METHODS OF MATHEMATICAL PHYSICS

By R. COURANT and D. HILBERT

First English edition
Translated and Revised from the German Original

VOLUME I

19 53

INTERSCIENCE PUBLISHERS, INC., NEW YORK
INTERSCIENCE PUBLISHERS LTD., LONDON

PREFACE

The first German edition of this volume was published by Julius Springer, Berlin, in 1924. A second edition, revised and improved with the help of K. O. Friedrichs, R. Luneburg, F. Rellich, and other unselfish friends, followed in 1930. The second volume appeared in 1938. In the meantime I had been forced to leave Germany and was fortunate and grateful to be given the opportunities open in the United States. During the Second World War the German book became unavailable and later was even suppressed by the National Socialist rulers of Germany. Thus the survival of the book was secured when the United States Government seized the copyright and licensed a reprint issued by Interscience Publishers, New York. Such a license also had to be obtained from the Alien Property Custodian for the present English edition.

This edition follows the German original fairly closely but contains a large number of additions and modifications. I have had to postpone a plan to completely rewrite and modernize the book in collaboration with K. O. Friedrichs, because the pressure for publication of an English "Courant-Hilbert" has become irresistible. Even so, it is hoped that the work in its present form will be useful to mathematicians and physicists alike, as the numerous demands from all sides seem to indicate.

The objective of the book can still today be expressed almost as in the preface to the first German edition. "Since the seventeenth century, physical intuition has served as a vital source for mathematical problems and methods. Recent trends and fashions have, however, weakened the connection between mathematics and physics; mathematicians, turning away from the roots of mathematics in intuition, have concentrated on refinement and emphasized the postulational side of mathematics, and at times have overlooked the unity of their science with physics and other fields. In many cases, physicists have ceased to appreciate the attitudes of mathematicians. This rift is unquestionably a serious threat to science as a whole; the broad stream of scientific development may split into smaller and smaller rivulets and dry out. It seems therefore important to direct our efforts toward reuniting divergent trends by clarifying the common features and interconnections of many distinct and diverse scientific facts. Only thus can the student attain some mastery of the material and the basis be prepared for further organic development of research.

"The present work is designed to serve this purpose for the field of mathematical physics. Mathematical methods originating in problems of physics are developed and the attempt is made to shape results into unified mathematical theories. Completeness is not attempted, but it is hoped that access to a rich and important field will be facilitated by the book.

"The responsibility for the present book rests with me. Yet the name of my teacher, colleague, and friend, D. Hilbert, on the title page seems justified by the fact that much material from Hilbert's papers and lectures has been used, as well as by the hope that the book expresses some of Hilbert's spirit, which has had such a decisive influence on mathematical research and education."

I am greatly indebted to many helpers in all phases of the task of preparing this edition: to Peter Ceike, Ernest Courant, and Anneli Lax, who provided most of the first draft of the translation; to Hanan Rubin and Herbert Kranzer, who have given constructive criticism; to Wilhelm Magnus, who is responsible for the appendix to Chapter VII; and to Natascha Artin and Lucile Gardner, who carried the burden of the editorial work. Most cordial thanks also are due to Interscience Publishers for their patient and helpful attitude and to my old friend and publisher, Dr. Ferdinand Springer in Heidelberg, the great pioneer of modern scientific publishing, for his sympathetic understanding of the situation, which has so greatly changed since the old days of our close cooperation.

R. COURANT

New Rochelle, New York
June 1953

CONTENTS

CHAPTER I

The Algebra of Linear Transformations and Quadratic Forms

In the present volume we shall be concerned with many topics in mathematical analysis which are intimately related to the theory of linear transformations and quadratic forms. A brief résumé of pertinent aspects of this field will, therefore, be given in Chapter I. The reader is assumed to be familiar with the subject in general.

§1. *Linear Equations and Linear Transformations*

1. Vectors. The results of the theory of linear equations can be expressed concisely by the notation of vector analysis. A system of n real numbers $x_1, x_2, \cdots, x_n$ is called an *n-dimensional vector* or a vector in n-dimensional space and denoted by the bold face letter $\mathbf{x}$; the numbers x_i $(i = 1, \cdots, n)$ are called the *components* of the vector $\mathbf{x}$. If all components vanish, the vector is said to be zero or the *null vector*; for $n = 2$ or $n = 3$ a vector can be interpreted geometrically as a "position vector" leading from the origin to the point with the rectangular coordinates x_i. For $n > 3$ geometrical visualization is no longer possible but geometrical terminology remains suitable.

Given two arbitrary real numbers λ and μ, the vector $\lambda\mathbf{x} + \mu\mathbf{y} = \mathbf{z}$ is defined as the vector whose components z_i are given by $z_i = \lambda x_i + \mu y_i$. Thus, in particular, the sum and difference of two vectors are defined.

The number

$$(1) \qquad \mathbf{x}\cdot\mathbf{y} = x_1y_1 + \cdots + x_ny_n = y_1x_1 + \cdots + y_nx_n = \mathbf{y}\cdot\mathbf{x}$$

is called the "*inner product*" of the vectors $\mathbf{x}$ and $\mathbf{y}$.

Occasionally we shall call the inner product $\mathbf{x}\cdot\mathbf{y}$ the *component of the vector* $\mathbf{y}$ *with respect to* $\mathbf{x}$ or vice versa.

If the inner product $\mathbf{x}\cdot\mathbf{y}$ vanishes we say that the vectors $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal*; for $n = 2$ and $n = 3$ this terminology has an immediate geometrical meaning. The inner product $\mathbf{x}\cdot\mathbf{x} = \mathbf{x}^2$ of a vector with itself plays a special role; it is called the *norm* of the vector. The positive square root of $\mathbf{x}^2$ is called the *length* of the vector and denoted by $|\mathbf{x}| = \sqrt{\mathbf{x}^2}$. A vector whose length is unity is called a *normalized vector* or *unit vector.*

The following inequality is satisfied by the inner product of two vectors $\mathbf{a} = (a_1, \cdots, a_n)$ and $\mathbf{b} = (b_1, \cdots, b_n)$:

$$(\mathbf{a}\cdot\mathbf{b})^2 \leq \mathbf{a}^2\mathbf{b}^2$$

or, without using vector notation,

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right),$$

where the equality holds if and only if the a_i and the b_i are proportional, i.e. if a relation of the form $\lambda\mathbf{a} + \mu\mathbf{b} = 0$ with $\lambda^2 + \mu^2 \neq 0$ is satisfied.

The proof of this "*Schwarz inequality*"[1] follows from the fact that the roots of the quadratic equation

$$\sum_{i=1}^{n} (a_i x + b_i)^2 = x^2 \sum_{i=1}^{n} a_i^2 + 2x \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} b_i^2 = 0$$

for the unknown x can never be real and distinct, but must be imaginary, unless the a_i and b_i are proportional. The Schwarz inequality is merely an expression of this fact in terms of the discriminant of the equation. Another proof of the Schwarz inequality follows immediately from the identity

$$\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2 - \left(\sum_{i=1}^{n} a_i b_i\right)^2 = \tfrac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} (a_j b_k - a_k b_j)^2.$$

Vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m$ are said to be *linearly dependent* if a set of numbers $\lambda_1, \lambda_2, \cdots, \lambda_m$ (not all equal to zero) exists such that the vector equation

$$\lambda_1\mathbf{x}_1 + \cdots + \lambda_m\mathbf{x}_m = 0$$

is satisfied, i.e. such that all the components of the vector on the left vanish. Otherwise the vectors are said to be *linearly independent.*

The n vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ in n-dimensional space whose com-

[1] This relation was, as a matter of fact, used by Cauchy before Schwarz.

ponents are given, respectively, by the first, second, $\cdots$, and n-th rows of the array

$$\begin{array}{cccc} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1, \end{array}$$

form a system of n linearly independent vectors. For, if a relation $\lambda_1 \mathbf{e}_1 + \cdots + \lambda_n \mathbf{e}_n = 0$ were satisfied, we could multiply[1] this relation by $\mathbf{e}_h$ and obtain $\lambda_h = 0$ for every h, since $\mathbf{e}_h^2 = 1$ and $\mathbf{e}_h \cdot \mathbf{e}_k = 0$ if $h \neq k$. Thus, systems of n linearly independent vectors certainly exist. However, for any $n + 1$ vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{n+1}$ (in n-dimensional space) there is at least one linear equation of the form

$$\mu_1 \mathbf{u}_1 + \cdots + \mu_{n+1} \mathbf{u}_{n+1} = 0,$$

with coefficients that do not all vanish, since n homogeneous linear equations

$$\sum_{i=1}^{n+1} u_{ik} \mu_i = 0 \qquad (k = 1, \cdots, n)$$

for the $n + 1$ unknowns $\mu_1, \mu_2, \cdots, \mu_{n+1}$ always have at least one nontrivial solution (cf. subsection 3).

2. Orthogonal Systems of Vectors. Completeness. The above "coordinate vectors" $\mathbf{e}_i$ form a particular system of *orthogonal unit vectors*. In general a system of n orthogonal unit vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ is defined as a system of vectors of unit length satisfying the relations

$$\mathbf{e}_h^2 = 1, \qquad \mathbf{e}_h \cdot \mathbf{e}_k = 0 \qquad (h \neq k)$$

for $h, k, = 1, 2, \cdots, n$. As above, we see that the n vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ are linearly independent.

If $\mathbf{x}$ is an arbitrary vector, a relation of the form

$$c_0 \mathbf{x} - c_1 \mathbf{e}_1 - \cdots - c_n \mathbf{e}_n = 0$$

with constants c_i that do not all vanish must hold; for, as we have seen, any $n + 1$ vectors are linearly dependent. Since the $\mathbf{e}_i$ are linearly independent, c_0 cannot be zero; we may therefore, without

[1] To multiply two vectors is to take their inner product.

loss of generality, take it to be equal to unity. Every vector $\mathbf{x}$ can thus be expressed in terms of a system of orthogonal unit vectors in the form

$$\mathbf{x} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n. \tag{2}$$

The coefficients c_i, the *components of* $\mathbf{x}$ *with respect to the system* $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$, may be found by multiplying (2) by each of the vectors $\mathbf{e}_i$; they are

$$c_i = \mathbf{x}\cdot\mathbf{e}_i.$$

From any arbitrary system of m linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$, we may, by the following *orthogonalization process* due to E. Schmidt, obtain a system of m orthogonal unit vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_m$: First set $\mathbf{e}_1 = \mathbf{v}_1/|\mathbf{v}_1|$. Then choose a number c_1' in such a way that $\mathbf{v}_2 - c_1'\mathbf{e}_1$ is orthogonal to $\mathbf{e}_1$, i.e. set $c_1' = \mathbf{v}_2\cdot\mathbf{e}_1$. Since $\mathbf{v}_1$ and $\mathbf{v}_2$, and therefore $\mathbf{e}_1$ and $\mathbf{v}_2$, are linearly independent, the vector $\mathbf{v}_2 - c_1'\mathbf{e}_1$ is different from zero. We may then divide this vector by its length obtaining a unit vector $\mathbf{e}_2$ which is orthogonal to $\mathbf{e}_1$. We next find two numbers c_1'', c_2'' such that $\mathbf{v}_3 - c_1''\mathbf{e}_1 - c_2''\mathbf{e}_2$ is orthogonal to both $\mathbf{e}_1$ and $\mathbf{e}_2$, i.e. we set $c_1'' = \mathbf{v}_3\cdot\mathbf{e}_1$ and $c_2'' = \mathbf{v}_3\cdot\mathbf{e}_2$. This vector is again different from zero and can, therefore, be normalized; we divide it by its length and obtain the unit vector $\mathbf{e}_3$. By continuing this procedure we obtain the desired orthogonal system.

For $m < n$ the resulting orthogonal system is called *incomplete*, and if $m = n$ we speak of a *complete orthogonal system*. Let us denote the components of a vector $\mathbf{x}$ with respect to $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_m$ by $c_1, c_2, \cdots, c_m$ as before. The self-evident inequality

$$(\mathbf{x} - c_1\mathbf{e}_1 - \cdots - c_m\mathbf{e}_m)^2 \geq 0$$

is satisfied. Evaluating the left side term by term according to the usual rules of algebra (which hold for vectors if the inner product of two vectors is used whenever two vectors are multiplied), we find

$$\mathbf{x}^2 - 2\mathbf{x}\cdot\sum_{i=1}^{m} c_i\mathbf{e}_i + \sum_{i=1}^{m} c_i^2 = \mathbf{x}^2 - 2\sum_{i=1}^{m} c_i^2 + \sum_{i=1}^{m} c_i^2 \geq 0$$

or

$$\mathbf{x}^2 \geq \sum_{i=1}^{m} c_i^2, \tag{3}$$

where $m \leq n$ and $c_i = \mathbf{x} \cdot \mathbf{e}_i$; the following equality holds for $m = n$:

$$\mathbf{x}^2 = \sum_{i=1}^{m} c_i^2. \tag{4}$$

Relations (3) and (4)—(4) expresses the theorem of Pythagoras in vector notation—have an intuitive significance for $n \leq 3$; they are called, respectively, *Bessel's inequality* and the *completeness relation*. Relation (4), if it is satisfied for every vector $\mathbf{x}$, does in fact indicate that the given orthogonal system is complete since (4) could not be satisfied for a unit vector orthogonal to all vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_m$, and such a vector necessarily exists if $m < n$.

The completeness relation may also be expressed in the more general form

$$\mathbf{x} \cdot \mathbf{x}' = \sum_{i=1}^{m} c_i c_i', \tag{5}$$

which follows from the orthogonality of the $\mathbf{e}_i$.

So far these algebraic relations are all purely formal. Their significance lies in the fact that they occur again in a similar manner in transcendental problems of analysis.

3. Linear Transformations. Matrices. A system of n linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= y_n; \end{aligned} \tag{6}$$

with coefficients a_{ik}, assigns a unique set of quantities $y_1, y_2, \cdots, y_n$ to every set of quantities $x_1, x_2, \cdots, x_n$. Such an assignment is called a *linear transformation* of the set $x_1, x_2, \cdots, x_n$ into the set $y_1, y_2, \cdots, y_n$, or, briefly, of the vector $\mathbf{x}$ into the vector $\mathbf{y}$. This transformation is clearly linear since the vector $\lambda_1\mathbf{y}_1 + \lambda_2\mathbf{y}_2$ corresponds to the vector $\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2$.

The most important problem in connection with linear transformations is the problem of inversion, the question, in other words, of the