, \Omega - \Omega') = 0$, we obtain

$$\int\limits_{\Omega'} [H^*(\bar{n}) + t\delta H^*(\bar{n})]\, F(H_0,\, d\omega) \geqslant \int\limits_{\Omega} H_t(\bar{n})\, F(H_0,\, d\omega). \tag{18}$$

By Lemma IV,

$$\int\limits_{\Omega'} H^*(\bar{n})\, F(H_0,\, d\omega) = \int\limits_{\Omega} H_0(\bar{n})\, F(H_0,\, d\omega). \tag{19}$$

Subtracting it from (18), we obtain

$$t\int\limits_{\Omega'} \delta H^*(\bar{n})\, F(H_0,\, d\omega) \geqslant \int\limits_{\Omega} H_t(\bar{n})\, F(H_0,\, d\omega) - \int\limits_{\Omega} H_0(\bar{n})\, F(H_0, d\,\omega). \quad (20)$$

The integrals on the right side are equal to n times the product of the mixed volume $V(H_t, H_0, \dots, H_0)$ and the volume of the body H_0, respectively. For the sake of brevity, let

$$V(\underbrace{H_t, \dots, H_t}_{m}, H_0, \dots, H_0) = V_m. \quad (21)$$

Dividing (20) by t and then making t tend to zero, we obtain

$$\int\limits_{\Omega'} \delta H^*(\bar{n}) F(H_0, d\omega) \geqslant n \varlimsup_{t=0} \frac{V_1 - V_0}{t}. \quad (22)$$

By Lemma IV,

$$\int\limits_{\Omega} [H^*(\bar{n}) + t\delta H^*(\bar{n})]\, F(H_t, d\omega) = \int\limits_{\Omega} H_t(\bar{n})\, F(H_t, d\omega). \quad (23)$$

And for an n falling into Ω', we have $H^*(\bar{n}) \geqslant H_0(\bar{n})$ and, since $F(H_t, \Omega - \Omega') = 0$, we have

$$\int\limits_{\Omega'} H^*(\bar{n})\, F(H_t, d\omega) \geqslant \int\limits_{\Omega} H_0(\bar{n})\, F(H_t, d\omega). \quad (24)$$

Subtracting it from (23), we obtain

$$t\int\limits_{\Omega'} \delta H^*(\bar{n}) F(H_t, d\omega) \leqslant \int\limits_{\Omega} H_t(\bar{n})\, F(H_t, d\omega) - \int\limits_{\Omega} H_0(\bar{n})\, F(H_t, d\omega). \quad (25)$$

The continuous function $\delta H^*(\bar{n})$ can be extended to the whole surface of the unit ball.[5] In so doing, its integral over $F(H_t, \omega)$ does not change because $F(H_t, \Omega - \Omega') = 0$. Clearly, the integral on the left side of (25) is the generalized mixed volume $nV(\delta H, H_t, \dots, H_t)$. By a lemma proved in § 6, Part I, this integral tends, for $t \to 0$, to

$$nV(\delta H^*, H_0, \dots, H_0) = \int\limits_{\Omega'} \delta H^*(\bar{n}) F(H_0, d\omega). \quad (26)$$

[5] See Hausdorff, Mengenlehre, § 41: The Brauer-Urysohn Theorem, Berlin-Leipzig, 1927.

Therefore, expressing the integrals on the right side of (25) through mixed volumes, and making t tend to zero, we obtain

$$\int\limits_{\Omega'} \delta H^*(\bar{n})\, F(H_0, d\omega) \leqslant n \varliminf_{t=0} \frac{V_n - V_{n-1}}{t}. \tag{27}$$

Subtracting V_n from the Brunn-Minkowski inequality

$$V_1^n \geqslant V_0^{n-1} V_n, \tag{28}$$

we obtain

$$V_1^n - V_0^n \geqslant V_0^{n-1}(V_n - V_0). \tag{29}$$

Now expanding the difference of the nth powers in the usual way and then dividing by V_0^{n-1}, we obtain

$$(V_1 - V_0) \sum_{k=0}^{n-1} \left(\frac{V_1}{V_0}\right)^k \geqslant V_n - V_0. \tag{30}$$

Since, H_t tends to H as t tends to zero, we infer $\lim_{t\to 0} V_1 = V_0$. Therefore, dividing inequality (30) by t and then passing to limit we obtain

$$n \varlimsup_{t=0} \frac{V_1 - V_0}{t} \geqslant \varlimsup_{t=0} \frac{V_n - V_0}{t}. \tag{31}$$

By analogous reasoning, from the inequality

$$V_{n-1}^n \geqslant V_n^{n-1} V_0, \tag{32}$$

we obtain another inequality

$$n \varliminf_{t=0} \frac{V_n - V_{n-1}}{t} \leqslant \varliminf_{t=0} \frac{V_n - V_0}{t}. \tag{33}$$

This inequality, together with (31), yields

$$n \varlimsup_{t=0} \frac{V_1 - V_0}{t} \geqslant \varlimsup_{t=0} \frac{V_n - V_0}{t} \geqslant \varliminf_{t=0} \frac{V_n - V_0}{t} \geqslant n \varliminf_{t=0} \frac{V_n - V_{n-1}}{t}. \tag{34}$$

Now comparing (22) and (27), we obtain

$$n \varlimsup_{t=0} \frac{V_1 - V_0}{t} \leqslant \int\limits_{\Omega'} \delta H^*(\bar{n}) F(H_0, d\omega) \leqslant n \varliminf_{t=0} \frac{V_n - V_{n-1}}{t}. \tag{35}$$

From comparison of these two chains of inequalities, it is seen, that

$$\lim_{t=0} \frac{V_n - V_0}{t}$$

exists and is equal to the variation of the volume; second, it is also seen that this limit is equal to the integral in the second chain of inequalities. Therefore

$$\delta V(H^*) = \int\limits_{\Omega'} \delta H^*(\bar{n}) F(H_0, d\omega). \tag{36}$$

This completes the proof of our theorem.

In § 3 we need the general results obtained here as applied to the case in which the set extends over the whole surface of the unit ball. Each positive continuous function on the unit sphere defines a certain convex body. The volume of this body can be called the *volume of this function.* We have shown that this volume depends continuously on this function and even has a first variation.

§ 2. Minimal Property of Bodies Circumscribed About a Sphere

On the unit sphere let us take a closed set Ω' not confined within only one hemisphere.

Lemma. *If $H_0^*(\bar{n})$ and $H_1^*(\bar{n})$ are two positive continuous functions given on a set Ω', then the functions $H_\vartheta^*(\bar{n}) = (1-\vartheta)H_0^* + \vartheta H^*(\bar{n})$ define the convex family of convex bodies H_ϑ joining the bodies H_0 and H_1 defined by the functions $H_0^*(\bar{n})$ and $H_1^*(\bar{n})$.*[6]

The convex body H_ϑ is defined as the intersection of the half-spaces

$$\bar{n}\bar{x} \leqslant H_\vartheta^*(\bar{n}) \quad (\bar{n} \in \Omega'), \tag{1}$$

i.e., it is the locus of a point $\bar{x}$ that satisfies all inequalities (1).

Take two values ϑ_1 and ϑ_2 of the parameter ϑ. Join the bodies H_{ϑ_1} and H_{ϑ_2} by the linear family $(1-\vartheta')H_{\vartheta_1} + \vartheta' H_{\vartheta_2}$. Each body of this family is the locus of a point $\bar{x}$ that satisfies the inequalities:

$$\bar{n}\bar{x} \leqslant (1-\vartheta')H_{\vartheta'}(\bar{n}) + \vartheta' H_{\vartheta_1}(\bar{n}), \tag{2}$$

here $\bar{n}$ is an arbitrary point on the unit sphere. On the other hand, the body $H_{(1-\vartheta')\vartheta_1+\vartheta'\vartheta_2}$ is the locus of a point $\bar{x}$ that obeys the inequalities

$$\bar{n}\bar{x} \leqslant (1-\vartheta')H_{\vartheta_1}^*(\bar{n}) + \vartheta' H_{\vartheta_2}^*(\bar{n}) \quad (\bar{n} \in \Omega'), \tag{3}$$

[6] A family of bodies H_ϑ is defined to be convex if for all ϑ_1 and ϑ_2, the body $H_{(1-\vartheta')\vartheta_1+\vartheta'\vartheta_2}$ includes the body $(1-\vartheta')H_{\vartheta_1}+\vartheta' H_{\vartheta_2}$, where $0 \leqslant \vartheta \leqslant 1$ and $0 \leqslant \vartheta' \leqslant 1$. In the lemma it is implied that $0 \leqslant \vartheta \leqslant 1$.

where n is a point of the set Ω'. Furthermore, on the set Ω',

$$H_{\vartheta_1}(\bar{n}) \leqslant H^*_{\vartheta_1}(\bar{n}) \text{ and } H_{\vartheta_2}(\bar{n}) \leqslant H^*_{\vartheta_2}(\bar{n}).$$

Consequently, each point $\bar{x}$ that satisfies (2) certainly satisfies (3). Hence the body $(1-\vartheta')H_{\vartheta_1} + \vartheta' H_{\vartheta_2}$ is included in the body $H_{(1-\vartheta')\vartheta_1+\vartheta'\vartheta_2}$ i.e., the family H_ϑ is convex.

Every convex family of bodies obeys the Brunn inequality. Hence the function

$$\Phi(\vartheta) = \sqrt[n]{V(H_\vartheta)} \tag{4}$$

is convex and $V(H_\vartheta)$ is the volume of the body H_ϑ.

In the previous section it was shown that the volume of the body defined by the function $H^*(\bar{n})$ has a first variation. So the function $\Phi(\vartheta)$ is differentiable. Since it is convex; its derivative $\Phi'(\vartheta)$ is a monotone decreasing function and is constant if and only if H_0 and H_1 are homothetic, because only in this case equality holds in the Brunn inequality.

On the set Ω' let us define a positive continuous function $H^*(\bar{n})$. It determines a convex body H. If $H^*(\bar{n}) = R$ is a constant, then all the supporting planes of the body H whose normals fall into Ω' touch a sphere of radius R. All other supporting planes of this body, by Lemma III of the previous section, pass through the singular points on its surface. Such a body is said to be *circumscribed about a sphere*.

Let H_0 be the body defined by the function $H^*(\bar{n})$. Consider the family of bodies H_ϑ defined by the function $H^*(\bar{n}) + \vartheta$ $(0 \leqslant \vartheta \leqslant 1)$. All these bodies H_ϑ have the same domain of definition for the support function. The ϑ-derivative of $\Phi(\vartheta)$ is

$$\Phi'(\vartheta) = \frac{F(H_\vartheta, \Omega)}{nV(H_\vartheta)^{\frac{n-1}{n}}}. \tag{5}$$

Indeed, in this case we have $\delta[H^*(\bar{n}) + \vartheta] = 1$ and, therefore by Lemma VI of the previous section, we obtain formula (5). Since $F(H_\vartheta, \Omega - \Omega') = 0$, we find $F(H_\vartheta, \Omega') = F(H_\vartheta, \Omega)$, i.e., equal to the area of the surface of the body H_ϑ.

The derivative $\Phi'(\vartheta)$ decreases steadily and remains constant if and only if the terminal bodies H_0 and H_1 of the family are homothetic. Therefore the ratio K of the nth power of the area surface of H_ϑ to the $(n-1)$th power of its volume has the same property. Similarly we can pass on from the body given by the function $H^*(\bar{n})+1$ to the body given by the function $H^*(\bar{n}) + 2$, and so on. The bodies thus obtained can be converted so that they all have the same volume. The ratio Kn defined above however does not change in this process. All the bodies obtained in this way converge to

a body circumscribed about a sphere. However, the ratio K decreases and remains constant if and only if the initial body is homothetic to all of them, i.e., homothetic to the limit body, i.e., if the initial body is circumscribed about a sphere. Thus we complete the proof of the following

Theorem. *Amongst all the convex bodies having the same domain of definition for the support functions, that body has the least surface with given volume which is circumscribed about a sphere and only such a body has this minimal property.*

This theorem includes the Lindelöf theorem which asserts that *amongst all polyhedra with given system of normals to faces, that polyhedron has least area with given volume which is circumscribed about a sphere*. It also includes, for example, the following: *amongst all right convex cylinders with given volume, that cylinder has least area which is circumscribed about a sphere*. Of course, we can give many other particular cases of this theorem which have simple intuitive meaning.

§ 3. Existence of a Convex Body with Given Area Function

In Part I we showed that the area function $F(H,\omega)$ of a convex body H is, first, nonnegative, second, countably additive, third, satisfies the condition:

$$\int\limits_{\Omega} \bar{n}\, F(H, d\omega) = 0,$$

and, finally, if H has interior points, then for an arbitrary unit vector $\bar{n}_0$

$$\int\limits_{\Omega} |\bar{n}_0 \bar{n}|\; F(H, d\omega) > 2a > 0,$$

where a is a constant the same for all $\bar{n}_0$. These four conditions are not only necessary, but also sufficient for a set function on the unit sphere to be the area function of some convex body having interior points.

Theorem. *Let $F(\omega)$ be a nonnegative and countably additive set function defined on the unit sphere and satisfying the two conditions*

$$\int\limits_{\Omega} \bar{n}\; F(d\omega) = 0, \tag{1}$$

and for an arbitrary unit vector $\bar{n}_0$

$$\int\limits_{\Omega} |\bar{n}_0 \bar{n}|\; F(d\omega) > 2a > 0, \tag{2}$$

where a is a constant the same for all $\bar{n}_0$.[7] *Then there exists a convex body (unique up to a translation) having interior points and the given $F(H, \omega)$ as its area function.*

Prior to proving the theorem, let us make two remarks.

Conditions (1) and (2) are easily given an intuitive interpretation. Let us assume that the surface of the unit ball is loaded with mass so that $F(\omega)$ is the mass of the set ω. This representation is admissible because the function $F(\omega)$ is nonnegative and countably additive. Condition (1) requires that the centroid of the system of masses be at the center of the ball. And condition (2) requires that the loading be not concentrated entirely over a single great circle or, in the n-dimensional case, not entirely over the single $(n-1)$-dimensional sphere obtained from the intersection of the unit sphere by a plane passing through its center. This follows from the fact that if $\bar{n}_0$ is taken as a pole, then $|\bar{n}_0\bar{n}|$ is zero only on the equator.

The theorem formulated above implies that it is possible to realize a nonnegative and countably additive set function on the unit sphere by means of a convex surface. For this purpose, let us add to such a function certain point loadings so that the "augmented" function satisfies conditions (1) and (2). It is always possible to find not more than $(n+1)$ such point loadings, say, for instance, taking them at the vertices of a regular simplex. The augmented function then becomes the area function of some convex body. The augmented point loadings are represented by the areas of certain plane domains on the surface of the body. Cutting out these plane domains, we obtain a convex surface with the property that the measure of the set of the points through which pass the supporting planes with normals falls in the set ω on the unit sphere is equal to the value of the given set function for the set ω.

We now take up the proof of our theorem. First note that we already proved in Part II of this work the uniqueness of a convex body having interior points and a given area function. To prove the particular case in hand there is no need to retrace all the steps in the proof of the general uniqueness theorem. Here it suffices to know the Brunn theorem and expression for the mixed volume in terms of the area functions found in Part I. Suppose that the convex bodies H_1 and H_2 have equal area functions: $F(H_1, \omega) = F(H_2, \omega)$. "Differentiating," and then multiplying by $H_1(\bar{n})$ and finally integrating over the unit ball, we obtain

$$V(H_1, \dots, H_1) = V(H_1, H_2, \dots, H_2). \tag{3}$$

From the Brunn-Minkowski inequality

$$V(H_1, H_2, \dots, H_2)^n \geqslant V(H_1, \dots, H_1)V(H_2, \dots, H_2)^{n-1}, \tag{4}$$

[7]Since $\int_\Omega |\bar{n}_0\bar{n}|\, F(d\omega)$ is a continuous function of $\bar{n}_0$, it suffices to require that it be positive; then there exists an a such that condition (2) is satisfied.

we find that

$$V(H_1, \ldots, H_1) \geqslant V(H_2, \ldots, H_2). \tag{5}$$

Since neither of the bodies H_1 and H_2 is to be preferred to the other, equality must hold here, i.e., there must be equality in (4). If the bodies H_1 and H_2 have interior points, then this is possible only when they are homothetic. However, as their volumes are equal, the bodies themselves are translates of one another.

We prove that a convex body with given area function $F(\omega)$ exists, by searching for a function among all positive continuous functions $H^*(\bar{n})$ with "volume" equal to one[8]

$$V(H^*) = 1, \tag{6}$$

that gives a minimum value to the integral

$$\int\limits_{\Omega} H^*(\bar{n})\, F(d\omega). \tag{7}$$

Note that it suffices to search for this minimum among the support functions of convex bodies. Indeed, if $H(\bar{n})$ is the support function of the body H given by the function $H^*(\bar{n})$, then, by the definition of "volume" of $H^*(\bar{n})$, we have $V(H^*) = V(H)$, while $H(\bar{n}) \leqslant H^*(\bar{n})$. So a minimum of integral (7) attained over the set of support functions will also be a minimum over the whole set of positive continuous functions.

Second, we can confine ourselves only to convex bodies with centroid at the origin. Indeed, the volume of a convex body does not change under translation, but the support function acquires a factor $\bar{a}\bar{n}$, where a is the translation vector. Integral (7) for the support function of the translated body is

$$\int\limits_{\Omega} (H(\bar{n}) + \bar{a}\ \bar{n})\, F(d\omega), \tag{8}$$

i.e., exactly equal to the integral of the body that is not translated because, by condition (1), we find that

$$\int\limits_{\Omega} \bar{a}\ \bar{n}\, F(d\omega) = 0. \tag{9}$$

We now prove that a minimum does exist. Take some convex body with unit volume and let its integral (7) be equal to M. Of course, we can confine

[8]According to the definition given in § 1, the "volume" of $H^*(\bar{n})$ is the volume of the convex body defined by $H^*(\bar{n})$. See also the final remark at the end of § 1.

ourselves only to the convex bodies for which

$$\int_{\Omega} H(\bar{n})\, F(d\omega) \leqslant M. \tag{10}$$

We show that all such bodies (with unit volume and centroid at the origin) are uniformly bounded. Let r be the length of a segment starting from the origin in the direction of n and lying entirely in one of the bodies H belonging to our set. The support function of such a segment, as is known and easily verified, is

$$\frac{r}{2}|\bar{n}_0\bar{n}| + \frac{r}{2}\bar{n}_0\bar{n},$$

and since the segment lies in the body H, we have

$$\frac{r}{2}|\bar{n}_0\bar{n}| + \frac{r}{2}\bar{n}_0\bar{n} \leqslant H(\bar{n}). \tag{11}$$

By assumption, the function $F(d\omega)$ is nonnegative; hence, by condition (1) [see formula (9)], we find

$$\frac{r}{2}\int_{\Omega} |\bar{n}_0\bar{n}|\, F(d\omega) \leqslant \int_{\Omega} H(\bar{n})\, F(d\omega) \leqslant M. \tag{12}$$

By the condition imposed on $F(\omega)$, we have

$$\int_{\Omega} |\bar{n}_0\bar{n}|\, F(dw) > 2a > 0. \tag{13}$$

Therefore

$$r < \frac{M}{a}. \tag{14}$$

Hence the lengths of the segments lying in our convex bodies are bounded. Consequently, the bodies themselves are uniformly bounded.

Blaschke's selection theorem asserts the existence of a minimum for every continuous functional defined on a uniformly bounded set of convex bodies.[9] Consequently, integral (7) actually attains its minimum for some body H_0 with unit volume.

Thus we arrive at the following result: the integral

$$\int_{\Omega} H^*(\bar{n})\, F(d\omega)$$

[9]In terms of functional analysis the *Blaschke selection theorem*, Auswahlzatz, reads: A bounded set of convex bodies is compact. And a continuous functional attains a minimum on every compact set [Blaschke, Kreis und Kugel, §18].

under the additional condition of $V(H^*) = 1$ is minimum for $H^*(\bar{n}) = H_0(\bar{n})$, where $H_0(\bar{n})$ is the support function of a convex body with unit volume and centroid at the origin. Since $H_0(\bar{n}) > 0$, Lemma VI of § 1 says that for $H^*(\bar{n}) - H_0(\bar{n})$, the "volume" of $H^*(\bar{n})$ has a first variation and

$$\delta V(H^*) = \int_\Omega \delta H^*(\bar{n})\, F(H_0, d\omega). \tag{15}$$

So we can apply variational techniques. Using the *Lagrange multiplier method*, we obtain

$$\int_\Omega \delta H^*(\bar{n})\, F(d\omega) = \lambda \int_\Omega \delta H^*(\bar{n})\, F(H_0, d\omega). \tag{16}$$

The continuous function $\delta H^*(\bar{n})$ being arbitrary, we find

$$F(\omega) = \lambda F(H_0, \omega).$$

On dilating the body H_0 by the similarity factor of $\sqrt[n-1]{\lambda}$, we find its area function increases λ times and becomes equal to $F(\omega)$. Consequently, the body $\sqrt[n-1]{\lambda} H_0$ is the one we are searching.

§ 4. Existence of a Convex Body with Given Curvature Function

It is natural to ask whether the theorem on the existence of a convex body with given area function could be extended to curvature functions of arbitrary order. If this is not possible then, by examining curvature functions of a general type, we may become able to prove that *for a positive analytic function $f(\bar{n})$ defined on the unit sphere and satisfying the condition*

$$\int_\Omega \bar{n}\, f(\bar{n})\, d\omega = 0,$$

there exists a convex body for which $f(\bar{n})$ is its mth curvature function ($m < n - 1$). This question seems to have a negative answer. Namely, we prove the following

Theorem. *There exist infinitely many positive analytic functions $f(\bar{n})$ on the unit sphere satisfying the condition*

$$\int_\Omega \bar{n} f(\bar{n})\, d\omega = 0, \tag{1}$$

which are not the mth curvature functions of any convex body whatsoever if $m < n-1$ (where n as usual is the dimension of the space).

This theorem by itself is not of great interest. Certain authors however believed in the truth of its converse that they even asserted the corresponding false theorem as proved for particular cases or attempted to prove it.[10] Thus, it was asserted that if a positive function $f(\bar{n})$ satisfied condition (1), then there existed a convex body for which $f(\bar{n})$ was the sum of the principal radii of curvature. That this theorem is false can be verified by the following simple example of a surface of revolution in three-dimensional space. Let the distance from the origin to the tangent plane of this surface with normal n be

$$H(\bar{n}) = \frac{3}{2} - \cos^2\theta, \tag{2}$$

where θ is the polar distance of the normal $\bar{n}$ from the axis of the surface. We know that the sum of the principal radii of curvature is:

$$R_1 + R_2 = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial H}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 H}{\partial\varphi^2} + 2H \tag{3}$$

(φ is the longitude). Substituting $H = \frac{3}{2} - \cos^2\theta$ we find

$$R_1 + R_2 = 4\cos^2\theta + 1 > 0. \tag{4}$$

However, the surface under consideration is not convex. The radius of curvature of a meridian of it can be found from the familiar relation:

$$R = H + \frac{d^2 H}{d\theta^2}, \tag{5}$$

as H depends only on θ. By simple calculation, we obtain

$$R = 3\cos^2\theta - \frac{1}{2}, \tag{6}$$

i.e., $R = -\frac{1}{2}$ for $\theta = \frac{\pi}{2}$ and $R = \frac{5}{2}$ for $\theta = 0$. Since the radius of curvature of a meridian changes its sign, the surface is not convex and has a regression edge if $R = 3\cos^2 - \frac{1}{2} = 0$. The fact that no other surface has this sum of principal radii of curvature is implied in the uniqueness of the solution of the differential equation (3) for a given sum $R_1 + R_2$ (uniqueness up to the factor $\bar{a}\bar{n}$ of translation).

[10] See Favard, Sur la détermination des surfaces convexes, Bull. de Cl. des Sci. Acad. Roy. Belgigue, (5), **19** (1933); Süss, Beschtimmung einer geschlosenen Konvexen Fläche durch die Summe ihrer Hauptkrümmungsradien, Math. Ann., **108** (1933); Bonnesen and Fenchel, Theorie der konvexen Körper, p. 123.

We now prove the theorem that is formulated above. This permits us to gain a somewhat deeper understanding of the cases in which the theorem on the existence of a convex body with given arbitrary curvature functions is false.

We say a set function $F(\omega)$ has a *point loading* at a point $\bar{n}$ if, on taking the point $\bar{n}_0$ as the set ω, we find $F(\omega) \neq 0$.

Lemma I. *Curvature functions of order less than* $(n-1)$ *have no point loadings.*

Assume the contrary and let $F_m(H,\omega)$ be a curvature function of the convex body H having a point loading at the point $\bar{n}_0$. Take just this point as the set ω_0. For this set the value of the area function of the body $H+\lambda E$ parallel to H[11] is

$$F_{n-1}(H+\lambda E,\omega_0) = \sum_{k=0}^{n-1} \lambda^{n-k-1} C_{n-1}^k F_k(H,\omega_0) > 0, \tag{7}$$

since, by assumption, $F_m(H,\omega_0) > 0$. So the body $H+\lambda E$ has a whole face with normal $\bar{n}_0$ and with area equal to $F_{n-1}(H+\lambda E,\omega_0)$. Returning back to the body H, we draw interior segments of length λ on the normals to this face. All these segments beginning at this face are parallel and, consequently, their endpoints sweep exactly a similar face on the surface of H. Therefore

$$F_{n-1}(H+\lambda E,\omega_0) = F_{n-1}(H,\omega_0),$$

which means

$$F_m(H,\omega_0) = 0 \quad \text{for } m < n-1.$$

We say a sequence of point functions $f_1(\bar{n}), f_2(\bar{n}), \dots$ *converges weakly* to the set function $F(\omega)$ if, for every continuous $Z(\bar{n})$

$$\lim_{k=0} \int_\Omega Z(\bar{n})\, f_k(\bar{n})\, d\omega = \int_\Omega Z(\bar{n})\, F(d\omega). \tag{8}$$

Lemma II. *For each nonnegative and countably additive set function* $F(\omega)$ *on the unit sphere satisfying the conditions*

$$\int_\Omega \bar{n}\, F(d\omega) = 0, \qquad \int_\Omega |\bar{n}_0 \bar{n}|\, F(d\omega) > 2a > 0, \tag{9}$$

it is possible to construct a sequence of positive analytic functions $f_k(\bar{n})$ *converging weakly to* $F(\bar{n})$ *and satisfying the condition*

$$\int_\Omega \bar{n} f_k(\bar{n})\, d\omega = 0. \tag{10}$$

[11]Here and in what follows, E is the unit ball.

If $F(\omega)$ satisfies these conditions; then, as has been shown, there exists a convex body H for which $F(\omega)$ is the area function. Minkowski proved that every convex body can be approximated with any accuracy by regular convex bodies with analytic support functions. Construct a sequence of bodies $H_1, H_2, \dots$ converging to H. The reciprocal Gaussian curvatures of the bodies H_k form a sequence of functions $f_k(\bar{n})$ converging weakly to $F(\omega)$ since

$$\frac{1}{n}\int\limits_{\Omega} Z(\bar{n})\, f_k(\bar{n})\, d\omega = V(Z, H_k, \dots, H_k), \tag{11}$$

and by the lemma proved in § 6 of Part I,

$$\lim_{k=\infty} V(Z, H_k, \dots, H_k) = V(Z, H, \dots, H) = \frac{1}{n}\int\limits_{\Omega} Z(\bar{n})\, F(d\omega). \tag{12}$$

Now assume, contrary to our theorem, that for each positive analytic function $f(\bar{n})$ on the unit sphere satisfying condition (1), there exists a convex body for which $f(\bar{n})$ is the mth curvature function in the usual sense. Let $F(\omega)$ be a set function with point loadings satisfying conditions (9). Construct a sequence of functions $f_k(\bar{n})$ converging weakly to $F(\omega)$. Let H_k be a convex body with $f_k(\bar{n})$ as its mth curvature function.

We now prove the bodies H_k all are bounded. Of course, we could as well assume their centroids lie at the origin. The integrals $\int_\Omega f_k(\bar{n})\, d\omega$ are uniformly bounded and as the functions $f_k(\bar{n})$ converge weakly to $F(\omega)$, we have

$$\lim_{k=\infty} \int\limits_{\Omega} f_k(\bar{n})\, d\omega = F(\Omega). \tag{13}$$

Moreover,

$$\frac{1}{n}\int\limits_{\Omega} f_k(\bar{n})\, d\omega = V_m(H_m) = V(\underbrace{H_k, \dots, H_k}_{m}, E, \dots, E), \tag{14}$$

because $f_k(\bar{n})$ is the mth curvature function of the body H_k. The inequality between the quermassintegrals, which we proved in § 4 of Part II of this work, yields

$$V_m(H_k)^{m+1} \geqslant V_{m+1}(H_k)^m V(E), \tag{15}$$

where $V(E)$ is the volume of the unit ball. Hence, the numbers $V_m(H_k)$ all being bounded, we obtain

$$V_{m+1}(H_k) < M. \tag{16}$$

Let r be the length of the segment that emanates from the origin in the direction $\bar{n}_0$ and lies entirely in the body H_k. Its support function is

$$\frac{r}{2}|\bar{n}_0\bar{n}| + \frac{r}{2}\bar{n}_0\bar{n} \leqslant H_k(\bar{n}_0). \tag{17}$$

By virtue of this inequality, using condition (10), we derive

$$\frac{r}{2}\int\limits_{\Omega} |\bar{n}_0\bar{n}|\, f_k(\bar{n})\, d\omega \leqslant \int\limits_{\Omega} H_k(\bar{n})\, f_k(\bar{n})\, d\omega. \tag{18}$$

The integral on the right side is equal to $nV_{m+1}(H_k)$. Hence from (10) we obtain

$$\frac{r}{2}\int\limits_{\Omega} |\bar{n}_0\bar{n}|\, f_k(\bar{n})\, d\omega < nM. \tag{19}$$

Since the functions $f_k(\bar{n})$ are positive and converge weakly to the function $F(\omega)$ satisfying conditions (9), the integral on the left side of (19) is bounded below for all $(\bar{n}_0)$ and the number k is bounded by a certain constant $2b$.[12] Therefore from (19) we obtain

$$r < \frac{nM}{b}. \tag{20}$$

Hence the lengths of segments all lying in the bodies H_k are uniformly bounded and, consequently, the bodies are bounded.

Now using the Blaschke selection theorem, we can choose a convergent sequence from the set of all bodies H_k. We denote the bodies of this sequence also by H_k.

Let H be the limit body of this sequence and $F_m(H,\omega)$ its mth curvature function. Since the functions $f_k(\bar{n})$ are the mth curvature functions, in the ordinary sense, of the bodies H_k, by the lemma proved in § 6, Part 1, for arbitrary continuous $Z(n)$, we have

$$\int\limits_{\Omega} Z(\bar{n})\, F_m(H, d\omega) = \lim_{k=\infty} \int\limits_{\Omega} Z(\bar{n})\, f_k(\bar{n})\, d\omega. \tag{21}$$

[12] This assertion is based on the lemma: *if the functions $F_k(\omega)$ are countably additive and converge weakly to $F_0(\omega)$ and $K(\bar{n}_0, n)$ is equicontinuous in $\bar{n}_0$ for all $(\bar{n})$, we find that $\varphi_k(\bar{n}_0)$ converges uniformly in $\bar{n}_0$ to $\varphi_0(\bar{n}_0)$, where $\int K(\bar{n}_0, n)\, F_k(d\omega) = \varphi_k(\bar{n}_0)$.*

Suppose that $\operatorname{Var} F_k(\omega) < M$. For every $\varepsilon > 0$ there exists a δ such that for $|\bar{n}_0' - \bar{n}_0''| < \delta$, we have

$$|\varphi_k(\bar{n}_0') - \varphi_k(\bar{n}_0'')| \leqslant \int\limits_{\Omega} |K(\bar{n}_0', \bar{n}) - K(\bar{n}_0'', \bar{n})| \operatorname{Var} F_k(d\omega) < \varepsilon M,$$

so the $\varphi_k(\bar{n}_0)$'s are equicontinuous and, since they converge to $\varphi_0(\bar{n}_0)$, this convergence is uniform.

And since the functions $f_k(\bar{n})$ converge weakly to $F(\omega)$,

$$\int\limits_{\Omega} Z(\bar{n})\, F(\omega) = \lim_{k=\infty} \int\limits_{\Omega} Z(\bar{n})\, f_k(\bar{n})\, d\omega. \tag{22}$$

Comparing (21) and (22) and since $Z(\bar{n})$ is an arbitrary continuous function, we find

$$F_m(H, \omega) = F(\omega). \tag{23}$$

However, if the function $F(\omega)$ has point loadings, then, for $m < n - 1$, this equality is impossible as implied in Lemma I.

Consequently, a contradiction is obtained and this proves the theorem. The positive functions $f(\bar{n})$, being rather too close to the set functions $F(\omega)$ with point loadings, cannot be the curvature function of an order less than $(n - 1)$ for any convex body.

Institute of Mathematics and Mechanics,
Leningrad State University

CHAPTER VI

TO THE THEORY OF MIXED VOLUMES OF CONVEX BODIES PART IV[1]

Matematicheskiĭ Sbornik, vol. 3 (45), No. 2, 227–249 (1938).

MIXED DISCRIMINANTS AND MIXED VOLUMES

This paper is devoted, first of all, to studying the invariants of a family of quadratic forms in n variables, the so-called mixed discriminants of these forms. The relation between these discriminants and the questions in the theory of convex bodies was established first by Minkowski and Weyl.[2] The algebraic part (§ 1–3) of the work is presented without any reference to convex bodies since it possibly has intrinsic interest. Furthermore, I demonstrate how one can, basing on the algebraic results, deduce the main theorems of the theory of mixed volumes of convex bodies, true, confining oneself only to bodies with twice differentiable support functions. In § 4 I derive the concept of mixed volume and its properties from an application of mixed discriminants. In § 5 I prove the uniqueness of a convex body with given curvature functions without taking recourse to the inequalities between mixed volumes. In § 6 I deduce the main inequality between mixed volumes by the *Hilbert method*.[3] Finally, in § 7 I generalize two theorems proved in Part II, namely, the theorem on the uniqueness of a convex body with given curvature functions and the theorem on the equality in the Brunn inequality. This last section stands somewhat alone and is rather closely related to Part II of this work.

[1]For the preceding parts of this work, see Chapters III, IV, and V in this monograph.

[2]Minkowski, Ges. Abh., II; Weyl, Sitz. Ber. Preuss. Akad., 1917. See also Bonnesen and Fenchel, Theorie der konvexen Körper.

[3]Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Berlin (1912).

§ 1. Basic Properties of Mixed Discriminants

Let $f_1, f_2, \dots, f_m$ be quadratic forms in n variables:

$$f_k = \sum_{i,j=1}^{n} a_{ij}^{(k)} x_i x_j. \tag{1}$$

Their linear combination

$$f = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_m f_m \tag{2}$$

is also a quadratic form in these n variables with coefficients

$$a_{ij} = \lambda_1 a_{ij}^{(1)} + \lambda_2 a_{ij}^{(2)} + \cdots + \lambda_m a_{ij}^{(m)}. \tag{3}$$

The *discriminant* $D(\underbrace{f, \dots, f}_{n})$ of f is a homogeneous polynomial of degree n in $\lambda_1, \dots, \lambda_m$:

$$D(f, \dots, f) = \sum_{k_1, \dots, k_n} \lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_n} D(f_{k_1}, f_{k_2}, \dots, f_{k_n}). \tag{4}$$

Here summation is taken over all $k_1, \dots, k_n$ which run independently from 1 to m. That the coefficient $D(f_{k_1}, \dots, f_{k_n})$ of the product $\lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_n}$ depends only on the forms $f_{k_1}, \dots, f_{k_n}$ can be proved by putting the remaining λ's equal to zero. The coefficients $D(f_{k_1}, \dots, f_{k_n})$ are chosen so that they do not depend on the order of occurrence of the forms $f_{k_1}, \dots, f_{k_n}$. These coefficients are called *mixed discriminants* of the forms $f_{k_1}, \dots, f_{k_n}$.

Expanding the determinant

$$D(f, \dots, f) = |\lambda_1 a_{ij}^{(1)} + \cdots + \lambda_m a_{ij}^{(m)}|, \tag{5}$$

we find

$$\begin{aligned} &D(\underbrace{f_1, \dots, f_1}_{p_1}, \underbrace{f_2, \dots, f_2}_{p_2}, \dots, \underbrace{f_m, \dots, f_m}_{p_m}) \\ &= \frac{p_1!\, p_2! \dots p_m!}{n!} \sum \begin{vmatrix} a_{11}^{(k_1)} a_{12}^{(k_1)} \dots a_{1n}^{(k_1)} \\ a_{21}^{(k_2)} a_{22}^{(k_2)} \dots a_{2n}^{(k_2)} \\ \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ a_{n1}^{(k_n)} a_{n2}^{(k_n)} \dots a_{nn}^{(k_n)} \end{vmatrix}, \end{aligned} \tag{6}$$

where summation is taken over all permutations of the superscripts k_i which take p_1 times the value one, p_2 times the value two and so on.

Note that the definitions and the properties of mixed discriminants presented in this paper, together with their proofs, can be extended without any modification to the case in which the forms are Hermitian.

A mixed discriminant has the following properties:

(1) If the forms $f_1, f_2, \dots, f_n$ are equal to a form f, then

$$D(f_1, f_2, \dots, f_n) = D(f, \dots, f)$$

is the discriminant of the form f.

(2) Let $f = \lambda_1 f^{(1)} + \lambda_2 f^{(2)} + \cdots + \lambda_p f^{(p)}$. Then

$$D(\underbrace{f, \dots, f}_{m}, f_1, \dots, f_{n-m}),$$

where $f^{(1)}, \dots, f^{(p)}$, $f_1, \dots, f_{n-m}$ are arbitrary quadratic forms, is a homogeneous polynomial of degree m in $\lambda_1, \lambda_2, \dots, \lambda_p$:

$$\begin{aligned} &D(f, \dots, f, f_1, \dots, f_{n-m}) \\ &= \sum_{k_1, \dots, k_m} \lambda_{k_1} \dots \lambda_{k_m} D(f^{(k_1)}, \dots, f^{(k_m)}, f_1, \dots, f_{n-m}), \end{aligned} \tag{7}$$

where summation is taken over all $k_1, \dots, k_m$ from 1 to p. This property is derived from (4) by equating coefficients.

(3) If the forms $f_1, \dots, f_n$ are subject to the same linear transformation with determinant D, then $D(f_1, \dots, f_n)$ is multiplied by a factor D^2.

(Here and in what follows, the forms $f_1, \dots, f_n$ all need not necessarily be distinct; n always stands for the number of variables.)

If the forms $f_1, \dots, f_n$ are subject to the same transformation, then irrespective of the values of the coefficients $\lambda_1, \lambda_2, \dots, \lambda_n$, the form $f = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n$ suffers the same transformation. Hence its discriminant acquires the factor D^2, in other words the coefficients of polynomial (4) each acquires the factor D^2.

Hence, in particular, it follows that a unimodular transformation of the forms $f_1, \dots, f_n$ does not change their mixed discriminant $D(f_1, \dots, f_n)$. We often use this property.

(4) $D(f_1, \dots, f_n)$ is a linear form of the coefficients of each of the forms $f_1, \dots, f_n$:

$$D(f_1, \dots, f_n) = \sum_{i,j=1}^{n} D(f_1, \dots, f_{n-1})_{ij}\, a_{ij}^{(n)}. \tag{8}$$

Here $D(f_1, \dots, f_{n-1})_{ij}$ is equal, up to an unimportant factor n, to the mixed discriminant of the forms obtained from the forms $f_1, \dots, f_{n-1}$ by deleting all coefficients with subscript i.

Assuming that the forms in (6) are all formally distinct and expanding the determinants over $a_{ij}^{(n)}$, we find our assertion is true.

(5) If $f_1, \ldots, f_n$ are all positive-definite forms, then $D(f_1, \ldots, f_n) > 0$.

For a form in one variable this assertion is trivial. Therefore, assuming that it holds for a form of $(n-1)$ variables, we prove that it also holds for a form of n variables.

Transform the forms $f_1, \ldots, f_n$ so that f_n is reduced to a sum of squares. We denote a new form by an overbar. If the transformation is orthogonal, then by properties (3) and (4), we infer

$$D(f_1, \ldots, f_n) = \sum_{i=1}^{n} D(\bar{f}_1, \ldots, \bar{f}_{n-1})_{ii}\, \bar{a}_{ii}^{(n)}, \tag{9}$$

where, by the induction assumption, $D(\bar{f}_1, \ldots, \bar{f}_{n-1}) > 0$ since deletion of all coefficients with subscript i from positive-definite forms yields positive-definite forms. In addition, the positive definiteness of the form f_n implies that all $\bar{a}_{ii}^{(n)} > 0$. Hence $D(f_1, \ldots, f_n) > 0$.

A well-known example of mixed discriminants is furnished by the coefficients of the characteristic equation of any quadratic form f. If $e = x_1^2 + x_2^2 + \cdots + x_n^2$, then the characteristic equation of the form f can be written as:

$$D(f - \lambda e, \ldots, f - \lambda e) = \sum_{m=0}^{n} (-1)^m \lambda^m C_n^m\, D(\underbrace{f, \ldots, f}_{n-m}, \underbrace{e, \ldots, e}_{m}) = 0.$$

§ 2. A Form Associated with (n − 1) Forms

Let $f = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n$ and let F be the matrix of the form f. We denote the determinant of the matrix by the symbol of the matrix in parentheses with indices showing, as always, the places of the elements in the matrix.

A matrix composed of the cofactors of a matrix F is denoted as

$$\|D(\underbrace{f, \ldots, f}_{n-1})_{ik}\| = |F|\, F^{-1}. \tag{1}$$

If a form f, when transformed with matrix T, turns into the form f' with matrix F'; then

$$F' = \widetilde{T} F T, \tag{2}$$

where the tilde stands for the transposed matrix

$$F'^{-1} = T^{-1} F^{-1} \widetilde{T}^{-1}. \tag{3}$$

Therefore

$$\|D(f',\ldots,f')_{ik}\| = |F'|\,{F'}^{-1} = |T|^2\,T^{-1}\,|F|\,F^{-1}\,\widetilde{T}^{-1}. \tag{4}$$

If $|T| = 1$, then

$$\|D(f',\ldots,f')_{ik}\| = T^{-1}\,\|D(f,\ldots,f)_{ik}\|\,\widetilde{T}^{-1} \tag{5}$$

and

$$D(f',\ldots,f')_{ik} = \sum_{j,i=1}^{n}(T^{-1})(T^{-1}_{ij})_{ki}\,D(f,\ldots,f)_{ji}. \tag{6}$$

We now extend this formula to "mixed minors" of $D(f_1,..,f_{n-1})_{ik}$

$$f = \sum_{q=1}^{m}\lambda_q f_q \tag{7}$$

and

$$D(f,\ldots,f) \;= \sum_{q_1,\ldots,q_n}\lambda_{q_1}\ldots\lambda_{q_n}D(f_{q_1},\ldots,f_{q_n}). \tag{8}$$

Expanding $D(f,\ldots,f)$ by the minors, we infer

$$D(f,\ldots,f)_{ik} \;= \sum_{q_i,\ldots,q_{n-i}}\lambda_{q_1}\ldots\lambda_{q_{n-1}}\,D(f_{q_1},\ldots,f_{q_{n-1}})_{ik}. \tag{9}$$

Since simultaneous transformation of the forms f_q induces the same transformation of the form f, irrespective of the values of λ_q; applying a unimodular transformation with matrix T to every f_q, by virtue of (6) and (9), we obtain

$$D(f'_1,\ldots,f'_{n-1})_{ik} = \sum_{i,j}^{n}(T^{-1})_{ij}\,(T^{-1})_{ki}\,D(f_1,\ldots,f_{n-1})_{ji}. \tag{10}$$

Consider the quadratic form

$$\sum_{i,k=1}^{n}D(f_1,\ldots,f_{n-1})_{ik}\,x_i x_k. \tag{11}$$

Applying a unimodular transformation with matrix S to (11), we find that the coefficients of the resulting form are

$$D'(f_1,\ldots,f_{n-1})_{ik} = \sum_{j,i=1}^{n}(S)_{ji}\,(S)_{ik}\,D(f_1,\ldots,f_{n-1})_{ji}. \tag{12}$$

Hence if $S = \widetilde{T}^{-1}$, then

$$D(f'_1,\ldots,f'_{n-1})_{ik} = D'(f_1,\ldots,f_{n-1})_{ik}. \tag{13}$$

This proves

Lemma I. *A unimodular transformation with matrix S applied to the form*

$$\sum_{i,k=1}^{n} D(f_1, \ldots, f_{n-1})_{ik}\, x_i x_k$$

has the same effect as the transformation with matrix $\widetilde{S}^{-1}$ applied to the forms $f_1, \ldots, f_{n-1}$.

Therefore it is always possible to transform the forms $f_1, \ldots, f_{n-1}$ in such a way that the form

$$\sum_{i,k=1}^{n} D(f_1, \ldots, f_{n-1})_{ik}\, x_i x_k$$

is reduced to a sum of squares, i.e., so that $D(f_1', \ldots, f_{n-1}')_{ik} = 0$ for $i \neq k$.

If the forms $f_1, \ldots, f_{n-1}$ are positive-definite; then , on transforming, them in such a way that $D(f_1', \ldots, f_{n-1})_{ik} = 0$ for $i \neq k$, we find that the form

$$\sum_{i,k=1}^{n} D(f_1', \ldots, f_{n-1}')_{ii}\, x_i^2$$

is positive-definite because by the lemma proved in § 1, $D(f_1', \ldots, f_{n-1}')_{ii} > 0$. However, the same result is obtained on transforming the form

$$\sum_{i,k=1}^{n} D(f_1, \ldots, f_{n-1})_{ik}\, x_i x_k$$

in an appropriate way. This proves

Lemma II. *If the forms $f_1, \ldots, f_{n-1}$ are positive-definite, then the form*

$$\sum_{i,k=1}^{n} D(f_1, \ldots, f_{n-1})_{ik}\, x_i x_k$$

is also positive-definite.

§ 3. Inequalities Between Mixed Discriminants

Theorem. *Let the forms $f_1, \ldots, f_{n-1}$*

$$f_m = \sum_{i,k=1}^{n} a_{ik}^{(m)}\, x_i x_k \quad (m = 1, \ldots, n-1)$$

be positive-definite and

$$g = \sum_{i,k=1}^{n} b_{ik}\, x_i x_k$$

be an arbitrary form, then

$$D(f_1, \ldots, f_{n-1}, g)^2 \geqslant D(f_1, \ldots, f_{n-1}, f_{n-1})\, D(f_1, \ldots, f_{n-2}, g, g),$$

where equality holds if and only if $g = \lambda f_{n-1}$ *(*λ *is a constant).*

Since $D(f_1, \ldots, f_{n-1}, f_{n-1}) > 0$, this theorem yields the next

Corollary. *If* g *satisfies the condition*

$$D(f_1, \ldots, f_{n-1}, g) = 0, \tag{1}$$

then

$$D(f_1, \ldots, f_{n-2}, g, g) \leqslant 0, \tag{2}$$

where equality holds if and only if g *is identically zero.*

We now prove the converse, namely, that the theorem follows from the corollary stated above. Suppose g to be an arbitrary form and λ to be chosen so that

$$D(f_1, \ldots, f_{n-1}, g) = \lambda D(f_1, \ldots, f_{n-1}, f_{n-1}). \tag{3}$$

Then, putting $g - \lambda f_{n-1} = g'$, we derive

$$D(f_1, \ldots, f_{n-1}, g') = 0, \tag{4}$$

and

$$D(f_1, \ldots, f_{n-2}, g', g') \leqslant 0. \tag{5}$$

Substituting $g' = g - \lambda f_{n-1}$, expanding and putting the value of λ from (3), we obtain

$$D(f_1, \ldots, f_{n-2}, g, g) - \frac{D(f_1, \ldots, f_{n-1}, g)^2}{D(f_1, \ldots, f_{n-1}, f_{n-1})} \leqslant 0, \tag{6}$$

and, since $D(f_1, \ldots, f_{n-1}, f_{n-1}) > 0$, we obtain the inequality stated in the theorem.

Equality holds in (6) if and only if it holds in (5), i.e., if and only if $g' = 0 = g - \lambda f_{n-1}$ (by assumption). In other words, $g = \lambda f_{n-1}$.

On the basis of this remark, we prove, not the theorem itself, but the corollary implied in it. First consider binary forms f and g. The form f is

positive-definite, so both f and g can be reduced to sums of squares and, in addition, in such a way that $D(f,g)$ and $D(g,g)$ remain unchanged. Then condition (1) is written as

$$2D(f,g) = a_{11}b_{11} + a_{22}b_{22} = 0,$$

and since $a_{11}, a_{22} > 0$, we have

$$D(g,g) = b_{11}b_{22} \leqslant 0,$$

and equality holds if and only if $b_{11} = b_{22} = 0$.

We now take up forms in n variables and assume that our theorem holds for forms in $(n-1)$ variables.

By property 4, § 1, of mixed discriminants, we find $D(f_1, \ldots, f_{n-2}, g, g)$ is a quadratic form in the coefficients b_{ik} of the form g. Let G denote this quadratic form:

$$G = \sum_{i,k=1}^{n} D(f_1, \ldots, f_{n-2}, g)_{ik}\, b_{ik} = \sum_{i,k,j,l=1}^{n} D(f_1, \ldots, f_{n-2})_{ik,jl}\, b_{ik} b_{jl}. \tag{7}$$

The meaning of the coefficients $D(f_1, \ldots, f_{n-2})_{ik,jl}$ can be inferred from property 4, § 1. Here $D(f_1, \ldots, f_{n-2}, g)_{ii}$ is the mixed discriminant of the forms $f_1^{(i)}, \ldots, f_{n-2}^{(i)}, g^{(i)}$ obtained from $f_1, \ldots, f_{n-2}, g$ by deleting the coefficients with subscript i.

And $D(f_1, \ldots, f_{n-2})_{ii,jj}$ $(i \neq j)$ is the mixed discriminant of the forms obtained from $f_1^{(i)}, \ldots, f_{n-2}^{(i)}$ by deleting the coefficients with subscripts j. Therefore

$$D(f_1, \ldots, f_{n-2})_{ii,jj} \begin{cases} > 0, & i \neq j, \\ = 0, & i = j. \end{cases}$$

We take $b_{ik} = b_{ki}$ to be a variable and build the proof of the theorem with the help of the eigenvalues of the form G.

For the form G to have zero eigenvalues, it is necessary and sufficient that the system of equations

$$D(f_1, \ldots, f_{n-2}, g)_{ik} = 0 \tag{9}$$

admit a nontrivial solution for the unknown variables b_{ik}. The form g_0 with coefficients b_{ik} which are the solutions to this system is called the *eigenform* corresponding to the zero eigenvalue of the form G.

Let us convert the forms $f_1, \ldots, f_{n-2}, g_0$ by the same unimodular transformation. Then, for the transformed forms $f_1', \ldots, f_{n-2}', g_0'$, the conditions

$$D(f_1', \ldots, f_{n-2}', g_0')_{ik} = 0 \tag{10}$$

are satisfied, where the subscripts now pertain to the new variable x. This follows directly from formula (10) of § 2. Hence, if $G = D(f_1, \dots, f_{n-2}, g, g)$ has a zero eigenvalue, then the form $G' = D(f'_1, \dots, f'_{n-2}, g', g')$ obtained from G by subjecting the forms $f_1, \dots, f_{n-2}$, all to a linear transformation, also has a zero eigenvalue. And the corresponding eigenform g'_0 is the transform of the eigenform g_0 of the form G.

Lemma I. *The form $G = D(f_1, \dots, f_{n-2}, g, g)$ has no zero eigenvalue.*

Let g_0 be the eigenform associated with a zero eigenvalue of the form G. Transform all the forms $f_1, \dots, f_{n-2}, g_0$ so that f_{n-2} and g_0 are reduced to sums of squares. This is clearly possible as the form f_{n-2} is positive-definite. For the transformed forms we have the equations

$$D(f'_1, \dots, f'_{n-2}, g'_0)_{ii} = 0 \quad (i = 1, 2, \dots, n). \tag{11}$$

Since $D(f'_1, \dots, f'_{n-2}, g'_0)_{ii}$ is the mixed discriminant of a form in $(n-1)$ variables, and assuming that the theorem holds for such a form, we find

$$D(f'_1, \dots, f'_{n-3}, g'_0, g'_0)_{ii} \leqslant 0 \quad (i = 1, \dots, n), \tag{12}$$

and equality holds for all i if and only if the form g'_0 vanishes identically.

Multiplying (12) by $a'^{(n-2)}_{ii}$ and summing over all i, we obtain

$$\sum_{i=1}^{n} D(f'_1, \dots, f'_{n-3}, g'_0, g'_0)_{ii}\, a'^{(n-2)}_{ii} = D(f'_1, \dots, f'_{n-2}, g'_0, g'_0) \leqslant 0, \tag{13}$$

because $a'^{(n-2)}_{ii} > 0$ and $a'^{(n-2)}_{ik} = 0$ for $i \neq k$.

However, g'_0 being the eigenform corresponding to zero eigenvalue, we find

$$D(f'_1, \dots, f'_{n-2}, g'_0, g'_0) = 0. \tag{14}$$

Consequently, equality holds in all inequalities (12) and the form g'_0 vanishes identically.

Lemma II. *The form $G = D(f_1, \dots, f_{n-2}, g, g)$ has only one positive eigenvalue.*[4]

Let $e = x_1^2 + x_2^2 + \cdots + x_n^2$, and let

$$nD(e, \dots, e, g) = b_{11} + \cdots + b_{nn} = 0. \tag{15}$$

[4]Wherever the number of eigenvalues is mentioned, it is implied that a *multiple eigenvalue* is to be counted as many times as its multiplicity.

Since

$$2\sum_{i<k=1}^{n} b_{ii}b_{kk} = \left(\sum_{i=1}^{n} b_{ii}\right)^2 - \sum_{i=1}^{n} b_{ii}^2, \tag{16}$$

we have

$$\frac{n(n-1)}{2}D(e,\dots,e,g,g) = \sum_{i<k=1}^{n} (b_{ii}b_{kk} - b_{ik}^2) \leqslant 0 \tag{17}$$

and is zero if and only if all b_{ik} vanish. On our imposing an additional condition $\sum b_{ik}^2 = 0$, form (17) attains maximum when $b_{ik} = 0$ for $i \neq k$ and $b_{11} = \dots = b_{nn} = \frac{1}{\sqrt{n}}$. Consequently, the maximal eigenvalue of the form $\frac{n(n-1)}{2}D(e,\dots,e,g,g)$ is equal to $\frac{1}{n}$. Thus condition (15) implies nothing other than the orthogonality of the admissible systems of the values of b_{ik} to the "eigenvector" corresponding to the eigenvalue $\frac{1}{n}$ (i.e., to the system of numbers $b_{ii} = \frac{1}{\sqrt{n}}$ and $b_{ik} = 0$ for $i \neq k$). Since, under this condition, the form $\frac{n(n-1)}{2}D(e,\dots,e,g,g)$ is not positive-definite, we find $\frac{1}{n}$ is the only positive eigenvalue.

Suppose as before that the forms $f_1,\dots,f_{n-2}$ are positive-definite. We now construct the family of *pencils* $(1-\vartheta)e+\vartheta f_1,\dots,(1-\vartheta)e+\vartheta f_{n-2}$. As ϑ increases continuously from 0 to 1, the forms of all pencils remain positive-definite and vary continuously from e to $f_1,\dots,f_{n-2}$. In this process the form $G = D((1-\vartheta)e+\vartheta f_1,\dots,g,g)$ and, consequently, its eigenvalues also vary continuously. By Lemma 1, none of them can pass through zero. Therefore the number of positive eigenvalues of the form G does not change and since, for $\vartheta = 0$, the form $G = D(e,\dots,e,g,g)$ has only one such eigenvalue, the form $G_1 = D(f_1,\dots,f_{n-2},g,g)$ has exactly one positive eigenvalue.

Now we are in a position to prove the theorem formulated at the beginning of this section. Suppose that the form f_{n-2} been reduced to a sum of squares. The form $G = D(f_1,\dots,f_{n-2},g,g)$ has only one positive eigenvalue. Assuming that all $b_{ik} = 0$ for $i \neq k$, we obtain from the form G a new quadratic form in the variables $b_{11}, b_{22},\dots,b_{nn}$. This new form is also denoted by G:

$$G = D(f_1,\dots,f_{n-2},g,g) = \sum_{i,k=1}^{n} D(f_1,\dots,f_{n-2})_{ii,kk}\, b_{ii}b_{kk}. \tag{18}$$

It has no more than one positive eigenvalue because, obviously, the number of positive eigenvalues of the form derived from a given form by equating certain variables to zero is not greater than the number of eigenvalues of the initially given form.

The limitations imposed stipulate that we are assuming f_{n-1} and g to be canonical forms. However, if we prove the theorem under this assumption, we also prove it for the general case because, the form f_{n-1} being positive, all the forms $f_1, \ldots, f_{n-1}, g$ (when g is given) can be transformed such that f_{n-1} and g are reduced to canonical form. Suppose that

$$D(f_1, \ldots, f_{n-1}, g) = 0, \tag{19}$$

and take the quadratic form

$$\mathcal{E} = \sum_{i=1}^{n} \frac{D(f_1, \ldots, f_{n-1})_{ii}}{a_{ii}^{(n-1)}} b_{ii}^2. \tag{20}$$

Since

$$D(f_1, \ldots, f_{n-1})_{ii} > 0, \quad a_{ii}^{(n-1)} > 0,$$

the form $\mathcal{E}$ is positive-definite. Therefore both the forms $\mathcal{E}$ and G can be reduced to canonical form by solving the equations

$$\sum_{k=1}^{n} D(f_1, \ldots, f_{n-2})_{ii,kk}\, b_{kk} = \lambda \frac{b_{ii}}{a_{ii}^{(n-1)}} D(f_1, \ldots, f_{n-1})_{ii}. \tag{21}$$

Since

$$D(f_1, \ldots, f_{n-1})_{ii} = \sum_{k=1}^{n} D(f_1, \ldots, f_{n-2})_{ii,kk}\, a_{kk}^{(n-1)}; \tag{22}$$

therefore, for $\lambda = 1$, the system of equations (21) admits the solution $b_{kk} = a_{kk}^{(n-1)}$. This means $\lambda = 1$ is the only positive eigenvalue of the form G.

Condition (19), which can be rewritten as

$$\sum_{i=1}^{n} \frac{D(f_1, \ldots, f_{n-1})_{ii}}{a_{ii}^{(n-1)}} a_{ii}^{(n-1)} b_{ii} = 0, \tag{23}$$

implies "weighted" orthogonality of the admissible system of the values b_{ii} to the system of numbers $a_{ii}^{(n-1)}$, i.e., to the eigenvector of the form G corresponding to the unique positive eigenvalue $\lambda = 1$ of the form G. Consequently, under condition (19), the form G is not positive-definite. Since it has no zero eigenvalues, it is negative-definite unless g vanishes identically.

Thus we proved the theorem. Its analogy with the result of Part II for mixed volumes is obvious. Here the positive-definite forms play the role

which analogous polyhedra play there. So, in the same way as we did for mixed volumes, from the inequality

$$D(f_1, \ldots, f_n)^2 \geqslant D(f_1, \ldots, f_{n-1}, f_{n-1})\, D(f_1, \ldots, f_{n-2}, f_n, f_n) \tag{24}$$

between mixed discriminants of positive-definite forms we can derive two corollaries.

(1) Suppose that $f_1, \ldots, f_{n-m}, f^{(0)}, f^{(1)}$ are given positive-definite forms in n variables and let $f^{(\vartheta)} = (1-\vartheta)f^{(0)} + \vartheta f^{(1)}$. Put

$$D(\vartheta) = D(f^{(\vartheta)}, \ldots, f^{(\vartheta)}, f_1, \ldots, f_{n-m}). \tag{25}$$

Then the inequality

$$\sqrt[m]{D(\vartheta)} \geqslant (1-\vartheta)\ \sqrt[m]{D(0)} + \vartheta\ \sqrt[m]{D(1)} \tag{26}$$

is valid, and equality holds if and only if the forms $f_{(0)}$ and $f_{(1)}$ are proportional: $f_{(0)} = \lambda f_{(1)}$.

(2) Let $f_1, \ldots, f_n$ be given positive-definite forms in n variables. Then the inequality

$$D(f_1, \ldots, f_n)^m \geqslant \prod_{k=1}^{m} D(f_k, \ldots, f_k, f_{m+1}, \ldots, f_n) \tag{27}$$

is valid, and equality holds if and only if the forms $f_1, \ldots, f_m$ all are proportional to one another.

In both cases m is an arbitrary number between 2 and n. Unlike the corresponding theorems on mixed volumes, the theorems here state the conditions for equality to hold. This is naturally explained by the fact that we know the condition for equality to hold in the initial inequality (24).

§ 4. Relation of Mixed Discriminants with Mixed Area Functions and Mixed Volumes

Let H be a convex body in n-dimensional space having the twice continuously differentiable support function $H(\bar{u})$ with $\bar{x}$ a point on the surface of H and $\bar{n}$ the normal to this surface. Direct the nth axis along $\bar{n}$ and the remaining ones perpendicular to $\bar{n}$. Then the coordinates of the vector $\bar{n}$ are

$$u_1 = \cdots = u_{n-1} = 0, \quad u_n = 1.$$

For $u_n > 0$, since the support function is positively homogeneous, we have

$$\frac{\partial}{\partial u_k} H(0, \ldots, 0, u_n) = \frac{\partial}{\partial u_k} H(0, \ldots, 0, 1), \tag{1}$$

because the first derivatives of $H(\bar{u})$ are homogeneous of degree zero. Therefore

$$\frac{\partial^2 H(0,\ldots,1)}{\partial u_k \partial u_n} = \frac{\partial^2 H^2(\bar{n})}{\partial u_k \partial u_n} = 0. \tag{2}$$

Hence, in our coordinate system, which we call *normal*, the second differential of $H(\bar{u})$ for $\bar{u} = \bar{n}$ is

$$d^2 H(\bar{n}) = \sum_{i,k=1}^{n-1} \frac{\partial^2 H(\bar{n})}{\partial u_k \partial u_i}\, du_i du_k. \tag{3}$$

We know that the eigenvalues of this quadratic form are the principal radii of curvature of the surface H at the point $\bar{x}$ and its principal directions coincide with the principal directions on the surface at the point $\bar{x}$. Therefore reducing (3) to the principal axes, we find

$$d^2 H(\bar{n}) = \sum_{i=1}^{n-1} R_i\, du_i^2, \tag{4}$$

where R_i are the principal radii of curvature. Since H is convex, they all are nonnegative.[5]

In what follows we always use a normal coordinate system. Since the second differential of the support function of a convex body in this system can be represented in the form of (3), we disregard the trivial zero eigenvalue and take that $d^2H(\bar{n})$ is a form in $(n-1)$ variables $du_1, du_2, \ldots, du_{n-1}$.

The discriminant $D(\underbrace{H,\ldots,H}_{n-1};\bar{n})$ of the form $d^2H(\bar{n})$ (in a normal coordinate system) is equal to the product of the principal radii of curvature as is clear from formula (4). Hence, if $d\omega$ is the element of the surface of the unit sphere, then the element of the surface of H is

$$dS = R_1 \ldots R_{n-1}\, d\omega = D(H,\ldots,H,\bar{n})\, d\omega. \tag{5}$$

Hence the volume of the body H is expressed by the formula

$$V(\underbrace{H,\ldots,H}_{n}) = \frac{1}{n} \int\limits_{\Omega} H(\bar{n})\, D(H,\ldots,H,\bar{n})\, d\omega, \tag{6}$$

where Ω is the whole surface of the unit ball.

[5]See, for example, Blaschke, Differentialgeometrie, vol. 1, § 94 or Bonnesen and Fenchel, Theorie der konvexen Körper, § 8.

If the body H is the mixing of the bodies $H_1, H_2, \dots, H_m$, then

$$H(\bar{n}) = \sum_{i=1}^{m} \lambda_i H_i(\bar{n})$$

and

$$d^2 H(\bar{n}) = \sum_{i=1}^{m} \lambda_i \, d^2 \, H_i(\bar{n}).$$

In this case $D(H, \dots, H, \bar{n})$ is a homogeneous polynomial of degree $(n-1)$ in $\lambda_1, \dots, \lambda_m$. So the volume of the body H is a homogeneous polynomial of degree n in $\lambda_1, \dots, \lambda_m$:

$$V(H, \dots, H) \; = \sum_{i_1, \dots, i_n} \lambda_{i_1}, \dots, \lambda_{i_n} \, V(H_{i_1}, \dots, H_{i_n}), \tag{7}$$

where subscripts $i_1, \dots, i_n$ run independently from 1 to m. The coefficients $V(H_{i_1}, \dots, H_{i_n})$ are defined in such a way that they do not depend on the order of occurrence of the of the subscripts $i_1, i_2, \dots, i_n$. They are just the mixed volumes of the bodies $H_1, H_2, \dots, H_m$.

Let H and H_n be two given convex bodies. The volume of the body $H + \lambda H_n$ is

$$V(H + \lambda H_n, \dots, H + \lambda H_n) = \sum_{m=0}^{n} \lambda^m \, C_n^m \, V(\underbrace{H, \dots, H}_{n-m}, \underbrace{H_n, \dots, H_n}_{m}). \tag{8}$$

The body $H + \lambda H_n$ is constructed by displacing the supporting planes of the body H through a distance λH_n in the direction of the outward normal [$H_n(\bar{n})$ is the distance from the origin to the supporting plane of the body H_n whose outward normal is $\bar{n}$]. If λ is infinitely small, then the variation of the volume of the body H is

$$\int\limits_{\Omega} H_n(\bar{n}) \, dS = \int\limits_{\Omega} H_n(\bar{n}) \, D(H, \dots H, \bar{n}) \, d\omega. \tag{9}$$

On the other hand, formula (8) gives that the variation is $nV(H_n, H, \dots, H)$. Hence

$$V(H_n, H, \dots, H) = \frac{1}{n} \int\limits_{\Omega} H_n(\bar{n}) \, D(H, \dots, H, \bar{n}) \, d\omega. \tag{10}$$

If $H = \sum_{i=1}^{n-1} \lambda_i H_i$, then on one hand,[6]

$$V(H_n, H, \dots, H) \; = \sum_{i_1, \dots, i_{n-1}} \lambda_{i_1}, \dots, \lambda_{i_{n-1}} \, V(H_n, H_{i_1}, \dots, H_{i_{n-1}}), \tag{11}$$

[6] To obtain this expansion it suffices to put $H = \sum_{i=1}^{n-1} \lambda_i H_i$ in (8), then expand $V(H + \lambda H_n, \dots, H + \lambda H_n)$ by the general formula (7) and finally compare the coefficients λ.

and, on the other hand,

$$D(H,\dots,H,\bar{n}) = \sum_{i_1,\dots,i_{n-1}} \lambda_i \dots \lambda_{i_{n-1}} D(H_{i_1},\dots,H_{i_{n-1}};\bar{n}), \tag{12}$$

where $D(H_{i_1},\dots,H_{i_{n-1}};\bar{n})$ is the mixed discriminant of the forms

$$d^2H_{i_1}(\bar{n}),\dots,d^2H_{i_{n-1}}(\bar{n}).$$

Therefore, substituting (11) and (12) into (10), and then comparing the coefficients of the product $\lambda_1\lambda_2\dots\lambda_{n-1}$, we find

$$V(H_n,H_1,\dots,H_{n-1}) = \frac{1}{n}\int\limits_{\Omega} H_n(\bar{n})\,D(H_1,\dots,H_{n-1};\bar{n})\,d\omega. \tag{13}$$

However, by definition, mixed volume does not depend on the order of occurrence of the bodies in it and so, for example,

$$V(H_1,\dots,H_n) = V(H_n,H_1,\dots,H_{n-1}).$$

Hence we could substitute H_1 for H_n and put $H = \sum_{i=2}^{n}\lambda_i H_i$. Then we would obtain

$$V(H_1,\dots,H_n) = \frac{1}{n}\int\limits_{\Omega} H_1(\bar{n})\,D(H_2,\dots,H_n;\bar{n})\,d\omega. \tag{14}$$

Now comparing (13) and (14), we find the bodies in the expression for the mixed volume (14) can be transposed arbitrarily.

Let $Z(\bar{u})$ be an arbitrary twice continuously differentiable function of the vector $\bar{u}$ that is positively homogeneous of degree one. Let $d^2Z(\bar{n})$ be its second differential in a normal coordinate system taken for the unit vector $\bar{n}$. Let $D(Z,\dots,Z;\bar{n})$ be the discriminant of $d^2Z(\bar{n})$. Defining the "volume of $Z(\bar{u})$" as

$$\frac{1}{n}\int\limits_{\Omega} Z(\bar{n})\,D(Z,\dots,Z;\bar{n})\,d\omega,$$

and considering a linear combination of such functions

$$\sum \lambda_i Z_i(\bar{u}),$$

we can introduce, in the same way as we did for the support functions of convex bodies, a "mixed volume" $V(Z_1,Z_2,\dots,Z_n)$ and prove that

$$V(Z_1,\dots,Z_n) = \frac{1}{n}\int\limits_{\Omega} Z_1(\bar{n})D(Z_2,\dots,Z_n;\bar{n})d\omega,$$

where $D(Z_2,\dots,Z_n;\bar{n})$ is the mixed discriminant of the second differentials

$$d^2Z_2(\bar{n}),\dots,d^2Z_n(\bar{n})$$

in a normal coordinate system. In formula (15) the functions

$$Z_1(\bar{n}),\dots,Z_n(\bar{n})$$

can be transposed arbitrarily.

§ 5. Uniqueness of a Convex Body with Given Curvature Function

A *curvature function* (in the usual sense) is defined to be an elementary symmetric function of the principal radii of curvature. Let H be a given convex body with twice differentiable support function and let E be the unit ball. From formula (4) of the previous section and from the definition of mixed discriminant it immediately follows that an elementary symmetric function of degree m in the principal radii of curvature, i.e., the mth curvature function, of a convex body H is

$$C_{n-1}^{m}\, D(\underbrace{H,\dots,H}_{m},E,\dots,E;\bar{n}).$$

It is regarded a function of the outward normal $\bar{n}$ of the surface of H.[7]

Theorem. *If two convex bodies with twice continuously differentiable support functions and nowhere vanishing principal radii of curvature have identical curvature functions of a given order, then these two bodies are translates of one another.*

Let H_1 and H_2 be two bodies that satisfy the conditions of the theorem and suppose that they have equal curvature functions of order one, so that

$$D(H_1,E,\dots,E;\bar{n}) = D(H_2,E,\dots,E;\bar{n}). \tag{1}$$

Assume that $H_1(\bar{u})-H_2(\bar{u}) = Z(\bar{u})$, where $H_1(\bar{u})$ and $H_2(\bar{u})$ are the support functions of the bodies H_1 and H_2, respectively. Then $d^2Z(\bar{n}) = d^2H_1(\bar{n}) - d^2H_2(\bar{n})$ and, by a property of mixed discriminants stated in § 1,

$$D(Z,E,\dots E;\bar{n}) = 0. \tag{2}$$

By virtue of the general inequality between mixed discriminants, it follows that

$$D(Z,Z,E,\dots,E,\bar{n}) \leqslant 0, \tag{3}$$

where equality holds if and only if $d^2Z(\bar{n})$ vanishes identically.

Multiplying (2) by $Z(\bar{n})$ and then integrating over the surface of the unit ball, we find

$$\int\limits_{\Omega} E(\bar{n})\, D(Z,Z,E,\dots,E;\bar{n})\, d\omega = 0. \tag{4}$$

[7] See Bonnesen and Fenchel, Theorie der konvexen Körper, § 8.

In the same way, multiplying (3) by $E(\bar{n}) = 1$, i.e., by the support function of the unit ball, we infer

$$\int_{\Omega} E(\bar{n})\, D(Z, Z, E, \dots, E, \bar{n})\, d\omega \leqslant 0 \tag{5}$$

and equality holds if and only if equality holds in (3) for all $\bar{n}$.

However, as demonstrated in the previous section, integrals (4) and (5) are equal. Therefore equality holds in (3) and $d^2 Z(\bar{n})$ vanishes identically for all $\bar{n}$. Since $Z(\bar{n})$ is a positively homogeneous function of degree one, we find it is linear, i.e.,

$$Z(\bar{u}) = \bar{a}\bar{u}, \tag{6}$$

where $\bar{a}$ is a constant vector. And $\bar{a}\bar{u}$ is just the factor which the support function of a convex body acquires under translation through the vector $\bar{a}$. Consequently, the bodies H_1 and H_2 are translates of one another.

Now suppose that the bodies H_1 and H_2 have equal curvature functions of order m, so that

$$D(\underbrace{H_1, \dots, H_1}_{n}, E, \dots, E; \bar{n}) = D(\underbrace{H_2, \dots, H_2}_{m}, E, \dots, E; \bar{n}). \tag{7}$$

Recalling that these expressions are interpreted as mixed discriminants of positive forms (because, by assumption, the principal radii of curvature of the bodies H_1 and H_2 are greater than zero), we can write

$$\begin{aligned} D(H_1, \underbrace{H_2, \dots, H_2}_{m-1}, E, \dots, E; \bar{n})^m \geqslant{}& D(\underbrace{H_1, \dots, H_1}_{m}, E, \dots, E; \bar{n}) \\ &\times D(\underbrace{H_2, \dots, H_2}_{m}, E, \dots, E; \bar{n})^{m-1}, \end{aligned} \tag{8}$$

which is a particular case of the general inequality (27) of § 3. By virtue of (7), from (8) we infer

$$D(H_1, H_2, \dots, H_2, E, \dots, E, \bar{n}) \geqslant D(H_1, \dots, H_1, E, \dots, E, \bar{n}). \tag{9}$$

We can of course assume that the origin is interior to the body H_2 so that $H_2(\bar{n}) > 0$.

Multiplying inequality (9) by $H_2(\bar{n})$ and then integrating over the surface of the unit ball, we obtain, by virtue of formula (14), § 4

$$V(H_1, \underbrace{H_2, \dots, H_2}_{m}) \geqslant V(H_2, H_1, \dots, H_1, E, \dots, E). \tag{10}$$

However, as the bodies H_1 and H_2 play completely identical roles, the opposite inequality must also hold. Consequently, equality must hold here. Hence equality must hold in (8) for all n. Therefore, by virtue of the remark made at the end of § 3, it follows that $d^2H_1(\bar{n})$ and $d^2H_2(\bar{n})$ are proportional and, moreover, as the initial equality (7) is true, then

$$d^2H_1(\bar{n}) = d^2H_2(\bar{n})$$

for all n. Hence it immediately follows that the bodies H_1 and H_2 are translates of one another.

Note that it is especially simple to prove the Minkowski theorem on the uniqueness of a convex body with given curvature by our method as the inequality between mixed discriminants needed in the proof is easily and directly established.

Indeed, let f and g be two positive quadratic forms. They can be simultaneously reduced to canonical forms so that

$$f = \sum_{i=1}^{n-1} a_i\, x_i^2, \qquad g = \sum_{i=1}^{n-1} b_i\, x_i^2,$$

and

$$D(\underbrace{f,\dots,f}_{n-1}) = a_1a_2\dots a_{n-1}, \qquad D(\underbrace{g,\dots,g}_{n-1}) = b_1b_2\dots b_{n-1}. \tag{11}$$

Therefore the inequality

$$D(g,\underbrace{f,\dots,f}_{n-2})^{n-1} \geqslant D(\underbrace{f,\dots,f}_{n-1})^{n-2}\, D(\underbrace{g,\dots,g}_{n-1}) \tag{12}$$

follows directly from the fact that for a given product $\frac{b_1}{a_1}\frac{b_2}{a_2}\dots\frac{b_{n-1}}{a_{n-1}}$ the sum $\sum_{i=1}^{n-1}\frac{b_i}{a_i}$ is maximum when the numbers $\frac{b_i}{a_i}$ are all equal.

If two convex bodies H_1 and H_2 have equal products of the principal radii of curvature at the points where their outward normals are parallel, then

$$D(H_1,\dots,H_1;\bar{n}) = D(H_2,\dots,H_2;\bar{n}), \tag{13}$$

and applying inequality (12), we obtain

$$D(H_2,H_1,\dots,H_1;\bar{n}) \geqslant D(H_2,\dots,H_2;\bar{n}), \tag{14}$$

and, multiplying by $H_1(\bar{n})$ and integrating, we obtain

$$V(H_2,H_1,\dots,H_1) \geqslant V(H_1,H_2,\dots,H_2), \tag{15}$$

after which we repeat the simple argument used before.

Probably, the following fact also deserves our attention: the proof given here for the theorem on the uniqueness of a convex body with given curvature functions coincides formally in the three-dimensional case with the proof given by Blaschke to demonstrate the rigidity of closed convex surfaces. In fact, if H is the support function of a closed convex surface and Z the "support" function of the corresponding "diagram of revolutions," then[8]

$$D(H, Z; \bar{n}) = 0. \tag{16}$$

At the same time, if two three-dimensional convex bodies have equal first curvature functions (sums of the principal radii of curvature), then the difference of their support functions satisfies the equation

$$D(E, Z; \bar{n}) = 0, \tag{17}$$

where E is the unit ball. Therefore the Christoffel theorem coincides formally with the rigidity theorem for spheres.

However, if two three-dimensional convex bodies H_1 and H_2 have equal second curvature functions, *reciprocal Gaussian curvatures*:

$$D(H_1, H_1, \bar{n}) = D(H_2, H_2, \bar{n}), \tag{18}$$

then, putting

$$H_1(\bar{n}) + H_2(\bar{n}) = H(\bar{n}), \quad H_1(\bar{n}) - H_2(\bar{n}) = Z(\bar{n}),$$

we obtain

$$D(H, Z, \bar{n}) = 0, \tag{19}$$

which is just the equation satisfied by the support function of the diagram of revolution for bending of the surface H.

§ 6. Proof of the Main Inequality Between Mixed Volumes by the Hilbert Method

The inequality between mixed volumes, which we call the *main inequality*, is of the form

$$V(H_1, \ldots, H_{n-1}, Z)^2 \geqslant V(H_1, \ldots, H_{n-1}, H_{n-1})\, V(H_1, \ldots, H_{n-2}, Z, Z), \tag{1}$$

[8]See Blaschke, Gott. Nachr., 607–610 (1912); Weyl, Sitz. Ber. Preuss. Akad., 250–266 (1917). The "support" function of the diagram of revolution gives distances from the origin to tangent planes.

where $H_1, \dots, H_{n-1}$ are convex bodies in n-dimensional space, and Z the difference between support functions of convex bodies.[9]

However, it is well-known that every convex body can be arbitrarily well approximated by convex bodies with analytic support functions and nowhere vanishing principal radii of curvature (such a body we call *regular*). Therefore it suffices to prove the main inequality (1) for the case in which the bodies $H_1, \dots, H_{n-1}$ are regular and Z is an analytic function.

In just the same way as we did in deriving the inequality between mixed discriminants (or in deriving the main inequality between mixed volumes of analogous polyhedra in Part II) the assertion of inequality (1) can be reduced to the following

Theorem. *Let $H_1, \dots, H_{n-1}$ be regular convex bodies and $Z(\bar{u})$ a twice continuously differentiable function of the vector $\bar{u}$ which is positively homogeneous of degree one and satisfies the condition*

$$V(H_1, \dots, H_{n-1}, Z) = 0, \tag{2}$$

then

$$V(H_1, \dots, H_{n-2}, Z, Z) \leqslant 0, \tag{3}$$

and is zero if and only if $Z(\bar{u})$ is the support function of a point (i.e., $Z(\bar{u}) = \bar{a}\bar{u}$).

Hilbert's idea lies in reducing the proof of this theorem to solving the extremal value problem for the functional

$$V(H_1, \dots, H_{n-2}, Z, Z) = \frac{1}{n} \int\limits_{\Omega} Z(\bar{n})\, D(H_1, \dots, H_{n-2}, Z; \bar{n})\, d\omega \tag{4}$$

under an additional condition

$$\frac{1}{n} \int\limits_{\Omega} \frac{D(H_1, \dots, H_{n-1}; \bar{n})}{H_{n-1}(\bar{n})}\, Z^2(\bar{n})\, d\omega = 1. \tag{5}$$

Without any loss of generality, we can here assume that the origin lies in the interior of all the bodies $H_1, \dots, H_{n-1}$, so that $H_1(\bar{n}), \dots, H_{n-1}(\bar{n}) > 0$. And since the second differentials

$$d^2 H_1(\bar{n}), \dots, d^2 H_{n-1}(\bar{n})$$

are positive, their mixed discriminant is also positive, so

$$\frac{D(H_1, \dots, H_{n-1}; \bar{n})}{H_{n-1}(\bar{n})} > 0. \tag{6}$$

The nature of the variational problem just formulated is made clear by

[9]This inequality was first proved by Fenchel, C. R., Paris, 203: 647 (1936), and independently by me (see Dokl. Akad. Nauk SSSR, No. 4 (1937)).

Lemma I. *$D(H_1, \ldots, H_{n-2}, Z; \bar{n})$ is a linear self-adjoint differential expression of second order in the function $Z(\bar{n})$ which is elliptic on the unit sphere.*

The self-adjoint nature of the expression is implied in the assertion proved in § 4 that the functions $Z(\bar{n})$ and $Y(\bar{n})$ can be transposed in the expression for the "mixed volume"

$$V(H_1, \ldots, H_{n-2}, Y, Z) = \frac{1}{n} \int\limits_{\Omega} Y(\bar{n})\, D(H_1, \ldots, H_{n-2}, Z, \bar{n})\, d\omega. \tag{7}$$

That $D(H_1, \ldots, H_2, Z, \bar{n})$ is a second order linear differential expression follows directly from its definition as a mixed discriminant of the second differentials $d^2H_1, \ldots, d^2H_{n-2}, d^2Z$. It only remains to prove that it is elliptic. For this, take a normal coordinate system at an arbitrary point n on the unit sphere. Then $(H_1, \ldots, H_{n-2}, Z, \bar{n})$ is the mixed discriminant of the forms

$$\sum_{i,k=1}^{n-1} \frac{\partial^2 H_1(\bar{n}_0)}{\partial u_i \partial u_k}\, du_i du_k \quad , \ldots, \sum_{i,k=1}^{n-1} \frac{\partial^2 Z_1(\bar{n}_0)}{\partial u_i \partial u_k}\, du_i du_k.$$

In § 2 it was shown that the coordinate axes $u_1, u_2, \ldots, u_{n-1}$ can be so chosen that

$$D(H_1, \ldots, H_{n-2}, Z; \bar{n}_0) = \sum D(H_1, \ldots, H_{n-2}; \bar{n}_0)_{ii} \frac{\partial^2 Z(\bar{n}_0)}{\partial u_i^2}, \tag{8}$$

where the expression $D(H_1, \ldots, H_{n-2}; \bar{n}_0)_{ii}$ has the same meaning as the expression $D(f_1, \ldots, f_{n-1})_{ii}$ in § 2. Therefore

$$D(H_1, \ldots, H_2; \bar{n}_0)_{ii} > 0 \tag{9}$$

because the second differentials $d^2H_1(\bar{n}_0), \ldots, d^2H_{n-2}(\bar{n})$ are positive forms.

Put $r = \sqrt{u_1^2 + \cdots + u_n^2}$ and take $\xi_1 = u_1, \ldots, \xi_{n-1} = u_{n-1}$ as the variables on the unit sphere in the neighborhood of the point $(\bar{n}_0)$. At the point $(\bar{n}_0)$, we have $r = 1$, $\xi_1 = \cdots = \xi_{n-1} = 0$. Since $Z(\bar{u})$ is homogeneous,

$$\frac{\partial Z(\bar{u})}{\partial r} = Z(\bar{n}).$$

Furthermore,

$$\frac{\partial \xi_i}{\partial u_i} = 1, \quad \frac{\partial \xi_i}{\partial u_k} = 0 \quad (i \neq k).$$

Therefore

$$\frac{\partial^2 Z(\bar{n}_0)}{\partial u_i^2} = \frac{\partial^2 Z(\bar{n}_0)}{\partial \xi_i^2} + \cdots, \tag{10}$$

where dots stand for the terms unimportant for us and not containing second derivatives. Thus, the coefficients being positive, we find that the expression

$$D(H_1, \ldots, H_{n-2}, Z; \bar{n})$$

is elliptic.

Now that the lemma has been proved, it is clear that our variational problem reduces to eigenvalue and eigenfunction problem for a self-adjoint linear differential equation of second order and of elliptic type

$$D(H_1, \ldots, H_{n-2}, Z; \bar{n}) + \lambda \frac{D(H_1, \ldots, H_{n-1}; \bar{n})}{H_{n-1}(\bar{n})} Z(\bar{n}) = 0. \tag{11}$$

The general theory of these equations predicts that equation (11) has a closed system of eigenfunctions and has only finitely many negative eigenvalues.[10]

Lemma II. *Equation (11) has the eigenvalue $\lambda = 0$ to which belong the eigenfunctions $Z(\bar{n}) = \bar{a}\bar{n}$ which are the support functions of points.*

In fact, suppose that

$$D(H_1, \ldots, H_{n-2}, Z; \bar{n}) = 0, \tag{12}$$

then

$$D(H_1, \ldots, H_{n-3}, Z, Z; \bar{n}) \leqslant 0. \tag{13}$$

Multiplying by $H_{n-2}(\bar{n}) > 0$ (the origin lies in the interior of the body H_{n-2}) and integrating over the surface of the unit ball, we obtain

$$V(H_1, \ldots, H_{n-2}, Z, Z) \leqslant 0. \tag{14}$$

Now multiplying (12) by $Z(\bar{n})$ and integrating over the surface of the unit ball, we find that equality holds in (14). Consequently, equality holds also in (13). Hence, by the theorem proved in § 3, $Z(\bar{u})$, which is positively homogeneous of degree one, is linear.

Lemma III. *Equation (11) has only one negative eigenvalue $\lambda = -1$ with eigenfunction $Z(\bar{n}) = H_{n-1}(\bar{n})$.*

That, for $\lambda = -1$, the function $H_{n-1}(\bar{n})$ does satisfy equation (11) is quite obvious. It now remains to prove that the negative eigenvalue is unique. For this purpose, consider the equation

$$D(E, \ldots, E, Z; \bar{n}) + \lambda Z(\bar{n}) = 0, \tag{15}$$

[10] Hilbert, Grundzuge einer allgemeinen Theorie . . . , Chapter 18. The "parametrix method" developed there can be extended to an equation on the n-dimensional sphere.

which is simply the Laplace equation on the sphere. Therefore its eigenfunctions are spherical functions on the n-dimensional sphere[11]. Hence, as is known, equation (15) has only one negative eigenvalue $\lambda = -1$ with eigenfunction $Z(\bar{n}) = 1$.

Let us construct a family of bodies $H_1^\vartheta = (1-\vartheta)E + \vartheta H_1, \ldots, H_{n-1}^\vartheta = (1-\vartheta)E + \vartheta H_{n-1}$. As ϑ is varies from 0 to 1, equation (15) passes into equation (11), after going through a sequence of equations of the type

$$D(H_1^\vartheta, \ldots, H_{n-1}^\vartheta, Z; \bar{n}) + \lambda \frac{D(H_1^\vartheta, \ldots, H_{n-1}^\vartheta)}{H_{n-1}^\vartheta(\bar{n})} Z(\bar{n}) = 0. \tag{16}$$

In this transformation the eigenvalues vary continuously and since, by virtue of Lemma II, the multiplicity of the eigenvalue $\lambda = 0$ remains always the same (equal to n), the number of negative eigenvalues and their multiplicity do not increase. This proves our lemma.

The theorem formulated above follows directly from Lemmas II and III. Indeed, condition (2)

$$V(H_1, \ldots, H_{n-1}, Z) = 0,$$

which can be rewritten as

$$\frac{1}{n} \int\limits_{\Omega} \frac{D(H_1, \ldots, H_{n-1}; \bar{n})}{H_{n-1}(\bar{n})} \, H_{n-1}(\bar{n}) \, Z(\bar{n}) \, d\omega = 0, \tag{17}$$

expresses the "weighted" orthogonality of the admissible functions $Z(\bar{n})$ to the eigenfunction $H_{n-1}(\bar{n})$ of the unique negative eigenvalue of equation (11). Hence, under condition (2), we obtain condition (3)

$$V(H_1, \ldots, H_{n-1}, Z, Z) \leqslant 0$$

because the eigenvalue of equation (11), taken with opposite signs, are assumed to be the extrema of this functional. In (3) equality holds if and only if $Z(\bar{n})$ is the eigenfunction corresponding to the zero eigenvalue, i.e., if and only if $Z(\bar{n}) = \bar{a}\bar{u}$.

[11] In a normal coordinate system $D(E, \ldots, E, Z; \bar{n}) = \sum_{i=1}^{n-1} \frac{\partial^2 Z(\bar{n})}{\partial u_i^2}$. See Bonnesen and Fenchel, Theorie der konvexen Körper, § 8 or Part II: On the Theory of Mixed Volumes of Convex Bodies, § 7, in this book $D(E, \ldots, E, \bar{n}) = 1, E(\bar{n}) = 1$.

§ 7. Equality in the Generalized Brunn Inequality

This section is closely related to § 7 of Part II of this work.[12] We proved two theorems there: one concerning equality in certain generalized Brunn inequalities and the other on the uniqueness of a convex body with given curvature function which is defined, in our general sense, as a set function on the unit sphere. In these theorems, the ball plays the part of the *fundamental body*, *Eichkörper*. However, here we take an arbitrary regular convex body as the fundamental body and deal with "relative" functions and integrals of curvature. The proofs thus obtained for the generalizations of the two theorems are essentially based on the same arguments as before, so whenever possible, we refer to the arguments given in § 7 of Part II. All considerations relate, as usual, to n-dimensional space. Let $E_1, E_2, \dots, E_{n-2}$ be regular convex bodies.

Theorem I. *If two convex bodies H_0 and H_1 of dimension not less than n satisfy the equality*

$$F(H_0, \dots, H_0, E_1, \dots, E_{n-m}; \omega) = F(H_1, \dots, H_1, E_1, \dots, E_{n-m}; \omega), \quad (1)$$

then they are translates of one another.

Theorem II. *In the inequality*

$$\begin{aligned} &\sqrt[m]{V(H_0, \dots, H_0, E_1, \dots, E_{n-m})} \\ &\geqslant (1-\vartheta)\sqrt[m]{V(H_0, \dots, H_0, E_1, \dots, E_{n-m})} \\ &+ \vartheta \sqrt[m]{V(H_1, \dots, H_1, E_1, \dots, E_{n-m})}, \end{aligned} \quad (2)$$

where H_0 and H_1 are convex bodies of dimension not less than m and $H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1$ $(0 \leqslant \vartheta \leqslant 1)$, equality holds if and only if H_0 and H_1 are homothetic.

The equivalence of these two theorems is proved in just the same way as the equivalence of the two theorems stated in § 7 of Part II. So, as in § 7, here too we have

Lemma I. *If*

$$F(H_0, \dots, H_0, E_1, \dots, E_{n-m}; \omega) = F(H_1, \dots, H_1, E_1, \dots, E_{n-m}; \omega), \quad (3)$$

then for the bodies $H_\vartheta = (1-\vartheta)H_0 + \vartheta H_1$ we have

$$F(H_\vartheta, \dots, H_\vartheta, E_1, \dots, E_{n-m}; \omega) = F(H_1, \dots, H_1, E_1, \dots, E_{n-m}; \omega). \quad (4)$$

[12] Chapter IV: Inequalities between mixed volumes and their applications.

In proving Lemma IV of § 7, Part II (validity of Theorem I for first curvature functions), it was essential that the equation

$$D(E, \ldots, E, Z, \bar{n}) + \lambda Z(\bar{n}) = 0 \tag{5}$$

have a closed system of eigenfunctions and that to zero eigenvalue there correspond eigenfunctions $Z(\bar{n}) = \bar{a}\bar{u}$ which represent translation.

However, under the conditions imposed on the bodies $E_1, \ldots, E_{n-2}$, the equation

$$D(E_1, \ldots, E_{n-2}, Z, \bar{n}) + \lambda\, D(E_1, \ldots, E_{n-2}, E, \bar{n})\, Z(\bar{n}) = 0 \tag{6}$$

also has a closed system of eigenfunctions $Z_1(\bar{n})$ satisfying the *orthogonality condition*

$$\int\limits_{\Omega} D(E_1, \ldots, E_{n-2}, E, \bar{n})\, Z_l(\bar{n})\, Z_k(\bar{n})\, d\omega = 0 \quad (i \neq k). \tag{7}$$

and, furthermore, to the eigenvalue $\lambda = 0$ there correspond eigenfunctions $\bar{a}\bar{n}$ which represent parallel translation (see § 5 and § 6). Hence the arguments used in proving Lemma IV of § 7, Part II are valid and so we can assert

Lemma II. *If*

$$F(H_0, E_1, \ldots, E_{n-2}; \omega) = F(H_1, E_1, \ldots, E_{n-2}; \omega), \tag{8}$$

then the bodies H_0 and H_1 are translates of one another.

It only remains to generalize the inductive reasoning that led us to the proofs of both theorems at the end of § 7, Part II.

Lemma II of § 5, Part II [formula (1)] implies that if the bodies $H_\vartheta = (1 - \vartheta)H_0 + \vartheta H_1$ all have equal mth "relative" curvature functions, the mth "relative" curvature integrals of their projections to the same plane are all equal. Assuming that Theorem II holds for all $(n-1)$-flats, we hence find that the projections of the bodies H_0 and H_1 to the same plane are translates of one another. Recalling Lemma I of § 5, Part II, we therefore find that H_0 and H_1 themselves are translates of one another.

It is possible to take one step forward in generalizing the theorems that were proved in the framework of the relative differential geometry of Minkowski. In place of the area function defined relative to the unit ball, we can consider some area function defined relative to a nondegenerate convex body E_{n-1} having the origin as an interior point. Such a *relative area function* of a body H is defined by the condition

$$V(Z, H, \ldots, H) = \frac{1}{n} \int\limits_{\Omega} \frac{Z(\bar{n})}{E_{n-1}(\bar{n})}\, F_{E_{n-1}(\bar{n})}(H, \ldots, H;\, d\omega), \tag{9}$$

i.e.,

$$F_{E_{n-1}}(H, \ldots, H;\, \omega) = \int\limits_{\Omega} E_{n-1}(\bar{n})\, F(H, \ldots, H;\, d\omega). \tag{10}$$

Then we can define the relative curvature functions

$$F_{E_{n-1}}(H, \ldots, H,\, E_1, \ldots, E_{n-m};\, \omega)$$

and prove that if one of them is given, then the convex body is defined to within a translation. This can be done with the help of the theorems proved above.

Institute of Mathematics and Mechanics,
Leningrad State University

Received July 8, 1937

CHAPTER VII

A GENERAL UNIQUENESS THEOREM FOR CLOSED SURFACES

Communicated by Academician I. M. Vinogradov on March 5, 1938

Doklady Akad. Nauk SSSR, Matematika, vol. 19, No. 4, 233–236 (1938).

Let $H(\bar{u})$ be a positively homogeneous function of degree one in vectors $\bar{u}$ in 3-dimensional space. Let H be the envelope of the family of planes $\bar{u}\bar{x} = H(\bar{u})$. Its parametric equation is $x_1 = \frac{\partial H}{\partial u_1}$, $x_2 = \frac{\partial H}{\partial u_2}$, $x_3 = \frac{\partial H}{\partial u_3}$. The eigenvalues R_1 and R_2 of the second differential $d^2H(\bar{u})$ for unit vectors $\bar{u}$ are the principal radii of curvature of the surface H. [R_1 and R_2 may be zero. We disregard trivial zero eigenvalues of $d^2H(\bar{u})$ [1].]

Theorem 1. *Let $f(R_1, R_2, \bar{n})$, a piecewise analytic function of R_1 and R_2, the points $\bar{n}$ being on the unit sphere, be defined in a domain D such that $R_1 \geqslant R_2$. Let $\frac{\partial f}{\partial R_1}$ and $\frac{\partial f}{\partial R_2}$ have like signs everywhere. If there exists a piecewise analytic function $H(\bar{u})$ with eigenvalues R_1 and R_2 for its second differential belonging to the domain D such that*

$$f(R_1, R_2, \bar{n}) = g(\bar{n}),$$

where $g(\bar{n})$ is a given function of $\bar{n}$, then $H(\bar{u})$ is unique up to a linear term $\bar{a}\bar{u}$, i.e., the corresponding surface H is unique up to a translation.

Studying the convex function $H(\bar{u})$ and appropriate closed convex surfaces, and defining a domain D so that $R_1 \geqslant R_2 > 0$, we can obtain a general uniqueness theorem that includes in itself the theorems of Minkowski and Christoffel concerning the definability of a convex surface when $R_1R_2 = g(\bar{n})$ and $R_1 + R_2 = g(\bar{n})$ are given [1].

Our theorem also implies the definability of a convex surface by the *mean curvature* $\frac{1}{R_1} + \frac{1}{R_2} = g(\bar{n})$, an assertion established earlier only for a constant mean curvature.

Theorem II. *Let $f(R_1, R_2; \bar{n})$ satisfy the conditions of Theorem I. If the function $H(\bar{u})$ varied so that*

$$\delta f(R_1, R_2; \bar{n}) = 0,$$

then $\delta H(\bar{u}) = \bar{a}\bar{u}$, i.e., if $f(R_1, R_2; \bar{n})$ is stationary, the surface H can suffer only an infinitely small translation.

Theorem II is proved along the same lines as Theorem I and is even simpler to prove; so we omit the proof.

Proof of Theorem I. Let $H'(\bar{u})$ and $H''(\bar{n})$ be two functions for which the functions $f(R_1, R_2; \bar{n}) = g(\bar{n})$ are identical. Put $H'(\bar{u}) - H''(\bar{u}) = Z(\bar{u})$. Now we show that $d^2 Z(\bar{u})$ vanishes everywhere and, consequently, $Z(\bar{u}) = \bar{a}\bar{u}$. For the sake of definiteness, take

$$\frac{\partial f}{\partial R_1}, \ \frac{\partial f}{\partial R_2} > 0. \tag{1}$$

1. *For every $\bar{u}$, the quadratic form $d^2 Z(\bar{u})$ admits values of unlike signs or vanishes identically.*

From condition (1), it is clear that for

$$f(R_1', R_2'; \bar{n}) = f(R_1'', R_2''; \bar{n}) \tag{2}$$

the differences $(R_1' - R_1'')$ and $(R_2' - R_2'')$ have either unlike signs or both vanish. However,

$$d^2 Z(\bar{u}) = d^2 H'(\bar{u}) - d^2 H''(\bar{u}). \tag{3}$$

Hence, from a well-known property of quadratic forms (bearing in mind the meaning of R_1 and R_2), we find that our assertion is true.

2. As is clear from the meaning of R_1 and R_2, the expression

$$f(R_1, R_2; \bar{n}) = g(\bar{n}) \tag{4}$$

is a second order partial differential equation for $H(\bar{u})$ on the sphere. *For an arbitrary choice of the parameters u and v, the equation*

$$f(R_1, R_2; \bar{n}) = F(H_{uu}, H_{uv}, H_{vv}, H_u, H_v, H; u, v) = g(u, v) \tag{5}$$

is solvable for H in a neighborhood of every point on the sphere.

Take $r = |\bar{u}|$, and choose the parameters u and v such that

$$H(\bar{u}) = rH(u, v). \tag{6}$$

For unit vectors $\bar{u} = \bar{n}$ and $r = 1$, we have

$$d^2H(\bar{n}) = H_{uu}\,du^2 + 2H_{uv}\,dudv + H_{vv}\,dv^2 + 2(H_u\,du + H_v\,dv)\,dr. \tag{7}$$

When H_{uu} increases, its eigenvalues do not decrease; at least one of them increases. So, let

$$\frac{\partial R_1}{\partial H_{uu}} > 0, \qquad \frac{\partial R_2}{\partial H_{uu}} \geqslant 0, \tag{8}$$

and by virtue of (1) and (8), we have

$$\frac{\partial f}{\partial H_{uu}} = \frac{\partial f}{\partial R_1}\frac{\partial R_1}{\partial H_{uu}} + \frac{\partial f}{\partial R_2}\frac{\partial R_2}{\partial H_{uu}} > 0. \tag{9}$$

Hence, from a well-known implicit function theorem, we find that if equation (5) is satisfied for certain values of $H_{uu}, H_{vv}, \ldots, u$ and v, it can be expressed in the neighborhood of these values as

$$H_{uu} = \Phi(H_{uv}, H_{uv}, H_u, H_v, H, u, v), \tag{10}$$

and this representation is unique.

3. Since, by assumption, H' and H'' are piecewise analytic, the function Z is also piecewise analytic and in even domain $d^2Z = d^2H' - d^2H''$ where Z is analytic, either everywhere, or on certain straight lines or at certain isolated points. This follows from the Weierstrass implicit function theorem.

If d^2Z vanishes on a straight line, then $Z(\bar{u}) = \bar{a}\bar{u}$.

Indeed, let C be such a line. We have $d^2H' = d^2H''$ on this line. Let us add a suitable term of the type $\bar{c}\bar{u}$ to $H''(\bar{u})$ (i.e., move the surface H'') so that $dH' = dH''$ and $H' = H''$ on C. Then, by virtue of what has been proved in Section 2, and taking C to be the line $u = \text{const}$, we find that near C the differentials $H'(u,v)$ and $H''(u,v)$ satisfy equation (10) under the same initial conditions as for the line $u = \text{const}$. Hence, in the neighborhood of C, by the Cauchy theorem, we have

$$H'(u,v) = H''(u,v).$$

If $d^2Z = 0$ in a domain where Z is analytic, then by what has been proved above, this equality can be extended to the neighboring domains across the boundary.

4. What now remains is to rule out the case of $d^2Z = 0$ only at isolated points. For this purpose, consider the envelope Z of the family of planes $\bar{u}\bar{x} = Z(\bar{u})$. The surface Z is bounded and, consequently, has supporting planes in every direction. At the points where $d^2Z(\bar{u}) \neq 0$, there is no supporting plane because the radii of curvature of the envelope at these

points are of unlike signs. A supporting plane of Z can pass only through a point that corresponds to $d^2Z = 0$. We however prove that if a point is isolated, no more than one supporting plane can pass through such a point. And this contradicts the assumption that Z has supporting planes in every direction.

5. Let $\bar{n}_0$ be a point on the unit sphere where $d^2Z(\bar{n}_0) = 0$, and O be the point on the surface Z that corresponds to $\bar{n}_0$. Let S be the part of the surface Z that corresponds to the neighborhood of the point $\bar{n}_0$ not containing other points where $d^2Z(\bar{n}) = 0$. Through O draw a plane E_0 with normal $\bar{n}_0$. Take a line a that is not parallel to any plane whose spherical image is included in $U(\bar{n}_0)$ and that passes through some point on S other than O. Obviously, such a point can be found.

Now translate the line a so that its intersection with E_0 traces a circle C about O.

After crossing the boundary of S, the line a no longer intersects S since S has no tangent parallel to the line a. Were this happen in an arbitrarily small circle C and in an arbitrary neighborhood $U(\bar{n}_0)$ included in the initial neighborhood, the point O, for all such neighborhoods $U(\bar{n}_0)$, would be the projection of the boundary points that correspond to S on E_0 in the direction of the line a. Then the surface S would include a straight segment in the direction of the line a with one endpoint at O. Therefore, S would have, contrary to the assumption, tangent planes parallel to the line a. Consequently, if C is a sufficiently small circle, then the line a always intersects the surface S. Therefore every plane that is parallel to the line a and that passes through the point O intersects the surface S. Since the neighborhood $U(\bar{n}_0)$ can be taken as small as we please, no supporting plane of S other than E_0 can pass through the point O.

We can also demonstrate that, if the curvature of S is everywhere negative, the plane E_0 intersects the surface S except at the point O. So E_0 is not a supporting plane of S.

Institute of Mathematics and Mechanics,
Leningrad State University

Received March 9, 1938

References

[1] Bonnesen T. and Fenchel W., *Theorie der konvexen Körper*; § 8, 1937.
[2] Blaschke W., *Differential Geometry*; § 94 and § 95.

CHAPTER VIII

UNIQUENESS THEOREMS FOR CLOSED SURFACES

Communicated by Academician I. M. Vinogradov on December 11, 1938

Doklady Akad. Nauk SSSR, Matematika, vol. 22, No. 3, 99–102 (1939).

We prove the following

Theorem. *Let a function $Z(u_1, u_2, u_3)$ defined in the space (u_1, u_2, u_3) be analytic and positively homogeneous of degree one [i.e., for $r \geqslant 0$, $Z(ru_1, ru_2, ru_3) = rZ(u_1, u_2, u_3)$]. Assume that its second differential is definite nowhere (i.e., at every point it is either an indefinite form of du_i or vanishes identically). Then $Z(u_1, u_2, u_3)$ is a linear function:*

$$Z(u_1, u_2, u_3) = a_1u_1 + a_2u_2 + a_3u_3.$$

To prove this theorem, we interpret the variables u_1, u_2 and u_3 as the components of a vector $\bar{u}$ in a rectangular coordinate system and study the envelope Z of the family of planes:

$$x_1u_1 + x_2u_2 + x_3u_3 = Z(u_1, u_2, u_3). \tag{1}$$

Since Z is a positively homogeneous function of degree one, we can restrict ourselves to unit vectors $\bar{u}, \dots, \bar{e}$, so that $\sqrt{u_1^2 + u_2^2 + u_3^2} = 1$. Accordingly, we can consider our function on the surface of the unit ball E (a point on E is denoted by the same symbol as for the corresponding unit vector). Clearly, family (1) depends on two parameters and can therefore be expressed in vector form as

$$\bar{n}\bar{x} = Z(\bar{n}). \tag{2}$$

The coordinates of a point on the surface of Z are

$$x_i = \frac{\partial Z(u_1, u_2, u_3)}{\partial u_i} \qquad (i = 1, 2, 3). \tag{3}$$

At a point where d^2Z does not vanish, the surface Z has a definite tangent plane. The eigenvalues of d^2Z are the principal radii of curvature of the surface Z. Therefore at a point where d^2Z is an indefinite form, the surface Z intersects the tangent plane [1].

However, being bounded, the surface Z has supporting planes in every direction. These planes touch the surface Z only at the points corresponding to those (u_1, u_2, u_3) where d^2Z vanishes identically, i.e.,

$$\frac{\partial^2 Z}{\partial u_i \partial u_k} = 0 \qquad (i, k = 1, 2, 3). \tag{4}$$

Let N denote the set of the points on the unit ball E where condition (4) is satisfied. For the set N we have the following lemma.

Lemma 1. *Let N be a set of points on the unit ball E where the analytic functions $f_1(\bar{n}), \dots, f_m(\bar{n})$ all vanish. The set N (1) is either empty, (2) or is the entire surface E, (3) or contains finitely many points, (4) or consists of curves that subdivide the surface E into finitely many domains in the interior of each the functions $f_1(\bar{n}), \dots, f_m(\bar{n})$ all vanish at most at finitely many points.*

This lemma is a simple corollary of the Weierstrass analytic implicit function theorem [2]. In our case N is not empty; otherwise the surface Z would have no supporting plane. However, if N is the whole surface E, the second differential d^2Z vanishes everywhere, i.e., the function $Z(u_1, u_2, u_3)$ is linear. What now remains is to rule out the last two alternatives stipulated in Lemma 1. For this purpose we require

Lemma 2. *If $\bar{n}_0$ where $d^2Z \equiv 0$ is an isolated point, then no other plane than the supporting plane with normal $\bar{n}_0$ passes through the corresponding point $\bar{x}_0$ on the surface Z.*

We now prove this lemma.

Let the set N be finite. Since the surface Z has supporting planes only at the points that correspond to the points $\bar{n}$ in the set N, by Lemma 2, the surface Z can have only finitely many supporting planes. This contradicts the assumption that Z is bounded. Consequently, the point set N is not bounded.

What now remains is the fourth case of Lemma 1. Let G be a domain on E bounded by a curve L belonging to the set N such that $d^2Z \equiv 0$ only at finitely many points in the domain G. Since $d^2Z(\bar{u}) \equiv 0$ on L; therefore,

$$x_1 = \frac{\partial Z}{\partial u_1} = a_1, \qquad x_2 = \frac{\partial Z}{\partial u_2} = a_2, \qquad x_3 = \frac{\partial Z}{\partial u_3} = a_3 \tag{5}$$

are constants on L. And since the function $Z(\bar{u})$ is homogeneous, we have

$$Z(\bar{u}) = u_1 \frac{\partial Z}{\partial u_1} + u_2 \frac{\partial Z}{\partial u_2} + u_2 \frac{\partial Z}{\partial u_3}. \tag{6}$$

Consequently,

$$Z(\bar{u}) = a_1 u_1 + a_2 u_2 + a_3 u_3 \tag{7}$$

on L.

Let $Z^*(u_1, u_2, u_3)$ be a positively homogeneous function of degree one that is equal to $Z(u_1, u_2, u_3)$ in the domain G and is equal to $a_1 u_1 + a_2 u_2 + a_3 u_3$ outside the domain G. The function Z^* is twice continuously differentiable because $d^2 Z(\bar{u}) \equiv 0$ on the boundary of the domain G.

Let us now construct the surface Z^* which is the envelope of the planes

$$x_1 u_1 + x_2 u_3 + x_3 u_3 = Z^*(u_1, u_2, u_3). \tag{8}$$

The surface Z^* has supporting planes in all directions. However, at a point $\bar{x}$ on Z corresponding to a point $\bar{n}$ in the domain G, by Lemma 2, there can only be finitely many supporting planes.

Therefore all the supporting planes of Z^* pass through only one point (a_1, a_2, a_3) corresponding to those u_1, u_2, u_3 for which

$$Z^*(u_1, u_2, u_3) = a_1 u_1 + a_2 u_3 + a_3 u_3. \tag{9}$$

[By formula (3) or (5), we find that a_1, a_2 and a_3 are precisely the coordinates of the point on Z^* that corresponds to (u_1, u_2, u_3) for which (9) holds.]

However, in this case the surface Z^* is reduced to a point (a_1, a_2, a_3) and the domain G vanishes because to every point $\bar{n}$ on Z^* where $d^2 Z^* \not\equiv 0$ there should correspond on Z a point where there is a tangent plane. This completes the proof of our theorem.

We now prove Lemma 2. Let $\bar{n}_0$ be an isolated point at which $d^2 Z(\bar{n}_0) \equiv 0$. Let $\bar{x}_0$ be the corresponding point on the surface Z.

Contrary to the assertion of the lemma, assume that at the point $\bar{x}_0$ there is a supporting plane P_1 with normal $\bar{n}_1 \neq \bar{n}_0$. Let U be a neighborhood of $\bar{n}_0$ which does not contain, either in its interior or on its boundary, other points where $d^2 Z(\bar{n}) \equiv 0$.

Let U be small so that it does not intersect any great circle passing through $\bar{n}_1$. Let V be the corresponding neighborhood of $\bar{x}_0$ on Z. A tangent plane exists at all points of V, except at $\bar{x}_0$. On the plane P_1 the neighborhood V has only one point $\bar{x}_0$, otherwise P_1 would not be a supporting plane of V.

Therefore, there exist planes P parallel to P_1 and intersecting Z along a closed curve L. All these curves L are smooth and have supporting lines in every direction, i.e., normal in every direction. The normal to L is the projection of the normal to Z at the same point on the plane P. Therefore, since L has a normal in every direction, Z contains the points with spherical image on every great circle passing through $\bar{n}_1$. However, the spherical image of V is the neighborhood U we chose for the point $\bar{n}_0$. Thus we arrive at a contradiction on the choice of U. This completes the proof of our lemma.

The proven theorem can be reformulated as follows:

Let H_1 and H_2 be two convex bodies with analytic support functions. If the Dupin indicatrices of H_1 and H_2 at the points where the normals are parallel cannot be included one within the other on bringing the points together by translation, then the bodies H_1 and H_2 are translates of one another.

It is clear that the theorem proved in this note is an analog of the theorem on polyhedra which I proved earlier in [3], namely,

If in two convex polyhedra no face of one can be embedded by a translation into a face of the other with the same outward normal, then the polyhedra are translates of one another.

(To every face of one polyhedron there always corresponds a face of the other with the same normal as that of the former; this face may however degenerate into an edge or vertex.)

From the theorem proved above we can easily derive a general uniqueness theorem for closed surfaces which I proved recently in [4]. Strictly speaking, the theorem stated below does not agree in full with the theorem proved in [4]: here we assume that the support function is analytic whereas in [4] it is piecewise analytic. Here the function $f(R_1, R_2; n)$ can be arbitrary.

Theorem. *Let $f(R_1, R_2; \bar{n})$ be a function, where $\bar{n}$ is a point on the unit sphere, and R_1 and R_2 vary in the domain $R_1 \geqslant R_2$. For every given $\bar{n}$, let the function f be strictly monotone in every argument. A surface H, whose spherical image uniquely covers the entire sphere once, with an analytic support function, is uniquely defined by the function $f(R_1, R_2; \bar{n})$ up to a translation for all normals $\bar{n}$, provided R_1 and R_2 are interpreted as the principal radii of curvature of H at the point where $\bar{n}$ is the normal.*

Let $H'(\bar{u})$ and $H''(\bar{u})$ be the support functions of two surfaces for which

$$f(R_1', R_2'; \bar{n}) = f(R_1'', R_2''; \bar{n}). \tag{10}$$

Since the function $f(R_1, R_2; \bar{n})$ is strictly monotone, we find that $R_1' - R_1''$ and $R_2' - R_2''$ are of unlike signs or both vanish. Therefore $d^2[H'(\bar{u}) - H''(\bar{u})]$

is either an indefinite form or vanishes identically. Consequently, by the theorem we proved earlier, the difference $H'(\bar{u}) - H''(\bar{u})$ is a linear function; i. e., the surfaces H' and H'' are translates of one another.

Steklov Institute of Mathematics
The USSR Academy of Sciences

Received December 13, 1938

References

[1] Blaschke W., Differential Geometry § **94**.
[2] Goursat E. J.-B., *A Course of Analysis*; vol. 2, Part II, Chapter XVII.
[3] Alexandrov A. D., Izv. Akad. Nauk SSSR Ser. Mat., **4** (1937), 597.
[4] ———, Dokl. Akad. Nauk, **19** (1938), no. 4, 233.

CHAPTER IX

ON THE AREA FUNCTION OF A CONVEX BODY

A remark on the paper

To the Theory of Mixed Volumes of Convex Bodies[1]

In my paper *To the theory of mixed volumes of convex bodies*,[2] I introduced the concept of the area function $F(H,\omega)$ of a convex body H. By definition, $F(H,\omega)$ is the area of the set of points on the surface of H through which the supporting planes pass whose outward normals fall into the set ω on the unit sphere Ω, if they are drawn from the center of the sphere. I also proved there that $F(H,\omega)$ is a countably additive function of the set ω and that the mixed volume of the body H and a convex body L can be expressed by a Lebesgue-Stieltjes integral taken over the surface Ω of the ball as

$$V(L,\underbrace{H,\dots,H}_{n-1}) = \frac{1}{n}\int\limits_{\Omega} L(\bar{n})F(H,d\omega). \tag{1}$$

Here and in what follows, n always denotes the dimension of space, $\bar{n}$ a unit normal or (if necessary) a point on the sphere Ω and $L(\bar{n})$ the support function of the body L. The concept of mixed area function and its applications in the theory of mixed volumes are largely based on the proposition expressed by formula (1).

However, as already noted in the paper where these concepts were introduced, I proved this proposition by means of some arguments that, although trivial, are very cumbersome in calculation. Unfortunately, an omission

[1] *Matematicheskiĭ Sbornik*, **6(38)**, No. 1, 167–173 (1939).

[2] This paper consists of four parts published serially in *Matematicheskiĭ Sbornik*, **2(44):** 5–6 (1937) and **3(45):** 1–2 (1938). In the sequel they are referred to as Parts I, II, III and IV, respectively. Here we are dealing with Part I with which this remark is closely related.

crept into the conclusion. One more evaluation must have been given under Lemma I in Part I. True, the omission is so obvious that it can easily be guessed. Furthermore, soon after my work, Fenchel and Jessen published their work,[3] formulating the same result in reverse order. They define the area function as the countably additive set function through which the mixed volume is expressed by formula (1). In the conclusion they prove that the area function thus defined gives the area (in the Minkowski sense) of the set on a body H having the spherical image ω. As a consequence of the foregoing, the arguments refined in this paper gain certain weight. Although the results given here were already obtained by me (and by Fenchel and Jessen as well), the new route by which they are derived here are undoubtedly simpler, elegant and more suitable.

Let H be a convex body with interior points and O an interior point of H, $\bar{r}$ a variable unit vector starting from O. The distance between O and a point on the surface Σ of H is the function $R(\bar{r})$ of $\bar{r}$.

If τ is a measurable set of vectors $\bar{r}$ (or of points on the unit sphere T), then the area $F(\sigma)$ of the corresponding set σ on Σ is expressed as usual by the formula[4]

$$F(\sigma) = \int\limits_{\tau} \frac{R(\bar{r})^{n-1}}{\cos\varphi(\bar{r})} F(d\tau), \tag{2}$$

where $\varphi(\bar{r})$ is the angle formed by the vector $\bar{r}$ with the normal $\bar{n}$ at the corresponding point on Σ (if the point is singular, we take the normal to one of the supporting planes passing through this point) and $F(\tau)$ is the area of the set τ. The points on Σ where there is more than one normal (more than one supporting plane) form a set of zero measure.[5] Since the measure of the set of vectors $\bar{r}$ for which $\varphi(\bar{r})$ is discontinuous (or is not single-valued) is zero, integral (2) can be interpreted as an ordinary *Riemann integral.*

Lemma 1. *Let $f(\bar{n})$ be a continuous function on the sphere Ω and H a body containing a point O in its interior. Then*

$$\int\limits_{\Omega} f(\bar{n}) F(H, d\omega) = \int\limits_{\Sigma} f(\bar{n}(\bar{r}))\, F(d\sigma) = \int\limits_{T} f(\bar{n}(\bar{r}))\, \frac{R(\bar{r})^{n-1}}{\cos\varphi(\bar{r})} F(d\tau), \tag{3}$$

where $f(\bar{n}(\bar{r}))$ is discontinuous for those $\bar{r}$ to which correspond points on Σ where there is more than one normal. These relations are obvious corollaries of the definitions of the integrals contained in them.

[3]Fenchel and Jessen, Mengenfunktionen und konvexen Körper, Danske Vid. Selsk., Math.-Fys. Medd., **16**, No. 3: 1–31 (1938).

[4]See Part I, § 1.

[5]See Rademacher, Über die singulären Randpunkte der konvexen Körper, Math. Ann., **83:** 116–118 (1921).

Let $\varepsilon > 0$. Subdivide the interval between $\min f(\bar{n})$ and $\max f(\bar{n})$ into less than $\frac{\varepsilon}{F(H,\Sigma)}$ segments, where $F(H,\Sigma) = F(\Sigma)$ is the area of the surface Σ of the body H. Let f_i be some point in the ith segment and σ_i the set on which $f(\bar{n})$ takes the values lying in the ith segment. Then, by the definition of the Lebesgue-Radon integral, we have

$$\left| \int_{\Sigma} f(\bar{n}) F(H, d\sigma) - \sum f_i F(H, \sigma_i) \right| < \varepsilon. \tag{4}$$

To each set ω_i there corresponds a set

$$\sigma_i = \sigma(\omega_i)$$

of points on Σ where there are normals directed into ω_i. The intersection of every pair of these sets has measure zero because the set of all points on Σ where there is more than one normal has measure zero.

Hence, by the definition of Lebesgue integral, we have

$$\left| \int_{\Sigma} f(\bar{n}(\bar{r})) F(d\sigma) - \sum f_i F(\sigma_i) \right| < \varepsilon. \tag{5}$$

However, by the definition of surface integral, we have

$$F(\sigma_i) = F(H, \omega_i). \tag{6}$$

Substituting this expression into (6), and then comparing it with (4), we find

$$\left| \int_{\Omega} f(\bar{n}) F(H, d\omega) - \int_{\Sigma} f(\bar{n}(\bar{r})) F(d\sigma) \right| < 2\varepsilon. \tag{7}$$

Since ε is arbitrary, we obtain the first expression in (3). Using (2), we obtain the second expression in (3).

Lemma II. *If the convex bodies $H_1, H_2, \ldots$ converge to a body H, then the corresponding area functions $F(H_k, \omega)$ converge weakly to the area function $F(H, \omega)$, i.e., for every continuous $f(\bar{n})$*

$$\lim_{k\to\infty} \int_{\Omega} f(\bar{n}) F(H_k, d\omega) = \int_{\Omega} f(\bar{n}) F(H, d\omega). \tag{8}$$

We distinguish two cases: Case 1, the body H has interior points, and Case 2, the body H has no interior points.

Case 1. Suppose that the body H has interior points and let O be an interior point of H. Obviously, we can also assume that O is an interior point of each of the bodies $H_1, H_2, ..., H_k, \dots$. (In case there are bodies that do not contain the point O in their interiors, we may discard a few terms at the head of the sequence.) Then using (3), we find that (4) is equivalent to

$$\lim_{k\to\infty} \int\limits_T f(\bar{n}_k(\bar{r}))\, \frac{R_k(\bar{r})^{n-1}}{\cos\varphi(\bar{r})} F(d\tau) = \int\limits_T f(\bar{n}(\bar{r}))\, \frac{R(\bar{r})^{n-1}}{\cos\varphi(\bar{r})} F(d\tau). \tag{9}$$

However, if the bodies H_k converge to H, then the functions $R_k(\bar{r})$ converge to $R(\bar{r})$ for all $\bar{r}$ and the normals $\bar{n}_k(\bar{r})$ converge to $\bar{n}(\bar{r})$ almost everywhere[6].

Since $f(\bar{n})$ is continuous, the functions $f[\bar{n}_k(\bar{r})]$ also converge to $f[\bar{n}(\bar{r})]$ almost everywhere. The same is true of $\cos\varphi_k(\bar{r})$; moreover, $\cos\varphi(\bar{r})^{-1}$ remains bounded.

By virtue of the Lebesgue dominated convergence theorem, integrals (9) converge. This completes the proof of formula (9).

Case 2. Let the body H have no interior points. If the area of its surface is zero, the limit of the areas of the surface of the bodies H_k is also zero. Therefore both the integral on the right side and the limit of integrals on the left side of (8) vanish.

However, if H is a plane domain, on mixing it with a segment l of length h perpendicular to H, we obtain a cylinder H^h. Mixing the bodies H_k with this segment, we obtain the bodies H_k^h which converge to H^h. For these bodies formula (8) has already been proved.

In this construction, the lateral surface of a cylinder formed by the generator l is added to the surfaces of the bodies H_k and H. If the length h of the segment l is sufficiently small, the lateral surface is arbitrarily small. Therefore, for sufficiently large k and sufficiently small h, we have, on one hand

$$\left| \int\limits_\Omega f(\bar{n}) F(H_k^h, d\omega) - \int\limits_\Omega f(\bar{n}) F(H^h, d\omega) \right| < \varepsilon, \tag{10}$$

and, on the other hand,

$$\left| \int\limits_\Omega f(\bar{n}) F(H_k^h, d\omega) - \int\limits_\Omega f(\bar{n}) F(H_k, d\omega) \right| < \varepsilon \tag{11}$$

[6]$\lim_{k\to\infty} \bar{n}_k(\bar{r}) = \bar{n}(\bar{r})$ for all $\bar{r}$ for which $\bar{n}(\bar{r})$ is single-valued. This is a direct corollary of the lemma: *If $H_k \to H$ and the points x_k on H_k converge to the point $\bar{x}$ on H, then every convergent subsequence of the normals to H_k at the points x_k converges to the normal to H at the point $\bar{x}$* (Part I, § 1, Lemma 1).

and

$$\left| \int_{\Omega} f(\bar{n}) F(H^h, d\omega) - \int_{\Omega} f(\bar{n}) F(H, d\omega) \right| < \varepsilon. \tag{12}$$

Hence

$$\left| \int_{\Omega} f(\bar{n}) F(H_k, d\omega) - \int_{\Omega} f(\bar{n}) F(H, d\omega) \right| < 3\varepsilon, \tag{13}$$

and, ε being arbitrary, we obtain formula (4).

Now we prove formula (1):

$$V(L, H, \ldots, H) = \frac{1}{n} \int_{\Omega} L(\bar{n}) F(H, d\omega).$$

Indeed, if H_k are polyhedra that converge to H, and F_{k_i} the areas of their faces with the normals $\bar{n}_{k_i}$, we have

$$V(L, H_k, ..., H_k) = \frac{1}{n} \sum_{i} L(\bar{n}_{k_i}) F_{k_i}, \tag{14}$$

which is a simple matter to prove. Expressing this sum as a Lebesgue-Stieltjes integral, we obtain

$$V(L, H_k, \ldots, H_k) \to \frac{1}{n} \int_{\Omega} L(\bar{n}) F(H_k, d\omega). \tag{15}$$

If the bodies H_k converge to H, we have

$$\lim_{k \to \infty} V(L, H_k, \ldots, H_k) = V(L, H, ..., H). \tag{16}$$

On the other hand, by Lemma II, the integrals in (15) converge to the integral in (1). This completes the proof of formula (1).

Now, by virtue of Lemma II, it is a simple matter to derive mixed area functions. Indeed, for the polyhedra $H_1, H_2, \ldots, H_{n-1}$, the area of a face of the polyhedron

$$H = \lambda_1 H_1 + \lambda_2 H_2 + \cdots + \lambda_{n-1} H_{n-1}, \tag{17}$$

by the well-known Minkowski theorem, is

$$F = \sum_{k_1, \ldots, k_{n-1}} \lambda_{k_1} \lambda_{k_2} \ldots \lambda_{k_{n-1}} F(H_{k_1}, H_{k_2}, \ldots, H_{k_{n-1}}). \tag{18}$$

In terms of set functions, we find that for every continuous $f(\bar{n})$:

$$\int\limits_{\Omega} f(\bar{n})F(H, d\omega) = \sum_{k_1,\ldots,k_{n-1}} \lambda_{k_1}\lambda_{k_2}\ldots\lambda_{k_{n-1}} \int\limits_{\Omega} f(\bar{n})F(H_{k_1}, H_{k_2}, \ldots, H_{k_{n-1}}; d\omega). \tag{19}$$

If the polyhedra converge to given convex bodies, then, by virtue of Lemma II, we find that formula (15) holds even in the limit. Therefore, if $H_1, \ldots, H_{n-1}$ are arbitrary convex bodies and $H = \lambda_1 H_1 + \cdots + \lambda_{n-1}H_{n-1}$, then

$$F(H, \omega) = \sum_{k_1,\ldots,k_{n-1}} \lambda_{k_1}\lambda_{k_2}\ldots\lambda_{k_{n-1}} F(H_{k_1}, \ldots, H_{k_{n-1}}; \omega). \tag{20}$$

The set functions on the right side of this expression are just the mixed area functions of our convex bodies.

On substituting the support function $H_n(n)$ of the convex body H_n for the function $f(\bar{n})$, we find, by virtue of (1), the left side of (19) is the mixed volume $V(H_n, H, \ldots, H)$ taken n times. Representing it in terms of the mixed volumes of the bodies $H_1, H_2, \ldots, H_{n-1}, H_n$, we obtain a polynomial in $\lambda_1, \ldots, \lambda_{n-1}$. Comparing its coefficients with the right side of (19), we obtain an expression for the general mixed volume

$$V(H_1, \ldots, H_n) = \frac{1}{n}\int\limits_{\Omega} H_n(\bar{n})F(H_1, \ldots, H_{n-1}; d\omega). \tag{21}$$

Now it is easy to derive the properties of a mixed area function listed under the theorem stated in Part I, § 5.

If the bodies $H_1, \ldots, H_{n-1}$, tend, on being modified, to certain limit bodies, then for all $\lambda_1, \ldots, \lambda_{n-1}$, the area function of the body

$$H = \lambda_1 H_1 + \cdots + \lambda_{n-1}H_{n-1}$$

(by Lemma II) converges weakly to the area function of the corresponding limit body. Hence, *the mixed area functions of the bodies $H_1, \ldots, H_{n-1}$ converge weakly to the corresponding mixed area functions of the limit bodies.*

Finally, note that Lemma II is useful in proving the existence of a convex body with given area function.[7]

Let $F(\omega)$ be a given nonnegative countably additive set function such that

$$\int\limits_{\Omega} \bar{n}F(d\omega) = 0, \tag{22}$$

[7]See Part III as well as Fenchel and Jessen, Mengenfunktionen und konvexen Körper, Danske Vid. Selsk., Math.-Fys. Medd., **16,** No. 3: 1–31 (1938).

and that for every $\bar{n}_0$

$$\int\limits_{\Omega} |\bar{n}_0\bar{n}| F(d\omega) > \alpha > 0. \tag{23}$$

Now we have to prove there exists a convex body having $F(\omega)$ as its area function.

Following the Minkowski method, let us take a sequence Ω of finer subdivisions of sets ω_{ij}, where i is the number of a subdivision and j the number of the set in a given subdivision. Let the center of gravity of the load defined by the function $F(\omega)$ on the set ω_{ij} be located at the point $r_{ij}\bar{n}_{ij}$. For each i, let us replace the function $F(\omega)$ by the function $F_i(\omega)$ having at the point $\bar{n}_{ij}$ the point load

$$F_{ij} = r_{ij} F(\omega_{ij}). \tag{24}$$

Clearly, the functions $F_i(\omega)$ converge weakly to $F(\omega)$. Using some conventional method, it is a simple matter to show that there exist, for sufficiently large i, polyhedra having face areas F_{ij} and normals to faces $\bar{n}_{ij}$, i.e., polyhedra with the area functions $F_i(\omega)$.

Furthermore, using some conventional method, we can show that these polyhedra all are uniformly bounded.[8] Extract a convergent subsequence of polyhedra from them and let H be the limit body of this subsequence. By Lemma II the area functions $F_i(\omega)$ of the polyhedra in the subsequence converge weakly to the area function $F(H, \omega)$ of H. Moreover, they also converge weakly to $F(\omega)$. Consequently,

$$F(H, \omega) = F(\omega). \tag{25}$$

This completes the proof.

In conclusion it is worth mentioning certain properties of the weak convergence of nonnegative countably additive set functions on the sphere that elucidate the geometric meaning of this abstract concept.

(1) *For $F_i(\omega)$ to converge weakly to $F(\omega)$, it is necessary and sufficient that, for every closed (or open) ω_0 on whose boundary $F(\omega)$ vanishes,*

$$\lim_{i=\infty} F_i(\omega_0) = F(\omega_0). \tag{26}$$

(2) *For $F_i(\omega)$ to converge weakly to $F(\omega)$, it is necessary and sufficient that for every closed ω there exist a decreasing sequence of sets ω_i*

[8] See Minkowski, Volumen und Oberfläche, Ges. Abh., Bd. II; Bonnesen and Fenchel, *Theorie der konvexen Körper;* Fenchel and Jessen, *Mengenfunktionen und Körper,* Danske Vid. Selsk., Math.-Fys. Medd., **16,** No. 3: 1–3 (1938).

containing ω *and converging to* ω *(i.e., a sequence of sets such that* $\omega_1 \supset \omega_2 \supset \omega_3 \supset \ldots,: \prod_{i=1}^{\infty} \omega_i = \omega$*) for which*

$$\lim_{i=\infty} F_i(\omega_i) = F(\omega). \tag{27}$$

The sets ω_i, while converging to ω, act as if they build up into the set ω a load it would ultimately experience in the limit. For an open set σ this assertion can be formulated as usual with the help of complements. I intend to prove these properties of weak convergence in a more general form in a separate note.

Steklov Mathematics Institute,
The USSR Academy of Sciences

Received February 20, 1939

CHAPTER X

INTRINSIC GEOMETRY OF AN ARBITRARY CONVEX SURFACE

(*Communicated by Academician I. M. Vinogradov*)

Doklady Akad. Nauk SSSR, Matematika, vol. 32, No. 7, 467–470 (1941).

1. A *convex surface* is defined to be a domain[1] of a bounded or an unbounded convex body, the latter being a closed convex set containing interior points. For all points x and y on a convex surface F, there exists a greatest lower bound for the lengths of the curves lying on F that join the points x and y. It is natural to take this infimum as the distance $\rho_F(x, y)$ between the points x and y on the surface F. In this case the surface F becomes a metric space. The objective of *intrinsic geometry* of a surface lies in studying this metric space as such, i.e., with regard for the fact that this space is somehow embedded in three-dimensional Euclidean space.

If a surface F has boundary, then the distance on it depends on the boundary. If a piece F_1 is cut out from the surface F, then the distance $\rho_{F_1}(x, y)$ between two points x and y belonging to the fragment F_1 in general becomes unequal to the distance $\rho_F(x, y)$. Since a piece can be cut out in an arbitrary manner, an indefinite number of intrinsic metrics arise. However, if the points x and y are sufficiently close to each other, then $\rho_{F_1}(x, y) = \rho_F(x, y)$.

Therefore the local properties of the metric on a convex surface do not depend on the boundary of the surface. Therefore, if we are interested only in the local properties, we can study a surface without boundary, i.e., the whole surface of a convex body. Studying a surface without boundary, we can fully characterize the *intrinsic metric* of the surface and thus completely characterize the local properties of the intrinsic metric of an arbitrary convex surface.

[1] A domain is a connected (relatively) open set.—Eds.

2. Convex surfaces presenting the (entire) boundaries of convex bodies may be of three topological types: a surface homeomorphic to the sphere, a surface homeomorphic to the plane, and a surface homeomorphic to some infinite circular cylinder. A surface of the last kind is itself a *cylinder* and can therefore be bent into a circular cylinder without altering the intrinsic metric. The question concerning the intrinsic metric of these surfaces is thus a resolved topic and can therefore be omitted from consideration.

We also classify plane closed convex surfaces as being of two types: bounded domains are of the first type and unbounded domains (except the plane and strips between two parallel lines) are of the second type. For each domain of these types we recognize two sides, and the distance between two points lying on *different sides of a domain* is reckoned along a curve passing from one side to the other through the boundary of the domain. The points that coincide but lie on opposite sides of a domain are regarded distinct.[2] This distinction between the opposite sides of a plane domain can of course be formulated in precise terms as follows.

Let F be a bounded domain and let S be a sphere subdivided into two hemispheres: the boundary of each hemisphere is also included in the hemisphere. We now topologically map each hemisphere onto F so that the respective mappings agree on the common equator, with the equator being mapped onto the boundary of F. The plane can also be mapped onto an unbounded convex domain in this manner. Basing on this construction, we consider a bounded domain as a domain homeomorphic to the sphere and an unbounded domain as a domain homeomorphic to the plane and call the two-sheeted mapping a *homeomorphism*.

3. The question that we intend to resolve reads: What are the necessary and sufficient conditions which a metric $\rho(x, y)$ defined on the sphere or on the plane should satisfy so that an extended convex surface realizing this metric may exist? A metric $\rho(x, y)$ is said to be defined on a manifold R, if the metric defines a proper topology on the manifold R. A surface F is said to *realize a metric* $\rho(x, y)$ given on a manifold R, if there exists a homeomorphism h that sends the manifold R onto the surface F such that to each pair of points $x, y \in R$, we have $\rho(x, y) = \rho_F(x, y)$, where ρ_F is the distance on F as defined above.

I resolved this problem for a convex polyhedron and a convex surface with a metric given by a quadratic form ds^2 as is usually done in differential geometry [1]. Here we solve this problem and certain other related questions without giving the proof in detail. The problem is presented in detail in a separate note.

4. We now take up certain definitions. The *length of a curve* $x(t)$ $(0 \leqslant t \leqslant 1)$ in a metrized manifold R is defined to be the least upper bound

[2]The author describes here a doubly covered plane domain.—Eds.

of the sum of the distances between the points $x(t_i)$ and $x(t_{i+1})$, where $0 = t_0 < t_1 < \cdots < t_n = 1$. A curve $x(t)$ $(0 \leqslant t \leqslant 1)$ is called a *geodesic*, if for every t there exists a connected neighborhood of the interval [0,1] such that the corresponding part of the curve $x(t)$ is the shortest distance between the endpoints of this part. A *geodesic triangle* is a closed domain homeomorphic to the disk and bounded by three geodesics. A triangle is called *normal*, if each side is a *shortest arc*, i.e., there is no other curve joining the two vertices shorter than the corresponding side. A *midline* corresponding to a side of triangle is the shortest arc joining the midpoints of other two sides.

5. On a two-dimensional manifold R let there be given a metric $\rho(x, y)$ satisfying, in addition to the usual three conditions, also the following conditions:

(1) For every $x \in R$ and for every $\varepsilon > 0$, there is a neighborhood $U(x, r)$ of radius r such that there exists a homeomorphism h that maps $U(x, r)$ into a plane or into a convex cone[3] so that for every pair $y, z \in U(x, r)$

$$|\rho(y, z) - \rho_0(h(y), h(z))| < \varepsilon r,$$

where ρ_0 is the distance on the plane or on the cone.

(2) For every pair $x, y \in R$, there exists a point z such that $\rho(x, z) = \rho(z, y) = \frac{1}{2}\rho(x, y)$. If $y_1 \neq y_2$ and z_1, z_2 are such that $\rho(x, z_i) = \rho(z_i, y_i) = \frac{1}{2}\rho(x, y_i)$ $(i = 1, 2)$, then $z_1 \neq z_2$.

(3) For every point $x \in R$, there exists a neighborhood $U(x)$ such that for every normal geodesic triangle in $U(x)$, the length of each midline of the triangle is not less than half the length of the corresponding side.

(4) By virtue of the properties of the metric $\rho(x, y)$, the manifold R is a complete metric space.

If the first three conditions are satisfied, then the metric is said to be *convex*. If, in addition, the fourth condition is also satisfied, then the metric is called *complete*. A metric is called *strictly convex*, if at least one of the triangles in condition (3) has a midline greater than half the corresponding side. If this not true of any triangle in condition (3), then the metric is said to be *not strictly convex*. Obviously, condition (2), when satisfied along with the condition of completeness, is equivalent to the existence of a shortest arc between two points of length equal to the distance between these points.

6. We now formulate our main results.

Theorem 1. *Every complete convex surface has metric convex and complete.*

Here the surface F is regarded a manifold on which the metric ρ_F is defined.

[3]The point x is sent to the vertex of the cone.

Theorem 2. *For every convex metric defined on the sphere,*[4] *there exists a convex surface realizing this metric.*

Theorem 3. *For every complete convex metric given on the plane, there exists a convex surface realizing this metric.*

In these theorems the surfaces are necessarily homeomorphic to the sphere or to the plane; here doubly covered plane domains are also admitted.

Theorem 4. *Every not strictly convex metric is the Euclidean metric in a sufficiently small domain.*

Theorem 5. *A complete strictly convex metric can be defined only on the following two-dimensional finitely connected manifolds: a sphere, a plane and a projective plane.*

Theorem 6. *For a metric defined by a quadratic form ds^2 with twice continuously differentiable coefficients to be convex, it is necessary and sufficient that the Gaussian curvature given by the form ds^2 be nonnegative everywhere.*

Theorem 7. *If a convex metric is defined in a manifold R, then for every point $x \in R$, there exists a neighborhood $U(x)$ of x such that the metric in $U(x)$ is realized by a certain convex surface.*

The main technique of proof is, on the one hand, the approximation of a convex surface by convex polyhedra and, on the other hand, the approximation of a convex metric by polyhedral convex metrics whose realizability is proved in [1].

7. Now we give a few results obtained in the course of proving the theorems formulated above.

Let a convex metric be given in a manifold R. Let two geodesics start from a point $x \in R$. We can construct a normal geodesic triangle, as small as we please, with vertex at x and sides on the given geodesics. Let a be the side opposite to x, and b and c the sides meeting at x. Drawing the midline a_1 that joins the midpoints of the sides b and c, by virtue of the third condition governing a convex metric, we find

$$\frac{a}{2} \leqslant a_1 \leqslant \frac{b+c}{2}.$$

Repeating this construction n times, we obtain

$$\frac{a_{n-1}}{2} \leqslant a_n \leqslant \frac{b+c}{2^n}.$$

[4]The sphere being compact, the condition of completeness is satisfied trivially.

Therefore the ratios $\frac{2^n a_n}{b}$ and $\frac{2^n a_n}{c}$ are bounded and vary monotonically. They have therefore a limit. So we can readily define the angle φ between the geodesics as $\cos\varphi = \lim_{n\to\infty} \frac{b^2+c^2-(2^n a_n)^2}{2bc}$. Using condition (1) of the definition of convex metric, we can easily show that the angle φ does not depend on the choice of the initial lengths of the sides b and c, provided they are sufficiently small. Thus, we obtain the following theorem.

Theorem 8. *In a convex metric there always exists an angle between two geodesics starting from a common point. This angle is always different from zero.*

I. M. Liberman has recently demonstrated that there always exists a semitangent to a geodesic on a convex surface at the starting point. The angle between two semitangents to geodesics emanating from a point as measured on the tangent cone at the point of start is equal to the angle between the geodesics in the sense of the intrinsic definition given above.

Theorem 9. *The sum of the angles of a geodesical n-gon (homeomorphic to the disk) in a manifold with a convex metric is greater than or equal to $(n-2)\pi$. The sum of the angles of a geodesic n-gon on a convex surface is greater than $(n-2)\pi$ by an amount equal to the area of the spherical image of its interior domain.*

The first assertion enables us to determine the intrinsic integral curvature as a function of a domain in a manifold with a convex metric. The second assertion shows that the intrinsic integral curvature of a domain on a convex surface is equal to the area of its spherical image. The spherical image is defined by the normals to supporting planes.

Theorem 10. *Every angle of a plane triangle whose sides are equal to the sides of a normal geodesic triangle on a convex surface is not greater than the corresponding angle of the geodesic triangle.*

Finally, using the conditions governing convex metric, we can establish several theorems concerning geodesic lines, geodesic triangles and geodesic circles. Thus we gain full knowledge about the intrinsic geometry of an arbitrary convex surface.

Steklov Mathematics Institute,
The USSR Academy of Sciences

Received June 21, 1941

References

[1] Alexandrov A. D., Dokl. Akad. Nauk SSSR **30** (1941), no. 2.

CHAPTER XI

EXISTENCE OF A CONVEX POLYHEDRON AND A CONVEX SURFACE WITH GIVEN METRIC

(*Communicated by Academician I. M. Vinogradov*)

Doklady Akad. Nauk SSSR, Matematika, vol. 30, No. 2, 103–106 (1941).

1. Let K be a topological complex of triangles homeomorphic to a sphere.[1] If the triangles in K are Euclidean, then the complex K is said to have *metrization* m or a *polyhedral metric* m is said to be defined on K. Such a metric is called *convex*, if the sum of the angles at every vertex is less than 2π. The aim of this note is to prove

Theorem 1. *For every convex polyhedral metric, there exists a unique polyhedron (up to a translation or a translation with a symmetry) realizing this metric.*

Or more precisely: let K be a complex of triangles homeomorphic to a sphere and m its convex metrization. Then there exists a convex polyhedron p that admits a triangulation forming the complex K and isometric to metrization m, i.e., a triangulation such that the triangles of the metrized complex K and the triangulation of the polyhedron p are isometric. (The triangles on p need not necessarily be planar, but their sides are invariably geodesics on p.)

A convex polygon, which is supposed to be doubly covered and in this sense homeomorphic to a sphere, is also assumed to be a convex polyhedron.

2. Lemma 1. *Two isometric convex polyhedra are congruent.*

Cauchy proved this lemma under the assumption that the polyhedra[2] are identical in structure. Lemma 1, which is somewhat more general, can

[1]By a *complex of triangles* we mean a set of triangles whose sides are incident pairwise. Here we do not except the case in which two triangles have more than one pair of incident sides or the case in which two vertices or sides of a triangle may be incident to one another.

[2]See, for example, [4].

be easily reduced to the Cauchy theorem. Indeed, let P_1 and P_2 be two isometric convex polyhedra. On P_1 construct geodesic segments that correspond to the edges of P_2 and vice versa. Thus we obtain finer subdivisions of P_1 and P_2. The complexes thus obtained have identical structures. Since the vertices of P_1 and P_2 are obviously identical (the sum of the angles at a vertex being less than 2π, the vertex is preserved under isometry), new "pseudovertices" may be formed on P_1 only due to the intersection of the edges of P_1 with the images of the edges of P_2 (pseudoedges) and on P_2 due to the intersection of the edges of P_2 with the images of the edges of P_1. At each pseudovertex the difference of the dihedral angles (dihedral angle at a pseudoedge is π) changes its sign exactly four times. However, the pseudoedges newly generated at vertices obviously preserve the sign changes of the differences of dihedral angles. Cauchy's method can therefore be applied to prove our lemma.

3. We prove the existence of a polyhedron with given metric by

Lemma 2. *Let P be a manifold consisting of elements p, and let M be a manifold of elements m and of dimension the same as P. Let P be mapped bijectively and continuously onto a subset M' of the manifold M such that if m_n is the image of p_n and the elements m_n converge to m, then the elements p_n converge to p, the inverse image of m. (Consequently, the mapping of P_n onto M is homeomorphic.) Furthermore, let every connected component of the manifold M contain the images of the elements of P. Then M' coincides with M, in other words, P is mapped onto M.*

The mapping of P onto M' is a homeomorphism; hence, by the domain invariance theorem, the submanifold M' is open in M. However, by the conditions of Lemma 2, M' is closed in M. Since the points of M' belong to every connected component of M, the submanifold M' is M.

4. We prove our theorem by induction on the number of vertices in the complex K. For a complex with three vertices our theorem is trivial: such a complex is realized by a twice covered triangle. Suppose that our theorem holds for a complex with $(n-1)$ vertices. We demonstrate that it also holds for a complex with n vertices.

Take an n-vertex topological complex K of triangles homeomorphic to a sphere and M_0 be the set of its metrizations. A metrization obviously is defined by the lengths of edges. The "triangle axiom" being the only condition governing the sides of a triangle, the set M_0 is the interior of a convex solid angle in the k-dimensional space in which coordinates are equal to the lengths of the edges of the complex K. Convex metrics in the manifold M_0 form the manifold M. Now we prove

Lemma 3. *In every simply connected component of the manifold M, there always exist realizable metrics.*

The boundary of M in M_0 consists of metrics in which the sum of the angles at every vertex is less than or equal to 2π, and there is at least one vertex where the sum is 2π.[3] By retriangulation we can easily eliminate the vertices where the sum of the angles is 2π. Therefore, the boundary of M consists of convex metrics having less than n vertices. However, this boundary obviously represents an analytic hypersurface F and the boundary of every connected component of the manifold M contains an ordinary point m_0 of the surface F. Moreover, such a point m_0 has the property that in a sufficiently small neighborhood of m_0 there are no points of other connected components of the surface F and consequently, there are no points of other connected components of the manifold M.

By the induction assumption, the metric m_0 is realized by some convex polyhedron p_0. On this polyhedron let us construct a triangulation forming the complex K and isometric to m_0. The vertices of the complex K, where the sum of the angles is 2π, are the pseudovertices of the polyhedron p_0. A sufficiently small displacement of these pseudovertices outward the polyhedron generates a convex polyhedron with n real vertices. This convex polyhedron admits a triangulation of K in a metric as close to m_0 as we please. The metric of the triangulation thus constructed for the polyhedron exists in a small neighborhood of the metric m_0, consequently, in the connected component of the manifold M. This completes the proof.

5. Now consider the manifold P of polyhedra with n-vertices each of which admits a triangulation K. We assume that each polyhedron of P has the triangulation, and we regard two polyhedra different not only when they are noncongruent but also when they are congruent but their triangulations differ. Two polyhedra are supposed to be congruent, if and only if they can be made coincident together with their triangulations by a translation or by a translation with a reflection. Of course, the vertices in triangulation are regarded different and their permutations are inadmissible.

Take three vertices of the complex K and the corresponding vertices a, b and c of the polyhedra in P. Take the vertex a at the origin, the vertex b on $x > 0$ and the vertex c on $y > 0$ of the plane xy. Take a polyhedron p from P with given triangulation K. On moving the vertices of this polyhedron through a small distance, we obtain a convex polyhedron with n vertices and a triangulation K. Because of the conditions imposed on the vertices a, b and c, we have in all $(3n - 6)$ coordinate variables. Different displacements of vertices give rise to different polyhedra.[4]

[3] It is a simple matter to prove that M_0 contains the metrics not contained in M, so M really has a boundary in M_0.

[4] By the conditions imposed on the vertices a, b and c, two neighboring polyhedra are congruent if the initial polyhedron lies wholly in the plane (a, b, c), i.e., degenerates into a doubly covered polygon. The polyhedra in this case are symmetric relative to the plane (a, b, c) but their triangulations are obviously different.

Consequently, the set P is a $(3n-6)-$dimensional manifold.

From the Euler theorem, it readily follows that the complex K has $k=(3n-6)$ edges. Consequently, the manifolds P and M have identical dimensions.

Every polyhedron with a triangulation K defines a trivially unique metric and, by Lemma 1, every realizable metric defines only one polyhedron.

To a continuous deformation of a polyhedron, there apparently corresponds a continuous variation of its metric.

Moreover, if realizable metrics m_n converge to a convex metric m, the polyhedra p_n realizing these metrics m_n are bounded and a convergent subsequence of polyhedra can be chosen from p_n. The metric of the limiting polyhedron p of this subsequence, as can be easily verified, is the metric m. Since there is only one such polyhedron, it is clear that the polyhedron p_n must converge to the polyhedron p with the metric m.

Thus, all the conditions of Lemma 2 are satisfied, particularly, the last condition by virtue of Lemma 3. Consequently, the manifold P of polyhedra is mapped onto the manifold M of convex metrics which are hence all realizable.

6. Theorem 2. *A metric of positive curvature defined on a sphere is realized by a closed convex surface.*

Take some metric m of positive curvature on a sphere and carry out a sufficiently fine geodesic triangulation of this metric. Replace each triangle by a Euclidean triangle with the same sides. The angles in this substitution obviously decrease by $\frac{1}{3}KF$ [1], and therefore we obtain a convex metric that approximates the metric m: the finer the triangulation, the more exact the approximation. Constructing the polyhedra that realize these polyhedral metrics, and choosing a convergent subsequence and taking the limit, we obtain a convex surface whose metric, as be easily verified, is the limiting metric, i.e., the given metric.

In 1915, Weyl speculated this theorem and even outlined its proof which was recently completed by Lewy [2]. However, his method of proof is quite different from mine.

7. Theorem 3. *For a metric of positive curvature defined on a plane E such that E is transformed into a complete (in the sense of metric space) differential geometric surface, there exists an infinite convex surface realizing this metric.*

On the plane E, let us construct a sequence of geodesic polygons covering the domains g_n such that the angles of the nth polygon in g_n are $\leqslant \pi$.

Cohn-Vossen asserts that such a construction is possible [3]. Let

$$g_n \subset g_{n+1} \text{ and } \sum_{n=1}^{n} g_n = E.$$

Let us take a sequence of $\varepsilon_n > 0$, $\varepsilon_n \to 0$. Construct geodesic triangulations of the domains g_n with sides of triangles equal to ε_n and then replace them by plane triangles with the same sides. Now complete the metrics thus obtained to the metrics given on a sphere by adding to each g_n an appropriately chosen convex polyhedron. Let P_n be the polyhedra realizing the convex metrics thus constructed. Clearly, we can choose a convergent subsequence from P_n such that the limiting surface is the surface we need.

Steklov Mathematics Institute,
The USSR Academy of Sciences

Received October 22, 1940

References

[1] Cartan H., *Geometry of Riemann Spaces*, 1936.
[2] Lewy H., Proc. Nat. Acad. Sci. USA **24** (1938).
[3] Cohn-Vossen S., Mat. Sb. **7** (1936), no. 2.
[4] Hadamard J. S., *Elements, Geometry* (Part II).

CHAPTER XII

ON TILING A SPACE WITH POLYHEDRA

Vestnik Leningradskogo Universiteta, No. 2, 33–43 (1954).

In his paper *On a class of Euclidean polyhedra* Venkov[1] studied a tiling of n-dimensional Euclidean space with translates of a convex polyhedron which are joined together at their faces so that a face of one translate is wholly covered by a face of another. And he derived the necessary and sufficient conditions for the space to be tiled in this fashion without mutual overlappings of translates, i.e., for the polyhedron to be a *parallelohedron*. In this paper we generalize the Venkov theorem and give even a simpler result, which is actually a corollary of a well-known theorem in topology, namely, a simply connected space (or polyhedron) has no other covering than the space itself.[2] Here we give all the necessary definitions in order to make our paper comprehensible without the need for citing Venkov in the sequel.

§ 1. Formulation of the Problem

We study a *tiling* of an n-dimensional simply connected complete space R^n of constant curvature (i.e., either Euclidean, spherical or Lobachevskiĭ space) with polyhedra. In simple words, we deal with the following construction. Suppose that finitely many n-dimensional polyhedra P_i are given. In the space R^n we take a polyhedron congruent to one of the given polyhedra P_i and place on its every $(n-1)$-dimensional face a polyhedron congruent to one of the given polyhedra P_i; again on all its free $(n-1)$-dimensional faces we place a polyhedron from the polyhedra P_i and continue this process. At

[1] See in the present issue of the journal. I had an opportunity to read his manuscript.
[2] See, H. Seifert and W. Threlfall, Topology, 1938 [Russian translation], p. 224.

each step, certain $(n-1)$-dimensional faces can be simply identified with one another, instead of placing a new polyhedron on them if each of them can superimposed on another completely and belongs to the polyhedra lying near them on opposite sides of their planes.

In this construction we do not except that polyhedra intersect.[3]

In the sequel we need the following rigorous definition. Let there be given a complex K^n formed by n-dimensional polyhedra. We assume that every face of a polyhedron belongs to the complex. The *complex* K^n is understood in an abstract sense, i.e., although every polyhedron in it is a polyhedron in R^n (Euclidean, spherical or Lobachevskiĭ space), the complex K^n is nevertheless not supposed to be embedded in R^n, but forms by itself an appropriate polyhedron $\widetilde{K}^n$. Furthermore, K^n is assumed to possess the following properties:

(1) Among the polyhedra in the complex there are only finitely many distinct polyhedra, i.e., geometrically not congruent to one another.

(2) Every $(n-1)$-dimensional face of a polyhedron P in K^n is also a face of one more and only one more polyhedron P' in K^n.

(3) "Strong connectedness" condition. If P and P' are two arbitrary polyhedra in the complex K^n, then they are joined by a chain, i.e., a finite collection of polyhedra in which every two neighboring members are adjacent at an $(n-1)$-dimensional face. A chain is said to join two polyhedra P and P', if the first and the last polyhedron in the chain are just the polyhedrons P and P'.

(4) Let Q^k be a k-dimensional face $(0 \leqslant k \leqslant n-2)$ of a polyhedron P in K^n. It belongs also to P', if and only if there exists a chain containing Q^k joining P and P' such that every two neighboring polyhedra are adjacent at an $(n-1)$-dimensional face containing the face Q^k.

A complex may be infinite or even locally infinite, i.e., certain faces Q^k $(0 \leqslant k \leqslant n-2)$ may belong to infinitely many polyhedra of the complex.

We assume that in addition to the complex K^n there is also a continuous mapping of the corresponding polyhedron $\widetilde{K}^n$ into a given space R^n satisfying two conditions:

(1) For every polyhedron this mapping is congruent (isometric).

(2) If two polyhedra P and P' are adjacent at a face Q^{n-1}, then their images $\bar{P}$ and $\bar{P}'$ lie on opposite sides in the neighborhood of the image $\bar{Q}^{n-1}$ of the face Q^{n-1}. (If the polyhedra are convex, then the images lie

[3]Our generalization, unlike Venkov's formulation, is not restricted only to translates of a convex polyhedron in Euclidean space. Starting with an arbitrary polyhedron P, we can build up our construction as follows: in the space R^n we place a polyhedron P_0 congruent to the polyhedron P. Then, taking the reflection of this polyhedron P_0 in the plane of some of its faces, we obtain the polyhedron P_1. Then we repeat this process indefinitely. Thus we obtain an ensemble of congruent polyhedra that are adjacent at $(n-1)$-dimensional complete faces but may possibly have intersections.

generally on opposite sides of the plane of the face $\bar{Q}^{(n-1)}$. If they are not convex, then this is required only in the neighborhood of this face.) Polyhedra are thus embedded in R^n.

We prove two main theorems.

Theorem 1. *The image of a polyhedron $\widetilde{K}^n$ in the space R^n is the entire space itself, i.e., on joining the polyhedra at the faces such that a face of one completely covers a face of the other, the entire space is tiled. Furthermore, for a bounded domain of R^n, there always exists finitely many polyhedra in K^n whose images cover this domain.*

A basic problem here is to find the conditions for tiling to be attained without the mutual intersections of polyhedra, i.e., the polyhedron $\widetilde{K}^n$ is mapped bijectively onto R^n.[4]

This question is resolved by

Theorem 2. *For a mapping $\widetilde{K}^n \to R^n$ to be one-to-one, it is necessary and sufficient that the mapping be one-to-one near every $(n-2)$-dimensional face of the complex K^n, i.e., the polyhedra converging at such a face, when mapped, do not intersect in an arbitrarily small neighborhood of the face. In other words, for the space R^n to be completely tiled with polyhedra without overlapping, it is sufficient that such a tiling be realized without overlapping "locally" in the neighborhood of every $(n-2)$-dimensional face.*

Necessity of the conditions is obviously trivial; so we only prove their sufficiency. The theorem implies that a one-to-one tiling of an entire space can be studied through a one-to-one local tiling in the neighborhood of $(n-2)$-dimensional faces. In particular, it gives a simple geometric characteristic of a parallelohedron due to Venkov.[5]

---

[4] In this case the author speaks of a *one-to-one tiling*.—Eds.

[5] We know that every n-dimensional (normal or nonnormal) parallelohedron has (1) a center of symmetry, (2) each of its $(n-1)$-dimensional faces has a center of symmetry, and (3) for every $(n-2)$-dimensional face the $(n-1)$-dimensional faces parallel to it form a closed hexahedral or a tetrahedral domain. Venkov's theorem asserts that these conditions are sufficient for a convex polyhedron to be normal, i.e., for a polyhedron to admit a one-to-one tiling of a space by joining the translates of the polyhedron at faces so that a face of one translate completely covers a face of another. (A convex polyhedron not admitting such a tiling in which a face of one translate is completely covered by a face of another is called *nonnormal parallelohedron*.) This is be easily proved by Theorem 2. Indeed, conditions (1)–(3) imply that the $(n-1)$-dimensional faces of a domain corresponding to some $(n-2)$-dimensional face Q^{n-2} cover, when extended along Q^{n-2}, either a hexahedral or a tetrahedral prism having a center of symmetry. These prisms, when their translates are joined at faces, tile in a one-to-one fashion a neighborhood of the face Q^{n-2}. Hence, the condition of Theorem 2 holds, so a polyhedron having properties (1)–(3) provides a one-to-one tiling of a space when its translates are joined at the faces, i.e., the polyhedron is normal. Since conditions (1)–(3) are also satisfied by a nonnormal parallelohedron, from a theorem proved elsewhere, we obtain another theorem due to Venkov: *every nonnormal parallelohedron is also a normal parallelohedron.*

The Venkov theorem also implies that for a space R^n to be tiled in a one-to-one fashion with regular polyhedra joined together at faces so that a face of one polyhedron completely covers a face of another, it is necessary and sufficient that the dihedral angles (i.e., the angle between the $(n-1)$-dimensional faces adjacent at an $(n-2)$-dimensional face) be an integral multiple of 2π. This is particularly true of a proper division of the n-dimensional sphere. In the same way, we obviously derive a condition for a one-to-one tiling of a space with polyhedra obtained from a given polyhedron P by consecutive reflections in $(n-1)$-dimensional faces. This condition implies that every dihedral angle of the polyhedron P should be an integral part of 2π, and if the dihedral angle at a face Q^{n-2} is an odd multiple of 2π, then P should have a plane of symmetry passing through the face Q^{n-2}.

Also notice that, as can be seen from the next section, the condition for joining polyhedra at faces so that one face fully covers some other, and the presence of only finitely many noncongruent faces are not absolutely necessary and can be weakened. The parts of a face at which polyhedra are adjacent may themselves be regarded as faces of polyhedra. Second, in place of the condition that there be only finitely many noncongruent polyhedra, it is suffices to require that there exist a number $a > 0$ such that every two nonadjacent faces of any dimension of a polyhedron be spaced at a distance not less than a. We however do not go into details as such generalizations are quite self-evident.

§ 2. Construction of Special Neighborhoods in the Polyhedron $\widetilde{K}^n$

We prove Theorems 1 and 2 by induction on the dimension of the space, noting a remark.

On some face Q^k $(k \neq 0)$ of the complex K^n, take an interior point A. If a polyhedron P of the complex K^n contains the point A, then we construct a spherical sector V with center at A such that the sector does not intersect any face, except Q^k and the faces of dimension greater than k adjacent to the whole face Q^k. Since, amongst the polyhedra P there are only finitely many noncongruent polyhedra, we can find a number $r > 0$ such that every polyhedron P containing the point A contains a spherical sector of radius r of the kind defined above. All these sectors taken together form a neighborhood of the point A; we call it a spherical neighborhood $S(A, r)$ of the point A in the polyhedron $\widetilde{K}^n$. (In the sequel a spherical neighborhood is understood in the sense of this definition.)

The surface of a spherical neighborhood $S(A, r)$ consists of $(n-1)$-dimensional spherical polyhedra truncated from the sphere by the sectors V. These polyhedra form a certain complex K^{n-1}. From conditions (1)–(4)

imposed on the complex K^n (in § 1), it is clear that the complex K^{n-1} also satisfies these conditions (of course, everywhere $(n-1)$ is to be substituted for n). In particular, strong connectedness of the complex (condition (3)) follows from the fact that, by condition (4), the face Q^k and, consequently, the point A, are common both to the polyhedra P and P', if and only if there exists a chain of polyhedra joining P and P' in which every two neighboring polyhedra are adjacent at an $(n-1)$-dimensional face containing the face Q^k.

The polyhedron $\widetilde{K}^{n-1}$ formed by the polyhedra of the complex K^{n-1} is none else than the surface of the spherical neighborhood $S(A,r)$.

Furthermore, a continuous mapping of the polyhedron $\widetilde{K}^n$ into the space R^n naturally defines a continuous mapping of the polyhedron $\widetilde{K}^{n-1}$. And, since the mapping of the polyhedron $\widetilde{K}^n$ is congruent on every polyhedron, the polyhedron $\widetilde{K}^{n-1}$ is mapped into the $(n-1)$-dimensional sphere of radius r about the point $\bar{A}$, the image of the point A. This mapping clearly satisfies the two conditions formulated in § 1.

Thus, the complex K^{n-1} satisfies all the conditions that were formulated for the complex K^n in § 1. And as the dimension of the complex K^{n-1} is $(n-1)$, we are justified in applying induction.

We may also notice that if a point A lies in the interior of a polyhedron P of the complex K^n, then it has consequently a spherical neighborhood lying in P. However, if the point A lies on an $(n-1)$-dimensional face at which the polyhedra P and P' are adjacent, then the point A has a spherical neighborhood consisting of two hemispheres, so the complex K^{n-1} in this case simply consists of two hemispheres.

This result can be briefly formulated as

Lemma 1. *Every point A of the polyhedron K^n has a spherical neighborhood $S(A,r)$ satisfying the following: If A lies on a face, then the surface $\widetilde{K}^{n-1}$ of the neighborhood $S(A,r)$ consists of $(n-1)$-dimensional spherical polyhedra forming the complex K^{n-1} and obeying conditions similar to those imposed on the complex K^n.*

We now prove one more lemma.

Lemma 2. *For a given complex K^n, there exists a number $a>0$ such that every point in the polyhedron $\widetilde{K}^n$ has an a-neighborhood (of radius a) included in a certain spherical neighborhood with center generally at some other point B.*

Proof. Let $P_1, P_2, \dots, P_m$ be polyhedra such that every polyhedron in the complex K^n is congruent to one of the polyhedra P_i. Since $P_1, \dots, P_m$ are mapped congruently onto the polyhedra of the complex K^n, every point

of the polyhedron $\widetilde{K}^n$ is the image of one (or many) points of the polyhedra $P_1, \dots, P_m$ and conversely every point in the polyhedra $P_1, \dots, P_m$ has images in the polyhedron $\widetilde{K}^n$.

Take a point A in the interior or on the boundary of one of the polyhedra P_i and consider all its images A' in the polyhedron $\widetilde{K}^n$. By Lemma 1, each point A' has a spherical neighborhood $S(A', r_{A'})$. Then there exists an $r_A > 0$ such that the points A' have spherical neighborhoods $S(A', r_A)$ of radius r_A the same for all neighborhoods.

If the point A lies in the interior of a polyhedron P_i, then the assertion is obvious; the radius of the neighborhood of the point A in the polyhedron P_i can be taken as the radius r_A.

Let the point A lie on the boundary of a polyhedron P_i but in the interior of some of its faces Q^k $(0 \leqslant k \leqslant n-1)$. The spherical neighborhood of any of its images A' in the polyhedron $\widetilde{K}^n$ consists of spherical sectors V each lying in one polyhedron of the complex K^n. However, every polyhedron of the complex $\widetilde{K}^n$ is congruent to one of the polyhedra $P_1, \dots, P_m$ and each of the polyhedra $P_1, \dots, P_m$ contains only finitely many images of the point A,[6] each image point having its own radius decreasing in the corresponding polyhedra. It suffices to take the least of these radii, and we have a radius r_A such that each image $A' \in \widetilde{K}^n$ of the point A has a spherical neighborhood of radius r_A.

About each point A in the polyhedra P_i let us draw a neighborhood of radius $\frac{1}{2}r_A$. For an interior point the neighborhood is a ball and, for a boundary point it is a spherical sector. By a well-known Borel theorem, the polyhedra are covered by finitely many neighborhoods with centers at some of the points $A_1, \dots, A_p$. Let a be the least radius of these neighborhoods.

Since the mappings of the polyhedra $P_1, \dots, P_m$ onto the polyhedra of the complex K^n are congruent, the polyhedron $\widetilde{K}^n$ is covered by the images of these neighborhoods.

If the point A lies on the boundary of a polyhedron P_i, then a neighborhood of it in P_i is only a spherical sector V of radius $\frac{1}{2}r_A$. If A' is the image of A in the polyhedron $\widetilde{K}^n$, then its neighborhood spherical sector V' can be supplemented to a whole spherical neighborhood. Therefore, around the point A', there is a whole spherical neighborhood of radius r_A (and, still more, of radius $\frac{1}{2}r_A$). Thus, the polyhedron $\widetilde{K}^n$ is covered by spherical neighborhoods $S(A'_i, \frac{1}{2}r_A)$ of radii $\frac{1}{2}r_A$.

Now let M be an arbitrary point in the polyhedron $\widetilde{K}^n$ and $S(A', \frac{1}{2}r_A)$

[6]Let the point A lie in the interior of a face Q^k. There are only finitely many faces of the polyhedra $P_1, \dots, P_m$ that are congruent to the face Q^k. And each of these faces has only finitely many admissible congruent mappings onto the face Q^k. Hence there are only finitely many images of the point A under such congruent mappings.

the spherical neighborhood with center some point A' that contains the point M.

By the definition of radius r_A, there exists a spherical neighborhood $S(A', r_A)$ of radius r_A about the point A', and the magnitude of the number a, by definition, is not greater than $\frac{1}{2}r_A$, so we have $r_A \geqslant 2a$. And, since the point M is contained in the neighborhood $S(A', \frac{1}{2}r_A)$, the a-neighborhood of the point M is completely included in the spherical neighborhood $S(A', r_A)$.

Thus, the number a has the required property, namely, every point M in the polyhedron $\widetilde{K}^n$ has an a-neighborhood included in some spherical neighborhood. This completes the proof of our lemma.

§ 3. Proof of Theorem 1

We now prove Theorem 1 which asserts that the image of the polyhedron $\widetilde{K}^n$ covers the entire space R^n and that every bounded domain in R^n is covered by the images of finitely many certain polyhedra in K^n. We prove the theorem by induction on the dimension n of the space. We start with $n = 1$. This case simply deals with the covering of a circle or a straight line by means of the segments joined to one another consequently. Obviously, the theorem is true for this case. (Indeed, for $n = 1$, the theorem is, strictly speaking, different from Theorem 1 formulated in § 1, at least, for the reason that a circle is not a simply connected space. However, as can be readily seen, this does not make any difference.)

Thus, we assume that the theorem holds for each $(n-1)$-dimensional space. On the basis of this assumption, we prove the following

Theorem. *If a point $A \in R^n$ is covered by the image of the polyhedron $\widetilde{K}^n$, then the whole ball of radius a about the point A is covered by finitely many images of certain polyhedra in K^n, where the number a does not depend on the point A and is the number a stated in Lemma 2.*

Let $A \in \widetilde{K}^n$ be the inverse image of the point $\bar{A}$ (or one of the inverse images, if $\bar{A}$ has several inverse images). By Lemma 2, the point A has a neighborhood of radius a included in a spherical neighborhood $S(C, r)$ of some point C. The surface of the neighborhood $S(C, r)$ consists of $(n-1)$-dimensional spherical polyhedra comprising the complex K^{n-1} which, by Lemma 1, is governed by the same conditions as for the complex K^n. The polyhedron $\widetilde{K}^{n-1}$ (the surface of the neighborhood $S(C, r)$) is mapped to the surface $\bar{S}^{n-1}$ of the ball S about the point $\bar{C} \in R^n$, the image of the point C.

Since, by our assumption, Theorem 1 holds for an $(n-1)$-dimensional space, the image $\widetilde{K}^{(n-1)}$ covers the entire sphere $\bar{S}^{(n-1)}$ and even a certain finite number of polyhedra from K^{n-1} suffices to cover it.

However, the mapping of the polyhedron $\widetilde{K}^n$ into the space R^n is congruent on every polyhedron. Therefore the ball $\bar{S}(C,r)$ is also covered by finitely many images of the spherical sectors forming the neighborhood $S(C,r)$. In short, the ball $\bar{S}$ is covered by finitely many polyhedra of the complex K^n. This completes the proof.

We now prove the first assertion of Theorem 1, namely, the *image of the polyhedron $\widetilde{K}^n$ covers the entire space R^n*.

Let $\bar{A}$ be an arbitrary point in R^n. Take some other point $\bar{B}$ in R^n that is covered by the image of the polyhedron $\widetilde{K}^n$, and draw the segment $\bar{A}\bar{B}$. Cover the segment $\bar{A}\bar{B}$ by finitely many segments $\bar{A}\bar{A}_1, \bar{A}\bar{A}_2, \dots, \bar{A}_p\bar{B}$, each of a length less than or congruent to a.

Since the point A is covered by the image of $\widetilde{K}^n$, by what has been proved above, the ball of radius a about the point A is also covered by the image of $\widetilde{K}^n$. Consequently, the point $\bar{A}_1$ is covered by the image of $\widetilde{K}^n$. Similarly, the point $\bar{A}_2$ is also covered by the image of $\widetilde{K}^n$, and so on. On reaching the point $\bar{B}$, we find that it is also covered by the image of $\widetilde{K}^n$. And as $\bar{B}$ is an arbitrary point, the whole space R^n is covered by the image of $\widetilde{K}^n$.

We now prove the second assertion: *every bounded domain of the space R^n is covered by the images of finitely many polyhedra of the complex K^n.*

Indeed, a bounded domain of the space R^n is covered by finitely many balls of radius a; but, by what has been proved above, each such ball is covered by the images of finitely many polyhedra in the complex K^n. Hence we obtain the second assertion. This completes the proof of Theorem 1.

Additionally, we can prove that *every segment in R^n is definitely covered by a finite "chain" of polyhedra from the complex K^n, i.e., by a sequence of polyhedra in which every two consecutive polyhedra are adjacent at an $(n-1)$-dimensional face.*

Let $\bar{A}\bar{B}$ be an segment in R^n. Starting from the endpoint $\bar{A}$, let us move towards $\bar{B}$, as we did in proving the first assertion of Theorem 1. Suppose that we arrive at a certain point $\bar{C}$ on the segment $\bar{A}\bar{B}$. If $\bar{C}$ is an interior point of a certain polyhedron, then we move ahead to its boundary. If the point $\bar{C}$ lies on a face $\bar{Q}^k$ of a polyhedron $\bar{P}$ and the segment $\bar{A}\bar{B}$ lies along this face, then we move up to a point where the segment $\bar{A}\bar{B}$ goes out of the polyhedron $\bar{P}$. What remains is the last case where the segment $\bar{A}\bar{B}$ goes out of the polyhedron $\bar{P}$ at a point beyond $\bar{C}$. Let P be the inverse image of the polyhedron $\bar{P}$, i.e., a polyhedron in the complex K^n, and let C be the inverse image of $\bar{C}$ in the polyhedron P. The point C has an a-neighborhood lying in the spherical neighborhood $S(D, r_D)$ of some point D. All those polyhedra that are intersected by the neighborhood S meet at some face Q^k containing the point D. The image of the neighborhood S covers the neighborhood of the point $\bar{C}$ and, consequently, amongst these polyhedra there is a polyhedron P' whose image $\bar{P}'$ contains a point of the

segment $\bar{A}\bar{B}$ that we encounter in moving from the point $\bar{A}$ to the point $\bar{B}$ beyond the point $\bar{C}$. By condition 4 imposed on the complex K^n, the polyhedra P and P' are joined together by a chain of polyhedra containing the face Q^k. To this chain we add the chain of polyhedra constructed in the course of moving to the point $\bar{C}$.

Thus, our construction gives a chain of polyhedra, all having a common point with the segment $\bar{A}\bar{B}$ and gradually cover the segment. That this construction takes us to the point $\bar{B}$ is clear from the existence of an segment of length a on which we can always move beyond any point C. Thus, the segment $\bar{A}\bar{B}$ is definitely covered by a finite chain of polyhedra, each having at least one common point with the segment $\bar{A}\bar{B}$.

§ 4. Proof of Theorem 2

We now prove

Lemma 3. *If each spherical neighborhood is mapped bijectively under a mapping of the polyhedron $\widetilde{K}^n$ onto the space R^n, then the whole polyhedron K^n is also mapped bijectively onto the space R^n.*

Proof. Let A be an arbitrary point of the polyhedron $\widetilde{K}^n$ and $\bar{A}$ its image in R^n. By Lemma 2, the point A has a neighborhood $U(A,a)$ of radius a included in a spherical neighborhood $S(B,a)$ of some point B.

According to Lemma 1, the surface $\widetilde{K}^{n-1}$ of S consists of spherical polyhedra of dimension $(n-1)$ forming the complex K^{n-1} which obeys the same conditions as the complex K^n.[7] The polyhedron (surface) $\widetilde{K}^{n-1}$ is mapped into the surface of a ball $\bar{S}$ about the point $\bar{B}$ (image of the point B) and, by Theorem 1, completely covers $\bar{S}$. Since the neighborhood S consists of spherical sectors that are mapped to R^n congruently, the image of S completely covers the ball $\bar{S}$.

Finally, since, by the condition of the theorem, the mapping of a spherical neighborhood is one-to-one, the mapping of S onto $\bar{S}$ is congruent as a whole. However, the neighborhood $U(A,a)$ of the point A is included in the ball S, and is therefore mapped congruently on a ball of radius a about the point A.

Thus, every neighborhood $U(A,a)$ about an arbitrary point A is mapped congruently onto a ball $\bar{U}$ in the space R^n.

Hence, it is clear that the mapping of the polyhedron $\widetilde{K}^n$ onto the space R^n is the so-called *covering map*: the polyhedron $\widetilde{K}^n$ covers the space R^n.

Indeed, the covering map is defined by three well-known requirements:

(1) Every point $\widetilde{M}$ of the space R^n has at least one inverse image in the polyhedron K^n.

[7] Hence the initial complex K^n is *locally finite*, i.e., every point of the complex belongs to finitely many polyhedra in the complex.

This condition holds, because, by Theorem 1, the image of $\widetilde{K}^n$ is the space R^n.

(2) Let the points $M_1, M_2, \dots$ be mapped to the point $\bar{M}$. Then there exists certain "labelled" neighborhoods $\bar{U}(M)$, $\bar{U}(M_1)$, $\bar{U}(M_2), \dots$ such that the neighborhoods $U(M_1), U(M_2), \dots$ are mapped topologically onto $\bar{U}(\bar{M})$, i.e., bijectively and continuously in both directions.

This condition is satisfied because as the "labelled" neighborhoods we can take neighborhoods of radius a that are mapped into R^n congruently.

(3) Every point N of the polyhedron $\widetilde{K}^n$, which is mapped to some point $\bar{N}$ in the labelled neighborhood $\bar{U}(\bar{M})$, belongs at least to one of the labelled neighborhoods $\bar{U}(M_1)$, $\bar{U}(M_2), \dots$.

This condition is also satisfied. Indeed, about the point N there is a neighborhood $U(N, a)$ mapped congruently to the space R^n. If the image $\bar{N}$ of the point N lies in a labelled neighborhood $\bar{U}(\bar{M}, a)$, then the point $\bar{M}$ therefore lies in the image of $U(N, a)$. And the mapping of this neighborhood into R^n being congruent, the neighborhood $U(N, a)$ contains one of the inverse images M_i of the point $\bar{M}$ and the neighborhood $U(M_i, a)$ contains the point N. This means the point N belongs to the labelled neighborhood $U(M_i, a)$, i.e., condition (3) is satisfied.

So the polyhedron $\widetilde{K}^n$ covers the space R^n. However, the space R^n is, by the condition of the theorem, simply connected. Therefore, by a well-known theorem, the mapping of $\widetilde{K}^n$ onto R^n is a topological map. This completes the proof of our theorem.

Moreover, under our conditions, the mapping of the complex $\widetilde{K}^n$ onto the space R^n is evidently a congruent map.

Now we turn our attention to

Theorem 2. *For a mapping of the polyhedron $\widetilde{K}^n$ into the space R^n to be one-to-one, it is sufficient that the mapping be one-to-one at every $(n-2)$-dimensional face.*

Proof. The theorem is true for $n = 2$. In this case the theorem implies that the mapping of the polyhedron $\widetilde{K}^2$ into R^2 is one-to-one around every vertex. Moreover, about every interior point of the polyhedron (in our case a polygon of the complex K^2) or around every interior point of a side, the mapping is one-to-one by the conditions imposed on this mapping.

Thus, a spherical (circular in our case) neighborhood of every point in K^2 is mapped bijectively into R^2. By Lemma 3, it follows that the mapping of $\widetilde{K}^2$ onto R^2 is one-to-one.

Assuming the theorem holds for the $(n-1)$-dimensional case, we demonstrate that it holds for an n-dimensional complex K^n.

Let A be a point of the polyhedron $\widetilde{K}^n$, S its spherical neighborhood and K^{n-1} the corresponding complex of $(n-1)$-dimensional polyhedra forming the surface $\widetilde{K}^{n-1}$ of the spherical neighborhood S. The polyhedron $\widetilde{K}^{n-1}$ is mapped onto the surface S^{n-1} of the ball $\bar{S}$ about the point $\bar{A}$ in the space R^n.

Every $[(n-1)-2]$-dimensional face of the complex K^{n-1} is simply the intersection of a certain $(n-2)$-dimensional face approaching the point A with a sphere having center at the point A. By the condition of the theorem, the mapping of the polyhedron $\widetilde{K}^n$ onto the space R^n is one-to-one about every $(n-2)$- dimensional face. Therefore the mapping of the polyhedron K^{n-1} onto S^{n-1} is one-to-one about every $[(n-1)-2]$-dimensional face. This means the conditions of the theorem are satisfied for the complex K^{n-1}. Since we assumed that the theorem holds for an $(n-1)$-dimensional complex, the mapping of the polyhedron $\widetilde{K}^{n-1}$ on the sphere S^{n-1} is one-to-one. Furthermore, the mapping of the spherical neighborhood S onto the sphere $\bar{S}$ is also obviously one-to-one. Thus we proved that the mapping of any spherical neighborhood is one-to-one and, consequently, the mapping of the entire polyhedron $\widetilde{K}^n$ onto the space R^n, by virtue of Lemma 3, is also one-to-one. This completes the proof of Theorem 2.

Received November 10, 1953

CHAPTER XIII

ON A GENERALIZATION OF RIEMANNIAN GEOMETRY

CONTENTS

1. Introduction

1.1. The geometry of a space which may be called a *space of curvature not greater than a given number* K, is the topic of this paper.[1] Briefly

[1]This paper is a revised and enlarged version of the lecture I read in March 1955 at the Humboldt University of Berlin. A part of the paper was published in *Trudy Mat. Inst. Steklov.*, 1951, No. 38: 5–23; but a greater part is being published for the first time.

speaking, a *space of curvature not greater than K* is a metric space in which two conditions are satisfied locally, i.e., in a neighborhood of every point:

(a) every two points are joinable by a *segment* or, as we say, by a *shortest arc*, i.e., a line of length equal to the distance between its endpoints;

(b) The sum of the angles defined appropriately between the sides of a triangle is not greater than the sum of the angles of a triangle with sides of the same length on a plane of constant curvature K. In a particular case where $K = 0$, the space is said to be a *space of nonpositive curvature*.

Every *Riemannian space* in which the curvature is bounded above by a number K is a space of curvature not greater than K. However, a space of curvature not greater than K is generally neither a Riemannian space nor a manifold. For instance, the figure formed by two plane (Euclidean) triangles meeting at a common vertex is a two-dimensional space of nonpositive curvature (provided the distance between two points is defined as the length of a shortest arc joining the points in this figure).

Nevertheless, all these spaces have many common properties and their geometry constitutes an extensive field for research. In this paper I expound the elements of this geometry; moreover, we derive some results that illustrate the general concepts. We deduce these results by geometric methods somewhat close to the methods of elementary geometry. Side by side, we mention a number of results, new in Riemannian geometry to my knowledge. There is not enough place in one paper to elaborate the theory in detail. The present theory utilizes the ideas that underlie the intrinsic geometry of surfaces and their development in my papers[2] [1, 2]. In the following few paragraphs we give the exact definitions of the basic concepts, formulate the main results and specify the relation between our theory and certain other works.

1.2. A shortest arc and a triangle. All our definitions hold for metric spaces. In the conventional notation, $\rho(x, y)$ stands for the distance between two points x and y in a given space. Often we denote by the symbol xy a shortest arc and the distance between x and y. We know that an analogy can be drawn between the length of a curve in a metric space and the usual definition of length. Namely, the *length of a curve L* represented by a continuous function $x_t[x \in R,\ t \in [0, t]]$ is

$$\rho(L) = \sup \sum_{k=0}^{n-1} \rho(x_{t_k}, x_{t_{k+1}}),$$

where $0 = t_0 < t_1 < \cdots < t_n = 1$.

[2]The number in square brackets refers to the paper in the References at the end of the paper.

A *shortest arc* is a curve whose length is equal to the distance between its endpoints. This is equivalent to the next assertion. A shortest arc, as a continuous image of the interval [0,1], has the following property: If x_t is a point of a shortest arc, then for every ordered triple $t_1 < t_2 < t_3$ ($t_i \in [0,1]$), we have

$$\rho(x_{t_1}, x_{t_2}) + \rho(x_{t_2}, x_{t_3}) = \rho(x_{t_1}, x_{t_3}).$$

Evidently, a segment of a shortest arc is also a shortest arc. By AB we denote the shortest arc that joins two points A and B, i.e., the shortest arc for which $x_0 = A$ and $x_1 = B$.

A *triangle* ABC is the union of the three shortest arcs AB, BC and CA which join three distinct points A, B and C in pairs. These points are called *vertices* and the shortest arcs AB, BC and CA, *sides* of the triangle. It is not excepted that the sides intersect or partially overlap; in particular, two sides may completely cover the third side: $AB + BC = AC$ so that the triangle degenerates into a shortest arc. In what follows a key role is played by the following construction. To a given triangle $T = ABC$, we put into correspondence a triangle T^K with sides of the same length on a surface of constant curvature K or, as we say, on a K-plane.

By a K-*plane* we understand a Euclidean plane if $K = 0$, a Lobachevskiĭ plane if $K < 0$ and an open hemisphere if $K > 0$. A triangle T^K on a K-plane with sides of the same length as a given triangle T in R is called a *triangle* on a K-plane which corresponds to the given triangle T. If $K \leqslant 0$, then to every triangle T there exists a triangle T^K (it is of course not excepted that T^K may degenerate into a segment). On the other hand if $K > 0$, then the triangle T^K exists only if the sum of the sides satisfies the inequality

$$AB + BC + CA < \frac{2\pi}{\sqrt{K}}.$$

In what follows we assume that this condition is always satisfied for $K > 0$. For a given $K > 0$, to this end it suffices to restrict ourselves to triangles included in a sufficiently small domain of the space R_K.

1.3. An angle. (The definitions given here are drawn from the paper [1].) Let L and M be two curves that start from a point O and let X and Y be two points different from O, one on L and the other on M. On a K-plane we now construct a triangle T^K with sides equal to the distances OX, OY, and XY (Fig. 1), respectively.

We denote the angle of this triangle at the vertex O' that corresponds to the point O by $\gamma^K_{LM}(X, Y)$ or simply by $\gamma(X, Y)$. In a particular case where L and M are shortest arcs, the positions of the points X and Y on them

FIG. 1

are defined by the distances $x = OX$ and $y = OY$, we therefore denote an angle also by $\gamma^K_{LM}(x, y)$. In what follows an angle thus defined is everywhere denoted by $\gamma^K_{LM}(X, Y)$ or $\gamma^K_{LM}(x, y)$ or simply by $\gamma(x, y)$.

The *upper angle* between two curves L and M is defined to be the upper limit of $\gamma(x, y)$ as $x, y \to 0$:

$$\alpha_{LM} = \varlimsup_{x,y\to 0} \gamma(x, y).$$

From the definition it is obvious that the upper angle always exists and that $0 \leqslant \alpha \leqslant \pi$. Since the angle in an infinitesimal triangle on a K-plane differs infinitesimally from the angle of the corresponding triangle on a Euclidean plane, the upper limit α_{LM} of the angles γ^K as $x, y \to 0$ does not at all depend on K and depends therefore only on the curves L and M alone.

If the limit $\alpha = \lim_{x,y\to 0} \gamma(x, y)$ exists, then we say there exists an *angle* between the curves L and M and it is equal to α.

The angle (upper angle) between the sides AB and AC of a triangle ABC is called the *angle (upper angle)* at the vertex A.

We now introduce the concept of *relative excess* $\delta_K(T)$ of a triangle T (excess relative to the curvature K). It is defined to be the difference between the sum of the upper angles of a triangle T and the sum of the angles of the corresponding triangle T^K.

1.4. Space of curvature $\leqslant K$. A space with curvature $\leqslant K$ is such that for every "sufficiently small" triangle T its *excess* $\delta_K(T)$ is less than or equal to zero. More precisely, *by R_K we denote a domain of a metric space that has the following properties:*

(a) *every two points in R_K can be joined by a shortest arc (not necessarily included in R_K);*

(b) *the relative excess* $\delta_K(T)$ *of an arbitrary triangle with vertices in* R_K *is nonpositive; and*

(c) *if* $K > 0$, *the sum of the sides of each triangle with vertices in* R_K *is less than* $2\pi/\sqrt{K}$.

Property (c) has been introduced so that a corresponding triangle on a K-plane may exist. We can simply assume that the diameter is sufficiently small.

By a space of curvature $\leqslant K$ we understand a metric space such that each of its points is contained in the domain R_K[3]. All our further reasoning applies only to the domain R_K: we are not studying the properties of spaces of curvature $\leqslant K$ as a whole. Let us note that a space is a space of constant curvature only if the relative excess of every triangle is not only nonpositive but is also zero. More precisely, as we prove in § 3, if the domain R_K is homeomorphic to a domain of the n-dimensional Euclidean space and the *relative excess* $\delta_K(T)$ of every triangle T in R_K is equal to zero, then the domain R_K is isometric to a domain in an n-dimensional space of constant curvature K.

The definition of a space of curvature $\leqslant K$ can be reformulated in a different form closely related to the usual concept of curvature. Namely, a space of curvature $\leqslant K$ is defined by the following two conditions:

(a) every point in the space has a neighborhood in which every two points can be joined by a shortest arc; and

(b) for an arbitrary sequence of triangles T which converges to a point, we have

$$\overline{\lim}\ \frac{\delta_0(T)}{S(T^0)} \leqslant K, \tag{1}$$

where $\delta_0(T)$ is the *absolute excess* of T, i.e., its excess relative to zero curvature: $\delta_0(T) = \alpha+\beta+\gamma-\pi$, and $S(T^0)$ is the area of the corresponding triangle on a plane.

At the end of § 3 we prove that this definition is equivalent to the previous one.

It should be noted that formula (1) admits the zero value for the area $S(T^0)$. In this case it is understood that for these triangles T (at least,

[3]It is readily seen that this definition is equivalent to the following: a space of curvature $\leqslant K$ is a metric space in which each of its points is contained in a neighborhood U such that (i) every pair of points in the neighborhood U can be joined by a shortest arc (not necessarily included in the neighborhood), and (ii) for every triangle T included in the neighborhood U, its excess $\delta_K(T)$ is less than or equal to zero. Incidentally, we prove that each point in R_K is contained in some *convex neighborhood*, i.e., in a neighborhood such that every two points in it can be joined by a shortest arc lying wholly in the neighborhood. So we can confine ourselves only to such neighborhoods.

for some sufficiently large number) the absolute excess $\delta_0(T)$ is less than or equal to zero.[4] We cannot exclude the case of $S(T^0) = 0$ just for the reason that the definition of a space of curvature K would otherwise fail to be equivalent to the previous definition. A trivial example of a space of nonpositive curvature is a *straight line*: every triangle lying on a line degenerates and has zero excess, so the condition $S(T^0) = \delta_0(T) = 0$ is satisfied.

Although, one-dimensional and even zero-dimensional spaces are formally spaces of curvature $\leqslant K$, nevertheless, they are of no real interest.

1.5. Main properties of R_K. The fundamental statement concerning a space of curvature $\leqslant K$, which we prove in Section 3 and which distinctly elucidates the properties of the space, asserts that *in the space R_K the angle $\gamma^K_{LM}(x, y)$ between every two shortest arcs L and M starting from a common point is a nondecreasing function of the distances x and y.*

Monotonicity of $\gamma(x, y)$ implies that $\lim_{x,y\to 0} \gamma(x, y)$ exists; hence *in the space R_K there exists a definite angle between every two shortest arcs starting from a common point.*

Furthermore, in the space R_K, as asserted by the definition, not only is the sum of the angles of a triangle T not greater than the sum of the angles of the corresponding triangle T^K on a K-plane, but *every angle of the triangle T is also not greater than the corresponding angle of the triangle T^K.*

Then the angle $\gamma(x, y)$ does not decrease with the increasing of x and y, and for arbitrary x and y we have $\gamma(x, y) \geqslant \lim_{x,y\to 0}(x, y)$. This means that the angle α_{LM} between two shortest arcs L and M obeys the inequality

$$\alpha_{LM} \leqslant \gamma^K_{LM}(x, y). \tag{2}$$

Let T be a triangle in R_K, with α the angle at its vertex A and α^K the angle of the corresponding triangle T^K. Obviously, α^K is none other than the angle $\gamma^K_{LM}(a, b)$, where L and M are the sides of the triangle T^K which pass through the vertex A and a and b are their lengths. Then inequality (2) implies $\alpha \leqslant \alpha^K$. However, this means that every angle of a triangle T in the space R_K is not greater than the corresponding angle of the triangle T^K.

The assertion that the angle $\gamma(x, y)$ is a nondecreasing function clearly demonstrates that the shortest arcs starting from a common point in R_K

[4] The equality $S(T^0) = 0$ implies that the sum of every two sides of the triangle is equal to its third side, say, $AB + BC = AC$. However, the sides AB and BC in this case merge into a shortest arc and the angle between them is therefore π, so that $\delta_0(T) \geqslant 0$. Indeed, if $S(T^0) = 0$, we have $\delta_0(T) = 0$.

everywhere deviate from one another not slower than on a K-plane. This can be explained as follows.

Let A and B be two points, one each on the shortest arcs L and M, starting from the point O (Fig. 2). The triangle $T^K = O'A'B'$ corresponds to the triangle $T = OAB$. Let $\gamma^K_{LM}(a, b)$ be the angle at the vertex O', where $a = OA$, $b = OB$. Now let X and Y be points on OA and OB; X' and Y'

FIG. 2

the corresponding points on $O'A'$ and $O'B'$, respectively, i.e., points such that $x = OX = O'X'$, $y = OY = O'Y'$, then $XY \leqslant X'Y'$.

Indeed, if $\gamma(x, y)$ denotes the angle of the triangle on a K-plane and x and y are its sides, $Z' = X'Y'$, etc. and since $\gamma(x, y) \leqslant \gamma(a, b)$ we have $XY \leqslant X'Y'$. Taking, for instance, X and Y at the midpoints of the sides OA and OB, we arrive at the assertion:

The midline of a triangle $T = OAB$ in R_K is not greater than the midline of the corresponding triangle T^K.

Since the shortest arcs in R_K starting from a common point deviate from one another not slower than the corresponding shortest arcs on a K-plane, a *shortest arc joining two points in R_K is unique* because the shortest arcs starting from a common point can no longer meet together as they deviate from one another no slower than on a K-plane.

However, the shortest arcs starting from a common point O may coincide on some part of OA and then start deviating from one another somewhere beyond A. This is true of a polyhedron at each vertex A the total angle about which is greater than 2π. A shortest arc passing through such a vertex A is not unique beyond this point.

In addition to these main properties of the space R_K in § 3–6 we derive a number of other properties. For instance, in § 5 we introduce the concept of the area of a surface and prove in particular the following assertion.

In R_K it is possible to span a surface on each closed curve of length l such that its area is not greater than the area of a disk of perimeter l on a K-plane. (If $K > 0$, then it is assumed that $l\sqrt{K} < 2\pi$, otherwise no such disk would exist.)

This proposition generalizes the maximal property of a disk as well as Carleman's theorem which asserts that the area of the minimal surface spanned on a given contour in a Euclidean space is not greater than the area of a disk of the same perimeter [6].

1.6. The proof of the theorem concerning the monotonicity of the angle $\gamma(x, y)$ and, consequently, the construction of a theory of spaces of curvature $\leqslant K$ are based on the following general proposition on the upper angles of a triangle.

Let ABC be a triangle in an arbitrary metric space such that every two points on the sides of the triangle ABC can be joined by a shortest arc. Let α be the upper angle at the vertex A of this triangle and α_K the corresponding angle of the corresponding triangle on a K-plane. (Here K is an arbitrary number such that $AB + BC + CA < 2\pi/\sqrt{K}$, for $K > 0$.) Finally, let ν be the upper bound of the relative excesses of the triangles AXY whose sides AX and AY lie on the sides AB and AC, respectively. Then the following inequality holds:

$$\alpha - \alpha_K \leqslant \nu.$$

Since, by definition, $\nu \leqslant 0$ for every triangle in the domain R_K, generalization of this statement implies that in R_K the angle α is always less than or equal to the angle α_K, i.e., the upper angle of a triangle T in R_K is not greater than the corresponding angle of the triangle T^K. This assertion is the starting point for our study into the properties of the domain R_K.

1.7. Space of curvature $\leqslant K$: Examples.

(a) If, in a domain R of a Riemannian space, the curvature is bounded above by a number K, then R is a space of curvature $\leqslant K$.

(b) Every closed convex set M in a Riemannian space is an R_K domain, if the curvature in M is bounded above and K is the upper bound of the curvatures in M. (A convex set is a set such that every two points in it can be joined by a geodesic included in the set.) In particular, a closed convex set M in the Euclidean space is an R_0 domain. However, a set other than a convex set can also be a domain. For instance, a closed set M in the Euclidean space which consists of two convex bodies touching at

a boundary point is an R_K domain, if the distance between two points is defined as the length of the shortest arc joining these points in M. The following question is of particular interest: what are the properties that are necessary and sufficient for a set M in a space of constant curvature K to be an R_K domain, if the distance XY is defined as the lower bound of the lengths of the curves joining the points X and Y in M?

(c) A polyhedron consisting of triangles in a K-plane is a space of curvature $\leqslant K$, if the sum of the angles at each vertex not lying on a boundary is $\geqslant 2\pi$. A similar statement holds for a multidimensional polyhedron consisting of n-dimensional simplexes in an n-dimensional space of constant curvature K, such that the sum of dihedral angles at every $(n-2)$-dimensional interior face is $\geqslant 2\pi$ (under certain additional constraints omitted here).

(d) In some sense we can assert that a space with a metric presenting the limit of the metrics of spaces of curvature $\leqslant K$ is itself a space of curvature $\leqslant K$. So, in particular, the limit of Riemannian metrics of curvature $\leqslant K$ is a metric of curvature $\leqslant K$. Here we do not determine the conditions under which this assertion holds in general; we only state the simplest but the most important particular case.

In a ball S included in n-dimensional Euclidean space, let there be given continuous functions $\rho_m(X,Y)$, $m = 1, 2, \ldots$, of pairs of points X and Y which obey the usual properties of a metric. Assume that the ball S in every metric is also a ball in the sense of the metric ρ_m and a domain R_K with the same value of K or all m. Furthermore, let the functions $\rho_m(X,Y)$ converge uniformly to a function $\rho(X,Y)$ which vanishes, if and only if $X = Y$. Then the ball S in the sense of the metric ρ is a ball and an R_K domain with the same K.

The proof of this assertion is based on the following remark: if the shortest arcs L_m and M_m, corresponding to the metric ρ_m and starting from the point O, converge to the lines L and M, then these lines are shortest paths in the limit metric ρ. Similarly, the angles $\gamma_m^K(x,y)$ between the shortest arcs L_m and M_m as defined in Section 3 converge to the angle $\gamma^K(x,y)$ between the shortest lines L and M in the limit metric ρ. Since the angles $\gamma_m^K(x,y)$ are nondecreasing functions of x and y, so is the angle $\gamma^K(x,y)$. In fact, this property of the angle $\gamma^K(x,y)$ holds at least for those shortest arcs L and M that can be represented as the limits of shortest arcs corresponding to the metrics ρ_m. What remains to prove is that every given shortest arc is the limit of shortest arcs corresponding to the metrics ρ_m. However, from the last corollary of example 5, it is evident that every space of curvature $\leqslant K$ has this property, so the limit metric ρ is also a metric of curvature $\leqslant K$.

We can therefore assert that the limit of Riemannian metrics of curvature $\leqslant K$ is a metric of curvature $\leqslant K$. Then an important question arises as

to whether the converse holds true or not. In other words, can a general metric of curvature $\leqslant K$ defined in a domain of n-dimensional Euclidean space be (under appropriate conditions) the limit of Riemannian metrics of curvature $\leqslant K$ defined on the same domain G? For a two-dimensional manifold the answer is positive and is proved largely in [2].

1.8. Relation with certain other works. In my studies on intrinsic geometry I start from the concept of a *space with intrinsic metric*, namely, such a metric space in which the distance $\rho(X, Y)$ between two points is equal to the greatest lower bound of the lengths of the curves joining these points, all the lengths being determined in the same metric ρ.

Assuming that such a space is locally compact in a neighborhood of each point, we find that every two points are joinable by a shortest arc. On the other hand, if there exist shortest arcs, then the metric is intrinsic because, by definition, the lengths of shortest arcs are equal to the distances between their endpoints. Therefore in any case, for a locally compact space the condition (presupposed in the definition of R_K domain) for the existence of shortest arcs in a "small domain" (i.e., in a certain neighborhood of every point) is equivalent to the condition for the metric to be intrinsic.

The basic property of an R_K space mentioned in § 1, namely, that the angle $\gamma^K_{LM}(x, y)$ for two arbitrary shortest arcs L and M starting from a common point is a nondecreasing function of x and y, was proved by me for a convex surface with "specific curvature" $\leqslant K$. I called this property the "K-concavity" as opposed to "K-convexity" which characterizes the metric of convex surfaces in a three-dimensional space of constant curvature K and it implies that the angle $\gamma(x, y)$ on the contrary is a nonincreasing function of x and y.

Earlier I characterized the intrinsic metric of a convex surface by the property that the midline of a triangle on a surface is not shorter than the midline of the corresponding triangle on a K-plane [3], provided there exists a tangent cone at every point. Naturally, an opposite condition arises which might elucidate the metric of a space with nonpositive curvature or of a space with curvature in which the midline of a triangle is not greater than the midline of the corresponding triangle on a K-plane.

In his extensive paper [7] Busemann defined a space with nonpositive curvature through the property that the midline of a small triangle is not greater than the half-length of the corresponding side of the triangle. On the basis of this statement (and together with certain other general properties like (1) the relative compactness of a bounded set, (2) existence and (3) uniqueness of prolongation of a shortest arc), he constructed a theory of such spaces and demonstrated that they have properties largely similar to those of Riemannian spaces with nonpositive curvature.

Busemann spaces, however, include a wide class of *Finsler spaces*. A *zero-*

curvature Busemann space is a Minkowski space of arbitrary metric, i.e., an affine space in which a ball is taken to be an arbitrary centrally symmetric closed bounded domain. In such a metric the midline of a triangle is half the corresponding side. However, for such metrics the sum of the upper angles is as a rule greater than π.

Therefore, we have a choice of two alternatives for the definition of a space of nonpositive curvature, or in general, a space of curvature $\leqslant K$:

(a) the midline of a triangle is not greater than the midline of the corresponding triangle on a K-plane, or

(b) the "corresponding relative" excess of a triangle is nonpositive.

As shown in Section 5, assertion (a) is a corollary of assertion (b). However, the converse is false as demonstrated by the Minkowski metric already given as an example.

It is therefore clear that the Busemann assertions are applicable to spaces of nonpositive curvature in our sense (in any case under the assumptions that a bounded set is relatively compact and that prolongation of a shortest arc is unique); many of his conclusions are valid for spaces of curvature $\leqslant K$, provided that the conditions pertaining to the midline of a triangle are appropriately generalized.

Our present paper, by its results, has hardly anything common with Busemann's paper because, first, we center our attention on the local properties related to the concept of angle and on certain other properties which directly generalize the properties of Riemannian spaces but are generally not typical of Finsler spaces. In this respect we note that our conclusions are not at all based on the assumption that prolongation of a shortest arc is unique. A shortest arc may admit several prolongations, say, on a polyhedron with a vertex about which the total angle is $> 2\pi$.

Finally, let us note the general proposition, formulated in Section 6, concerning the upper angles of a triangle underlies our study of two-dimensional metric manifolds; I call the manifolds of bounded curvature [2]. However, it seems that a manifold admits a more general definition than that given in [2]. Namely, a manifold R of bounded curvature is defined by the following conditions:

(a) a manifold R is a two-dimensional metric manifold with intrinsic metric, and

(b) in a neighborhood of every point the sum of the excesses of every finite system of pairwise nonoverlapping triangles is bounded above

$$\sum \delta(T_i) < N,$$

where $N < \infty$ depends only on the neighborhood. The excess is defined to be $\alpha+\beta+\gamma-\pi$, where α, β and γ are the upper angles of a triangle. Instead of these excesses, we can take the relative excesses $\delta_K(T_i)$ for an arbitrary

constant K. Unlike in the definition given in [2] which requires that the sum of the absolute values of the excesses be bounded; in our generalization the boundedness of the sum of excesses from below is implied in the boundedness from above. This assertion was proved by V. Zalgaller.

These remark shows that in the two-dimensional case, the theory of manifolds of curvature $\leqslant K$ is a particular case of the theory of manifolds of bounded curvature.

2. General Propositions Concerning Upper Angles

2.1. First we compare certain general propositions concerning the upper angles (as defined in Section 1, § 3) in an arbitrary metric space.

Proposition 1. *Let α_{12}, α_{23} and α_{12} be the upper angles between the curves L_1, L_2 and L_3 starting from a common point. Then*

$$\alpha_{12} + \alpha_{23} \geqslant \alpha_{13}.$$

This general proposition is proved in [1]. Clearly, if L_1 and L_2 are the branches of the same shortest arc, then $\alpha_{13} = \pi$. Therefore, from Proposition 1 we derive

Proposition 2. *The sum of the upper adjacent angles between curves is not less than π.*

2.2. Proposition 3. *The upper angle α between two shortest arcs L and M starting from a common point O obeys the equality*

$$\alpha = \sup_{x \to 0} \gamma^K_{LM}(x, y), \tag{1}$$

where γ^K_{LM} is the angle defined in Section 1, § 3 (Fig. 1).

By definition, $\alpha = \sup \lim_{x,y \to 0} \gamma^K(x, y)$; in (1) the supremum is taken only for x tending to zero, whereas y can vary arbitrarily, i.e., the point X on L tends to zero, while the point Y on M varies arbitrarily. Of course, here X and Y can be interchanged.

The proof of this proposition is based on the following

Lemma. *For an arbitrary K the following holds*

$$\cos \gamma^K = \frac{y - z}{x} + \varepsilon, \tag{2}$$

where $\varepsilon \to 0$ as $\frac{x}{y} \to 0$ and x, y, z and γ^K have the same meaning as in Proposition 3, namely, $OX = x$, $OY = y$, $XY = z$ and X and Y lie on L

and M, respectively. (Since $Y \in M$, y is bounded and, therefore, $\frac{x}{y} \to 0$ implies $x \to 0$.)

Let, for instance, $K < 0$; we take $K = -k^2$. Then for a triangle T^K with sides x, y and z, by a well-known formula in Lobachevskiĭ geometry, we obtain the equality

$$\cosh kz = \cosh kx \cosh ky - \sinh kx \sinh ky \cos \gamma^k.$$

Hence, denoting, for simplicity sake, kx, ky and kz by x, y and z, we obtain

$$\cos \gamma^K = \frac{\cosh y - \cosh z}{\sinh x \sinh y} + \frac{\cosh y\,(\cosh x - 1)}{\sinh x \sinh y}. \tag{3}$$

Since

$$\cosh y - \cosh z = 2 \sinh \frac{y-z}{2} \sinh \frac{y+z}{2},$$
$$\cosh x - 1 = 2 \sinh^2 \frac{x}{2}, \qquad \sinh x = 2 \sinh \frac{x}{2} \cosh \frac{x}{2},$$

by virtue of (3), we derive

$$\cos \gamma^K = \frac{2 \sinh \frac{y-z}{2}}{\sinh x} \cdot \frac{\sinh \frac{y+z}{2}}{\sinh y} + \frac{\cosh y \sinh \frac{x}{2}}{\sinh y \cosh \frac{x}{2}}. \tag{4}$$

When x and $\frac{x}{y}$ tend to zero, the second term on the right-hand side tends to zero. Furthermore, by the triangle inequality we have $|y - z| \leqslant x$ and, therefore, $\frac{\sinh \frac{(y+z)}{2}}{\sinh y} \to 1$ when $\frac{x}{y}$ tends to zero; $2 \sinh \frac{y-z}{2}$ and $\sinh x$ are equivalent to $(y - z)$ and x, respectively. Hence, equality (4) yields

$$\cos \gamma^K = \frac{y-z}{x} + \varepsilon,$$

where ε tends to zero as x and $\frac{x}{y}$ tend to zero.

Now we prove Proposition 3, i.e., we demonstrate that the upper angle

$$\alpha = \sup \lim_{x \to 0} \gamma^K(x, y).$$

However, by definition, we have $\alpha = \sup \lim_{x,y \to 0} \gamma^K(x, y)$. Therefore it suffices to show that

$$\alpha = \sup \lim_{\substack{x \to 0 \\ y \geqslant a > 0}} \gamma^K(x, y), \tag{5}$$

where the limit is taken for $x \to 0$ and y is always greater than a given positive number.

So, we can assume that the point X on the shortest arc L tends to the point O and the point on the shortest arc M remains at a finite distance from the point O. On M we now take a variable point Y' tending to the point O such that

$$\frac{x}{y'} \to 0,$$

where $OY' = y'$. Let $XY' = z'$. Then, by the triangle inequality (Fig. 3), we have

$$YY' \geqslant XY - XY', \text{ i.e., } y - y' \geqslant z - z' \quad \text{or} \quad y - z \geqslant y' - z'.$$

FIG. 3

By virtue of the lemma proved above (formula (2)), the above inequality yields

$$\cos\gamma^K(x,y) \; + \varepsilon \geqslant \cos\gamma^K(x,y) \; + \varepsilon',$$

or

$$\gamma^K(x,y) \; \leqslant \gamma^K(x,y') + \varepsilon''. \tag{6}$$

However, x and y' tend to zero; so, by the definition of upper angle, we have

$$\alpha \; \geqslant \gamma^K(x,y') - \varepsilon''',$$

where $\varepsilon, \varepsilon''$ and ε''' tend to zero as x and y' tend to zero.

Therefore, from (6) we obtain

$$\alpha \geqslant \gamma^K(x,y) \quad \text{and} \quad d \geqslant \sup \lim_{x\to 0} \gamma^K(x,y).$$

This completes the proof.

2.3. Proposition 4. *Under the conditions and notation of Proposition 3, we have*

$$\cos\alpha \leqslant \lim_{\frac{x}{y}\to 0} \frac{y-z}{x}. \tag{7}$$

Proof. By Proposition 3, we have $\alpha \geqslant \sup\lim_{\frac{x}{y}\to 0}\gamma^K(x,y)$ and, consequently,

$$\cos\alpha \leqslant \lim_{\frac{x}{y}\to 0} \cos\gamma^K(x,y),$$

but, by formula (2), we have

$$\cos\gamma^K(x,y) = \frac{y-x}{x} + \varepsilon.$$

Hence we infer (7).

Corollary. Let L be a given shortest arc, with X a variable point on it and x the length of the segment from the point O to the point X over the shortest arc L (Fig. 4)

FIG.4

Let $Z(x)$ be the length of the shortest arc joining a fixed point A with the point X. Finally, let ξ be the upper angle between the segment OX and a shortest arc AX. (One of such shortest arcs. It is assumed that at least

one such shortest arc exists.) Then, *the left lower derivative of $Z(x)$ with respect to x is governed by the inequality*

$$\left(\frac{dZ}{dx}\right)_{l.l} \geqslant \cos\xi,$$

where l.l stands for "left lower."

To prove this inequality, it suffices to substitute in (7) ξ, $Z(x), Z(x+\Delta x)$ and $-\Delta x$ for α, y, z and x.

Remark. The function $Z(x)$ satisfies the condition

$$|Z(x+\Delta x) - Z(x)| \leqslant \Delta x,$$

that follows directly from the triangle inequality. Therefore, by virtue of a well-known theorem, this function has almost everywhere a derivative. In Euclidean space and, generally, in a Riemannian space, the derivative $\frac{dZ}{dx}$ always exists and is equal to $\cos\xi$.

2.4. We now return to the proposition formulated in § 6 of Section 1 which concerns the upper angles of a triangle.

Let us consider a triangle ABC in an arbitrary metric space under only one assumption: namely, every two points on the sides of a triangle can be joined by a shortest arc. Let α be the upper angle between the sides AB and AC. The problem now is to evaluate the difference between the angle α and the corresponding angle α_K of the triangle $A'B'C'$ on a K-plane with sides of the same length as of the triangle ABC.

In order that all the triangles considered in what follows may exist, we assume $K < 0$.[5]

Let us take two points X and Y, one on each side AB and AC, respectively, of the triangle ABC and consider the triangle AXY whose sides AX and AY are segments of the sides AB and AC, respectively (Fig. 5). We put $AX = x, AY = y$ and $XY = z$; let ξ and η denote the upper angles between AX, XY and AY, XY. On a K-plane let us construct the corresponding triangle $A'X'Y'$ (i.e., a triangle with sides of lengths equal to x, y and z); let γ_K, ξ_K and η_K be its angles that correspond to γ, ξ and η, respectively. The angles γ_K is a function of $x = AX$ and $y = AY$.

2.5. Our next problem lies in evaluating the left lower derivatives.

[5] For $K \geqslant 0$ the conclusion is the same as for $K < 0$. However, for $K > 0$, we have to assume that the perimeter of the triangle ABC is less than $2\pi/\sqrt{K}$; if this constraint is satisfied, all the triangles considered on a K-plane exist.

FIG. 5

Lemma 1. *If none of the sides of the triangle OXY is equal to the sum of the other two sides so that the angles ξ_K and ν_K are neither zero nor π, then*

$$\frac{\partial \gamma_K}{\partial x} \geqslant \frac{\cos \xi - \cos \xi_K}{\sin \xi_K} \cdot \frac{k}{\sinh kx}, \tag{9}$$

where $\frac{\partial \gamma_K}{\partial x}$ stands (here and in what follows) for the left lower derivative and $k^2 = -K$.

A similar equality holds naturally for $\frac{\partial \gamma_K}{\partial y}$ as well. If $K = 0$, then $\frac{1}{x}$ is to be substituted for $\frac{k}{\sinh kx}$ and if $K > 0$, then $\frac{k}{\sin kx}$ is to be substituted, where $k = \sqrt{K}$. The corresponding conclusions are deduced, word by word, in the same way as in the case of formula (9).

Proof. From a well-known formula in Lobachevskiĭ geometry, we find

$$\cosh kz = \cosh kx \, \cosh ky - \sinh kx \, \sinh ky \, \cos \gamma_K.$$

Or, substituting x, y, z, and γ for kx, ky, kz, and γ_K, we obtain

$$\cosh z = \cosh x \cosh y - \sinh x \sinh y \cos \gamma. \tag{10}$$

For the left lower derivative, as usual, we have

$$\sinh z \, \frac{\partial z}{\partial x} = \sinh x \, \cosh y - \cosh x \, \sinh y \, \cos \gamma + \sinh x \, \sinh y \, \sin \gamma \, \frac{\partial \gamma}{\partial x} \tag{11}$$

for those cases in which $\sinh z$ and $\sinh x \sinh y \sin\gamma$ are nonnegative and continuous functions. Let us now transform the first two terms on the right-hand side of (11), expressing $\sinh y \cos\gamma$ through formula (10). Thus, we obtain

$$\sinh x \cosh y - \cosh x \sinh y \cos\gamma = \frac{\cosh x \cosh z - \cosh y}{\sinh x}. \tag{a}$$

From a formula similar to (10) we infer

$$\frac{\cosh x \cosh z - \cosh y}{\sinh x} = \sinh z \cos\xi_K. \tag{b}$$

Now transforming the last term in (11) with the help of the law of sines, we find

$$\sinh y \sin\gamma = \sinh z \sin\xi_K. \tag{c}$$

By virtue of (a), (b) and (c), after dividing (11) by $\sinh z$, we obtain

$$\frac{\partial z}{\partial x} = \cos\xi_K + \sinh x \sin\xi_K \frac{\partial\gamma}{\partial x}$$

or, reverting back from x, z and γ to kx, kz and γ_K, we find

$$\frac{\partial z}{\partial x} = \cos\xi_K + \frac{\sinh kx}{k} \sin\xi_K \frac{\partial\gamma}{\partial x}. \tag{12}$$

By virtue of the corollary [formula (8)] of Proposition 4, we deduce

$$\frac{\partial z}{\partial x} \geqslant \cos\xi.$$

Therefore from (12) we obtain (9)

$$\frac{\partial\gamma_K}{\partial x} \geqslant \frac{\cos\xi - \cos\xi_K}{\sin\xi_K} \cdot \frac{k}{\sinh kx},$$

and this completes the proof.

2.6. Now we prove a lemma by which it is a simple matter to demonstrate the general proposition of § 6 in Section 1, which concerns the angles of a triangle.

Lemma 2. *If none of the sides a triangle AXY is equal to the sum of the other two sides and $\xi_K - \xi \geqslant \varepsilon > 0$, then there exists an $x' < x$ such that*

$$\gamma(x, y) - \gamma(x', y) > a \log \frac{x}{x'},$$

where $a > 0$ depends only on ε, K and the diameter of the triangle ABC.

To prove this lemma, first let us rewrite (9) for $\frac{\partial \gamma}{\partial x}$. Namely, if $\xi_K - \xi \geqslant \varepsilon > 0$, we have

$$\frac{\cos \xi - \cos \xi_K}{\sin \xi_K} \geqslant \frac{\cos(\xi_K - \varepsilon) - \cos \xi_K}{\sin \xi_K} = \sin \varepsilon - (1 - \cos \varepsilon) \cot \xi_K.$$

Since $\xi_K \geqslant \varepsilon$, we find $-\cot \xi_K > -\cot \varepsilon$ and, therefore,

$$\frac{\cos \xi - \cos \xi_K}{\sin \xi_K} \geqslant \sin \varepsilon - (1 - \cos \varepsilon) \cot \varepsilon = \frac{1 - \cos \varepsilon}{\sin \varepsilon} = \tan \frac{\varepsilon}{2}.$$

Using this inequality, from (9) we obtain

$$\frac{\partial \gamma_K}{\partial x} \geqslant \tan \frac{\varepsilon}{2} \cdot \frac{k}{\sinh kx}. \tag{13}$$

The function $\frac{kx}{\sinh kx}$ is continuous and positive in the closed interval $[0, d]$, where d is the diameter of the triangle ABC; therefore it is bounded from below by a positive number, so that

$$\frac{kx}{\sinh kx} \geqslant b > 0; \qquad \frac{k}{\sinh kx} \geqslant \frac{b}{x}.$$

(Since $\frac{kx}{\sinh kx}$ decreases with x increasing x, it suffices to take $b = \frac{kd}{\sinh kd}$.)

Hence, by virtue of (13), we can write[6]

$$\frac{\partial \gamma}{\partial x} \geqslant 2a \frac{1}{x} = 2a \frac{d \log x}{dx},$$

where $2a = b \tan \frac{\varepsilon}{2}$.

Since $\frac{\partial \gamma}{\partial x}$ is the left lower derivative, we can definitely find an $x' < x$ such that

$$\gamma(x, y) - \gamma(x', y) > a(\log x - \log x') = a \log \frac{x}{x'}.$$

This completes the proof of our lemma.

---

[6]If $k > 0$, we have

$$\frac{\partial \gamma}{\partial x} \geqslant \tan \frac{\varepsilon}{2} \cdot \frac{k}{\sin kx} \geqslant \frac{1}{x} \tan \frac{\varepsilon}{2}.$$

2.7. We prove a general proposition on the upper angles of a triangle.

Proposition 5. *If α is the upper angle between the sides AB and AC of the triangle ABC and α_K the corresponding angle in a triangle with sides of the same length (as the sides of the triangle ABC) on a K-plane, then*

$$\alpha - \alpha_K \leqslant \nu, \tag{14}$$

where ν is the upper bound of the relative excesses of the triangle AXY, i.e., $(\alpha + \xi + \eta) - (\gamma + \xi_K + \eta_K)$, and ξ, η, γ, ξ_K and η_K have the same meaning as defined before.

Proof. By the definition of the upper bound ν, we have

$$(\alpha - \gamma) + (\xi - \xi_K) + (\eta - \eta_K) \leqslant \nu.$$

The triangle ABC itself becomes the triangle AXY when X and Y coincide with B and C, respectively. So, we have $\gamma = \gamma_K$ and therefore

$$(\alpha - \alpha_K) + (\xi - \xi_K) + (\eta - \eta_K) \leqslant \nu. \tag{15}$$

Put $AB = x_0$ and $AC = y_0$.

For every two sides of the triangle ABC, we have a dilemma: either the sum of these two sides is equal to the third side or not.

For the first case, we show that estimate (14) is definitely satisfied. Let, for instance, $x_0 + y_0 = z_0$, i.e., $AB + AC = BC$, so that AB and AC together form a shortest arc. Then $\alpha = \pi$ and the corresponding triangle on a plane degenerates into a segment, so we have $\alpha_K = \pi$ and $\xi_K = \eta_K = 0$. However, as ξ and η are $\geqslant 0$, from (15) we derive

$$\nu \geqslant (\alpha - \alpha_K) + (\xi - \xi_K) + (\eta - \eta_K) \geqslant 0 = \alpha - \alpha_K.$$

The case of $x_0 + z_0 = y_0$, i.e., $AB + BC = AC$ (or $y_0 + z_0 = x_0$) is studied on similar lines. Thus, we obtain $\xi = \xi_K = \pi$, $\eta_K = 0$ and $\eta \geqslant 0$; so $\nu > (\alpha - \alpha_K) + (\xi - \xi_K) + (\eta - \eta_K) \geqslant \alpha - \alpha_K$.

2.8. What remains is to study the general case in which none of the sides of the triangle ABC is equal to the sum of the other two sides.

Suppose that inequality (14) does not hold, so that $\alpha - \alpha_K > 0$ which is equivalent to

$$\alpha - \alpha_K \geqslant \nu + 2\varepsilon, \tag{16}$$

where ε, for example, is equal to $\frac{1}{2}(\alpha - \alpha_K - \nu)$. Then, from (15) we obtain

$$(\xi_K - \xi) + (\eta_K - \eta) \geqslant (\alpha - \alpha_K) - \nu \geqslant 2\varepsilon.$$

Therefore, at least one of the numbers $(\xi_K - \xi)$, $(\eta_K - \eta)$ is greater than ε.

Let, for instance,

$$\xi_K - \xi \geqslant \varepsilon.$$

Then, by Lemma 2, we can take a point $X'(AX' = x < x_0)$ on the side AB such that

$$\gamma(x_0, y_0) - \gamma(x_0', y_0) > a \log \frac{x_0}{x'}.$$

If $\eta_K - \eta \geqslant \varepsilon$, then

$$\gamma(x_0, y_0) - \gamma(x_0, y') > a \log \frac{y_0}{y'}.$$

Combining both these two cases, we can assert that it is possible to take two points $x' < x_0$ and $y' < y_0$ such that

$$\gamma(x_0, y_0) - \gamma(x', y') > a \log \frac{x_0 y_0}{x' y'}. \tag{17}$$

We now consider the triangle $AX'C$ or ABY' or, in general, $AX'Y'$; for such a triangle the angle $\gamma(x', y')$ plays the part of the angle $\alpha_K = \gamma(x_0, y_0)$. By virtue of (17), we have

$$\gamma(x', y') < \gamma(x_0, y_0).$$

Therefore, by virtue of (16), we obtain

$$\alpha - \gamma(x', y') \geqslant \alpha - \alpha_K \geqslant \nu + 2\varepsilon. \tag{18}$$

For the triangle $AX'Y'$ this inequality plays the same part as inequality (16) for the triangle ABC and, consequently, the situation is the same as for the triangle ABC. Indeed, the value of ν for the "smaller" triangle $AX'Y'$ can only be less than the angle for the "greater" triangle ABC; therefore for the triangle $AX'Y'$ we can use ν in (18). Moreover, inequality (17) implies that the upper bound of the angle does not hold for the triangle $AX'Y'$. Therefore, the sum of every two sides of the triangle cannot be equal to its third side because the required estimate holds. Consequently, it is clear that for the triangle $AX'Y'$ inequality (18) has exactly the same meaning as inequality (18) for the triangle ABC.

Hence we can repeat our reasoning and can thus find $x'' < x'$ and $y'' < y'$ such that

$$\gamma(x', y') - \gamma(x'', y'') > a \log \frac{x' y'}{x'' y''}.$$

On combining this inequality with inequality (17), we obtain

$$\gamma(x_0, y_0) - \gamma(x'', y'') > a \log \frac{x_0 y_0}{x'' y''}.$$

Obviously, we can now repeat our reasoning for the triangle $AX''Y''$.

Let us consider all the x's and y's for which

$$\gamma(x_0, y_0) - \gamma(x, y) \geqslant a \log \frac{x_0 y_0}{xy}. \tag{19}$$

Then we find $\alpha_K = \gamma(x_0, y_0) \geqslant a \ln \frac{x_0 y_0}{xy}$ and, consequently, the product of those x's and y's for which (19) holds is bounded below by a positive number: $xy > c > 0$, where $c = c_0 y_0 \, e^{-\alpha_k / a}$.

Let ρ be the greatest lower bound for those xy's for which (19) holds so that $\rho \geqslant c > 0$. Then there exist x_n and y_n $(n = 1, 2, \dots)$ such that

(1) inequality (19) holds for them;

(2) $x_n y_n$ tends to ρ; and

(3) x_n and y_n converge to certain points $\bar{x}$ and $\bar{y}$, respectively.

Since the logarithm and the angle γ, being functions of x, y, are continuous, inequality (19) holds for $\bar{x}$ and $\bar{y}$ as well. This means what is true of the triangle ABC is also true of the corresponding triangle $A\bar{X}\bar{Y}$. So we can find $x' \leqslant \bar{x}$, $y' \leqslant \bar{y}$ such that strict inequality holds in at least one of these above inequalities for which (19) holds. Moreover, we have $x'y' < \bar{x}\bar{y} = \rho$, i.e., ρ, contrary to its definition, is not the lower bound of the product of those x's and y's for which (19) holds.

This contradiction proves the estimate for the difference $(\alpha - \alpha_K)$.

3. Main Properties of the Domain R_K

3.1. Let us recall the definition of the domain R_K given in § 4 of Section 1, namely, R_K is a domain in a metric space having the following properties:

(a) every two points in R_K can be joined by a shortest arc,

(b) every triangle with vertices in the domain R_K has a nonpositive excess relative to the curvature K, and

(c) for $K > 0$, the perimeter of a triangle with vertices lying in R_K is less than $\frac{2\pi}{\sqrt{K}}$.

3.2. All the conclusions derived here and in what follows apply to the domain R_K.

From the general proposition concerning the upper angles of a triangle proved in § 2, we obtain

Lemma 1. *The upper angles α, β, and γ of every triangle in R_K are not greater than the corresponding angles α_K, β_K, and γ_K of the triangle T^K on a K-plane.*

Since the general proposition mentioned above implies $\alpha - \alpha_K \leqslant \nu$ and since the relative excesses are nonpositive, we find that the upper bound of the excesses ν is less than or equal to zero and, therefore, $\alpha \leqslant \alpha_K$.

An elementary geometric proposition plays a key role in the sequel.

Lemma 2. *Let a polygon Q on a K-plane be bounded by three convex broken lines AB, BC and CA concave relative to the interior of Q (Fig. 6).*

FIG. 6

Let T be the triangle that is obtained from the polygon on straightening out the broken lines, i.e., a triangle whose sides are of the same length as the broken lines AB, BC, and CA (the case where one or two broken lines are straight segments is not excepted).

Then the angles at the vertices A, B, and C of the polygon Q are less than the corresponding angles of the triangle T.[7]

Proof. We demonstrate this assertion for the simplest case; namely, for a quadrangle $ABCD$ with the vertex D lying inside the triangle ABC (Fig. 7). We prove that the angles of this quadrangle are less than the corresponding angles of a triangle T with sides equal to AB, BC, and $AD + DC$. For the angle at B, this assertion is self-evident. We only have to prove the assertion for the angle at the vertex A.

Let us prolong the side AD of our quadrangle to E so that $DE = DC$ (Fig. 7). The triangles DBC and DBE have a common side DE, but the angle at D of the triangle DBE is less than the angle at D of the triangle DBC. Therefore $BE < BC$.

Hence the angle at the vertex A of the triangle ABE (i.e., the angle at the vertex A of our quadrangle) is less than the angle at the vertex A of the triangle with sides $AB, BC, AE = AD + DC$. However, this triangle is just the triangle T; so this completes the proof of our assertion.

Now we prove the lemma for the general case by induction on the number of vertices of the polygon Q.

[7] The triangle T exists because each of the broken lines AB, BC and CA is less than the sum of the other two. Only if $K > 0$, it is required that the perimeter of Q be less than $\frac{2\pi}{\sqrt{K}}$.

FIG. 7

Draw the diagonal D_2E to truncate a quadrangle AD_1D_2E from the polygon Q (Fig. 6). By straightening the broken line AD_1D_2, we transform the quadrangle into a triangle, thereby reducing the number of vertices D_i by one. By what has been proved above, the angles at the vertices D_2 and E_1 become larger and therefore the broken lines AB and AC remain convex. The angle at the vertex A also becomes larger. Repeating this construction, we once again increase the angle at the vertex A but decrease the number of vertices. Eliminating all the vertices on the broken lines AB, BC and CA, we find that the angle at the vertex A of the polygon Q is less than the corresponding angle of the triangle T.

3.3. Now we prove the main proposition concerning the domain R_K.

Proposition 1. *The angle $\gamma^K_{LM}(x,y)$ for every two shortest arcs L and M in R_K starting from a common point is a nondecreasing function of x and y.*

(The angle $\gamma(x,y)$ is defined in § 3 of Section 1.)

Proof. Let the shortest arcs L and M start from a common point O. Take a point Y on M and points X and X_1 on L such that $OX < OX_1$ (Fig. 8). Let us denote OY by y, OX by x and OX_1 by x_1.

Draw the shortest arcs XY and X_1Y. Thus we obtain the triangles $T = OXY$ and $T_1 = X_1XY$. On a K-plane let us construct the corresponding triangles $T^K = O'X'Y'$ and $T_1^K = X_1'X'Y'$, and join them together at the sides $X'Y'$ so as to obtain the quadrangle $Q = O'X'X_1'Y$.

By Lemma 1, the angles of the triangles T^K and T_1^K are not less than the angles of the triangles T and T_1, respectively. However, by Proposition 2 (Section 2), the sum of the upper angles at the vertex X of the triangles T and T_1 are not less than π because they are adjacent angles. The sums of the corresponding angles of the triangles T^K and T_1^K are therefore not less

FIG. 8

than π. These angles together with the angle at X' are the angles of the quadrangle Q. Consequently, the arrowhead of the angle at the vertex X' of the quadrangle Q is directed into the interior of the quadrangle and, in the limit case, where the angle at X' is π, the quadrangle forms a triangle.

On straightening out the broken line $O'X'X_1'$, the quadrangle Q is transformed into a triangle T_2^K, so the angle at the vertex O', by Lemma 2, becomes larger (or does not change, if the quadrangle is initially a triangle).

The corresponding angle of the triangle T_2^K is none else than $\gamma(x,y)$. However, the angle at the vertex O' of the quadrangle Q is also the angle of the triangle T^K, i.e., is none else than $\gamma(x,y)$. Therefore

$$\gamma(x,y) \leqslant \gamma(x_1,y),$$

and this completes the proof of our proposition.

Proposition 1 can be reformulated, as already mentioned in § 5 of Section 1, as follows:

Proposition 2. *Let X and Y be points on the sides AB and AC of the triangle $T = ABC$ in R_K and X' and Y' the corresponding points on the sides of the corresponding triangle $T^K = A'B'C'$ (i.e., $A'X' = AX, A'Y' = AY$), then*

$$XY \leqslant X'Y'.$$

Proposition 3. *There always exists an angle between every two shortest arcs starting from a common point in R_K and for arbitrary x and y, we have $\alpha_{LM} \leqslant \gamma^K_{LM}(x, y)$.*

Proof. Since, by Proposition 1, the angle γ^K_{LM} is a nondecreasing function, the limit $\lim_{x,y\to 0} \gamma^K_{LM}(x, y)$ exists, i.e., the angle α_{LM} exists and is not greater than the angle $\gamma^K_{LM}(x, y)$ for any x and y.

3.4. Proposition 4. *Every angle of a triangle in R_K is not greater than the corresponding angle of the corresponding triangle on a K-plane.*

This proposition is a direct corollary of Lemma 1 and Proposition 3.

Let us note that Lemma 2 is only a particular case of Proposition 4. This is a consequence of the fact that the polygon Q on a K-plane with its intrinsic metric forms the domain R_K. If, in addition, Q is bounded by three convex broken lines with concavity facing the interior of Q, then these broken lines are the shortest arcs in the polygon Q. Therefore Q, together with its intrinsic metric, is a triangle in the domain R_K and, consequently, each of its angles, by Proposition 4, is not greater than the corresponding angle of the corresponding triangle T^K.

3.5. Proposition 5. *A shortest arc joining every two points in R_K is unique.*

Assume that two shortest arcs L and M starting from a point O have a common point A. Denoting OA by a, we evidently find

$$\gamma^K_{LM}(a, a) = 0.$$

Since γ is a nondecreasing function, for an arbitrary $x \leqslant a$, we have

$$\gamma^K_{LM}(x, x) = 0.$$

Obviously, this implies that the arcs L and M coincide.

Proposition 6. *A shortest arc in R_K depends continuously on its endpoints a and b, i.e., if A_n tends to A and B_n tends to B, then the shortest arcs (A_nB_n) tend to the shortest arc (AB).*

Proof. Let A_n tend to A and B_n to B. Take an arbitrary point C on the shortest arc AB and arbitrary points C_n on the shortest arcs A_nB_n such that

$$AC : AB = A_nC_n : A_nB_n. \tag{1}$$

Draw the shortest arc AB_n and take a point D_n on it such that D_n divides AB_n in the same proportion (Fig. 9) as (1).

Applying Proposition 2 to the triangle ABB_n, and since $AC : AB = AD_n : AB_n = t$, we find that CD_n tends to zero as BB_n tends to zero.

FIG. 9

Similarly, D_nC_n tends to zero as AA_n tends to zero. However, $CC_n \leqslant CD_n + D_nC_n$; consequently, CC_n tends to zero when AA_n and BB_n tend to zero. Hence, C being an arbitrary point on the shortest arc AB, we find the shortest arcs A_nB_n converge to AB, when A_n and B_n converge to A and B, respectively.

3.6. Proposition 7. *If a point O in the domain R_K has a neighborhood U homeomorphic to the n-dimensional ball, and r is the distance between O and the boundary of U, then every shortest arc starting from O can be prolonged to a shortest arc of length r.*

Proof. Let $S(r)$ be the sphere of radius r with center the point O, i.e., the set of all points X which are spaced from O at the distance of $\rho(O, X) = r$. From the conditions of the proposition, we find $S(r) \subset U$; mapping the neighborhood U into the disk in n-dimensional Euclidean space, we can assume that $S(r)$ is embedded in Euclidean space and represents the boundary of a closed domain (image of the r-neighborhood of the point O).

We define a *deformation* of the sphere $S(r)$ as follows:

By virtue of Proposition 5, to every point $X \in S(r)$, there corresponds a unique shortest arc OX. For every $t\,(0 \leqslant t \leqslant 1)$, let us send a point X into a point X_t lying on the shortest arc OX such that

$$\rho(O, X_t) = tr = t\rho(O, X).$$

The deformation thus defined is continuous because the shortest arc OX is a continuous function of the point X (Proposition 6). Assume that in the r-neighborhood of the point O, there exists a point A such that the shortest arc OA cannot be extended to the shortest length r, i.e., to the sphere $S(r)$. Such a point does not lie on any of the radii of the sphere $S(r)$. Therefore under the deformation of the sphere $S(r)$, the point A does not

exist on the image of the sphere $S(r)$. Nevertheless, prior to deformation, the point A was inside the domain bounded by the sphere $S(r)$; but for small t, it definitely lies outside the sphere $S(tr)$. This contradicts a well-known proposition in topology [8, Chapter VI, proposition (10)]. And this completes the proof of our assertion.

3.7. Proposition 8. *For $K > 0$, a sphere in R_K is convex, if its radius is less than $\frac{\pi}{2\sqrt{K}}$. This implies that in R_K for every two points A and B spaced from a point O at a distance less than r such that $OA \leqslant OB \leqslant r$, then every point on the shortest arc AB is also spaced from O at a distance less than r. (For $K > 0$, it is assumed that $r < \frac{\pi}{2\sqrt{K}}$.)*

Proof. Consider the triangle OAB and its corresponding triangle $O'A'B'$ on a K-plane. Take a point X on AB and the corresponding point X' on $A'B'$, i.e., a point X' such that $A'X' = AX$. Then, as a consequence of Proposition 2, $O'K'$, is not greater than the greater of the two sides $O'A'$ and $O'B'$. For $K \leqslant 0$ this assertion holds for every triangle on a K-plane. For $K > 0$, it is definitely satisfied if the greater of the sides $O'A'$ and $O'B'$ is not greater than $\frac{\pi}{2\sqrt{K}}$, i.e., not greater than the inner radius (half the meridian) of the hemisphere which in this case is a K-plane. Therefore,

$$OX' \leqslant O'X' \leqslant O'B' = OB = r.$$

This completes the proof.

3.8. The definition of the domain R_K requires that the relative excesses of an arbitrary triangle be nonpositive. It would be desirable to weaken this condition so that nonpositivity of the relative excesses is required only for sufficiently small triangles. Obviously, this requirement is necessary but not sufficient: for instance, a closed cylinder has nonpositive curvature in this sense; nevertheless, it is not an R_0 domain because it contains points which can be joined by two shortest arcs. However, requiring additionally that a shortest arc be a continuous function of its endpoints, we obtain the conditions that define the domain R_K.

This assertion is the content of Proposition 9 formulated below. Furthermore, we may add that in a compact domain the uniqueness of a shortest arc joining two points guarantees the continuity of the shortest arc as a function of its endpoints.[8]

Proposition 9. *Let a domain G of a metric space satisfy three conditions:*

(1) *every two points in G can be joined by a unique shortest arc;*

[8]In a compact domain every sequence of arcs with uniformly bounded lengths includes a converging subsequence. If, A_n tends to A and B_n to B, then the limit of every converging subsequence of the shortest arcs A_nB_n is the shortest arc AB. Therefore, if this limit is unique, then A_nB_n tends to AB.

(2) *the shortest arc depends continuously on its endpoints, i.e., if* A_n *tends to* A *and* B_n *to* B, *then* A_nB_n *tends to* AB; *and*
(3) *every point has a neighborhood in which the relative excess of every triangle relative to* K *is nonpositive.*

Then the relative excess of every triangle in G *is also nonpositive.*

Proof. Since the conditions defining the domain R_K, by the assumptions of the proposition, are satisfied in every neighborhood of a point, all the propositions concerning the domain R_K which were formulated earlier also hold in such a neighborhood. Therefore, there exists an angle between every two shortest arcs in the domain G. Furthermore, every angle of a "sufficiently small" triangle is not greater than the corresponding angle of the corresponding triangle on a K-plane (Propositions 3 and 4). We make use of these two properties.

Now let $T = ABC$ be an arbitrary triangle in the domain G. We prove that every angle of the triangle ABC, say, the angle α at the vertex A, is not greater than the corresponding angle of the triangle T^K with sides of the same length as the triangle ABC. This would complete the proof of our assertion.

On BC let us take the points $D_0 = B, D_1, D_2, \dots, D_n = C$ and draw the shortest arcs $AD_0, AD_1, \dots, AD_n$. Thus we obtain a series of "narrow" triangles $T_i = AD_{i-1}D_i$. Since the shortest arc AD, by condition (2), depends continuously on the point D, the adjacent shortest arcs AD_{i-1} and AD_i lie sufficiently close to each other if the points D_i are packed densely enough on BC. Let us take the points $E_{i_1}, E_{i_2}, \dots, E_{i_m}$ on the shortest arc AD_i. Joining these points as shown in Fig. 10, we obtain "small" triangles T_{ij}.

FIG. 10

Sending each triangle T_{ij} into a corresponding triangle T_{ij}^K on a K-plane, we obtain the development Q consisting of the triangles T_{ij}^K. The development Q also consists of "narrow" polygons Q_i corresponding to the triangles T_i (i.e., Q_i is composed of the triangles $T_{i_1}^K, T_{i_2}^K, \dots$). If the points D_i and E_i are packed densely enough, then the triangles T_{ij} would be quite small so that each of their angles is not greater than the corresponding angle of the triangles T_{ij}^K.

There are three types of angles of a triangle T_{ij} (and accordingly of the triangle T_{ij}^K):

(1) the angles α_i and accordingly the angles α_i^K, at the vertex A,
(2) the angles at the vertices E_{ij} inside the shortest arc AD_i, and
(3) the angles at the vertex D_i.

If α is the angle at the vertex A of the initial triangle T, then by virtue of Proposition 1 of § 2, we find $\alpha \leqslant \sum \alpha_i$. However, since $\alpha_i \leqslant \alpha_i^K$, we obtain

$$\alpha < \sum \alpha_i^K. \tag{1}$$

Furthermore, the sum of the angles at a vertex of adjacent triangles T_{ij} lying on the shortest arc AD_i or AD_{i-1} that are inscribed in the narrow triangle T_i is not less than π (as a consequence of Propositions 1 and 2 of § 2). Consequently, the sum of angles at a vertex of adjacent triangles T_{ij}^K is not less than π. This means the broken lines $A'O_i'$ and $A'D_{i-1}'$ that bound the polygon Q_i are convex and are bent inward into this polygon.

Therefore we can apply Lemma 2. By this lemma the angles of the polygon Q_i are not greater than the angles of the triangle T_i^K which is obtained by straightening out the broken lines bounding the polygons Q_i. (Each such triangle T_i^K is obviously none else than the triangle corresponding to the narrow triangle T_i and having, consequently, sides of the same length as T_i.) Let α_i denote the angle of the triangle T_K at the vertex corresponding to A; then, we have $\alpha_i^K \leqslant \bar{\alpha}_i^K$. Hence, (1) yields

$$\alpha < \sum \bar{\alpha}_i^K. \tag{2}$$

Moreover, the sum of the angles of the small triangles adjacent to a side lying inside the side BC is not less than π (this is also a consequence of Propositions 1 and 2 of § 2). Therefore the sum of the corresponding angles of the triangles T_{ij}^K is all the more not less than π and, consequently, the sum of the angles of the triangles T_i^K and T_{i+1}^K adjacent to the vertex D_i is not less than π.

This means the polygon P obtained by joining the triangles T_i^K to one another is bounded by the convex broken line $B''C''$ bent inward into P. Obviously, straightening out this broken line $B''C''$, we obtain the triangle

T^K which corresponds to the initial triangle T (and, consequently, has sides of the same length as T). Now, applying Lemma 2, we find that the angle at the vertex A'' of the polygon P is not greater than the corresponding angle of the triangle T^K, i.e.,

$$\sum \bar{\alpha}_i^K \leqslant \alpha^K.$$

Comparing this inequality with (2), we find $\alpha \leqslant \alpha^K$. This completes the proof.

3.9. In the end of this section we demonstrate the equivalence of the two definitions given in § 5 of Section 1 for a space of curvature $\leqslant K$.

Proposition 10. *For a metric space to be a space of curvature $\leqslant K$ in the sense that every point has a neighborhood R_K, it is necessary and sufficient that the following two conditions hold:*

(1) *every point has a neighborhood in which every two points can be joined by a unique shortest arc; and*

(2) *for every sequence of triangles T converging to a point,*

$$\overline{\lim}\, \frac{\delta_0(T)}{S(T^0)} \leqslant K, \tag{3}$$

where $\delta_0(T) = \alpha + \beta + \gamma - \pi$ is the excess of the triangle T and $S(T^0)$ is the area of the corresponding triangle T^0 on the Euclidean plane.

(It is assumed that $\delta_0(T) \leqslant 0$, if $S(T^0) = 0$, at least for a sequence of a sufficiently large number of triangles so that the expression under the limit sign in (3) is equal to minus infinity or zero, therefore it is less than or equal to K.)

Proof. Necessity. Obviously, it only suffices to show the necessity of the second condition. For this purpose, let us note that the relative excess

$$\delta_K(T) = (\alpha + \beta + \gamma) - (\alpha^K + \beta^K + \gamma^K)$$

obeys the equality

$$\delta_K(T) = (\alpha + \beta + \gamma - \pi) - (\alpha^K + \beta^K + \gamma^K - \pi) = \delta_0(T) - \delta_0(T^K), \tag{4}$$

i.e., the relative excess of the triangle T is equal to the difference between the "absolute" excesses of the triangles T and T^K.

We know that the excess $\delta_0(T^K)$ of a triangle on a K-plane is proportional to its area $s(T^K)$, i.e.,

$$\delta_0(T^K) = K\, s(T^K). \tag{5}$$

From (4) and (5) it follows that the condition

$$\delta_K(T) \leqslant 0$$

is equivalent to the condition

$$\delta_0(T) \leqslant K\, s(T^K). \tag{6}$$

If the triangle T^K is small, then the relative difference between the areas of the triangles T^K and T^0 is small; and, in particular, $S(T^K) = 0$, if and only if $s(T^0) = 0$. Therefore we can assert that

$$S(T^K) = A(T)\, s(T^0)$$

and $A(T)$ tends to one as the sides of the triangle T tend to zero.

Therefore, by virtue of (6), we can write

$$\delta_0(T) < K\, s(T^0)\, A(T). \tag{7}$$

From this inequality, it follows that for every sequence of triangles converging to a point, since A tends to one, we have

$$\overline{\lim}\, \frac{\delta_0(T)}{s(T^0)} \leqslant K.$$

Here it is admitted that $s(T^0) = 0$, then, as a consequence of (7), we obtain $\delta_0(T) \leqslant 0$.

Sufficiency. The second condition of the proposition implies that if $K' > K$, then for an arbitrary point O, there exists a neighborhood U in which for every triangle T, we have

$$\delta_0(T) \leqslant K' S(T^0). \tag{8}$$

By virtue of the first condition, the neighborhood U can be chosen such that every two points in U can be joined by a unique shortest arc. We have[9]

$$\begin{aligned} S(T^0) &> s(T^{K'}) \quad \text{for } K' < 0, \\ S(T^0) &< s(T^{K'}) \quad \text{for } K' > 0. \end{aligned}$$

Therefore, from (8) we obtain

$$\delta_0(T) < K' s(T^{K'}).$$

[9]In this connection we note that this result is contained in Proposition 1 (of § 5) concerning the area of a triangle in R_K.

However, since $K's(T^{K'}) = \delta_0(T^{K'})$, it follows that

$$\delta_{K'}(T) = \delta_0(T) - \delta_0(T^{K'}) < 0,$$

i.e., in the neighborhood U the relative excess of a triangle relative to K is nonpositive.

Since all the assertions for an arbitrary $K' > K$ hold in a neighborhood of each point, we can say that our space is a space of curvature $\leqslant K'$ for an arbitrary $K' > K$. Hence it follows that our space is a space of curvature $\leqslant K$. In order to verify this assertion, let us return back to the neighborhood U of the point O. Let us take a neighborhood V of the point O such that every two points in V can be joined by a shortest arc lying in U.

Then, by virtue of what has been proved above, the excess of a triangle with vertices in K' is nonpositive. This implies that V is an R_K domain and, in particular, a shortest arc in this neighborhood depends continuously on its endpoints.

Thus the first two conditions of Proposition 9, namely, the existence of a shortest arc and the continuity of a shortest arc as a function of its endpoints, are satisfied in the neighborhood V. Moreover, by virtue of the assertion proved above, every point in V, for an arbitrary $K' > K$, has a neighborhood in which the excess of a triangle relative to K' is nonpositive. This means that for a $K' > K$ the third condition of Proposition 9 is also satisfied in the neighborhood V.

By Proposition 9, for an arbitrary $K' > K$, for every triangle T in V, we also have

$$\delta_{K'}(T) \leqslant 0.$$

Since this is true of every $K' > K$, obviously

$$\delta_K(T) \leqslant 0.$$

(In fact, $\delta_K(T) = (\alpha+\beta+\gamma) - (\alpha^{K'} + \beta^{K'} + \gamma^{K'})$ and for a given triangle T the angles $\alpha^{K'}$, $\beta^{K'}$ and $\gamma^{K'}$ obviously tend to α^K, β^K and γ^K, respectively, as K' tends to K; i.e., $\delta_{K'}(T)$ tends to $\delta_K(T)$.) Hence, the neighborhood V is an R_K domain and, since every point O has a neighborhood of this kind, our space is a space of curvature $\leqslant K$.

4. The Direction of a Curve and the Angle of the Direction Cone

4.1. The direction of a curve. The following definition of "direction" is sound in every metric space.

We say that a curve starting from a point O has a *definite direction* at this point, if the upper angle which it forms with itself is zero.

From the definition of shortest arc we readily obtain

Proposition 1. *Every shortest arc has a definite direction at its initial point.*

We can prove that in a Riemannian space the existence of a direction as defined above is equivalent to the existence of a geodesic tangent.

We say that two curves L_1 and L_2 starting from a common point have a *common direction* at this point, if the upper angle between them is zero. Then each curve has a definite direction because, by Proposition 1 of Section 2, we have

$$\alpha_{12} + \alpha_{21} \geqslant \alpha_{11},$$

and therefore for $\alpha_{12} = \alpha_{21} = 0$, the angle α_{11} vanishes.

Lemma 1. *If each of two curves L_1 and L_2 starting from a sole point has a common direction with a third curve L_3, then they have a common direction with one another.*

Indeed, since

$$\alpha_{12} \leqslant \alpha_{13} + \alpha_{23},$$

then $\alpha_{13} = \alpha_{32} = 0$ implies $\alpha_{12} = 0$.

Lemma 2. *If two curves L_1 and L_2 have a common direction, then for a third curve L_3 the upper angle α_{13} between L_1 and L_3 equals the upper angle α_{23} between L_2 and L_3.*

Proof. By Proposition 1 of Section 2, we have $\alpha_{13} + \alpha_{12} \geqslant \alpha_{23}$, so $\alpha_{13} \geqslant \alpha_{23}$ if $\alpha_{12} = 0$. Similarly, $\alpha_{23} \geqslant \alpha_{13}$. Therefore, $\alpha_{13} = \alpha_{23}$. This completes the proof of our lemma.

By virtue of Lemma 1, the set of curves starting from a point O and having a definite direction at O decomposes into groups of curves such that the curves in a group have a common direction. This gives a method for defining the concept of direction at a given point without taking recourse to a curve. And on the basis of Lemma 2, we can speak about the upper angle or simply the angle between two directions at a point.

If α_{12}, α_{23} and α_{31} are the angles between the directions D_1, D_2 and D_3, then, by Proposition 1 of Section 2, it immediately follows that

$$\alpha_{12} + \alpha_{23} \geqslant \alpha_{13}.$$

Hence, we have

Proportion 2. *The directions at a given point form a metric space in which the distance between two directions is defined by the upper angle between them.*

4.2. Direction in the domain R_K. In the domain R_K every shortest arc has a definite direction at the initial point, and there exists a continuous dependence between shortest arcs and their directions. In other words, if the shortest arcs L_n starting from a point O converge to a shortest arc L, then their directions converge to the direction of the shortest arc L. If in R_K the point O has a neighborhood homeomorphic to a ball, then a shortest arc emanates definitely from O into each direction. Moreover, there may exist continuum shortest arcs emanating from O, having a given direction and failing to be a part of one another. These assertions are contained in the following propositions.

Proposition 3. *In R_K if the shortest arcs L_n and M_n starting from a point O converge to the shortest arcs L and M, respectively, then the angles $\alpha(L_n, M_n)$ converge to the angle $\alpha(L, M)$. In particular, if L_n tend to L, then the angles $\alpha(L_n, L)$ tend to zero.*

It suffices to put $L = M_1 = M_2 = \dots$. Since, by definition, the angle between the directions of shortest arcs is equal to the angle between the shortest arcs themselves, we find this statement asserts the convergence of the directions of a sequence of converging shortest arcs.

Proof. Consider the shortest arcs starting from a given point O in the domain R_K. Let the shortest arcs L_n tend to a shortest arc L. Let us take equi-spaced points A_n and A on the arcs L_n and L, respectively, such that $OA_n = OA = a$. If, now, $\gamma_n(a, a)$ denotes the angle of the triangle on a K-plane that corresponds to the triangle OAA_n, then, by Proposition 1 of Section 3, we find

$$\gamma_n(a, a) \geqslant \alpha(L, L_n).$$

However, as the arcs L_n tend to L, we find $\gamma_n(a, a)$ tends to zero. Therefore

$$\alpha(L, L_n) \to 0.$$

Now let the arcs L_n tend to the arc L and the arcs M_n to M. From the inequality governing the angles of a triangle, we obtain

$$|\alpha(L, M) - \alpha(L_n, M_n)| \leqslant \alpha(L, L_n) + \alpha(M, M_n).$$

From what has been proved for the angle $\alpha(L, L_n)$, it follows that the angle $\alpha(M, M_n)$ tends to zero, so we have

$$\alpha(L_n, M_n) \to \alpha(L, M).$$

This completes the proof.

Lemma 3. *If a curve starting from a point O in the domain R_K has a direction at O, then the upper angle between the curve L and its secants OX, i.e., the shortest arcs $OX, X \in L$, tends to zero as X tends to the point O. In other words, the direction of a curve is the limit of the directions of its secants OX.*

Proof. By the definition of upper angle, we have

$$\alpha(L, OX) = \overline{\lim_{Y,Z\to 0}}\, \gamma(Y, Z),$$

where $Y \in L$ and $Z \in OX$. And, since by Proposition 1 of Section 3, the angle γ is a nondecreasing function of OZ, the following inequality holds

$$\gamma(Y, Z) \leqslant \gamma(Y, X).$$

Consequently, we have

$$\alpha(L, OX) \leqslant \overline{\lim_{Y\to 0}}\, \gamma(Y, X).$$

However, as the arc has a direction, we find

$$\overline{\lim_{X,Y\to 0}}\, \gamma(Y, X) = 0,$$

consequently,

$$\overline{\lim_{Y\to 0}}\, \alpha(L, OX) = 0.$$

Proposition 4. *If a point O in R_K has a neighborhood homeomorphic to a ball, then a shortest arc can be drawn in every direction from O.*

Proof. Let D be a certain direction at O and L an arc starting from O in the given direction D. On L take the points $X_n \to 0$ and draw the shortest arcs OX_n. By Proposition 7 of Section 3 each of these arcs may be prolonged to a shortest arc of length r. From the shortest arcs M_n thus obtained, let us choose a convergent subsequence. The limit arc is a shortest arc that has the given direction D at the point O because, by Proposition 3, the directions of the shortest arcs M_n converge to the direction D_M of the limit shortest arc M. On the other hand, by Lemma 3, they converge to the direction D of the arc L. Therefore D_M coincides with D.

Remark 1. The following simple example shows that the condition of Proposition 4 is essential. The usual (closed) disk on a plane is evidently a space of curvature $\leqslant 0$. However, no shortest arc starts from a point on the boundary in the direction of the bounding curve.

2. From the proof of Proposition 4 we immediately derive the following: If a point O in R_K has a neighborhood homeomorphic to a ball and a shortest arc starting from O has the direction L, then the L has a tangential shortest arc because every sequence of secants includes a subsequence converging to the shortest arc of the same direction as L. However, as shown by simple examples, this tangential arc is not unique. In general, there is a continuum of tangential shortest arcs in a given direction.

Proposition 5. *If the curves L and M starting from O in R_K each has a direction at the point O, then there exists an angle between them; and it is equal to the limit of the angles between the secants. If X and Y are points on L and M, then the following equality holds:*

$$\alpha(L, M) = \lim_{Y\to O} \alpha(L, OY) = \lim_{X\to O} \alpha(OX, M) = \lim_{X,Y\to O} \alpha(OX, OY).$$

(Assuming, in particular, $M = L$, we obtain the assertion of Proposition 3, namely, for $X \in L$, we always find that $\alpha(L, CX) \to 0$ as $X \to O$.)

Proof. Let $\bar{\alpha}(L, M)$ be the upper angle between L and M. By definition,

$$\bar{\alpha}(L, M) = \overline{\lim_{X,Y\to O}}\, \gamma(X, Y). \tag{1}$$

On the other hand, by the triangle inequality, for the upper angles, we have

$$\bar{\alpha}(L, M) \leqslant \alpha(OX, OY) + \bar{\alpha}(L, OX) + \bar{\alpha}(M, OY).$$

However, since by Lemma 3,

$$\lim \bar{\alpha}(L, OX) = \lim \bar{\alpha}(M, OY) = 0,$$

we obtain

$$\bar{\alpha}(L, M) \leqslant \underline{\lim_{X,Y\to O}}\, \alpha(OX, OY). \tag{2}$$

Furthermore, by Proposition 1 of Section 3,

$$\alpha(OX, OY) \leqslant \gamma(X, Y). \tag{3}$$

Comparing (1), (2) and (3), we obtain

$$\alpha(L, M) = \lim \gamma(X, Y) = \lim \alpha(OX, OY),$$

i.e., first, the limit $\lim \gamma(X, Y)$ exists, in other words, there exists an angle between L and M; second, it is equal to the limit of the angles $\alpha(OX, OY)$. The other equalities of the proposition are proved on similar lines.

4.3. The angle of the direction cone. Let us consider the space consisting of the directions at a given point O, O being a point in some metric space. The cone C formed by the directions D at O is defined to be a curve in the space of directions, i.e., it is given by a continuous function $D(t)$, $a \leqslant t \leqslant b$.

The angle of such a cone is determined by the length of this curve in the metric of angles, i.e., the angle β of the cone $C = [D(t), a \leqslant t \leqslant b]$ is

$$\beta(C) = \sup \sum_{i=1}^{n} \alpha[D(t_{i-1}), D(t_i)],$$

where α is the angle between the directions and $a = t_0 < t_1 < \cdots < t_n = b$.

Since every shortest arc starting from O has a direction at O, the cone formed by the shortest arcs $L(t)$ ($a \leqslant t \leqslant b$) starting from the point O defines a cone of directions. Then the angle of the cone of shortest arcs at O is the angle of the direction cone (in a two-dimensional manifold the angle of this cone is essentially the same as the angle of a sector as defined in [1]). However, even if at least one shortest arc starts from O in each direction, then the direction cone does not in general uniquely determine the cone of shortest arcs, namely, not in every direction a unique shortest arc starts from O. However, if only one shortest arc starts from O in each direction and the shortest arcs depend continuously on the directions, then, just like in R_K, every direction cone defines the cone of shortest arcs.

For two given directions D_1 and D_2 at a point in some metric space, the *shortest cone* is defined as usual in the same way as a shortest arc joining two points, as the direction cone joining the directions D_1 and D_2 which has the least angle. Such a "shortest path" resembles a plane sector, thus giving us a definition of plane direction in an arbitrary metric space. Naturally, such a cone need not necessarily exist in an arbitrary metric space, but if it does exist, its angle may prove to be larger than the angle between the directions D_1 and D_2.

However, as demonstrated by Proposition 7 proved in this section, this is excepted in R_K at least for those directions in which shortest arcs start from a given point. First we would like to study the cone composed of all shortest arcs joining a given point A with the points on a given shortest arc BC.

4.4. In the R_K domain let us take an arbitrary triangle ABC and join the point A with all the points on the side BC by shortest arcs AX. By Proposition 6 of Section 3, the shortest arc AX, when the endpoint X is displaced continuously, varies continuously. So the shortest arcs AX form a special surface.[10] We call this surface the *surface triangle ABC spanned*

[10]Note that different shortest arcs AX may coincide, even in simple cases, over some length as shown in Fig. 11.

from the vertex A. This is the *cone of the shortest arcs* AX.

Such a surface triangle can be represented as a continuous image of a plane triangle as follows. Take a plane triangle $A'B'C'$ with sides of the same length as the sides of the triangle ABC. Send a point X on BC into a point X' on $B'C'$ such that $B'C' = BX$; then the shortest arc AX corresponds to the segment $A'X'$. Now send a point Y on AX into a point Y' on $A'X'$ such that $A'Y' : A'X' = AY : AX$. In this way the triangle $A'B'C'$ is mapped onto the triangle ABC. This mapping is continuous because a shortest arc is a continuous function of its endpoints (Proposition 6 of Section 3). The mapping so defined we call the *canonical mapping*.

Since the shortest arcs AX form a continuous set, their directions at the point A, by Proposition 3, also form a continuous set.

The angle of the sector formed by the shortest arcs AX is, by the definition given above, the least upper bound of the sum of the angles between the shortest arcs AX_{m-1} and AX_m which consecutively join the point A with the points $B, X_1, \ldots, X_n, C$. It is natural to call this angle the *angle at the vertex* A of the surface triangle.

FIG. 11

4.5. Proposition 6. *Let* $T = ABC$ *be the surface triangle spanned from the vertex* A *and* $T^K = A'B'C'$ *a triangle on a* K*-plane with sides of the same length as the triangle* ABC*. If* α *and* α_K *are the angles at the vertices* A *and* A' *of the triangles* T *and* T^K*, respectively, in the sense of the definition given for the angle of the surface triangle, then*

$$\alpha \leqslant \alpha_K \tag{4}$$

and $\alpha = \alpha_K$ *if and only if the triangle* T *is isometric to* T^K.

(Since the angle of a surface triangle is obviously not less than the angle between its sides, inequality (4) makes the Proposition 4 of Section 3 stronger. If the angle between the sides AB and AC is α_K then, from what has been said above, we have $\alpha = \alpha_K$; but the proposition asserts that in this case the triangles T and T^K are isometric. Therefore the necessary condition for the angles between the sides of the triangles ABC and $A'B'C'$ to be equal is contained here.)

We now prove inequality (4). Take an arbitrary $\varepsilon > 0$. Take a set of consecutive points $X_0 = B, X_1, X_2, \ldots, X_n, X_{n+1} = C$ on the side BC of the triangle ABC. Let $\xi_0, \xi_1, \ldots, \xi_n$ be the angles between AB and AX_1, AX_1 and AX_2 and so on (Fig. 12).

Fig. 12

The points X_i can be taken close enough such that

$$\alpha - \varepsilon < \sum_{i=0}^{n} \xi_i. \tag{5}$$

On a K-plane let us construct the triangles $T_i^K = A'X_i'X_{i+1}'$ with sides of the same length as the triangle $T_i = AX_iX_{i+1}$. Let ξ_i^K denote the angle at the vertex A' of the triangles T_i^K. Then, by Proposition 4 of Section 3, we have

$$\xi_i^K \geqslant \xi_i. \tag{6}$$

Joining the triangles T_i^K to one another in the same order in which the triangles T_i occur, we obtain a polygon $P = A'B'X_1' \ldots X_n'C'$. The angle

at the vertex A' of the polygon is $\sum \xi_i^K$, and, by virtue of (5) and (6), the following inequality holds:

$$\sum \xi_i^K > \alpha - \varepsilon. \tag{7}$$

At the same time, the angles at the vertices X_i' of the polygon P are not less than π. In fact, each such angle consists of the angles η_i^K and ζ_{i+1}^K of two adjacent triangles T_i^K and T_{i+1}^K. By Proposition 4 of Section 3, these angles are not less than the corresponding angles η_i and ζ_{i+1} of the triangles T_i and T_{i+1}, respectively,

$$\eta_i^K \geqslant \eta_i, \quad \eta_{i+1}^K \geqslant \eta_{i+1}.$$

Furthermore, the angles η_i and ζ_{i+1} are adjacent angles and therefore $\eta_i + \zeta_{i+1}^K \geqslant \pi$. Consequently, $\eta_i^K + \zeta_{i+1}^K \geqslant \pi$.

Further, Lemma 1 is applicable to our polygon: on unfolding it into a triangle, we increase the angle at the vertex A'. And the result is none else than the triangle $T^K = A'B'C'$ and we are therefore dealing with its angle α_K. Now the angle at the vertex A of the polygon P is $\sum \xi_i^K$, so

$$\sum \xi_i^K \leqslant \alpha_K.$$

Along with (7), we obtain

$$\alpha_K > \alpha - \varepsilon$$

and, ε being arbitrary, we find

$$\alpha_K \geqslant \alpha.$$

4.6. Now we prove that if $\alpha_K = \alpha$, the triangles T and T^K are isometric. Take an arbitrary point D on BC (Fig. 13) and consider the surface triangles $T_1 = ABD$, $T_2 = ACD$ and their corresponding triangles T_1^K and T_2^K on a K-plane.

If $\alpha', \alpha'', \alpha_K'$ and α_K'' are the angles of these triangles, then

$$\alpha' + \alpha'' = \alpha, \tag{8}$$

$$\alpha_K' \geqslant \alpha', \quad \alpha_K'' > \alpha''. \tag{9}$$

Moreover, by Lemma 2 of Section 3, the angle δ_K at the vertex D' in the quadrilateral composed of the triangles T_1^K and T_2^K is greater than or equal to π, i.e.,

$$\delta_K \geqslant \pi.$$

Fig. 13

Therefore, by Lemma 1 of Section 3, we have

$$\alpha'_K + \alpha''_K \leqslant \alpha_K, \tag{10}$$

and $\alpha'_K + \alpha''_K = \alpha_K$ if and only if $\delta_K = \pi$. However, from (8), (9) and (10), it follows that

$$\alpha_K \geqslant \alpha'_K + \alpha''_K \geqslant \alpha.$$

Then the equality $\alpha = \alpha_K$ implies $\alpha'_K + \alpha''_K = \alpha_K$, so we have $\delta_K = \pi$, i.e., the quadrilateral $T_1^K + T_2^K$ coincides with the triangle T_K.

This means, first, that $AD = A'D'$ for the point D' on $B'C'$ for which $B'D' = BD$. Second, the angle of the sector between AB and AD is equal to the angle between $A'B'$ and $A'D'$. Since D is an arbitrary point on BC, for the canonical mapping of the triangle T^K onto T, we have

$$AX = A'X'$$

and the angle (AB, AX) is equal to the angle $(A'B', A'X')$.

Now let P and Q be two arbitrary points lying on certain shortest arcs AX and AY, respectively $(X, Y \in BC)$, of the triangle T. Let P', Q', X' and Y' be the corresponding points in the triangle T^K. From what has been proved above, the triangles AXY and $A'X'Y'$ have equal sides and equal angles at the vertices A and A'. Therefore, by Proposition 2 of Section 3, it follows that $PQ = P'Q'$. Consequently, the triangle T is isometric to the triangle T^K (in the sense of the metric in the domain R_K and, of course, in the sense of the general intrinsic metric as well).

4.7. Proposition 7. *Let O be a point in R_K, and L and M two shortest arcs which start from O. If the angle $\alpha(L, M)$ between L and M is less than or equal to π, then there exists a cone of shortest arcs joining L and M with an angle differing from $\alpha(L, M)$ by an arbitrarily small quantity.*

Furthermore, if the point O has a neighborhood homeomorphic to the ball, there exists a cone, joining L and M, of shortest arcs starting from O with the angle equal to $\alpha(L, M)$.

Proof. Let X and Y be points on L and M. If $\alpha(L, M)$ is less than π, then the shortest arc XY does not pass through O, even if X and Y lie arbitrarily close to O, and does not therefore lie wholly on $L + M$. Consequently, the shortest arcs OZ joining the point O with the points $Z \in XY$ form a cone C_{XY} between L and M that is at the same time the surface triangle OXY spanned from O.

By Proposition 6, the angle $\beta(C_{XY})$ of the cone C_{XY} (of the surface triangle) satisfies the inequality

$$\beta(C_{XY}) \leqslant \gamma(X, Y).$$

However, obviously, we have

$$\beta(C_{XY}) \geqslant \alpha(M, N),$$

and, by definition,

$$\alpha(L, M) = \lim_{X,Y \to 0} \gamma(X, Y).$$

Hence, it follows that

$$\alpha(L, M) = \lim \beta(C_{XY}).$$

This completes the proof of the first assertion of our proposition.

If the point O has a neighborhood homeomorphic to the ball; then, first, each shortest arc starting from O prolongs to a shortest arc of a certain fixed length $r > 0$. Therefore, the cone can be included into a cone composed of shortest ars of length at least r. Second, from these cones we can choose a converging subsequence of cones. The angle of the limit cone is equal to $\alpha(L, M)$.

4.8. Applying Proposition 6, we can prove the following

Proposition 8. *If a point O in R_K has a neighborhood U homeomorphic to the n-dimensional ball such that the excess (relative to K) of every*

triangle in the neighborhood U is zero, then the neighborhood U of the point O is isometric to a domain in an n-dimensional space of constant curvature.

Proof. Let a point O in R_K have a neighborhood U satisfying the conditions of our proposition. We can naturally assume that $n \geqslant 2$. By Proposition 6, every surface triangle in U is isometric to a plane triangle, i.e., a triangle on a K-plane. This remark underlies our proof.

Furthermore, by Proposition 7 of Section 3, every shortest arc starting from O prolongs beyond the point O. This prolongation is unique, else we would obtain a triangle with two sides partially overlapping one another. Such a triangle is definitely not isometric to a plane triangle.

Let L_1 and L_2 be two shortest arcs starting from O, none of which being the prolongation of the other. Let L_1' and L_2' be their prolongations beyond O. Then each of the four angles between L_1 and L_2, L_1' and L_2', etc. is less than π. Were the angle between, say, L_1 and L_2 equal to π, then L_1, as can be easily shown, would be the prolongation of L_2, so the prolongation of L_2 would not be unique. Taking the points A_1, A_2, A_1' and A_2' on the shortest arcs L_1, L_2, L_1' and L_2', respectively, let us construct four surface triangles OA_1A_2, OA_2A_1' and so on (Fig. 14).

FIG. 14

By what has been proved above, their angles at the point O are less than π, so none of the triangles degenerates into a line. Consequently, by Proposition 6, each of these triangles is isometric to a plane triangle. Moreover, with prolongation of a shortest arc being unique, it is a simple matter to verify that the sum of every pair of adjacent angles at the vertex O of these triangles is π.

Hence, it follows that the triangles so constructed form a surface Q^2 isometric to a quadrilateral on a K-plane.

If the neighborhood U of the point O is two-dimensional ($n = 2$), then the plane surface Q^2 is the neighborhood U of the point O (because the neighborhood U of the point O is, by our proposition, homeomorphic to the ball, i.e., a disk in case $n = 2$). If, on the contrary, $n > 2$; then the surface Q^2 is not a neighborhood of the point O, so we can draw a shortest arc L_3 from O prolonging beyond the surface Q^2. Let L_3' be the prolongation of this shortest arc beyond the point O. Take the points A_3 and A_3' on L_3 and L_3', respectively. The shortest arcs L_1, L_2 and L_3 together with their prolongations, form a coordinate triple with origin at O (Fig. 15). On these curves take the points A_1, A_2, A_3 etc. and construct tetrahedra with a common vertex at O and with other vertices at the points A_1, etc. Construct the surface triangle $A_1A_2A_3$; it is isometric to a triangle on a K-plane and, consequently, it is unique, regardless of the vertex from which it is spanned. Drawing shortest arcs from O to the points on the triangle $A_1A_2A_3$, we obtain the tetrahedron $T = OA_1A_2A_3$ isometric to a tetrahedron in a space of constant curvature K.

To prove this assertion, in a space of constant curvature K, let us construct a tetrahedron $T^K = O''A_1''A_2''A_3''$ with edges of the same length as the tetrahedron T. Its faces are isometric to the faces of the tetrahedron T. By virtue of the isometry between the bases $A_1A_2A_3$ and $A_1''A_2''A_3''$ there exists a correspondence between the shortest arcs OX and $O''X''$ joining the vertices O and O'' with the corresponding points on the bases. We want to prove that these shortest arcs are of equal length, i.e., $OX = O''X''$ (Fig. 15).

For this purpose, let us draw a plane through $O''X''$ and $O''A_1''$; it intersects the tetrahedron T^K alone a certain triangle $O''A_1''B''$. Because of the isometry between the faces of the tetrahedra T and T^K, to the triangle $O''A_1''B''$ there corresponds a triangle OA_1B with sides of the same length as the sides of the triangle $O''A_1''B''$. Consequently, the surface triangle OA_1B is isometric to the triangle $O''A_1''B''$. Hence $OX = O''X''$. Therefore the tetrahedra T and T^K are formed by equal segments OX and $O''X''$ drawn between the corresponding points of the isometric triangles $A_1A_2A_3$ and $A_1''A_2''A_3''$. Now we may assert that these tetrahedra are isometric.

The reasoning we applied to the tetrahedron $T = OA_1A_2A_3$ can also be applied to each of the eight tetrahedra $OA_1A_2A_3$, $OA_1'A_2'A_3'$, etc. Therefore the point O is surrounded by eight tetrahedra which are isometric to the tetrahedra in a space of constant curvature K. Furthermore, it is a simple matter to prove that these tetrahedra form, together with the figure Q^3, an isometric octahedron of constant curvature K.

If the dimension n of the neighborhood U of the point O is 3, then

Fig. 15

the octahedron Q^3 is a neighborhood of the point O. On the contrary, if $n > 3$, let us draw a shortest arc L_4 from O passing beyond Q^3 and draw its prolongation L'_4. Take the points A_4 and A'_4 on L_4 and L'_4, respectively. Thus we obtain a coordinate quadruple formed by shortest arcs and their prolongations. Now consider the simplex $OA_1A_2A_3A_4$, etc. (in all there are sixteen such simplexes) and prove as before that they are isometric to the simplexes in a space of constant curvature K (of course, first we demonstrate this for the bases $A_1A_2A_3A_4$ and so on).

Continuing this construction till we obtain the neighborhood U of the point O, we find that U is isometric to a domain in an n-dimensional space of constant curvature K.

5. The Area of a Surface and the Isoperimetric Inequality in R_K

5.1. The area of a surface: its definition. In a metric space let us consider a surface F which is the image of a plane Jordan domain D under a certain continuous mapping f. Let M be a triangulated polygon or, in other words, a complex of triangles included in D. Take an arbitrary triangle T_i of the complex M. Its vertices A, B, and C correspond to the points $f(A), f(B)$, and $f(C)$ on the surface F. Construct a plane triangle T_i^0 with sides equal to the distances between the points $f(A), f(B)$ and $f(C)$. Let $S(T_i^0)$ be the area of this plane triangle. Let us evaluate the sum

of the areas $S(T_i^0)$ of all triangles T_i^0 corresponding to the triangles T_i:

$$S_0(M) = \sum_i S(T_i^0).$$

We define the *area of a surface* that is the image of the domain D under a continuous mapping as the greatest lower bound of the quantity $S_0(M)$, provided that the vertices of the complex M condense indefinitely in the domain D and the polygon M gradually fills the interior of the domain D

$$S(F, D, f) = \underline{\lim}\, S_0(M).$$

A surface can be generated by different mappings f of different domains D. In order to answer the question whether the area of a surface depends on its representation, it suffices to define the area of a surface F by the formula

$$S(F) = \inf S(F, D, f),$$

i.e., as the greatest lower bound of $S(F, D, f)$ for all mappings f of the domain D, corresponding to the one and the same surface.

This definition is in full analogy with the definition of the area of a surface as the greatest lower bound of inscribed polygons.

In the definition we used plane triangles T_i^0, but we could as well have used the triangles T_i^K in an arbitrary K-plane. Namely, if the sides of the triangles T_i^0 tend to zero, then the ratio of the areas of the triangles T_i^K and T_i^0 with equal sides tends to one, and therefore

$$\underline{\lim} \sum S(T_i^0) = \underline{\lim}\, S(T_i^K).$$

This remark is of great help in studying the area of a surface in R_K; in our case it also permits us to use the triangles on a K-plane.

For the complex M encountered in the definition of the area of a surface F, it is desirable to relate the complex M with the following construction.

To every triangle $T_i = ABC$ of the complex M there corresponds a triangle T_i^K with sides equal to $\rho(F(A), f(B))$, etc. The triangles T_i^K naturally form a complex P^K isomorphic to the complex M. The complex P^K, also called a *development*, is an analog of a polygon inscribed in the surface F, the only difference being in the intrinsic metric. We can say that the complex M determines the development P^K or (an abstract) polygon P^K inscribed in the surface F.

The quantity $S_K(M) = \sum S(T_i^K)$ is none else than the area of the development P^K, so the area of a surface is defined with the help of an *abstract inscribed polygon* P^K.

5.2. Proposition 1. *The area of a surface triangle T is not greater than the area of the corresponding triangle T^K on a K-plane and is equal to it, if and only if the triangle T is isometric to the triangle T^K.*

To prove this proposition, we need the following lemma (which is a simple particular case of Proposition 1).

Lemma. *Let a polygon P on a K-plane be bounded by three convex broken lines $\widehat{AB}$, $\widehat{AC}$ and $\widehat{BC}$ bent into the interior of the polygon; the case in which one or two of the broken lines are straight segments is not excluded (Fig. 16). Then the triangle obtained from the polygon by straightening out these broken lines with sides of the same length as the broken lines $\widehat{AB}$, $\widehat{AC}$ and $\widehat{BC}$ has area greater than the area of the polygon P.*

Fig. 16

Since each of the broken lines $\widehat{AB}$, $\widehat{BC}$ and $\widehat{AC}$ is convex and the convexity of each line faces the other two, each of them is shorter than the sum of the other two lines. In other words, the sides made of broken lines obey the triangle inequality. Consequently, the triangle T does exist.

Proof. First, consider the case in which the polygon P has, in addition to the vertices A, B and C, only one more vertex D, i.e., the polygon is a quadrilateral $ABCD$. Subdivide it into triangles ABD and ACD. By Lemma 2 of Section 3, on transforming the quadrilateral $ABCD$ into a triangle T by straightening out the broken line BDC, we increase the angles at the vertices B, C and A. Therefore, superimposing the triangle ABD on the triangle T so that the side AB coincides with the corresponding side of the triangle T, and doing the same with the triangle ACD, we imbed

both the triangles ABD and ACD inside the triangle T; note that the triangles ABD and ACD do not intersect (Fig. 17). Therefore the area of the quadrilateral P is less than the area of the triangle T.

FIG. 17

Now for a polygon with more than four vertices D_i, the lemma is proved by induction on the number of vertices. Indeed, if a polygon P has more than four vertices, let us truncate a quadrilateral, say, $ADEF$ by drawing a diagonal across the polygon P (Fig. 18).

FIG. 18

Straightening out the broken line ADE, we reduce the number of vertices by one but increase the area of the polygon P because the quadrilateral $ADEF$ is converted into a triangle with greater area. Since, in doing so, the angle at the vertex E can only increase (Lemma 2 of Section 3), the broken line $\widehat{AB}$ still remains convex in the direction of C. Consequently, we obtain a polygon having the same properties but with fewer vertices. This completes the proof of our lemma.

5.3. *Proof of Proposition 1.* Let T be a surface triangle in R_K spanned on the triangle ABC from the vertex A. Let T^K be the corresponding triangle on a K-plane. Take a sequence of consecutive points $B = D_0, D_1, \ldots, D_n$, $D_{n+1} = C$ on BC and subdivide the triangle T into "narrow" triangles $T_i = AD_iD_{i+1}$ by drawing the shortest arcs AD_i. To each narrow triangle T_i let us assign a corresponding triangle T_i^K on a K-plane. These triangles form some polygon Q bounded by the segments $A'B'$, $A'C'$ and the broken line $B'D_1' \ldots D_n'C$. This broken line $B'D_1' \ldots D_n'C$ is convex towards the interior of the polygon Q. The proposition concerning the adjacent angles asserts that the sum of the angles of narrow triangles T_{i-1} and T_i at both the vertices is not less than π, and that when the polygon is converted into triangles T_{i-1}^K and T_i^K, the angles can only increase. Therefore the angles at the vertices D_i' of the polygon Q are not less than π, i.e., the broken line $\widehat{B'C'}$ is bent towards the interior of the polygon Q.

FIG. 19

Consequently, the lemma proved above can be applied to the polygon Q. For the areas, this lemma yields

$$\operatorname{Area} Q = \operatorname{Area} \sum T_i^K \leqslant \operatorname{Area} T^K. \tag{1}$$

Equality holds, if and only if the polygon Q coincides with the triangle T^K.

Now on the shortest arcs AD_i let us take the points E_{ij} ($j = 1, \dots, m$; m being the same for all i). These points on the sides of each narrow triangle T_i can be joined by shortest arcs (Fig. 19). As a result of joining, we obtain "small" triangles T_{ik}, starting from the triangle at the vertex A and ending with the triangle adjacent to the side BC.

Now, putting a triangle on a K-plane into correspondence with each small triangle T_{ik}, we construct a surface T of an abstract inscribed polygon P as exposed in § 1.

The polygon P consists of "narrow" polygons P_i'' corresponding to narrow triangles. Here the shortest arcs AD_{i-1} and AD_i correspond to the broken lines $\widehat{AD}_{i-1}$ and $\widehat{AD}_i$, respectively, which together with the segments $\widehat{D_{i-1}D_i}$ bound the polygon P_i. The "narrow" polygon is composed of the triangles corresponding to small triangles T_{ik}. By the proposition concerning the adjacent angles (Proposition 2 of Section 2), the sum of the angles of the triangles T_{ik} at the vertex E_{ij} is not less than π. For the corresponding triangles on a K-plane, the angles do not increase and, consequently, their sum at each vertex is greater than or equal to π. This means the broken lines $\widehat{AD}_{i-1}$ and $\widehat{AD}_i$ are convex towards the interior of the polygon P_i. Therefore, by the lemma proved above, the area of the narrow triangle P_i is not greater than the area of the corresponding triangle. However, this triangle is simply the triangle T_i^K corresponding to the "narrow" triangle T_i. Therefore, we have

$$\operatorname{Area} P_i \leqslant \operatorname{Area} T_i^K. \tag{2}$$

Equality holds if and only if the polygon P_i coincides with the triangle T_i^K. Summing inequalities (2), we obtain the inequality

$$\operatorname{Area} P \leqslant \operatorname{Area} \sum T_i^K$$

governing the area of the abstract inscribed polygon P. By virtue of (1), we find

$$\operatorname{Area} P \leqslant \operatorname{Area} T^K. \tag{3}$$

By definition, the area of the triangle T is the least lower bound of the areas of the inscribed polygons; therefore (3) yields

$$\operatorname{Area} T \leqslant \operatorname{Area} T^K. \tag{4}$$

This completes the proof of Proposition 1.

5.4. What now remains is to prove that in (4) equality holds if and only if the triangles T and T^K are isometric. Furthermore, let us note that in (4) equality holds if and only if equality holds in (1) and (2).

To prove this assertion, on the side AD of the narrow triangle T_i, let us take one more point F between the points E_{ij} and $E_{i,j+1}$. In constructing the narrow polygon P_i, the triangle which corresponds to the small triangle with the side $(E_{ij}, E_{i,j+1})$ is replaced by a quadrilateral with sides equal to $E_{ij}F$ and $FE_{i,j+1}$ and an angle greater than or equal to π at the vertex F. On converting the quadrilateral into a triangle, its area increases.

Hence it follows that the introduction of new vertices on the shortest arcs AD_i, and on the same base, and the introduction of vertices D' on the side BD can only decrease the area of the corresponding inscribed polygon P. Therefore, the least lower bound of the areas of the polygons P is not less than the area of the polygon itself, i.e.,

$$\operatorname{Area} T \leqslant \operatorname{Area} P.$$

Comparing this inequality with inequality (3), we find

$$\operatorname{Area} T \leqslant \operatorname{Area} F \leqslant \operatorname{Area} T^K,$$

and, consequently, if the area T is equal to the area T^K, then

$$\operatorname{Area} P = \operatorname{Area} T^K.$$

In other words, if Area T=Area T^K, then equality in (3) also holds; then equality in (2) and (1) also hold. However, the equality in (1) holds, as we already mentioned, if and only if each of the narrow polygons Q_i is a triangle T_i^K and the polygon $Q = \sum T_i^K$ is the triangle T^K. From these considerations it follows that the assertion should hold for arbitrarily chosen points D_i on the side DC and for arbitrarily chosen points E_{ij} on the shortest arcs AD_i.

In the triangle T take two points X_1 and X_2 on the shortest arcs AD_1 and AD_2 $(D_1, D_2 \in BC)$, respectively. These shortest arcs divide the triangle T into three "narrow" triangles. In turn subdivide one of these narrow triangles into small triangles, X_1 and X_2 being certain vertices of small triangles (Fig. 20).

The corresponding narrow polygons Q_i must be equal to the triangles T_i^K and the polygon Q composed of the polygons Q_i must be equal to the triangle T^K.

This implies the following: first, if the points D_1' and D_2' on the side $B'C'$ of the triangle T^K correspond to the points D_1 and D_2 (i.e., $B'D_i' = BD_i$), then $A'D_1' = AD_1$ and $A'D_2' = AD_2$. Second, if the points X_1' and X_2' on the

FIG. 20

segments $A'D'_1$ and $A'D'_2$ correspond to the points X_1 and X_2, respectively, (i.e., $A'X'_i = AX_i$), then

$$X'_1X'_2 = X_1X_2.$$

However, X_1 and X_2 being arbitrarily chosen points, the last equality implies that the triangles T and T^K are isometric. This completes the proof.

5.5. From Proposition 1 we can easily deduce the following generalization.

Proposition 2. *Let L be a closed broken line in R_K and P the polygon "spanned on L," i.e., the surface formed by the shortest arcs which join a vertex A on L with all the points on L. The area of the polygon P is not greater than the area of the polygon P^K on a K-plane with sides of the same length as P; here the polygon P^K is inscribed in a curve of constant curvature. The polygon P has the same area as the polygon P^K if and only if both the polygons are isometric.*

(We could as well consider a polygon P formed by the shortest arcs that start, instead of from one vertex, from several vertices of the broken line L. If, say, a broken line has a sequence of consecutive vertices A, B, C and D, then the vertex A can be joined with all the points on the segment BC by shortest arcs, and the vertex C with all the points on the segment AD by shortest arcs.)

Proof. Let us divide the polygon P by means of its diagonals, i.e., by the shortest arcs which pass through A and other vertices. Replacing each such a triangle T_i by its corresponding triangle T_i^K on a K-plane and joining all

these triangles to one another in the same sequence in which the triangles T_i occur, we obtain the polygon $Q = \sum T_i^K$.

Since, by Proposition 1,

$$\operatorname{Area} T_i \leqslant \operatorname{Area} T_i^K,$$

and therefore it follows that

$$\operatorname{Area} P \leqslant \operatorname{Area} Q.$$

Equality holds if and only if each of the triangles T_i is isometric to its corresponding triangle T_i^K, i.e., if the polygon P is isometric to the polygon Q.

Since on a K-plane the polygon P^K (with sides of the same length as the sides of P) that is inscribed in a curve of constant curvature has greatest area, the areas of P^K and P are different; the areas are equal, if and only if the polygons P and P^K are isometric.

A special case of Proposition 2 is that in which P is a polygon on a surface of curvature $\leqslant K$ or a polygon on the development composed of pieces of a K-plane and having negative curvature at every vertex; the total angle Θ at each of its (interior) vertex is greater than or equal to 2π.

5.6. Proposition 3. *On a rectifiable curve in R_K it is possible to span a surface of area not greater than the area of the disk C in a K-plane of the perimeter equal to the length of the curve.*

Taking an arbitrary point O on a given curve L and joining O with all the points on L by shortest arcs, we obtain the surface F whose area is less than the area of a disk C, the only exception is the case in which F is isometric to C. (If $K > 0$, then the length l of the curve must not be greater than $2\pi/\sqrt{K}$; otherwise no disk C would exist.) We only prove that the area of the surface F is not greater than the area of the disk C. The case of equal areas requires a reasoning far from elementary, so we omit it.

Proof. (A) First, let us notice that, a shortest arc OX being a continuous function of its endpoint X on the curve L, the surface F can be represented as the image of the disk under a continuous mapping, so F is in this sense indeed a surface. Now let us construct an abstract polygon P inscribed in the surface F. (For the sake of simplicity in constructing this polygon, instead of using triangulation of the disk whose image is the surface F, we consider the surface F itself. Let us take, in addition to the point A, certain other points $A_1, A_2, \dots, A_n$ on L. Similarly, let us take certain other points $B_{i1}, B_{i2}, \dots, B_{im_i}$ on each of the shortest arcs OA_i. Now consider the triangles whose vertices lie on adjacent shortest arcs OA_i and OA_{i+1}. These triangles are put into correspondence with the triangles with sides of

the same length on a K-plane that form the development or the inscribed abstract polyhedron P.

The polyhedron P is bounded by a closed broken line whose length is not greater than the length of the curve L; but the length of the broken line differs arbitrarily little from the length of L, if the points A_i on are chosen sufficiently close to one another.

(B) We prove that the curvature at every interior vertex of the development P^K is nonpositive, i.e. the total angle Θ is not less than 2π.

Every interior vertex D corresponds to some point B_{ij} on a certain shortest arc OA_i. The triangles $T_1^K, \ldots, T_p^K$ adjoining the vertex D correspond to the triangles $T_1, \ldots, T_p$ adjoining the point B_{ij}. By Proposition 4 of Section 3, the angles of the triangles T_l^K are not less than the angles of the triangles T_l

$$\alpha_l^K \geqslant \alpha_l$$

and, consequently, the total angle $\Theta(B)$ about the point B is

$$\Theta(B) = \sum_i \alpha_l^K \geqslant \sum \alpha_l. \tag{5}$$

Since the point B_{ij} lies in the interior of the shortest arc OA_i, the triangles $T_1, \ldots, T_p$ fall into two groups: the vertices of the first group lie on the shortest arc OA_{i-1} adjacent to OA_i. The vertices of the triangles of the other group lie on the other adjacent shortest arc OA_{i+1}. The sum of the angles of the triangles of one group at the point B_{ij}, by Proposition 1 of Section 2, is not less than the angle between the branches of the shortest arc OA_i starting from B_{ij}, i.e., not less than π. Therefore their total sum is not less than 2π, i.e.,

$$\sum_l \alpha_l \geqslant 2\pi.$$

However, by virtue of (5), it follows that

$$\Theta(B) \geqslant 2\pi.$$

(C) Since the curvature at each point of the development P is nonpositive, we can apply Proposition 2 to the development P, more precisely, its corollary stated at the end of Section 5. Therefore the area of P is not greater than the area of P^K on a K-plane with sides of the same length as the sides of

$$\operatorname{Area} P \leqslant \operatorname{Area} P^K.$$

However, the area of P^K is definitely less than the area of the disk C' with the same perimeter, and is less than the area of the disk C with the

perimeter equal to the length of the initial curve L (i.e., the length L is not less than the perimeter of the polygon P). Therefore

$$\operatorname{Area} P \leqslant \operatorname{Area} C.$$

However, by definition of the area of a surface,

$$\operatorname{Area} F < \varliminf \operatorname{Area} P.$$

Hence it follows that

$$\operatorname{Area} F \leqslant \operatorname{Area} C.$$

This completes the proof.

6. Supplements to Previous Results

6.1. An angle in the strong sense. In this section we state most of the results without proof. Certain results concerning the domain R_K in many respects supplement the results of Section 3 to 5. A key role in a number of these results is played by the following proposition which strengthens the Proposition 3 of Section 3 on the existence of the angle between two shortest arcs.

Proposition 1. *Between two shortest arcs starting from a point in R_K there always exists an angle "in the strong sense," i.e., not only does the angle*

$$\alpha = \lim_{x,y\to 0} \alpha(x,y)$$

exist, but there also exists the angle

$$\alpha = \lim_{x\to 0} \gamma(x,y) = \lim_{y\to 0} \gamma(x,y) \tag{1}$$

In other words, the limit of the angle $\gamma(x, y)$ exists under a single assumption that x (or y) tends to zero, while y (or x) varies in an absolutely arbitrary[11] fashion.

Proof. We prove, say, that $\lim_{x\to 0} \gamma(x, y)$ exists and is equal to α. For this purpose, we apply Proposition 3 of Section 2. According to the proposition, the upper angle α is

$$\alpha = \sup \lim_{x\to 0} \gamma(x,y). \tag{2}$$

[11]The definition of angle in the strong sense was proposed earlier in [1] where it played an important part. The definition given in [1], in the case of R_K, coincides completely with the definition (1).

In our case there exists, as already demonstrated, even an angle (Proposition 3 of Section 3). Therefore we denote this angle by α

$$\alpha = \lim_{x,y\to 0} \gamma(x, y).$$

On the other hand, by Proposition 4 of Section 3, for arbitrary x and y

$$\alpha \leqslant \gamma(x, y),$$

therefore

$$\alpha \leqslant \inf \lim_{x\to 0} \gamma(x, y). \tag{3}$$

From (2) and (3) we find that the $\lim_{x\to 0} \gamma(x, y)$ exists and is equal to α.

Proposition 2. *Let a point A and a shortest arc OB be given in R_K. Let X be a point on OB and put $OX = x$ and $AX = Z(x)$; let ξ be the angle between OX and AX. Then there exists a left derivative of Z with respect to x such that*

$$\left(\frac{dZ}{dx}\right)_{\text{left}} = \cos\xi.$$

By virtue of Proposition 1, Proposition 2 follows directly from a weaker assertion proved in Section 2, namely, that the left lower derivative

$$\left(\frac{dZ}{dx}\right)_{\text{left lower}} \geqslant \cos\xi.$$

Proposition 1 can be generalized for the angle between a shortest curve and a curve. Accordingly, Proposition 2 can be extended to the case in which the shortest arc OB is replaced by a curve satisfying certain appropriate conditions. These propositions admit a variety of applications pertaining to, first of all, the distance between a point and a shortest arc or a curve.

6.2. Space of curvature $\geqslant K'$. Proposition 1 is particularly essential in what follows because if there exist an angle in the strong sense, the angles of a triangle admit evaluation; in a sense this estimate is opposite of the estimate enunciated in the proposition proved in Section 3. Thus, we have

Proposition 3. *Let ABC be a triangle in a certain metric space; every two points on the sides of the triangle ABC can definitely be joined by a unique shortest arc; furthermore, there always exists an angle in the strong sense between these shortest arcs and the corresponding segments of the shortest arcs. Let K' be an arbitrary number and μ'_K, the lower bound of the relative excesses relative to K' of the triangles AXY with vertices at X and Y on the sides AB and AC, respectively, of the triangle ABC.*

Under these assumptions, the angle α at the vertex A of the triangle ABC and the corresponding angle α_K, of the triangle with sides of the same length on the K'-plane obey the inequality:

$$\alpha - \alpha_K \geqslant \mu_{K'}.$$

It is proved on the same lines as the corresponding proposition proved in Section 2. It has been demonstrated in [1, Chapter IV] for $\mu = 0$ and $K' = 0$ in a two-dimensional space. The same reasoning can be applied in a more general form to prove Proposition 3 which may be used as the starting point in the study of spaces of curvature $\geqslant K'$.

Since Proposition 3 holds in R_K, it can be applied to a domain in R_K, particularly, to investigate the properties of a domain of R_K in which the curvature is not less than a given $K' \leqslant K$. This is such a domain in R_K in which the excess of an arbitrary triangle relative to K' is nonnegative. Then, we have $\mu_{K'} \geqslant 0$ and therefore Proposition 3 yields

Proposition 4. *In R_K if the curvature is not less than a given $K' \leqslant K$, then the angle α of a triangle in R_K satisfies the inequality:*

$$\alpha_{K'} \leqslant \alpha \leqslant \alpha_K,$$

where $\alpha_{K'}$ and α_K are the corresponding angles of the triangles on the K' and K planes.

Furthermore, we have

Proposition 5. *If in R_K the curvature is not less than $K' \leqslant K$, then the angle $\gamma^{K'}(x, y)$ for every two shortest arcs starting from a common point is a nonincreasing function of x and y.*

This proposition is proved on similar lines as the Proposition 1 of Section 3 and is demonstrated in Chapter XI of [1]. We can also state other propositions. Their meaning lies in that the properties of a space of curvature $\leqslant K$ and $\geqslant K'$ are represented, so to say, as something in between the properties of spaces of constant curvatures K and K'. In [1] these propositions are proved for convex surfaces, they however can be generalized.

6.3. A ruled surface in R_K. By a *ruled surface* we understand a surface formed by shortest arcs; however, we assume that each point on the surface has a neighborhood in which every pair of points can be joined by a shortest line lying on the surface. We can prove that this property is exhibited by a surface formed by shortest arcs in R_K whose endpoints continuously trace two rectifiable curves. If one of the curves degenerates into a point, then the surface is a cone.

An important result concerning the ruled surfaces is the following

Proposition 6. *A ruled surface in R_K is, in respect of the intrinsic metric, a two-dimensional space of curvature $\leqslant K$.*

This proposition is a direct generalization of a well-known theorem which asserts that a ruled surface in Euclidean space has nonpositive curvature.

Proposition 6 is proved with the use of a finite sequence of shortest arcs and the application of Proposition 2.

Let us note that the surface triangle defined in § 4 of Section 4 is a ruled surface and is traced by the triangle itself. Therefore, by virtue of Proposition 6, Proposition 5 stated in Section 4 concerning the angle of a surface triangle is a simple corollary of Proposition 3 of Section 3 on the angles of a triangle in an arbitrary R_K space.

Likewise, by virtue of Proposition 6, Proposition 2 of Section 5 concerning the area of a surface spanned on a given contour is a corollary of

Proposition 7. *In a two-dimensional manifold of curvature $\leqslant K$, a simply connected domain G bounded by a closed curve of length l has area not greater than the area of a disk C of perimeter l on a K-plane; the area of the domain G is equal to the area of the disk C, if and only if the domain G is isometric to the disk C.*

(For $K > 0$, it is assumed that $l < 2\pi/\sqrt{K}$, otherwise no disk C would exist.)

The first assertion of the proposition is proved just on the same lines as the Proposition 2 of Section 5 is proved. This proposition is demonstrated in [3] under different assumptions.

By virtue of Proposition 6, not only does Proposition 3 follow as a direct corollary of Proposition 7, and so does its generalization:

Every ruled surface F in R_K spanned on a rectifiable curve of length l has area not greater than the area of the corresponding disk C on a K-plane; the area of the surface F is equal to the area of the disk C, if and only if the surface is isometric to the disk C.

6.4. A cone in R_K. In the space R_K let there be given a point O and a rectifiable curve L; let X be a point on the curve L. The shortest arcs OX form some cone C with vertex the point O and directrix L.

We say thet a cone C is *developed into* a cone C_K (more precisely, into a sector) on a K-plane, if there exist a mapping of C onto C^K such that

(1) the vertex O of the cone C is sent into the vertex O' of the cone C^K,

(2) to every generator OX of the cone C, there corresponds a generator $O'X'$ of the cone C^K such that $O'X' = OX$; the mapping of the generator OX onto $O'X'$ preserves the lengths of the corresponding segments,

(3) the directrix of the cone C is mapped onto the directrix of the cone C^K, the lengths being preserved, and
(4) as the point X moves over the directrix L, the shortest arc $O'X'$ steadily rotates about the point O'.

It is understood, such a mapping is always possible and is unique up to a rotation of the cone C^K about the vertex O'.

If the directrix L is a shortest arc, then the cone C is a surface triangle spanned from the vertex O. The figure generated by developing the triangle has the following properties.

Proposition 8. *Let $T = OAB$ be a surface triangle in R_K obtained by joining O with all the points on the side AB. When the triangle is developed, the figure $T' = O'A'B'$ is traced on a K-plane which is bounded by the segments $O'A'$, $O'B'$ and the convex arc $A'B'$ bent inside T'. The arc $A'B'$ here is a segment, if and only if the surface triangle T is isometric to the triangle T^K. If this is not true, then*

(1) *the angle α' between the segments $O'A'$ and $O'B'$ is less than the corresponding angle α_K of the triangle T^K and*
(2) *the area S' of the figure T' is less than the area of the triangle T^K.*

In Sections 4 and 5 we proved certain propositions concerning the angles and the area of a surface triangle, and Proposition 2 in Section 5 on the area of a cone spanned on a closed curve in such a way that the curve forms the directrix of a cone with vertex lying on the curve. These propositions are simple corollaries of the following general proposition on the development C^K of an arbitrary cone C in R^K.

Proposition 9. *A cone C in R_K and the cone C^K generated by developing the cone C in a K-plane are governed by the following relations:*

(1) *If α and α_K are the angles at the vertices of the cones C and C^K, then*

$$\alpha \leqslant \alpha_K$$

and $\alpha = \alpha_K$, if and only if the cone C is isometric to the cone C^K.

(2) *If M is a curve on the cone C and M^K the corresponding curve on the cone C^K, then their lengths ρ obey the inequality:*

$$\rho(M) \leqslant \rho(M^K).$$

If M is a rectifiable curve meeting every generator of the cone C (i.e., has with each generator a common point different from the endpoints of the generators), then

$$\rho(M) = \rho(M^K)$$

if and only if the cone C is isometric to the cone C^K.

(3) *If S and S^K are the areas of the cones C and C^K, then*

$$S \leqslant S^K$$

and $S = S^K$ if and only if the cone C is isometric to the cone C^K.

Assertion (1) corresponds in a sense to Proposition 4 of Section 3 on the angles of a triangle. Proposition 6 of Section 4 is a particular case of Proposition 8. Assertion (2) corresponds to Proposition 2 of Section 3.

Assertion (3) has obviously a corollary which is Proposition 2 of Section 5 on the area of a surface spanned on a closed contour. This assertion is related to Proposition 8 as well as to Proposition 1 of Section 5 on the area of a surface triangle.

6.5. Deviation of a curve from a shortest arc. The results of Section 3 concerning a shortest arc in R_K, viz., Propositions 5 and 6 (of Section 3) on the uniqueness of a shortest arc in R_K and its continuous dependence on its endpoints are supplemented and made stronger by

Proposition 10. *If the length of a curve in R_K differs little from the distance between the points A and B which it joins, then the curve differs little from the shortest arc AB.*

An obvious corollary of this assertion is the uniqueness of the shortest arc AB and the closeness of the locations of the shortest arcs when their endpoints are near one another. Namely, if the points A_n and B_n are close to the points A and B, then the broken line $AA_n + A_nB_n + B_nB$ differs little in length from AB and, consequently, deviates little from AB.

More precisely, Proposition 10 can be reformulated as follows:

In R_K let the points A and B, spaced at a distance r, be joined by a curve L of length l. Then the deviation of the curve L from the shortest arc AB (i.e., the maximum of the distances of the points on the curve L from AB)[12] *is not greater than the height of an isosceles triangle on a K-plane on base of length r and equal sides of length l.*

In particular, for $K < 0$, i.e., in a space of nonpositive curvature, the deviation h of the curve L from the shortest arc AB satisfies the inequality:

$$h^2 \leqslant \frac{1}{4}[l^2 - r^2].$$

If $K > 0$, then it is natural to assume that the isosceles triangle mentioned above exists, so

$$l + r < \frac{2\pi}{\sqrt{K}}$$

[12]This is the so-called *Fréchet distance*. —Eds.

(in a K-plane each of the equal sides of the isosceles triangle is a curve spaced farthest, for a given length l of the sides, from the segment joining its endpoints).

Proof. Let a curve L join two given points A and B. Joining the point A with the points on L by shortest arcs, we obtain a cone which will be developed on a K-plane. Then the curve L is mapped onto a curve L^K on a K-plane: its length and the distance between its endpoints A' and B' are preserved.

From assertion (2) of our proposition we find that the deviation of the curve L^K from the segment $A'B'$ is not less than the deviation of the curve L from the shortest arc AB. Considering an ellipse of perimeter equal to the length of the curve L^K with foci at A' and B' and with major axis of length equal to the sum of the focal distances, we find that the deviation is maximum only when L^K consists of two equal segments. This completes the proof of our proposition.

This proof is based on Proposition 9 which we did not prove. It is relatively complicated to prove Proposition 9. However, Proposition 10 is far simpler to prove for $K \leqslant 0$ or for $K > 0$, at least, under the assumption that the distance r between the points A and B satisfies the inequality

$$r < \frac{\pi}{2\sqrt{K}}.$$

In this case the proof can be built on the following lines.

On a curve L joining the points A and B, let us choose a point C spaced farthest from the shortest arc AB. If D is a point on AB closest to C, then the distance CD is exactly equal to the deviation of the curve from L the shortest arc AB:

$$CD = h.$$

Draw the shortest arcs AC and CB. If l is the length of the curve L, then

$$AC + CB = l.$$

On a K-plane let us construct a triangle $A'B'C'$ which corresponds to the triangle ABC. The deviation of its sides $A'C'$ and $B'C'$ from the base $A'B'$ is exactly equal to the distance between the vertex C' and the base $A'B'$ (for $K < 0$, this is self-evident, but for $K > 0$ this is not necessarily true. However, if

$$A'B' < \frac{\pi}{2\sqrt{K}},$$

this is true).

By Proposition 2 of Section 3, the distance of the points on the sides of the triangle $A'B'C'$ is not less than the distance between the corresponding points on the sides of the triangle ABC. Therefore the deviation of the broken line $A'C' + C'B'$ from the base $A'B'$ is not less than the deviation of the broken line $AC + CB$ from the shortest arc AB. Consequently, it is definitely not less than $h = CD$. Hence, h is not greater than the distance of the vertex C' from the base $A'B'$. However, for a given base $A'B'$ and a given sum of the sides $A'C' + C'B'$, the distance between the vertex C' and the base $A'B'$ is largest in an isosceles triangle. Since, in addition,

$$A'C' + C'B' < l,$$

we find h is not greater than the height of an isosceles triangle on base AB with sum of sides equal to l.

This estimate derived for the deviation of a curve from a shortest arc directly yields an estimate for the Fréchet distance d between two curves joining two given points on AB. Namely, the deviation of one curve from the other is not greater than the sum of the deviations from the shortest arc AB. Hence we obtain an estimate for the deviation in terms of the distance $r = AB$ and the lengths l_1 and l_2 of the curves. Furthermore, this estimate depends naturally on K. Particularly, for $K = 0$, we have

$$d^2 + r^2 \leqslant \frac{1}{2}\,[\,l_1^2 + l_2^2\,].$$

For a two-dimensional manifold this estimate was derived by Beurling [5] in an entirely different way. Not only does our reasoning generalize this estimate, but it also reveals, in our opinion, its simple geometric meaning.

References

[1] Alexandrov A. D., *Intrinsic geometry of surfaces*, Translations Math. monographs **15**; Amer. Math. Soc. (1967).

[2] ———, *Die innere Geometric der konvexen Flachen.*, Berlin, A-V; Supplement, 1955.

[3] ———, *Intrinsic metric of a convex surface in a space of constant curvature*, Dokl. Akad. Nauk SSSR, **45** (1944), 3–6.

[4] ———, *Isoperimetric inequalities for curved surfaces*, Dokl. Akad. Nauk SSSR **45** (1945), 239–242.

[5] Beurling A., *Sur la geometric metrique des surfaces a curbure totale* < 0, Meddanlanden Zunds. Univ. Mat. Sem., Supplement band (1952), 7–11.

[6] Blaschke W., *Differential geometry*, 1933.

[7] Busemann H., *Spaces with nonpositive curvature*, Acta Math., **80** (1948), 259–310.

[8] Hurewicz W. and Wallman H., *Dimension Theory*, Moscow, 1948; Russian translation.

CHAPTER XIV

THE DIRICHLET PROBLEM FOR THE EQUATION $\det\|z_{ij}\| = \varphi(z_1, \dots, z_n, z, x_1, \dots, x_n)$ PART I

Vestnik Leningradskogo Universiteta, Mat., No. 1, 5–24 (1958).

§ 1. Formulation of the Problem

1. Our aim is to study the existence and uniqueness of the solutions to the equation

$$\det\|z_{ij}\| = \varphi(z_1, \dots, z_n, z, x_1, \dots, x_n),$$

where $\det\|z_{ij}\|$ is the determinant consisting of the second derivatives of an unknown function z; with $z_1, \dots, z_n$ standing its first derivatives, and $x_1, \dots, x_n$ for independent variables. We assume that the function φ is nonnegative. For the reasons stated below, it is more convenient to study the equation in a slightly different form. We use the following notation: x stands for the set of variables $x_1, \dots, x_n$; ζ, for the set of first derivatives $z_1, \dots, z_n$, which is simply equivalent to $\zeta = \operatorname{grad} z(x)$. In this notation, we study an equation of the type

$$f(\zeta, z, x) \det\|z_{ij}\| = h(x), \tag{1.1[1]}$$

where h and f are assumed to be greater than or equal to zero. The other conditions imposed on them are formulated below.

We search for generalized solutions of this equation in the class of convex functions, and so we express the equation in certain set functions. Thereafter we return to equation (1.1) in its initial form.

When there are only two variables, the solution of (1.1), for f and $h > 0$, is invariably a convex surface. If there are a larger number of variables,

[1]The first number denotes the section number and the second the serial number of the formula in the section. When referring to a formula within the same section, we only cite its serial number.

then this cannot be guaranteed because det $\|z_{ij}\|$ may be positive when its eigenvalues and, accordingly, the principal radii of curvature of the surface are of unlike signs, the only condition being negative signs must occur in pairs.

Nevertheless, we restrict ourselves only to convex solutions, disregarding all other types.

2. We apply geometric methods; accordingly, we use intuitive geometric concepts in $(n+1)$-dimensional Euclidean space with rectangular coordinates $x_1, \dots, x_{n+1} = z$. We denote the plane $z = 0$ by X, a point on this plane by x which is also used to denote its coordinates. A point in the space is denoted by its coordinates (z, x). The axis z is supposed to be vertical, so we say that a point (z_1, x) lies below a point (z_2, x), if $z_1 < z_2$.

By a domain D we understand, unless otherwise stated, a bounded domain in a plane X.

By a surface S we understand, unless otherwise specified, a convex surface uniquely projectable onto a given domain D and having convexity turned downward. In the language of geometry, instead of saying "a function $z(x)$" we say "the surface S given by the equation $z = z(x)$ satisfying (1.1)."

We now introduce an auxiliary plane Z (the n-dimensional space) with rectangular coordinates. A point on this plane, or the vector joining the origin to this point, is denoted by ζ. When a differentiable function $z(x)$ is given, to each x we assign a vector $\zeta = \operatorname{grad} z(x)$.

The integral of a function $g(x)$ over a set M on the plane X (multiple Lebesgue integral) and the integral of a function $g(\zeta)$ over a set N on the plane Z are simply denoted as

$$\int_M g(x)\, dX, \qquad \int_N g(\zeta)\, dZ.$$

Every function is assumed to be measurable and summable over every bounded closed subset of its domain of definition.

Finally, we define a set function to be a nonnegative countably additive set function defined for all Borel sets M that all are included, together with their closures, in a certain given domain D. It is implied that this function is finite on every such subset. Such a set function can certainly be extended to every Borel set $M \subset D$, but it may become infinite on a set M if the closure of the set is not included in D.

3. In what follows we always deal with a function $f(\zeta, z, x)$ related with a certain given domain D and obeying the following conditions.

Condition I. *The function $f(\zeta, z, x)$ is well-defined for all ζ, z and all $x \in D$; furthermore, it is always nonnegative and may take the infinite value.*

Condition II. *In every closed bounded domain R of the variables ζ, z and x, there exists a summable function $f_0(\zeta)$ such that for every $(\zeta, z, x) \in R$*

$$f(\zeta, z, x) \leqslant f_0(\zeta).$$

Condition III. *There exist a z_0 and a function $f_1(\zeta) \geqslant 0$ such that*

$$\int\limits_Z f_1(\zeta)\, dZ > 0$$

(the infinite value not excepted) and for all $x \in D$ and $z < z_0$

$$f(\zeta, z, x) \geqslant f_1(\zeta).$$

Condition IV. *For almost all ζ, the function $f(\zeta, z, x)$ is continuous in z and x.*

Conditions II, III and IV are definitely satisfied, say, if $f(\zeta, z, x) \geqslant 0$ and is continuous, and if there exists a $f_1(\zeta) \leqslant f(\zeta, z, x)$ such that its integral over Z is greater than 0.

4. We now define some concept that is essential in the sequel, namely, a *normal mapping* φ_s of a domain $D \subset X$ into a plane Z, determined by the surface S. Suppose that S is a surface projectable onto a domain D. Let z_i be the coefficients in the equation of its supporting plane at some point, or the coefficients of the "current coordinates" x_i, if the equation has been solved for z. If the supporting plane is a tangent plane, then z_i are the derivatives of z with respect to x.

In our notation, we can say that to every supporting plane P of the surface, there corresponds a point $\zeta = (z_1, \dots, z_n)$ in the plane Z. Since ζ completely determines the direction of the plane P, we say that the plane P has the direction ζ.

To every set $Q \subset S$, there corresponds a set $\varphi(Q)$ on the plane Z, namely, the set of all points ζ that correspond to the supporting planes of S passing through the points of the set Q. This defines the "normal mapping" of the surface S into the plane Z similar to the spherical mapping. The connection between these two mappings is quite apparent. It suffices to draw a plane touching the Gaussian sphere at one of the poles and then extend the radii of the sphere till they intersect this plane. In this way a spherical mapping is converted to a normal mapping. So, instead of a normal mapping, we could as well use a spherical mapping by introducing a unit normal $\bar{n}$ in place of ζ. However, in analysis it is easier to handle $\zeta = (z_1, \dots, z_n)$ than a unit normal and its connection with differential equations is simpler.

Now suppose that M is a set in the domain D and Q a set on the surface S whose projection on D is the set M. Then the normal mapping

of Q obviously defines a mapping of M into the plane Z. This mapping is denoted by φ_s and is also called *normal*.

Thus, the surface S defines the "normal" mapping φ_s of the domain D into the plane Z. [This mapping and its inverse are in general not unique. However, uniqueness is violated only on a set of measure zero.]

5. Suppose that a domain D and a function $f(\zeta, z, x)$ obeying Conditions I–IV (of § 3) are given. Basing on this function, we define a set function $\omega_f(M; S)$ for a surface S projectable into D. Let $x(\zeta)$ and $z(\zeta)$ be the coordinates of a point on S through which passes the supporting plane of direction ζ. The mapping $\zeta \to [x(\zeta), z(\zeta)]$ is just the inverse of our normal mapping. The functions $x(\zeta)$ and $z(\zeta)$ are not unique for those points ζ for which the supporting plane of direction ζ touches the surface at more than one point. However, the set of such points ζ has zero measure.

Thus, the function $f[\zeta, z(\zeta), x(\zeta)]$ is uniquely defined for almost all ζ. For every given $M \subset D$, we now define a function $\omega_f(M; S)$, putting

$$\omega_f(M; S) = \int\limits_{\varphi_s(M)} f(\zeta, z(\zeta), x(\zeta))\, dZ, \tag{1.2}$$

i.e., the integral of f over the normal mapping $\varphi_s(M)$ of the set M.

By virtue of Conditions I and II imposed on f, the function ω_f definitely is nonnegative and, in addition, finite for every closed $M \subset D$. Indeed, for an arbitrary closed $M \subset D$, there is no supporting plane arbitrarily close to the vertical plane at a point of S lying above M. Hence the set $\varphi_s(M)$ is bounded and, by Condition II, integral (1.2) is finite. Incidentally, for

$$f(1 + \zeta^2)^{-(n+1)/2},$$

the function $\omega_f(M; S)$ is simply the area of the spherical map of that set on the surface S whose projection is the set M.

So the function $\omega_f(M; S)$ may be called a *relative curvature* or a *relative integral curvature of S relative to the plane X*.

6. We now transform equation (1.1) into some form with a set function. Let a regular surface S projectable onto the domain D satisfy (1.1). Then, for each set $M \subset D$ we definitely have

$$\int\limits_M f(\zeta, z, x)\ \det \|z_{ij}\|\ dX = \int\limits_M h(x)\ dX. \tag{1.3}$$

However, $\det \|z_{ij}\|$ is the Jacobian of the transformation $(x_1, \dots, x_n) \to (z_1, \dots, z_n)$, i.e., the normal mapping φ_s. Therefore in our notation we can write $\det \|z_{ij}\|\, dX = dZ$, and equality (1.3) is equivalent to the following

$$\int\limits_{\varphi_s(M)} f(\zeta, z(\zeta), x(\zeta))\ dZ = \int\limits_M h(x)\ dX. \tag{1.4}$$

Using definition (1.2) of $\omega_f(M; S)$ and substituting the set function $\nu(M)$ for the integral of h, we obtain

$$\omega_f(M; S) = \nu(M). \tag{1.5}$$

This is just the equation in set functions we need.

Since the function ω_f is defined for all convex surfaces S, there is no need to suppose that S is regular. In the same way, we can also interpret $\nu(M)$ to be an arbitrary set function. So our problem of finding the solution of (1.5), can be reformulated as: *To find a surface S for which $\omega_f(M; S)$ is a given set function $\nu(M)$.*

If, in addition, a surface S yielding the solution of our problem, is sufficiently smooth; then it also gives the solution of (1.1), because the function h in (1.1) is defined by the function $\nu(M)$ as its "density," the area derivative.

7. As already noted, for the particular case of

$$f = (1 + \zeta^2)^{-\frac{n+1}{2}}, \tag{1.6}$$

the function $\omega_f(M; S)$ is the area of the spherical map (integral curvature) of the set on the surface S whose projection is M.

However, in general, if f depends only on ζ [so that $f = f(\zeta)$], the function $\omega_f(M; S)$, as can be easily verified, is the area of a spherical map with weight

$$\delta(\zeta) = f(\zeta)(1 + \zeta^2)^{-\frac{n+1}{2}}.$$

In other words, we can imagine it is the Gaussian sphere of "density" $\delta(\zeta) = \delta(n)$, where n is the point on the sphere that corresponds to a point ζ on the plane Z. Then the integral of this density taken over the spherical image of a set $Q \subset S$ having M as its projection gives the function $\omega_f(M; S)$.[2]

In [1–4] I proved certain theorems concerning the existence and uniqueness of a convex surface with given integral curvature. They also include the theorems on the existence and uniqueness of a solution to equation (1.5) and the existence and uniqueness of a generalized solution to (1.1) if f is a function of the type (1.4).

[2]A sphere or a plane Z could be assigned a certain "mass" distribution by defining a set function $\mu(N)$ on it and then determine the function $\omega_\mu(M; S) = \mu(\varphi_s(M_f))$ for the surface S. However, even if the product $M_1 M_2$ vanishes, the product $\varphi_s(M_1)\varphi_s(M_2)$ can fail to vanish. Therefore $\omega_\mu(M; S)$ can fail to be an additive function. However, we know that if $M_1 M_2$ vanishes, the set $\varphi_s(M_1)\varphi_s(M_2)$ has zero measure, so additivity of the function ω_μ is guaranteed, provided $\mu(N) = 0$ for every set N of measure zero. However, in this case $\mu(N)$ is the indefinite integral of the corresponding density. Hence, for $\omega_|\mu S$ to be additive, it is necessary and sufficient that $\mu(N)$ be representable in the form of such an integral.

Bakel′man noticed that my methods [1–4] are equally applicable to a more general case of $f = f(\zeta)$. The assumption made in [8], namely, that $f(\zeta)$ be continuous and bounded, is not essential as the methods [1–4] are in no way connected with it.

In this paper we apply analogous methods to a general function $f(\zeta, z, x)$ obeying Conditions I–IV. We also derive certain new results on the solution of Dirichlet problems for particular cases.

8. Our proof for the existence of a solution to equation (1.5) is based on approximating a surface by polyhedra and, accordingly, on approximating a set function by functions "consisting of finitely many point loadings." Therefore, in Section 2 we first prove certain theorems to justify the applicability of our method.

In Section 3 we demonstrate the existence of a solution to equation (1.5) under certain general conditions imposed on the function $\nu(M)$. Unlike in [1, 2, 5], we use an elegant tool, due to Pogorelov [11], to prove the existence of polyhedra. In addition to simplicity, its main merit is that it does not rely on any existence theorem. Essentially, it is an effective method for finding approximate solutions. The theorem on the existence of a solution to equation (1.5) also implies the existence of a generalized solution to equation (1.1). For sufficiently regular functions f and h, we cannot say anything regarding the regularity of the solution in the general case of n variables. However, for two variables, we may cite [9, 11, 12, 6] which contain important, although not exhaustive results.

In Section 4 we establish general sufficiency conditions for the solvability of the Dirichlet problem for equation (1.5) and, hence, of equation (1.1).

In Section 5, imposing an additional condition on $f(\zeta, z, x)$, i.e., requiring that it be a nonincreasing function of ζ, we prove the solution of the Dirichlet problem is unique.

In Section 6 we derive results that are analogous to those found in Sections 3–5, concerning the solutions of equations (1.1) and (1.5) in unbounded domains.

In all these case we somewhat examine the necessity of the conditions for the existence and uniqueness of a solution.

The present note contains Sections 1, 2 and 3; Sections 4, 5 and 6 are published in the next issue of the journal.

§ 2. Compactness Theorems

1. Let D be a fixed domain on a plane X and $f(\zeta, z, x)$ a function satisfying Conditions I–IV, § 1. By virtue of Condition III, there exist a z_0 and a function $f_1(\zeta)$ such that, for all $x \in D$ and $z < z_0$,

$$f(\zeta, z, x) \geqslant f_1(\zeta)$$

and

$$\int_Z f_1(\zeta)\, dZ > 0.$$

We now define a quantity $A(f)$ as follows

$$A(f) = \sup \int_Z f_1(\zeta)\, dZ,$$

where the supremum is taken over all possible z_0 and over all possible functions $f_1(\zeta)$. It is not excepted that the integral of f_1 and, consequently, the supremum $A(f)$ may be infinite.

The supremum $A(f)$ plays a key role in the conditions for the main equation (1.5) to be solvable. This role is determined by

Theorem 1. *Convex surfaces S projectable onto a domain D and having a common border L are uniformly bounded, if they all obeying the inequalities*

$$\omega_f(D; S) \leqslant C < A(f), \tag{2.1}$$

where C is a constant.

Proof. By the definition of the supremum $A(f)$, for a given $C < A(f)$ there clearly exist a z_0 and a function $f_1(\zeta)$ such that

$$f(\zeta, z, x) \geqslant f_1(\zeta) \geqslant 0 \qquad (z < z_0, x \in D) \tag{2.2}$$

and

$$\int_Z f_1(\zeta)\, dZ > C. \tag{2.3}$$

Of course, we can assume that z_0 is such that the common border L of the surfaces S lies above the plane $z = z_0$. Then this plane cuts off a "cap" $\bar{S}$ from every "large" surface. It suffices to demonstrate that the "caps" are bounded.

Let $\bar{D}$ denote the projection of the cap $\bar{S}$. Then, clearly

$$\omega_f(\bar{D}; \bar{S}) \leqslant \omega_f(D; S). \tag{2.4}$$

By the definition of the function ω_s and by virtue of inequality (2.2), however

$$\omega_f(\bar{D}; \bar{S}) = \int_{\varphi_s(\bar{D})} f(\zeta, z, x)\, dZ \geqslant \int_{\varphi_s(\bar{D})} f_1(\zeta)\, dZ, \tag{2.5}$$

where $\varphi_s(\bar{D})$ is the normal image of the cap $\bar{S}$.

Since the common border of the surfaces S is fixed, the cap $\bar{S}$ grows with the surface S. And the normal image of the cap also increases and eventually covers the entire plane Z. Therefore, by virtue of (2.3), for a sufficiently large cap, we have

$$\int\limits_{\varphi_s(\bar{D})} f_1(\zeta)\, dZ > C. \tag{2.6}$$

Now comparing inequalities (6), (5) and (4), we find that they contradict (1.1). This demonstrates that under the conditions of the theorem, the caps $\bar{S}$ cannot be arbitrarily large. So the surfaces $\bar{S}$ are bounded.

2. By improving this result, we can easily find an estimate for the possible height of the cap $\bar{S}$ and, thereby, the possible height of the surfaces S.

Let z_0 and $f_1(\zeta)$ denote the same quantities as before. Suppose that Z_ρ is the ball in the plane Z of radius ρ and center the point $\zeta = 0$, and

$$\sigma(\rho) = \int\limits_{Z_\rho} f_1(\zeta)\, dZ. \tag{2.7}$$

Let d be the diameter of the domain D. Now, if (2.1) holds for the surface S, then the height $\bar{p}$ of the cap $\bar{S}$ cut out from the surface S by the plane $z = z_0$ is given by the inequality:

$$\sigma\left(\frac{\bar{p}}{d}\right) \leqslant C. \tag{2.8}$$

This inequality gives an estimate for the cap height and the surface height.[3]

To prove inequality (2.8), take a point A on the cap $\bar{S}$ that is farthest from the plane $z = z_0$. Projecting the border from this point A, we obtain a cone K. The normal image of the cone K is obviously included in the normal image of the cap $\bar{S}$ and the latter is included in the normal image of the surface S. Therefore

$$\varphi(K) \subset \varphi(\bar{S}) \subset \varphi_s(D). \tag{2.9}$$

[3]Since $f_1(\zeta)$ is nonnegative, the function σ is monotone. Therefore, we can write (8) as $\bar{p} \leqslant \sigma^{-1}(C)d$, where σ^{-1} is the inverse of σ. If p_1 denotes the greatest of the distances of boundary points of the surface S from the plane $z = z_1$, then the height of the surface S is given by the estimate $p \leqslant \sigma^{-1}(C)d + p_1$. The condition $C < A(f)$ is implicitly implied here as, by the definition of $A(f)$, we have $\sigma(\rho) < A(f)$ and therefore $\sigma^{-1}(C)$ is defined, if and only if $C < A(f)$.

Let $\bar{D}$ denote the base of the cap $\bar{S}$ on the plane $z = z_0$; it is also the base of the cone on this plane. The diameter of $\bar{D}$ is obviously not greater than that of the domain D. Therefore the base $\bar{D}$ is certainly included in the ball (disk) D_0 of radius d described on the plane $z = z_0$ about the point A_1, the projection of the point A. Projecting this disk D_0 from the point A, we obtain a right circular (spherical) cone K_0. The normal image of the cone K_0 is included in the normal image of the cone K. It is a ball $Z_{\frac{\bar{p}}{d}}$ of radius $\bar{p}/d$ lying on the plane Z. Thus $Z_{\frac{\bar{p}}{d}} = \varphi(K_0) \subset \varphi(K)$. Hence, from (2.9), we obtain $Z_{\frac{\bar{p}}{d}} = \varphi_S(D)$. Therefore

$$\int\limits_{Z_{\frac{\bar{p}}{d}}} f_1(\zeta)\, dZ < \int\limits_{\varphi_S(D)} f_1(\zeta)\, dZ. \tag{2.10}$$

By virtue of (2.7) and (2.1), we derive (2.8). This completes the proof.

3. We now elucidate the sense in which the relative curvatures $\omega_f(M; S_m)$ of the surfaces S_m converge to $\omega_f(M; S)$ when the surfaces S_m converge to the surface S. When we say that the surfaces S_m projectable onto D converge to a surface S, we simply understand the functions $Z_m(x)$ defining S_m converge to the function $Z(x)$ defining S.

The set functions $\mu_n(M)$ defined for the sets $M \subset D$ are said to *converge weakly* in the interior of D to a function $\mu(M)$, if for every continuous function $g(x)$ nonzero only on a set G which is included, together with its closure $\bar{G}$, in D

$$\lim_{m\to\infty} \int\limits_D g(x)\mu_m\,(dM) = \int\limits_D g(x)\mu\,(dM). \tag{2.11}$$

Essentially, the function μ, to which the given functions μ_n converge weakly in the interior of D, is unique.[4]

[4]Let G be an open set, which is included together with its closure $\bar{G}$ in D. Let Γ be a set of continuous functions $g(x)$ such that $0 \le g(x) \leqslant 1$ and $E_x[g(x) > 0] \subset G$. Since the set functions are assumed to be nonnegative, we have $\mu(G) = \sup_{G\in\Gamma} \int_D g(x)\mu\,(dM)$ and for every M, we have $\mu(M) = \inf_{G\ni M} \mu(G)$. Therefore the function μ is uniquely defined by these integrals. Refer to [7] for greater detail regarding this question and weak convergence. By the definition given therein, μ_n converge weakly to μ, if (2.11) holds for arbitrary continuous bounded functions $g(x)$ defined on D. It is easy to prove with examples that in Theorem 2 we cannot substitute simple weak convergence in D for the weak convergence in the interior of D. The reason is that as S_m converge to S, the functions $\omega_f(M; S_m)$ may have a "slipping load." According to [7], for a sequence of set functions to converge weakly in D, it is necessary and sufficient that the sequence be weakly convergent in the interior of D and have no slipping loads. We may note that Theorem 2, if f depends only on ζ, is truly a direct corollary of a more general theorem established in [7].

Theorem 2. *If the surfaces S_m projectable onto a given domain D converge to a surface S, then their relative curvatures $w_f(M; S_n)$ converge weakly in the interior of D to $\omega_f(M; S)$.*

Proof. By virtue of definition (2.11), as $S_n \to S$,

$$\int_D g(x)\omega_f\,(dM; S_n) \to \int_D g(x)\omega_f\,(dM; S) \tag{2.12}$$

for every continuous function $g(x)$ is nonzero only on a certain set G that is included together with its closure $\bar{G}$ in D. Let us take such a function $g(x)$ and prove that

$$\int_D g(x)\,\omega_f\,(dM; S) \;=\int_{\varphi_S(D)} g(x(\zeta))\,f(\zeta, z(\zeta), x(\zeta))\,dZ. \tag{2.13}$$

By the definition of relative curvature $\omega_f(M; S)$, the functions $x = x(\zeta)$ and $z = z(\zeta)$ on the right side are, for a given ζ, the coordinates of the point on the surface S through which passes the supporting plane of direction ζ.

Since the function $g(x)$ is nonzero only on G, the integral on the left side of (2.13), which we denote by J, can be taken only on G. Since $g(x)$ is continuous, J is Riemann integrable and can be found by subdividing the domain D into "smaller" sets M. Therefore

$$J = \int_G g(x)\,\omega_f\,(dM; S) = \lim \sum_i g(x_i)\,\omega_f\,(M_i; S).$$

Now, using the definition of $\omega_f(M; S)$, we find

$$J = \lim \sum_i g(x_i) \int_{\omega_S(M_i)} f(\zeta, z, x)\,dZ. \tag{2.14}$$

As already mentioned in § 1, the functions $x(\zeta)$ and $z(\zeta)$, by the definition of ω_f, are unique almost everywhere. Therefore in determining the integral J we can discard the points ζ where they are not unique. So, in calculating the integral J over $\omega_S(M_i)$, we can take $x \in M_i$.

Now, from (2.14), we easily obtain

$$J = \int_{\omega_S(M)} g(x)\,f(\zeta, z, x)\,dZ + \lim \sum_i \int_{\omega_S(M_i)} [g(x_i) - g(x)]\,f(\zeta, z, x)\,dZ. \tag{2.15}$$

Here the second term on the right is zero. Indeed, since $g(x)$ is continuous, if the sets M_i are made smaller and smaller, the difference $[g(x_i) - g(x)]$ tends to zero (because x can be regarded an element of M_i). Moreover, by Condition II, § 1, we can take $f(\zeta, z, x) < f_0(\zeta)$, where $f_0(\zeta)$ is summable over a bounded domain. However, the sets $\varphi_S(M_i)$ all are included in $\varphi_S(G)$. And since $\bar{G} \subset D$, the set $\varphi_S(G)$ is obviously bounded. Hence, the second term on the right of (2.15) vanishes. This completes the proof of (2.13).

A similar equality holds for the function $\omega_f(M; S_m)$ as well. Thus, in place of (2.12), we have to prove that

$$\int\limits_{\varphi_{S_m(G)}} g(x_m(\zeta))\, f(\zeta, z_m(\zeta), x_m(\zeta))\, dZ \;\to \int\limits_{\varphi_S(G)} g(x(\zeta))\, f(\zeta, z(\zeta), x(\zeta))\, dZ. \tag{2.16}$$

4. We now prove relation (2.16). For this, let us introduce a function $\vartheta(\zeta)$ defined for all ζ on the plane Z as

$$\vartheta(\zeta) = \begin{cases} g(x(\zeta)f(\zeta, z(\zeta), x(\zeta)) & \zeta \in \varphi_s(G) \\ 0 & \zeta \notin \varphi_s(G), \end{cases}$$

and define a function $\vartheta_n(\zeta)$ in the same way. Then, (2.16) is reduced to

$$\int\limits_Z \vartheta_n(\zeta)\, dZ \to \int\limits_Z \vartheta(\zeta)\, dZ. \tag{2.17}$$

To prove this relation, first we show that $\vartheta_n(\zeta)$ tends to $\vartheta(\zeta)$ almost everywhere. We distinguish two cases:

(1) $\zeta \in \varphi_S(G)$ for $\vartheta = gf$, and
(2) $\zeta \notin \varphi_S(G)$ for $\vartheta(\zeta) = 0$.

Case 1. By Condition IV of § 1, $f(\zeta, z, x)$ is continuous in z and x for almost all ζ. We can therefore consider a $\zeta_0 \subset \varphi_S(G)$ so that $f(\zeta, z, x)$ is continuous in z and x.

The set of the directions of the supporting planes that touch our surface at more than one point has measure zero. So we can choose a ζ_0 such that the supporting plane P_0 of direction ζ_0 touches the surface S at only one point A. The projection of this point certainly belongs to G. In this case a cap $\bar{S}$ projectable into the domain G can be cut off from the surface S by a plane P parallel to P_0. Since $S_m \to S$, then, from the surfaces S_m with large m the plane P truncates caps also projectable into the domain G. Each of them contains a point A_m through which passes a supporting plane parallel to P.

This proves that for the ζ_0 chosen, if m is sufficiently large, on each surface S_m there is a point A_m projectable into G through which passes a

supporting plane of direction ζ_0. Hence, for large m, the coordinates $x_m(\zeta_0)$ and $z_m(\zeta_0)$ are well defined where $x_m(\zeta_0) \in G$, so we have $\zeta_0 \in \varphi_{S_m}(G)$. We can restrict ourselves only to those m's for which this holds.

If the supporting plane P_0 touches the surface S at only one point A, then we know that each of the surfaces S_m converging to S, contains a point A_m through which passes a supporting plane. These points A_m, where the supporting planes are parallel, converge to the point A. This means for the ζ_0 chosen

$$x_m(\zeta_0) \to x(\zeta_0), \qquad z_m(\zeta_0) \to z(\zeta_0). \tag{2.18}$$

Since ζ_0 belongs to $\varphi_S(G)$ and to $\varphi_{S_m}(G)$, we have

$$\begin{aligned} \vartheta(\zeta_0) &= g(x(\zeta_0))\, f(\zeta_0, z(\zeta_0), x(\zeta_0)), \text{ and} \\ \vartheta(\zeta_0) &= g(x_m(\zeta_0))\, f(\zeta_0, z_m(\zeta_0), x_m(\zeta_0)). \end{aligned}$$

By the choice of ζ_0, the function $f(\zeta_0, z, x)$ is continuous in z, x, and by the condition of the theorem, $g(x)$ is continuous. Therefore, from (18), it follows that $\vartheta_m(\zeta_0) \to \vartheta(\zeta_0)$.

Here $\zeta_0 \in \varphi_S(G)$ being arbitrary up to a set of measure zero, $\vartheta_m(\zeta)$ converges to $\vartheta(\zeta_0)$ almost everywhere on $\varphi_S(G)$.

Case 2. Here $\zeta \notin \varphi_S(G)$, so $\vartheta(\zeta) = 0$. If, in addition, ζ_0 does not belong to $\varphi_{S_m}(G)$, then, by the definition of ϑ_m, we have $\vartheta_m(\zeta_0) = 0$. If this holds for all sufficiently large m, then $\vartheta_m(\zeta_0) \to \vartheta(\zeta_0)$. Therefore, what remains is to consider a ζ_0 which belongs to $\varphi_S(G)$ but is not included in $\varphi_{S_m}(G)$, starting from an arbitrarily large m. So we can assume

$$\zeta_0 \in \varphi_S(G), \quad \zeta_0 \in \varphi_{S_m}(G). \tag{2.19}$$

Since we are interested only in almost everywhere convergence, and since the functions $x_m(\zeta)$ and $z_m(\zeta)$ are unique almost everywhere, it suffices to examine a ζ_0 for which our assertion holds. Then, for every given m, the point $x_m(\zeta)$ is unique and the second relation in (2.19) yields $x_m(\zeta_0) \in G$. Now we prove the point set $x_m(\zeta_0)$ has no condensation points in the domain G. Assuming the contrary, let x_0 be a condensation point such that $x_m(\zeta_0) \to x_0 \in G$. Some point A_{m_i} on the surface S_{m_i} corresponds to the point $x_{m_i}(\zeta_0)$. Since the surfaces S_{m_i} converge to the surface S, the points A_{m_i} converge to some point A on S and A is projectable into $x_0 \in G$.

The supporting plane of the surface S_{m_i} of direction ζ_0 passes through the point A_{m_i}. Hence the supporting plane of the surface S of the same direction ζ_0 passes through the limit point A. As the point x_0 belongs to G, the point ζ_0 belongs to $\varphi_S(G)$ contrary to condition (2.12).

Thus the set of points $x_m(\zeta_0)$ fails to have a condensation point in the domain G. Therefore, these points approach the boundary of the domain

with m increasing. However, the function $g(x)$, by the condition of the theorem, is continuous and vanishes on the boundary of G. Therefore $g[x_m(\zeta_0)]$ tends to zero. Since ζ_0 is arbitrary and satisfies condition (2.19) everywhere except on a set of measure zero, the function $g[x_m(\zeta)]$ tends to zero for almost all ζ under condition (2.19).

By definition, the functions $\vartheta_m(\zeta)$ clearly tend to zero for almost all ζ. Since $\vartheta(\zeta)$ vanishes for the values of ζ being studied, the functions $\vartheta_m(\zeta)$ tend to $\vartheta(\zeta)$.

We have studied all the possible choices for ζ and have thus proved the functions $\vartheta_m(\zeta)$ converge to $\vartheta(\zeta)$ almost everywhere.

We now prove that $\vartheta_m(\zeta)$ all are dominated by the same summable function $\vartheta_0(\zeta)$.

Since $\bar{G} \in D$ and the surfaces S_m converge, there are no points on these surfaces over the domain G where the supporting planes are as close to the vertical as we please. Therefore the plane Z includes a bounded set U such that $\varphi_{S_m}(G) \subset U$ for all m.

By definition, we have

$$\vartheta_m(\zeta) = 0 \quad \text{for} \ \ \zeta \notin U. \tag{2.20}$$

And, by Condition II of § 1, there clearly exists a function $f_0(\zeta)$ summable over each bounded domain such that, for all m,

$$f(\zeta, x_m(\zeta), z_m(\zeta)) \leqslant f_0(\zeta).$$

Therefore, by definition, we have

$$\vartheta_m(\zeta) \leqslant Bf_0(\zeta), \tag{2.21}$$

where B is the supremum of $g(x)$. Now put

$$\vartheta_0(\zeta) = \begin{cases} Bf_0(\zeta), & \zeta \in U \\ \quad 0 \quad , & \zeta \notin U. \end{cases}$$

It is summable over the plane Z, and from (2.20) and (2.21), we find that $\vartheta_m(\zeta) \leqslant \vartheta_0(\zeta)$.

Thus, $\vartheta_m(\zeta) \to \vartheta(\zeta)$ almost everywhere and the functions ϑ_m are dominated by a summable function. Hence, by virtue of a well-known theorem, their integrals converge and therefore (17) holds. This completes the proof of Theorem 2.

§ 3. Existence of a Solution

1. Let D be a convex domain in a plane X and Γ its boundary. Not every closed $(n-1)$-dimensional surface projectable onto Γ can be the boundary of a convex surface S projectable onto D. (The border of the surface S is simply the set of its limit points not belonging to S.) A surface L that could be the boundary of some other surface S is called an *admissible surface* or an *admissible border*. For the Dirichlet problem to be solvable, it is necessary that the surface L be admissible. It is a simple matter to formulate the necessary and sufficient conditions for a surface L to be admissible.

Suppose that C is the closed cylinder developed by a vertical generator moving over a directrix Γ. Let L be a closed $(n-1)$-dimensional surface projectable onto Γ and thus lying on the cylinder C. The surface L subdivides the cylinder C into two parts: the "upper cylinder" C_1 and the "lower cylinder" C_2. Let us construct the convex hull of the upper cylinder C_1. Its lateral surface may include some part of the cylinder C that does not belong to C_1. In this case the surface L is not the boundary of the lateral surface of our convex hull. However, the properties of the convex hull assert that L cannot at all be the border of any surface S, i.e., any convex surface projectable onto D and having convexity turned downward.

On the other hand, if the lateral surface of our convex hull converges to the upper cylinder C_1, the surface L is the boundary of this lateral surface. The remaining portion of the convex hull is a convex surface S_L of zero curvature, i.e., a convex surface with zero area of its spherical image projectable onto D and having convexity turned downward. Its border is just L. We say the surface S_L is spanned over L from below.

Thus, for L to be an admissible border of surfaces S, it is necessary and sufficient that there exist a surface S_L of zero curvature that could be spanned on L from below.

If the domain D is strictly convex, i.e., each of its supporting planes touches its boundary Γ at only one point, then each closed $(n-1)$-dimensional surface L uniquely projectable onto Γ is an admissible boundary. However, unique projectability onto Γ is not essential. For instance, L might include vertical segments, yet being an admissible border. In terms of analysis, the boundary conditions imposed by the surface L might not be continuous.

2. Now we turn to our main equation

$$\omega_f(M; S) = \nu(M). \tag{3.1}$$

We assume that the convex domain D is fixed and a function $f(\zeta, z, x)$ is given that satisfies Conditions I–IV of § 1. Let $A(f)$ be the quantity defined at the beginning of § 2.

Theorem 3. *Let D be a polyhedral convex domain in a plane X and L an admissible $(n-1)$-dimensional polyhedral surface projectable onto the boundary of D. Let $\nu(M)$ be a set function defined on D and consisting of point loadings.*[5] *Then, if*

$$\nu(D) < A(f), \tag{3.2}$$

there exists a convex polyhedral surface S having L as its border and satisfying equation (3.1) with the given $\nu(M)$ on the right side.

In simpler terms, this theorem can be formulated as follows.
Let the loadings of the function $\nu(M)$ be concentrated at the points $x_1, \dots,$ x_m, and be $\nu_1, \dots, \nu_m$, respectively. Then condition (3.2) is reduced to

$$\sum_{i=1}^{m} \nu_i < A(f). \tag{3.3}$$

Let us consider polyhedral surfaces S that have possibly no other vertices than those that are projectable into the points $x_1, \dots, x_m$. The area of the normal image of a set not containing vertices is zero. Therefore the function $\omega_f(M;S)$ consists of only point loadings $\omega_f(x_i;S)$. If z_i's are the coordinates z of the vertex A_i projected into x_i, then, by definition (1.2), we have

$$\omega_f(x_i;S) = \int\limits_{\varphi_S(x_i)} f(\zeta, z_i, x_i)\, dZ. \tag{3.4}$$

In this integral of z_i, the coordinates x_i are constants. Therefore $\omega_f(x_i;S)$ is simply the "area" of the normal image of the vertex $A_i(z_i, x_i)$ with "weight" $f(\zeta, z_i, x_i)$.

Let D and L have the meaning as stated in Theorem 3. Then the theorem can be improved as follows.

Under condition (3.3), there exists a surface amongst the polyhedral surfaces S such that for $i = 1, \dots, m$

$$\omega_f(x_i;S) = \nu_i, \tag{3.5}$$

where S is projected onto D and L is its border.

Proof. Let us span the surface S_L on L from below. Clearly, S_L is a polyhedral surface with no vertices and L as its border. So we can assert it has no vertices except perhaps those that are projectable into the points $x_1, \dots, x_m$. For this surface, $\omega_f(x_i;S_L) = 0$ $(i = 1, \dots, m)$.

[5] We say $\nu(M)$ consists of the point loadings concentrated at the points $x_1, \dots, x_m$, if $\nu(M)$ vanishes on every set M not containing any one of the points x_i and $\nu(x_i) \neq 0$ for every x_i.

We now consider all polyhedral surfaces S that have L as a border and that have no vertices possibly other than those that are projectable into the points $x_1, \dots, x_m$ such that m

$$\omega_f(x_i; S) \leqslant \nu_i \qquad (i = 1, \dots, m). \tag{3.6}$$

By virtue of condition (3.3), from Theorem 1 of § 2, these surfaces all are bounded. Therefore, amongst them there is a surface S^0 for which $\sum x_i$ (the sum of the coordinates of its vertices) is minimal.[6] We prove S^0 is just the surface that we need.

Assume the contrary. Then, for a certain point

$$\omega_f(x_k; S^0) < \nu_k. \tag{3.7}$$

Let A_k be a point on S^0 that is projectable into the point x_k. A priori, the point A_k may or may fail to be a vertex in the true meaning of the word. Move A_k vertically down through a small distance $A_k A'_k$. Construct a convex hull of the point A'_k, the vertices of S^0 and a border L. That part of this convex hull with downward convexity is a convex polyhedral surface S' having L as a border and whose vertices A'_i are projectable into the points x_i.

We now show that if the displacement $A_k A'_k$ is small, inequality (3.6) is satisfied on S'. So S' is one of the surfaces under consideration.

Let A'_i be a vertex of S' different from A'_k. From the construction of the surface S', it is clear that A'_k is the vertex A_k of the surface S^0 and the polyhedral angle of the surface S' at this vertex includes the polyhedral angle of the surface S^0 (the angles may also be equal). Therefore for the normal images of these surfaces we have an opposite inclusion

$$\varphi_{S'}(x_i) \subset \varphi_S(x_i).$$

Now applying formula (3.4), and as the coordinates z_k and x_k are the same in both the cases, we find that

$$\omega_f(x_i; S') \leqslant \omega_f(x_i; S^0),$$

so, by virtue of (3.6)

$$\omega_f(x_i; S') \leqslant \nu_i. \tag{3.8}$$

A priori it may happen that the point A' on S' projectable into x_k is not a vertex. Then, $\omega_f(x_i; S') = 0$ because (3.8) is definitely satisfied.

---

[6]Strictly speaking, boundedness of the surfaces S only implies there exists a limit surface S^0 on which $\sum z_i$ is minimal. However, S^0 is clearly a polyhedral surface also bounded by L and satisfies relation (3.6).

Now let us examine the shifted vertex A'_k. By (3.4), its relative curvature is

$$\omega_f(x_k; S') = \int\limits_{\varphi_{S'}(x_k)} f(\zeta, z'_k, x_k)\, dZ. \tag{3.9}$$

We prove that it varies continuously as the point A_k is shifted continuously from its initial position. The domain of integration of $\omega_{S'}(x_k)$ is the normal image of the vertex A'_k. This image is a convex polyhedron in the plane Z, and it varies continuously as the point A'_k is varied continuously. The normal image of the point A'_k is a convex polyhedron in the plane Z and varies continuously as the point is varied continuously. By Condition IV, the function $f(\zeta, z, x)$ is continuous in z and x for almost all ζ. Therefore, as the point A'_k is varied continuously, the function $f(\zeta, z'_k, x_k)$ varies continuously for almost all ζ. And by Condition II, the function $f(\zeta, z'_k, x_k)$ is dominated by some function $f_0(\zeta)$ summable over a bounded domain.

From these observations and from the convergence of Lebesgue integrals, we conclude that integral (3.8) and, consequently, $\omega_f(x_k; S')$ indeed continuously depends on the position of the point A'_k. Therefore, for sufficiently small displacements of $A_kA'_k$, the value of $\omega_f(x_k; S')$ differs little from $\omega_f(x_k; S^0)$. Thus, by virtue of (3.7),

$$\omega_f(x_k; S') < \nu_k. \tag{3.10}$$

So $\omega_f(x_i; S) \leqslant \nu_i$ for all x_i. However, as A'_k lies below A_k, we have $\sum z_i < \sum z_j^0$. This is contrary to the definition of the surface S^0. So, S^0 is the surface we are searching for.

From the proof, clearly, for the surface S^0 not only does the sum $\sum z_i$ attain its minimum, but also the coordinate z_i of each vertex attains its minimum separately.

3. Theorem 4. *Let D be a convex domain on a plane X, and $f(\zeta, z, x)$ and $A(f)$ have the usual meaning. Then, for every set function $\nu(M)$ defined on D and satisfying an inequality of the type*

$$\nu(D) < A(f),$$

there exists a convex surface S satisfying equation (3.1).

Proof. Let D_i $(i = 1, 2, \dots)$ be polyhedral convex domains included in and converging to a domain D. Let L be an $(n-1)$-dimensional surface admissible as the border of the surfaces S projectable onto D and L_i be a sequence of polyhedral surfaces converging to L and let each L_i be an admissible border of D_i. It is a simple matter to prove that such L_i's do exist.

The function $f(\zeta, z, x)$ is defined for an $x \in D_i$. A function $f(\zeta, z, x)$ whose domain of definition is restricted to D_i is denoted by f_i. Accordingly we now define $A(f_i)$. Clearly,

$$A(f_i) \geqslant A(f). \tag{3.11}$$

By definition,

$$A(f) = \sup \int\limits_Z f_1(\zeta)\, dZ,$$

where supremum is taken over those $f_1(\zeta)$ for which $f(\zeta, z, x) \geqslant f(\zeta)$ for all z less than a certain z_i and for all $x \in D$. Therefore $A(f)$ can only increase with the decreasing D.

Let us construct an infinite sequence of diminishing subdivisions R_j partitioning the domain D into disjoint sets M_{js}. Take a point x_{js} one in each set. Basing on the function $\nu(M)$, define for each subdivision R_j a set function $\bar{\nu}_j(M)$ consisting of point loadings equal to $\nu(M_{js})$ and concentrated at the points x_{js}. Clearly, these functions converge weakly to $\nu(M)$ in the interior of D. In each D_i define a function $\nu_i(M)$, putting

$$\nu_i(M) = \bar{\nu}_i(M) \quad \text{for} \quad M \subset D_i.$$

Then, for every domain G included in D together with its closure $\bar{G}$, the functions ν_i are defined for sufficiently large i and the function ν is also defined. And the functions ν_i converge weakly to ν in the interior of G.

Moreover, $\nu_i(D_i) \leqslant \nu(D)$, and therefore, by virtue of (3.11) and (3.2), there exists a C such that for all i

$$\nu_i(D_i) \leqslant C < A(f_i). \tag{3.12}$$

Hence, by virtue of Theorem 3, we can assert that for every D_i there exists a polyhedral surface S_i having L_i as its border and satisfying the equation

$$\omega_{f_i}(M; S_i) = \nu_i(M_i). \tag{3.13}$$

By virtue of (3.12), from Theorem 1 § 1, we deduce that all these surfaces S_i are uniformly bounded. So from these surfaces we can choose a convergent subsequence which we also denote by S_i and its limit surface by S. The latter satisfies equation (3.1).

Let G be a certain domain included in D together with its closure $\bar{G}$. For sufficiently large i each surface S_i has a part $\bar{S}_i$ projectable onto G and $\bar{S}_i \to \bar{S}$ which is a piece of the surface S projected onto G. Hence, by virtue of Theorem 2 of § 2, the functions $\omega_{f_i}(M; S)$ converge weakly to $\nu(M)$ in

the interior of G. However, the functions $\nu_i(M)$ converge weakly to $\nu(M)$ in the interior of G. As $\omega_f = \nu_i$, we have

$$\omega_f(M; S) = \nu(M).$$

And G being an arbitrary domain, we also have

$$\omega_f(M; S) = \nu(M).$$

This completes the proof of our theorem.

4. Theorem 4 does not assert that the surface S has the given L as an admissible border. This cannot be guaranteed, although by the proof, the surface S is the limit of the surfaces S_i whose borders converge to L.

The point is that in the vicinity of the border the surfaces S_i may become more and more oblique. Then the surface that is their limit in the geometric sense (i.e., in the sense of the topological limit of a sequence of sets) may have a vertical belt, so a part of S projectable onto D may fail to have L as its border. Its border lies lower. This can be illustrated by elementary examples even for the simplest case in which L coincides with the boundary of the domain D.

In other words, the conditions of Theorem 4 are not adequate to guarantee the solvability of the Dirichlet problem for equation (1.1) in the domain D.

Later in Section 4 we examine this question in greater detail and give the sufficiency conditions for the surface S to definitely have a given L as an admissible border.

5. Now we turn to differential equation (1.1)

$$f(\zeta, z, x) \det \|z_{ij}\| = h(x) \tag{3.14}$$

Every set function is surface differentiable almost everywhere. So, for an arbitrarily given $\nu(M)$, we can determine

$$h(x) = \frac{d\nu}{dX}$$

almost everywhere. In the same way, the derivative of $\omega_f(M; S)$ exists almost everywhere. Moreover, every convex function is twice differentiable almost everywhere and the derivative of $\omega_f(M; S)$ is equal to

$$\frac{d\omega_f}{dX} = f(\zeta, z, x) \det \|z_{ij}\|$$

almost everywhere (see [5] whence this result is easily obtained).

From the foregoing it follows that *a surface S satisfying the equation $\omega_f(M;S) = \nu(M)$ almost everywhere satisfies the differential equation*

$$f(\zeta, z, x) \det \|z_i j\| = \frac{d\nu}{dx}.$$

However, this is hardly useful in practice. On adding some point or linear loadings to the function $\nu(M)$, its derivative remains unchanged almost everywhere. Therefore different surfaces, satisfying essentially different equations in set functions, satisfy the same differential equation almost everywhere.

We have to require at least that the set functions be uniquely defined by their derivatives, i.e., be absolutely continuous.

From definition (1.2), it is clear that the requirement of absolute continuity of the function $\omega_f(M;S)$ is equivalent to the requirement of absolute continuity of the area of the normal image of $\varphi_S(M)$ as a function of sets $M \subset D$. This is equivalent to the absolute continuity of the area of the spherical image as a set function defined in the surface S. The area of normal image is simply $\omega_1(M;S)$ as it is given by integral (3.12) for $f = 1$.

Furthermore, we define a generalized solution of the differential equation (3.14) as follows:

A function $z(x)$ is said to be a generalized solution of equation (3.14), if the surface s given by the equation $z = z(x)$ is convex and the area of its normal image $\omega_1(M;S)$ is absolutely continuous.

This definition is equivalent to another definition formulated in pure analytical terms. Such a substitution is justified by the following

Lemma. *Let a surface S be given by the equation $z = z(x)(x \in D)$ and let $\omega_1(D;S)$ be finite-valued. For $\omega_1(M;S)$ to be absolutely continuous, it is necessary and sufficient that there exist a sequence of regular convex functions $z^n(x)$ converging to $z(x)$ and a sequence of increasing closed domains $\bar{G}_p \subset D$ filling D (i.e., $\sum \bar{G}_p = D$) such that, for every $\bar{G}_p$*

$$\lim_{n\to\infty} \int\limits_{G_p} \det \|z_{ij}^{(n)}\| \, dZ = \int\limits_{G_p} \det \|z_{ij}\| \, dX. \tag{3.15}$$

Indeed, in general,[7]

$$\int\limits_M \det \|z_{ij}\| \, dZ \leqslant \omega_1(M;S),$$

[7]Cf. [5] for integral representation of the area of normal (spherical) image, and cf. [1, 3, 7] for weak convergence.

and for $\omega_1(M;S)$ to be absolutely continuous in G, it is necessary and sufficient that

$$\omega_1(G;S) = \int\limits_G \det \|z_{ij}\| \, dX. \tag{3.16}$$

Since the surfaces $S^{(n)}$ are regular, equality (3.16) holds for them for all G_p. And as the area of the normal (or spherical) image converges weakly, the domains G_p can always be chosen such that for each of them

$$\lim_{n\to\infty} \omega_1(G_p;S^{(n)}) = \omega_1(G_p;S).$$

Hence, in addition to condition (3.16) of absolute continuity of $\omega_1(M;S)$, the condition of the lemma is indeed necessary and sufficient for the absolute continuity of $\omega_1(M;S)$.

Now we can formulate a theorem on the existence of a generalized solution to the differential equation (3.14) derived directly from its definition and Theorem 4.

Theorem 5. *Let D, $f(\zeta,z,x)$ and $A(f)$ have the same meaning as in the previous theorems. For a function $h(x)$ defined on a domain D such that $h(x) > 0$ and*

$$\int\limits_D h(x)\, dX < A(f), \tag{3.17}$$

there is a generalized solution to equation (3.14) in D.

Putting

$$\nu(M) = \int\limits_M h(x)\, dX,$$

we find that (3.17) is equivalent to the condition $\nu(D) = A(f)$. Therefore, Theorem 5 follows directly from Theorem 4.

In the next section we study the existence and uniqueness of a solution given certain boundary conditions.

6. For regular functions $f(\zeta,z,x)$ and $h(x)$, the question pertains to the regularity of the solution of equation (3.14). If there are more than two variables; then we know nothing, nor do we know anything about the smoothness of the solution, i.e., one-time differentiability of the solution.

When there are two variables, the solution is smooth, if $h(x)$ is bounded in every closed domain $\bar{G} \subset D$. This follows directly from the smoothness of a convex surface with bounded curvature [6].

If there are more than two variables then no such assertion can be made, as illustrated by examples.

For two variables, some results were found in [9–12] concerning the regularity of solutions to certain special types of equation (3.14). They can probably be easily extended to a more general type of equation (3.14) containing two variables, provided $f_z(\zeta, z, x) \leqslant 0$.

7. Theorems 3–5 stipulate the same sufficient integral condition $\nu(D) < A(f)$ for a solution to exist. In a general case, however, hardly anything can be said about the necessity of this condition. However, under more stringent conditions imposed on the function $f(\zeta, z, x)$, particularly, when it depends only on ζ, the question can in a sense be resolved. We take up this question later in Section 5.

Received June 25, 1957

References

[1] Alexandrov A. D., *Application of the domain invariance theorem to proving an existence theorem*, Izv. Akad. Nauk SSSR Ser. Mat., **3** (1939).

[2] ———, *Existence and uniqueness of a convex surface with given integral curvature*, Dokl. Akad. Nauk SSSR **35** (1942), no. 5.

[3] ———, *Additive set functions in the theory of convex surfaces*, Uchen. Zap. LGU (Annals of Leningrad State University) Ser. Mat., **15** (1948).

[4] ———, *Convex polyhedra*, Gostekhizdat., 1950.

[5] ———, *Almost everywhere existence of the differential of a convex function*, ibid **37** (1939).

[6] ———, *Smoothness of a convex surface with bounded Gaussian curvature*, Dokl. Akad. Nauk SSSR **36** (1942), no. 7.

[7] ———, *Additive set functions in abstract spaces*, Mat. Sb., **13 (55)** (1943), no. 2–3.

[8] Bakel′man I. Ya., *Generalized solutions of the Monge-Ampère equation*, Dokl. Akad. Nauk SSSR **114** (1957), no. 5.

[9] Bernshtein S. N., *Study and integration of second order partial differential equations of elliptic type*, Soobshcheniya Khar′kovskogo Mat. Obshchestva, Vtoraya Ser. **XI** (1908–1909), 1–164.

[10] Pogorelov A. V., *Bending convex surfaces*, Gostekhizdat., 1951.

[11] ———, *Regularity of a convex surface with given Gaussian curvature*, Mat. Sb. **31 (73)** (1952), no. 1.

[12] Bakel′man I. Ya., *A priori estimates and regularity of generalized solutions of the Monge-Ampère equation*, Dokl. Akad. Nauk SSSR **116** (1957), no. 5.

CHAPTER XV

A GENERAL METHOD FOR DOMINATING SOLUTIONS OF THE DIRICHLET PROBLEM

Sibirskiĭ Matematicheskiĭ Zhurnal, vol. 7, No. 3, 486–498 (1966).

1. Principles of the Method

1. For second order differential equations of a general type (generally speaking, elliptic equations) we construct functions that dominate solutions to the Dirichlet problem. A dominant here depends on the domain and certain integral characteristics of the equations, norm of coefficients and, possibly, on the analogous characteristics of the solution itself as well as on the boundary conditions. Concurrently, we derive the necessary conditions for the existence of nonzero solutions to homogeneous problems and certain other results. Our approach is a generalization of the technique applied in [1–3] to quasilinear elliptic equations which was partly generalized in [4]. As is usually done in the works (e.g., [6–8]) dealing with these questions, the papers [1–4] only give an estimate for the maximum modulus of the solution rather than the possible value of the solution at a point in the domain of existence. In this note we precisely derive such estimates.[1]

It suffices to search for a lower bound as an upper bound can be deduced from it by reversing the sign of the solution and suitably modifying the conditions. So, in place of equations, we can consider the inequalities:

$$F(u_{ij}, u_i, u, x) \leqslant 0, \tag{1}$$

where $x = (x^1, \dots, x^n)$ and $n \geqslant 1$ is an integer. Unless otherwise specified, the derivatives $u_1 = u_{x_i}$ and $u_{ij} = u_{x_i x_j}$ are understood to be approximate derivatives which include both the classical and the *generalized*[2] *derivatives*.

[1] Such estimates for linear equations with two variables were studied in [9].

[2] A generalized derivative is the same as a weak derivative.—Eds.

Moreover, in the relations involving these derivatives, a set of zero measure can be discarded. By (u_{ij}) we denote a matrix that consists of the derivatives u_{ij} and $(u_{ij}) > 0$, if the eigenvalues of the matrix are greater than or equal to zero (but not all are zeros).

Let us determine the class of inequalities (1) governed by

Condition (F). *For a function F there exists a function $K(\nabla x, u, x)$ such that if $(u_{ij}) > 0$ and $F \leqslant 0$, then*

$$w \leqslant K(\nabla u, u, x) \tag{2}$$

with $w = \det(u_{ij})$.

As usual, it suffices to restrict ourselves to the values of u_i, u and x for the solutions being estimated (see, Subsection 6 of § 2).

Furthermore, we impose the next

Condition (K). *There exist functions $X(x, u)$ and $U(\nabla u) \geqslant 0$ such that $K \leqslant XU$.*

This condition is almost always satisfied. Then, (2) yields

$$w \leqslant X(x, u)\, U(\nabla u). \tag{3}$$

2. Inequality (1) is assumed to be defined in a bounded domain G and the functions U which are admissible solutions of (3) are assumed to be bounded and lower semicontinuous.

Let $\bar{u}$ be the *convex envelope* of u, i.e., the greatest of the convex functions v such that $v(x) \leqslant u(x)$ everywhere in G. It is defined on the convex hull G^* of the domain G. We say a point x_0 is a *convexity point of the function* u, if $u(x_0) = \bar{u}(x_0)$. At convexity points $(u_{ij}) \geqslant 0$ and $w \geqslant \bar{w} = \det(u_{ij})$ and at other points $\bar{w} = 0$. Therefore $\bar{u}$ also satisfies inequality (3) and estimating the solutions of inequality (1) under Conditions (F) and (K) lies in estimating the convex functions satisfying inequality (3) in G^*.

Therefore it suffices to require that the function u satisfy inequality (3) (and inequality (1)) only at convexity points. The approximate total differentials du and d^2u exist at almost all convexity points for every function. Indeed, if M is the set of convexity points of the function u, then $u = \bar{u}$ on M and there exists a subset $M' \subset M$ such that $\operatorname{mes} M = \operatorname{mes} M'$ and $\bar{u}$ is twice differentiable on M'. Therefore if x is a condensation point of M', then there exist approximate differentials du, d^2u ($du = d\bar{u}$, $d^2u = d^2\bar{u}$) at x. Hence, if we restrict ourselves only to convexity points, the requirement for the existence of derivatives becomes superfluous. For these points it suffices to impose Conditions (F) and (K). Then, if there are no convexity points, inequality (3) is superfluous; but the function u is clearly to be estimated from its values on the boundary of the domain G.

The function u, as in [1], is supposed to obey

Condition (A). *The convex envelope u has an absolutely continuous normal (or, as we say, support) image*[3].

In the sequel this condition is supposed to be satisfied without any reservation. In particular, it is satisfied for functions having generalized derivatives u_{ij} with summable nth powers in every closed subdomain $D \subset G$ [5].

3. We now define the boundary values of $u\big|_\Gamma$ as follows:*For $x \in \Gamma$, $u(x)$ takes the lower limit $\lim_{x'\to x} u(x)$ over all x' tending to $x, x' \in G$.*

We use the notation: $u_{(x)} = \min[u(x), 0]$; v is a point on the unit sphere w in E_n or the corresponding unit vector; $h(x, v)$ denotes the distance between a point x and the supporting plane to $(G \to \Gamma)$ with its outward normal v; $\nabla u = pv,\ p = |\nabla v|$.

Theorem 1. *Let a function u, defined in a domain G and obeying the condition $u\big|_\Gamma \geqslant 0$, satisfy inequality (3) at the convexity points of the function u_-. Then, for the points x where $u(x) < 0$,*

$$\int\limits_{w}^{|u(x)|/h(x,\nu)} \left(\int\limits_0 U^{-1}(p\nu)\, p^{n-1}\, dp \right) d\nu \quad < \int\limits_{G(u<0)} X(x,\, u(x))\, dx, \qquad (4)$$

if the integral on the right, which is taken over that part of G where $u < 0$, is finite.

Proof. By the conditions of the theorem, we can consider the function u_- instead of u, assuming that u_- does not vanish identically. Rewriting inequality (3) as $U^{-1}w \leqslant X$,let us integrate it over the set M of convexity points of u_-. Noting that $w = \frac{\partial(u_1,\ldots,u_n)}{\partial(x^1,\ldots,x^n)}$ and using Condition (A), we can express the integral of $U^{-1}w$ in terms of the variables u_i. Thus we obtain

$$\int\limits_{\psi_{u_-}(M)} U^{-1} du, \ldots, dU_n = \int\limits_M U^{-1} w\, dx \leqslant \int\limits_M X dx \leqslant \int\limits_{G(u<0)} X(x,\, u(x))\, dx. \qquad (5)$$

[3]This condition can be expressed without introducing the convex function $\bar{u}$. A convexity point x of the function u has the following property: there exists an affine function $p_i x^i + q$, called the "support function from below" of the function u, such that $u(x) \geqslant p_i x^i + q$, $u(x_0) = p_i x_0^i + q$. To each such function we assign the point $(p_1, \ldots, p_n)$. The set $\psi_u(M)$ of all points from M at which the functions $p_i x^i + q$ are support functions is called the (lower) *support image* of M for u. Condition (A) is equivalent to the following: if $\operatorname{mes} M = 0$, then $\operatorname{mes} \psi_u(M) = 0$. Note that Condition (A) is global: it is satisfied for the function u on G; but it does not imply that the condition is satisfied for a restriction of u to $G' \subset G$.

Choosing a point x where $u(x) < 0$, let us construct in the space $(x_1, \dots, x_n, z)$, a cone with vertex at $(x, u(x))$ that projects the boundary of the convex hull of the domain G from the vertex. Geometric considerations show the support image of the cone is included in $\psi_{u_-}(M)$ and is the domain bounded by the surface $p = |u(x)|/h(x, \nu)$. Hence, the left side integral in (5) is greater than the integral taken over a similar domain. The latter integral is exactly the left side integral in (4). Hence (4) follows from (5). This completes the proof.

The integral on the right side of (4) is clearly the upper bound of the integral of $X(x, u(x))$ taken over the entire domain G. For this integral to be finite, we have to impose a suitable condition on X. For instance, we may require that for every bounded $v(x)$ the function $K(x, v(x))$ be summable in G. Then inequality (4) will clearly contain the estimate of $|u(x)|$, although implicitly. Even in this form the estimate has a meaning because the solution at a given point is evaluated by means of some of its integral characteristic. However, if X does not depend on u, which happens, say, when F in (1) does not increase with u, then an unconditional estimate is easily obtained (of course, if the left side integral increases indefinitely with $|u(x)|$. In general our method for dominating the solutions of differential equations lies in applying Theorem 1.

4. We now generalize Conditions (F) and (K). Let E_m be an m-dimensional plane, $1 \leqslant m \leqslant n$. If $m = n$, then $E_m = E_n$ represents the entire space and what is said below reduces to what has been said above. For the time being, we omit the index m. Rotating the axes appropriately, we assume that the axes $x^1, \dots, x^m$ all lie in E. Let x_E denote the projection of a point x to E, G_E the projection of G.

The projection of u is defined to be a function u_E on G_E such that $u_E(x) = \inf_{x'_E = x} u(x')$ for every $x \in G_F$. By w_E we denote the determinant of the projection of the form d^2u; if $d^2u \geqslant 0$ this projection is a positive definite quadratic form. The determinant w_E is calculated algebraically from the derivatives u_{ij}. We now introduce

Condition ($\mathrm{F_E}$). *For the function F in (1) and a given space E there exists a function $K(\nabla u, u, x_E)$ such that*

$$w_E \leqslant K(\nabla u, u, x_E), \tag{6}$$

provided $(u_{ij}) > 0$ and $F \leqslant 0$.

First we determine a function $K_0(\nabla u, u, x)$ with the same property as that of K and then take $K = \sup K_0$ over all x with a given projection x_E.

Let $M(u, E)$ denote the set of convexity points of the function u such that $u(x) = u_E(x_E)$. The corresponding points x_E are the convexity points

of u_E. We can assume that u satisfies inequality (6), if the inequality is satisfied almost everywhere on $M(u,E)$ (the set $M(u,E)$ may be empty; then u is clearly estimated from below by $|u_\Gamma|$).

If the function u is differentiable at a point $x \in M(u,E)$, then $\nabla u = \nabla u_E$. Therefore, applying (6) to $M(u,E)$, in K we can substitute u_E for u and ∇u_E for ∇u. This can also be carried out in Condition ($\mathrm{F_E}$). Then (6) takes the form

$$w_E \leqslant K(\nabla u_E, u_E, x_E). \tag{7}$$

Introducing Condition ($\mathrm{K_E}$) similarly as Condition (K), we find

$$w_E \leqslant X(x_E, u_E)\, U(\nabla u_E). \tag{8}$$

Here the right-hand terms all pertain to G_E. However, w_E may fail to exist on a zero-measure set $x \in M(u,E)$ with the projection in E of positive measure, and (8) may be meaningless. In such a case Conditions ($\mathrm{F_E}$) and ($\mathrm{K_E}$) and inequality (8) are applicable to pencils of the planes E.

Let $\{E_m\}_0$ be the set of all planes E_m passing through a given plane E_{m-1} (a point if $m=1$). A *pencil* is a subset of $\{E_m\}_0$, if it has positive measure in the sense of the natural measure in $\{E_m\}_0$.

5. The following theorem, which includes Theorem 1, holds true.

Theorem 2. *For a given function u with $u\big|_\Gamma \geqslant 0$, let the planes E belonging to a pencil $\{E\}$ obey inequalities (8) with X and U, which are in general different for different planes E, at least at the points $x \in M(u_-, E)$. Then, for an x for which $u(x) < 0$, almost all the planes E of the pencil $\{E\}$ obey the inequalities*

$$\int\limits_{\omega_E} \int\limits_0^{|u(x)|/h(x,\nu)} U^{-1}(p\nu)\, p^{n-1}\, dp d\nu \quad < \int\limits_{G_E(u<0)} X(x,\, u_E(x))\, dx, \tag{9}$$

if the integrals on the right are finite. Here ω_E is the unit sphere in E and ν, p and $h(x,\nu)$ have the same meaning as defined in item 3. (If $m=n$, then the "pencil" is reduced to a single "plane" E_n and Theorem 2 is reduced to Theorem 1.)

Supplement. *If the function u is everywhere differentiable and is also twice differentiable and satisfies inequality (8) for a certain plane E everywhere except, at most, at a countable set of points, then inequality (9) is satisfied for this plane E.* (Here it suffices to require that the stipulated conditions be satisfied only on the set $M(u,E)$ and there is no need to suppose that Condition (A) is satisfied.)

Theorem 2 and its supplement are proved by reducing them to Theorem 1. Let w'_E denote $\det(u_{E_{ij}})$, if the axes $x^1, \dots, x^m$ lie in E, i.e., w'_E for u_E is the same as the w for u. If the functions u and u_E are twice differentiable at a point $x \in M(u_-, E)$ and at its projection x_E so that $w_E(x)$ and $w'_E(x_E)$ exist, then, as can be easily verified, $w_E(x) \geqslant w'_E(x_E)$. Therefore, from (8) it follows that

$$w'_E(x_E) \leqslant X(x_E, u_E)\, U(\nabla u_E). \tag{10}$$

Now this inequality is applicable to u_E and resembles inequality (3) which governs u. Therefore, if the conditions of Theorem 1 are satisfied for u_E, then, applying the theorem, we obtain (9). The conditions are 1) u_E must satisfy Conditions (A) and 2) inequality (10) must hold almost at all convexity points of u_E. The latter condition is clearly satisfied, if inequality (8) is not satisfied at most on a point set $x \in M(u_-, E)$ whose projection on E has zero measure in E.

Under the conditions stipulated in the supplement to Theorem 2, condition (2) is clearly satisfied. The first condition is also satisfied. Indeed, the convex function spanned on (u_E) is everywhere differentiable except on a countable point set and has finite upper second derivatives. Hence it satisfies Condition (A) as it clearly follows from a well-known theorem on the absolute continuity of a set function having a finite derivative everywhere (see, [10]).

Thus, the supplement to Theorem 2 has been proved. That the conditions imposed on u_E in the theorem are satisfied is implied in

Lemma 1. *Let a function u in G with $u\big|_\Gamma \geqslant 0$ satisfy Condition (A). Let M be the set of convexity points of u and P a certain subset of M such that* $\operatorname{mes} P = \operatorname{mes} M$. *Then, for almost all the planes E in each pencil,*

(1) *(u_E) satisfies Condition (A) and*
(2) *almost all the convexity points of $(u_E)_-$, where $w'_E > 0$, are the projections of the points $x \in M[(u_E)_-, E) \cap P$.*

Setting aside the proof of the lemma, let us take P to be the set of convexity points where u is twice differentiable and satisfies inequalities (8). Then, for the planes E obeying both the assertions of the lemma, the function $(u_E)_-$ satisfies Condition (A) but satisfies inequality (10) at almost all convexity points. At a point where $w'_E > 0$, this is true by virtue of the second assertion of the lemma and at a point where $w'_E = 0$ this is true simply because $XU \geqslant 0$. This completes the proof of Theorem 2.

6. We now prove Lemma 1. For a given E, suppose that $(u_E)_-$ does not satisfy Condition (A). Then the support image $\psi_{(u_E)-}(G_E)$ includes a set N_E of positive measure with the inverse image Q_E in G_E of measure zero. The set $R_E = \{x \in M(u_-, E)\}$ corresponds to the set Q_E. We may assume

that u is not twice differentiable everywhere on Q_E, even although one of its upper second derivative is infinite (because on a set where all the upper second derivatives are finite the support image is absolutely continuous). Then the function u itself would not be twice differentiable at the points $x \in R_E$.

Clearly, the support image $\psi_{(U_E)_-}(G_E)$ is equal to $\psi_{u^-}(G) \cap E$. Therefore, were the situation mentioned above hold for the planes E of a certain pencil, then in $\psi_{u_-}(G)$ there would be a set $N = \sum N_E$ of positive outer measure on whose inverse image Q the function u would be twice differentiable. However, as the function u has an absolutely continuous support image, its inverse image must have positive outer measure. This contradicts, as shown in 2, that the function u is almost everywhere twice differentiable on the set of its convexity points. This completes the proof of the first assertion of the lemma.

We now prove the second assertion. For a certain plane E let there be a set $Q_E \subset G_E$ of positive measure on which $w_E' > 0$ but the corresponding set $R_E \subset M(u_-, E)$ has no common points with the given set P. Since $w_E' > 0$ on Q_E, the support image $N_E = \psi_{(u_E)_-}(Q_E)$ has positive measure. Were this hold true of the planes E of a certain pencil, then reasoning as above, we would find that the inverse image of the set $N = \sum N_E$ has positive measure and does not intersect the set P. This contradiction completes the proof of the second assertion.

§ 2. General Scheme of the Method

1. Using Theorem 2, we can establish certain versions of the maximum principle. The simplest among them is

Theorem 3. *Under the conditions of Theorem 2, let the functions X and U be such that*

(1) *for every bounded set $v(x)$, $x \in G_E$, the function $X(x, v(x))$ is summable in every closed domain $D \subset G_E$ and*
(2) *the integral of $U^{-1}(pv)$ over every neighborhood of the point p is infinite. Then the function u, not necessarily with $u|_\Gamma \geqslant 0$, attains its greatest lower boundary on Γ*[4].

Suppose that the theorem false. Choose a number $u_0 > \inf u$ so that $u > u_0$ outside a certain closed domain $D \subset G$. Then the function $u' = u - u_0$ satisfies the condition $u'\big|_\Gamma \geqslant 0$ and the inequality $w_E' \leqslant X(x_E, u_E)\, U(\nabla w_E')$.

[4]This means there exists a sequence of points x tending to Γ such that the functions $u(x)$ tend to inf u. If we assume inf $u < 0$, then it suffices to put $v \geqslant 0$ for the condition imposed on the function u.

Moreover, $G(u' < 0) \subset D$. Therefore, applying Theorem 2 to u', we arrive at, by virtue of conditions (1) and (2), a contradiction. This completes the proof of the theorem.

2. Suppose that the function U obeys Condition (U) contrary to Theorem 3, namely, the integral of $U^{-1}(p\nu)$ over an arbitrary bounded domain is finite. This can always be achieved by substituting $(U + \varepsilon)$ for U. (After evaluating $U + \varepsilon$, we can thereafter put $\varepsilon = 0$.) Under this condition, the function

$$V(\xi, x) = \int_{\omega_E} \int_0^{\xi/h(x,\nu)} U^{-1}(p\nu)\, p^{m-1}\, dp\, dv \tag{11}$$

is well defined for all $\xi > 0$ and $x \in G_E$. Using the equality $\eta = V(\xi, x)$, we obtain an increasing function of η

$$\xi = Y(\eta, x), \quad 0 < \eta < V_\infty = V(\infty, x), \tag{12}$$

where V_∞ may either be finite or infinite. Using the function Y, we can rewrite inequality (9) as

$$-u(x) < Y(P, x), \qquad P = \int_{G_E(u<0)} X(x, u(x))\, dx. \tag{13}$$

Now assume that for almost all fixed x the function $X(x, v)$ is bounded for bounded $v \leqslant 0$. (Incidentally, this holds under condition (1) of Theorem 2, if $v \leqslant 0$.) Then we can introduce the following function

$$\bar{X}(x, u) = \sup_{u \leqslant v \leqslant 0} X(x, u'), \tag{14}$$

that is defined for $u \leqslant 0$ and does not increase with u.

In our theorems, it suffices that inequality (8) be satisfied only where $u \leqslant 0$. Hence in (8) we can substitute $\bar{X}$ for X. Making such a substitution, we reduce our problem to the case in which the function X obeys

Condition (X): *The function X does not increase with u (if $u \leqslant 0$) and for every $u_0 = \text{const} \leqslant 0$, the function $X(x, x_0)$ is summable in G_E.*

Theorem 4. *If, under the conditions of Theorem 2, the functions U and X satisfy Conditions (U) and (X); then the integral on the right side of (9), i.e., P in (3), is finite and satisfies the inequality*

$$P \leqslant \int_{G_E} X(x, -Y(P, x))\, dx. \tag{15}$$

If, in addition, X is a strictly decreasing function of u for every fixed x in a certain set of positive measure, then (15) is a strict inequality.

That F is finite is obvious from Condition (X). By this condition the function X does not increase with u, so formula (13) implies $X(x, u(x)) \leqslant X(x, -Y(P, x))$, if $u < 0$. Hence, integrating the left side over G_E $(u < 0)$ and the right side over G_E, we obtain (15).

From Theorem 4 it follows that if the solutions of inequalities (15) are bounded above by a positive number $P_0 < V_\infty$, then (13) yields an estimate: $-u(x) < Y(P_0, x)$, where P_0 no longer depends on u. Thus we obtain unconditional estimates for the solutions of the initial inequality (11) under Conditions $(\mathrm{F_E})$, $(\mathrm{K_E})$, (U) and (X).

However, if inequality (15) does not admit positive solutions, then (13) is not possible, i.e., u cannot take negative values. This "maximum principle" is equivalent to

Theorem 5. *If all Conditions* $(\mathrm{F_E})$ *to (X) are satisfied for a function F, then for the inequality to admit a solution $u(x)$ taking negative values with $u\big|_\Gamma \geqslant 0$, it is necessary that all inequalities (15) admit positive solutions. This assertion also holds true of a strict inequality for those E for which the functions are the same as in the second part of Theorem 4.*

This theorem establishes the necessary conditions for the existence of nontrivial solutions to homogeneous problems and, accordingly, gives a lower bound of eigenvalues. It also establishes sufficient conditions for the solutions to be unique for the Dirichlet problem, at least, stated for linear equations.

3. In case we are only interested in an approximate bound for $\min u(x)$, then in (9) we can substitute the diameter d_E of the domain G_E for the function $h(x, \nu)$. (Since (9) contains only $\nu \in \omega_E$, we have $h(x, \nu) = h(x_E, \nu)$, where $h(x_E, \nu)$ is the distance of x_E from the supporting plane of G_E.) After such a substitution, replacing $|u(x)|$ by $u_0 = |\min u|$, we find the integral on the left side of (9) is reduced to

$$\bar{V}(u_0/d_E) = \varkappa_m \int_0^{u_0/d_E} U^{-1}(p\nu)\, p^{m-1}\, dp, \tag{16}$$

where $\varkappa_m$ is the area of ω_E. Now if Y is the inverse of $\bar{V}$, then, in place of (13), we find

$$u_0 = -\min u(x) < \bar{Y}(P)\, d_E, \qquad P = \int_{G_E(u<0)} X(x, u(x))\, dx. \tag{17}$$

Accordingly, inequality (15) is also simplified.

4. So far we have dealt only with the solutions for which $u\big|_\Gamma \geqslant 0$, but the method can be easily extended to arbitrary boundary conditions. Let u be a solution of inequality (1) and v a convex function such that $(u-v)_\Gamma \geqslant 0$. Then, the function $u' = u - v$ has the following properties

(I) $u'\big|_\Gamma \geqslant 0$;

(II) its convexity points are the convexity points of u; and

(III) if $(u'_{ij}) > 0$, then $(u_{ij}) > 0$ and $w_E \geqslant w'_E$.

From (II) it follows that u' satisfies Condition (A) of absolute continuity of the support image, and from (III) it follows that if (6) is satisfied for u (when $(u_{ij}) > 0$), then

$$w'_E \leqslant K(\nabla u' + \nabla v,\, u' + v,\, x_E) = K'(\nabla u', u', x_E)$$

for $(u'_{ij}) > 0$. Therefore, considering only this inequality, it remains to apply the following conclusions. In particular, if ∇v is not contained in K, because either v is constant or K in general does not depend on ∇u, then inequality (8) takes the form: $w'_E \leqslant X(x, u' + v)\, U(\nabla u')$, so there is no need to search for new functions X and U for K'. We can take the convex envelope $u\big|_\Gamma$ as v. Then the estimate of $u' = u - v$ gives an estimate for the difference between u and its boundary values $u\big|_\Gamma$.

5. In this section we give a general method for finding the estimates of solutions and the conditions for the existence of nontrivial solutions to an inequality of the type (1) under Condition ($\mathrm{F_E}$). If the function F is given; then to find K and thereby to verify whether Condition ($\mathrm{F_E}$) is satisfied, we have first to find $\min F(u_{ij}, u_i, u, x)$ for $(u_{ij}) > 0$ and given $w_E, \nabla u$, u and x_E. Suppose that this minimum is $H(w_E, \nabla u, u, x_E)$. Then $K(\nabla u, u, x_E)$ can be determined as the least upper bound of those w_E for which $H \leqslant 0$ (for given ∇u, u and x_E). Accordingly, if $w_E > K$, then we have $F \geqslant 0$ for $w_E > F$, i.e., $w_E \leqslant K$ for $F \leqslant 0$.

As has already been noted, Condition (U) imposed on the function U can always be realized. Condition (X) is equivalent to restricting the function F so that certain norms of the coefficients may be finite. It is a simple matter to impose such restrictions.

It only remains to examine Condition ($\mathrm{K_E}$). There is no general test to verify whether or not this condition is satisfied, but in practical cases it is evidently satisfied. If it is satisfied, then the choice of the functions X and U is generally not unique. Therefore, to derive more accurate estimates, the choice has to be made suitably. Once the functions X and U were chosen and accordingly the functions V and Y are determined, it only remains to apply Theorem 4 and its corollaries.

6. We now give certain examples that illustrate the generality of Condition ($\mathrm{F_E}$).

In fact, we encounter a much stronger Condition ($\mathrm{F}_{\mathrm{E}}^*$): the function F is such that for every $f(\nabla u, u, x)$, the sum $F + f$ satisfies Condition (F_{E}). This is equivalent to the existence of a function $H(w_E, \nabla u, u, x_E)$ such that H tends to infinity as w_E tends to infinity and if $(u_{ij}) > 0$ the function F is greater than H.

(I) For $(u_{ij}) > 0$ let the function F be differentiable with respect to u_{ij} and let $(F_{u_{ij}}) \geqslant (a^{ij}) > 0$, where $a^{ij} = a^{ij}(\nabla u, u, x)$, i.e., F is elliptic and may possibly degenerate. Now if $(u_{ij}) > 0$, for every E where the axes $x^1, \dots, x^m$ lie, we have

$$F(u_{ij}, u_i, u, x) \geqslant m a_m^{1/m} w_E^{1/m} + F(0, u_i, u, x),$$

where $a_m = \det(a^{ij})$ $(i, j < m)$. Therefore Condition ($\mathrm{F}_{\mathrm{E}}^*$) is satisfied if $a_m \neq 0$. However, if $a_n \neq 0$, then ($\mathrm{F}_{\mathrm{E}}^*$) is satisfied for all E_m $(1 \leqslant m \leqslant n)$.

(II) Let $w_{i_1 \dots i_k j_1 \dots j_k}$ be the minor of the matrix (u_{ij}) formed from its rows $i_1 < \dots < i_k$ and columns $j_1 < \dots < j_k$. There are $(C_n^k)^2$ such minors. Let us denote them by $\underset{k}{w}_{ij}$, where $i, j = 1, \dots, C_n^k$. In particular, for $k = 1$, we have $\underset{1}{w}_{ij} = u_{ij}$ and for $k = n$, we have $\underset{n}{w}_{ij} = w$. Consider the function $F = a^{ij} \underset{k}{w}_{ij}$, $a^{ij} = a^{ij}(\nabla u, u, x)$, under the condition that $a^{ij} \underset{k}{w}_{ij} \geqslant 0$ if $(u_{ij}) > 0$ (for this to hold, it suffices $(a^{ij}) > 0$). Then for $(u_{ij}) > 0$ and $m \geqslant k$, we have

$$F = a^{ij} \underset{k}{w}_{ij} \geqslant C_m^k \, a_m^{1/C_m^k} \, w_E{}_m^{k/m}, \qquad a_m = \det(a^{ij}),\ i_s,\ j_s \leqslant m.$$

Therefore Condition ($\mathrm{F}_{\mathrm{E}}^*$) is satisfied if $a > 0$.

(III) Condition ($\mathrm{F}_{\mathrm{E}}^*$) may be satisfied even if F is not elliptic, provided $(u_{ij}) > 0$. Example: let $F = (a^{ij} \underset{k}{w}_{ij})^2 - c(a^{ij} \underset{k}{w}_{ij})$ under the same conditions as in the previous example, and $c = c(\nabla u, u, x)$.

§ 3. Estimates Under Certain Special Assumptions

1. Suppose that the function $U(p\nu)$ obeys the condition (U^k): $U(p\nu)$ is dominated by a function $U(p)$, which depends only on p, such that for a certain k the function $U(p)p^{k-m}$ does not increase. Then in the previous results we can substitute $U(p)$ for $U(p\nu)$. (Such a condition is equivalent to suitably restricting the growth of $K(\nabla u, u, x)$ with $|\nabla u|$). This happens, say, when in the quasilinear equation $a^{ij} u_{ij} + f = 0$ the growth of $a^{-1/n} f$, $a = \det(a^{ij})$ with ∇u is restricted.

Lemma 2. *If the function $U(p)p^{k-m}$ does not increase, then*

$$V(\xi, x) = \int\limits_{\omega_E} \int\limits_0^{\xi/h(x,\nu)} U^{-1}(p)\, p^{m-1}\, dp\, d\nu \;\geqslant\; \varkappa_m \int\limits_0^{\xi/h_k(x)} U^{-1}(p)\, p^{m-1}\, dp, \quad (18)$$

where $h_k(x)$, for $k \neq 0$, is the power mean of the distances $h(x, \nu)$:

$$h_k(x) = \left[\frac{1}{\varkappa_m} \int\limits_{\omega_E} h^{-k}(x, \nu)\, d\nu \right]^{1/k} \quad (19)$$

and, for $k = 0$, is the geometric mean

$$h_0(x) = \exp \frac{1}{\varkappa_m} \int\limits_{\omega_E} \ln h(x, \nu)\, d\nu. \quad (20)$$

Since h_k depends on E, we should have written $(h_k)_E$.

To prove this lemma, denoting the right side of (18) by $\bar{V}(\xi/h_k(x))$, let us consider the function $\bar{V}(\zeta^{1/k})$ for $k \neq 0$. Calculating its derivative with respect to ζ, we find that, by virtue of the condition imposed on $U(p)$, this derivative is a nondecreasing function of ζ. Consequently, the function $\bar{V}(\zeta^{-1/k})$ itself is convex. And the mean of a convex function is not less than its value at the corresponding mean of the argument. This precisely leads to (18). For $k = 0$ the same reasoning is to be applied to $V(e^{\zeta})$.

2. In the previous results, substituting $V(\xi, x)$ by $\bar{V}(\xi/h_k(x)$ and denoting the inverse of $\bar{V}$ by $\bar{Y}$, we find that inequalities (13) and (15) acquire a simpler form:

$$-u(x) < \bar{Y}(P) h_k(x), \quad P \;=\; \int\limits_{G_E(u<0)} X(x, u(x))\, dx, \quad (21)$$

$$P \leqslant \int\limits_{G_E} X(x, -\bar{Y}(P) h_k(x))\, dx. \quad (22)$$

Note the function $\bar{Y}$ here is the same as that in (17). Thus we arrive at

Theorem 6. *Under Conditions* $(\mathrm{F_E})$–$(\mathrm{U^k})$, *estimate (21) is a solution of inequality (1) obeying the boundary condition* $u|_\Gamma > 0$, *where the function* $h_k(x)$ *depends, for a given* k, *only on the domain* G_E, *and satisfies inequality (22).*

3. The simplest example is the case in which $U(p\nu) <$ const, so we can take $U(p) = 1$ by assigning a suitable factor to the function X in (8). Then the function $V(\xi, x)$ is calculated in explicit form: $V = \tau_m \xi^m / h_m^m(x)$, where $\tau_m = m^{-1}\varkappa_m$ is the volume of the unit ball. Therefore inequality (21) takes the form

$$-u(x) < \tau_m^{-1/m}\, p^{1/m}\, h_m(x). \tag{23}$$

This particular case clearly holds, if F in (1) does not depend on u_i. However, a more general case can also be reduced to this case, since our theorems require that inequality (8) be satisfied only at the convexity points of the function u_-. Hence we have

Lemma 3. *For an arbitrary fixed* u_E *and almost every fixed* x_E, *let the function* $K(\nabla u_E, u_E, x_E)$ *in (7) be bounded on every bounded set of* $|\nabla u_E|$. *Then (7) implies that at those convexity points of the functions* u_- *where* $u(x) = u_E(x_E)$ *the following inequality is satisfied:*

$$w_E \leqslant X(x_E, u_E), \tag{24}$$

where the function is defined as follows. Assuming that (the subscript E *is omitted)*

$$\sup_{|\nabla u| \leqslant p} K(\nabla u, u, x) = M(p, u, x), \tag{25}$$

and denoting by $r(x)$ *the distance of the point* $x = x_E$ *from the boundary of the convex hull of the domain* G_E, *we take*

$$X(x, u) = M(|u_-| \,/\, r(x), u, x). \tag{26}$$

From the definition of the function M it is seen that, if $w_E \leqslant K$, then $w_E \leqslant M(|\nabla u|, u, x)$. However, at the convexity points x of the function $(u_-)_E$, we have

$$|\nabla u| = |\nabla u_-| \leqslant \frac{|u_-(x)|}{r(x)}. \tag{27}$$

Hence, by virtue of definition (26), we obtain (24).

Owing to the denominator $r(x)$ in the arguments of M, Condition (X) for the function X defined by (26) becomes very strong. Such a strengthening is however useful in certain problems.

4. Notice that the functions h_{kE} are the same for the domain G and for its convex hull G^*. And for a given E the value of $h_{kE}(x)$ is the same for all x with a given projection x_E (since this property is exhibited by the distance $h(x,\nu)$, $\nu \in w_E$). Therefore the function h_{kE} can be regarded as belonging to G^* or its projections G^*_E.

We study the properties of the function h_k in a different note. Here we only mention that the bounds of h_{kE} are expressed through more illustrative geometric characteristics of the domain G^*, say, estimate of $\max h_{nE}$ through the volume of G^*_E. Our results give suitable estimates for the minimum of the solution of u. Second, it is a simple matter to prove that for $k \geqslant -1$ the function h_k is concave: $d^2 h_k < 0$. Furthermore, we study what are the points on the boundary of the domain G^* on approaching which the function $h_k(x)$ clearly tends to zero and what are the convergence rates. In combination with our estimates, this would give the rate with which the solution approaches its boundary values.

§ 4. Generalization to Unbounded Domains

1. The requirement that the domain G where inequality (1) holds can be replaced by a more general requirement: the domain G should be reducible to a bounded domain by some projective transformation. It is natural to call such a domain *projectively bounded*.

Accordingly, the requirement that the function u be bounded below is replaced by a requirement that the function u be *projectively bounded below*, i.e., under a projective transformation of the space $(x^1,\dots,x^n,u)$ which sends the space $(x^1,\dots,x^n)$ into itself and the domain G into a bounded domain, the function u should be transformed to a function bounded from below. The condition $u\big|_\Gamma \geqslant 0$ is also to be interpreted projectively: under a projective transformation the function u is transformed into a function $\widetilde{u}$ such that $\widetilde{u}\big|_{\widetilde{\Gamma}} \geqslant 0$ where $\widetilde{\Gamma}$ is the boundary of the bounded domain $\widetilde{G}$ into which the domain G is transformed.

These conditions can be expressed differently. The condition of projective finiteness of the domain G, as is readily seen, is equivalent to the requirement that its convex hull $\widetilde{G}$ should not include straight lines. If this condition is satisfied, then the projective boundedness from below of the function u is equivalent to the existence of a constant such that $u(x) > A(|x| + 1)$. Projectively, the boundary condition $u\big|_\Gamma \geqslant 0$ implies the following. For every $\varepsilon > 0$ there exists a closed bounded domain $D \subset G$ such that the function $u(x)$ is greater than $-\varepsilon(|x|+1)$ in D. Finally, we introduce some natural consistency condition. If $G + \Gamma$ has no supporting plane with outward normal, then $h(x,\nu)$ is assumed to be infinite.

Theorem 7. *Theorems 1, 2, 4 and 5 hold valid for a projectively bounded domain G, if the boundary conditions are interpreted projectively as defined*

above, all other conditions being left unchanged. Only such planes E are admissible for which G_E is also projectively finite. In Condition (X) an affine function is to be substituted for the constant u_0.

The spherical image of the boundary of a convex domain that does not include straight lines is a convex domain σ on the unit sphere ω that may include the entire or a part of the boundary. Since for $\nu \notin \sigma$ we suppose $h(x, \nu) = \infty$, the integrals on the left sides of inequalities (4) and (9) are actually taken on the domain σ, $\sigma \cap E_m$, where σ is the spherical image of the boundary G^*.

Theorem 7 is proved first by proving Theorem 1 under the conditions adopted here. Since the convexity properties are preserved under a projective transformation, the geometric part of the proof remains unchanged. However, the analytical part is in general not related to the boundedness of the domain.

After generalizing Theorem 1 in this way, we can generalize Theorem 2 by simply repeating the above procedure. Of course, we have to restrict ourselves to the planes E for which G_E is projectively bounded; otherwise it would not be possible to apply Theorem 1. Thereafter Theorems 4 and 5 are readily generalized.

2. Let, as before, σ be the spherical image of the boundary G^* and C an infinite cone that projects the domain σ from the center of the sphere ω. Let this center be at the origin. Since the integrals on the left sides of (4) and (9) are integrated with respect to u over $\sigma \cap E$, the terms $p\nu = \nabla u$ in these integrals belong to the cone C. Hence it suffices to consider the function $U(p\nu)$ on the cone C.

Theorem 8. *If the function $U(p\nu)$ is restricted only to the values of $p\nu$ belonging to the cone C, then Theorems 3 and 6 hold for a projectively bounded domain under the conditions of Theorem 7, provided the functions h_k in Theorem 6 are well defined (i.e., not infinite).*

Theorem 8 means the following. Theorem 3 requires the integral of $U^{-1}(p\nu)$ taken over an arbitrary neighborhood of the point $p = 0$ be infinite. Here it is assumed that the neighborhood of this point lies in the cone C. Under this condition, Theorem 3 follows directly from the generalized Theorem 2.

Furthermore, Lemma 2 implies that the mean $h_k(x)$ of the distances $h_k(x, \nu)$ is taken over $\sigma \cap E$ but not over the sphere ω_E. The integrals are taken over $\sigma \cap E$ and divided by the area of $\sigma \cap E$. Accordingly, the volume τ_m of the unit ball in item 3 of § 3 is to be replaced by the volume of the spherical sector resting on $\sigma \cap E$.

Under such an interpretation, all the conclusions of § 3 remain unchanged, of course, only for those G_E for which the function $h_k(x)$ is well defined

(finite). This is true, if $k > 0$. However, for $k \leqslant 0$, the function h_k may be infinite. It is a simple matter to show that $h_k(x)$ is either finite for all $x \in G^*$ or infinite.

Received June 30, 1965

References

[1] Alexandrov A. D., *Uniqueness conditions and estimates of the solution of the Dirichlet problem*, Vestnik LGU **3** (1963), no. 13, 5–30.

[2] ———, *Certain estimates concerning the Dirichlet problem*, Dokl. Akad. Nauk SSSR **134** (1960), no. 5, 1001–1004.

[3] Bakel′man I. Ya., *To the theory of quasilinear elliptic equations*, Sibirsk. Mat. Zh. **2** (1961), no. 2, 179–186.

[4] Alexandrov A. D., *The support image method for studying solutions of boundary value problems*, Materialy Sovetsko-Amerikanskogo Simpoziuma po Uraveniyam v Chastnykh Proizvodnykh (Proceedings of the American-Soviet Symposium on Partial Differential Equations) (1963), Novosibirsk.

[5] ———, *An investigation into the maximum principle. Part V*, Izv. Vyssh. Uchebn. Zaved. (1960), no. 5, 16–26.

[6] Maz′ya V. G., *Certain estimates for the solutions of second-order elliptic equations*, Dokl. Akad. Nauk SSSR **137** (1961), no. 5, 1057–1059.

[7] Stampacchia G., *Contributi alla regolarizzazione delle soluzioni dei problemi al contorno*, Ann. Scuola Norm. Sup. (1958), no. 12, Pisa, 88–92.

[8] Weinberger H. F., *Symmetrization in uniformly elliptic problems*, Studies in Math. Analysis and Related Topics (1962), University Press, Stanford, 88–92.

[9] Frasca M., *Un problema variazionale per operatori ellittici*, Matematiche (1963), no. 18, 1–11.

[10] Saks S., *Theory of the Integral*, Izdat. Inostr. Lit., Moscow, 1949; Russian translation.

CHAPTER XVI

ON THE PRINCIPLES OF RELATIVITY THEORY

Dedicated to S. V. Vallander

Vestnik Leningradskogo Universiteta, Mat., No. 19, 5–28 (1976).

In the fifties, while investigating into the principles of relativity, I often discussed with Sergeĭ Vasil′evich Vallander the problems set forth here and my conclusions. For this there were greater opportunities as we could then meet regularly in our office: in the spring of 1952 I was appointed as Rector of Leningrad University and invited Sergeĭ V. Vallander to assist me as Deputy Rector in charge of academic research. Amidst the overwhelming flow of administrative drudgeries, we found a few moments for our scientific work.

Sergeĭ Vasil′evich was a man of great acumen capable of delving deep to the bottom of problems. In addition, he was disposed to criticism rather than to compliments, so discussion with him was interesting and helpful, a piercing and fault finding critic is indeed a friend. A topic for particular discussion were furnished by my works [2, 3]. Here we dwell on [3].

As early as in 1949 (cf. [1]), I derived the Lorentz transformations exclusively from the law of constancy of the velocity of light (without any assumptions on continuity, etc.), or which is the same thing, from the conservation of a system of light cones. I greatly improved this result in [3] where the Lorentz transformations was also deduced from the conservation of a system of light cones K_a in the four-dimensional space-time. In other words, in the event space a cone K_a with vertex a is formed by those events that an event a influences (or may influence), i.e., an event at the "world" point a. "Influence" may be understood as transfer of energy-momentum. The conservation of a system of cones implies the conservation of influence relationship or, in other words, the causality principle. So, we can say that the Lorentz transformation was derived from the cause-and-effect relation-

ship.

Both for "influence cones" and for light cones, it was assumed that the question was about a one-to-one mapping of four-dimensional space onto itself which preserves the system of cones (a cone is mapped onto a cone of the same type).

That I work only with mappings of the entire space onto itself, Sergeĭ Vasil′evich believed as a demerit of my results, emphasizing the importance of studying a mapping of a bounded domain rather than the whole space. His viewpoint is that when we study a mapping of the whole space we actually treat the universe as entity and this is rather doubtful both from empirical and from philosophical viewpoints.

In this way my discussions with Sergeĭ Vasil′evich brought forth the problem: what are those one-to-one mappings of a domain of four-dimensional (in general, n-dimensional) space that send a circular cone onto a circular cone; of course, we do not mean the entire cone but only a part of the cone in a given domain and, consequently, in that domain onto which it (i.e., the part of the cone) is mapped. (The exact formulation of the problem and the results are given later.)

More than twenty years have elapsed but the question still remains unresolved. The invitation to contribute to Vallander's commemorative volume stimulated me to write down the ideas I had in mind and thus to complete the solution of this problem. This solution constitutes the topic of this note.

My results [1, 3] were also reported later by others [5, 6] who did not notice them. In [6] the authors, being impressed by their results, i.e., the "Lorentz group follows solely from the principle of constancy of the velocity of light," italicized these words and put an exclamation point. However, twenty years ago I had reported how I derived this result [1].

Although in this paper we essentially derive a local result for bounded domains; nonetheless, we obtain from it a global consequence concerning the possible models of the universe.

§ 1. Formulation of the Results

1.1. We consider a Minkowski space, a pseudo-Euclidean space R in whose metric form, one square is of one sign and all others are of the opposite sign. We suppose its dimension $\dim R \geqslant 3$, not excepting $\dim R = \infty$. In view of this assumption we give the definition of the space R. This assumption however does not complicate the proof.

An *affine space* A ($\dim A = \infty$ is not excluded) is simply a linear space over the field of real numbers, provided the translation $x \to x + a$ is defined in it (accordingly $\dim A$ is understood as the maximum cardinality of the set of independent vectors). Therefore, taking an arbitrary point $O \in A$ as the origin, i.e., assigning it the zero vector, we obtain a mapping of the

space A onto a linear space: to every point $x \in A$ there corresponds a vector O_x. Furthermore, we assume that the origin O is fixed and the symbols x, y, a, etc. denote both the points of the space A and the vectors O_x, O_y, O_a, etc. A locally convex topology is assumed in the space A.

Through the origin O let us draw a line L and a plane E of codimension 1 which does not include the line L. Every vector is uniquely decomposed along L and E:

$$x = x_L + x_E.$$

Suppose that a scalar product $x_E y_E$ is defined in the plane E. Introduce an affine coordinate x_0 on the line L: we can assume that the square of the vector x_L^2 is equal to x_0^2. We define the *pseudosquare* of a vector $x = x_L + x_E$ as follows:

$$x^2 = x_L^2 - x_E^2 = x_0^2 - x_E^2.$$

Accordingly, the *pseudoscalar product* is

$$xy = x_0 y_0 - x_E y_E.$$

Thus we obtain a pseudo-Euclidean space R. The same space is also obtained if x^2 is multiplied by a constant factor $\lambda \neq 0$, $\lambda < 0$ being not excepted. We choose the sign so that one square is positive and retain it in the sequel.

1.2. We consider six relations between two points $x, y \in R$ of which three are symmetric:

$$(x-y)^2 = 0, \tag{I}$$

$$(x-y)^2 \geqslant 0, \tag{II}$$

$$(x-y)^2 > 0, \tag{III}$$

and other three (I^+)–(III^+) derived from them by adding the condition $x_0 \leqslant y_0$ are antisymmetric. In particular, relation (I) signifies that the vector xy is *isotropic*, whereas (I^+) additionally implies that its x_0-component is nonnegative.

The problem we solve in this note lies in studying the mappings that preserve these relations together with their negations, i.e., for instance, $(x-y)^2 = 0$ and $(x-y) \neq 0$.

We can distinguish four kinds of elementary transformations possessing this property:

(1) A homothety H: $x' = \lambda x + a$ ($\lambda \neq 0$, $\lambda < 0$ being not excluded); assuming $\lambda = 1$, we put translations under homothety.

(2) A Lorentz transformation L, i.e., a one-to-one linear mapping that preserves x^2 and the inequality $x_0 > 0$ for $x^2 > 0$ (that does not invert time).

(3) An inversion I: an inversion with center a is the transformation:

$$x' = I_a(x) = \frac{x-a}{(x-a)^2} + a.$$

(4) A *special double inversion*, *sd*-inversion, J:

$$x' = J_{ca}(x) = \frac{(x-a) + c\,(x-a)^2}{1 + 2c\,(x-a)} + a$$

where the vector $c \neq 0$ is such that $c^2 = 0$ and in other respects is arbitrary. The vector a, as in 3), is arbitrary. Transformation 4 is called special because it differs from the product of two inversions with centers a and $(a-c)$ only in a translation (cf. 2.2).

Unlike transformations (1) and (2) defined on the entire space, the inversion I_a is not defined on the cone C_a: $(x-a)^2 = 0$, while the *sd*-inversion J_{ca} is not defined on the plane P_{ca}: $1 + 2c\,(x-a) = 0$. If $\dim R < \infty$, then Lorentz transformations map the space R onto itself; but if $\dim R = \infty$, then they may map the space R onto a linear subspace.

The properties of these transformations we need are stated in § 2. Here we only note

Assertion 1. *A mapping f that is a combination of transformations (1)–(4) preserves relation (I) and its negation for all pairs (x, y) if and only if $f(x)$ and $f(y)$ are defined, and preserves relations (II) and (III) together with their negations, if the points x, y belong to a simply connected component of the set on which f is defined. Under the last condition relations (I$^+$), (II$^+$) and (III$^+$) along with their negations are preserved if and only if there is an even number of inversions and homotheties with $\lambda < 0$ in the mapping f.*

1.3. Its converse is also true, namely,

Theorem 1. *If a one-to-one transformation f: $G \to R$ of a domain $G \subset R$ preserves one of the six relations* (I)–(III$^+$) *along with its negation, then the mapping f is a combination of transformations (1)–(4), i.e., f is either a Lorentz transformation with a homothety or can possibly be represented as such a transformation with the addition of a simple inversion or an sd-inversion.*

In other words, the mapping f is one of the three types: HL, HLI or HLJ. In the last two cases the mapping f can also be reduced to the type IHL or JHL (with other H, I and J, except when HL is an identity. Also notice that $HL = LH$).

Relations (I^+)–(III^+) are preserved, if the homothety coefficient λ in HL and HLJ is greater than zero, and in the case of HLI, if $\lambda < 0$.

The representation of f in the form of HL or HLI (or IHL) is unique. However, the representation HLJ (just like JHL) is not unique: in addition to HLJ, any other representation $H'L'J'$ is also possible, where J' has the same singular plane as before P: $1 + 2c\,(x - a) = 0$ (so $c' = \lambda c$ and $1 - 2c'a' = \lambda(1 - 2ca)$).

The theorem for (I^+)–(III^+) is a direct corollary of the theorem for relations (I)–(III). Indeed, for every pair (x, y) we have either $x_0 \leqslant y_0$ or $y_0 \leqslant x_0$. Therefore, if the relations, say, $(x-y)^2 = 0$ and $(x-y)^2 \neq 0$ along with the inequality $x_0 \leqslant y_0$ are preserved, they are also preserved regardless of the inequality because if $y_0 \leqslant x_0$ for a given (x, y), then substituting x for y we obtain $x_0 \leqslant y_0$.

On the strength of this remark, we could as well discard the cases (I^+)–(III^+). However, they have a bearing on the physical interpretation that we give in Subsection 1.6.

We demonstrate Theorem 1 for relation (1) in § 3 and for (II) in § 4 and 5; the proof for relation (III) is omitted due to lack of space.

1.4. We now formulate Theorem 1 in geometrical terms. Defining a relation between a pair of points x, y is equivalent to assigning to each point x the set of all y's that correspond to x by the given relation. To relations (I)–(III), there correspond the following cones:

($\mathrm{I_M}$)	$C_x = [\,y\colon (x-y)^2 = 0\,]$	the isotropic cone with apex x,
($\mathrm{II_M}$)	$K_x = [\,y\colon (x-y)^2 \geqslant 0\,]$,	and
($\mathrm{III_M}$)	$Q_x = [\,y\colon (x-y)^2 > 0\,] \cup (x)$	an open cone; if its apex x is included, (III) is supplemented with the relation $y = x$.

The sets C_x^+, K_x^+ and Q_x^+ that correspond to relations (I^+)–(III^+) are obtained by adding the condition $x_0 \leqslant y_0$, so we have ($\mathrm{I_M^+}$) $C_x^+ = [y\colon y \in C_x,\, y_0 \geqslant x_0]$. The cases ($\mathrm{II_M^+}$) K_x^+ and ($\mathrm{III_M^+}$) Q_x^+ are defined analogously. These sets are also cones but they are "single" cones unlike the "double" cones C_x, K_x and Q_x. We say the former are the "nappes" of C_x, K_x and Q_x, those nappes on which $y_0 \geqslant x_0$.

Theorem 1 is reformulated in terms of the cones ($\mathrm{I_M^+}$)–($\mathrm{III_M^+}$) as:

Theorem 2. *Assume that one of the six relations* ($\mathrm{I_M}$)–($\mathrm{III_M^+}$) *is selected with M_x standing for the corresponding cone with vertex x. If $f :\rightarrow \mathbb{R}$ is a one-to-one mapping such that*

$$f(M_x \cap G) = M_{f(x)} f(G) \tag{A}$$

then f is a mapping of one of the types listed in Theorem 1.

Theorem 2 is equivalent to Theorem 1 on the strength of the following observation. Suppose that $\mathcal{R}$ is one of the relations under study and $M_x = [\,y\colon \mathcal{R}(x,y)\,]$, then the conservation of the relation $\mathcal{R}$ under the mapping f implies

$$f(M_x \cap G) \subset M_{f(x)} \cap f(G).$$

However, the conservation of the negation of R implies the same for the inverse map f^{-1}, i.e. $M_{f(x)} \cap f(G) \subset f(M_x \cap G)$. Thus, (A) is equivalent to the conservation of R and its negation and, consequently, we find that Theorem 2 is equivalent to Theorem 1.

1.5. A theorem on conformal space. A space can be completed with a sl cone at infinity, the image of its singular cone under every inversion (cf. [7]) in which dim $R < \infty$; for $\dim R = \infty$ the completion is defined literally as we did). Thus we obtain a space C in which transformations (1)–(4) act as one-to-one mappings (which are continuous, if a natural topology is defined in C).

The space C is called *conformal* because conformal mappings of domains into R (these are transformations (1)–(4) and their combinations) are just the mappings HL, HLI and HLJ, so the completion of R to C regularizes conformal mappings. We know that for $\dim R < \infty$ the transformations HL, HLI and HLJ are conformal (cf. [7]). Theorem 2 is helpful in proving this for $\dim R = \infty$ and thereby in deriving even a stronger result, if the conformality condition is radically weakened. However, this is beyond the scope of our note.

When R is completed to C, the cones C_x and K_x are also completed, so these complete cones are supposed to exist in the space C. Every isotropic line is completed with a point "at infinity" and so becomes closed.

The space C, on excluding some cone C_x, is transformed into the space R. A mapping which in R is a reflection in a plane is an inversion in C. Observe that *sd*-inversions lose meaning as singular inversions because inversions have no singularities[1] in C. It is therefore natural to distinguish on C only three types of mappings: (1) homotheties H^c, (2) Lorentz transformations L^c without reflections and (3) inversions I^c, i.e., (homeomorphic) mappings of C onto itself which, when a suitable cone is excluded from C, are transformed to inversions. (For $\dim C < \infty$ this is the mapping of C onto itself.)

The following theorem holds:

[1]In the space R *sd*-inversions are distinguished as singular in the sense that they are defined on half-spaces, while inversions are defined on the domains into which the space is subdivided by singular cones.

Theorem 3. *A mapping $f\colon G \to C$ of a domain $G \subset C$ that satisfies the conditions of Theorem 2 is either $H^c L^c$, or $H^c L^c$ with the addition of one or two inversions I^c.*

In the finite-dimensional case the mapping f is represented as the product of finitely many inversions (not more than $n+2$ if $\dim C = n$).

Theorem 3 is easily derived from Theorem 2 (cf. § 6). Theorem 2, in turn, directly follows from Theorem 3: it suffices to complete the space R stated in Theorem 2 to C and then apply Theorem 3.

Theorem 3 needs the following remark. Since isotropic lines in C are closed, the nappes C_x^+, K_x^+, Q_x^+ of the cones C_x, K_x, Q_x can be distinguished only locally. Accordingly, the condition (A) of Theorem 2 is applicable only locally for these nappes. Further, as already noted, these nappes in our theorems are unimportant. Moreover, the condition (A) as applied to C_x is equivalent to a combination of condition (A) for C_x together with the conservation of orientation or direction of passage, which is first specified for some isotropic line and then extended by the continuity condition to other isotropic lines, the generators of the cones C_x. In the same way we can reformulate the condition for the cones K_x^+ and Q_x^+, if a local order is introduced on the cones K_x and Q_x that corresponds to the direction of passage over the generators of the cone C_x.

The space C is homeomorphic to the product of a sphere and a circle.[2] Therefore there exists a covering space $\bar{C}$ homeomorphic to the product of a sphere and a line. Naturally, a geometry that locally coincides with the geometry in C is induced in $\bar{C}$. Isotropic lines in $\bar{C}$ are no longer closed, so it its possible to distinguish the nappes C_x^+, K_x^+ and Q_x^+. The mappings H^c, L^c and I^c are extended onto $\bar{C}$. And from Theorem 3 we derive

Theorem 4. *Theorem 3 also holds for the space $\bar{C}$ and the condition governing the nappes has not only local but also global meaning.*

1.6. A physical interpretation. In order to ascribe a physical meaning to relations (I)–(III+) and to Theorems 1–3, let us consider the event space, or which is the same thing, the space-time continuum in which each point is associated with some event. Accordingly, x and y denote events, while the coordinate x_0 stands for the time in a certain reference frame.

Relations (I)–(III$^+$) signify that x influences or may possibly influence y, i.e., energy-momentum is transferred or may be transferred from x to y. To be more exact, (I$^+$) signifies influence by direct (i.e., unscattered) light, (II$^+$) any conceivable influence and (III$^+$) mechanical influence (transfer

[2]If $\dim R = \infty$ then we may take the sphere obtained from the plane E (in the definition of Subsection 1.1) on adding an "at infinity" point; the topology is defined in the same way as for $\dim R < \infty$, since x_E^2 is given.

of energy-momentum from particles of nonzero mass at rest by scattered light).

Consequently, the cones C_x^+, K_x^+ and Q_x^+ are the sets of those events that affect x or may be affected from x by light or some mechanical or arbitrary influence.

The symmetric relations (I) and (II) signify a connection or a possible connection between x and y for energy-momentum transfer, no matter whether from x to y or vice versa: (I) implies connection due to direct light, (II) arbitrary connection and (III) mechanical connection.

Our theorems signify that the conservation of any one of the six relations together with its negation without additional conditions is just sufficient to attribute a definite nature to the possible space-time transformations. (Under obvious requirement that they form a group, there is no need to require the conservation of negation along with the relation, because conservation of negation is guaranteed by the conservation of the relation under inverse transformation.) However, it is natural to restrict ourselves only to indefinitely extendible transformations; then inversions and *sd*-inversions become superfluous in a plane Minkowski space-time R and there only remain Lorentz transformations with homotheties. However, in a conformal space C or in its covering space $\bar{C}$, all locally admissible transformations are infinitely extendible and hence groups of all conformal transformations (we assume $\dim C < \infty$) act in them.

The requirement that transformations form a group yields the same result as the requirement that transformations be extendible. Notice that the last requirement does not imply that universe be considered as a whole in the same way as extendibility of the natural series does not necessarily entrain that the series be studied *ad infinitum*.

Assuming, as is generally accepted in geometry, that the space-time geometry is defined by a transformation group, we conclude from our theorems that each of the six relations (I)–(III$^+$) defines in this sense a space-time geometry. Relation (II) is the most important: it shows that the most general symmetric cause-effect relation for momentum-energy transfer is determined by the space-time geometry.

In addition to the Minkowski space R, we have two more models C and $\bar{C}$ for the space-time in which geometry is determined by one of these general relations. These models were examined by Segal [8] in relation to cosmology.

In the space C relations (I$^+$)–(III$^+$) can be defined only locally, so the conventional directional causality connection has only a local meaning in C. However, in the covering space $\bar{C}$ it has also a global meaning: continuation of a local causality connection does not yield closed causality chains as it does in C. Such chains are believed to be inadmissible and impossible "because they contradict the causality concept." This belief is unfounded as

it is baseless to believe that nature should act in unison with our concepts. The causality principle fits well with the local structure of nature from which it was drawn: this however does not mean it needs no modification elsewhere.

1.7. Supplement. The following theorem holds:

Theorem 5. *Let $f\colon\ G \to R$ be a one-to-one mapping of a domain $G \subset R$ such that 1) $f(G)$ is an open set and 2) for a cone M_x of one of the six types* (I)–(III$^+$)*, for every $x \in G$, there exists a $y \in f(G)$ such that*

$$f(M_x \cap G) = M_y \cap f(G).$$

Then f is the same mapping as in Theorem 1. If an analogous requirement is imposed on a domain $G \subset C$ or on $G \subset \bar{C}$, then f is the same mapping as in Theorems 3 and 4. The same is true of a cone M_x without a vertex, i.e., for the set $M_x \setminus (x)$ or, which is the same thing, for the complementary cone $N_x = [\,R \setminus M_x\,] \cup (x)$.

If we require, as in Condition (A) of Theorem 2, that $y = f(x)$, i.e., that the vertex of the cone M_x be mapped onto the vertex of the cone M_y, then both M_x and $M_x \setminus (x)$ give the same result, but they differ since it is *a priori* admitted that $y \neq f(x)$. On the other hand, a condition is imposed, superfluous although, i.e. $f(G)$ be open.

We do not prove Theorem 5 due to lack of room. For $G = R$ it asserts that f is a Lorentz transformation with a homothety. This result was established in [4] and the cases (I), (I$^+$) and (II$^+$) in [3] which I discussed with Vallander. For this reason I cite Theorem 5 here as generalization of an old result.

§ 2. Inversions and Lorentz Transformations

In what follows we use the definitions and conditions given in Subsections 1.1 and 1.2.

2.1. We prove Assertion 1 formulated in Subsection 1.2. We can take the origin at the center of symmetry or at the point a in an *sd*-inversion and then, if necessary, replace x by $(x - a)$.

2.1.1. Elementary calculation yields

$$(I_0(x) - I_0(y))^2 = \frac{(x-y)^2}{x^2 y^2}, \tag{1}$$

$$(J_{c0}(x) - J_{c0}(y))^2 = \frac{(x-y)^2}{(1+2cx)\,(1+2cy)}. \tag{2}$$

From (1) it follows, first, that an inversion preserves the relation $(x-y)^2 = 0$ and its negation wherever the inversion is defined. Second, on a connected

set where the inversion is defined, x^2 does not vanish and consequently does not change its sign. Therefore on such a set we have $x^2y^2 > 0$ and, hence, the inversion preserves the sign of $(x-y)^2$. From (2) an analogous conclusion follows for *sd*-inversions.

From these remarks and, since homotheties and Lorentz transformations everywhere preserve the sign of $(x-y)^2$, we obtain the first part of Assertion 1.

2.1.2. The second part of Assertion 1 is concerned with the conservation of relations (I^+)–(III^+), i.e., conservation of the sign of $(x-y)^2$ along with the condition $x_0 \leqslant y_0$. This is tantamount to that the halves of each cone $K_x\colon (x-y)^2 \geqslant 0$, where $x_0 \leqslant y_0$ and $x_0 \geqslant y_0$, are transformed into similar halves; otherwise, the cone is "inverted."

The inversion $I_0\colon x \to \frac{x}{x^2}$ induces the transformation $x_0 \to \frac{1}{x_0}$ on the axis x_0, so each of the half-axes $x_0 < 0$ and $x_0 > 0$ is "inverted." Every cone with vertex in the region $x > 0$, $x_0 > 0$ (or $x_0 < 0$) intersects the half-axis $x_0 > 0$ (or $x_0 < 0$) with its two halves, so it is inverted along with the axis.

In the region where $x^2 < 0$, the inversion I_0 reverts the direction of the vector x, hence, it obviously inverts a cone with vertex in this region.

An *sd*-inversion does not invert a cone since, as is shown below, it is, up to a translation, the product of two inversions.

Of the remaining transformations H and L, only a homothety with $\lambda < 0$ inverts a cone. On the strength of these observations we obtain the second part of Assertion 1.

2.2. Let us consider an *sd*-inversion. Take a vector $c \neq 0$ such that $c^2 = 0$. Then

$$I_{-c}I_0(x) = \frac{\frac{x}{x^2} + c}{(\frac{x}{x^2} + c)^2} - c = \frac{x + cx^2}{1 - 2cx} - c = J_{c0}(x) - c.$$

Substituting $(x-a)$ for x, and hence, $(a-c)$ and a for $-c$ and 0, respectively; we obtain

$$I_{a-c}I_a(x) = J_{ca}(x) - c. \tag{3}$$

It is a simple matter to verify that $J_{ca}^{-1} = J_{-ca}$. In the same way it is a simple matter to verify that two *sd*-inversions with a common singular plane differ only in the transformation HL. (It suffices to verify that if $1 + 2c\,(x-a) = \lambda(1 + 2c_1 x)$, then $J_{ca}J_{-c_1 a} = HL$.)

2.3. Assertion 2. *The product of the mappings (1)–(4) in any combination is reduced to one of the three representations: HL, HLI, HLJ (and in the last two cases, to IHL and JHL as well). The representations HL and HLI are unique, and $HLJ = HLJ'$, where J' is an arbitrary sd-inversion with the same singular plane as for J'.*

We do not prove the first part of the assertion in full as it follows readily from the theorem proved in § 3. For this we only require a particular case of the first part of Assertion 2.

2.3.1. Lemma. *The product of the mappings H, L and I that contains not more than two inversions is reducible to one of the three types: HL, HLI or HLJ (or to IHL or JHL as well).*

Proof. It is a simple matter to verify the following formula for arbitrary H, L and I (noting that $L(x-a))^2 = (x-y)^2$):

$$LI_a = I_{L(a)}L, \quad HI_a = I_{H(a)}H'. \tag{4}$$

Furthermore, it is easy to derive

$$I_bI_a = I_{I_b(a)}HL = H'LI_{I_a(b)}, \tag{5}$$

provided $(a-b)^2 \neq 0$, so the points $I_b(a)$ and $I_a(b)$ are defined.[3] However, if $(a-b)^2 = 0$, then from (3) we derive

$$I_bI_a = H''J_{ca}, \tag{6}$$

where $c = a - b$ and H'' is the translation by $-c$. It is understood that $a - b \neq 0$, otherwise I_bI_a would be an identity.

Since every product of H and L is obviously equal to HL, formulas (4)–(6) reduce the product of the mappings H, L and I with not more than two inversions to one of the types: HL, HLI, HLJ or to IHL, JHL.

2.3.2. The second part of Assertion 2 is derived from the following observation.

The representation HL is clearly unique (as the inversion of the inequality $x_0 > 0$ for $x^2 > 0$ is excluded from L and, hence, the homothety with $\lambda = -1$ is also excluded).

The representation HLI_a is unique since this mapping is not defined on the singular cone C_a of the inversion I_a but the inversion with this singular cone is unique. The mapping I_aHL has a singular cone C_b, where $b = L^{-1}H^{-1}(a)$.

A mapping that is representable as HLJ is not defined on the singular plane of the sd-inversion J. Therefore it may be represented as $H'L'J'$ only with the same singular plane. Therefore, as stated in Subsection 2.2, for any sd-inversion J' with the same singular plane, we have $J = H''L''J'$, so $HLI = H'L'J'$.

A qualitative, particularly, topological, distinction between HL, HLI and HLJ lies, first, in that they are defined on different sets: HL on the entire space R, HLI on the three domains truncated by the singular cone of inversion and HLJ on two half-spaces.

[3]If we take $(a-b) = c$, then in (5) we have $L(x) = x - 2\frac{c\,(cx)}{c^2}$, which is a reflection in the plane $cx = 0$. And $H(x) = c^2x + d$, $H'(x) = c^2x + d'$.

2.4. In § 3 we need the following

Lemma. *Under an inversion with center a, the following assertions hold up to a set lying in the singular cone C_a:*

(1) *Every plane passing through a is mapped onto itself and the domains into which the cone C_a subdivides the space are mapped onto themselves.*
(2) *Every cone C_b with $b \in C_a$ is mapped onto a plane and the cones with vertices lying on the same generator of the cone C_a are mapped onto parallel planes.*
(3) *Every hyperboloid containing a and having isotropic generators is mapped onto a plane, a two-dimensional "2-plane."*

Taking the point a as the origin, we obtain the inversion $x \to \frac{x}{x^2}$.

2.4.1. Assertion (1) is self-evident as the inversion $x \to \frac{x}{x^2}$ maps every line passing through O onto itself (minus the origin O).

2.4.2. We prove Assertion (2). The equation of the cone C_b $b \in C_0$, i.e., $b^2 = 0$, is

$$(x - b)^2 = x^2 - 2bx = 0.$$

When inverted, this equation yields the plane $2bx = 1$. If b and b_1 lie on the same generator, then $b_1 = \lambda b$ and the planes $2bx = 1$ and $2b_1x = 1$ are parallel.

2.4.3. We now prove Assertion (3). Suppose that H is a hyperboloid with isotropic generators passing through the origin O. Its asymptotic cone consists of isotropic lines passing although the center c of the hyperboloid H and lying in the three-dimensional space (plane) T containing H. Therefore its equation can be written as $(x-c)^2|_T = 0$, where $(x-c)^2|_T$ is the function $(x - c)^2$ on the plane T. Accordingly, the equation of the hyperboloid is

$$(x - c)^2|_T = p = \text{const}.$$

Consequently, H is the intersection of the surface S: $(x - c)^2 = p$ and the plane T. Since H passes through the origin O, we have $O \in S$. Therefore, the equation of the surface must not contain any free term. Hence, the equation is

$$x^2 - 2cx = 0.$$

When inverted, it gives the plane W:

$$2cx = 1.$$

And the plane T is mapped onto itself as $O \in H \subset T$. Therefore the hyperboloid $H = S \cap T$ is mapped onto of $W \cap T$, that is to say, onto a plane. This completes the proof.

2.5. An affine mapping is also a one-to-one linear mapping, possibly with a translation.

Lemma. *If an affine mapping preserves the equality $(x-y)^2 = 0$, then it is a HL mapping.*

To prove this lemma we first have to establish the following.

2.5.1. *If a quadratic form $q(x)$ is such that $q(x) = 0$ for $x^2 = 0$, then $q(x) = ax^2$, where a is a constant (by definition $q(x) = r(x,x)$, where $r(x,y)$ is symmetric and linear in each argument).*

Proof. Just as in defining x^2, on decomposing x along a line L and a plane E, we obtain $x = x_L + x_E,\ x^2 = x_0^2 - x_E^2,\ q(x) = ax_0^2 + 2x_0 l(x_E) + k(x_E)$. If $q(x) = 0$ for $x^2 = 0$, i.e., for $x_0 = \pm\sqrt{x_E^2}$, then $ax_E^2 \pm 2\sqrt{x_E^2}\, l(x_E) + k(x_E) = 0$. This equality must hold for all $x_E \in E$ because for an arbitrary x_E, x_0^2 may be equal to x_E^2. Therefore for all x_E, we have

$$l(x_E) = 0, \quad ax_E^2 + k(x_E) = 0.$$

Consequently,

$$q(x) = ax_0^2 - ax_E^2 = ax^2.$$

2.5.2. We now prove Lemma 2.5.

Translation does not alter $(x-y)^2$, so the lemma is reduced to the following.

Suppose that a linear mapping g is such that $x^2 = 0$ implies $g(x)^2 = 0$, then $g = HL$. Since $g(x)^2$ is a quadratic form, applying 2.5.1, we find that $g(x)^2 = ax^2$. Here $a > 0$ (because there is a plane on which $x^2 < 0$ but there is no plane on which $x^2 > 0$. And since g is linear the same must be true of $g(x)^2$ as well.) Now putting $g_1 = \frac{g}{\sqrt{a}}$, i.e., adding to g a homothety (contraction by $\sqrt{a}$ times), we obtain a linear mapping such that $g_1(x)^2 = x^2$. It is just the mapping L because, if it alters the sign of x_0 for $x^2 > 0$, then adding the homothety $x \to -x$, we find $-g_1$ is the mapping L. Thus we have $g = \pm\sqrt{aL} = HL$. This completes the proof.

§ 3. Proof of Theorem 1 for Relation (I)

Suppose that the conditions of Theorem 1 are satisfied for relation (I), i.e., in the space R we are given a domain G whose one-to-one mapping $f\colon G \to R$ preserves the relations

$$(x-y)^2 = 0 \quad (\mathrm{I}) \qquad \text{and} \qquad (x-y)^2 \neq 0 \quad (\bar{\mathrm{I}}).$$

We assume that the domain G is convex. (By proving this theorem under this assumption, we indeed establish it for a convex neighborhood of every point in an arbitrary domain. Then the representation thus obtained for f is extendible to the whole of the domain as the domain is connected.)

3.1. We show that f *sends an isotopic line into an isotropic line*, i.e., if L is an isotropic line, then there exists an isotropic line l' such that

$$f(l \cap g) = l' \cap f(G).$$

This obviously follows from the following assertion.

Three points x, y and z lie on an isotropic line if and only if

$$(x - y)^2 = (y - z)^2 = (z - x)^2 = 0. \tag{1}$$

Proof. If x, y, and z lie on an isotropic line, then (1) is obviously satisfied. Suppose that (1) is satisfied. If x, y and z do not lie on a single line, then there would be three independent isotropic vectors xy, yz and zx on the plane spanned on them. However, a 2-plane can contain at most two independent isotropic vectors, two isotropic lines. Consequently, the points x, y and z lie on one line.

3.2. Suppose that P is a 2-plane containing two intersecting isotropic lines; in other words, the plane intersects some cone C_a with $a \in P$ along two generators. We assert that the *set $f(P \cap G)$ is included either in a 2-plane or in a hyperboloid.*

Proof. For every $x \in P$, the intersection $P \cap C_x$ consists of two generators of the cone C_x. Hence P is covered by two families of parallel generators of C_x, i.e., isotropic lines.

As has already been proved, f maps (within G) an isotropic line into an isotropic line. Therefore the set $P' = f(P \cap G)$ is included in a ruled surface with two families of linear generators and this, as is known, may be only a surface of one of the three types: 1) a plane, 2) a hyperboloid and 3) a hyperbolic paraboloid.

All the generators of a paraboloid that belong to a family are parallel to a 2-plane. In our case they are the generators of the cones C_x, while on P' this is only a part of the generators included in $f(G)$. However, a cone has only two generators parallel to a 2-plane and these generators of different cones are parallel in pairs. This means there are only two directions for the generators parallel to the same 2-plane. Therefore P' cannot be a part of a paraboloid and, hence is included either in a plane or in a hyperboloid. This completes the proof.

3.3. Now suppose that a is a point of the set G. Consider a mapping g obtained from the given f by adding an inversion with centers a and $f(a)$:

$$g = I_{f(a)}\, f\, I_a.$$

We prove that *for every 2-plane P passing through a and intersecting the cone C_a along two generators, the mapping g is affine on the domain*

$$V = I_a(U), \quad U = P \cap G \cap Q_a^+,$$

where Q_a^+ is one nape of the interior of the cone C_a. Of course, the same is also true of the other nape of the interior.

3.3.1. Take a point a and a plane P as described above. Now apply the inversion I_a. Since $a \in P$, the plane P is mapped into itself. And since C_a does not intersect U, the inversion I_a is defined everywhere on U. Hence we obtain the domain

$$V = I_a(U) \subset P.$$

According to Lemma 2.4 (item (2)), the cones C_b with $b \in C_a$ are mapped under the inversion I_a to a plane $I(C_b)$. Hence the intersections $C_b \cap U$ are mapped to straight lines. Call them the lines l. When the point b moves over a generator of the cone C_a, the intersection $C_b \cap U$ sweeps U and, by Lemma 2.3, the plane $I(C_b)$ translates. So the line l translates and sweeps the domain V.

Thus the domain V is covered by a continuous ensemble of families of parallel lines l: lines of one family are obtained when the point b moves over a generator of the cone C_a. This family may also be generated when the point b moves over another generator. However, there is a continuous ensemble of different families, i.e., interval of directions of the lines l.

Indeed, taking the point a as the origin, we find from 2.4.2 that the equation of the plane $I(C_b)$ is $2bx = 1$. The vectors b fill the cone C_a. The lines l are given by the intersection $P \cap I(C_b)$. Therefore their directions vary in a certain interval.

3.3.2. We now discuss this conclusion in greater detail. Applying a Lorentz transformation, move the axis x_0 onto the plane P and then take an axis x_1 on it so that $x^2 = x_0^2 - x_1^2$ for vector $x \in P$ and $x^2 = x_0^2 - x_E^2 \leqslant x_0^2 - x_1^2$ for an arbitrary vector x. The equation of the line l is $2(b_0x_0 - b_1x_1) = 1$. Since $b \in C_a$, we have $b^2 = 0$, so $b_0^2 \geqslant b_1^2$. Hence, the directions of the lines l are not arbitrary but have an angular coefficient with modulus $\left|\frac{b_0}{b_1}\right| \geqslant 1$. Second, through every point (x_0, x_1) of the domain $P \cap Q_a^+$ there passes a line l with a preassigned slope $\frac{b_0}{b_1}$ such that $\left|\frac{b_0}{b_1}\right| \geqslant 1$. Indeed, $x^2 > 0$ in the domain Q_a^+, so $x_0^2 > x_1^2$ and $x_0 > 0$. Take a vector $d \in C_a^+$ so that $d_0 > 0$ and $|d_1| \leqslant d_0$. Then for every point $(c_0, c_1) \in P \cap Q_a^+$, we have $(d_0c_0 - d_1c_1) > 0$. Therefore, putting

$$b_i = \frac{d_i}{2(d_0c_0 - d_ic_i)}, \quad i = 1, 2,$$

we find that the line $l = P \cap I(C_b)$, i.e.,

$$2(b_0x_0 - b_ix_i) = 1$$

passes through the point (c_0, c_1).

3.4. Now to continue the proof of (3.3), we apply the mapping f and the inversion $I_{f(a)}$.

3.4.1. According to 3.2, the set $f(P \cap G)$ is included either in a plane or in a hyperboloid. The point $f(a)$ belongs to $f(P \cap G)$. Therefore *If the set $f(P \cap G)$ is included in a plane, then the plane passes through the center of the inversion $I_{f(a)}$ and, consequently, is transformed into itself. So, $I_{f(a)}f(P \cap G)$ is included in the plane.*

Assume that $f(P \cap G)$ is included in a hyperboloid. Its generators are isotropic lines. Therefore, by Lemma 2.4, this hyperboloid is transformed into a plane under an inversion with center $f(a)$ lying on the hyperboloid. Hence the set $I_{f(a)}f(P \cap G)$ is included in a plane.

Thus, the domain $U = P \cap G \cap Q_a^+$ is mapped to a plane. Put $I_{f(a)}f(U) = V'$. Since $U \cap C_a$ is not empty and $f(C_a \cap G) = C_{f(a)} \cap f(G)$, the intersection $f(U) \cap C_{f(a)}$ is also not empty. Therefore the inversion $I_{f(a)}$ is defined everywhere on $f(U)$.

3.4.2. The inversion I_a introduced earlier maps U onto a domain $V \subset P$. Since an inversion itself is reversible, we can schematically represent our mappings as follows:

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, leftrightarrow, "I_a"'] & U' \arrow[d, leftrightarrow, "I_{f(a)}"] \\ V \arrow[r] & V' \end{tikzcd}$$

$$g = I_{f(a)}\, f\, I_a$$

The cones C_b with vertices on the cone C_a are transformed under the mapping f into cones $C_{f(a)}$ with vertices on $C_{f(a)}$.[4] However, under the inversion $I_{f(a)}$, these cones pass into planes just like the cones C_b under the inversion I_a.

Therefore, for the domain $V' = I_{f(a)}f(U)$ we obtain the same picture as for the domain $V = I_a(V)$ described in Subsection 3.3.1. The domain V' is

[4] $f(C_b \cap G) = C_{f(b)} \cap f(G)$, and, if $b \in C_a$, then $f(b) \in C_{f(a)}$: we used Condition (A) of Subsection 1.4.

covered by families of parallel lines l'. The subsets of these lines that are included in V' are the images of the corresponding subsets of the lines l: $l' \cap V' = g\,(l \cap U) = I_{f(a)} f\, I_a(l \cap V)$.

Under the mapping f a generator of the cone C_x is transformed to a generator. And cones with vertices lying on the same generator of C_a (or $C_{f(a)}$) are generated by parallel lines l (or l'). Therefore parallel lines l' correspond to parallel lines l.

3.4.3. Thus we arrive at the

Assertion. *On the plane P there exist a domain V and a continuous ensemble of families F of parallel lines l, each family being a covering of V; there also exists a mapping g that sends V into a plane so that the sets $l \cap V$ are mapped into lines in such a way that the sets including parallel lines are sent to parallel lines.*

We prove assertion 3.3, namely, that the mapping g is affine. Since $g(V)$ is definitely not included in a line (at least for the reason that different generators of a cone are transformed to different generators) it follows that the mapping g is affine as is implied in the following lemma which is in no way associated with cones.

3.5. Lemma. *If the above assertion holds for a plane domain V, for the families F and for a mapping g, and $g(V)$ is not included in a line, then the mapping g is affine.*

(There is no need here to assume that g is a one-to-one mapping; if however V is mapped uniquely into a line, then all the conditions are satisfied; moreover, the set $g(V)$ is not included in a line, so this condition is necessary).

3.5.1. *Proof.* First, note that the "lines," i.e., the lines included in V, from different families F cannot be mapped into parallel lines. Suppose that the images of "lines" from two different families F_1 and F_2 are parallel.

It is possible to take l_1, $l'_1 \in F_1$ and $l_2 \in F_2$ such that l_2 intersects l_1 and l'_1. Since the images of all these three "lines" are parallel, they lie on a line. Hence we deduce that all the "lines" from F_1 and F_2, and along with them, the whole of the domain V, are mapped to one line. However, this contradicts the conditions of our theorem.

3.5.2. In V let us introduce affine coordinates u and v such that the lines $u = \text{const}$ and $v = \text{const}$ are the members of two different families F and that the origin (0,0) lies in V. The lines of the families F, except those that are parallel to the axis v, are given by

$$v = cu + d. \tag{3}$$

Since the families F form a continuous set, the coefficient c can take any value on the half-axis $[0, \infty)$ or on $(-\infty, 0]$. Of course, we may take $c \geqslant 0$. Furthermore, since each family from the set F covers the domain V, the free term d can take any value in some interval (d_1, d_2). The choice can be made such that for an arbitrary $d \in (d_1, d_2)$, the coordinates u and v take a value in a certain interval near zero.

On the plane P' into which the domain V is mapped, let us choose the affine coordinates u' and v' such that the origin is the image of the origin of the u, v-coordinate frame and the lines $u' = \text{const}$ and $v' = \text{const}$ include the images of the lines $u = \text{const}$ and $u = \text{const}$, respectively. This can be done because these lines belong to two different families of F and are therefore mapped to parallel lines.

In the new coordinate system, the mapping g: $V \to P'$ is given by the formulas

$$u' = p(u), \quad v' = q(v), \tag{4}$$

and g is affine means the functions p and q are linear.

3.5.3. The lines $u = \text{const}$ are mapped to the lines $u' = \text{const}$ in the u', v'-coordinates. And by what has been proved above, no line from other families can be mapped to the lines $u' = \text{const}$. Thus, each “line” (3) is mapped to the line

$$v' = c'u' + d'. \tag{5}$$

Since lines from one family, i.e., with the same c, are mapped to parallel lines, we have

$$c' = r(c). \tag{6}$$

Using (6), (4) and (3), we obtain from (5)

$$q(cu + d) = r(c)\, p(u) + d'.$$

Since $p(0) = 0$, putting $u = 0$, we find $d' = q(d)$. Hence

$$q(cu + d) = r(c)\, p(u) + q(d). \tag{7}$$

Furthermore, since $q(0) = 0$, putting $d = 0$, we obtain

$$r(c)p(u) = q(cu). \tag{8}$$

Therefore, (7) yields

$$q(z + d) = q(z) + q(d). \tag{9}$$

That q is linear is still not evident as no assumption was made beforehand regarding the nature of the function.

3.5.4. We take a number $a_0 \neq 0$ such that $r(a_0)$ and $p(a_0)$ are defined. Furthermore, we restrict ourselves to the values of u in an interval near zero so that $r(u)$, $p(u)$ and $q(a_0 u)$ are defined. In such a case, from (8), we derive

$$r(a_0)p(u) = q(a_0 u), \quad r(u)p(a_0) = q(a_0 u), \tag{10}$$

hence

$$r(u) = kp(u), \quad k = \frac{r(a_0)}{p(a_0)}.$$

Therefore, in (7) putting $c = u = w$, we find

$$q(w^2 + d) = kp(w)^2 + q(d).$$

Hence, for $z > 0$,

$$q(d + z) = q(d) + kp(\sqrt{z}\,)^2.$$

So q is monotone. Under this condition, equality (9) shows that q is linear: $q(v) = q_0 v$. Therefore, from (10) it follows that $p(u)$ is linear.

3.5.5. Thus we showed that the mapping g is affine at least in the neighborhood of (0,0) for which our last conclusion is true. However, as the origin can be taken arbitrarily at any point in the domain V, q is affine over the entire domain V. This completes the proof.

3.6. Thus we proved Lemma 3.5 and along with it assertion 3.5 as well. From this lemma we find that *the mapping* $g = I_{f(a)}\, f\, I_a$ *is affine on the domain* $W = I_a(G \cap Q_a^+)$.

By assertion 3.3, the mapping g is affine on every plane domain $V = W \cap P$, where P is a 2-plane passing through the point a and intersecting the cone C_a along two generators.

Let l be a ray emanating from the point a and passing through some point in W, and l_x a ray parallel to l and passing through a point $x \in W$. The plane P on which g is affine passes through these rays. So the rays $g(l)$ and $g(l_x)$ are mapped to parallel rays and g is affine on them. Therefore any two rays l_x and l_y parallel to l are also sent into parallel rays and g is affine on them.

Now take an arbitrary finite-dimensional plane T intersecting the domain W and let S be a plane passing through T and a point a. On S let us introduce affine coordinates with origin a such that the axes are directed into the interior of Q_a^+. Later we move the origin to some other point $b \in W \cap S$.

From what has been demonstrated for rays, it follows that a coordinate grid is transformed into an affine system under g. So the mapping g is affine on $W \cap S$. And since S includes an arbitrary finite-dimensional subspace T, the mapping g is affine on W.

3.7. We now prove the assertion of Theorem 1, i.e., f is either HL or HLI or HLJ.

First, we notice that the mapping $g = I_{f(a)} f I(a)$ transforms the cones C_x (within the limits of the domain W where g is defined) into cones of the same type as this holds for the components of its cofactors. So, for every $x \in W$, we have

$$g(C_x \cap W) = C_{g(x)} \cap g(W).$$

And, since g has been demonstrated to be affine, by Lemma 2.5, we have $g = HL$.

Thus, for the mapping f on the domain $G \cap Q_a^+$, we have

$$f = I_{f(a)} g I_a = I_{f(a)} HLI_a.$$

Therefore, by Subsection 2.3, we can write

$$f = I_{f(a)} U_{g(a)} H' L'.$$

Hence, if $f(a) = g(a)$, then $f = H'L'$; otherwise, by Subsection 2.3, $f = IH''L''$ or $JH''L''$.

We proved this for the mapping f on the domain $G \cap Q_a^+$. However, as the point $a \in G$ is arbitrary, the same holds true of every domain $G \cap Q_x^+$, $x \in G$.

Hence, G being convex, we conclude that the mapping f can be represented as HL or IHL or JHL everywhere on the domain G.

Thus we proved Theorem 1 for relation (I), assuming that G is a convex domain. However, if the domain is arbitrary, applying the assertion for a convex neighborhood of a point in the domain, and since the domain is connected, we find that Theorem 1 holds for an arbitrary domain as well.

Other assertions of Theorem 1, namely, that the representation in the form of HL is unique, etc. are contained in Assertion 2 of § 2.

§ 4. Proof of Theorem 1 for Relation (II$^+$)

Instead of demonstrating Theorem 1 for relation (II), we prove Theorem 2, an equivalent of Theorem 1 for the cones K_x. First, we prove the theorem for a single cone K_x^+ and then in § 5 we extend it to the double cone K_x. Incidentally, in view of the interpretation given in Subsection 1.5, single cones K_x are of independent interest. As in § 3, here too we suppose that the domain G is convex.

4.1. Let $f\colon G \to R$ be a one-to-one mapping of a convex $G \subset R$ such that for any $x \in G$:

$$f(K_x^+ \cap G) = K_{f(x)}^+ \cap f(G). \tag{1}$$

Together with single cones K_x^+, we consider the cones C_x^+ and Q_x^+ that are the boundary and the interior of K_x^+ (the vertex x does not belong to Q_x^+), respectively.

We prove that for every $x \in G$

$$f(C_x^+ \cap G) = C_{f(x)}^+ \cap f(G), \tag{2}$$

i.e., the condition of Theorem 2 holds for the cone C_x^+. In this case the theorem is proven as it reduced, as shown in Subsection 1.5, to the case of cones C_x, an assertion equivalent to Theorem 1 for relation (I) demonstrated in § 3.

4.2. In the space R we now define a *partial order*, putting $y \geqslant x$ if and only if $y \in K_x^+$. The cone symmetric with K_x^+ relative to the point x (the "second nape" of the double cone K_x) is the cone $K_x^- = \{y\colon y \leqslant x\}$.

In terms of this order, condition (1) signifies that the mapping f and its inverse f^{-1} preserve the ordering; for every x, we have $y \in G$, $x \leqslant y \iff f(x) \leqslant f(y)$.

In what follows we write K, C, Q in place of K^+, C^+ and Q^+, respectively.

4.2.1. In this order the interval $[ab]$ is, by definition, the set:

$$[ab] = \{x\colon a \leqslant x \leqslant b\} = K_a \cap K_b^-.$$

It is immediately clear that if $b \in C_a$, the interval $[ab]$ is an segment on the generator of the cone C_a and it is linearly ordered. However, if $b \in Q_a$, the interval $[ab]$ is not linearly ordered, i.e., there are points $u, v \in [ab]$ such that neither $u \leqslant v$ nor $v \leqslant u$.

From this observation on intervals and since the order is preserved, under f, we derive:

4.2.2. *The interior Q_x of a cone K_x is mapped to an interior*

$$f(Q_x \cap G) \subset Q_{f(x)}. \tag{3a}$$

Indeed, for an $y \in Q$, if $f(y)$ were an element of $C_{f(x)}$, then the interval $[f(x)f(y)]$ would be linearly ordered, although it is the image of a not linearly ordered interval $[xy]$.

4.2.3. Now it suffices to demonstrate that

$$f(C_x \cap G) \subset C_{f(x)}. \tag{3b}$$

Then from (1) and (3a) and since $K_x = Q_x \cup C_x$, we arrive at (2).

4.3. We need the following lemma. It presupposes a topology in R in which the interiors of intervals, i.e., the sets $Q_x \cap Q_y^-$, form a neighborhood base. On a set $M \subset R$ we assume the topology that is induced by the above topology.

Lemma. *An isotonic mapping of a linearly ordered set is continuous everywhere on the set, except at most at a countable set of points.*

4.3.1. *Proof.* Suppose that M is a linearly ordered set and g a monotone mapping of it. Let p denote the projection to the axis x_0 along the plane E stated in the definition of x^2 (see Subsection 1.1). Since the set M is linearly ordered, the projection p is a one-to-one mapping of M into x_0. Hence, its inverse

$$x = p^{-1}(x_0), \quad x \in p(M),$$

is defined. Accordingly, we define on $p(M)$ a function

$$h = pgp^{-1}.$$

Since g is monotone, the function h is monotone, and, hence, is everywhere continuous on $p(M)$, except at most on a countable set of points.

Further, we show that the mapping g is continuous at every point $x \in M$ for which h is continuous at the corresponding point $p(x)$. This completes the proof.

4.3.2. Suppose that $p(a)$ is a point of continuity of the function h, $a \in M$. Take a interval U containing the point $b = g(a)$ in the interior. Through b let us draw a line parallel to the axis x_0 and take the points c, $d \in U$, $c < d$ on this line such that b subdivides the segment cd into two halves. Then the interval $[cd] = K_c \cap K_d^-$ consists of two cones with vertices c, d and a common base on the plane E_b. (By E_x we denote a plane parallel to E and passing through the point x.) This common base of the cones is obviously a ball $S = K_c \cap E_b$ and the point b is its center.

Suppose that the point b is such that $b \geqslant y$ and the plane E_y is at least twice closer to E_b than the distance between the planes E_c and E_b. Then we find $y \in K_c$.

Since $b \geqslant y$ and, consequently, $b \in K_y$, the ball $S_y = K_y \in E$ contains the point b. And since the plane E_y is at least twice closer to E_b than the distance between E_c and E_b, the diameter of the ball S_y is at least twice less than the diameter of the ball $S = K_c \cap E_b$. Therefore $S \supset S_y$ and, consequently, $K_c \supset K_y$; hence, $y \in K_c$.

Moreover, $y \in K_d^-$ as $y \leqslant b$ and $b \leqslant d$. Therefore

$$y \in K_c \cap K_d^- = [cd].$$

In the same way we find that if $y \geqslant b$ and the plane E_y is at least twice closer to the plane E_b than the plane E_d, then $y \in [cd]$.

Since the set M is linearly ordered, and the mapping g is isotonic, then for every $x \in M$, we have either $y = f(x) \leqslant b = f(a)$, or $y \geqslant b$.

Therefore from the previous reasoning we find that if the point $y \in f(m)$ is such that the plane E_y is at least twice closer to E_b than the planes E_c and E_d, then $y \in [cd]$.

4.3.3. If $p(a)$ is a point of continuity of h and $p(x_i) \to p(a)$, then $hp(x_i) \to hp(a)$, i.e., $pg(x_i) \to pg(a)$. Or, since $g(a) = b$, and assuming $g(x_i) = y_i$, we obtain $p(y_i) \to p(b)$.

However, if $p(y_i) \to p(b)$, then for sufficiently large i the planes E_{y_i} are spaced from the plane E_b at least at a distance equal to half the distance between the planes E_c and E_d. Therefore for such i, we have

$$y_i \in [cd] \subset U.$$

So the mapping g is continuous at the point a. And this completes the proof of our lemma.

4.4. We now prove (3b), i.e., for every $x \in G$, we have $f(C_x \cap G) \subset C_{f(x)}$.

Assume the contrary, so there exist points $a \in G$ and $b \in C_a \cap G$ such that $f(b) \notin C_{f(a)}$. Since $f(K_a) \subset K_{f(a)}$, the point $f(a)$ lies in the interior of $K_{f(a)}$:

$$f(b) \in Q_{f(a)}. \tag{4}$$

The segment of the line ab that is included in G is a linearly ordered set. Its map g is a restriction of the mapping f to this segment, is monotone. Therefore, by virtue of Lemma 4.3, this segment contains a point $c \leqslant a$ where g is continuous.

The point c lies on the continuation of the generator ab of the cone C_a. Therefore $b \in C_c$. Since $c \leqslant a$, we have $f(c) \leqslant f(a)$ and hence from (4) we find $f(b) \in Q_{f(c)}$. Thus,

$$b \in C_c, \quad f(b) \in Q_{f(c)}. \tag{5}$$

Since $b \in C_c$, the interval

$$[cd] = K_c \cap K_b^- = cb$$

is cd.

Since the order is preserved and (1) holds, it follows that, for every $x \in K_c \cap G$, i.e., for an $x \geqslant c$, $x \in G$, we have

$$f(|cx| \cap G) = |f(c)f(x)| \cap f(G). \tag{6}$$

Applying the last equality to the interval $[cd]$, we find that

$$f(cb) = |f(c)f(b)| \cap f(G). \tag{7}$$

Therefore the interval $[f(c)f(b)]$ does not contain the images of any point either from Q_c or from C_c, except the interval $[cb]$. Furthermore, as the mapping g of the line ab is continuous at the point c, every neighborhood of the point $f(c)$ contains points of the set $f(cb)$ in addition to the point $f(c)$ itself. These points, in view of (7), are included in the interval $[f(c)f(b)]$.

We now reduce this to a statement contradicting the assertion concerning the interval $[f(c)f(b)]$ and thereby prove (3b).

4.5. Let l denote a generator cb of the cone C_c minus the point c. Since every neighborhood of $f(c)$ contains points from $f(l)$ and $f(l) \subset K_{f(c)}$, we find for every $y \in Q_{f(c)}$

$$f(l) \cap K_y^- \neq \varnothing. \tag{8}$$

4.5.1. We prove that

$$f(C_c \diagdown l) \subset C_{f(c)}. \tag{9}$$

Assume the contrary. Then, as $f(C_c) \subset K_{f(c)}$, we would have a point $x \in C_c \diagdown l$ such that $f(x) \in Q_{f(c)}$. And by (8), we have

$$f(l) \cap K_{f(x)}^- \neq \varnothing.$$

Since order is preserved, this would signify that the line l contains points from K_x^-. Since $l \subset K_c$, these points would be the points of the interval $[cx]$. However, as $x \in C_c$, this interval is the segment cx which has no common points with l. This contradiction proves (8).

4.5.2. We now show that every generator of the cone C_x minus l is mapped to a generator of the cone $C_{f(x)}$ and different generators are mapped into different generators. Indeed, for an arbitrary $x \in C_c \diagdown l$, the interval $[cx]$ is $cx \subset C_c$. However, by 4.5.1, we find $f(x) \in C_{f(c)}$, so the interval $[f(c)f(x)]$ is $f(c)f(x)$. And it is the image of the segment cx (as is implied in (6)). Thus, segments of the generators of the cone $C_c \diagdown l$ are mapped to segments on the generators of the cones $C_{f(c)}$. Furthermore, if the points x and y lie on different generators of C_c, obviously neither $x \leqslant y$ nor $y \leqslant x$. Therefore, neither $f(x) \leqslant f(y)$, nor $f(y) \leqslant f(x)$. Consequently, the points $f(x)$ and $f(y)$ lie on different generators of $C_{f(c)}$.

Hence, together with the previous conclusion regarding the segments on the generators, we obtain our assertion 4.5.2.

4.6. Let l_1 and l_2 be two distinct generators of the cone $C_c \diagdown l$ minus the point c. Moreover, 4.5.2 implies their images $f(l_1)$ and $f(l_2)$ lie on different generators l', l'' of the cone $C_{f(c)}$. Furthermore, $f(l_1)$ and $f(l_2)$ do not contain the points of the interval $[f(c)f(b)]$, because, as it follows from (7), $cb = 1$ and $l_1, l_2 \subset C_c \diagdown l$. Therefore, l' and l'' contain points a' and a''

different from $f(c)$ such that the segments $a'f(c)$ and $a''f(c)$ do not contain points from $f(l_1)$ and $f(l_2)$.

Let d be the midpoint of the segment $a'a''$. Since a' and a'' lie on different generators of $C_{f(c)}$, the point d lies in the interior of $K_f(c)$: $d \in Q_{f(c)}$. Therefore $f(c) \in Q_d^-$.

Every neighborhood of the point $f(c)$ contains a point of $f(l)$. Therefore there exists a point $e \in l$ such that

$$f(e) \in Q_d^-.$$

Moreover, we can take a point $p \in Q_c$ such that $e \notin K_p^-$. Hence

$$f(e) \notin K_{f(p)}^-.$$

However, the cone K_p^- intersects the generators l_1 and l_2. Therefore the cone $K_{f(p)}^-$ intersects the sets $f(l_1)$ and $f(l_2)$, so certainly contains the points a' and a'', and along with them, the point d as well.

Consequently, we find that

$$Q_d^- \subset K_{f(p)}^-.$$

Hence, from (10), it follows that

$$f(e) \in K_{f(p)}^-.$$

However, this contradicts (11). This proves inclusion (3b).

4.7. Thus we proved that

$$f(Q_x) \subset Q_{f(x)}, \quad f(C_x) \subset C_{f(x)}.$$

Hence, and from the main relation (1) governing the cones $K_x = Q_x \cup C_x$, it follows that relation (2) holds for the cones C_x. Theorem 1 has already been demonstrated for these cones, thereby it has been proved for the cones K_x as well.

We assume that the domain G is convex. However, if G is arbitrary, it suffices to apply the result thus obtained to a convex neighborhood of its points and then make use of the connectedness of the domain G.

§ 5. Proof of Theorems 1 and 2 for Relation (II)

We now prove Theorem 2 for a double cone K_x by reducing it to the single cone K_x. The domain G is assumed convex. Reduction is carried out by

5.1. Assertion. *If the points x and y lie on different halves of the cone K_a, say, $x \in K_a^+$, $y \in K_a^-$, then*

$$K_a \subset K_x \cup K_y.$$

However, if the points x and y lie on the same nape and are different from the point a, then

$$K_a \subset K_x \cup K_y.$$

If, in addition, x and y belong to G, then there exist points $z_i \in K_a \cap G$, $i = 1 \dots m$, such that $x = z_1$, $y = z_m$ and for every i

$$K_a \cap G \subset K_{z_i} \cap K_{z_{i+1}} \cap G.$$

Note that, as can be easily verified by simple examples, that

(1) if G is not convex, then the assertion is in general false, and
(2) for a convex domain G, possibly $K_a \cap G \subset K_x \cap K_y \cap G$, if x and y lie on the same nape and $x, y \neq a$.

5.1.1. We prove the first statement of our assertion. Let $x \in K_a^+$, $y \in K_a^-$. Then, $a \in K_x^-$ and $a \in K_y^+$. Hence

$$K_a^- \subset K_x^-, \quad K_a^+ \subset K_y^+.$$

Consequently,

$$K_a = K_a^+ \cup K_a^- \subset K_y^+ \cup K_x^- \subset K_y \cup K_x.$$

5.1.2. We prove the second statement of our assertion. Let $x, y \in K_a^+ \diagdown (a)$. Through the points a, x and y, let us draw a 2-plane P. The intersections $V_a = K_a \cap P$, $V_x = K_x \cap P$ and $V_y = K_y \cap P$ are translated double angles (pairs of mutually opposite vertical angles) with vertices a, x and y, respectively. Here $x, y \in V_a^+ = K_a^+ \cap P$. Simple consideration shows that $V_a \subset V_x \cup V_y$. Therefore, it readily follows that $K_a \subset K_x \cup K_y$.

5.1.3. We prove the last statement of our assertion. Again let $x, y \in K_a^+ \diagdown (a)$ and V_a, V_x, V_y and V_a^+ have the same meaning as before, so $x, y \in V_a^+$. However, here we restrict ourselves to a convex domain G and $x, y \in G$. Therefore, in place of the angles V, we consider their intersections with G:

$$W_a = V_a \cap G, \quad \text{etc.}$$

Let us draw the segment xy. It is included in W_a^+ since $W_a^+ = V_a^+ \cap G$ and the domain G and the angle V_a^+ are convex.

Now we may note that if $z \in W_a^+ \diagdown (a)$, then, as can be easily seen, $W_z \supset W_a$. Hence every given z has a neighborhood U such that for every $v \in U \cap W_a^+$, so we have

$$W_a \subset W_z \cup W_v.$$

Applying this observation to the points z of the segment xy, we find, by virtue of the Borel lemma, the third statement of our assertion is true.

5.2. Now suppose that $f\colon G \to R$ is a mapping that satisfies the conditions of Theorem 1 for relation (II), so for every $x \in G$

$$f(K_x \cap G) = K_{f(x)} \cap f(G)$$

Then, as is easily derived from assertion 5.1, the nappes of the cone K_x are mapped into nappes, so for every x, we find that $f(K_x^+ \cap G)$ is included either in $K_{f(x)}^+$ or in $K_{f(x)}^-$. However, we have to show this is either the same K^+ or the same K^- for all x. We may take that for a given point a

$$f(K_a^+ \cap G) = K_{f(x)} \cap f(G), \tag{1}$$

assuming, if necessary, that f is symmetric at the point a. We now prove this also holds for every $b \in G$.

5.2.1. Let b be an element of K_a and assume that

$$f(K_b^+ \cup G) = K_{f(b)}^- \cap f(G). \tag{2}$$

Let $b \in K_a^+$, so that $f(b) \in K_{f(a)}^+$. Then

$$K_b^+ \subset K_a^+, \quad K_{f(b)}^- \supset K_{f(a)}^+.$$

Together with (2), this gives

$$f(K_a^+ \cap G) \supset f(K_b^+ \cap G) = K_{f(b)}^- \cap f(G) \supset K_{f(a)}^- \cap f(G).$$

Hence, applying (1), we obtain

$$K_{f(a)}^+ \cap f(G) \supset K_{f(a)}^- \cap f(G),$$

which is impossible.

Let $b \in K_a^-$, so that $f(b) \in K_{f(a)}^-$. Then

$$K_b^+ \supset K_a^+, \quad K_{f(b)}^- \subset K_{f(a)}^-.$$

Therefore, using (2), we obtain

$$K_{f(a)}^- \cup f(G) \supset K_{f(b)}^- \cap f(G) = f(K_b^+ \cap G) \supset f(K_a^+ \cap G).$$

Hence with the help of (1), we find

$$K_{f(a)}^- \cap f(G) \supset K_{f(a)}^+ \cap f(G),$$

which is impossible.

Thus, if $b \in K_a$, then (2) is impossible and therefore (1) holds both for K_b and for K_a.

5.2.2. Now suppose that $b \notin K_a$ so that the segment ab has no common points with K_a, except the point a. It has no common points with K_x, $x \in ab$, except the point x. If the points $x, y \in ab$ are sufficiently close to each other, then, as can be easily seen, the like halves of the cones K_x and K_y intersect in G, i.e.,

$$K_x^+ \cap K_y^+ \cap G \neq \varnothing, \quad K_x^- \cap K_y^- \cap G \neq \varnothing. \tag{3}$$

Consequently, their images intersect. However, since $y \notin K_x$, then $f(y)$ is also not contained in $K_{f(x)}$. In this case unlike halves of the cones $K_{f(x)}$ and $K_{f(y)}$ obviously do not intersect.

Consequently, the images of intersecting halves of the cones K_x and K_y must be like halves. And, by virtue of (3), this means the like halves are mapped to like halves. This is true of every pair of points x and y sufficiently close to each other on the segment ab. Therefore for all $x \in ab$, particularly for b, the relation that holds for the point a also holds for the point b.

5.3. Thus, we proved that like halves of the cones K_x are mapped into like halves. Therefore, together with (1), this condition is satisfied for single cones K_x^+. Theorem 2 was established for these cones in § 4, so it is proven also for the double cones K_x.

§ 6. Proof of Theorems III and IV for Relations (I) and (II)

6.1. Suppose that G is a domain in a conformal space C and f is a mapping $f : G \to C$ satisfying the conditions of Theorem 3 for the cones C_x in C so that for $x \in G$

$$f(C_x \cap G) = C_{f(x)} \cap f(G). \tag{1}$$

6.1.1. Take a point $a \in G$. Excluding the cone $C_{f(a)}$ from the space C, we obtain a space R. Consider the set

$$G' = G \diagdown (C_a \cup C_{f(a)}).$$

Obviously, G' is open and is included in R. Moreover, if f' is a restriction of the mapping f to G', then $f' : G' \to R$.

Therefore, if G_1 is some connected component of G', then Theorem 2 is applicable to the mapping $f_1 = f' \diagup G$, since f_1 is either HL or HLI or HLJ and is extended into C in the form of $H^C L^C$ or $H^C L^G$ with the addition of one or two inversions.

6.1.2. Suppose that g is the mapping f_1 extended to G. Consider the mapping $h = g^{-1} f$ of the domain G. Relation (1) is satisfied for this mapping and it is an identity on G_1. We show that h is an identity on G.

Take a point $x \in G_1$ on the boundary of G_1 and a point $b \in G_1$ such that $x \notin C_b$. Excluding the cone C_b from the space C, we obtain a space R and an open set $G'' = G \diagdown G_b \subset R$.

By virtue of its properties, the mapping h is defined on G''. So it is also defined, particularly on the component G_2 of G'' containing the point x and, hence, intersects G_1. On the intersection $G_1 \cap G_2$ the mapping h is an identity. Applying Theorem 2 to the domain $G_2 \subset R$, we conclude that the mapping h is an identity on G_2. Hence it is an identity in the neighborhood of the point x. However, $x \in G$ is an arbitrary point on the boundary of the domain G_1. Therefore h is an identity on a certain domain $G_3 \subset G$ such that $G_1 \cap G \subset G_3$.

Since this conclusion can be applied to an arbitrary domain included in G on which h is an identity, we find h is an identity on G. This means $f = g$ on G. And this completes the proof of Theorem 3 for the cones C_x.

6.2. Now suppose that the mapping $f\colon G \to C$ satisfies the conditions of Theorem 3 for the cones K_x so that for every $x \in G$

$$f(K_x \cap G) = K_{f(x)} \cap f(G). \tag{2}$$

Take a point $a \in G$. Excluding the cone $C_{f(a)}$ from the space C, we obtain a space R. Consider the set

$$G' = G \diagdown (K_a \cup C_{f(a)}).$$

Obviously, G' is open and included in R, but is not empty, because $a \in G$ is the vertex of the cone K_a. Furthermore, if f' is a restriction of the mapping f to G', then $f'\colon G \to R$ (because $K_a \supset C$, so G' does not contain the points from C_a and hence f' maps G' into $C \diagdown G_{f(a)} = R$). Now taking some connected component G_1 of the set G', we literally arrive at the conclusion of Subsection 6.1, namely, f is either $H^C I^C$ or $H^G L^C$ with one or two inversions, that is what Theorem 3 asserts.

6.3. We do not prove Theorem 4 for the covering space $\bar{C}$ as it obviously follows from Theorem 3.

Received April 2, 1976

References

[1] Alexandrov A. D., *On Lorentz transformations*, Uspekhi Mat. Nauk **5** (1950), no. 3, 1–187; in Russian.

[2] Alexandrov A. D., *Relativity theory: its essence*, Vestnik LGU (1948), no. 8, 108–128; in Russian.

[3] Alexandrov A. D. and Ovchinnikova V. V., *A remark on the relativity principles*, Vestnik LGU (1953), no. 11, 95–110; in Russian.

[4] Alexandrov A. D., *Mappings of a space with families of cones and space-time transformations*, Ann. Mat. Pura Appl. **103** (1975), 229–257.

[5] Zeeman E. C., *Causality implies the Lorentz group*, J. Math. Phys. **5** (1964), no. 4, 490–493.

[6] Borchers H. J. and Hegerfeld G. C., *The structure of space-time transformations*, Comm. Math. Phys. **28** (1972), no. 3, 259–266.

[7] Rozenfel′d B. A., *Non-Euclidean Geometry*, Moscow, 1955.

[8] Segal I., *Covariant chronogeometry and extreme distances*, J. Astron. & Astrophys. **18** (1972), 143–148.

INDEX