HILBERT'S FOURTH PROBLEM

Aleksei V. Pogorelov

HILBERT'S FOURTH PROBLEM

Aleksei Vasil'evich Pogorelov
USSR Academy of Sciences

Translated by
Richard A. Silverman

Edited by
Irwin Kra
SUNY at Stony Brook
in cooperation with
Eugene Zaustinskiy
SUNY at Stony Brook

V. H. WINSTON & SONS
Washington, D.C.

A HALSTED PRESS BOOK

JOHN WILEY & SONS
New York Toronto London Sydney

V. H. Winston & Sons, a Division of Scripta Technica, Inc., Publishers
1511 K Street, N.W., Washington, D.C. 20005

Distributed solely by Halsted Press, a Division of John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data:

Pogorelov, Alekseĭ Vasil'evich, 1919–
Hilbert's fourth problem.

(Scripta series in mathematics)
Translation of Chetvertaia problema Gil'berta.
Bibliography: p.
1. Geometry–Foundations. I. Kra, Irwin. II. Zaustinskiy, Eugene. III. Title. IV. Series.
QA681.P5913 516'.1 79-14508
ISBN 0-470-26735-6 (Halsted)

Composition by Isabelle K. Sneeringer, Scripta Technica, Inc.

CONTENTS

PREFACE TO THE AMERICAN EDITION

Hilbert's Fourth Problem, as presented by Hilbert himself (see the Introduction), is stated in rather broad and general terms. In brief, the problem consists in the investigation of metric spaces which admit a geodesic mapping onto a projective space or a domain of such a space.[1*] Hilbert's Fourth Problem is related to the foundations of geometry, the calculus of variations, and differential geometry.[2] We shall consider Hilbert's Problem as a problem in the foundations of geometry and, in this regard, following Hilbert, we formulate the problem more precisely as follows.

Suppose we take the system of Axioms for Euclidean geometry, drop those axioms involving the concept of angle, and then supplement the resulting system with the "triangle inequality," regarded as an axiom. The resulting system of axioms is incomplete and there exist infinitely many geometries, in addition to Euclidean geometry, which satisfy these axioms. Hilbert's Problem consists in describing all

*Note. Superscripts refer to Notes beginning on page 88.

possible geometries satisfying this system of axioms. The present work is devoted to the solution of the problem as stated in this form. The problem will be considered from the standpoint of all three classical geometries, namely those of Euclid, Lobachevski and Riemann.

It turns out that the solution of the problem, as formulated here, reduces to the determination of all of the so-called Desarguesian metrics[3] in projective space; that is, metrics for which the geodesics are straight lines. Such metrics were obtained by Hamel (see [12]), under the assumption of sufficient smoothness. However, a complete solution of the problem requires the determination of all Desarguesian metrics, without assuming smoothness, subject only to the condition of continuity which is guaranteed by the axioms.

The occasion for the present investigation is a remarkable idea due to Herbert Busemann, which I learned about from his report to the International Congress of Mathematicians at Moscow in 1966. Busemann gave an extremely simple and very general method of constructing Desarguesian metrics by using a nonnegative completely additive set function on the sets of planes and defining the length of a segment as the value of this function on the set of planes intersecting the segment.

I suspected that all continuous Desarguesian metrics could be obtained by this method. The proof of this in the two-dimensional case strengthened my belief in this conjecture and I announced a general theorem in [18]. However, it turned out later, on making a detailed investigation of the three-dimensional case, that the completely additive set function figuring in Busemann's construction may not satisfy the condition of nonnegativity. Therefore, the result given here, while preserving its original form, assumes that other conditions are satisfied.

This book is addressed to a wide circle of readers and, accordingly, it begins with a review of the basic facts of the geometry of projective space (Sections 1, 2). A description of the axiom systems of the classical geometries is given in Section 10. A detailed exposition of the core of the problem, as formulated here, is given in Section 11, together with illustrative examples.

I regard it as my pleasant duty to thank the publisher V. H. Winston & Sons for the interest shown in my work.

A. V. Pogorelov

TRANSLATION EDITOR'S FOREWORD

It has been slightly more than seventy-five years since David Hilbert presented a list of twenty-three outstanding and important problems to the Second International Congress of Mathematicians held in Paris in 1900. Surprisingly, very few books have appeared about this list of problems; despite the tremendous progress during the last three quarters of a century toward the solutions of practically the entire list of problems.

Hilbert's fourth problem (find all geometries in which the "ordinary lines" are the "geodesics") is particularly attractive. The problem is elementary enough that it can certainly be understood and appreciated by a beginning graduate student of mathematics. However, its solution though of a generally elementary character brings together ideas and tools from many diverse and interesting branches of mathematics: geometry, analysis (especially ordinary and partial differential equations), and the calculus of variations.

A partial solution, under strong assumptions, to Hilbert's fourth problem was already obtained by Georg Hamel in 1901. This present work by A. V. Pogorelov originated from a remarkable idea of Herbert Busemann (in his report to the International Congress of Mathematicians of Moscow in 1966). Pogorelov slightly reformulates Hilbert's problem, and proceeds on the basis of this new idea to give an extremely elegant solution—a real mathematical gem.

This book is extremely well written. Most of the prerequisites (with the exception of standard portions of advanced calculus) are developed as needed. The reader who studies this volume will not only discover how one particular problem is solved, but will also pick up a lot of interesting mathematics on the way.

The English translation was reviewed by Eugene Zaustinsky, who also supplied a very useful set of notes that guide the reader to more literature on the subject.

Pogorelov's book is a welcome addition to the mathematical literature. It will especially be appreciated by those interested in geometry and the foundations of geometry.

Irwin Kra

INTRODUCTION

In the year 1900, at the Second International Congress of Mathematicians in Paris, David Hilbert formulated a number of problems whose investigation would, in his opinion, greatly stimulate the further development of mathematics. His fourth problem was devoted to the foundations of geometry, and consists of the following, as stated by Hilbert himself ([13], pp. 449–451):

"If, from among the axioms necessary to establish ordinary euclidean geometry, we exclude the axiom of parallels, or assume it as not satisfied, but retain all other axioms, we obtain, as is well-known, the geometry of Lobachevski (hyperbolic geometry). We may therefore say that this is a geometry standing next to euclidean geometry. If we require further that that axiom be not satisfied whereby, of three points of a straight line, one and only one lies between the other two, we obtain Riemann's (elliptic) geometry, so that this geometry appears to be the next after Lobachevsky's. If we wish to carry out a similar investigation with respect to the axiom of Archimedes, we must look upon this as not satisfied, and we arrive thereby at the non-archimedean geometries which have been investigated by Veronese and myself [14]. A more general question now arises: Whether from other suggestive standpoints geometries may not be devised which, with equal right, stand next to euclidean geometry. Here I should like to direct your attention to a theorem which has, indeed, been employed by many authors as a definition of a straight line, viz., that the straight line is the shortest distance between two points. The essential content of this statement reduces to the theorem of Euclid that in a triangle the sum of two sides is always greater than the third side—a theorem which, as easily seen, deals solely with elementary concepts, i.e., with such as are derived directly from the axioms, and is therefore more accessible to logical investigation. Euclid proved this theorem, with the help of the theorem of the exterior angle, on the basis of the congruence theorems. Now it is readily shown that this theorem of Euclid cannot be proved solely on the basis of those congruence theorems which relate to the application of segments and angles, but that one of the theorems on the congruence of triangles is necessary. We are asking, then, for a geometry in which all the axioms of ordinary euclidean goemetry hold, and in particular all the congruence axioms except the one of the congruence of triangles (or all except the theorem of the equality of the base angles in the isosceles triangle), and in which, besides, the proposition that in every triangle the sum of two sides is greater than the third is assumed as a particular axiom.

One finds that such a geometry really exists and is none other than that which Minkowski constructed in his book, Geometrie der Zahlen [16] and made the basis of his arithmetical investigations. Minkowski's Geometry is therefore also a geometry standing next to the ordinary euclidean geometry; it is essentially characterized by the following stipulations:

1. The points which are at equal distances from a fixed point O lie on a convex closed surface of the ordinary euclidean space with O as a center.

2. Two segments are said to be equal when one can be carried into the other by a translation of the ordinary euclidean space.

In Minkowski's geometry the axiom of parallels also holds. By studying the theorem of the straight line as the shortest distance between two points, I arrived ([15] and [14], Appendix I) at a geometry in which the parallel axiom does not hold, while all other axioms of Minkowski's geometry are satisfied. The theorem of the straight line as the shortest distance between two points and the essentially equivalent theorem of Euclid about the sides of a triangle, play an important part not only in number theory but also in the theory of surfaces and the calculus of variations. For this reason, and because I believe that the thorough investigation of the conditions for the validity of this theorem will throw a new light upon the idea of distance, as well as upon other elementary ideas, *e.g.*, upon the idea of the plane, and the possibility of its definition by means of the idea of the straight line, *the construction and systematic treatment of the geometries here possible seem to me desirable.*

In the case of the plane and under the assumption of the continuity axiom, the indicated problem leads to the question treated by Darboux ([10] p. 59): Find all variational problems in the plane for which the solutions are all the straight lines of the plane—a question which seems to me capable and worthy of far-reaching generalizations [16]."

This book is devoted to Hilbert's fourth problem [13] and contains its solution when formulated as follows: Find to within an isomorphism all realizations of the axiom systems of the classical geometries (Euclidean, Lobachevskian and elliptic) if, in these systems, we drop the axioms of congruence involving the concept of angle and supplement the systems with the "triangle inequality," regarded as an axiom.

The first and, indeed, the only work devoted to Hilbert's problem in this formulation is due to Hamel [12] the other works being devoted to the study of special Desarguesian spaces. Hamel showed that every solution of Hilbert's problem can be represented in a projective space, or a convex domain of such a space, if congruence of segments is defined as equality of their lengths in a special metric, for which the lines of the space are geodesics. (Such metrics are called Desarguesian metrics.) Thus, the solution of Hilbert's problem was reduced to the problem of the constructive definition of all Desarguesian metrics. Hamel solved this problem under the assumption of a sufficiently regular metric. However, as simple examples show, regular plane metrics by no means exhaust the class of all plane metrics, and the axioms of the geometries under consideration imply only continuity of the metrics. Therefore, a complete solution of Hilbert's problem entails a constructive definition of all continuous Desarguesian metrics and this is the problem to which the present work is devoted.

A. V. Pogorelov

§1. PROJECTIVE SPACE

By a *point* of projective space we mean an ordered quadruple of real numbers $x = (x_1, x_2, x_3, x_4)$, which are not all zero. Proportional quadruples are regarded as *equivalent*, and define the same point of space. The numbers x_1, x_2, x_3, x_4 are called *homogeneous coordinates*.[4]

By a *plane* we mean the set of points satisfying a linear equation

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = 0,$$

and by a *line* we mean the intersection of two distinct planes. Thus, a line is specified by a system of two equations

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = 0, \\ a_1'x_1 + a_2'x_2 + a_3'x_3 + a_4'x_4 = 0, \end{cases}$$

where the rank of the matrix

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ a_1' & a_2' & a_3' & a_4' \end{pmatrix}$$

equals two. It is convenient to use vector notation for the equations of lines and planes. We set

$$a \cdot x = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4,$$

and can then write the equation of a plane in the form $a \cdot x = 0$, and the equation of a line in the form $a \cdot x = 0, a' \cdot x = 0$, where a and a' are linearly independent vectors. We now note some properties of lines and planes.

Let x' and x'' be two distinct points of a line. *Then every point of the line has a representation* $x = \lambda' x' + \lambda'' x''$, *where* λ' *and* λ'' *are real numbers, not both zero.* Conversely, every point with this representation belongs to the line. In fact, the line is specified by a system of equations

$$a' \cdot x = 0, \quad a'' \cdot x = 0 \tag{1}$$

and the points x' and x'' satisfy this system. Since the rank of the system is two, every solution x of the system is a linear combination of the independent solutions x' and x'', i.e., $x = \lambda' x' + \lambda'' x''$. The fact that every point x with this representation satisfies the system (1) is obvious. In view of the indicated representation of a point x of the line, in terms of two given distinct points x' and x'', we conclude that a line is uniquely determined by any two of its distinct points. Hence, *no more than one line passes through two distinct points.*

We next show that *there is a line passing through any two distinct points*. Let the given points be x' and x'' and consider the system of equations

$$a \cdot x' = 0, \quad a \cdot x'' = 0 \tag{2}$$

in a. Since the points x' and x'' are distinct, the rank of the matrix

$$\begin{pmatrix} x_1' & x_2' & x_3' & x_4' \\ x_1'' & x_2'' & x_3'' & x_4'' \end{pmatrix}$$

of the system is two. Therefore the system has two independent solutions a' and a'', and the line specified by the equations

$$a' \cdot x = 0, \quad a'' \cdot x = 0$$

passes through the points x' and x''. Q.E.D.

On every line there are two distinct points. In fact, let $a' \cdot x = 0$, $a'' \cdot x = 0$ be the equations of the line. This system of equations $a' \cdot x = 0$, $a'' \cdot x = 0$ in x has two linearly independent solutions x' and x'', since the rank of the system is two. These solutions give two distinct points of the line determined by the intersection of the planes $a' \cdot x = 0$, $a'' \cdot x = 0$. Q.E.D.

Three distinct points x', x'', x''' lie on the same line if and only if the rank of the matrix

$$\begin{pmatrix} x_1' & x_2' & x_3' & x_4' \\ x_1'' & x_2'' & x_3'' & x_4'' \\ x_1''' & x_2''' & x_3''' & x_4''' \end{pmatrix}$$

is two. In fact, if the points x' and x'' are distinct, then $x''' = \lambda' x' + \lambda'' x''$, by what has been proven. The rank of the matrix is less than three because its rows are linearly dependent. Conversely, if the rank of the matrix is less than three, it must be two since the points x' and x'' are distinct. But, then, the third row can be expressed as a linear combination of the first and second rows; i.e., $x''' = \lambda' x' + \lambda'' x''$. This means that the point x''' lies on the line passing through the points x' and x''. Q.E.D.

There is one and only one plane passing through three noncollinear points. In fact, let x', x'', x''' be the given points and consider the system of equations

$$a \cdot x' = 0, \qquad a \cdot x'' = 0, \qquad a \cdot x''' = 0 \tag{3}$$

in a. The rank of the system is three since the points x', x'', and x''' are noncollinear. The system (3), therefore, has a nontrivial solution a, which is uniquely determined up to a nonzero factor. The plane $a \cdot x = 0$ passes through the given points and is unique, by the uniqueness of the solution of the system (3). Q.E.D.

There is one and only one plane passing through a line and a point not lying on the line. Let us mark two distinct points on the line and draw a plane through them and the given point. This plane contains the given line and passes through the given point. Every plane passing through the given line contains the two distinct points which we marked on it. Our plane is unique because three distinct noncollinear points of a plane determine the plane uniquely. Q.E.D.

A line that does not lie in a plane intersects the plane in one and only one point. Let $a' \cdot x = 0$ be the equation of the plane and let $a'' \cdot x = 0$ and $a''' \cdot x = 0$ be the equations of the line. The system of homogeneous equations $a' \cdot x = 0$, $a'' \cdot x = 0$, $a''' \cdot x = 0$ always has a nontrivial solution x, and this solution gives a point lying on the intersection of the line and the plane. If there were two distinct such points, then, by what has been proven, the line would lie in the plane, contrary to hypothesis. Q.E.D.

Two distinct lines, lying in the same plane, intersect in one and only one point. Let α be the plane and let g_1 and g_2 be the distinct lines lying in the plane α. Take a point x that does not lie in the plane α, and draw planes α_1 and α_2 through x and the lines g_1 and g_2, respectively. The planes α, α_1, α_2 intersect and the point of intersection belongs to the lines g_1 and g_2. Since the lines g_1 and g_2 are distinct, they cannot have other points of intersection, by what has already been proven. Q.E.D.

Let x and y be two points on a line. Then, every point of the line, other than x and y, has a representation $\lambda x + \mu y$, $\lambda\mu \neq 0$. A set of points of the line, for which $\lambda\mu$ has a fixed sign, is called a (*projective line*) *interval with endpoints x and y*. On a projective line there are two intervals with endpoints x and y. We have $\lambda\mu > 0$ on one of these intervals and $\lambda\mu < 0$ on the other. This definition of line interval is obviously independent of the normalization of the coordinates of the endpoints x and y.

Let x^1, x^2, x^3 be three noncollinear points. The figure consisting of these points and three intervals joining them in pairs will be called a *projective triangle* if there exists a plane which does not intersect the triangle.[5] The points x^1, x^2, x^3 are called the *vertices* of the triangle and the intervals joining them are called the *sides* of the triangle.

A projective triangle satisfies the *Axiom of Pasch*: *If a plane does not pass through the vertices of a triangle and intersects one of its sides, then it intersects one and only one of the other two sides of the triangle.* The proof is as follows. Let $a \cdot x = 0$ be the equation of a plane which does not intersect the triangle. Then $a \cdot x^i \neq 0, i = 1, 2, 3$, since this plane does not pass through the vertices of the triangle. We may assume, without loss of generality, that $a \cdot x^i > 0$, as this can always be achieved through a suitable normalization of the points x^i. Because $a \cdot x^k > 0$, the points of the side of the triangle joining the vertices x^i and x^j have a representation $u = \lambda x^i + \mu x^j$, for which $\lambda\mu > 0$. In fact, if we had $\lambda\mu < 0$, we could find values of λ and μ for which $a \cdot u = \lambda a \cdot x^i + \mu a \cdot x^j = 0$; that is, the side $x^i x^j$ of the triangle