ALGEBRA

A TEXT-BOOK OF DETERMINANTS, MATRICES, AND ALGEBRAIC FORMS

BY

W. L. FERRAR

M.A., F.R.S.E.

FELLOW AND TUTOR OF HERTFORD COLLEGE
OXFORD

SECOND EDITION

OXFORD UNIVERSITY PRESS

Oxford University Press, Amen House, London E.C.4
GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS KARACHI KUALA LUMPUR
CAPE TOWN IBADAN NAIROBI ACCRA

FIRST EDITION 1941
Reprinted lithographically in Great Britain
by LOWE & BRYDONE, PRINTERS, LTD., LONDON
from sheets of the first edition 1942, 1946,
1948, 1950, 1953
SECOND EDITION 1957
REPRINTED 1960

PREFACE TO SECOND EDITION

In revising the book for a second edition I have corrected a number of slips and misprints and I have added a new chapter on latent vectors.

W. L. F.

HERTFORD COLLEGE, OXFORD,
1957.

PREFACE TO FIRST EDITION

In writing this book I have tried to provide a text-book of the more elementary properties of determinants, matrices, and algebraic forms. The sections on determinants and matrices, Parts I and II of the book, are, to some extent, suitable either for undergraduates or for boys in their last year at school. Part III is suitable for study at a university and is not intended to be read at school.

The book as a whole is written primarily for undergraduates. University teaching in mathematics should, in my view, provide at least two things. The first is a broad basis of knowledge comprising such theories and theorems in any one branch of mathematics as are of constant application in other branches. The second is incentive and opportunity to acquire a detailed knowledge of some one branch of mathematics. The books available make reasonable provision for the latter, especially if the student has, as he should have, a working knowledge of at least one foreign language. But we are deplorably lacking in books that cut down each topic, I will not say to a minimum, but to something that may reasonably be reckoned as an essential part of an undergraduate's mathematical education.

Accordingly, I have written this book on the same general plan as that adopted in my book on convergence. I have included topics commonly required for a university honours course in pure and applied mathematics: I have excluded topics appropriate to post-graduate or to highly specialized courses of study.

Some of the books to which I am indebted may well serve as a guide for my readers to further algebraic reading. Without pretending that the list exhausts my indebtedness to others, I may note the following: Scott and Matthews, *Theory of Determinants*; Bôcher, *Introduction to Higher Algebra*; Dickson, *Modern Algebraic Theories*; Aitken, *Determinants and Matrices*; Turnbull, *The Theory of Determinants, Matrices, and Invariants*; Elliott, *Introduction to the Algebra of Quantics*; Turnbull and Aitken, *The Theory of Canonical Matrices*; Salmon, *Modern*

Higher Algebra (though the 'modern' refers to some sixty years back); Burnside and Panton, *Theory of Equations.* Further, though the reference will be useless to my readers, I gratefully acknowledge my debt to Professor E. T. Whittaker, whose invaluable 'research lectures' on matrices I studied at Edinburgh many years ago.

The omissions from this book are many. I hope they are, all of them, deliberate. It would have been easy to fit in something about the theory of equations and eliminants, or to digress at one of several possible points in order to introduce the notion of a group, or to enlarge upon number rings and fields so as to give some hint of modern abstract algebra. A book written expressly for undergraduates and dealing with one or more of these topics would be a valuable addition to our stock of university text-books, but I think little is to be gained by references to such subjects when it is not intended to develop them seriously.

In part, the book was read, while still in manuscript, by my friend and colleague, the late Mr. J. Hodgkinson, whose excellent lectures on Algebra will be remembered by many Oxford men. In the exacting task of reading proofs and checking references I have again received invaluable help from Professor E. T. Copson, who has read all the proofs once and checked nearly all the examples. I am deeply grateful to him for this work and, in particular, for the criticisms which have enabled me to remove some notable faults from the text.

Finally, I wish to thank the staff of the University Press, both on its publishing and its printing side, for their excellent work on this book. I have been concerned with the printing of mathematical work (mostly that of other people!) for many years, and I still marvel at the patience and skill that go to the printing of a mathematical book or periodical.

W. L. F.

HERTFORD COLLEGE, OXFORD,
September 1940.

CONTENTS

PART I

DETERMINANTS

PRELIMINARY NOTE: NUMBER, NUMBER FIELDS, AND MATRICES 1

I. ELEMENTARY PROPERTIES OF DETERMINANTS . . . 5

II. THE MINORS OF A DETERMINANT 28

III. THE PRODUCT OF TWO DETERMINANTS . . . 41

IV. JACOBI'S THEOREM AND ITS EXTENSIONS . . . 55

V. SYMMETRICAL AND SKEW-SYMMETRICAL DETERMINANTS 59

PART II

MATRICES

VI. DEFINITIONS AND ELEMENTARY PROPERTIES . . . 67

VII. RELATED MATRICES 83

VIII. THE RANK OF A MATRIX 93

IX. DETERMINANTS ASSOCIATED WITH MATRICES . . . 109

PART III

LINEAR AND QUADRATIC FORMS

X. ALGEBRAIC FORMS: LINEAR TRANSFORMATIONS . 119

XI. THE POSITIVE-DEFINITE FORM 135

XII. THE CHARACTERISTIC EQUATION AND CANONICAL FORMS 144

XIII. ORTHOGONAL TRANSFORMATIONS 160

XIV. INVARIANTS AND COVARIANTS 169

XV. LATENT VECTORS 200

INDEX . . . .

PART I

PRELIMINARY NOTE; CHAPTERS ON DETERMINANTS

PRELIMINARY NOTE

1. Number

In its initial stages algebra is little more than a generalization of elementary arithmetic. It deals only with the positive integers, 1, 2, 3,.... We can all remember the type of problem that began 'let x be the number of eggs', and if x came to $3\frac{1}{2}$ we knew we were wrong.

In later stages x is permitted to be negative or zero, to be the ratio of two integers, and then to be any real number either rational, such as $3\frac{1}{3}$ or $-\frac{1}{4}$, or irrational, such as π or $\sqrt{3}$. Finally, with the solution of the quadratic equation, x is permitted to be a complex number, such as $2+3i$.

The numbers used in this book may be either real or complex and we shall assume that readers have studied, to a greater or a lesser extent, the precise definitions of these numbers and the rules governing their addition, subtraction, multiplication, and division.

2. Number rings

Consider the set of numbers

$$0,\ \pm 1,\ \pm 2,\ \dots. \tag{1}$$

Let r, s denote numbers selected from (1). Then, whether r and s denote the same or different numbers, the numbers

$$r+s,\quad r-s,\quad r\times s$$

all belong to (1). This property of the set (1) is shared by other sets of numbers. For example,

$$\text{all numbers of the form } a+b\sqrt{5}, \tag{2}$$

where a and b belong to (1), have the same property; if r and s belong to (2), then so do $r+s$, $r-s$, and $r\times s$. A set of numbers having this property is called a RING of numbers.

3. Number fields

3.1. Consider the set of numbers comprising 0 and every number of the form p/q, where p and q belong to (1) and q is not zero, that is to say,

$$\text{the set of all rational real numbers.} \tag{3}$$

Let r, s denote numbers selected from (3). Then, when s is not zero, whether r and s denote the same or different numbers, the numbers

$$r+s, \quad r-s, \quad r\times s, \quad r\div s$$

all belong to the set (3).

This property characterizes what is called a FIELD of numbers. The property is shared by the following sets, among many others:—

the set of all complex numbers; (4)

the set of all real numbers (rational and irrational); (5)

the set of all numbers of the form $p+q\surd 3$, where p and q belong to (3). (6)

Each of the sets (4), (5), and (6) constitutes a field.

DEFINITION. *A set of numbers, real or complex, is said to form a* FIELD OF NUMBERS *when, if r and s belong to the set and s is not zero,*

$$r+s, \quad r-s, \quad r\times s, \quad r\div s$$

also belong to the set.

Notice that the set (1) is not a field; for, whereas it contains the numbers 1 and 2, it does not contain the number $\frac{1}{2}$.

3.2. Most of the propositions in this book presuppose that the work is carried out within *a* field of numbers; what *particular* field is usually of little consequence.

In the early part of the book this aspect of the matter need not be emphasized: in some of the later chapters the essence of the theorem is that all the operations envisaged by the theorem can be carried out within the confines of any given field of numbers.

In this preliminary note we wish to do no more than give a formal definition of a field of numbers and to familiarize the reader with the concept.

4. Matrices

A set of mn numbers, real or complex, arranged in an array of m columns and n rows is called a matrix. Thus

$$\begin{matrix} a_{11} & a_{12} & . & . & a_{1m} \\ a_{21} & a_{22} & . & . & a_{2m} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nm} \end{matrix}$$

is a matrix. When $m = n$ we speak of a square matrix of order n.

Associated with any given square matrix of order n there are a number of algebraical entities. The matrix written above, with $m = n$, is associated

(i) with the determinant

$$\begin{vmatrix} a_{11} & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & a_{2n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nn} \end{vmatrix};$$

(ii) with the form

$$\sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} x_r x_s,$$

of degree 2 in the n variables $x_1, x_2, \ldots, x_n$;

(iii) with the bilinear form

$$\sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} x_r y_s$$

in the $2n$ variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$;

(iv) with the Hermitian form

$$\sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} x_r \bar{x}_s,$$

where x_s and $\bar{x}_s$ are conjugate complex numbers;

(v) with the linear transformations

$$x_r = \sum_{s=1}^{n} a_{rs} X_s \quad (r = 1, \ldots, n),$$

$$L_s = \sum_{r=1}^{n} a_{rs} l_r \quad (s = 1, \ldots, n).$$

The theories of the matrix and of its associated forms are closely knit together. The plan of expounding these theories that I have adopted is, roughly, this: Part I develops properties of the determinant; Part II develops the algebra of matrices, referring back to Part I for any result about determinants that may be needed; Part III develops the theory of the other associated forms.

CHAPTER I

ELEMENTARY PROPERTIES OF DETERMINANTS

1. Introduction

1.1. In the following chapters it is assumed that most readers will already be familiar with determinants of the second and third orders. On the other hand, no theorems about such determinants are assumed, so that the account given here is complete in itself.

Until the middle of the last century the use of determinant notation was practically unknown, but once introduced it gained such popularity that it is now employed in almost every branch of mathematics. The theory has been developed to such an extent that few mathematicians would pretend to a knowledge of the whole of it. On the other hand, the range of theory that is of constant application in other branches of mathematics is relatively small, and it is this restricted range that the book covers.

1.2. Determinants of the second and third orders. Determinants are, in origin, closely connected with the solution of linear equations.

Suppose that the two equations

$$a_1 x + b_1 y = 0, \qquad a_2 x + b_2 y = 0$$

are satisfied by a pair of numbers x and y, one of them, at least, being different from zero. Then

$$b_2(a_1 x + b_1 y) - b_1(a_2 x + b_2 y) = 0,$$

and so $$(a_1 b_2 - a_2 b_1)x = 0.$$

Similarly, $(a_1 b_2 - a_2 b_1)y = 0$, and so $a_1 b_2 - a_2 b_1 = 0$.

The number $a_1 b_2 - a_2 b_1$ is a simple example of a determinant; it is usually written as

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}. \tag{1}$$

The term $a_1 b_2$ is referred to as 'the leading diagonal'. Since there are two rows and two columns, the determinant is said to be 'of order two', or 'of the second order'.

The determinant has one obvious property. If, in (1), we interchange simultaneously a_1 and b_1, a_2 and b_2, we get

$$b_1 a_2 - b_2 a_1 \quad \text{instead of} \quad a_1 b_2 - a_2 b_1.$$

That is, the interchange of two columns of (1) reproduces the same terms, namely $a_1 b_2$ and $a_2 b_1$, but in a different order and with the opposite signs.

Again, let numbers x, y, and z, not all zero, satisfy the three equations

$$a_1 x + b_1 y + c_1 z = 0, \tag{2}$$

$$a_2 x + b_2 y + c_2 z = 0, \tag{3}$$

$$a_3 x + b_3 y + c_3 z = 0; \tag{4}$$

then, from equations (3) and (4),

$$(a_2 b_3 - a_3 b_2)x - (b_2 c_3 - b_3 c_2)z = 0,$$

$$(a_2 b_3 - a_3 b_2)y - (c_2 a_3 - c_3 a_2)z = 0,$$

and so, from equation (2),

$$z\{a_1(b_2 c_3 - b_3 c_2) + b_1(c_2 a_3 - c_3 a_2) + c_1(a_2 b_3 - a_3 b_2)\} = 0.$$

We denote the coefficient of z, which may be written as

$$a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1,$$

by Δ; so that our result is $z\Delta = 0$.

By similar working we can show that

$$x\Delta = 0, \qquad y\Delta = 0.$$

Since x, y, and z are not all zero, $\Delta = 0$.

The number Δ is usually written as

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix},$$

in which form it is referred to as a 'determinant of order three' or a 'determinant of the third order'. The term $a_1 b_2 c_3$ is referred to as 'the leading diagonal'.

1.3. It is thus suggested that, associated with n linear equations in n variables, say

$$a_1 x + b_1 y + \ldots + k_1 z = 0,$$

$$a_2 x + b_2 y + \ldots + k_2 z = 0,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$a_n x + b_n y + \ldots + k_n z = 0,$$

there is a certain function of the coefficients which must be zero if all the equations are to be satisfied by a set of values x, y,..., z which are not all zero. It is suggested that this function of the coefficients may be conveniently denoted by

$$\begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix},$$

in which form it may be referred to as a determinant of order n, and $a_1\, b_2 \ldots k_n$ called the leading diagonal.

Just as we formed a determinant of order three (in § 1.2) by using determinants of order two, so we could form a determinant of order four by using those of order three, and proceed step by step to a definition of a determinant of order n. But this is not the only possible procedure and we shall arrive at our definition by another path.

We shall first observe certain properties of determinants of the third order and then define a determinant of order n in such a way that these properties are preserved for determinants of every order.

1.4. Note on definitions. There are many different ways of defining a determinant of order n, though all the definitions lead to the same result in the end. The only particular merit we claim for our own definition is that it is easily reconcilable with any of the others, and so makes reference to other books a simple matter.

1.5. Properties of determinants of order three. As we have seen in § 1.2, the determinant

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

stands for the expression

$$+a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1. \qquad (1)$$

The following facts are all but self-evident:

(I) The expression (1) is of the form

$$\sum \pm a_r b_s c_t,$$

wherein the sum is taken over the six possible ways of assigning to r, s, t the values 1, 2, 3 in some order and without repetition.

(II) The leading diagonal term $a_1 b_2 c_3$ is prefixed by $+$.

(III) As with the determinant of order 2 (§ 1.2), the interchange of any two letters throughout the expression (1) reproduces the same set of terms, but in a different order and with the opposite signs prefixed to them. For example, when a and b are interchanged† in (1), we get

$$+b_1 a_2 c_3 - b_1 a_3 c_2 + b_2 a_3 c_1 - b_2 a_1 c_3 + b_3 a_1 c_2 - b_3 a_2 c_1,$$

which consists of the terms of (1), but in a different order and with the opposite signs prefixed.

2. Determinants of order n

2.1. Having observed (§ 1.5) three essential properties of a determinant of the third order, we now define a determinant of order n.

DEFINITION. *The determinant*

$$\Delta_n \equiv \begin{vmatrix} a_1 & b_1 & . & . & j_1 & k_1 \\ a_2 & b_2 & . & . & j_2 & k_2 \\ . & . & . & . & . & . \\ a_n & b_n & . & . & j_n & k_n \end{vmatrix} \tag{1}$$

is that function of the a's, b's,..., k's which satisfies the three conditions:

(I) *it is an expression of the form*

$$\sum \pm a_r b_s \dots k_\theta, \tag{2}$$

wherein the sum is taken over the $n!$ possible ways of assigning to r, s,..., θ the values 1, 2,..., *n in some order, and without repetition;*

(II) *the leading diagonal term, $a_1 b_2 \dots k_n$, is prefixed by the sign $+$;*

(III) *the sign prefixed to any other term is such that the interchange of any two letters‡ throughout* (2) *reproduces the same*

---

† Throughout we use the phrase 'interchange a and b' to denote the simultaneous interchanges a_1 and b_1, a_2 and b_2, a_3 and b_3, ..., a_n and b_n.

‡ See previous footnote. The interchange of p and q, say, means the simultaneous interchanges

p_1 and q_1, p_2 and q_2,..., p_n and q_n.

set of terms, but in a different order of occurrence, and with the opposite signs prefixed.

Before proceeding we must prove that the definition yields one function of the a's, b's,..., k's and one only. The proof that follows is divided into four main steps.

First step. Let the letters $a, b, \ldots, k$, which correspond to the columns of (1), be written down in any order, say

$$\ldots,\ d,\ g,\ a,\ \ldots,\ p,\ \ldots,\ q,\ \ldots. \tag{A}$$

An interchange of two letters that stand next to each other is called an ADJACENT INTERCHANGE. Take any two letters p and q, having, say, m letters between them in the order (A). By $m+1$ adjacent interchanges, in each of which p is moved one place to the right, we reach a stage at which p comes next after q; by m further adjacent interchanges, in each of which q is moved one place to the left, we reach a stage at which the order (A) is reproduced save that p and q have changed places. This stage has been reached by means of $2m+1$ adjacent interchanges.

Now if, in (2), we change all the signs $2m+1$ times, we end with signs opposite to our initial signs. Accordingly, if the condition (III) of the definition is satisfied for adjacent interchanges of letters, it is automatically satisfied for every interchange of letters.

Second step. The conditions (I), (II), (III) fix the value of the determinant (of the second order)

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

to be $a_1 b_2 - a_2 b_1$. For, by (I) and (II), the value must be

$$+a_1 b_2 \pm a_2 b_1,$$

and, by (III), the interchange of a and b must change the signs, so that we cannot have $a_1 b_2 + a_2 b_1$.

Third step. Assume, then, that the conditions (I), (II), (III) are sufficient to fix the value of a determinant of order $n-1$.

By (I), the determinant Δ_n contains a set of terms in which a has the suffix 1; this set of terms is

$$a_1 \sum \pm b_s c_t \dots k_\theta, \tag{3}$$

wherein, by (I), as applied to Δ_n,

(i) the sum is taken over the $(n-1)!$ possible ways of assigning to $s, t, \dots, \theta$ the values $2, 3, \dots, n$ in some order and without repetition.

Moreover, by (II), as applied to Δ_n,

(ii) the term $b_2 c_3 \dots k_n$ is prefixed by $+$.

Finally, by (III), as applied to Δ_n,

(iii) an interchange of any two of the letters $b, c, \dots, k$ changes the signs throughout (3).

Hence, by our hypothesis that the conditions (I), (II), (III) fix the value of a determinant of order $n-1$, the terms of (2) in which a has the suffix 1 are given by

$$a_1 \begin{vmatrix} b_2 & c_2 & . & . & k_2 \\ b_3 & c_3 & . & . & k_3 \\ . & . & . & . & . \\ b_n & c_n & . & . & k_n \end{vmatrix}. \tag{3 a}$$

This, on our assumption that a determinant of order $n-1$ is defined by the conditions (I), (II), (III), fixes the signs of all terms in (2) that contain $a_1 b_2, a_1 b_3, \dots, a_1 b_n$.

Fourth step. The interchange of a and b in (2) must, by condition (III), change all the signs in (2). Hence the terms of (2) in which b has the suffix 1 are given by

$$-b_1 \begin{vmatrix} a_2 & c_2 & . & . & k_2 \\ a_3 & c_3 & . & . & k_3 \\ . & . & . & . & . \\ a_n & c_n & . & . & k_n \end{vmatrix}; \tag{3 b}$$

for (3 b) fixes the sign of a term $b_1 a_s c_t \dots k_\theta$ to be the opposite of the sign of the term $a_1 b_s c_t \dots k_\theta$ in (3 a).

The adjacent interchanges b with c, c with d, ..., j with k now show that (2) must take the form

$$a_1\begin{vmatrix} b_2 & c_2 & . & . & k_2 \\ b_3 & c_3 & . & . & k_3 \\ . & . & . & . & . \\ b_n & c_n & . & . & k_n \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 & . & . & k_2 \\ a_3 & c_3 & . & . & k_3 \\ . & . & . & . & . \\ a_n & c_n & . & . & k_n \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 & . & . & k_2 \\ a_3 & b_3 & . & . & k_3 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix} -$$

$$-\ldots+(-1)^{n-1}k_1\begin{vmatrix} a_2 & b_2 & . & . & j_2 \\ a_3 & b_3 & . & . & j_3 \\ . & . & . & . & . \\ a_n & b_n & . & . & j_n \end{vmatrix}. \qquad (4)$$

That is to say, if conditions (I), (II), (III) define uniquely a determinant of order $n-1$, then they define uniquely a determinant of order n. But they do define uniquely a determinant of order 2, and hence, by induction, they define uniquely a determinant of any order.

2.2. Rule for determining the sign of a given term. If in a term $a_r b_s \ldots k_\theta$ there are λ_n suffixes less than n that come after n, we say that there are λ_n inversions with respect to n. For example, in the term $a_2 b_3 c_4 d_1$, there is one inversion with respect to 4. Similarly, if there are λ_{n-1} suffixes less than $n-1$ that come after $n-1$, we say that there are λ_{n-1} inversions with respect to $n-1$; and so on. The sum

$$N = \lambda_1+\lambda_2+\ldots+\lambda_n$$

is called the total number of inversions of suffixes. Thus, with $n = 6$ and the term

$$a_4 b_3 c_2 d_6 e_1 f_5, \qquad (5)$$

$\lambda_6 = 2$, since the suffixes 1 and 5 come after 6,

$\lambda_5 = 0$, since no suffix less than 5 comes after 5,

$\lambda_4 = 3$, since the suffixes 3, 2, 1 come after 4,

$\lambda_3 = 2,\ \lambda_2 = 1,\ \lambda_1 = 0$;

the total number of inversions is $2+3+2+1 = 8$.

If $a_r b_s \ldots k_\theta$ has λ_n inversions with respect to n, then, leaving the order of the suffixes $1, 2,\ldots, n-1$ unchanged, we can make n to be the suffix of the nth letter of the alphabet by λ_n adjacent interchanges of letters and, on restoring alphabetical order, make n the last suffix. For example, in (5), where $\lambda_n = 2$, the two adjacent interchanges f with e and f with d give, in succession,

$$a_4 b_3 c_2 d_6 f_1 e_5, \qquad a_4 b_3 c_2 f_6 d_1 e_5.$$

On restoring alphabetical order in the last form, we have $a_4 b_3 c_2 d_1 e_5 f_6$, in which the suffixes 4, 3, 2, 1, 5 are in their original order, as in (5), and the suffix 6 comes at the end.

Similarly, when n has been made the last suffix, λ_{n-1} adjacent interchanges of letters followed by a restoration of alphabetical order will then make $n-1$ the $(n-1)$th suffix; and so on.

Thus $\lambda_1+\lambda_2+...+\lambda_n$ adjacent interchanges of letters make the term $a_r b_s ... k_\theta$ coincide with $a_1 b_2 ... k_n$. By (III), the sign to be prefixed to any term of (2) is $(-1)^N$, where N, i.e. $\lambda_1+\lambda_2+...+\lambda_n$, is the total number of inversions of suffixes.

2.3. The number N may also be arrived at in another way. Let $1 \leqslant m \leqslant n$. In the term

$$a_r b_s ... k_\theta$$

let there be μ_m suffixes greater than m that come before m. Then the suffix m comes after each of these μ_m greater suffixes and, in evaluating N, accounts for one inversion with respect to each of them. It follows that

$$N = \sum_{m=1}^{n} \mu_m. \tag{6}$$

3. Properties of a determinant

3.1. THEOREM 1. *The determinant* Δ_n *of* § 2.1 *can be expanded in either of the forms*

(i) $$\sum (-1)^N a_r b_s ... k_\theta,$$

where N *is the total number of inversions in the suffixes* $r, s,..., \theta$;

(ii) $$a_1 \begin{vmatrix} b_2 & c_2 & . & . & k_2 \\ b_3 & c_3 & . & . & k_3 \\ . & . & . & . & . \\ b_n & c_n & . & . & k_n \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 & . & . & k_2 \\ a_3 & c_3 & . & . & k_3 \\ . & . & . & . & . \\ a_n & c_n & . & . & k_n \end{vmatrix} + ... +$$
$$+(-1)^{n-1} k_1 \begin{vmatrix} a_2 & b_2 & . & . & j_2 \\ a_3 & b_3 & . & . & j_3 \\ . & . & . & . & . \\ a_n & b_n & . & . & j_n \end{vmatrix}.$$

This theorem has been proved in § 2.

THEOREM 2. *A determinant is unaltered in value when rows and columns are interchanged; that is to say*

$$\begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & . & . & a_n \\ b_1 & b_2 & . & . & b_n \\ . & . & . & . & . \\ k_1 & k_2 & . & . & k_n \end{vmatrix}.$$

By Theorem 1, the second determinant is

$$\sum (-1)^M \alpha_1 \beta_2 \dots \kappa_n, \tag{7}$$

where α, β,..., κ are the letters a, b,..., k in some order and M is the total number of inversions of letters.

[There are μ_n inversions with respect to k in $\alpha\beta...\kappa$ if there are μ_n letters after k that come before k in the alphabet; and so on: $\mu_1+\mu_2+...+\mu_n$ is the total number of inversions of letters.]

Now consider any one term of (7), say

$$(-1)^M \alpha_1 \beta_2 \dots \kappa_n. \tag{8}$$

If we write the product with its letters in alphabetical order, we get a term of the form

$$(-1)^M a_r b_s \dots j_t k_\theta. \tag{9}$$

In (8) there are μ_n letters that come after k, so that in (9) there are μ_n suffixes greater than θ that come before θ. There are μ_{n-1} letters that come before j in the alphabet but after it in (8), so there are μ_{n-1} suffixes greater than t that come before it in (9); and so on. It follows from § 2.3 that M, which is defined as $\sum \mu_n$, is equal to N, where N is the total number of inversions of suffixes in (9).

Thus (7), which is the expansion of the second determinant of the enunciation, may also be written as

$$\sum (-1)^N a_r b_s \dots k_\theta,$$

which is, by Theorem 1, the expansion of the first determinant of the enunciation; and Theorem 2 is proved.

THEOREM 3. *The interchange of two columns, or of two rows, in a determinant multiplies the value of the determinant by* -1.

It follows at once from (III) of the definition of Δ_n that an interchange of two columns, i.e. an interchange of two letters, multiplies the determinant by -1.

Hence also, by Theorem 2, an interchange of two rows multiplies the determinant by -1.

Corollary. *If a column (row) is moved past an even number of columns (rows), then the value of the determinant is unaltered; in particular*

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \begin{vmatrix} c_1 & a_1 & b_1 & d_1 \\ c_2 & a_2 & b_2 & d_2 \\ c_3 & a_3 & b_3 & d_3 \\ c_4 & a_4 & b_4 & d_4 \end{vmatrix}.$$

If a column (row) is moved past an odd number of columns (rows), then the value of the determinant is thereby multiplied by -1.

For a column can be moved past an even (odd) number of columns by an even (odd) number of adjacent interchanges. In the particular example, $abcd$ can be changed into $cabd$ by first interchanging b and c, giving $acbd$, and then interchanging a and c.

3.11. The expansion (ii) of Theorem 1 is usually referred to as the expansion by the first row. By the corollary of Theorem 3, there is a similar expansion by any other row. For example,

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \begin{vmatrix} a_3 & b_3 & c_3 & d_3 \\ a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}$$

and, on expanding the second determinant by its first row, the first determinant is seen to be equal to

$$a_3\begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_4 & c_4 & d_4 \end{vmatrix} - b_3\begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_4 & c_4 & d_4 \end{vmatrix} + c_3\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_4 & b_4 & d_4 \end{vmatrix} - d_3\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_4 & b_4 & c_4 \end{vmatrix}.$$

Similarly, we may show that the first determinant may be written as

$$-a_2\begin{vmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} + b_2\begin{vmatrix} a_1 & c_1 & d_1 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} - c_2\begin{vmatrix} a_1 & b_1 & d_1 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix} + d_2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{vmatrix}.$$

Again, by Theorem 2, we may turn columns into rows and rows into columns without changing the value of the determinant. Hence there are corresponding expansions of the determinant by each of its columns.

We shall return to this point in Chapter II, § 1.

3.2. THEOREM 4. *If a determinant has two columns, or two rows, identical, its value is zero.*

The interchange of the two identical columns, or rows, leaves its value unaltered. But, by Theorem 3, if its value is x, its value after the interchange of the two columns, or rows, is $-x$. Hence $x = -x$, or $2x = 0$.

THEOREM 5. *If each element of one column, or row, is multiplied by a factor K, the value of the determinant is thereby multiplied by K.*

This is an immediate corollary of the definition, for

$$\sum \pm K a_r b_s \dots k_\theta = K \sum \pm a_r b_s \dots k_\theta.$$

3.3. THEOREM 6. *The determinant of order n,*

$$\begin{vmatrix} a_1+\alpha_1 & b_1+\beta_1 & . & . & k_1+\kappa_1 \\ a_2+\alpha_2 & b_2+\beta_2 & . & . & k_2+\kappa_2 \\ . & . & . & . & . \\ a_n+\alpha_n & b_n+\beta_n & . & . & k_n+\kappa_n \end{vmatrix},$$

is equal to the sum of the 2^n determinants corresponding to the 2^n different ways of choosing one letter from each column; in particular,

$$\begin{vmatrix} a_1+\alpha_1 & b_1+\beta_1 \\ a_2+\alpha_2 & b_2+\beta_2 \end{vmatrix}$$

is the sum of the four determinants

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & \beta_1 \\ a_2 & \beta_2 \end{vmatrix}, \quad \begin{vmatrix} \alpha_1 & b_1 \\ \alpha_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix}.$$

This again is obvious from the definition; for

$$\sum \pm (a_r+\alpha_r)(b_s+\beta_s)\dots(k_\theta+\kappa_\theta)$$

is the algebraic sum of the 2^n summations typified by taking one term from each bracket.

3.4. THEOREM 7. *The value of a determinant is unaltered if to each element of one column (or row) is added a constant multiple of the corresponding element of another column (or row); in particular,*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+\lambda b_1 & b_1 & c_1 \\ a_2+\lambda b_2 & b_2 & c_2 \\ a_3+\lambda b_3 & b_3 & c_3 \end{vmatrix}.$$

Proof. Let the determinant be Δ_n of § 2.1, x and y the letters of two distinct columns of Δ_n. Let Δ'_n be the determinant formed from Δ_n by replacing each x_r of the x column by $x_r+\lambda y_r$, where λ is independent of r. Then, by Theorem 1,

$$\begin{aligned} \Delta'_n &= \sum (-1)^N a_r b_s \ldots (x_u+\lambda y_u) \ldots y_t \ldots k_\theta \\ &= \begin{vmatrix} a_1 & b_1 & \cdot & x_1 & \cdot & y_1 & \cdot & k_1 \\ a_2 & b_2 & \cdot & x_2 & \cdot & y_2 & \cdot & k_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & \cdot & x_n & \cdot & y_n & \cdot & k_n \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & \cdot & \lambda y_1 & \cdot & y_1 & \cdot & k_1 \\ a_2 & b_2 & \cdot & \lambda y_2 & \cdot & y_2 & \cdot & k_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & \cdot & \lambda y_n & \cdot & y_n & \cdot & k_n \end{vmatrix}. \end{aligned}$$

But the second of these is zero since, by Theorem 5, it is λ times a determinant which has two columns identical. Hence $\Delta'_n = \Delta_n$.

COROLLARIES OF THEOREM 7. There are many extensions of Theorem 7. For example, by repeated applications of Theorem 7 it can be proved that

We may add to each column (or row) of a determinant fixed multiples of the SUBSEQUENT *columns (or rows) and leave the value of the determinant unaltered; in particular,*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+\lambda b_1+\mu c_1 & b_1+\nu c_1 & c_1 \\ a_2+\lambda b_2+\mu c_2 & b_2+\nu c_2 & c_2 \\ a_3+\lambda b_3+\mu c_3 & b_3+\nu c_3 & c_3 \end{vmatrix}.$$

There is a similar corollary with PRECEDING *instead of* SUBSEQUENT.

Another extension of Theorem 7 is

We may add multiples of any ONE *column (or row) to every other column (or row) and leave the value of the determinant unaltered; in particular,*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+\lambda b_1 & b_1 & c_1+\mu b_1 \\ a_2+\lambda b_2 & b_2 & c_2+\mu b_2 \\ a_3+\lambda b_3 & b_3 & c_3+\mu b_3 \end{vmatrix}.$$

There are many others. But experience rather than set rule is the better guide for further extensions. The practice of adding multiples of columns or rows at random is liable to lead to error unless each step is checked for validity by appeal to Theorem 6. For example,

$$\Delta = \begin{vmatrix} a_1+\lambda b_1 & b_1+\mu c_1 & c_1+\nu a_1 \\ a_2+\lambda b_2 & b_2+\mu c_2 & c_2+\nu a_2 \\ a_3+\lambda b_3 & b_3+\mu c_3 & c_3+\nu a_3 \end{vmatrix}$$

is, by Theorem 6,

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \lambda b_1 & \mu c_1 & \nu a_1 \\ \lambda b_2 & \mu c_2 & \nu a_2 \\ \lambda b_3 & \mu c_3 & \nu a_3 \end{vmatrix};$$

all the other determinants envisaged by Theorem 6, such as

$$\begin{vmatrix} \lambda b_1 & b_1 & c_1 \\ \lambda b_2 & b_2 & c_2 \\ \lambda b_3 & b_3 & c_3 \end{vmatrix},$$

being zero in virtue of Theorems 4 and 5. Hence the determinant Δ is equal to

$$(1+\lambda\mu\nu)\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

NOTE. One of the more curious errors into which one is led by adding multiples of rows at random is a fallacious proof that $\Delta = 0$. In the example just given, 'subtract second column from first, add second and third, add first and third', corresponds to $\lambda = -1, \mu = 1, \nu = 1$, a choice of values that will (wrongly, of course) 'prove' that $\Delta = 0$: the manipulation of the columns has not left the value unaltered, but multiplied the determinant by a zero factor, $1+\lambda\mu\nu$.

3.5. Applications of Theorems 1–7.

As a convenient notation for the application of Theorem 7 and its corollaries, we shall use

$$r_k' = lr_1+mr_2+...+tr_n$$

to denote that, starting from a determinant Δ_n, we form a new determinant Δ_n' whose kth row is obtained by taking l times the first row plus m times the second row plus ... plus t times the nth row of Δ_n. The notation

$$c_n' = \lambda c_1 + \mu c_2 + ... + \kappa c_n$$

refers to a similar process applied to columns.

EXAMPLES I

1. Find the value of

$$\Delta = \begin{vmatrix} 87 & 42 & 3 & -1 \\ 45 & 18 & 7 & 4 \\ 50 & 17 & 3 & -5 \\ 91 & 9 & 6 & 0 \end{vmatrix}.$$

The working that follows is an illustration of how one may deal with an isolated numerical determinant. For a systematic method of computing determinants (one which reduces the order of the determinant from, say, 6 to 5, from 5 to 4, from 4 to 3, from 3 to 2) the reader is referred to Whittaker and Robinson, *The Calculus of Observations* (London, 1926), chapter v.

On writing $c_1' = c_1 - 2c_2 - c_3$, $c_3' = c_3 + 3c_4$,

$$\Delta = \begin{vmatrix} 0 & 42 & 0 & -1 \\ 2 & 18 & 19 & 4 \\ 13 & 17 & -12 & -5 \\ 67 & 9 & 6 & 0 \end{vmatrix}.$$

$$= -42 \begin{vmatrix} 2 & 19 & 4 \\ 13 & -12 & -5 \\ 67 & 6 & 0 \end{vmatrix} + \begin{vmatrix} 2 & 18 & 19 \\ 13 & 17 & -12 \\ 67 & 9 & 6 \end{vmatrix},$$

by Theorem I. On expanding the third-order determinants,

$$\Delta = -42\{2(30) - 13(-24) + 67(-95+48)\} +$$
$$+ \{2(102+108) - 13(108-171) + 67(-216-323)\}, \quad \text{etc.}$$

2. Prove that†

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \beta\gamma & \gamma\alpha & \alpha\beta \end{vmatrix} = (\alpha-\beta)(\beta-\gamma)(\gamma-\alpha).$$

On taking $c_2' = c_2 - c_1$, $c_3' = c_3 - c_2$, the determinant becomes

$$\begin{vmatrix} 1 & 0 & 0 \\ \alpha & \beta-\alpha & \gamma-\beta \\ \beta\gamma & \gamma(\alpha-\beta) & \alpha(\beta-\gamma) \end{vmatrix},$$

---

† This determinant, and others that occur in this set of examples, can be evaluated quickly by using the Remainder Theorem. Here they are intended as exercises on §§ 1–3.

which, by Theorems 1 and 5, is equal to

$$(\alpha-\beta)(\beta-\gamma)\begin{vmatrix} -1 & -1 \\ \gamma & \alpha \end{vmatrix},$$

i.e. $$(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha).$$

3. Prove that

$$\Delta \equiv \begin{vmatrix} 0 & (\alpha-\beta)^2 & (\alpha-\gamma)^2 & (\alpha-\delta)^2 \\ (\beta-\alpha)^2 & 0 & (\beta-\gamma)^2 & (\beta-\delta)^2 \\ (\gamma-\alpha)^2 & (\gamma-\beta)^2 & 0 & (\gamma-\delta)^2 \\ (\delta-\alpha)^2 & (\delta-\beta)^2 & (\delta-\gamma)^2 & 0 \end{vmatrix} = 0.$$

When we have studied the multiplication of determinants we shall see (Examples IV, 2) that the given determinant is 'obviously' zero, being the product of two determinants each having a column of zeros. But, at present, we treat it as an exercise on Theorem 7. Take

$r_1' = r_1-r_2$ and remove the factor $\alpha-\beta$ from r_1' (Theorem 5),

$r_2' = r_2-r_3$ and remove the factor $\beta-\gamma$ from r_2',

$r_3' = r_3-r_4$ and remove the factor $\gamma-\delta$ from r_3'.

Then $\Delta/(\alpha-\beta)(\beta-\gamma)(\gamma-\delta)$ is equal to

$$\begin{vmatrix} -\alpha+\beta & \alpha-\beta & \alpha+\beta-2\gamma & \alpha+\beta-2\delta \\ \beta+\gamma-2\alpha & -\beta+\gamma & \beta-\gamma & \beta+\gamma-2\delta \\ \gamma+\delta-2\alpha & \gamma+\delta-2\beta & -\gamma+\delta & \gamma-\delta \\ (\delta-\alpha)^2 & (\delta-\beta)^2 & (\delta-\gamma)^2 & 0 \end{vmatrix}.$$

In this take $c_1' = c_1-c_2$, $c_2' = c_2-c_3$, $c_3' = c_3-c_4$ and remove the factors $\beta-\alpha$, $\gamma-\beta$, $\delta-\gamma$; it becomes

$$(\beta-\alpha)(\gamma-\beta)(\delta-\gamma)\begin{vmatrix} 2 & 2 & 2 & \alpha+\beta-2\delta \\ 2 & 2 & 2 & \beta+\gamma-2\delta \\ 2 & 2 & 2 & \gamma-\delta \\ 2\delta-\alpha-\beta & 2\delta-\beta-\gamma & \delta-\gamma & 0 \end{vmatrix}.$$

In the last determinant take $r_1' = r_1-r_2$, $r_2' = r_2-r_3$; it becomes

$$\begin{vmatrix} 0 & 0 & 0 & \alpha-\gamma \\ 0 & 0 & 0 & \beta-\delta \\ 2 & 2 & 2 & \gamma-\delta \\ 2\delta-\alpha-\beta & 2\delta-\beta-\gamma & \delta-\gamma & 0 \end{vmatrix},$$

which, on expanding by the first row, is zero since the determinant of order 3 that multiplies $\alpha-\gamma$ is one with a complete row of zeros.

4. Prove that

$$\begin{vmatrix} 0 & (\alpha-\beta)^3 & (\alpha-\gamma)^3 \\ (\beta-\alpha)^3 & 0 & (\beta-\gamma)^3 \\ (\gamma-\alpha)^3 & (\gamma-\beta)^3 & 0 \end{vmatrix} = 0.$$

✓5. Prove that

$$\begin{vmatrix} \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \\ \beta+\gamma & \gamma+\alpha & \alpha+\beta \end{vmatrix} = (\alpha+\beta+\gamma)(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta).$$

HINT. Take $r_3' = r_1+r_3$.

✓ 6. Prove that

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \beta+\gamma & \gamma+\delta & \delta+\alpha & \alpha+\beta \\ \delta & \alpha & \beta & \gamma \end{vmatrix} = 0.$$

✓ 7. Prove that

$$\begin{vmatrix} 0 & \alpha^2-\beta^2 & \alpha^2-\gamma^2 & \alpha^2-\delta^2 \\ \beta^2-\alpha^2 & 0 & \beta^2-\gamma^2 & \beta^2-\delta^2 \\ \gamma^2-\alpha^2 & \gamma^2-\beta^2 & 0 & \gamma^2-\delta^2 \\ \delta^2-\alpha^2 & \delta^2-\beta^2 & \delta^2-\gamma^2 & 0 \end{vmatrix} = 0.$$

HINT. Use Theorem 6.

✓ 8. Prove that

$$\begin{vmatrix} 2\alpha & \alpha+\beta & \alpha+\gamma & \alpha+\delta \\ \beta+\alpha & 2\beta & \beta+\gamma & \beta+\delta \\ \gamma+\alpha & \gamma+\beta & 2\gamma & \gamma+\delta \\ \delta+\alpha & \delta+\beta & \delta+\gamma & 2\delta \end{vmatrix} = 0.$$

✓ 9. Prove that

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \beta\gamma\delta & \gamma\delta\alpha & \delta\alpha\beta & \alpha\beta\gamma \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \end{vmatrix} = (\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta).$$

✓ 10. Prove that

$$\Delta \equiv \begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = (af-be+cd)^2.$$

The determinant may be evaluated directly if we consider it as

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix},$$

whose expansion is $\sum (-1)^N a_r b_s c_t d_u$, the value of N being determined by the rule of § 2.2. There are, however, a number of zero terms in the expansion of Δ. The *non-zero* terms are

a^2f^2 $[a_2 b_1 c_4 d_3$ and so is prefixed by $(-1)^2]$,

b^2e^2, c^2d^2, each prefixed by $+$, two terms $aebf$ [one $a_2 b_4 c_1 d_3$, the other $a_3 b_1 c_4 d_2$, and so each prefixed by $(-1)^3$],

two terms $adfc$ prefixed by $+$, and two terms $cdbe$ each prefixed by $-$.

Hence $$\Delta = a^2f^2+b^2e^2+c^2d^2-2aebf+2afcd-2becd.$$

✓11. Prove that the expansion of

$$\begin{vmatrix} 0 & c & b & d \\ c & 0 & a & e \\ b & a & 0 & f \\ d & e & f & 0 \end{vmatrix}$$

is $a^2d^2+b^2e^2+c^2f^2-2bcef-2cafd-2abde$.

✓12. Prove that the expansion of

$$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & z^2 & y^2 \\ 1 & z^2 & 0 & x^2 \\ 1 & y^2 & x^2 & 0 \end{vmatrix}$$

is $\sum x^4-2\sum y^2z^2$.

✓13. (*Harder.*) Express the determinant

$$\begin{vmatrix} x+a & b & c & d \\ b & x+c & d & a \\ c & d & x+a & b \\ d & a & b & x+c \end{vmatrix}$$

as a product of factors.

HINT. Two linear factors $x+a+c\pm(b+d)$ and one quadratic factor.

✓14. From a determinant Δ, of order 4, a new determinant Δ' is formed by taking

$$c_1' = c_1+\lambda c_2, \quad c_2' = c_2+\mu c_3, \quad c_3' = c_3+\nu c_1, \quad c_4' = c_4.$$

Prove that $\Delta' = (1+\lambda\mu\nu)\Delta$.

4. Notation for determinants

The determinant Δ_n of § 2.1 is sufficiently indicated by its leading diagonal and it is often written as $(a_1 b_2 c_3 \ldots k_n)$.

The use of the double suffix notation, which we used on p. 3 of the Preliminary Note, enables one to abbreviate still further. The determinant that has a_{rs} as the element in the rth row and the sth column may be written as $|a_{rs}|$.

Sometimes a determinant is sufficiently indicated by its first row; thus $(a_1 b_2 c_3 \dots k_n)$ may be indicated by $|a_1 b_1 c_1 \dots k_1|$, but the notation is liable to misinterpretation.

5. Standard types of determinant

5.1. The product of differences: alternants.

THEOREM 8. *The determinant*

$$\Delta_4 \equiv \begin{vmatrix} \alpha^3 & \beta^3 & \gamma^3 & \delta^3 \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \\ \alpha & \beta & \gamma & \delta \\ 1 & 1 & 1 & 1 \end{vmatrix}$$

is equal to $(\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta)$; *the corresponding determinant of order* n, *namely* $(\alpha^{n-1}\beta^{n-2}\dots 1)$, *is equal to the product of the differences that can be formed from the letters* $\alpha, \beta, \dots, \kappa$ *appearing in the determinant, due regard being paid to alphabetical order in the factors. Such determinants are called* ALTERNANTS.

On expanding Δ_4, we see that it may be regarded

(*a*) as a homogeneous polynomial of degree 6 in the variables $\alpha, \beta, \gamma, \delta$, the coefficients being ± 1;

(*b*) as a non-homogeneous polynomial of highest degree 3 in α, the coefficients of the powers of α being functions of β, γ, δ.

The determinant vanishes when $\alpha = \beta$, since it then has two columns identical. Hence, by the Remainder Theorem applied to a polynomial in α, the determinant has a factor $\alpha-\beta$. By a similar argument, the difference of any two of α, β, γ, δ is a factor of Δ_4, so that

$$\Delta_4 = K(\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta). \tag{10}$$

Since Δ_4 is a homogeneous polynomial of degree 6 in $\alpha, \beta, \gamma, \delta$, the factor K must be independent of $\alpha, \beta, \gamma, \delta$ and so is a numerical constant. The coefficient of $\alpha^3\beta^2\gamma$ on the right-hand side of (10) is K, and in Δ_4 it is 1. Hence $K = 1$ and

$$\Delta_4 = (\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta).$$

The last product is conveniently written as $\zeta(\alpha, \beta, \gamma, \delta)$, and the corresponding product of differences of the n letters $\alpha, \beta, \dots, \kappa$ as $\zeta(\alpha, \beta, \dots, \kappa)$. This last product contains

$$(n-1)+(n-2)+\dots+1$$

factors and the degree of the determinant $(\alpha^{n-1}\beta^{n-2}\dots 1)$ is also $(n-1)+(n-2)+\dots+1$, so that the argument used for Δ_4 is readily extended to Δ_n.

5.2. The circulant.

THEOREM 9. *The determinant*

$$\Delta \equiv \begin{vmatrix} a_1 & a_2 & a_3 & . & . & a_n \\ a_n & a_1 & a_2 & . & . & a_{n-1} \\ a_{n-1} & a_n & a_1 & . & . & a_{n-2} \\ . & . & . & . & . & . \\ a_2 & a_3 & a_4 & . & . & a_1 \end{vmatrix} = \prod (a_1+a_2\omega+\dots+a_n\omega^{n-1}),$$

where the product is taken over the n-th roots of unity. Such a determinant is called a CIRCULANT.

Let ω be any one of the n numbers

$$\omega_k = \cos\frac{2k\pi}{n}+i\sin\frac{2k\pi}{n} \quad (k = 1, 2, \dots, n).$$

In Δ replace the first column c_1 by a new column c_1', where

$$c_1' = c_1+\omega c_2+\omega^2 c_3+\dots+\omega^{n-1}c_n.$$

This leaves the value of Δ unaltered.

The first column of the new determinant is, on using the fact that $\omega^n = 1$ and writing $a_1+a_2\omega+\dots+a_n\omega^{n-1} \equiv \alpha$,

$$a_1+a_2\omega+a_3\omega^2+\dots+a_n\omega^{n-1} = \alpha,$$

$$a_n\omega^n+a_1\omega+a_2\omega^2+\dots+a_{n-1}\omega^{n-1} = \omega\alpha,$$

$$a_{n-1}\omega^n+a_n\omega^{n+1}+a_1\omega^2+\dots+a_{n-2}\omega^{n-1} = \omega^2\alpha,$$

$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$

$$a_2\omega^n+a_3\omega^{n+1}+\dots+a_n\omega^{2n-2}+a_1\omega^{n-1} = \omega^{n-1}\alpha.$$

Hence the first column of the new determinant has a factor

$\alpha \equiv a_1+a_2\omega+\ldots+a_n\omega^{n-1}$, which is therefore (Theorem 5) a factor of Δ. This is true for $\omega = \omega_k$ $(k = 1, 2,\ldots, n)$, so that

$$\Delta = K\prod_{k=1}^{n}(a_1+a_2\omega_k+\ldots+a_n\omega_k^{n-1}). \tag{11}$$

Moreover, since Δ and $\prod$ are homogeneous of degree n in the variables $a_1, a_2,\ldots, a_n$, the factor K must be independent of these variables and so is a numerical constant. Comparing the coefficients of a_1^n on the two sides of (11), we see that $K = 1$.

6. Odd and even permutations

6.1. The n numbers $r, s,\ldots, \theta$ are said to be a PERMUTATION of 1, 2,..., n if they consist of 1, 2,..., n in some order. We can obtain the order $r, s,\ldots, \theta$ from the order 1, 2,..., n by suitable interchanges of pairs; but the set of interchanges leading from the one order to the other is not unique. For example, 4 3 2 6 1 5 becomes 1 2 3 4 5 6 after the interchanges denoted by

$$\begin{pmatrix}4&3&2&6&1&5\\1&3&2&6&4&5\end{pmatrix},\quad \begin{pmatrix}1&3&2&6&4&5\\1&2&3&6&4&5\end{pmatrix},\quad \begin{pmatrix}1&2&3&6&4&5\\1&2&3&4&6&5\end{pmatrix},\quad \begin{pmatrix}1&2&3&4&6&5\\1&2&3&4&5&6\end{pmatrix},$$

whereby we first put 1 in the first place, then 2 in the second place, and so on. But we may arrive at the same final result by first putting 6 in the sixth place, then 5 in the fifth, and so on: or we can proceed solely by adjacent interchanges, beginning by

$$\begin{pmatrix}4&3&2&6&1&5\\4&3&2&1&6&5\end{pmatrix},\quad \begin{pmatrix}4&3&2&1&6&5\\4&3&1&2&6&5\end{pmatrix},$$

as first steps towards moving 1 into the first place. In fact, as the reader will see for himself, there is a wide variety of sets of interchanges that will ultimately change 4 3 2 6 1 5 into 1 2 3 4 5 6.

Now suppose that there are K interchanges in any ONE way of going from $r, s,\ldots, \theta$ to 1, 2,..., n. Then, by condition III in the definition of a determinant (*vide* § 2.1), the term $a_r b_s \ldots k_\theta$ in the expansion of $\Delta_n \equiv (a_1 b_2 \ldots k_n)$ is prefixed by the sign $(-1)^K$. But, as we have proved, the definition of Δ_n by the conditions I, II, III is unique. Hence the sign to be prefixed to any given term is uniquely determined, and therefore the sign $(-1)^K$ must be the same whatever set of interchanges is used in going from $r, s,\ldots, \theta$ to 1, 2,..., n. Thus, if one way of going from $r, s,..., \theta$ to $1, 2,..., n$ involves an odd (even) number of interchanges, so does every way of going from $r, s,..., \theta$ to $1, 2,..., n$.

Accordingly, every permutation may be characterized either as EVEN, when the change from it to the standard order $1, 2,..., n$ can be effected by an even number of interchanges, or as ODD, when the change from it to the standard order $1, 2,..., n$ can be effected by an odd number of interchanges.

6.2. It can be proved, without reference to determinants, that, if one way of changing from $r, s,..., \theta$ to $1, 2,..., n$ involves an odd (even) number of interchanges, so does every way of effecting the same result. When this has been done it is legitimate to *define* $\Delta_n \equiv (a_1 b_2 ... k_n)$ as $\sum \pm a_r b_s ... k_\theta$, where the plus or minus sign is prefixed to a term according as its suffixes form an even or an odd permutation of $1, 2,..., n$. This is, in fact, one of the common ways of defining a determinant.

7. Differentiation

When the elements of a determinant are functions of a variable x, the rule for obtaining its differential coefficient is as follows. If Δ denotes the determinant

$$\begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix}$$

and the elements are functions of x, $d\Delta/dx$ is the *sum* of the n determinants obtained by differentiating the elements of *one* row (or column) of Δ and leaving the elements of the other $n-1$ rows (or columns) unaltered. For example,

$$\frac{d}{dx}\begin{vmatrix} x^2 & x & 1 \\ x^3 & x^2 & x \\ x^4 & x^3 & x^2 \end{vmatrix}$$

$$= \begin{vmatrix} 2x & 1 & 0 \\ x^3 & x^2 & x \\ x^4 & x^3 & x^2 \end{vmatrix} + \begin{vmatrix} x^2 & x & 1 \\ 3x^2 & 2x & 1 \\ x^4 & x^3 & x^2 \end{vmatrix} + \begin{vmatrix} x^2 & x & 1 \\ x^3 & x^2 & x \\ 4x^3 & 3x^2 & 2x \end{vmatrix}.$$

The proof of the rule follows at once from Theorem 1; since

$$\Delta = \sum (-1)^N a_r b_s \dots k_\theta,$$

the usual rules for differentiating a product give

$$\frac{d\Delta}{dx} = \sum (-1)^N \left\{ \frac{da_r}{dx} b_s \dots k_\theta + a_r \frac{db_s}{dx} \dots k_\theta + \dots + a_r b_s \dots j_t \frac{dk_\theta}{dx} \right\}$$

Since, for example,

$$\sum (-1)^N \frac{da_r}{dx} b_s \dots k_\theta = \begin{vmatrix} \frac{da_1}{dx} & . & . & \frac{da_n}{dx} \\ b_1 & . & . & b_n \\ . & . & . & . \\ k_1 & . & . & k_n \end{vmatrix},$$

$d\Delta/dx$ is the sum of n determinants, in each of which one row consists of differential coefficients of a row of Δ and the remaining $n-1$ rows consist of the corresponding rows of Δ.

EXAMPLES II

✓1. Prove that

$$\begin{vmatrix} 1 & \beta+\gamma & \beta^2+\gamma^2 \\ 1 & \gamma+\alpha & \gamma^2+\alpha^2 \\ 1 & \alpha+\beta & \alpha^2+\beta^2 \end{vmatrix} = \begin{vmatrix} 1 & \alpha & \alpha^2 \\ 1. & \beta & \beta^2 \\ 1 & \gamma & \gamma^2 \end{vmatrix} = -\zeta(\alpha,\beta,\gamma).$$

HINT. Take $c_2' = c_2 - s_1 c_1$, $c_3' = c_3 - s_2 c_1$, where $s_r = \alpha^r + \beta^r + \gamma^r$.

✓ 2. Prove that

$$\begin{vmatrix} 1 & \alpha & \beta\gamma \\ 1 & \beta & \gamma\alpha \\ 1 & \gamma & \alpha\beta \end{vmatrix} = -\zeta(\alpha,\beta,\gamma).$$

3. Prove that

$$\begin{vmatrix} 1 & \beta+\gamma & (\beta+\gamma)^2 \\ 1 & \gamma+\alpha & (\gamma+\alpha)^2 \\ 1 & \alpha+\beta & (\alpha+\beta)^2 \end{vmatrix} = \zeta(\alpha,\beta,\gamma).$$

✓ 4. Prove that each of the determinants of the fourth order whose first rows are (i) 1, $\beta+\gamma+\delta$, α^2, α^3; (ii) 1, α, $\beta^2+\gamma^2+\delta^2$, $\beta\gamma\delta$, is equal to $\pm\zeta(\alpha,\beta,\gamma,\delta)$. Write down other determinants that equal $\pm\zeta(\alpha,\beta,\gamma,\delta)$.

✓ 5. Prove that the 'skew' circulant

$$\begin{vmatrix} a_1 & a_2 & a_3 & . & . & . & a_n \\ -a_n & a_1 & a_2 & . & . & . & a_{n-1} \\ -a_{n-1} & -a_n & a_1 & . & . & . & a_{n-2} \\ . & . & . & . & . & . & . \\ -a_2 & -a_3 & -a_4 & . & . & . & a_1 \end{vmatrix} = \prod (a_1 + a_2\omega + \dots + a_n\omega^{n-1}),$$

where ω runs through the n roots of -1.

6. Prove that the circulant of order $2n$ whose first row is

$$a_1 \quad a_2 \quad a_3 \quad \dots \quad a_{2n}$$

is the product of a circulant of order n (with first row a_1+a_{n+1}, $a_2+a_{n+2},\dots$) and a skew circulant of order n (with first row a_1-a_{n+1}, $a_2-a_{n+2},\dots$).

7. Prove that the circulant of the fourth order with a first row a_1, a_2, a_3, a_4 is equal to

$$\{(a_1+a_3)^2-(a_2+a_4)^2\}\{(a_1-a_3)^2+(a_2-a_4)^2\}.$$

8. When $Q(x) \equiv (x-a_1)(x-a_2)\dots(x-a_n)$, prove that

$$\frac{1}{Q(x)} = \begin{vmatrix} 1 & \cdot & \cdot & \cdot & 1 \\ a_1 & \cdot & \cdot & \cdot & a_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^{n-2} & \cdot & \cdot & \cdot & a_n^{n-2} \\ (x-a_1)^{-1} & \cdot & \cdot & \cdot & (x-a_n)^{-1} \end{vmatrix} \div \begin{vmatrix} 1 & \cdot & \cdot & \cdot & 1 \\ a_1 & \cdot & \cdot & \cdot & a_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^{n-2} & \cdot & \cdot & \cdot & a_n^{n-2} \\ a_1^{n-1} & \cdot & \cdot & \cdot & a_n^{n-1} \end{vmatrix}$$

9. By putting $x = y^{-1}$ in Example 8 and expanding in powers of y, show that ${}_nH_p$, the sum of homogeneous products of a_1, $a_2,\dots$, a_n of degree p, is equal to

$$\begin{vmatrix} 1 & \cdot & \cdot & \cdot & 1 \\ a_1 & \cdot & \cdot & \cdot & a_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^{n-2} & \cdot & \cdot & \cdot & a_n^{n-2} \\ a_1^{n+p-1} & \cdot & \cdot & \cdot & a_n^{n+p-1} \end{vmatrix} \div \begin{vmatrix} 1 & \cdot & \cdot & \cdot & 1 \\ a_1 & \cdot & \cdot & \cdot & a_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^{n-2} & \cdot & \cdot & \cdot & a_n^{n-2} \\ a_1^{n-1} & \cdot & \cdot & \cdot & a_n^{n-1} \end{vmatrix}.$$

10. Prove that

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^3 & \beta^3 & \gamma^3 \end{vmatrix} = (\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)(\alpha+\beta+\gamma).$$

HINT. The first three factors are obtained as in Theorem 8. The degree of the determinant in α, β, γ is *four*; so the remaining factor must be linear in α, β, γ and it must be unaltered by the interchange of any two letters (for both the determinant and the product of the first three factors are altered in sign by such an interchange).

Alternatively, consider the coefficient of δ^2 in the alternant $\zeta(\alpha,\beta,\gamma,\delta)$.

11. Prove that

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha^2 & \beta^2 & \gamma^2 \\ \alpha^3 & \beta^3 & \gamma^3 \end{vmatrix} = (\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)(\beta\gamma+\gamma\alpha+\alpha\beta).$$

12. Extend the results of Examples 10 and 11 to determinants of higher order.

CHAPTER II

THE MINORS OF A DETERMINANT

1. First minors

1.1. In the determinant of order n,

$$\Delta_n = \begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix},$$

the determinant of order $n-1$ obtained by deleting the row and the column containing a_r is called the minor of a_r; and so for other letters. Such minors are called FIRST MINORS; the determinant of order $n-2$ obtained by deleting the two rows with suffixes r, s and the two columns with letters a, b is called a SECOND MINOR; and so on for third, fourth,... minors. We shall denote the first minors by α_r, β_r,... .

We can expand Δ_n by any row or column (Chap. I, § 3.11); for example, on expanding by the first column,

$$\Delta_n = a_1\alpha_1 - a_2\alpha_2 + ... + (-1)^{n-1}a_n\alpha_n, \qquad (1)$$

or, on expanding by the second row,

$$\Delta_n = -a_2\alpha_2 + b_2\beta_2 - ... + (-1)^n k_2\kappa_2. \qquad (2)$$

1.2. The above notation requires a careful consideration of sign in its use and it is more convenient to introduce CO-FACTORS. They are defined as the numbers

$$A_r,\ B_r,\ ...,\ K_r \quad (r = 1, 2,..., n)$$

such that

$$\begin{aligned} \Delta_n &= a_1 A_1 + a_2 A_2 + ... + a_n A_n \\ &= b_1 B_1 + b_2 B_2 + ... + b_n B_n \\ & . \quad . \quad . \quad . \quad . \quad . \quad . \\ &= k_1 K_1 + k_2 K_2 + ... + k_n K_n, \end{aligned}$$

these being the expansions of Δ_n by its various columns, and

$$\begin{aligned} \Delta_n &= a_1 A_1 + b_1 B_1 + ... + k_1 K_1 \\ & . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ &= a_n A_n + b_n B_n + ... + k_n K_n, \end{aligned}$$

these being the expansions of Δ_n by its various rows.

It follows from the definition that A_r, B_r,... are obtained by prefixing a suitable sign to α_r, β_r,... .

1.3. It is a simple matter to determine the sign for any particular minor. For example, to determine whether C_2 is $+\gamma_2$ or $-\gamma_2$ we observe that, on interchanging the first two rows of Δ_n,

$$\Delta_n = -\begin{vmatrix} a_2 & b_2 & . & . & k_2 \\ a_1 & b_1 & . & . & k_1 \\ a_3 & b_3 & . & . & k_3 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix}$$
$$= -a_2\alpha_2+b_2\beta_2-c_2\gamma_2+...,$$

while, by the definition of the co-factors

$$\Delta_n = a_2 A_2+b_2 B_2+c_2 C_2+...;$$

so that $C_2 = -\gamma_2$.

Or again, to determine whether D_3 is $+\delta_3$ or $-\delta_3$ we observe that Δ_n is unaltered (Theorem 3, Corollary) if we move the third row up until it becomes the first row, so that, on expanding by the first row of the determinant so formed,

$$\Delta_n = a_3\alpha_3-b_3\beta_3+c_3\gamma_3-d_3\delta_3+... .$$

But $$\Delta_n = a_3 A_3+b_3 B_3+c_3 C_3+d_3 D_3+...,$$

and so $D_3 = -\delta_3$.

1.4. If we use the double suffix notation (Chap. I, § 4), the co-factor A_{rs} of a_{rs} in $|a_{rs}|$ is, by the procedure of § 1.3, $(-1)^{r+s-2}$ times the minor of a_{rs}; that is, A_{rs} is $(-1)^{r+s}$ times the determinant obtained by deleting the rth row and sth column.

2. We have seen in § 1 that, with $\Delta_n \equiv (a_1\, b_2\, ...\, k_n)$,

$$\Delta_n = a_r A_r+b_r B_r+...+k_r K_r. \tag{3}$$

If we replace the rth row of Δ_n, namely

$$a_r \quad b_r \quad ... \quad k_r,$$

by $$a_s \quad b_s \quad ... \quad k_s,$$

where s is one of the numbers 1, 2,..., n other than r, we thereby get a determinant having *two* rows $a_s\, b_s\, ...\, k_s$; that is, we get a determinant equal to zero. But A_r, B_r,..., K_r are unaffected by such a change in the rth row of Δ_n. Hence, for the new determinant, (3) takes the form

$$0 = a_s A_r + b_s B_r + \ldots + k_s K_r. \tag{4}$$

A corresponding result for columns may be proved in the same way, and the results summed up thus:

The sum of elements of a row (or column) multiplied by their own co-factors is Δ; the sum of elements of a row multiplied by the corresponding co-factors of another row is zero; the sum of elements of a column multiplied by the corresponding co-factors of another column is zero.

We also state these important facts as a theorem.

THEOREM 10. *The determinant* $\Delta \equiv (a_1\, b_2 \ldots k_n)$ *may be expanded by any row or by any column: such expansions take one of the forms*

$$\Delta = a_r A_r + b_r B_r + \ldots + k_r K_r, \tag{5}$$

$$\Delta = x_1 X_1 + x_2 X_2 + \ldots + x_n X_n, \tag{6}$$

where r is any one of the numbers 1, 2,..., n *and x is any one of the letters a, b,..., k. Moreover,*

$$0 = a_s A_r + b_s B_r + \ldots + k_s K_r, \tag{7}$$

$$0 = y_1 X_1 + y_2 X_2 + \ldots + y_n X_n, \tag{8}$$

where r, s are two different numbers taken from 1, 2,..., n *and x, y are two different letters taken from a, b,..., k.*

3. Preface to §§ 4–6

We come now to a group of problems that depend for their full discussion on the implications of 'rank' in a matrix. This full discussion is deferred to Chapter VIII. But even the more elementary aspects of these problems are of considerable importance and these are set out in §§ 4–6.

4. The solution of non-homogeneous linear equations

4.1. The equations

$$a_1 x + b_1 y + c_1 z = 0, \qquad a_2 x + b_2 y + c_2 z = 0$$

are said to be homogeneous linear equations in x, y, z; the equations $a_1 x + b_1 y = -c_1, \qquad a_2 x + b_2 y = -c_2$

are said to be non-homogeneous linear equations in x, y.

4.2. If it is possible to choose a set of values for $x, y,..., t$ so that the m equations

$$a_r x+b_r y+...+k_r t = l_r \quad (r = 1,2,...,m)$$

are all satisfied, these equations are said to be CONSISTENT. If it is not possible so to choose the values of $x, y,..., t$, the equations are said to be INCONSISTENT. For example, $x+y = 2$, $x-y = 0$, $3x-2y = 1$ are consistent, since all three equations are satisfied when $x = 1$, $y = 1$; on the other hand, $x+y = 2$, $x-y = 0$, $3x-2y = 6$ are inconsistent.

4.3. Consider the n non-homogeneous linear equations, in the n variables $x, y,..., t$,

$$a_r x+b_r y+...+k_r t = l_r \quad (r = 1,2,...,n). \tag{9}$$

Let $\Delta \equiv (a_1 b_2 ... k_n)$ and let A_r, B_r,... be the co-factors of a_r, b_r,... in Δ.

Since, by Theorem 10, $\sum a_r A_r = \Delta$, $\sum b_r A_r = 0$,..., the result of multiplying each equation (9) by its corresponding A_r and adding is

$$\Delta x+0.y+...+0.t = l_1 A_1+...+l_n A_n;$$

that is,
$$\Delta x = (l_1 b_2 c_3 ... k_n), \tag{10}$$

the determinant obtained by writing l for a in Δ.

Similarly, the result of multiplying each equation (9) by its corresponding B_r and adding is

$$\Delta y = l_1 B_1+...+l_n B_n = (a_1 l_2 c_3 ... k_n), \tag{11}$$

and so on.

When $\Delta \neq 0$ the equations (9) have a unique solution given by (10), (11), and their analogues. In words, the solution is '$\Delta.x$ *is equal to the determinant obtained by putting* l *for* a *in* Δ; $\Delta.y$ *is equal to the determinant obtained by putting* l *for* b *in* Δ; *and so on*'.†

When $\Delta = 0$ the non-homogeneous equations (9) are inconsistent unless each of the determinants on the right-hand sides of (10), (11), and their analogues is also zero. When all such determinants are zero the equations (9) may or may not be consistent: we defer consideration of the problem to a later chapter.

---

† Some readers may already be familiar with different forms of setting out this result—the differences of form are unimportant.

5. The solution of homogeneous linear equations

THEOREM 11. *A necessary and sufficient condition that values, not all zero, may be assigned to the n variables x, y,..., t so that the n homogeneous equations*

$$a_r x+b_r y+...+k_r t = 0 \quad (r = 1,2,...,n) \tag{12}$$

hold simultaneously is $(a_1\, b_2\, ...\, k_n) = 0$.

5.1. NECESSARY. Let equations (12) be satisfied by values of $x, y,..., t$ not all zero. Let $\Delta \equiv (a_1\, b_2\, ...\, k_n)$ and let A_r, B_r,... be the co-factors of a_r, b_r,... in Δ. Multiply each equation (12) by its corresponding A_r and add; the result is, as in § 4.3, $\Delta x = 0$. Similarly, $\Delta y = 0$, $\Delta z = 0$,..., $\Delta t = 0$. But, by hypothesis, at least one of x, y,..., t is not zero and therefore Δ must be zero.

5.2. SUFFICIENT. Let $\Delta = 0$.

5.21. In the first place suppose, further, that $A_1 \neq 0$. Omit $r = 1$ from (12) and consider the $n-1$ equations

$$b_r y+c_r z+...+k_r t = -a_r x \quad (r = 2,3,...,n), \tag{13}$$

where the determinant of the coefficients on the left is A_1. Then, proceeding exactly as in § 4.3, but with the determinant A_1 in place of Δ, we obtain

$$A_1 y = -(a_2\, c_3\, ...\, k_n)x = B_1 x,$$
$$A_1 z = -(b_2\, a_3\, d_4\, ...\, k_n)x = C_1 x,$$

and so on. Hence the set of values

$$x = A_1\xi, \quad y = B_1\xi, \quad z = C_1\xi, \quad ...,$$

where $\xi \neq 0$, is a set, not all zero (since $A_1\xi \neq 0$), satisfying the equations (13). But

$$a_1 A_1+b_1 B_1+... = \Delta = 0,$$

and hence this set of values also satisfies the omitted equation corresponding to $r = 1$. This proves that $\Delta = 0$ is a sufficient condition for our result to hold provided also that $A_1 \neq 0$.

If $A_1 = 0$ and some other first minor, say C_s, is not zero, an interchange of the letters a and c, of the letters x and z, and of the suffixes 1 and s will give the equations (12) in a slightly changed notation and, in this new notation, A_1 (the C_s of the old notation) is not zero. It follows that if $\Delta = 0$ and if *any* one first minor of Δ is not zero, values, not all zero, may be assigned to $x, y, \ldots, t$ so that equations (12) are all satisfied.

5.22. Suppose now that $\Delta = 0$ and that every first minor of Δ is zero; or, proceeding to the general case, suppose that all minors with more than R rows and columns vanish, but that at least one minor with R rows and columns does not vanish. Change the notation (by interchanging letters and interchanging suffixes) so that one non-vanishing minor of R rows is the first minor of a_1 in the determinant $(a_1\, b_2\, \ldots\, e_{R+1})$, where e denotes the $(R+1)$th letter of the alphabet.

Consider, instead of (12), the $R+1$ equations

$$a_r x + b_r y + \ldots + e_r \lambda = 0 \quad (r = 1, 2, \ldots, R+1), \tag{12'}$$

where λ denotes the $(R+1)$th variable of the set $x, y, \ldots$. The determinant $\Delta' \equiv (a_1\, b_2\, \ldots\, e_{R+1}) = 0$, by hypothesis, while the minor of a_1 in Δ' is not zero, also by hypothesis. Hence, by § 5.21, the equations (12′) are satisfied when

$$x = A_1', \quad y = B_1', \quad \ldots, \quad \lambda = E_1', \tag{14}$$

where $A_1', B_1', \ldots$ are the minors of $a_1, b_1, \ldots$ in Δ'. Moreover, $A_1' \neq 0$.

Further, if $R+1 < r \leqslant n$,

$$a_r A_1' + b_r B_1' + \ldots + e_r E_1',$$

being the determinant formed by putting $a_r, b_r, \ldots$ for $a_1, b_1, \ldots$ in Δ', is a determinant of order $R+1$ formed from the coefficients of the equations (12); as such its value is, by our hypothesis, zero. Hence the values (14) satisfy not only (12′), but also

$$a_r x + b_r y + \ldots + e_r \lambda = 0 \quad (R+1 < r \leqslant n).$$

Hence the n equations (12) are satisfied if we put the values (14) for $x, y, \ldots, \lambda$ and the value zero for all variables in (12) other than these. Moreover, the value of x in (14) is not zero.

6. The minors of a zero determinant

THEOREM 12. *If* $\Delta = 0$,

$$A_r B_s = A_s B_r, \quad A_r C_s = A_s C_r, \quad \ldots \quad \begin{pmatrix} r = 1, 2, \ldots, n \\ s = 1, 2, \ldots, n \end{pmatrix}.$$

If $\Delta = 0$ *and if, further, no first minor of* Δ *is zero, the cofactors of the r-th row (column) are proportional to those of the s-th row (column); that is,*

$$\frac{A_r}{A_s} = \frac{B_r}{B_s} = \dots = \frac{K_r}{K_s}.$$

Consider the n equations

$$a_r x + b_r y + \dots + k_r t = 0 \quad (r = 1, 2, \dots, n). \tag{15}$$

They are satisfied by $x = A_1$, $y = B_1, \dots,$ $t = K_1$; for, by Theorem 10,

$$a_1 A_1 + b_1 B_1 + \dots + k_1 K_1 = \Delta = 0,$$

$$a_r A_1 + b_r B_1 + \dots + k_r K_1 = 0 \quad (r = 2, 3, \dots, n).$$

Now let s be any one of the numbers $2, 3, \dots, n$, and consider the $n-1$ equations

$$b_r y + c_r z + \dots + k_r t = -a_r x \quad (r \neq s).$$

Proceeding as in § 4.3, but with the determinant A_s in place of Δ, we obtain

$$(b_1 c_2 \dots k_n) y = -(a_1 c_2 \dots k_n) x,$$

$$(b_1 c_2 \dots k_n) z = -(b_1 a_2 d_3 \dots k_n) x,$$

and so on, there being no suffix s. That is, we have

$$A_s y = B_s x, \quad A_s z = C_s x, \quad \dots.$$

These equations hold whenever $x, y, \dots, t$ satisfy the equations (15), and therefore hold when

$$x = A_1, \quad y = B_1, \quad \dots, \quad t = K_1.$$

Hence $$A_s B_1 = A_1 B_s, \quad A_s C_1 = A_1 C_s, \quad \dots. \tag{16}$$

The same method of proof holds when we take

$$x = A_r, \quad y = B_r, \quad \dots, \quad t = K_r,$$

with $r \neq 1$, as a solution of (15). Hence our general theorem is proved.

If no first minor is zero, we may divide the equation $A_r B_s = A_s B_r$ *by* $A_s B_s$, etc., and so obtain

$$\frac{A_r}{A_s} = \frac{B_r}{B_s} = \dots = \frac{K_r}{K_s}. \tag{17}$$

Working with columns instead of rows, we have

$$\frac{A_1}{B_1} = \frac{A_2}{B_2} = \dots = \frac{A_n}{B_n},$$

and so for other columns.

EXAMPLES III

✓1. By solving the equations

$$a^r x + b^r y + c^r z + d^r t = \alpha^r \quad (r = 0, 1, 2, 3),$$

prove that

$$\sum_a \frac{(\alpha-b)(\alpha-c)(\alpha-d)}{(a-b)(a-c)(a-d)} a^r = \alpha^r \quad (r = 0, 1, 2, 3).$$

Obtain identities in a, b, c, d by considering the coefficients of powers of α.

✓2. If Δ denotes the determinant

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$$

and G, F, C are the co-factors of g, f, c in Δ, prove that

$$aG^2 + 2hFG + bF^2 + 2gGC + 2fFC + cC^2 = C\Delta.$$

HINT. Use $aG + hF + gC = 0$, etc.

3. If the determinant Δ of Example 2 is equal to zero, prove that $BC = F^2$, $GH = AF$,..., and hence that, when $C \neq 0$,

$$a(ax^2 + 2hxy + by^2 + 2gx + 2fy + c) = (ax+hy+g)^2 + \frac{1}{C}(Cy - F)^2.$$

Prove also that, if a,..., h are real numbers and C is negative, then so also are A and B. [$A = bc - f^2$, $F = gh - af$, etc.]

✓ 4. Prove that the three lines whose cartesian equations are

$$a_r x + b_r y + c_r = 0 \quad (r = 1, 2, 3)$$

are concurrent or parallel if $(a_1 b_2 c_3) = 0$.

HINT. Make equations homogeneous and use Theorem 11.

5. Find the conditions that the four planes whose cartesian equations are $a_r x + b_r y + c_r z + d_r = 0$ $(r = 1, 2, 3, 4)$ should have a finite point in common.

6. Prove that the equation of the circle cutting the three given circles

$$x^2 + y^2 + 2g_r x + 2f_r y + c_r = 0 \quad (r = 1, 2, 3)$$

orthogonally is

$$\begin{vmatrix} x^2+y^2 & -x & -y & 1 \\ c_1 & g_1 & f_1 & 1 \\ c_2 & g_2 & f_2 & 1 \\ c_3 & g_3 & f_3 & 1 \end{vmatrix} = 0.$$

By writing the last equation as

$$\begin{vmatrix} c_1+g_1x+f_1y & g_1+x & f_1+y \\ c_2+g_2x+f_2y & g_2+x & f_2+y \\ c_3+g_3x+f_3y & g_3+x & f_3+y \end{vmatrix} = 0,$$

prove that the circle in question is the locus of the point whose polars with respect to the three given circles are concurrent.

7. Express in determinantal form the condition that four given circles may have a common orthogonal circle.

8. Write down

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & x \\ \alpha^2 & \beta^2 & \gamma^2 & x^2 \\ \alpha^3 & \beta^3 & \gamma^3 & x^3 \end{vmatrix},$$

as a product of factors, and, by considering the minors of x, x^2 in the determinant and the coefficients of x, x^2 in the product of the factors, evaluate the determinants of Examples 10, 11 on p. 27.

9. Extend the results of Example 8 to determinants of higher orders.

7. Laplace's expansion of a determinant

7.1. The determinant

$$\Delta_n = \begin{vmatrix} a_1 & b_1 & c_1 & . & . & k_1 \\ a_2 & b_2 & c_2 & . & . & k_2 \\ . & . & . & . & . & . \\ a_n & b_n & c_n & . & . & k_n \end{vmatrix}$$

can be expressed in the form

$$\sum (-1)^N a_r b_s c_t \ldots k_\theta, \tag{18}$$

where the sum is taken over the $n!$ ways of assigning to $r, s, \ldots, \theta$ the values $1, 2, \ldots, n$ in some order and N is the total number of inversions in the suffixes $r, s, \ldots, \theta$.

We now show that the expansions of Δ_n by its rows and columns are but special cases of a more general procedure. The terms in (18) that contain $a_p b_q$, when p and q are *fixed*, constitute a sum

$$a_p b_q (AB)_{pq} \equiv a_p b_q \sum \pm c_t \ldots k_\theta \tag{19}$$

in which the sum is taken over the $(n-2)!$ ways of assigning to $t, \ldots, \theta$ the values $1, 2, \ldots, n$, in some order, excluding p and q. Also, an interchange of any two letters throughout (19) will reproduce the same set of terms, but with opposite signs prefixed, since this is true of (18). Hence, by the definition of a determinant, either $+(AB)_{pq}$ or $-(AB)_{pq}$ is equal to the determinant of order $n-2$ obtained by deleting from Δ_n the first and second columns and also the pth and qth rows. Denote this determinant by its leading diagonal, say $(c_t d_u \dots k_\theta)$, where $t, u, \dots, \theta$ are $1, 2, \dots, n$ *in that order* but excluding p and q.

Again, since (18) contains a set of terms $a_p b_q (AB)_{pq}$ and an interchange of a and b leaves $(AB)_{pq}$ unaltered, it follows (from the definition of a determinant) that (18) also contains a set of terms $-a_q b_p (AB)_{pq}$. Thus (18) contains a set of terms

$$(a_p b_q - a_q b_p)(AB)_{pq}.$$

But we have seen that $(AB)_{pq} = \pm(c_t d_u \dots k_\theta)$, and the determinant $a_p b_q - a_q b_p$ may be denoted by $(a_p b_q)$, so that a typical set of terms in (18) is

$$\pm(a_p b_q)(c_t d_u \dots k_\theta).$$

Moreover, *all* the terms of (18) are accounted for if we take all possible pairs of numbers p, q from $1, 2, \dots, n$. Hence

$$\Delta_n = \sum \pm(a_p b_q)(c_t d_u \dots k_\theta),$$

where the sum is taken over all possible pairs p, q.

In this form the fixing of the sign is a simple matter, for the leading diagonal term $a_p b_q$ of the determinant $(a_p b_q)$ is prefixed by plus, as is the leading diagonal term $c_t d_u \dots k_\theta$ of the determinant $(c_t d_u \dots k_\theta)$. Hence, on comparison with (18),

$$\Delta_n = \sum (-1)^N (a_p b_q)(c_t d_u \dots k_\theta), \tag{20}$$

where

(i) the sum is taken over all possible pairs p, q,
(ii) $t, u, \dots, \theta$ are $1, 2, \dots, n$ in that order but excluding p and q,
(iii) N is the total number of inversions of suffixes in $p, q, t, u, \dots, \theta$.

The sum (20) is called the expansion of Δ_n by its first two columns.

7.2. When once the argument of § 7.1 has been grasped it is intuitive that the argument extends to expansions such as

$$\Delta_n = \sum (-1)^N (a_p b_q c_r)(d_\lambda \dots k_\theta), \tag{21}$$

$$= \sum (-1)^N (a_p b_q c_r d_s)(e_\mu \dots k_\theta), \tag{22}$$

and so on. We leave the elaboration to the reader.

The expansion (20) is called an expansion of Δ_n by second minors, (21) an expansion by third minors, (22) an expansion by fourth minors, and so on.

Various modifications of § 7.1 are often used. In § 7.1 we expanded Δ_n by its first two columns: we may, in fact, expand it by any two (or more) columns or rows. For example, Δ_n is equal to

$$\begin{vmatrix} c_1 & d_1 & a_1 & b_1 & . & . & k_1 \\ c_2 & d_2 & a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . & . & . \\ c_n & d_n & a_n & b_n & . & . & k_n \end{vmatrix},$$

as we see by interchanging a and c and also b and d. The Laplace expansion of the latter determinant by its first two columns may be considered as the Laplace expansion of Δ_n by its third and fourth columns.

7.3. Determination of sign in a Laplace expansion

The procedure of p. 29, § 1.4, enables us to calculate with ease the sign appropriate to a given term of a Laplace expansion. Consider the determinant $\Delta \equiv |a_{rs}|$, that is, the determinant having a_{rs} as the element in its rth row and sth column, and its Laplace expansion by the s_1th and s_2th columns ($s_1 < s_2$). This expansion is of the form

$$\sum \pm \begin{vmatrix} a_{r_1 s_1} & a_{r_1 s_2} \\ a_{r_2 s_1} & a_{r_2 s_2} \end{vmatrix} \times \Delta(r_1, r_2; s_1, s_2), \tag{23}$$

where (i) $\Delta(r_1, r_2; s_1, s_2)$ denotes the determinant obtained by deleting the r_1th and r_2th rows as also the s_1th and s_2th columns of Δ, and (ii) the summation is taken over all possible pairs r_1, r_2 ($r_1 < r_2$).

Now, by § 1.4 (p. 29), the terms in the expansion of Δ that involve $a_{r_1 s_1}$ are given by

$$(-1)^{r_1+s_1} a_{r_1 s_1} \Delta(r_1; s_1),$$

where $\Delta(r_1; s_1)$ is the determinant obtained by deleting the r_1th row and s_1th column of Δ. Moreover, since $r_1 < r_2$ and $s_1 < s_2$, the element $a_{r_2 s_2}$ appears in the (r_2-1)th row and (s_2-1)th column of $\Delta(r_1;s_1)$; hence the terms in the expansion of $\Delta(r_1;s_1)$ that involve $a_{r_2 s_2}$ are given by

$$(-1)^{r_2+s_2-2}a_{r_2 s_2}\Delta(r_1,r_2;s_1,s_2),$$

for $\Delta(r_1,r_2;s_1,s_2)$ is obtained from $\Delta(r_1;s_1)$ by deleting the row and column containing $a_{r_2 s_2}$.

Thus the expansion of Δ contains the term

$$(-1)^{r_1+r_2+s_1+s_2}a_{r_1 s_1}a_{r_2 s_2}\Delta(r_1,r_2;s_1,s_2),$$

and therefore the Laplace expansion of Δ contains the term

$$(-1)^{r_1+r_2+s_1+s_2}\begin{vmatrix} a_{r_1 s_1} & a_{r_1 s_2} \\ a_{r_2 s_1} & a_{r_2 s_2} \end{vmatrix}\times\Delta(r_1,r_2;s_1,s_2). \tag{24}$$

That is to say, *the sign to be prefixed to a term in the Laplace expansion is* $(-1)^\sigma$, *where* σ *is the sum of the row and column numbers of the elements appearing in the first factor of the term.*

The rule, proved above for expansions by second minors, easily extends to Laplace expansions by third, fourth,..., minors.

7.4. As an exercise to ensure that the import of (20) has been grasped, the reader should check the following expansions of determinants.

(i)
$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}\times\begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}\times\begin{vmatrix} c_2 & d_2 \\ c_4 & d_4 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix}\times\begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix} +$$

$$+ \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}\times\begin{vmatrix} c_1 & d_1 \\ c_4 & d_4 \end{vmatrix} - \begin{vmatrix} a_2 & b_2 \\ a_4 & b_4 \end{vmatrix}\times\begin{vmatrix} c_1 & d_1 \\ c_3 & d_3 \end{vmatrix} + \begin{vmatrix} a_3 & b_3 \\ a_4 & b_4 \end{vmatrix}\times\begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix}.$$

The signs are, by (20), those of the leading diagonal term products

$$a_1 b_2 c_3 d_4,\quad a_1 b_3 c_2 d_4,\quad a_1 b_4 c_2 d_3,\quad \ldots.$$

Alternatively, the signs are, by (24),

$$(-1)^{1+2+1+2},\quad (-1)^{1+3+1+2},\quad (-1)^{1+4+1+2},\quad \ldots.$$

(ii)
$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 & e_1 \\ 0 & 0 & c_2 & d_2 & e_2 \\ 0 & 0 & c_3 & d_3 & e_3 \\ 0 & 0 & c_4 & d_4 & e_4 \\ a_5 & b_5 & c_5 & d_5 & e_5 \end{vmatrix} = -(a_1 b_5)(c_2 d_3 e_4).$$

The sign is most easily fixed by (24), which gives it to be $(-1)^{1+5+1+2}$.

(iii) $$0 = \begin{vmatrix} a_1 & b_1 & c_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 & b_4 & c_4 \\ 0 & 0 & 0 & b_4 & c_4 \end{vmatrix} = \Delta_1 A_2^3 + \Delta_2 A_1^3 - \Delta_3 A_1^2,$$

where $\Delta_1 \equiv (a_2 b_3 c_4)$, $\Delta_2 \equiv (a_3 b_4 c_1)$, $\Delta_3 \equiv (a_4 b_1 c_2)$, and A_r^s is the co-factor of a_r in Δ_s.

CHAPTER III

THE PRODUCT OF TWO DETERMINANTS

1. The summation convention

It is an established convention that a sum such as

$$\sum_{r=1}^{n} a_r b_r \tag{1}$$

may be denoted by the single term

$$a_r b_r. \tag{2}$$

The convention is that when a literal suffix is repeated in the single term (as here r is repeated) then the single term shall represent the sum of all the terms that correspond to different values of r. The convention is applicable only when, by the context, one knows the range of values of the suffix. In what follows we suppose that each suffix has the range of values $1, 2, \ldots, n$.

Further examples of the convention are

$$a_{rs} x_s, \quad \text{which denotes} \quad \sum_{s=1}^{n} a_{rs} x_s, \tag{3}$$

$$a_{rj} b_{js}, \quad \text{which denotes} \quad \sum_{j=1}^{n} a_{rj} b_{js}, \tag{4}$$

$$a_{rs} x_r x_s, \quad \text{which denotes} \quad \sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} x_r x_s. \tag{5}$$

In the last example both r and s are repeated and so we must sum with regard to both of them.

The repeated suffix is often called a 'DUMMY SUFFIX', a curious, but almost universal, term for a suffix whose presence implies the summation. The nomenclature is appropriate because the meaning of the symbol as a whole does not depend on what letter is used for the 'dummy'; for example, both

$$a_{rs} x_s \quad \text{and} \quad a_{rj} x_j$$

stand for the same thing, namely,

$$a_{r1} x_1 + a_{r2} x_2 + \ldots + a_{rn} x_n.$$

Any suffix that is not repeated is called a '**FREE SUFFIX**'. On most occasions when we are using the summation convention, we regard

$$a_{rs}x_s \tag{6}$$

not as a single expression in which r has a fixed value, but as a typical one of the n expressions

$$\sum_{s=1}^{n} a_{rs}x_s \quad (r = 1, 2,..., n).$$

Thus, in (6), r is to be thought of as a suffix free to take any one of the values 1, 2,..., n; and no further summation of all such expressions is implied.

In the next section we use the convention in considering linear transformations of a set of n linear equations.

2. Consider n variables x_i $(i = 1, 2,..., n)$ and the n linear forms†

$$a_{ri}x_i \quad (r = 1, 2,..., n). \tag{7}$$

When we substitute

$$x_i = b_{is}X_s \quad (i = 1, 2,..., n), \tag{8}$$

the forms (7) become

$$a_{ri}b_{is}X_s \quad (r = 1, 2,..., n). \tag{9}$$

Now consider the n equations

$$c_{rs}X_s = 0, \tag{10}$$

where

$$c_{rs} \equiv a_{ri}b_{is}. \tag{11}$$

If the determinant‡ $|c_{rs}| = 0$, then (10) is satisfied by a set of values X_i $(i = 1, 2,..., n)$ which are not all zero (Theorem 11). But, when the X_i satisfy equations (10), we have the n equations

$$a_{ri}x_i = a_{ri}b_{is}X_s = c_{rs}X_s = 0, \tag{12}$$

and so

EITHER all the x_i are zero,
OR the determinant $|a_{rs}|$ is zero (Theorem 11).

---

† The reader who is unfamiliar with the summation convention is recommended to write the next few lines in full and to pick out the coefficient of X_s in

$$\sum_{i=1}^{n} a_{ri}\left(\sum_{s=1}^{n} b_{is}X_s\right).$$

‡ Compare Chap. I, §4 (p. 21).

In the former case the n equations

$$b_{is} X_s = 0$$

are satisfied by a set of values X_i which are not all zero; that is, $|b_{rs}| = 0$.

Hence if $|c_{rs}| = 0$, at least one of the determinants $|a_{rs}|$, $|b_{rs}|$ is zero.

The determinant $|c_{rs}|$, i.e. $|a_{ri} b_{is}|$, is of degree n in the a's and of degree n in the b's, and so it is indicated that

$$|c_{rs}| = k|a_{rs}| \times |b_{rs}|, \tag{13}$$

where k is independent of the a's and b's.

By considering the particular case $a_{rs} = 0$ when $r \neq s$, $a_{rr} = 1$, it is indicated that

$$|c_{rs}| = |a_{rs}| \times |b_{rs}|. \tag{14}$$

The foregoing is an indication rather than a proof of the important theorem contained in (14). In the next section we give two proofs of this theorem. Of the two proofs, the second is perhaps the easier; it is certainly the more artificial.

3. Proofs of the rule for multiplying determinants

THEOREM 13. *Let $|a_{rs}|$, $|b_{rs}|$ be two determinants of order n; then their product is the determinant $|c_{rs}|$, where*

$$c_{rs} = \sum_{i=1}^{n} a_{ri} b_{is}.$$

3.1. *First proof*—a proof that uses the double suffix notation.

In this proof, we shall use only the Greek letter α as a dummy suffix implying summation: a repeated Roman letter will *not* imply summation. Thus

$a_{1\alpha} b_{\alpha 1}$ stands for $a_{11} b_{11} + a_{12} b_{21} + \dots$.

$a_{1r} b_{r1}$ stands for the single term; e.g. if $r = 2$, it stands for $a_{12} b_{21}$.

Consider the determinant $|c_{pq}|$ written in the form

$$\begin{vmatrix} a_{1\alpha} b_{\alpha 1} & a_{1\alpha} b_{\alpha 2} & \cdot & \cdot & a_{1\alpha} b_{\alpha n} \\ a_{2\alpha} b_{\alpha 1} & a_{2\alpha} b_{\alpha 2} & \cdot & \cdot & a_{2\alpha} b_{\alpha n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n\alpha} b_{\alpha 1} & a_{n\alpha} b_{\alpha 2} & \cdot & \cdot & a_{n\alpha} b_{\alpha n} \end{vmatrix},$$

and pick out the coefficient of

$$a_{1r}a_{2s}a_{3t}\dots a_{nz} \tag{15}$$

in the expansion of this determinant. This we can do by considering all a_{1x} other than a_{1r} to be zero, all a_{2x} other than a_{2s} to be zero, and so on. Doing this, we see that the terms of $|c_{pq}|$ that contain (15) as a factor are given by†

$$\begin{vmatrix} a_{1r}b_{r1} & a_{1r}b_{r2} & . & . & a_{1r}b_{rn} \\ a_{2s}b_{s1} & a_{2s}b_{s2} & . & . & a_{2s}b_{sn} \\ . & . & . & . & . \\ a_{nz}b_{z1} & a_{nz}b_{z2} & . & . & a_{nz}b_{zn} \end{vmatrix}$$

$$= a_{1r}a_{2s}\dots a_{nz}\begin{vmatrix} b_{r1} & b_{r2} & . & . & b_{rn} \\ b_{s1} & b_{s2} & . & . & b_{sn} \\ . & . & . & . & . \\ b_{z1} & b_{z2} & . & . & b_{zn} \end{vmatrix}.$$

Unless the numbers $r, s,\dots, z$ are all different, the 'b' determinant is zero, having two rows identical. When $r, s,\dots, z$ are the numbers 1, 2,..., n in some order, the 'b' determinant is equal to $(-1)^N|b_{pq}|$, where N is the total number of inversions of the suffixes in $r, s,\dots, z$ (p. 11, § 2.2, applied to rows). Hence

$$|c_{pq}| = \sum (-1)^N a_{1r}a_{2s}\dots a_{nz}|b_{pq}|, \tag{16}$$

where the summation is taken over the $n!$ ways of assigning to $r, s,\dots, z$ the values 1, 2,..., n in some order and N is the total number of inversions of the suffixes $r, s,\dots, z$. But the factor multiplying $|b_{pq}|$ in (16) is merely the expanded form of $|a_{pq}|$, so that

$$|c_{pq}| = |a_{pq}|\times|b_{pq}|.$$

3.2. *Second proof*—a proof that uses a single suffix notation. Consider, in the first place,

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \qquad \Delta' = \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix}.$$

Then

$$\Delta\Delta' = \begin{vmatrix} a_1 & b_1 & 0 & 0 \\ a_2 & b_2 & 0 & 0 \\ -1 & 0 & \alpha_1 & \beta_1 \\ 0 & -1 & \alpha_2 & \beta_2 \end{vmatrix}, \tag{17}$$

---

† Note that $a_{1r}b_{r1}$ stands for a *single* term.

as we see by using Laplace's method and expanding the determinant by its last two columns.

In the determinant (17) take

$$r_1' = r_1+a_1 r_3+b_1 r_4,$$
$$r_2' = r_2+a_2 r_3+b_2 r_4,$$

using the notation of p. 17, § 3.5. Then (17) becomes

$$\Delta\Delta' = \begin{vmatrix} 0 & 0 & a_1\alpha_1+b_1\alpha_2 & a_1\beta_1+b_1\beta_2 \\ 0 & 0 & a_2\alpha_1+b_2\alpha_2 & a_2\beta_1+b_2\beta_2 \\ -1 & 0 & \alpha_1 & \beta_1 \\ 0 & -1 & \alpha_2 & \beta_2 \end{vmatrix},$$

and so, on expanding the last determinant by its first two columns,

$$\Delta\Delta' = \begin{vmatrix} a_1\alpha_1+b_1\alpha_2 & a_1\beta_1+b_1\beta_2 \\ a_2\alpha_2+b_2\alpha_2 & a_2\beta_1+b_2\beta_2 \end{vmatrix}. \tag{18}$$

The argument extends readily to determinants of order n. Let $\Delta = (a_1 b_2 \ldots k_n)$, $\Delta' = (\alpha_1 \beta_2 \ldots \kappa_n)$. The determinant

$$\begin{vmatrix} a_1 & . & . & k_1 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ a_n & . & . & k_n & 0 & . & . & 0 \\ -1 & . & . & 0 & \alpha_1 & . & . & \kappa_1 \\ . & . & . & . & . & . & . & . \\ 0 & . & . & -1 & \alpha_n & . & . & \kappa_n \end{vmatrix},$$

in which the bottom left-hand quarter has -1 in each element of its principal diagonal and 0 elsewhere, is unaltered if we write

$$r_1' = r_1+a_1 r_{n+1}+b_1 r_{n+2}+\ldots+k_1 r_{2n},$$
$$r_2' = r_2+a_2 r_{n+1}+b_2 r_{n+2}+\ldots+k_2 r_{2n},$$
$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$
$$r_n' = r_n+a_n r_{n+1}+b_n r_{n+2}+\ldots+k_n r_{2n}.$$

Hence

$$\Delta\Delta' = \begin{vmatrix} 0 & . & . & 0 & a_1\alpha_1+\ldots+k_1\alpha_n & . & . & a_1\kappa_1+\ldots+k_1\kappa_n \\ . & . & . & . & . & . & . & . \\ 0 & . & . & 0 & a_n\alpha_1+\ldots+k_n\alpha_n & . & . & a_n\kappa_1+\ldots+k_n\kappa_n \\ -1 & . & . & 0 & \alpha_1 & . & . & \kappa_1 \\ . & . & . & . & . & . & . & . \\ 0 & . & . & -1 & \alpha_n & . & . & \kappa_n \end{vmatrix}.$$

When we expand the last determinant by its first n columns, the only non-zero term is (for sign see p. 39, § 7.3)

$$(-1)^n \begin{vmatrix} -1 & 0 & . & 0 \\ 0 & -1 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & -1 \end{vmatrix} . \begin{vmatrix} a_1\alpha_1+...+k_1\alpha_n & . & a_1\kappa_1+...+k_1\kappa_n \\ a_2\alpha_1+...+k_2\alpha_n & . & a_2\kappa_1+...+k_2\kappa_n \\ . & . & . \\ a_n\alpha_1+...+k_n\alpha_n & . & a_n\kappa_1+...+k_n\kappa_n \end{vmatrix}$$

$$= (-1)^{2n} \begin{vmatrix} a_1\alpha_1+...+k_1\alpha_n & . & . & a_1\kappa_1+...+k_1\kappa_n \\ . & . & . & . \\ a_n\alpha_1+...+k_n\alpha_n & . & . & a_n\kappa_1+...+k_n\kappa_n \end{vmatrix}.$$

Moreover, $2n$ is an even number and so the sign to be prefixed to the determinant is always the plus sign. Hence

$$\begin{vmatrix} a_1 & . & k_1 \\ . & . & . \\ a_n & . & k_n \end{vmatrix} . \begin{vmatrix} \alpha_1 & . & \kappa_1 \\ . & . & . \\ \alpha_n & . & \kappa_n \end{vmatrix} = \begin{vmatrix} a_1\alpha_1+...+k_1\alpha_n & . & a_1\kappa_1+...+k_1\kappa_n \\ . & . & . \\ a_n\alpha_1+...+k_n\alpha_n & . & a_n\kappa_1+...+k_n\kappa_n \end{vmatrix}, \tag{19}$$

which is another way of stating Theorem 13.

4. Other ways of multiplying determinants

The rule contained in Theorem 13 is sometimes called the *matrix*† *rule* or *the rule for multiplication of rows by columns*. In forming the first row of the determinant on the right of (19) we multiply the elements of the first row of Δ by the elements of the successive columns of Δ'; in forming the second row of the determinant we multiply the elements of the second row of Δ by the elements of the successive columns of Δ'. The process is most easily fixed in the mind by considering (18).

A determinant is unaltered in value if rows and columns are interchanged, and so we can at once deduce from (18) that, if

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \qquad \Delta' = \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix},$$

then (18) may also be written in the forms

$$\Delta\Delta' = \begin{vmatrix} a_1\alpha_1+b_1\beta_1 & a_1\alpha_2+b_1\beta_2 \\ a_2\alpha_1+b_2\beta_1 & a_2\alpha_2+b_2\beta_2 \end{vmatrix}, \tag{20}$$

$$\Delta\Delta' = \begin{vmatrix} a_1\alpha_1+a_2\alpha_2 & b_1\alpha_1+b_2\alpha_2 \\ a_1\beta_1+a_2\beta_2 & b_1\beta_1+b_2\beta_2 \end{vmatrix}. \tag{21}$$

---

† See Chapter VI.

In (20) the elements of the rows of Δ are multiplied by the elements of the rows of Δ'; in (21) the elements of the columns of Δ are multiplied by the elements of the columns of Δ'. The first of these is referred to as *multiplication by rows*, the second as *multiplication by columns*.

The extension to determinants of order n is immediate: several examples of the process occur later in the book. In the examples that follow multiplication is by rows: this is a matter of taste, and many writers use multiplication by columns or by the matrix rule. There is, however, much economy of thought if one consistently uses the same method whenever possible.

EXAMPLES IV

A number of interesting results can be obtained by applying the rules for multiplying determinants to particular examples. We shall arrange the examples in groups and we shall indicate the method of solution for at least one example in each group.

1. $$\begin{vmatrix} 0 & (\alpha-\beta)^2 & (\alpha-\gamma)^2 \\ (\beta-\alpha)^2 & 0 & (\beta-\gamma)^2 \\ (\gamma-\alpha)^2 & (\gamma-\beta)^2 & 0 \end{vmatrix} = 2\{(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)\}^2.$$

The determinant is the product by rows of the two determinants

$$\begin{vmatrix} \alpha^2 & -2\alpha & 1 \\ \beta^2 & -2\beta & 1 \\ \gamma^2 & -2\gamma & 1 \end{vmatrix}, \qquad \begin{vmatrix} 1 & \alpha & \alpha^2 \\ 1 & \beta & \beta^2 \\ 1 & \gamma & \gamma^2 \end{vmatrix}.$$

By Theorem 8, the first of these is equal to $2(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)$ and the second to $(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)$.

[In applying Theorem 8 we write down the product of the differences and adjust the numerical constant by considering the diagonal term of the determinant.]

✓2. The determinant of Example 3, p. 19, is the product by rows of the two determinants

$$\begin{vmatrix} \alpha^2 & -2\alpha & 1 & 0 \\ \beta^2 & -2\beta & 1 & 0 \\ \gamma^2 & -2\gamma & 1 & 0 \\ \delta^2 & -2\delta & 1 & 0 \end{vmatrix}, \qquad \begin{vmatrix} 1 & \alpha & \alpha^2 & 0 \\ 1 & \beta & \beta^2 & 0 \\ 1 & \gamma & \gamma^2 & 0 \\ 1 & \delta & \delta^2 & 0 \end{vmatrix},$$

and so is zero.

3. Prove that the determinant of order n which has $(a_r-a_s)^2$ as the element in the rth row and sth column is zero when $n > 3$.

4. Evaluate the determinant of order n which has $(a_r-a_s)^3$ as the element in the rth row and sth column: (i) when $n = 4$, (ii) when $n > 4$.

5. Extend the results of Examples 3 and 4 to higher powers of $a_r - a_s$.

6. (*Harder.*) Prove that

$$\begin{vmatrix} (a-x)^3 & (a-y)^3 & (a-z)^3 & a^3 \\ (b-x)^3 & (b-y)^3 & (b-z)^3 & b^3 \\ (c-x)^3 & (c-y)^3 & (c-z)^3 & c^3 \\ x^3 & y^3 & z^3 & 0 \end{vmatrix}$$

$$= -9abcxyz(b-c)(c-a)(a-b)(y-z)(z-x)(x-y).$$

7. (*Harder.*) Prove that

$$\begin{vmatrix} (a-x)^2 & (a-y)^2 & (a-z)^2 & a^2 \\ (b-x)^2 & (b-y)^2 & (b-z)^2 & b^2 \\ (c-x)^2 & (c-y)^2 & (c-z)^2 & c^2 \\ x^k & y^k & z^k & 0 \end{vmatrix}$$

is zero when $k = 1, 2$ and evaluate it when $k = 0$ and when $k \geqslant 3$.

✓ 8. Prove that, if $s_r = \alpha^r + \beta^r + \gamma^r + \delta^r$, then

$$\begin{vmatrix} s_0 & s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 & s_4 \\ s_2 & s_3 & s_4 & s_5 \\ s_3 & s_4 & s_5 & s_6 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \\ \alpha^3 & \beta^3 & \gamma^3 & \delta^3 \end{vmatrix}^2.$$

Hence (by Theorem 8) prove that the first determinant is equal to the product of the squares of the differences of α, β, γ, δ; i.e. $\{\zeta(\alpha, \beta, \gamma, \delta)\}^2$.

✓ 9. Prove that $(x-\alpha)(x-\beta)(x-\gamma)(x-\delta)$ is a factor of

$$\Delta \equiv \begin{vmatrix} s_0 & s_1 & s_2 & s_3 & s_4 \\ s_1 & s_2 & s_3 & s_4 & s_5 \\ s_2 & s_3 & s_4 & s_5 & s_6 \\ s_3 & s_4 & s_5 & s_6 & s_7 \\ 1 & x & x^2 & x^3 & x^4 \end{vmatrix}$$

and find the other factor.

Solution. The arrangement of the elements s_r in Δ indicates the product by rows (*vide* Example 8) of

$$\Delta_1 = \begin{vmatrix} 1 & 1 & 1 & 1 & 0 \\ \alpha & \beta & \gamma & \delta & 0 \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 & 0 \\ \alpha^3 & \beta^3 & \gamma^3 & \delta^3 & 0 \\ ? & ? & ? & ? & ? \end{vmatrix}, \qquad \Delta_2 = \begin{vmatrix} 1 & 1 & 1 & 1 & ? \\ \alpha & \beta & \gamma & \delta & ? \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 & ? \\ \alpha^3 & \beta^3 & \gamma^3 & \delta^3 & ? \\ \alpha^4 & \beta^4 & \gamma^4 & \delta^4 & ? \end{vmatrix}.$$

If we have anything but 0's in the first four places of the last row of Δ_1, then we shall get unwanted terms in the last row of the product $\Delta_1 \Delta_2$. So we try 0, 0, 0, 0, 1 as the last row of Δ_1 and then it is not hard to see that, in order to give $\Delta_1 \Delta_2 = \Delta$, the last column of Δ_2 must be $1, x, x^2, x^3, x^4$.

The factors of Δ follow from Theorem 8.

✓10. Prove that, when $s_r = \alpha^r+\beta^r+\gamma^r$,

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^3 & \beta^3 & \gamma^3 \end{vmatrix} \times \begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} = \begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_3 & s_4 & s_5 \end{vmatrix}$$

Note how the sequence of indices in the columns of the first determinant affects the sequence of suffixes in the columns of the last determinant.

✓11. Find the factors of

$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_2 & s_3 & s_4 \\ s_4 & s_5 & s_6 \end{vmatrix}, \quad \begin{vmatrix} s_0 & s_2 & s_4 \\ s_2 & s_4 & s_6 \\ s_4 & s_6 & s_8 \end{vmatrix}, \quad \begin{vmatrix} s_0 & s_1 & s_2 & s_4 \\ s_1 & s_2 & s_3 & s_5 \\ s_3 & s_4 & s_5 & s_7 \\ 1 & x & x^2 & x^4 \end{vmatrix},$$

where $s_r = \alpha^r+\beta^r+\gamma^r$.

12. Extend the results of Examples 8–11 to determinants of higher orders.

---

13. Multiply the determinants of the fourth order whose rows are given by

$$x_r^2+y_r^2 \quad -2x_r \quad -2y_r \quad 1; \qquad 1 \quad x_r \quad y_r \quad x_r^2+y_r^2$$

and $r = 1, 2, 3, 4$. Hence prove that the determinant $|a_{rs}|$, where $a_{rs} = (x_r-x_s)^2+(y_r-y_s)^2$, is zero whenever the four points (x_r, y_r) are concyclic.

14. Find a relation between the mutual distances of five points on a sphere.

---

15. By considering the product of two determinants of the form

$$\begin{vmatrix} a+ib & c+id \\ -(c-id) & a-ib \end{vmatrix},$$

prove that the product of a sum of four squares by a sum of four squares is itself a sum of four squares.

16. Express $a^3+b^3+c^3-3abc$ as a circulant (p. 23, § 5.2) of order three and hence prove that $(a^3+b^3+c^3-3abc)(A^3+B^3+C^3-3ABC)$ may be written in the form $X^3+Y^3+Z^3-3XYZ$.

Prove, further, that if $A = a^2-bc$, $B = b^2-ca$, $C = c^2-ab$, then $A^3+B^3+C^3-3ABC = (a^3+b^3+c^3-3abc)^2$.

---

17. If $X = ax+hy$, $Y = hx+by$; (x_1, y_1), (x_2, y_2) two sets of values of (x, y); $S_{11} = x_1X_1+y_1Y_1$, $S_{12} = x_1X_2+y_1Y_2$, etc.; prove that $S_{12} = S_{21}$ and that

$$\begin{vmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{vmatrix} = \begin{vmatrix} a & h \\ h & b \end{vmatrix} \times \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}^2.$$

18 A. (*Harder.*) Using the summation convention, let $S \equiv a_{rs}x^r x^s$, where $x^1, x^2,\ldots, x^n$ are n independent variables (and *not* powers of x), be a given quadratic form in n variables. Let $X_{r,\lambda} = a_{rs}x^s_\lambda$; $S_{\lambda\mu} = x^r_\mu X_{r,\lambda}$, where $(x^1_\lambda, x^2_\lambda,\ldots, x^n_\lambda)$, $\lambda = 1,\ldots, n$, denotes n sets of values of the variables $x^1, x^2,\ldots, x^n$. Prove that

$$|S_{\lambda\mu}| = |a_{rs}| \times |x^s_\lambda|^2.$$

18 B. (*Easier.*) Extend the result of Example 17 to three variables (x, y, z) and the coordinates of three points in space, (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3).

5. Multiplication of arrays

5.1. First take two arrays

$$\begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{matrix} \qquad \begin{matrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \end{matrix} \tag{A}$$

in which the number of columns exceeds the number of rows, and multiply them by rows in the manner of multiplying determinants. The result is the determinant

$$\Delta \equiv \begin{vmatrix} a_1\alpha_1+b_1\beta_1+c_1\gamma_1 & a_1\alpha_2+b_1\beta_2+c_1\gamma_2 \\ a_2\alpha_1+b_2\beta_1+c_2\gamma_1 & a_2\alpha_2+b_2\beta_2+c_2\gamma_2 \end{vmatrix}. \tag{1}$$

If we expand Δ and pick out the terms involving, say, $\beta_1\gamma_2$ we see that they are $\beta_1\gamma_2(b_1c_2-b_2c_1)$. Similarly, the terms in $\beta_2\gamma_1$ are $-(b_1c_2-b_2c_1)$. Hence Δ contains a term

$$(b_1c_2)\times(\beta_1\gamma_2),$$

where (b_1c_2) denotes the determinant formed by the last two columns of the first array in (A) and $(\beta_1\gamma_2)$ the corresponding determinant formed from the second array.

It follows that Δ, when expanded, contains terms

$$(b_1c_2)(\beta_1\gamma_2)+(c_1a_2)(\gamma_1\alpha_2)+(a_1b_2)(\alpha_1\beta_2); \tag{2}$$

moreover (2) accounts for *all* terms that can possibly arise from the expansion of Δ. Hence Δ is equal to the sum of terms in (2); that is, *the sum of all the products of corresponding determinants of order two that can be formed from the arrays in* (A).

Now consider $a, b,\ldots, k,\ldots, t$, supposing k to be the nth letter and t the Nth, where $n < N$; consider also a corresponding notation in Greek letters.

The determinant, of order n, that is obtained on multiplying by rows the two arrays

$$\begin{array}{ccccccccc} a_1 & . & . & k_1 & . & . & t_1 \\ . & . & . & . & . & . & . \\ a_n & . & . & k_n & . & . & t_n \end{array} \qquad \begin{array}{ccccccccc} \alpha_1 & . & . & \kappa_1 & . & . & \tau_1 \\ . & . & . & . & . & . & . \\ \alpha_n & . & . & \kappa_n & . & . & \tau_n \end{array}$$

is

$$\Delta \equiv \begin{vmatrix} a_1\alpha_1+\ldots+t_1\tau_1 & . & . & a_1\alpha_n+\ldots+t_1\tau_n \\ . & . & . & . \\ a_n\alpha_1+\ldots+t_n\tau_1 & . & . & a_n\alpha_n+\ldots+t_n\tau_n \end{vmatrix}. \tag{3}$$

This determinant, when expanded, contains a term (put all letters after k, the nth letter, equal to zero)

$$\begin{vmatrix} a_1\alpha_1+\ldots+k_1\kappa_1 & . & . & a_1\alpha_n+\ldots+k_1\kappa_n \\ . & . & . & . \\ a_n\alpha_1+\ldots+k_n\kappa_1 & . & . & a_n\alpha_n+\ldots+k_n\kappa_n \end{vmatrix}, \tag{4}$$

which is the product of the two determinants $(a_1\, b_2\, \ldots\, k_n)$ and $(\alpha_1\, \beta_2\, \ldots\, \kappa_n)$. Moreover (4) includes every term in the expansion of Δ containing the letters $a, b, \ldots, k$ and no other roman letter. Thus the full expansion of Δ consists of a sum of such products, the number of such products being ${}^N C_n$, the number of ways of choosing n distinct letters from N given letters. We have therefore proved the following theorem.

THEOREM 14. *The determinant Δ, of order n, which is obtained on multiplying by rows two arrays that have N columns and n rows, where $n < N$, is the sum of all the products of corresponding determinants of order n that can be formed from the two arrays.*

5.2. THEOREM 15. *The determinant of order n which is obtained on multiplying by rows two arrays that have N columns and n rows, where $n > N$, is equal to zero.*

The determinant so formed is, in fact, the product by rows of the two zero determinants of order n that one obtains on adding $(n-N)$ columns of ciphers to each array. For example,

$$\begin{vmatrix} a_1\alpha_1+b_1\beta_1 & a_1\alpha_2+b_1\beta_2 & a_1\alpha_3+b_1\beta_3 \\ a_2\alpha_1+b_2\beta_1 & a_2\alpha_2+b_2\beta_2 & a_2\alpha_3+b_2\beta_3 \\ a_3\alpha_1+b_3\beta_1 & a_3\alpha_2+b_3\beta_2 & a_3\alpha_3+b_3\beta_3 \end{vmatrix}$$

is obtained on multiplying by rows the two arrays

$$\begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{array} \qquad \begin{array}{cc} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \end{array}$$

and is also obtained on multiplying by rows the two zero determinants

$$\begin{vmatrix} a_1 & b_1 & 0 \\ a_2 & b_2 & 0 \\ a_3 & b_3 & 0 \end{vmatrix}, \qquad \begin{vmatrix} \alpha_1 & \beta_1 & 0 \\ \alpha_2 & \beta_2 & 0 \\ \alpha_3 & \beta_3 & 0 \end{vmatrix}.$$

EXAMPLES V

1. The arrays

$$\begin{array}{cccc} 1 & 0 & 0 & 0 \\ x_1^2+y_1^2 & -2x_1 & -2y_1 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ x_4^2+y_4^2 & -2x_4 & -2y_4 & 1 \end{array} \qquad \begin{array}{cccc} 0 & 0 & 0 & 1 \\ 1 & x_1 & y_1 & x_1^2+y_1^2 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_4 & y_4 & x_4^2+y_4^2 \end{array}$$

have 5 rows and 4 columns. By Theorem 15, the determinant obtained on multiplying them by rows vanishes identically. Show that this result gives a relation connecting the mutual distances of any four points in a plane.

2. Obtain the corresponding relation for five points in space.

---

3. If α, β, γ,... are the roots of an equation of degree n, and if s_r denotes the sum of the rth powers of the roots, prove that

$$\begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix} = \sum (\alpha-\beta)^2,$$

$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = \sum (\beta-\gamma)^2(\gamma-\alpha)^2(\alpha-\beta)^2.$$

HINT. Compare Example 8, p. 48, and Theorem 14.

4. Obtain identities by evaluating (in two ways) the determinant obtained on squaring (by rows) the arrays

$$\text{(i)} \quad \begin{array}{ccc} a & b & c \\ a' & b' & c' \end{array} \qquad \text{(ii)} \quad \begin{array}{cccc} a & b & c & d \\ a' & b' & c' & d' \end{array}$$

5. In the notation usually adopted for conics, let

$$S \equiv ax^2+\ldots+2fyz+\ldots,$$

$$X \equiv ax+hy+gz, \qquad Y \equiv hx+by+fz, \qquad Z \equiv gx+fy+cz,$$

$$S_{rs} = x_r X_s + y_r Y_s + z_r Z_s;$$

let A, B,..., H be the co-factors of a, b,..., h in the discriminant of S. Let $\xi = y_1 z_2 - y_2 z_1$, $\eta = z_1 x_2 - z_2 x_1$, $\zeta = x_1 y_2 - x_2 y_1$. Then

$$\begin{vmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{vmatrix} = A\xi^2 + B\eta^2 + C\zeta^2 + 2F\eta\zeta + 2G\zeta\xi + 2H\xi\eta.$$

Solution. The determinant is

$$\begin{vmatrix} x_1 X_1 + ... + z_1 Z_1 & x_1 X_2 + ... + z_1 Z_2 \\ x_2 X_1 + ... + z_2 Z_1 & x_2 X_2 + ... + z_2 Z_2 \end{vmatrix},$$

which is the product of arrays

$$\begin{matrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{matrix} \qquad \begin{matrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{matrix}$$

and so is $$\sum (y_1 z_2)(Y_1 Z_2).$$

But $(Y_1 Z_2)$, when written in full, is

$$\begin{vmatrix} hx_1 + by_1 + fz_1 & gx_1 + fy_1 + cz_1 \\ hx_2 + by_2 + fz_2 & gx_2 + fy_2 + cz_2 \end{vmatrix},$$

which is the product of arrays

$$\begin{matrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{matrix} \qquad \begin{matrix} h & b & f \\ g & f & c \end{matrix}$$

and so is $A\xi + H\eta + G\zeta$.

Hence the initial determinant is the sum of three terms of the type $\xi(A\xi + H\eta + G\zeta)$.

6. The multiplication of determinants of different orders

It is sometimes useful to be able to write down the product of a determinant of order n by a determinant of order m. The method† of doing so is sufficiently exemplified by considering the product of $(a_1\, b_2\, c_3)$ by $(\alpha_1\, \beta_2)$.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 \alpha_1 + b_1 \beta_1 & a_1 \alpha_2 + b_1 \beta_2 & c_1 \\ a_2 \alpha_1 + b_2 \beta_1 & a_2 \alpha_2 + b_2 \beta_2 & c_2 \\ a_3 \alpha_1 + b_3 \beta_1 & a_2 \alpha_2 + b_3 \beta_2 & c_3 \end{vmatrix}$$

We see that the result is true by considering the first and last determinants when expanded by their last columns. The first member of the above equation is

$$\{\sum \pm c_r (a_s\, b_t)\} \times (\alpha_1\, \beta_2),$$

---

† I learnt this method from Professor A. L. Dixon.

while the second is, by the rule for multiplying two determinants of the same order,

$$\sum \pm c_r (a_s b_t) \times (\alpha_1 \beta_2),$$

the signs having the same arrangement in both summations.

A further example is

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \times \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix}$$

$$= \begin{vmatrix} a_1\alpha_1+b_1\beta_1 & a_1\alpha_2+b_1\beta_2 & c_1 & d_1 \\ a_2\alpha_1+b_2\beta_1 & a_2\alpha_2+b_2\beta_2 & c_2 & d_2 \\ a_3\alpha_1+b_3\beta_1 & a_3\alpha_2+b_3\beta_2 & c_3 & d_3 \\ a_4\alpha_1+b_4\beta_1 & a_4\alpha_2+b_4\beta_2 & c_4 & d_4 \end{vmatrix},$$

a result which becomes evident on considering the Laplace expansion of the last determinant by its third and fourth columns.

The reader may prefer the following method:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} \alpha_1 & \beta_1 & 0 \\ \alpha_2 & \beta_2 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a_1\alpha_1+b_1\beta_1 & a_1\alpha_2+b_1\beta_2 & c_1 \\ a_2\alpha_1+b_2\beta_1 & a_2\alpha_2+b_2\beta_2 & c_2 \\ a_3\alpha_1+b_3\beta_1 & a_3\alpha_2+b_3\beta_2 & c_3 \end{vmatrix}.$$

This is, perhaps, easier if a little less elegant.

CHAPTER IV

JACOBI'S THEOREM AND ITS EXTENSIONS

1. Jacobi's theorem

THEOREM 16. *Let* A_r, B_r,... *denote the co-factors of* a_r, b_r,... *in a determinant*

$$\Delta \equiv (a_1\, b_2\, \ldots\, k_n).$$

Then $$\Delta' \equiv (A_1\, B_2\, \ldots\, K_n) = \Delta^{n-1}.$$

If we multiply by rows the two determinants

$$\Delta = \begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix}, \qquad \Delta' = \begin{vmatrix} A_1 & B_1 & . & . & K_1 \\ A_2 & B_2 & . & . & K_2 \\ . & . & . & . & . \\ A_n & B_n & . & . & K_n \end{vmatrix},$$

we obtain $$\Delta\Delta' = \begin{vmatrix} \Delta & 0 & . & . & 0 \\ 0 & \Delta & . & . & 0 \\ . & . & . & . & . \\ 0 & 0 & . & . & \Delta \end{vmatrix};$$

for $a_r A_s + b_r B_s + \ldots + k_r K_s$ is equal to zero when $r \neq s$ and is equal to Δ when $r = s$.

Hence $$\Delta\Delta' = \Delta^n,$$

so that, when $\Delta \neq 0$, $$\Delta' = \Delta^{n-1}.$$

But when Δ is zero, Δ' is also zero. For, by Theorem 12, if $\Delta = 0$, then $A_r B_s - A_s B_r = 0$, so that the Laplace expansion of Δ' by its first two columns is a sum of zeros.

Hence when $\Delta = 0$, $\Delta' = 0 = \Delta^{n-1}$.

DEFINITION. Δ' *is called the* ADJUGATE *determinant of* Δ.

THEOREM 17. *With the notation of Theorem* 16, *the co-factor of* A_1 *in* Δ' *is equal to* $a_1 \Delta^{n-2}$; *and so for other letters and suffixes.*

When we multiply by rows the two determinants

$$\begin{vmatrix} 1 & 0 & . & . & 0 \\ A_2 & B_2 & . & . & K_2 \\ . & . & . & . & . \\ A_n & B_n & . & . & K_n \end{vmatrix}, \qquad \begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix}$$

we obtain

$$\begin{vmatrix} a_1 & a_2 & . & . & a_n \\ 0 & \Delta & . & . & 0 \\ . & . & . & . & . \\ 0 & 0 & . & . & \Delta \end{vmatrix},$$

wherein all terms to the left of and below the leading diagonal are zero. Thus

$$\Delta \begin{vmatrix} B_2 & . & . & K_2 \\ . & . & . & . \\ B_n & . & . & K_n \end{vmatrix} = a_1 \Delta^{n-1}.$$

When $\Delta \neq 0$, this gives the result stated in the theorem. When $\Delta = 0$, the result follows by the argument used when we considered Δ' in Theorem 16; each $B_r C_s - B_s C_r$ is zero.

Moreover, if we wish to consider the minor of J_r in Δ', where J is the sth letter of the alphabet, we multiply by rows

$$\begin{vmatrix} A_1 & . & . & J_1 & . & . & K_1 \\ . & . & . & . & . & . & . \\ A_{r-1} & . & . & J_{r-1} & . & . & K_{r-1} \\ 0 & . & . & 1 & . & . & 0 \\ A_{r+1} & . & . & J_{r+1} & . & . & K_{r+1} \\ . & . & . & . & . & . & . \\ A_n & . & . & J_n & . & . & K_n \end{vmatrix}, \qquad \begin{vmatrix} a_1 & . & . & j_1 & . & . & k_1 \\ . & . & . & . & . & . & . \\ a_r & . & . & j_r & . & . & k_r \\ . & . & . & . & . & . & . \\ a_n & . & . & j_n & . & . & k_n \end{vmatrix},$$

to obtain

$$\begin{vmatrix} \Delta & . & . & 0 & . & . & 0 \\ . & . & . & . & . & . & . \\ j_1 & . & . & j_s & . & . & j_n \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & . & . & \Delta \end{vmatrix}, \tag{1}$$

where the j's occur in the rth row and the only other non-zero terms are Δ's in the leading diagonal. The value of (1) is $j_r \Delta^{n-1}$.

2. General form of Jacobi's theorem

2.1. Complementary minors. In a determinant Δ, of order n, the elements of any given r rows and r columns, where $r < n$, form a minor of the determinant. When these same rows and columns are excluded from Δ, the elements that are left form a determinant of order $n-r$, which, when the appropriate sign is prefixed, is called the complementary minor of the original minor. The *sign* to be prefixed is determined by the rule that a Laplace expansion shall always be of the form

$$\sum +\mu_r \gamma_{n-r},$$

where γ_{n-r} is the minor complementary to μ_r. Thus, in

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}, \tag{1}$$

the complementary minor of

$$\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} \quad \text{is} \quad +\begin{vmatrix} b_2 & d_2 \\ b_4 & d_4 \end{vmatrix}, \tag{2}$$

since the Laplace expansion of (1) by its first and third columns involves the term $+(a_1\,c_3)(b_2\,d_4)$.

THEOREM 18. *With the notation of Theorem* 16, *let* $\mathfrak{M}_r$ *be a minor of* Δ' *having r rows and columns and let* γ_{n-r} *be the complementary minor of the corresponding minor of* Δ. *Then*

$$\mathfrak{M}_r = \gamma_{n-r}\,\Delta^{r-1}. \tag{3}$$

In particular, if Δ *is the determinant* (1) *above, then, in the usual notation for co-factors,*

$$\begin{vmatrix} A_1 & C_1 \\ A_3 & C_3 \end{vmatrix} = \begin{vmatrix} b_2 & d_2 \\ b_4 & d_4 \end{vmatrix} \Delta.$$

Let us first dispose of the easy particular cases of the theorem. If $r = 1$, (3) is merely the definition of a first minor; for example, A_1 in (1) above is, by definition, the determinant $(b_2\,c_3\,d_4)$. If $r > 1$ and $\Delta = 0$, then also $\mathfrak{M}_r = 0$, as we see by Theorem 12. Accordingly, (3) is true whenever $r = 1$ and whenever $\Delta = 0$.

It remains to prove that (3) is true when $r > 1$ and $\Delta \neq 0$. So let $r > 1$, $\Delta \neq 0$; further, let μ_r be the minor of Δ whose elements correspond in row and column to those elements of Δ' that compose $\mathfrak{M}_r$. Then, by the definition of complementary minor, due regard being paid to the sign, one Laplace expansion of Δ must contain the term $+\mu_r \gamma_{n-r}$. Hence Δ may be written in the form

$$\left|\begin{array}{c:c} \mu_r & \text{other elements} \\ \hdashline \text{other elements} & \gamma_{n-r} \end{array}\right|.$$

It is thus sufficient to prove the theorem when $\mathfrak{M}$ is formed from the first r rows and first r columns of Δ'; this we now do.

Let $A,\ldots, E$ be the first r letters, and let $F,\ldots, K$ be the next $n-r$ letters of the alphabet. When we form the product

$$\left|\begin{array}{ccccccccc} A_1 & . & . & E_1 & F_1 & . & . & K_1 \\ . & . & . & . & . & . & . & . \\ A_r & . & . & E_r & F_r & . & . & K_r \\ 0 & . & . & 0 & 1 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ 0 & . & . & 0 & 0 & . & . & 1 \end{array}\right| \times \left|\begin{array}{cccccccc} a_1 & . & . & e_1 & f_1 & . & . & k_1 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ a_n & . & . & e_n & f_n & . & . & k_n \end{array}\right|,$$

where the last $n-r$ elements of the leading diagonal of the first determinant are 1's and all other elements of the last $n-r$ rows are 0's, we obtain

$$\left|\begin{array}{cccccccc} \Delta & . & . & 0 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ 0 & . & . & \Delta & 0 & . & . & 0 \\ f_1 & . & . & f_r & f_{r+1} & . & . & f_n \\ . & . & . & . & . & . & . & . \\ k_1 & . & . & k_r & k_{r+1} & . & . & k_n \end{array}\right|$$

That is to say,

$$\mathfrak{M}_r \times \Delta = \Delta^r \gamma_{n-r},$$

and so (3) is true whenever $r > 1$ and $\Delta \neq 0$.

CHAPTER V

SYMMETRICAL AND SKEW-SYMMETRICAL DETERMINANTS

1. The determinant $|a_{rs}|$ is said to be symmetrical if $a_{rs} = a_{sr}$ for every r and s.

The determinant $|a_{rs}|$ is said to be skew-symmetrical if $a_{rs} = -a_{sr}$ for every r and s. It is an immediate consequence of the definition that the elements in the leading diagonal of such a determinant must all be zero, for $a_{rr} = -a_{rr}$.

2. Theorem 19. *A skew-symmetrical determinant of odd order has the value zero.*

The value of a determinant is unaltered when rows and columns are interchanged. In a skew-symmetrical determinant such an interchange is equivalent to multiplying each row of the original determinant by -1, that is, to multiplying the whole determinant by $(-1)^n$. Hence the value of a skew-symmetrical determinant of order n is unaltered when it is multiplied by $(-1)^n$; and if n is odd, this value must be zero.

3. Before considering determinants of even order we shall examine the first minors of a skew-symmetrical determinant of odd order.

Let A_{rs} denote the co-factor of a_{rs}, the element in the rth row and sth column of $|a_{rs}|$, a skew-symmetrical determinant of odd order; then A_{rs} is $(-1)^{n-1}A_{sr}$, that is, $A_{sr} = A_{rs}$. For A_{rs} and A_{sr} differ only in considering column for row and row for column in Δ, and, as we have seen in § 2, a column of Δ is (-1) times the corresponding row; moreover, in forming A_{rs} we take the elements of $n-1$ rows of Δ.

4. We require yet another preliminary result: this time one that is often useful in other connexions.

THEOREM 20. *The determinant*

$$\begin{vmatrix} a_{11} & a_{12} & . & . & a_{1n} & X_1 \\ a_{21} & a_{22} & . & . & a_{2n} & X_2 \\ . & . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nn} & X_n \\ Y_1 & Y_2 & . & . & Y_n & S \end{vmatrix} \tag{1}$$

$$= |a_{rs}|S - \sum_{r=1}^{n} \sum_{s=1}^{n} A_{rs} X_r Y_s, \tag{2}$$

where A_{rs} *is the co-factor of* a_{rs} *in* $|a_{rs}|$.

The one term of the expansion that involves no X and no Y is $|a_{rs}|S$. The term involving $X_r Y_s$ is obtained by putting S, all X other than X_r, and all Y other than Y_s to be zero. Doing this, we see that the coefficient of $X_r Y_s$ is $\pm A_{rs}$. But, considering the Laplace expansion of (1) by its rth and $(n+1)$th rows, we see that the coefficients of $X_r Y_s$ and of $a_{rs} S$ are of opposite sign. Hence the coefficient of $X_r Y_s$ is $-A_{rs}$. Moreover, when Δ is expanded as in Theorem 1 (i) every term must involve either S or a product $X_r Y_s$. Hence (2) accounts for *all* the terms in the expansion of Δ.

5. THEOREM 21. *A skew-symmetrical determinant of order* $2n$ *is the square of a polynomial function of its elements.*

In Theorem 20, let $|a_{rs}|$ be a skew-symmetrical determinant of order $2n-1$; as such, its value is zero (Theorem 19). Further, let $Y_r = -X_r$, $S = 0$, so that the determinant (1) of Theorem 20 is now a skew-symmetrical determinant of order $2n$. Its value is

$$\sum_{r=1}^{n} \sum_{s=1}^{n} A_{rs} X_r X_s. \tag{3}$$

Since $|a_{rs}| = 0$, $A_{rs} A_{sr} = A_{rr} A_{ss}$ (Theorem 12), or $A_{rs}^2 = A_{rr} A_{ss}$, and $A_{r1}/A_{11} = A_{rs}/A_{1s}$, or $A_{rs} = A_{1r} A_{1s}/A_{11}$. Hence (3) is

$$(X_1 \surd A_{11} \pm X_2 \surd A_{22} \pm \dots \pm X_n \surd A_{nn})^2, \tag{4}$$

where the sign preceding X_r is $(-1)^p$, chosen so that

$$A_{1r} = (-1)^p \surd(A_{11} A_{rr}).$$

Now $A_{11}, \dots, A_{nn}$ are themselves skew-symmetrical determinants of order $2n-2$, and if we suppose that each is the square of a polynomial function of its elements, of degree $n-1$, then (4) will be the square of a polynomial function of degree n. Hence our theorem is true for a determinant of order $2n$ if it is true for one of order $2n-2$.

But

$$\begin{vmatrix} 0 & a_{11} \\ -a_{11} & 0 \end{vmatrix} = a_{11}^2,$$

and is a perfect square. Hence our theorem is true when $n = 1$.

It follows by induction that the theorem is true for all integer values of n.

6. The Pfaffian

A polynomial whose square is equal to a skew-symmetrical determinant of even order is called a Pfaffian. Its properties and relations to determinants have been widely studied.

Here we give merely sufficient references for the reader to study the subject if he so desires. The study of Pfaffians, interesting though it may be, is too special in its appeal to warrant its inclusion in this book.

G. Salmon, *Lessons Introductory to the Modern Higher Algebra* (Dublin, 1885), Lesson V.

S. Barnard and J. M. Child, *Higher Algebra* (London, 1936), chapter ix, § 22.

Sir T. Muir, *Contributions to the History of Determinants*, vols. i–iv (London, 1890–1930). This history covers every topic of determinant theory; it is delightfully written and can be recommended as a reference to anyone who is seriously interested. It is too detailed and exacting for the beginner.

Examples VI (Miscellaneous)

1. Prove that

$$\begin{vmatrix} a_{11} & a_{12} & . & . & . & a_{1n} & X_1 \\ a_{21} & a_{22} & . & . & . & a_{2n} & X_2 \\ . & . & . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & . & a_{nn} & X_n \\ X_1 & X_2 & . & . & . & X_n & 0 \end{vmatrix}$$

$$= -\{A_{11}X_1^2+...+(A_{rs}+A_{sr})X_r X_s+...\},$$

where A_{rs} is the co-factor of a_{rs} in $|a_{rs}|$.

If $a_{rs} = a_{sr}$ and $|a_{rs}| = 0$, prove that the above quadratic form may be written as

$$-(A_{11}X_1+...+A_{1n}X_n)^2/A_{11}.$$

2. If $S = \sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} x_r x_s$, $2X_u = \partial S/\partial x_u$, $a_{rs} = a_{sr}$, and $k < n$, prove that

$$J = \begin{vmatrix} a_{11} & a_{12} & . & . & . & a_{1k} & X_1 \\ a_{21} & a_{22} & . & . & . & a_{2k} & X_2 \\ . & . & . & . & . & . & . \\ a_{k1} & a_{k2} & . & . & . & a_{kk} & X_k \\ X_1 & X_2 & . & . & . & X_k & S \end{vmatrix}$$

is independent of $x_1, x_2,..., x_k$.

HINT. Use Example 1, consider $\partial J/\partial x_u$, and use Theorem 10.

3. $\Delta_1 = (a_2 b_3 c_4)$, $\Delta_2 = (a_3 b_4 c_1)$, $\Delta_3 = (a_4 b_1 c_2)$, $\Delta_4 = (a_1 b_2 c_3)$, and A_r^s is the co-factor of a_r in Δ_s. Prove that $A_3^1 = -A_1^3$, that $A_4^1 = A_1^4$, and that

$$\begin{vmatrix} A_1^2 & A_2^3 & A_3^1 \\ B_1^2 & B_2^3 & B_3^1 \\ C_1^2 & C_2^3 & C_3^1 \end{vmatrix} = 0, \qquad \begin{vmatrix} A_1^4 & A_2^3 & A_3^1 \\ B_1^4 & B_2^3 & B_3^1 \\ C_1^4 & C_2^3 & C_3^1 \end{vmatrix} = \Delta_1 \Delta_3.$$

HINT. $B_2^3 C_1^3 - B_1^3 C_2^3 = -a_4 \Delta_3$ (by Theorem 17): use Theorem 10.

4. Δ is a determinant of order n, a_{rs} a typical element of Δ, A_{rs} the co-factor of a_{rs} in Δ, and $\Delta \neq 0$. Prove that

$$\begin{vmatrix} A_{11}+\Delta/x & A_{12} & . & . & . & A_{1n} \\ A_{21} & A_{22}+\Delta/x & . & . & . & A_{2n} \\ . & . & . & . & . & . \\ A_{n1} & A_{n2} & . & . & . & A_{nn}+\Delta/x \end{vmatrix} = 0$$

whenever

$$\begin{vmatrix} a_{11}+x & a_{12} & . & . & . & a_{1n} \\ a_{21} & a_{22}+x & . & . & . & a_{2n} \\ . & . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & . & a_{nn}+x \end{vmatrix} = 0.$$

5. Prove that

$$\begin{vmatrix} 1 & \sin\alpha & \cos\alpha & \sin 2\alpha & \cos 2\alpha \\ 1 & \sin\beta & \cos\beta & \sin 2\beta & \cos 2\beta \\ 1 & \sin\gamma & \cos\gamma & \sin 2\gamma & \cos 2\gamma \\ 1 & \sin\delta & \cos\delta & \sin 2\delta & \cos 2\delta \\ 1 & \sin\epsilon & \cos\epsilon & \sin 2\epsilon & \cos 2\epsilon \end{vmatrix} = 256 \prod \sin \tfrac{1}{2}(\alpha-\beta),$$

with a proper arrangement of the signs of the differences.

6. Prove that if $\Delta = |a_{rs}|$ and the summation convention be used (Greek letters only being used as dummies), then the results of Theorem 10 may be expressed as

$$a_{\alpha s} A_{\alpha r} = 0 = a_{r\alpha} A_{s\alpha} \quad \text{when } r \neq s,$$

$$a_{\alpha s} A_{\alpha r} = \Delta = a_{r\alpha} A_{s\alpha} \quad \text{when } r = s.$$

Prove also that if n variables x are related to n variables X by the transformation

$$X_r = a_{r\alpha} x_\alpha,$$

then

$$\Delta x_s = A_{\alpha s} X_\alpha.$$

7. Prove Theorem 9 by forming the product of the given circulant by the determinant $(\omega_1, \omega_2^2, ..., \omega_n^n)$, where $\omega_1, ..., \omega_n$ are distinct nth roots of unity.

8. Prove that, if $f_r(a) = g_r(a) = h_r(a)$ when $r = 1, 2, 3$, then the determinant

$$\begin{vmatrix} f_1(x) & g_1(x) & h_1(x) \\ f_2(x) & g_2(x) & h_2(x) \\ f_3(x) & g_3(x) & h_3(x) \end{vmatrix},$$

whose elements are polynomials in x, contains $(x-a)^2$ as a factor.

HINT. Consider a determinant with $c_1' = c_1 - c_2$, $c_2' = c_2 - c_3$ and use the remainder theorem.

9. If the elements of a determinant are polynomials in x, and if r columns (rows) become equal when $x = a$, then the determinant has $(x-a)^{r-1}$ as a factor.

10. Verify Theorem 9 when $n = 5$, $a_1 = a_4 = a_5 = 1$, $a_2 = a$, and $a_3 = a^2$ by independent proofs that both determinant and product are equal to

$$(a^2+a+3)(a-1)^4(a^4+3a^3+4a^2+2a+1).$$

PART II

MATRICES

CHAPTER VI

DEFINITIONS AND ELEMENTARY PROPERTIES

1. Linear substitutions

We can think of n numbers, real or complex, either as separate entities $x_1, x_2,..., x_n$ or as a single entity x that can be broken up into its n separate pieces (or components) if we wish to do so. For example, in ordinary three-dimensional space, given a set of axes through O and a point P determined by its coordinates with respect to those axes, we can think of the vector OP (or of the point P) and denote it by x, or we can give prominence to the components x_1, x_2, x_3 of the vector along the axes and write x as (x_1, x_2, x_3).

When we are thinking of x as a single entity we shall refer to it as a '*number*'. This is merely a slight extension of a common practice in dealing with a complex number z, which is, in fact, a pair of real numbers x, y.[†]

Now suppose that two '*numbers*' x, with components $x_1, x_2,...,$ x_n, and X, with components $X_1, X_2,..., X_n$, are connected by a set of n equations

$$X_r = a_{r1}x_1+a_{r2}x_2+...+a_{rn}x_n \quad (r = 1,..., n), \tag{1}$$

wherein the a_{rs} are given constants.

[For example, consider a change of axes in coordinate geometry where the x_r and X_r are the components of the same vector referred to different sets of axes.]

Introduce the notation A for the set of n^2 numbers, real or complex,

$$\begin{array}{cccccc} a_{11} & a_{12} & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & a_{nn} \end{array}$$

and understand by the notation Ax a '*number*' whose components are given by the expressions on the right-hand side of equations (1). Then the equations (1) can be written symbolically as

$$X = Ax. \tag{2}$$

[†] Cf. G. H. Hardy, *Pure Mathematics*, chap. iii.

The reader will find that, with but little practice, the symbolical equation (2) will be all that it is necessary to write when equations such as (1) are under consideration.

2. Linear substitutions as a guide to matrix addition

It is a familiar fact of vector geometry that the sum of a vector X, with components X_1, X_2, X_3, and a vector Y, with components Y_1, Y_2, Y_3, is a vector Z, or $X+Y$, with components X_1+Y_1, X_2+Y_2, X_3+Y_3.

Consider now a '*number*' x and a related '*number*' X given by

$$X = Ax, \tag{2}$$

where A has the same meaning as in § 1. Take, further, a second '*number*' Y given by

$$Y = Bx, \tag{3}$$

where B symbolizes the set of n^2 numbers

$$\begin{matrix} b_{11} & b_{12} & . & . & b_{1n} \\ . & . & . & . & . \\ b_{n1} & b_{n2} & . & . & b_{nn} \end{matrix}$$

and (3) is the symbolic form of the n equations

$$Y_r = b_{r1}x_1+b_{r2}x_2+...+b_{rn}x_n \quad (r = 1, 2,..., n).$$

We naturally, by comparison with vectors in a plane or in space, denote by $X+Y$ the '*number*' with components X_r+Y_r. But

$$X_r+Y_r = (a_{r1}+b_{r1})x_1+(a_{r2}+b_{r2})x_2+...+(a_{rn}+b_{rn})x_n,$$

and so we can write

$$X+Y = (A+B)x$$

provided that we interpret $A+B$ to be the set of n^2 numbers

$$\begin{matrix} a_{11}+b_{11} & a_{12}+b_{12} & . & . & a_{1n}+b_{1n} \\ . & . & . & . & . \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & . & . & a_{nn}+b_{nn} \end{matrix}$$

We are thus led to the idea of defining the sum of two symbols such as A and B. This idea we consider with more precision in the next section.

3. Matrices

3.1. The sets symbolized by A and B in the previous section have as many rows as columns. In the definition that follows we do not impose this restriction, but admit sets wherein the number of rows may differ from the number of columns.

DEFINITION 1. *An array of mn numbers, real or complex, the array having m rows and n columns,*

$$\begin{bmatrix} a_{11} & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & a_{2n} \\ . & . & . & . & . \\ a_{m1} & a_{m2} & . & . & a_{mn} \end{bmatrix},$$

is called a MATRIX. *When $m = n$ the array is called a* SQUARE MATRIX *of order n.*

In writing, a matrix is frequently denoted by a single letter A, or α, or by any other symbol one cares to choose. For example, a common notation for the matrix of the definition is $[a_{rs}]$. The square bracket is merely a conventional symbol (to mark the fact that we are *not* considering a determinant) and is conveniently read as 'the matrix'.

As we have seen, the idea of matrices comes from linear substitutions, such as those considered in § 2. But there is one important difference between the A of § 2 and the A that denotes a matrix. In substitutions, A is thought of as operating on some '*number*' x. The definition of a matrix deliberately omits this notion of A in relation to something else, and so leaves matrix notation open to a wider interpretation.

3.2. We now introduce a number of definitions that will make precise the meaning to be attached to such symbols as $A+B$, $A-B$, $2A$, when A and B denote matrices.

DEFINITION 2. *Two matrices A, B are* CONFORMABLE FOR ADDITION *when each has the same number of rows and each has the same number of columns.*

DEFINITION 3. ADDITION. *The sum of a matrix A, with an element a_{rs} in its r-th row and s-th column, and a matrix B, with an element b_{rs} in its r-th row and s-th column, is defined only when*

A, B are conformable for addition and is then defined as the matrix having $a_{rs}+b_{rs}$ as the element in its r-th row and s-th column.

The sum is denoted by $A+B$.

The matrices
$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad \begin{bmatrix} a_1 & 0 \\ a_2 & 0 \end{bmatrix}$$
are distinct; the first cannot be added to a square matrix of order 2, but the second can; e.g.
$$\begin{bmatrix} a_1 & 0 \\ a_2 & 0 \end{bmatrix}+\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} = \begin{bmatrix} a_1+b_1 & c_1 \\ a_2+b_2 & c_2 \end{bmatrix}.$$

DEFINITION 4. *The matrix $-A$ is that matrix whose elements are those of A multiplied by -1.*

DEFINITION 5. SUBTRACTION. *$A-B$ is defined as $A+(-B)$. It is called the difference of the two matrices.*

For example, the matrix $-A$ where A is the one-rowed matrix $[a, b]$ is, by definition, the matrix $[-a, -b]$. Again, by Definition 5,
$$\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}-\begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}+\begin{bmatrix} -a_2 & -b_2 \\ -c_2 & -d_2 \end{bmatrix},$$
and this, by Definition 3, is equal to
$$\begin{bmatrix} a_1-a_2 & b_1-b_2 \\ c_1-c_2 & d_1-d_2 \end{bmatrix}.$$

DEFINITION 6. *A matrix having every element zero is called a* NULL MATRIX, *and is written* 0.

DEFINITION 7. *The matrices A and B are said to be equal, and we write $A = B$, when the two matrices are conformable for addition and each element of A is equal to the corresponding element of B.*

It follows from Definitions 6 and 7 that '$A = B$' and '$A-B = 0$' mean the same thing, namely, each a_{rs} is equal to the corresponding b_{rs}.

DEFINITION 8. MULTIPLICATION BY A NUMBER. *When r is a number, real or complex, and A is a matrix, rA is defined to be the matrix each element of which is r times the corresponding element of A.*

For example, if

$$A = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}, \quad \text{then} \quad 3A = \begin{bmatrix} 3a_1 & 3b_1 \\ 3c_1 & 3d_1 \end{bmatrix}.$$

In virtue of Definitions 3, 5, and 8, we are justified in writing

$$2A \quad \text{instead of} \quad A+A,$$

$$3A \quad \text{instead of} \quad 5A-2A,$$

and so on. By an easy extension, which we shall leave to the reader, we can consider the sum of several matrices and write, for example, $(4+i)A$ instead of $A+A+A+(1+i)A$.

3.3. Further, since the addition and subtraction of matrices is based directly on the addition and subtraction of their elements, which are real or complex numbers, the laws that govern addition in ordinary algebra also govern the addition of matrices. The laws that govern addition and subtraction in algebra are: (i) the associative law, of which an example is

$$(a+b)+c = a+(b+c),$$

either sum being denoted by $a+b+c$;

(ii) the commutative law, of which an example is

$$a+b = b+a;$$

(iii) the distributive law, of which examples are

$$r(a+b) = ra+rb, \qquad -(a-b) = -a+b.$$

The matrix equation $(A+B)+C = A+(B+C)$ is an immediate consequence of $(a+b)+c = a+(b+c)$, where the small letters denote the elements standing in, say, the rth row and sth column of the matrices A, B, C; and either sum is denoted by $A+B+C$. Similarly, the matrix equation $A+B = B+A$ is an immediate consequence of $a+b = b+a$. Thus the associative and commutative laws are satisfied for addition and subtraction, and we may correctly write

$$\begin{aligned} A+B+(C+D) &= A+B+C+D \\ &= (B+C)+(A+D); \end{aligned}$$

and so on.

On the other hand, although we may write

$$r(A+B) = rA+rB$$

when r is a number, real or complex, on the ground that

$$r(a+b) = ra+rb$$

when r, a, b are numbers, real or complex, we have not yet assigned meanings to symbols such as $R(A+B)$, RA, RB when R, A, B are all matrices. This we shall do in the sections that follow.

4. Linear substitutions as a guide to matrix multiplication

Consider the equations

$$\left.\begin{aligned} x_1 &= a_{11}y_1+a_{12}y_2, & y_1 &= b_{11}z_1+b_{12}z_2, \\ x_2 &= a_{21}y_1+a_{22}y_2, & y_2 &= b_{21}z_1+b_{22}z_2, \end{aligned}\right\} \tag{1}$$

where the a and b are given constants. They enable us to express x_1, x_2 in terms of the a, b, and z_1, z_2; in fact,

$$\left.\begin{aligned} x_1 &= (a_{11}b_{11}+a_{12}b_{21})z_1+(a_{11}b_{12}+a_{12}b_{22})z_2, \\ x_2 &= (a_{21}b_{11}+a_{22}b_{21})z_1+(a_{21}b_{12}+a_{22}b_{22})z_2. \end{aligned}\right\} \tag{2}$$

If we introduce the notation of § 1, we may write the equations (1) as

$$x = Ay, \qquad y = Bz. \tag{1 a}$$

We can go direct from x to z and write

$$x = ABz, \tag{2 a}$$

provided we interpret AB in such a way that (2 a) is the symbolic form of equations (2); that is, we must interpret AB to be the matrix

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix}. \tag{3}$$

This gives us a direct lead to the formal definition of the product AB of two matrices A and B. But before we give this formal definition it will be convenient to define the 'scalar product' of two '*numbers*'.

5. The scalar product or inner product of two numbers

Definition 9. *Let* x, y *be two* 'numbers', *in the sense of* § 1, *having components* $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$. *Then*

$$x_1y_1+x_2y_2+\dots+x_ny_n$$

is called the INNER PRODUCT *or the* SCALAR PRODUCT *of the two numbers. If the two numbers have m, n components respectively and $m \neq n$, the inner product is not defined.*

In using this definition for descriptive purposes, we shall permit ourselves a certain degree of freedom in what we call a '*number*'. Thus, in dealing with two matrices $[a_{rs}]$ and $[b_{rs}]$, we shall speak of

$$a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32}+...+a_{1n}b_{n2} \tag{1}$$

as the inner product of the first row of $[a_{rs}]$ by the second column of $[b_{rs}]$. That is to say, we think of the first row of $[a_{rs}]$ as a '*number*' having components $(a_{11}, a_{12},..., a_{1n})$ and of the second column of $[b_{rs}]$ as a '*number*' having components $(b_{12}, b_{22},..., b_{n2})$.

6. Matrix multiplication

6.1. DEFINITION 10. *Two matrices A, B are* CONFORMABLE FOR THE PRODUCT AB *when the number of columns in A is equal to the number of rows in B.*

DEFINITION 11. PRODUCT. *The product AB is defined only when the matrices A, B are conformable for this product: it is then defined as the matrix whose element in the i-th row and k-th column is the inner product of the i-th row of A by the k-th column of B.*

It is an immediate consequence† of the definition that AB has as many rows as A and as many columns as B.

The best way of seeing the necessity of having the matrices conformal is to try the process of multiplication on two non-conformable matrices. Thus, if

$$A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix},$$

AB cannot be defined. For a row of A consists of one letter and a column of B consists of two letters, so that we cannot

---

† The reader is recommended to work out a few products for himself. One or two examples follow in §§ 6.1, 6.3.

form the inner products required by the definition. On the other hand,

$$BA = \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$= \begin{bmatrix} b_1 a_1 + c_1 a_2 \\ b_2 a_1 + c_2 a_2 \end{bmatrix},$$

a matrix having 2 rows and 1 column.

6.2. THEOREM 12. *The matrix AB is,* in general, *distinct from the matrix BA.*

As we have just seen, we can form *both* AB and BA only if the number of columns of A is equal to the number of rows of B and the number of columns of B is equal to the number of rows of A. When these products are formed the elements in the ith row and kth column are, respectively,

in AB, the inner product of the ith row of A by the kth column of B;

in BA, the inner product of the ith row of B by the kth column of A.

Consequently, the matrix AB is not, in general, the sa[illegible] matrix as BA.

6.3. Pre-multiplication and post-multiplication. Since AB and BA are usually distinct, there can be no precision in the phrase 'multiply A by B' until it is clear whether it shall mean AB or BA. Accordingly, we introduce the terms† post-multiplication and pre-multiplication (and thereafter avoid the use of such lengthy words as much as we can!). The matrix A post-multiplied by B is the matrix AB; the matrix A pre-multiplied by B is the matrix BA.

Some simple examples of multiplication are

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} = \begin{bmatrix} a\alpha+b\beta & a\gamma+b\delta \\ c\alpha+d\beta & c\gamma+d\delta \end{bmatrix},$$

$$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax+hy+gz \\ hx+by+fz \\ gx+fy+cz \end{bmatrix}.$$

---

† Sometimes one uses the terms *fore* and *aft* instead of *pre* and *post*.

6.4. The distributive law for multiplication. One of the laws that govern the product of complex numbers is the distributive law, of which an illustration is $a(b+c) = ab+ac$. This law also governs the algebra of matrices.

For convenience of setting out the work we shall consider only square matrices of a given order n, and we shall use A, B,... to denote $[a_{ik}]$, $[b_{ik}]$,....

We recall that the notation $[a_{ik}]$ sets down the element that is in the ith row and the kth column of the matrix so denoted. Accordingly, by Definition 11,

$$[a_{ik}]\times[b_{ik}] = \Big[\sum_{\lambda=1}^{n} a_{i\lambda}b_{\lambda k}\Big].$$

In this notation we may write

$$\begin{aligned} A(B+C) &= [a_{ik}]\times([b_{ik}]+[c_{ik}]) \\ &= [a_{ik}]\times[b_{ik}+c_{ik}] \quad \text{(Definition 3)} \\ &= \Big[\sum_{\lambda=1}^{n} a_{i\lambda}(b_{\lambda k}+c_{\lambda k})\Big] \quad \text{(Definition 11)} \\ &= \Big[\sum_{\lambda=1}^{n} a_{i\lambda}b_{\lambda k}\Big]+\Big[\sum_{\lambda=1}^{n} a_{i\lambda}c_{\lambda k}\Big] \quad \text{(Definition 3)} \\ &= AB+AC \quad \text{(Definition 11).} \end{aligned}$$

Similarly, we may prove that

$$(A+B)C = AC+BC.$$

6.5. The associative law for multiplication. Another law that governs the product of complex numbers is the associative law, of which an illustration is $(ab)c = a(bc)$, either being commonly denoted by the symbol abc.

We shall now consider the corresponding relations between three square matrices A, B, C, each of order n. We shall show that

$$(AB)C = A(BC)$$

and thereafter we shall use the symbol ABC to denote either of them.

Let $B = [b_{ik}]$, $C = [c_{ik}]$ and let $BC = [\gamma_{ik}]$, where, by Definition 11,

$$\gamma_{ik} = \sum_{j=1}^{n} b_{ij}c_{jk}.$$

Then $A(BC)$ is a matrix whose element in the ith row and kth column is

$$\begin{aligned}\sum_{l=1}^{n} a_{il}\gamma_{lk} &= \sum_{l=1}^{n} a_{il} \sum_{j=1}^{n} b_{lj} c_{jk} \\ &= \sum_{l=1}^{n} \sum_{j=1}^{n} a_{il} b_{lj} c_{jk}. \qquad (1)\end{aligned}$$

Similarly, let $AB = [\beta_{ik}]$, where, by Definition 11,

$$\beta_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Then $(AB)C$ is a matrix whose element in the ith row and kth column is

$$\sum_{l=1}^{n} \beta_{il} c_{lk} = \sum_{l=1}^{n} \sum_{j=1}^{n} a_{ij} b_{jl} c_{lk}. \qquad (2)$$

But the expression (2) gives the same terms as (1), though in a different order of arrangement; for example, the term corresponding to $l = 2$, $j = 3$ in (1) is the same as the term corresponding to $l = 3$, $j = 2$ in (2).

Hence $A(BC) = (AB)C$ and we may use the symbol ABC to denote either.

We use A^2, A^3,... to denote AA, AAA, etc.

6.6. The summation convention. When once the principle involved in the previous work has been grasped, it is best to use the summation convention (Chapter III), whereby

$$b_{ij} c_{jk} \quad \text{denotes} \quad \sum_{j=1}^{n} b_{ij} c_{jk}.$$

By extension,

$$a_{il} b_{lj} c_{jk} \quad \text{denotes} \quad \sum_{l=1}^{n} \sum_{j=1}^{n} a_{il} b_{lj} c_{jk},$$

and once the forms (1) and (2) of § 6.5 have been studied, it becomes clear that any repeated suffix in such an expression is a 'dummy' and may be replaced by any other; for example,

$$a_{il} b_{lj} c_{jk} \equiv a_{ij} b_{jl} c_{lk}.$$

Moreover, the use of the convention makes obvious the law of formation of the elements of a product $ABC...Z$; this product of matrices is, in fact,

$$[a_{il} b_{lj} c_{jm} \ldots z_{tk}],$$

the i, k being the only suffixes that are not dummies.

7. The commutative law for multiplication

7.1. The third law that governs the multiplication of complex numbers is the commutative law, of which an example is $ab = ba$. As we have seen (Theorem 22), this law does not hold for matrices. There is, however, an important exception.

DEFINITION 12. *The square matrix of order n that has unity in its leading diagonal places and zero elsewhere is called* THE UNIT MATRIX *of order n. It is denoted by I.*

The number n of rows and columns in I is usually clear from the context and it is rarely necessary to use distinct symbols for unit matrices of different orders.

Let C be a square matrix of order n and I the unit matrix of order n; then

$$IC = CI = C.$$

Also

$$I = I^2 = I^3 = \dots.$$

Hence I has the properties of unity in ordinary algebra. Just as we replace $1 \times x$ and $x \times 1$ by x in ordinary algebra, so we replace $I \times C$ and $C \times I$ by C in the algebra of matrices.

Further, if k is any number, real or complex, $kI.C = C.kI$ and each is equal to the matrix kC.

7.2. It may also be noted that if A is a matrix of n rows, B a matrix of n columns, and I the unit matrix of order n, then $IA = A$ and $BI = B$, even though A and B are not square matrices. If A is not square, A and I are not conformable for the product AI.

8. The division law

Ordinary algebra is governed also by the division law, which states that when the product xy is zero, either x or y, or both, must be zero. This law does not govern matrix products. Use O to denote the null matrix. Then $AO = OA = O$, but the equation $AB = O$ does not necessarily imply that A or B is the null matrix. For instance, if

$$A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} b & 2b \\ -a & -2a \end{bmatrix},$$

the product AB is the zero matrix, although neither A nor B is the zero matrix.

Again, AB may be zero and BA not zero. For example, if

$$A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} b & 0 \\ -a & 0 \end{bmatrix},$$

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad BA = \begin{bmatrix} ab & b^2 \\ -a^2 & -ab \end{bmatrix}.$$

9. Summary of previous sections

We have shown that the symbols A, B,..., representing matrices, may be added, multiplied, and, in a large measure, manipulated as though they represented ordinary numbers. The points of difference between ordinary numbers a, b and matrices A, B are

(i) whereas $ab = ba$, AB is not usually the same as BA;

(ii) whereas $ab = 0$ implies that either a or b (or both) is zero, the equation $AB = 0$ does not necessarily imply that either A or B is zero. We shall return to this point in Theorem 29.

10. The determinant of a square matrix

10.1. When A is a square matrix, the determinant that has the same elements as the matrix, and in the same places, is called the determinant of the matrix. It is usually denoted by $|A|$. Thus, if A has the element a_{ik} in the ith row and kth column, so that $A = [a_{ik}]$, then $|A|$ denotes the determinant $|a_{ik}|$.

It is an immediate consequence of Theorem 13, that if A and B are square matrices of order n, and if AB is the product of the two matrices, *the determinant of the matrix AB is equal to the product of the determinants of the matrices A and B*; that is,

$$|AB| = |A| \times |B|.$$

Equally, $$|BA| = |B| \times |A|.$$

Since $|A|$ and $|B|$ are numbers, the commutative law holds for their product and $|A| \times |B| = |B| \times |A|$.

Thus $|AB| = |BA|$, although the matrix AB is not the same as the matrix BA. This is due to the fact that the value of a *determinant* is unaltered when rows and columns are interchanged, whereas such an interchange does not leave a general matrix unaltered.

10.2. The beginner should note that in the equation $|AB| = |A| \times |B|$, of § 10.1, both A and B are matrices. It is not true that $|kA| = k|A|$, where k is a number.

For simplicity of statement, let us suppose A to be a square matrix of order three and let us suppose that $k = 2$. The theorem

$$\text{‘}\,|A+B| = |A|+|B|\,\text{’}$$

is manifestly false (Theorem 6, p. 15) and so $|2A| \neq 2|A|$. The true theorem is easily found. We have

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

say, so that

$$2A = \begin{bmatrix} 2a_{11} & 2a_{12} & 2a_{13} \\ 2a_{21} & 2a_{22} & 2a_{23} \\ 2a_{31} & 2a_{32} & 2a_{33} \end{bmatrix} \quad \text{(Definition 8).}$$

Hence $|2A| = 2^3|A|$.

The same is evident on applying § 10.1 to the matrix product $2I \times A$, where $2I$ is twice the unit matrix of order three and so is

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

EXAMPLES VII

Examples 1–4 are intended as a kind of mental arithmetic.

1. Find the matrix $A+B$ when

(i) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$;

(ii) $A = \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{bmatrix}$, $\quad B = \begin{bmatrix} 3 & -3 \\ 4 & -4 \\ 5 & -5 \end{bmatrix}$;

(iii) $A = [1 \quad 2 \quad 3]$, $\quad B = [4 \quad 5 \quad 6]$.

Ans. (i) $\begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$, (ii) $\begin{bmatrix} 4 & -4 \\ 6 & -6 \\ 8 & -8 \end{bmatrix}$, (iii) $[5 \quad 7 \quad 9]$.

2. Is it possible to define the matrix $A+B$ when

(i) A has 3 rows, B has 4 rows,
(ii) A has 3 columns, B has 4 columns,
(iii) A has 3 rows, B has 3 columns?

Ans. (i), (ii), No; (iii) only if A has 3 columns and B 3 rows.

3. Is it possible to define the matrix BA and the matrix AB when A, B have the properties of Example 2?

Ans. AB (i) if A has 4 cols.; (ii) if B has 3 rows; (iii) depends on the number of *columns* of A and the number of *rows* of B.

BA (i) if B has 3 cols.; (ii) if A has 4 rows; (iii) always.

4. Form the products AB and BA when

(i) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$; (ii) $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 5 \\ 5 & 6 \end{bmatrix}$.

Examples 5–8 deal with *quaternions* in their matrix form.

5. If ι denotes $\sqrt{(-1)}$ and I, i, j, k are defined as

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad i = \begin{bmatrix} \iota & 0 \\ 0 & -\iota \end{bmatrix}, \quad j = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 & \iota \\ \iota & 0 \end{bmatrix},$$

then $ij = k, \quad jk = i, \quad ki = j; \quad ji = -k, \quad kj = -i, \quad ik = -j.$

Further, $\quad i^2 = j^2 = k^2 = -I.$

6. If $Q = aI+bi+cj+dk$, $Q' = aI-bi-cj-dk$, a, b, c, d being numbers, real or complex, then

$$QQ' = (a^2+b^2+c^2+d^2)I.$$

7. If $P = \alpha I+\beta i+\gamma j+\delta k$, $P' = \alpha I-\beta i-\gamma j-\delta k$, α, β, γ, δ being numbers, real or complex, then

$$\begin{aligned} QPP'Q' &= Q.(\alpha^2+\beta^2+\gamma^2+\delta^2)I.Q' \\ &= QQ'.(\alpha^2+\beta^2+\gamma^2+\delta^2)I \qquad [\text{see } \S\, 7] \\ &= (a^2+b^2+c^2+d^2)I.(\alpha^2+\beta^2+\gamma^2+\delta^2)I \\ &= (a^2+b^2+c^2+d^2)(\alpha^2+\beta^2+\gamma^2+\delta^2)I \\ &= Q'P'PQ. \end{aligned}$$

8. Prove the results of Example 5 when I, i, j, k are respectively

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

9. By considering the matrices

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \iota = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

together with matrices $aI+b\iota$, $aI-b\iota$, where a, b are *real* numbers, show that the usual complex number may be regarded as a sum of matrices whose elements are real numbers. In particular, show that

$$(aI+b\iota)(aI-b\iota) = (a^2+b^2)I,$$
$$(aI+b\iota)(cI+d\iota) = (ac-bd)I+(ad+bc)\iota.$$

10. Prove that

$$[x \quad y \quad z].\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}.\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [ax^2+by^2+cz^2+2fyz+2gzx+2hxy],$$

that is, the usual quadratic form as a matrix having only one row and one column.

11. Prove that the matrix equation

$$A^2-B^2 = (A-B)(A+B)$$

is true only if $AB = BA$ and that $aA^2+2hAB+bB^2$ cannot, in general, be written as the product of two linear factors unless $AB = BA$.

12. If λ_1, λ_2 are numbers, real or complex, and A is a square matrix of order n, then

$$A^2-(\lambda_1+\lambda_2)A+\lambda_1\lambda_2 I = (A-\lambda_1 I)(A-\lambda_2 I),$$

where I is the unit matrix of order n.

HINT. The R.H.S. is $A^2-\lambda_2 AI-\lambda_1 IA+\lambda_1\lambda_2 I^2$ (by the distributive law).

13. If $f(\lambda) \equiv p_0\lambda^n+p_1\lambda^{n-1}+...+p_n$, where p_r, λ are numbers, and if $f(A)$ denotes the matrix

$$p_0 A^n+p_1 A^{n-1}+...+p_n,$$

then $$f(A) = p_0(A-\lambda_1 I)...(A-\lambda_n I),$$

where $\lambda_1,..., \lambda_n$ are the roots of the equation $f(\lambda) = 0$.

14. If $B = \lambda A+\mu I$, where λ and μ are numbers, then $BA = AB$.

The use of submatrices

15. The matrix $$P \equiv \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

may be denoted by $$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},$$

where P_{11} denotes $\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$, P_{12} denotes $\begin{bmatrix} p_{13} \\ p_{23} \end{bmatrix}$,

P_{21} denotes $[p_{31} \quad p_{32}]$, P_{22} denotes $[p_{33}]$.

Prove that, if Q, Q_{11}, etc., refer to like matrices with q_{ik} instead of p_{ik}, then

$$P+Q = \begin{bmatrix} P_{11}+Q_{11} & P_{12}+Q_{12} \\ P_{21}+Q_{21} & P_{22}+Q_{22} \end{bmatrix},$$

$$PQ = \begin{bmatrix} P_{11}Q_{11}+P_{12}Q_{21} & P_{11}Q_{12}+P_{12}Q_{22} \\ P_{21}Q_{11}+P_{22}Q_{21} & P_{21}Q_{12}+P_{22}Q_{22} \end{bmatrix}.$$

NOTE. The first 'column' of PQ as it is written above is merely a shorthand for the first two columns of the product PQ when it is obtained directly from $[p_{ik}]\times[q_{ik}]$.

16. Prove Example 8 by putting for 1 and ι in Example 5 the two-rowed matrices of Example 9.

17. If each P_{rs} and Q_{rs} is a matrix of two rows and two columns, prove that, when $r = 1, 2, 3$, $s = 1, 2, 3$,

$$[P_{rs}]\times[Q_{rs}] = \left[\sum_{k=1}^{3} P_{rk}\,Q_{ks}\right].$$

Elementary transformations of a matrix

18. I_{ij} is the matrix obtained by interchanging the ith and jth rows of the unit matrix I. Prove that the effect of pre-multiplying a square matrix A by I_{ij} will be the interchange of two rows of A, and of post-multiplying by I_{ij}, the interchange of two columns of A.

Deduce that $\qquad I_{ij}^2 = I, \qquad I_{ik}\,I_{kj}\,I_{ji} = I_{kj}.$

19. If $H = I+[h_{ij}]$, the unit matrix supplemented by an element h in the position indicated, then HA affects A by replacing row_i by $\text{row}_i + h\,\text{row}_j$ and AH affects A by replacing col._j by $\text{col.}_j + h\,\text{col.}_i$.

20. If H is a matrix of order n obtained from the unit matrix by replacing the rth unity in the principal diagonal by k, then HA is the result of multiplying the rth row of A by k, and AH is the result of multiplying the rth column of A by k.

Examples on matrix multiplication

21. Prove that the product of the two matrices

$$\begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}, \qquad \begin{bmatrix} \cos^2\phi & \cos\phi\sin\phi \\ \cos\phi\sin\phi & \sin^2\phi \end{bmatrix}$$

is zero when θ and ϕ differ by an odd multiple of $\frac{1}{2}\pi$.

22. When $(\lambda_1,\lambda_2,\lambda_3)$ and (μ_1,μ_2,μ_3) are the direction cosines of two lines l and m, prove that the product

$$\begin{bmatrix} \lambda_1^2 & \lambda_1\lambda_2 & \lambda_1\lambda_3 \\ \lambda_1\lambda_2 & \lambda_2^2 & \lambda_2\lambda_3 \\ \lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2 \end{bmatrix} \times \begin{bmatrix} \mu_1^2 & \mu_1\mu_2 & \mu_1\mu_3 \\ \mu_1\mu_2 & \mu_2^2 & \mu_2\mu_3 \\ \mu_1\mu_3 & \mu_2\mu_3 & \mu_3^2 \end{bmatrix}$$

is zero if and only if the lines l and m are perpendicular.

23. Prove that $L^2 = L$ when L denotes the first matrix of Example 22.

CHAPTER VII

RELATED MATRICES

1. The transpose of a matrix

1.1. DEFINITION 13. *If A is a matrix of n columns, the matrix which, for $r = 1, 2,..., n$, has the r-th column of A as its r-th row is called the* TRANSPOSE OF A (*or the* TRANSPOSED MATRIX OF A). *It is denoted by A'.*

If $A = A'$, A is said to be SYMMETRICAL.

The definition applies to all rectangular matrices. For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{ is the transpose of } \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix},$$

for the rows of the one are the columns of the other.

When we are considering square matrices of order n we shall, throughout this chapter, denote the matrix by writing down the element in the ith row and kth column. In this notation, if $A = [a_{ik}]$, then $A' = [a_{ki}]$; for the element in the ith row and kth column of A' is that in the kth row and ith column of A.

1.2. THEOREM 23. LAW OF REVERSAL FOR A TRANSPOSE. *If A and B are square matrices of order n,*

$$(AB)' = B'A';$$

that is to say, the transpose of the product AB is the product of the transposes in the reverse order.

In the notation indicated in § 1.1, let

$$A = [a_{ik}], \qquad B = [b_{ik}],$$

so that

$$A' = [a_{ki}], \qquad B' = [b_{ki}].$$

Then, denoting a matrix by the element in the ith row and kth column, and using the summation convention, we have

$$AB = [a_{ij} b_{jk}], \qquad (AB)' = [a_{kj} b_{ji}].$$

Now the ith row of B' is $b_{1i}, b_{2i},..., b_{ni}$ and the kth column of A' is $a_{k1}, a_{k2},..., a_{kn}$; the inner product of the two is

$$\sum_{j=1}^{n} b_{ji} a_{kj} = \sum_{j=1}^{n} a_{kj} b_{ji}.$$

Hence, on reverting to the use of the summation convention, the product of the matrices B' and A' is given by

$$\begin{aligned} B'A' = [b_{ji} a_{kj}] &= [a_{kj} b_{ji}] \\ &= (AB)'. \end{aligned}$$

COROLLARY. *If A, B,..., K are square matrices of order n, the transpose of the product $AB...K$ is the product $K'...B'A'$.*

By the theorem,

$$\begin{aligned} (ABC)' &= C'(AB)' \\ &= C'B'A', \end{aligned}$$

and so, step by step, for any number of matrices.

2. The Kronecker delta

The symbol δ_{ik} is defined by the equations

$$\delta_{ik} = 0 \text{ when } i \neq k, \qquad \delta_{ik} = 1 \text{ when } i = k.$$

It is a particular instance of a class of symbols that is extensively used in tensor calculus. In matrix algebra it is often convenient to use $[\delta_{ik}]$ to denote the unit matrix, which, as we have already said, has 1 in the leading diagonal positions (when $i = k$) and zero elsewhere.

3. The adjoint matrix and the reciprocal matrix

3.1. Transformations as a guide to a correct definition.

Let '*numbers*' X and x be connected by the equation

$$X = Ax, \tag{1}$$

where A symbolizes a matrix $[a_{ik}]$.

When $\Delta = |a_{ik}| \neq 0$, x can be expressed uniquely in terms of X (Chapter II, § 4.3); for when we multiply each equation

$$X_i = \sum_{k=1}^{n} a_{ik} x_k \tag{1 a}$$

by the co-factor A_{it} and sum for $i = 1, 2,..., n$, we obtain

$$\begin{aligned}\sum_{i=1}^{n} A_{it} X_i &= \sum_{k=1}^{n} x_k \sum_{i=1}^{n} a_{ik} A_{it} \\ &= \sum_{k=1}^{n} x_k \delta_{kt} \Delta \quad \text{(Theorem 10)} \\ &= x_t \Delta.\end{aligned}$$

That is,

$$x_t = \sum_{i=1}^{n} (A_{it}/\Delta) X_i$$

or

$$x_i = \sum_{k=1}^{n} (A_{ki}/\Delta) X_k. \tag{2 a}$$

Thus we may write equations (2 a) in the symbolic form

$$x = A^{-1}X, \tag{2}$$

which is the natural complement of (1), provided we interpret A^{-1} to mean the matrix $[A_{ki}/\Delta]$.

3.2. Formal definitions. Let $A = [a_{ik}]$ be a square matrix of order n; let Δ, or $|A|$, denote the determinant $|a_{ik}|$; and let A_{rs} denote the co-factor of a_{rs} in Δ.

DEFINITION 14. *$A = [a_{ik}]$ is a* SINGULAR MATRIX *if $\Delta = 0$; it is an* ORDINARY MATRIX *or a* NON-SINGULAR MATRIX *if $\Delta \neq 0$.*

DEFINITION 15. *The matrix $[A_{ki}]$ is the* ADJOINT, *or the* ADJUGATE, MATRIX *of $[a_{ik}]$.*

DEFINITION 16. *When $[a_{ik}]$ is a non-singular matrix, the matrix $[A_{ki}/\Delta]$ is* THE RECIPROCAL MATRIX *of $[a_{ik}]$.*

Notice that, in the last two definitions, it is A_{ki} and not A_{ik} that is to be found in the ith row and kth column of the matrix.

3.3. Properties of the reciprocal matrix. The reason for the name '*the reciprocal matrix*' may be found in the properties to be enunciated in Theorems 24 and 25.

THEOREM 24. *If $[a_{ik}]$ is a non-singular square matrix of order n and if A_{rs} is the co-factor of a_{rs} in $\Delta = |a_{ik}|$, then*

$$[a_{ik}] \times [A_{ki}/\Delta] = I, \tag{1}$$

and

$$[A_{ki}/\Delta] \times [a_{ik}] = I, \tag{2}$$

where I is the unit matrix of order n.

The element in the ith row and kth column of the product $[a_{ik}]\times[A_{ki}/\Delta]$ is

$$\frac{1}{\Delta}\sum_{j=1}^{n} a_{ij}A_{kj},$$

which (by Theorem 10,† p. 30) is zero when $i \neq k$ and is unity when $i = k$. Thus the product is $[\delta_{ik}] = I$. Hence (1) is proved.

Similarly, using now the sum convention, we have

$$\begin{aligned}[A_{ki}/\Delta]\times[a_{ik}] &= [A_{ji}a_{jk}/\Delta]\\ &= [\delta_{ik}],\end{aligned}$$

and so (2) is proved.

3.4. In virtue of Theorem 24 we are justified, when A denotes $[a_{ik}]$, in writing

$$[A_{ki}/\Delta] = A^{-1};$$

for we have shown by (1) and (2) of Theorem 24 that, with such a notation,

$$AA^{-1} = I \quad \text{and} \quad A^{-1}A = I.$$

That is to say, a matrix multiplied by its reciprocal is equal to unity (the unit matrix).

But, though we have shown that A^{-1} is *a* reciprocal of A, we have not yet shown that it is the *only* reciprocal. This we do in Theorem 25.

THEOREM 25. *If A is a non-singular square matrix, there is only one matrix which, when multiplied by A, gives the unit matrix.*

Let R be *any* matrix such that $AR = I$ and let A^{-1} denote the matrix $[A_{ki}/\Delta]$. Then, by Theorem 24, $AA^{-1} = I$, and hence

$$A(R-A^{-1}) = AR-AA^{-1} = I-I = 0.$$

It follows that (note the next two steps)

$$A^{-1}A(R-A^{-1}) = A^{-1}.0 = 0 \quad (\text{p. } 75,\ \S 6.5),$$

that is (by (2) of Theorem 24),

$$I(R-A^{-1}) = 0, \quad \text{i.e. } R-A^{-1} = 0 \quad (\text{p. } 77,\ \S 7.1).$$

Hence, if $AR = I$, R must be A^{-1}.

---

† For the precise form used here compare Example 6, p. 62.

Similarly, if $RA = I$, we have

$$(R-A^{-1})A = RA-A^{-1}A = I-I = 0,$$

and so $$(R-A^{-1})AA^{-1} = 0.A^{-1} = 0;$$

that is $$(R-A^{-1})I = 0, \quad \text{i.e. } R-A^{-1} = 0.$$

Hence, if $RA = I$, R must be A^{-1}.

3.5. In virtue of Theorem 25 we are justified in speaking of A^{-1} not merely as *a* reciprocal of A, but as *the* reciprocal of A.

Moreover, the reciprocal of the reciprocal of A is A itself, or, in symbols,

$$(A^{-1})^{-1} = A.$$

For since, by Theorem 24, $AA^{-1} = A^{-1}A = I$, it follows that A is *a* reciprocal of A^{-1} and, by Theorem 25, it must be *the* reciprocal.

4. The index law for matrices

We have already used the notations A^2, A^3,... to stand for AA, AAA,.... When r and s are positive integers it is inherent in the notation that

$$A^r \times A^s = A^{r+s}.$$

When A^{-1} denotes the reciprocal of A and s is a positive integer, we use the notation A^{-s} to denote $(A^{-1})^s$. With this notation we may write

$$A^r \times A^s = A^{r+s} \tag{1}$$

whenever r, s are positive or negative integers, provided that A^0 is interpreted as I.

We shall not prove (1) in every case; a proof of (1) when $r > 0$, $s = -t$, and $t > r > 0$ will be enough to show the method.

Let $t = r+k$, where $k > 0$. Then

$$A^r A^{-t} = A^r A^{-r-k} = A^r A^{-r} A^{-k}.$$

But $$AA^{-1} = I \quad \text{(Theorem 24)},$$

and, when $r > 1$,

$$A^r A^{-r} = A^{r-1} A A^{-1} A^{-r+1} = A^{r-1} I A^{-r+1} = A^{r-1} A^{-r+1},$$

so that $$A^r A^{-r} = A^1 A^{-1} = I.$$

Hence $$A^r A^{-t} = IA^{-k} = A^{-k} = A^{r-t}.$$

Similarly, we may prove that

$$(A^r)^s = A^{rs} \tag{2}$$

whenever r and s are positive or negative integers.

5. The law of reversal for reciprocals

THEOREM 26. *The reciprocal of a product of factors is the product of the reciprocals of the factors in the reverse order; in particular,*

$$(AB)^{-1} = B^{-1}A^{-1},$$
$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1},$$
$$(A^{-1}B^{-1})^{-1} = BA.$$

Proof. For a product of two factors, we have

$$\begin{aligned}(B^{-1}A^{-1}).(AB) &= B^{-1}A^{-1}AB \quad \text{(p. 75, §6.5)}\\ &= B^{-1}IB\\ &= B^{-1}B = I \quad \text{(p. 77, §7.1)}.\end{aligned}$$

Hence $B^{-1}A^{-1}$ is *a* reciprocal of AB and, by Theorem 25, it must be *the* reciprocal of AB.

For a product of three factors, we then have

$$\begin{aligned}(C^{-1}B^{-1}A^{-1})(ABC) &= C^{-1}IC\\ &= C^{-1}C = I,\end{aligned}$$

and so on for any number of factors.

THEOREM 27. *The operations of reciprocating and transposing are commutative; that is,*

$$(A')^{-1} = (A^{-1})'.$$

Proof. The matrix A^{-1} is, by definition,

$$[A_{ki}/\Delta],$$

so that
$$(A^{-1})' = [A_{ik}/\Delta].$$

Moreover,
$$A' = [a_{ki}]$$

and so, by the definition of a product and the results of Theorem 10,
$$A'.(A^{-1})' = [\delta_{ik}] = I.$$

Hence $(A^{-1})'$ is a reciprocal of A' and, by Theorem 25, it must be *the* reciprocal.

6. Singular matrices

If A is a singular matrix, its determinant is zero and so we cannot form the reciprocal of A, though we can, of course, form the adjugate, or adjoint, matrix. Moreover,

THEOREM 28. *The product of a singular matrix by its adjoint is the zero matrix.*

Proof. If $A = [a_{ik}]$ and $|a_{ik}| = 0$, then

$$\sum_{j=1}^{n} a_{ij} A_{kj} = 0, \qquad \sum_{j=1}^{n} A_{ji} a_{jk} = 0$$

both when $i \neq k$ and when $i = k$. Hence the products

$$[a_{ik}] \times [A_{ki}] \quad \text{and} \quad [A_{ki}] \times [a_{ik}]$$

are both of them null matrices.

7. The division law for non-singular matrices

7.1. THEOREM 29. (i) *If the matrix A is non-singular, the equation $AB = 0$ implies $B = 0$.*

(ii) *If the matrix B is non-singular, the equation $AB = 0$ implies $A = 0$.*

Proof. (i) Since the matrix A is non-singular, it has a reciprocal A^{-1}, and $A^{-1}A = I$.

Since $AB = 0$, it follows that

$$A^{-1}AB = A^{-1} \times 0 = 0.$$

But $A^{-1}AB = IB = B$, and so $B = 0$.

(ii) Since B is non-singular, it has a reciprocal B^{-1}, and $BB^{-1} = I$.

Since $AB = 0$, it follows that

$$ABB^{-1} = 0 \times B^{-1} = 0.$$

But $ABB^{-1} = AI = A$, and so $A = 0$. .

7.2. The division law for matrices thus takes the form

'*If $AB = 0$, then either $A = 0$, or $B = 0$, or* BOTH *A and B are singular matrices.*'

7.3. THEOREM 30. *If A is a given matrix and B is a non-singular matrix (both being square matrices of order n), there is one and only one matrix X such that*

$$A = BX, \tag{1}$$

and there is one and only one matrix Y such that

$$A = YB. \tag{2}$$

Moreover, $\qquad X = B^{-1}A, \qquad Y = AB^{-1}.$

Proof. We see at once that $X = B^{-1}A$ satisfies (1); for $BB^{-1}A = IA = A$. Moreover, if $A = BR$,

$$0 = BR-A = B(R-B^{-1}A),$$

and

$$\begin{aligned} 0 &= B^{-1}B(R-B^{-1}A) \\ &= I(R-B^{-1}A), \end{aligned}$$

so that $R-B^{-1}A$ is the zero matrix.

Similarly, $Y = AB^{-1}$ is the only matrix that satisfies (2).

7.4. The theorem remains true in certain circumstances when A and B are not square matrices of the same order.

COROLLARY. *If A is a matrix of m rows and n columns, B a non-singular square matrix of order m, then*

$$A = BX$$

has the unique solution $X = B^{-1}A$.

If A is a matrix of m rows and n columns, C a non-singular square matrix of order n, then

$$A = YC$$

has the unique solution $Y = AC^{-1}$.

Since B^{-1} has m columns and A has m rows, B^{-1} and A are conformable for the product $B^{-1}A$, which is a matrix of m rows and n columns. Also

$$B.B^{-1}A = IA = A,$$

where I is the unit matrix of order m.

Further, if $A = BR$, R must be a matrix of m rows and n columns, and the first part of the corollary may be proved as we proved the theorem.

The second part of the corollary may be proved in the same way.

7.5. The corollary is particularly useful in writing down the solution of linear equations. If $A = [a_{ik}]$ is a non-singular matrix of order n, the set of equations

$$b_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1,..., n) \tag{3}$$

may be written in the matrix form $b = Ax$, where b, x stand for matrices with a single column.

The set of equations

$$b_i = \sum_{k=1}^{n} a_{ki} y_k \quad (i = 1,..., n) \tag{4}$$

may be written in the matrix form $b' = y'A$, where b', y' stand for matrices with a single row.

The solutions of the sets of equations are

$$x = A^{-1}b, \qquad y' = b'A^{-1}.$$

An interesting application is given in Example 11, p. 92.

EXAMPLES VIII

1. When $A = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$, $B = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}$,

form the products AB, BA, $A'B'$, $B'A'$ and verify the result enunciated in Theorem 23, p. 83.

HINT. Use $X = ax+hy+gz$, $Y = hx+by+fz$, $Z = gx+fy+cz$.

2. Form these products when A, B are square matrices of order 2, neither being symmetrical.

3. When A, B are the matrices of Example 1, form the matrices A^{-1}, B^{-1} and verify by direct evaluations that $(AB)^{-1} = B^{-1}A^{-1}$.

4. Prove Theorem 27, that $(A')^{-1} = (A^{-1})'$, by considering Theorem 23 in the particular case when $B = A^{-1}$.

5. Prove that the product of a matrix by its transpose is a symmetrical matrix.

HINT. Use Theorem 23 and consider the transpose of the matrix AA'.

6. The product of the matrix $[a_{ik}]$ by its adjoint is equal to $|a_{ik}|I$.

Use of submatrices (compare Example 15, p. 81)

7. Let $L = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, $M = \begin{bmatrix} \mu & 1 \\ 0 & \mu \end{bmatrix}$, $N = [\nu]$, where λ, μ, ν are non-zero numbers, and let

$$A = \begin{bmatrix} L & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & N \end{bmatrix}.$$

Prove that

$$A^2 = \begin{bmatrix} L^2 & 0 & 0 \\ 0 & M^2 & 0 \\ 0 & 0 & N^2 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} L^{-1} & 0 & 0 \\ 0 & M^{-1} & 0 \\ 0 & 0 & N^{-1} \end{bmatrix}.$$

HINT. In view of Theorem 25, the last part is proved by showing that the product of A and the last matrix is I.

8. Prove that if $f(A) = p_0 A^n + p_1 A^{n-1} + ... + p_n$, where $p_0, ..., p_n$ are numbers, then $f(A)$ may be written as a matrix

$$\begin{bmatrix} f(L) & 0 & 0 \\ 0 & f(M) & 0 \\ 0 & 0 & f(N) \end{bmatrix},$$

and that, when $f(A)$ is non-singular, $\{f(A)\}^{-1}$ or $1/f(A)$ can be written in the same matrix form, but with $1/f(L)$, $1/f(M)$, $1/f(N)$ instead of $f(L)$, $f(M)$, $f(N)$.

9. If $f(A)$, $g(A)$ are two polynomials in A and if $g(A)$ is non-singular, prove the extension of Example 8 in which f is replaced by f/g.

Miscellaneous exercises

10. Solve equations (4), p. 91, in the two forms

$$y = (A')^{-1}b, \qquad y' = b'A^{-1}$$

and then, using Theorem 23, prove that

$$(A')^{-1} = (A^{-1})'.$$

11. A is a non-singular matrix of order n, x and y are single-column matrices with n rows, l' and m' are single-row matrices with n columns;

$$y = Ax, \qquad l'x = m'y.$$

Prove that $m' = l'A^{-1}$, $m = (A^{-1})'l$.

Interpret this result when a linear transformation $y_i = \sum_{k=1}^{3} a_{ik} x_k$ $(i = 1, 2, 3)$ changes $l_1 x_1 + l_2 x_2 + l_3 x_3$ into $m_1 y_1 + m_2 y_2 + m_3 y_3$ and (x_1, x_2, x_3), (y_1, y_2, y_3) are regarded as homogeneous point-coordinates in a plane.

12. If the elements a_{ik} of A are real, and if $AA' = 0$, then $A = 0$.

13. The elements a_{ik} of A are complex, the elements $\bar{a}_{ik}$ of $\bar{A}$ are the conjugate complexes of a_{ik}, and $A\bar{A}' = 0$. Prove that $A = \bar{A} = 0$.

14. A and B are square matrices of the same order; A is symmetrical. Prove that $B'AB$ is symmetrical. [Chapter X, § 4.1, contains a proof of this.]

CHAPTER VIII

THE RANK OF A MATRIX

[Section 6 is of less general interest than the remainder of the chapter. It may well be omitted on a first reading.]

1. Definition of rank

1.1. The minors of a matrix. Suppose we are given any matrix A, whether square or not. From this matrix delete all elements save those in a certain r rows and r columns. When $r > 1$, the elements that remain form a square matrix of order r and the determinant of this matrix is called a minor of A of order r. Each individual element of A is, when considered in this connexion, called a minor of A of order 1. For example,

$$\text{the determinant} \quad \begin{vmatrix} a & c \\ g & i \end{vmatrix} \quad \text{is a minor of} \quad \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

of order 2; a, b,... are minors of order 1.

1.2. Now any minor of order $r+1$ $(r \geqslant 1)$ can be expanded by its first row (Theorem 1) and so can be written as a sum of multiples of minors of order r. Hence, if *every* minor of order r is zero, then every minor of order $r+1$ must also be zero.

The converse is not true; for instance, in the example given in § 1.1, the only minor of order 3 is the determinant that contains all the elements of the matrix, and this is zero if $a = b$, $d = e$, and $g = h$; but the minor $ai-gc$, of order 2, is not necessarily zero.

1.3. Rank. Unless every single element of a matrix A is zero, there will be at least one set of minors of the same order which do not *all* vanish. For any given matrix with n rows and m columns, other than the matrix with every element zero, there is, by the argument of § 1.2, a definite positive integer ρ such that

EITHER ρ is less than both m and n, not all minors of order ρ vanish, but all minors of order $\rho+1$ vanish, OR ρ is equal to[†] $\min(m,n)$, not all minors of order ρ vanish, and no minor of order $\rho+1$ can be formed from the matrix.

This number ρ is called the RANK of the matrix. But the definition can be made much more compact and we shall take as our working definition the following:

DEFINITION 17. *A matrix has rank r when r is the largest integer for which the statement 'not* ALL *minors of order r are zero' is valid.*

For a non-singular square matrix of order n, the rank is n; for a singular square matrix of order n, the rank is less than n.

It is sometimes convenient to consider the null matrix, all of whose elements are zero, as being of zero rank.

2. Linear dependence

2.1. In the remainder of this chapter we shall be concerned with linear dependence and its contrary, linear independence; in particular, we shall be concerned with the possibility of expressing one row of a matrix as a sum of multiples of certain other rows. We first make precise the meanings we shall attach to these terms.

Let $a_1,\ldots, a_m$ be '*numbers*', in the sense of Chapter VI, each having n components, real or complex numbers. Let the components be

$$(a_{11},\ldots, a_{1n}), \quad \ldots, \cdot \quad (a_{m1},\ldots, a_{mn}).$$

Let F be any field[‡] of numbers. Then $a_1,\ldots, a_m$ are said to be linearly dependent with respect to F if there is an equation

$$l_1 a_1+\ldots+l_m a_m = 0, \tag{1}$$

wherein all the l's belong to F and not all of them are zero. The equation (1) implies the n equations

$$l_1 a_{1r}+\ldots+l_m a_{mr} = 0 \quad (r = 1,\ldots, n).$$

The contrary of linear dependence with respect to F is linear independence with respect to F.

The '*number*' a_1 is said to be a SUM OF MULTIPLES of the

[†] When $m = n$, $\min(m,n) = m = n$;
when $m \neq n$, $\min(m,n) =$ the smaller of m and n.

[‡] Compare Preliminary Note, p. 2.

'*numbers*' $a_2,\ldots, a_m$ with respect to F if there is an equation of the form

$$a_1 = l_2 a_2 + \ldots + l_m a_m, \tag{2}$$

wherein all the l's belong to F. The equation (2) implies the n equations

$$a_{1r} = l_2 a_{2r} + \ldots + l_m a_{mr} \quad (r = 1,\ldots, n).$$

2.2. Unless the contrary is stated we shall take F to be the field of all numbers, real or complex. Moreover, we shall use the phrases 'linearly dependent', 'linearly independent', 'is a sum of multiples' to imply that each property holds with respect to the field of all numbers, real or complex. For example, 'the "*number*" a is a sum of multiples of the "*numbers*" b and c' will imply an equation of the form $a = l_1 b + l_2 c$, where l_1 and l_2 are not necessarily integers but may be any numbers, real or complex.

As on previous occasions, we shall allow ourselves a wide interpretation of what we shall call a '*number*'; for example, in Theorem 31, which follows, we consider each row of a matrix as a '*number*'.

3. Rank and linear dependence

3.1. THEOREM 31. *If A is a matrix of rank r, where $r \geqslant 1$, and if A has more than r rows, we can select r rows of A and express every other row as the sum of multiples of these r rows.*

Let A have m rows and n columns. Then there are two possibilities to be considered: either (i) $r = n$ and $m > n$, or (ii) $r < n$ and, also, $r < m$.

(i) Let $r = n$ and $m > n$.

There is at least one non-zero determinant of order n that can be formed from A. Taking letters for columns and suffixes for rows, let us label the elements of A so that one such non-zero determinant is denoted by

$$\Delta \equiv \begin{vmatrix} a_1 & b_1 & . & . & k_1 \\ a_2 & b_2 & . & . & k_2 \\ . & . & . & . & . \\ a_n & b_n & . & . & k_n \end{vmatrix}$$

With this labelling, let $n+\theta$ be the suffix of any other row.

Then (Chap. II, § 4.3) the n equations

$$\left.\begin{aligned} a_1 l_1+a_2 l_2+\ldots+a_n l_n &= a_{n+\theta}, \\ b_1 l_1+b_2 l_2+\ldots+b_n l_n &= b_{n+\theta}, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \quad \cdot \quad \cdot \quad \cdot \\ k_1 l_1+k_2 l_2+\ldots+k_n l_n &= k_{n+\theta} \end{aligned}\right\} \qquad (1)$$

have a unique solution in $l_1, l_2,\ldots, l_n$ given by

$$l_1 = (a_{n+\theta} b_2 \ldots k_n)/(a_1 b_2 \ldots k_n),$$
$$l_2 = (a_1 b_{n+\theta} c_3 \ldots k_n)/(a_1 b_2 \ldots k_n),$$

and so on. Hence the row† with suffix $n+\theta$ can be expressed as a sum of multiples of the n rows of A that occur in Δ. This proves the theorem when $r = n < m$.

Notice that the argument breaks down if $\Delta = 0$.

(ii) Let $r < n$ and, also, $r < m$.

There is at least one non-zero minor of order r. As in (i), take letters for columns and suffixes for rows, and label the elements of A so that one non-zero minor of order r is denoted by

$$M = \begin{vmatrix} a_1 & b_1 & \cdot & \cdot & p_1 \\ a_2 & b_2 & \cdot & \cdot & p_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_r & b_r & \cdot & \cdot & p_r \end{vmatrix}.$$

With this labelling, let the remaining rows of A have suffixes $r+1$, $r+2,\ldots, m$.

Consider now the determinant

$$D = \begin{vmatrix} a_1 & \cdot & \cdot & p_1 & \alpha_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_r & \cdot & \cdot & p_r & \alpha_r \\ a_{r+\theta} & \cdot & \cdot & p_{r+\theta} & \alpha_{r+\theta} \end{vmatrix},$$

where θ is any integer from 1 to $m-r$ and where α is the letter of ANY column of A.

If α is one of the letters $a,\ldots, p$, then D has two columns identical and so is zero.

If α is not one of the letters $a,\ldots, p$, then $\pm D$ is a minor of A of order $r+1$; and since the rank of A is r, D is again zero.

---

† Using ρ_t to denote the row whose letters have suffix t, we may write equations (1) as

$$\rho_{n+\theta} = l_1 \rho_1 + l_2 \rho_2 + \ldots + l_n \rho_n.$$

Hence D is zero when α is the letter of ANY column of A, and if we expand D by its last column, we obtain the equation

$$M\alpha_{r+\theta}+\lambda_1\alpha_1+\lambda_2\alpha_2+...+\lambda_r\alpha_r = 0, \tag{2}$$

where $M \neq 0$ and where M, $\lambda_1,..., \lambda_r$ are all independent of α. Hence (2) expresses the row with suffix $r+\theta$ as a sum of multiples of the r rows in M; symbolically,

$$\rho_{r+\theta} = -\frac{\lambda_1}{M}\rho_1-\frac{\lambda_2}{M}\rho_2-...-\frac{\lambda_r}{M}\rho_r.$$

This proves the theorem when r is less than m and less than n.

Notice that the argument breaks down if $M = 0$.

3.2. THEOREM 32. *If A is a matrix of rank r, it is impossible to select q rows of A, where $q < r$, and to express every other row as a sum of multiples of these q rows.*

Suppose that, contrary to the theorem, we can select q rows of A and then express every other row as a sum of multiples of them.

Using suffixes for rows and letters for columns, let us label the elements of A so that the selected q rows are denoted by $\rho_1, \rho_2,..., \rho_q$. Then, for $k = 1,..., m$, where m is the number of rows in A, there are constants λ_{tk} such that

$$\rho_k = \lambda_{1k}\rho_1+...+\lambda_{qk}\rho_q; \tag{3}$$

for if $k = 1,..., q$, (3) is satisfied when $\lambda_{tk} = \delta_{tk}$,

and if $k > q$, (3) expresses our hypothesis in symbols.

Now consider ANY (arbitrary) minor of A of order r. Let the letters of its columns be $\alpha,..., \theta$ and let the suffixes of its rows be $k_1,..., k_r$. Then this minor is

$$\begin{vmatrix} \lambda_{1k_1}\alpha_1+...+\lambda_{qk_1}\alpha_q & . & . & \lambda_{1k_1}\theta_1+...+\lambda_{qk_1}\theta_q \\ . & . & . & . \\ \lambda_{1k_r}\alpha_1+...+\lambda_{qk_r}\alpha_q & . & . & \lambda_{1k_r}\theta_1+...+\lambda_{qk_r}\theta_q \end{vmatrix},$$

which is the product by rows of the two determinants

$$\begin{vmatrix} \lambda_{1k_1} & . & . & \lambda_{qk_1} & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ \lambda_{1k_r} & . & . & \lambda_{qk_r} & 0 & . & . & 0 \end{vmatrix}, \qquad \begin{vmatrix} \alpha_1 & . & . & \alpha_q & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ \theta_1 & . & . & \theta_q & 0 & . & . & 0 \end{vmatrix},$$

each of which has $r-q$ columns of zeros.

Accordingly, if our supposition were true, every minor of A of order r would be zero. But this is not so, for since the rank of A is r, there is at least one minor of A of order r that is not zero. Hence our supposition is not true and the theorem is established.

4. Non-homogeneous linear equations

4.1. We now consider

$$\sum_{t=1}^{n} a_{it} x_t = b_i \quad (i = 1,..., m), \tag{1}$$

a set of m linear equations in n unknowns, $x_1,..., x_n$. Such a set of equations may either be consistent, that is, at least one set of values of x may be found to satisfy all the equations, or they may be inconsistent, that is, no set of values of x can be found to satisfy all the equations. To determine whether the equations are consistent or inconsistent we consider the matrices

$$A = \begin{bmatrix} a_{11} & . & . & a_{1n} \\ . & . & . & . \\ a_{m1} & . & . & a_{mn} \end{bmatrix}, \qquad B = \begin{bmatrix} a_{11} & . & . & a_{1n} & b_1 \\ . & . & . & . & . \\ a_{m1} & . & . & a_{mn} & b_m \end{bmatrix}.$$

We call B the augmented matrix of A.

Let the ranks of A, B be r, r' respectively. Then, since every minor of A is a minor of B, either $r = r'$ or $r < r'$ [if A has a non-zero minor of order r, then so has B; but the converse is not necessarily true].

We shall prove that *the equations are consistent when* $r = r'$ *and are inconsistent when* $r < r'$.

4.2. Let $r < r'$. Since $r' \leqslant m$, $r < m$. We can select r rows of A and express every other row of A as a sum of multiples of these r rows.

Let us number the equations (1) so that the selected r rows become the rows with suffixes $1,..., r$. Then we have equations of the form

$$a_{r+\theta,t} = \lambda_{1\theta} a_{1t} + ... + \lambda_{r\theta} a_{rt} \quad (t = 1,..., n) \tag{2}$$

wherein the λ's are independent of t.

Let us now make the hypothesis that the equations (1) are consistent. Then, on multiplying (2) by x_t and summing from $t = 1$ to $t = n$, we get

$$b_{r+\theta} = \lambda_{1\theta} b_1 + \ldots + \lambda_{r\theta} b_r. \tag{3}$$

The equations (2) and (3) together imply that we can express any row of B as a sum of multiples of the first r rows, which is contrary to the fact that the rank of B is r' (Theorem 32). Hence the equations (1) cannot be consistent.

4.3. Let $r' = r$. If $r = m$, (1) may be written as (5) below.

If $r < m$, we can select r rows of B and express every other row as a sum of multiples of these r rows. As before, number the equations so that these r rows correspond to $i = 1,\ldots, r$. Then every set of x that satisfies

$$\sum_{t=1}^{n} a_{it} x_t = b_i \quad (i = 1,\ldots, r) \tag{4}$$

satisfies *all* the equations of (1).

Now let $\mathbf{a}$ denote the matrix of the coefficients of x in (4). Then, as we shall prove, at least one minor of $\mathbf{a}$ of order r must have a value distinct from zero. For, since the kth row of A can be written as

$$\lambda_{1k} \rho_1 + \ldots + \lambda_{rk} \rho_r,$$

every minor of A of order r is the product of a determinant $|\lambda_{ik'}|$ by a determinant $|a_{ik''}|$, wherein i has the values $1,\ldots, r$ (put $q = r$ in the work of § 3.2); that is, every minor of A of order r has a minor of $\mathbf{a}$ of order r as a factor. But at least one minor of A of order r is not zero and hence one minor of $\mathbf{a}$ of order r is not zero.

Let the suffixes of the variables x be numbered so that the first r columns of $\mathbf{a}$ yield a non-zero minor of order r. Then, in this notation, the equations (4) become

$$\sum_{t=1}^{r} a_{it} x_t = b_i - \sum_{t=r+1}^{n} a_{it} x_t \quad (i = 1,\ldots, r), \tag{5}$$

wherein the determinant of the coefficients on the left-hand side is not zero. The summation on the right-hand side is present only when $n > r$.

If $n > r$, we may give $x_{r+1},\ldots, x_n$ arbitrary values and then solve equations (5) uniquely for $x_1,\ldots, x_r$. Hence, if $r = r'$ and

$n > r$, the equations (1) are consistent and certain sets of $n-r$ of the variables x can be given arbitrary values.[†]

If $n = r$, the equations (5), and so also the equations (1), are consistent and have a unique solution. This unique solution may be written symbolically (Chap. VII, § 7.5) as

$$x = A_1^{-1}b,$$

where A_1 is the matrix of the coefficients on the left-hand side of (5), b is the single-column matrix with elements $b_1,\ldots, b_n$, and x is the single-column matrix with elements $x_1,\ldots, x_n$.

5. Homogeneous linear equations

The equations

$$\sum_{t=1}^{n} a_{it}x_t = 0 \quad (i = 1,\ldots, m), \tag{6}$$

which form a set of m homogeneous equations in n unknowns $x_1,\ldots, x_n$, may be considered as an example of equations (1) when all the b_i are zero. The results of § 4 are immediately applicable.

Since all b_i are zero, the rank of B is equal to the rank of A, and so equations (6) are always consistent. But if r, the rank of A, is equal to n, then, in the notation of § 4.3, the only solution of equations (6) is given by

$$x = A_1^{-1}b, \qquad b = 0.$$

Hence, when $r = n$ the only solution of (6) is given by $x = 0$, that is, $x_1 = x_2 = \ldots = x_n = 0$.

When $r < n$ we can, as when we obtained equations (5) of § 4.3, reduce the solution of (6) to the solution of

$$\sum_{t=1}^{r} a_{it}x_t = -\sum_{t=r+1}^{n} a_{it}x_t \quad (i = 1,\ldots, r), \tag{7}$$

wherein the determinant of the coefficients on the left-hand side is not zero.[‡] In this case, all solutions of (6) are given by

---

[†] Not *all* sets of $n-r$ of the variables may be given arbitrary values; the criterion is that the coefficients of the remaining r variables should provide a non-zero determinant of order r.

[‡] The notation of (7) is not necessarily that of (6): the labelling of the elements of A and the numbering of the variables x has followed the procedure laid down in § 4.3.

assigning arbitrary values to $x_{r+1},\dots, x_n$ and then solving (7) for $x_1,\dots, x_r$. Once the values of $x_{r+1},\dots, x_n$ have been assigned, the equations (7) determine $x_1,\dots, x_r$ uniquely.

In particular, if there are more variables than equations, i.e. $n > m$, then $r < n$ and the equations have a solution other than $x_1 = \dots = x_n = 0$.

6. Fundamental sets of solutions

6.1. When r, the rank of A, is less than n, we obtain $n-r$ distinct solutions of equations (6) by solving equations (7) for the $n-r$ sets of values

$$\left.\begin{array}{llll} x_{r+1} = 1, & x_{r+1} = 0, & \dots, & x_{r+1} = 0, \\ x_{r+2} = 0, & x_{r+2} = 1, & \dots, & x_{r+2} = 0, \\ \cdot\quad\cdot\quad\cdot & \cdot\quad\cdot\quad\cdot & \cdot\quad\cdot & \cdot\quad\cdot\quad\cdot \\ x_n = 0, & x_n = 0, & \dots, & x_n = 1. \end{array}\right\} \tag{8}$$

We shall use $\mathbf{X}$ to denote the single-column matrix whose elements are $x_1,\dots, x_n$. The particular solutions of (6) obtained from the sets of values (8) we denote by

$$\mathbf{X}_k \quad (k = 1, 2,\dots, n-r). \tag{9}$$

As we shall prove in § 6.2, the general solution $\mathbf{X}$ of (6), which is obtained by giving $x_{r+1},\dots, x_n$ arbitrary values in (7), is furnished by the formula

$$\mathbf{X} = \sum_{k=1}^{n-r} x_{r+k}\, \mathbf{X}_k. \tag{10}$$

That is to say, (*a*) every solution of (6) is a sum of multiples of the particular solutions $\mathbf{X}_k$. Moreover, (*b*) no one $\mathbf{X}_k$ is a sum of multiples of the other $\mathbf{X}_k$. A set of solutions that has the properties (*a*) and (*b*) is called a FUNDAMENTAL SET OF SOLUTIONS.

There is more than one fundamental set of solutions. If we replace the numbers on the right of the equality signs in (8) by the elements of any non-singular matrix of order $n-r$, we are led to a fundamental set of solutions of (6). If we replace these numbers by the elements of a singular matrix of order $n-r$, we are led to $n-r$ solutions of (6) that do not form a fundamental set of solutions. This we shall prove in § 6.4.

6.2. Proof† of the statements made in §6.1. We are concerned with a set of m homogeneous linear equations in n unknowns when the matrix of the coefficients is of rank r and $r < n$. As we have seen, every solution of such a set is a solution of a set of equations that may be written in the form

$$\sum_{t=1}^{r} a_{it} x_t = -\sum_{t=r+1}^{n} a_{it} x_t \quad (i = 1,..., r), \tag{7}$$

wherein the determinant of the coefficients on the left-hand side is not zero. We shall, therefore, no longer consider the original equations but shall consider only (7).

As in §6.1, we use $\mathbf{X}$ to denote $x_1,..., x_n$ considered as the row elements of a matrix of one column, $\mathbf{X}_k$ to denote the particular solutions of equations (7) obtained from the sets of values (8). Further, we use $\mathbf{I}_k$ to denote the kth column of (8) considered as a one-column matrix, that is, $\mathbf{I}_k$ has unity in its kth row and zero in the remaining $n-r-1$ rows. Let A_1 denote the matrix of the coefficients on the left-hand side of equations (7) and let A_1^{-1} denote the reciprocal of A_1.

In one step of the proof we use $\mathbf{X}_a^{a+s}$ to denote $x_a,..., x_{a+s}$ considered as the elements of a one-column matrix.

The general solution of (7) is obtained by giving arbitrary values to $x_{r+1},..., x_n$ and then solving the equations (uniquely, since A_1 is non-singular) for $x_1,..., x_r$. This solution may be symbolized, on using the device of sub-matrices, as

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1^r \\ \mathbf{X}_{r+1}^n \end{bmatrix} = \begin{bmatrix} -A_1^{-1} \sum\limits_{k=1}^{n-r} \begin{bmatrix} a_{1,r+k} \\ \cdot \quad \cdot \quad \cdot \\ a_{r,r+k} \end{bmatrix} x_{r+k} \\ \sum\limits_{k=1}^{n-r} \mathbf{I}_k x_{r+k} \end{bmatrix}.$$

But the last matrix is

$$\sum_{k=1}^{n-r} x_{r+k} \begin{bmatrix} -A_1^{-1} \begin{bmatrix} a_{1,r+k} \\ \cdot \quad \cdot \quad \cdot \\ a_{r,r+k} \end{bmatrix} \\ \mathbf{I}_k \end{bmatrix} = \sum_{k=1}^{n-r} x_{r+k} \mathbf{X}_k,$$

---

† Many readers will prefer to omit these proofs, at least on a first reading. The remainder of § 6 does not make easy reading for a beginner.

and so we have proved that every solution of (7) may be represented by the formula

$$\mathbf{X} = \sum_{k=1}^{n-r} x_{r+k}\,\mathbf{X}_k. \tag{10}$$

6.3. A fundamental set contains $n-r$ solutions.

6.31. We first prove that a set of less than $n-r$ solutions cannot form a fundamental set.

Let $\mathbf{Y}_1,\ldots, \mathbf{Y}_\rho$, where $\rho < n-r$, be any ρ distinct solutions of (7). Suppose, contrary to what we wish to prove, that $\mathbf{Y}_i$ form a fundamental set. Then there are constants λ_{ik} such that

$$\mathbf{X}_k = \sum_{i=1}^{\rho} \lambda_{ik}\mathbf{Y}_i \quad (k = 1,\ldots, n-r),$$

where the $\mathbf{X}_k$ are the solutions of § 6.1 (9). By Theorem 31 applied to the matrix $[\lambda_{ik}]$, whose rank cannot exceed ρ, we can express every $\mathbf{X}_k$ as a sum of multiples of ρ of them (at most). But, as we see from the table of values (8), this is impossible; and so our supposition is not a true one and the $\mathbf{Y}_i$ cannot form a fundamental set.

6.32. We next prove that in any set of more than $n-r$ solutions one solution (at least) is a sum of multiples of the remainder.

Let $\mathbf{Y}_1,\ldots, \mathbf{Y}_\rho$, where $\rho > n-r$, be any ρ distinct solutions of (7). Then, by (10),

$$\mathbf{Y}_i = \sum_{k=1}^{n-r} \mu_{ik}\mathbf{X}_k \quad (i = 1,\ldots, \rho),$$

where μ_{ik} is the value of x_{r+k} in $\mathbf{Y}_i$. Exactly as in § 6.31, we can express every $\mathbf{Y}_i$ as a sum of multiples of $n-r$ of them (at most).

6.33. It follows from the definition of a fundamental set and from §§ 6.31 and 6.32 that a fundamental set must contain $n-r$ solutions.

6.4. Proof of statements made in § 6.1 (*continued*).

6.41. If $[c_{ik}]$ is a non-singular matrix of order $n-r$, and if $\mathbf{X}'_i$ is the solution of (7) obtained by taking the values

$$x_{r+1} = c_{i1},\quad x_{r+2} = c_{i2},\quad \ldots,\quad x_n = c_{i,n-r},$$

then, by what we have proved in § 6.2,

$$\mathbf{X}'_i = \sum_{k=1}^{n-r} c_{ik}\,\mathbf{X}_k \quad (i = 1,\ldots, n-r).$$

On solving these equations (Chap. VII, § 3.1), we get

$$\mathbf{X}_i = \sum_{k=1}^{n-r} \gamma_{ik}\,\mathbf{X}'_k \quad (i = 1,\ldots, n-r), \tag{11}$$

where $[\gamma_{ik}]$ is the reciprocal of $[c_{ik}]$.

Hence, by (10), every solution of (7) may be written in the form

$$\mathbf{X} = \sum_{i=1}^{n-r} x_{r+i} \sum_{k=1}^{n-r} \gamma_{ik}\,\mathbf{X}'_k$$
$$= \sum_{k=1}^{n-r}\Big(\sum_{i=1}^{n-r} x_{r+i}\,\gamma_{ik}\Big)\mathbf{X}'_k.$$

Moreover, it is not possible to express any one $\mathbf{X}'_k$ as a sum of multiples of the remainder. For if it were, (11) would give equations of the form†

$$\mathbf{X}_i = \sum_{k=1}^{n-r-1} \beta_{ik}\,\mathbf{X}'_k \quad (i = 1,\ldots, n-r)$$

and these equations, by the argument of § 6.31, would imply that one $\mathbf{X}_i$ could be expressed as a sum of multiples of the others.

Hence the solutions $\mathbf{X}'_i$ form a fundamental set.

6.42. If $[c_{ik}]$ is a singular matrix of order $n-r$, then, with a suitable labelling of its elements, there are constants $\lambda_{\theta s}$ such that

$$c_{\rho+\theta,k} = \sum_{s=1}^{\rho} \lambda_{\theta s}\, c_{sk} \quad (k = 1,\ldots, n-r),$$

where ρ is the rank of the matrix, and so $\rho < n-r$.

But, by (10),

$$\mathbf{X}'_{\rho+\theta} = \sum_{k=1}^{n-r} c_{\rho+\theta,k}\,\mathbf{X}_k \quad (\theta = 1,\ldots, n-r-\rho)$$

and

$$\mathbf{X}'_s = \sum_{k=1}^{n-r} c_{sk}\,\mathbf{X}_k \quad (s = 1,\ldots, \rho).$$

---

† We have supposed the solutions to be labelled so that $\mathbf{X}'_{n-r}$ can be written as a sum of multiples of the rest.

Hence there is an equation

$$\mathbf{X}'_{\rho+\theta} = \sum_{s=1}^{\rho} \lambda_{\theta s} \mathbf{X}'_s$$

and so the $\mathbf{X}'_i$ cannot form a fundamental set of solutions.

6.5. A set of solutions such that no one can be expressed as a sum of multiples of the others is said to be LINEARLY INDEPENDENT.

By what we have proved in §§ 6.41 and 6.42, a set of solutions that is derived from a non-singular matrix $[c_{ik}]$ is linearly independent and a set of solutions that is derived from a singular matrix $[c_{ik}]$ is not linearly independent.

Accordingly, if a set of $n-r$ solutions is linearly independent, it must be derived from a non-singular matrix $[c_{ik}]$ and so, by § 6.41, such a set of solutions is a fundamental set.

Hence, *any linearly independent set of* $n-r$ *solutions is a fundamental set.*

Moreover, *any set of* $n-r$ *solutions that is not linearly independent is not a fundamental set*, for it must be derived from a singular matrix $[c_{ik}]$.

EXAMPLES IX

Devices to reduce the calculations when finding the rank of a matrix

1. Prove that the rank of a matrix is unaltered by any of the following changes in the elements of the matrix:

(i) the interchange of two rows (columns),
(ii) the multiplication of a row (column) by a non-zero constant,
(iii) the addition of any two rows.

HINT. Finding the rank depends upon showing which minors of the matrix are non-zero determinants. Or, more compactly, use Examples VII, 18–20, and Theorem 34 of Chapter IX.

2. When every minor of $[a_{ik}]$ of order $r+1$ that contains the first r rows and r columns is zero, prove that there are constants λ_{ik} such that the kth row of $[a_{ik}]$ is given by

$$\rho_k = \sum_{i=1}^{r} \lambda_{ik} \rho_i.$$

Prove (by the method of § 3.2, or otherwise) that every minor of $[a_{ik}]$ of order $r+1$ is then zero.

3. By using Example 2, prove that if a minor of order r is non-zero and if every minor obtained by bordering it with the elements from an arbitrary column and an arbitrary row are zero, then the rank of the matrix is r.

4. **Rule for symmetric square matrices.** If a principal minor of order r is non-zero and all principal minors of order $r+1$ are zero, then the rank of the matrix is r. (See p. 108.)

Prove the rule by establishing the following results for $[a_{ik}]$.

(i) If A_{ik} denotes the co-factor of a_{ik} in

$$C = \begin{vmatrix} a_{11} & \cdot & \cdot & a_{1r} & a_{1s} & a_{1t} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & \cdot & \cdot & a_{rr} & a_{rs} & a_{rt} \\ a_{s1} & \cdot & \cdot & a_{sr} & a_{ss} & a_{st} \\ a_{t1} & \cdot & \cdot & a_{tr} & a_{ts} & a_{tt} \end{vmatrix} \qquad (a_{ik} = a_{ki}),$$

then $A_{ss}A_{tt}-A_{ts}A_{st} = MC$, where M is the complementary minor of $a_{ss}a_{tt}-a_{ts}a_{st}$ in C (Theorem 18).

(ii) If every $C = 0$ and every corresponding $A_{ss} = 0$, then every $A_{st} = A_{ts} = 0$ and (Example 2) every minor of the complete matrix $[a_{ik}]$ of order $r+1$ is zero.

(iii) If all principal minors of orders $r+1$ and $r+2$ are zero, then the rank is r or less. (May be proved by induction.)

Numerical examples

5. Find the ranks of the matrices

$$\text{(i)} \begin{bmatrix} 1 & 3 & 4 & 7 \\ 2 & 4 & 5 & 8 \\ 3 & 1 & 2 & 3 \end{bmatrix}, \qquad \text{(ii)} \begin{bmatrix} 4 & 5 & 6 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{bmatrix},$$

$$\text{(iii)} \begin{bmatrix} 2 & 1 & 3 & 4 \\ 5 & 8 & 1 & 4 \\ 6 & 5 & 8 & 1 \\ 3 & 8 & 7 & 2 \end{bmatrix}, \qquad \text{(iv)} \begin{bmatrix} 6 & 1 & 3 & 8 \\ 4 & 2 & 6 & -1 \\ 10 & 3 & 9 & 7 \\ 16 & 4 & 12 & 15 \end{bmatrix}$$

Ans. (i) 3, (ii) 2, (iii) 4, (iv) 2.

Geometrical applications

6. The homogeneous coordinates of a point in a plane are (x, y, z). Prove that the three points (x_r, y_r, z_r), $r = 1, 2, 3$, are collinear or non-collinear according as the rank of the matrix

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}$$

is 2 or 3.

7. The coordinates of an arbitrary point can be written (with the notation of Example 6) in the form

$$x = \lambda x_1 + \mu x_2 + \nu x_3,$$

$$y = \lambda y_1 + \mu y_2 + \nu y_3,$$

$$z = \lambda z_1 + \mu z_2 + \nu z_3,$$

provided that the points 1, 2, 3 are not collinear.

HINT. Consider the matrix of 4 rows and 3 columns and use Theorem 31.

8. Give the analogues of Examples 6 and 7 for the homogeneous coordinates (x, y, z, w) of a point in three-dimensional space.

9. The configuration of three planes in space, whose cartesian equations are

$$a_{i1}x + a_{i2}y + a_{i3}z = b_i \quad (i = 1, 2, 3),$$

is determined by the ranks of the matrices $A \equiv [a_{ik}]$ and the augmented matrix B (§ 4.1). Check the following results.

(i) When $|a_{ik}| \neq 0$, the rank of A is 3 and the three planes meet in a finite point.

(ii) When the rank of A is 2 and the rank of B is 3, the planes have no finite point of intersection.

The planes are all parallel only if every A_{ik} is zero, when the rank of A is 1. When the rank of A is 2,

either (*a*) no two planes are parallel,

or (*b*) two are parallel and the third not.

Since $|a_{ik}| = 0$,

$$A_{11} : A_{12} : A_{13} = A_{21} : A_{22} : A_{23} = A_{31} : A_{32} : A_{33}$$

(Theorem 12), and if the planes 2 and 3 intersect, their line of intersection has direction cosines proportional to A_{11}, A_{12}, A_{13}. Hence in (*a*) the planes meet in three parallel lines and in (*b*) the third plane meets the two parallel planes in a pair of parallel lines.

The configuration is three planes forming a triangular prism; as a special case, two faces of the prism may be parallel.

(iii) When the rank of A is 2 and the rank of B is 2, one equation is a sum of multiples of the other two. The planes given by these two equations are not parallel (or the rank of A would be 1), and so the configuration is three planes having a common line of intersection.

(iv) When the rank of A is 1 and the rank of B is 2, the three planes are parallel.

(v) When the rank of A is 1 and the rank of B is 1, all three equations represent one and the same plane.

10. Prove that the rectangular cartesian coordinates (X_1, X_2, X_3) of the orthogonal projection of (x_1, x_2, x_3) on a line through the origin having direction cosines (l_1, l_2, l_3) are given by

$$X = \begin{bmatrix} l_1^2 & l_1 l_2 & l_1 l_3 \\ l_1 l_2 & l_2^2 & l_2 l_3 \\ l_1 l_3 & l_2 l_3 & l_3^2 \end{bmatrix} x,$$

where X, x are single column matrices.

11. If L denotes the matrix in Example 10, and M denotes a corresponding m matrix, show that

$$L^2 = L \quad \text{and} \quad M^2 = M$$

and that the point $(L+M)x$ is the projection of x on a line in the plane of the lines l and m if and only if l and m are perpendicular.

NOTE. A 'principal minor' is obtained by deleting corresponding rows and columns.

CHAPTER IX

DETERMINANTS ASSOCIATED WITH MATRICES

1. The rank of a product of matrices

1.1. Suppose that A is a matrix of n columns, and B a matrix of n rows, so that the product AB can be formed. When A has n_1 rows and B has n_2 columns, the product AB has n_1 rows and n_2 columns.

If a_{ik}, b_{ik} are typical elements of A, B,

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

is a typical element of AB.

Any t-rowed minor of AB may, with a suitable labelling of the elements, be denoted by

$$\Delta = \begin{vmatrix} c_{11} & c_{12} & . & . & c_{1t} \\ c_{21} & c_{22} & . & . & c_{2t} \\ . & . & . & . & . \\ c_{t1} & c_{t2} & . & . & c_{tt} \end{vmatrix}.$$

When $t = n$ (this assumes that $n_1 \geqslant n$, $n_2 \geqslant n$), Δ is the product by rows of the two determinants

$$\begin{vmatrix} a_{11} & . & . & a_{1n} \\ . & . & . & . \\ a_{n1} & . & . & a_{nn} \end{vmatrix}, \qquad \begin{vmatrix} b_{11} & . & . & b_{n1} \\ . & . & . & . \\ b_{1n} & . & . & b_{nn} \end{vmatrix};$$

that is, Δ is the product of a t-rowed minor of A by a t-rowed minor of B.

When $t \neq n$, Δ is the product by rows of the two arrays

$$\begin{matrix} a_{11} & . & . & a_{1n} \\ . & . & . & . \\ a_{t1} & . & . & a_{tn} \end{matrix} \qquad \begin{matrix} b_{11} & . & . & b_{n1} \\ . & . & . & . \\ b_{1t} & . & . & b_{nt} \end{matrix}$$

Hence (Theorem 15), when $t > n$, Δ is zero; and when $t < n$, Δ is the sum of all the products of corresponding determinants of order t that can be formed from the two arrays (Theorem 14).

Hence every minor of AB of order greater than n is zero, and every minor of order $t \leqslant n$ is the sum of products of a t-rowed minor of A by a t-rowed minor of B.

Accordingly,

THEOREM 33. *The rank of a product AB cannot exceed the rank of either factor.*

For, by what we have shown, (i) every minor of AB with more than n rows is zero, and the ranks of A and B cannot exceed n, (ii) every minor of AB of order $t \leqslant n$ is the product or the sum of products of minors of A and B of order t. If all the minors of A or if all the minors of B of order r are zero, then so are all minors of AB of order r.

1.2. There is one type of multiplication which gives a theorem of a more precise character.

THEOREM 34. *If A has m rows and n columns and B is a non-singular square matrix of order n, then A and AB have the same rank; and if C is a non-singular square matrix of order m, then A and CA have the same rank.*

If the rank of A is r and the rank of AB is ρ, then $\rho \leqslant r$. But $A = AB.B^{-1}$ and so r, the rank of A, cannot exceed ρ and n, which are the ranks of the factors AB and B^{-1}. Hence $\rho = r$. The proof for CA is similar.

2. The characteristic equation; latent roots

2.1. Associated with every square matrix A of order n is the matrix $A-\lambda I$, where I is the unit matrix of order n and λ is any number, real or complex. In full, this matrix is

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22}-\lambda & . & . & a_{2n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nn}-\lambda \end{bmatrix}$$

The determinant of this matrix is of the form

$$f(\lambda) \equiv (-1)^n(\lambda^n+p_1\lambda^{n-1}+...+p_n), \tag{1}$$

where the p_r are polynomials in the n^2 elements a_{ik}. The roots of the equation $f(\lambda) = 0$, that is, of

$$|A-\lambda I| = 0, \tag{2}$$

are called the LATENT ROOTS of the matrix A and the equation itself is called the CHARACTERISTIC EQUATION of A.

2.2. THEOREM 35. *Every square matrix satisfies its own characteristic equation; that is, if*

$$|A-\lambda I| = (-1)^n(\lambda^n+p_1\lambda^{n-1}+...+p_n),$$

then $$A^n+p_1A^{n-1}+...+p_{n-1}A+p_nI = 0.$$

The adjoint matrix of $A-\lambda I$, say B, is a matrix whose elements are polynomials in λ of degree $n-1$ or less, the coefficients of the various powers of λ being polynomials in the a_{ik}. Such a matrix can be written as

$$B_0+B_1\lambda+...+B_{n-2}\lambda^{n-2}+B_{n-1}\lambda^{n-1}, \tag{3}$$

where the B_r are matrices whose elements are polynomials in the a_{ik}.

Now, by Theorem 24, the product of a matrix by its adjoint = determinant of the matrix × unit matrix. Hence

$$\begin{aligned}(A-\lambda I)B &= |A-\lambda I|\times I\\ &= (-1)^n(\lambda^n+p_1\lambda^{n-1}+...+p_n)I,\end{aligned}$$

on using the notation of (1). Since B is given by (3), we have

$$\begin{aligned}(A-\lambda I)(B_0+B_1\lambda+...+B_{n-2}\lambda^{n-2}+B_{n-1}\lambda^{n-1})&\\ = (-1)^n(\lambda^n+p_1\lambda^{n-1}+...+p_n)I&.\end{aligned}$$

This equation is true for all values of λ and we may therefore equate coefficients[†] of λ; this gives us

$$\begin{aligned}-B_{n-1} &= (-1)^nI \quad (IB_{n-1} = B_{n-1}),\\ -B_{n-2}+AB_{n-1} &= (-1)^np_1I,\\ -B_{n-3}+AB_{n-2} &= (-1)^np_2I,\\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\\ -B_0+AB_1 &= (-1)^np_{n-1}I,\\ AB_0 &= (-1)^np_nI.\end{aligned}$$

[†] We may regard the equation as the conspectus of the n^2 equations

$$\begin{aligned}(a_{ik}-\lambda\delta_{ik})(b_{ik}^{(0)}+...+b_{ik}^{(n-1)}\lambda^{n-1})&\\ = (-1)^n(\lambda^n+p_1\lambda^{n-1}+...+p_n)\delta_{ik} \quad (i,\, k = 1,..., n)&,\end{aligned}$$

so that the equating of coefficients is really the ordinary algebraical procedure of equating coefficients, but doing it n^2 times for each power of λ.

Hence we have

$$\begin{aligned}&A^n+p_1A^{n-1}+\ldots+p_{n-1}A+p_nI\\ &\quad= A^nI+A^{n-1}p_1I+\ldots+A\,.\,p_{n-1}I+p_nI\\ &\quad= (-1)^n\{-A^nB_{n-1}+A^{n-1}(-B_{n-2}+AB_{n-1})+\ldots+\\ &\qquad\qquad +A(-B_0+AB_1)+AB_0\}\\ &\quad= 0.\end{aligned}$$

2.3. If $\lambda_1, \lambda_2,\ldots, \lambda_n$ are the latent roots of the matrix A, so that

$$\begin{aligned}|A-\lambda I| &= (-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n)\\ &= (-1)^n(\lambda^n+p_1\lambda^{n-1}+\ldots+p_n),\end{aligned}$$

it follows that (Example 13, p. 81),

$$(A-\lambda_1 I)(A-\lambda_2 I)\ldots(A-\lambda_n I) = A^n+p_1A^{n-1}+\ldots+p_nI = 0.$$

It does *not* follow that any one of the matrices $A-\lambda_r I$ is the zero matrix.

3. A theorem on determinants and latent roots

3.1. THEOREM 36. *If $g(t)$ is a polynomial in t, and A is a square matrix, then the determinant of the matrix $g(A)$ is equal to the product*

$$g(\lambda_1)g(\lambda_2)\ldots g(\lambda_n),$$

where $\lambda_1, \lambda_2,\ldots, \lambda_n$ are the latent roots of A.

Let $$g(t) = c(t-t_1)\ldots(t-t_m),$$
so that $$g(A) = c(A-t_1 I)\ldots(A-t_m I).$$
Then, by Chapter VI, § 10,

$$\begin{aligned}|g(A)| &= c^n|A-t_1 I|\times\ldots\times|A-t_m I|\\ &= (-1)^{mn}c^nf(t_1)\ldots f(t_m)\\ &= c^n\prod_{r=1}^{n}(\lambda_r-t_1)\ldots\prod_{r=1}^{n}(\lambda_r-t_m)\\ &= \prod_{r=1}^{n}g(\lambda_r).\end{aligned}$$

The theorem also holds when $g(A)$ is a rational function $g_1(A)/g_2(A)$ provided that $g_2(A)$ is non-singular.

3.2. Further, let $g_1(A)$ and $g_2(A)$ be two polynomials in the matrix A, $g_2(A)$ being non-singular. Then the matrix

$$\lambda I-\{g_1(A)/g_2(A)\},$$

that is $$\{\lambda g_2(A)-g_1(A)\}/g_2(A),$$

is a rational function of A with a non-singular denominator $g_2(A)$. Applying the theorem to this function we obtain, on writing $g_1(A)/g_2(A) = g(A)$,

$$|\lambda I-g(A)| = \prod_{r=1}^{n} \{\lambda-g(\lambda_r)\}.$$

From this follows an important result:

THEOREM 36. COROLLARY. *If $\lambda_1,..., \lambda_n$ are the latent roots of the matrix A and $g(A)$ is of the form $g_1(A)/g_2(A)$, where g_1, g_2 are polynomials and $g_2(A)$ is non-singular, the latent roots of the matrix $g(A)$ are given by $g(\lambda_r)$.*

4. Equivalent matrices

4.1. Elementary transformations of a matrix to standard form.

DEFINITION 18. *The* ELEMENTARY TRANSFORMATIONS *of a matrix are*

(i) *the interchange of two rows or columns,*

(ii) *the multiplication of each element of a row (or column) by a constant other than zero,*

(iii) *the addition to the elements of one row (or column) of a constant multiple of the elements of another row (or column).*

Let A be a matrix of rank r. Then, as we shall prove, it can be changed by elementary transformations into a matrix of the form

$$\left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right],$$

where I denotes the unit matrix of order r and the zeros denote null matrices, in general rectangular.

In the first place,† elementary transformations of type (i) replace A by a matrix such that the minor formed by the elements common to its first r rows and r columns is not equal to zero. Next, we can express any other row of A as a sum of multiples of these r rows; subtraction of these multiples of the r rows from the row in question will therefore give a row of zeros. These transformations, of type (iii), leave the rank of the

---

† The argument that follows is based on Chapter VIII, § 3.

matrix unaltered (cf. Examples IX, 1) and the matrix now before us is of the form

$$\left[\begin{array}{c|c} P & Q \\ \hline 0 & 0 \end{array}\right],$$

where P denotes a matrix, of r rows and columns, such that $|P| \neq 0$, and the zeros denote null matrices.

By working with columns where before we worked with rows, this can be transformed by elementary transformations, of type (iii), to

$$\left[\begin{array}{c|c} P & 0 \\ \hline 0 & 0 \end{array}\right].$$

Finally, suppose P is, in full,

$$\begin{bmatrix} a_1 & b_1 & c_1 & . & . & k_1 \\ a_2 & b_2 & c_2 & . & . & k_2 \\ . & . & . & . & . & . \\ a_r & b_r & c_r & . & . & k_r \end{bmatrix}.$$

Then, again by elementary transformations of types (i) and (iii), P can be changed successively to

$$\begin{bmatrix} a_1 & 0 & 0 & . & . & 0 \\ a_2 & \beta_2 & \gamma_2 & . & . & \kappa_2 \\ . & . & . & . & . & . \\ a_r & \beta_r & \gamma_r & . & . & \kappa_r \end{bmatrix},$$

$$\begin{bmatrix} a_1 & 0 & 0 & . & . & 0 \\ 0 & \beta_2 & 0 & . & . & 0 \\ . & . & . & . & . & . \\ a_r' & \beta_r & c_r' & . & . & k_r' \end{bmatrix},$$

and so, step by step, to a matrix having zeros in all places other than the principal diagonal and non-zeros† in that diagonal.

A final series of elementary transformations, of type (ii), presents us with I, the unit matrix of order r.

4.2. As Examples VII, 18–20, show, the transformations envisaged in § 4.1 can be performed by pre- and post-multiplication of A by non-singular matrices. Hence we have proved that

---

† If a diagonal element were zero the rank of P would be less than r: all these transformations leave the rank unaltered.

when A is a matrix of rank r, there are non-singular matrices B and C such that

$$BAC = I_r',$$

where I_r' denotes the unit matrix of order r bordered by null matrices.

4.3. Definition 19. *Two matrices are* EQUIVALENT *if it is possible to pass from one to the other by a chain of elementary transformations.*

If A_1 and A_2 are equivalent, they have the same rank and there are non-singular matrices B_1, B_2, C_1, C_2 such that

$$B_1 A_1 C_1 = I_r' = B_2 A_2 C_2.$$

From this it follows that

$$\begin{aligned} A_1 &= B_1^{-1} B_2 A_2 C_2 C_1^{-1} \\ &= L A_2 M, \end{aligned}$$

say, where $L = B_1^{-1} B_2$ and so is non-singular, and $M = C_2 C_1^{-1}$ and is also non-singular.

Accordingly, when two matrices are equivalent each can be obtained from the other through pre- and post-multiplication by non-singular matrices.

Conversely, if A_2 is of rank r, L and M are non-singular matrices, and

$$A_1 = L A_2 M;$$

then, as we shall prove, we can pass from A_2 to A_1, or from A_1 to A_2, by elementary transformations. Both A_2 and A_1 are of rank r (Theorem 34). We can, as we saw in § 4.1, pass from A_2 to I_r' by elementary transformations; we can pass from A_1 to I_r' by elementary transformations and so, by using the inverse operations, we can pass from I_r' to A_1 by elementary transformations. Hence we can pass from A_2 to A_1, or from A_1 to A_2, by elementary transformations.

4.4. The detailed study of equivalent matrices is of fundamental importance in the more advanced theory. Here we have done no more than outline some of the immediate consequences of the definition.†

---

† The reader who intends to pursue the subject seriously should consult H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices* (London and Glasgow, 1932).

PART III

LINEAR AND QUADRATIC FORMS; INVARIANTS AND COVARIANTS

CHAPTER X

ALGEBRAIC FORMS: LINEAR TRANSFORMATIONS

1. Number fields

We recall the definition of a number field given in the preliminary note.

DEFINITION 1. *A set of numbers, real or complex, is called a* FIELD *of numbers, or a number field, when, if r and s belong to the set and s is not zero,*

$$r+s,\quad r-s,\quad r\times s,\quad r\div s$$

also belong to the set.

Typical examples of number fields are

(i) the field of all rational real numbers ($\mathfrak{F}_r$ say);
(ii) the field of all real numbers;
(iii) the field of all numbers of the form $a+b\sqrt{5}$, where a and b are rational real numbers;
(iv) the field of all complex numbers.

Every number field must contain each and every number that is contained in $\mathfrak{F}_r$ (example (i) above); it must contain 1, since it contains the quotient α/α, where α is any number of the set; it must contain 0, 2, and every integer, since it contains the difference $1-1$, the sum $1+1$, and so on; it must contain every fraction, since it contains the quotient of any one integer by another.

2. Linear and quadratic forms

2.1. Let $\mathfrak{F}$ be any field of numbers and let a_{ij}, b_{ij},... be determinate numbers of the field; that is, we suppose their values to be fixed. Let x_r, y_r,... denote numbers that are not to be thought of as fixed, but as free to be any, arbitrary, numbers from a field $\mathfrak{F}_1$, not necessarily the same as $\mathfrak{F}$. The numbers a_{ij}, b_{ij},... we call CONSTANTS in $\mathfrak{F}$; the symbols x_r, y_r,... we call VARIABLES in $\mathfrak{F}_1$.

An expression such as

$$\sum_{j=1}^{n} a_{ij} x_j \tag{1}$$

is said to be a LINEAR FORM in the variables x_j; an expression such as

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_i y_j \tag{2}$$

is said to be a BILINEAR FORM in the variables x_i and y_j; an expression such as

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \quad (a_{ij} = a_{ji}) \tag{3}$$

is said to be a QUADRATIC FORM in the variables x_i. In each case, when we wish to stress the fact that the constants a_{ij} belong to a certain field, $\mathfrak{F}$ say, we refer to the form as one with coefficients in $\mathfrak{F}$.

A form in which the variables are necessarily real numbers is said to be a 'FORM IN REAL VARIABLES'; one in which both coefficients and variables are necessarily real numbers is said to be a 'REAL FORM'.

2.2. It should be noticed that the term 'quadratic form' is used of (3) only when $a_{ij} = a_{ji}$. This restriction of usage is dictated by experience, which shows that the consequent theory is more compact when such a restriction is imposed.

The restriction is, however, more apparent than real: for an expression such as

$$\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_i x_j,$$

wherein b_{ij} is not always the same as b_{ji}, is identical with the quadratic form

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \quad (a_{ij} = a_{ji})$$

when we define the a's by means of the equations

$$a_{ii} = b_{ii}, \qquad a_{ij} = a_{ji} = \tfrac{1}{2}(b_{ij}+b_{ji}).$$

For example, $x^2+3xy+2y^2$ is a quadratic form in the two variables x and y, having coefficients

$$a_{11} = 1, \qquad a_{22} = 2, \qquad a_{12} = a_{21} = \tfrac{3}{2}.$$

2.3. Matrices associated with linear and quadratic forms. The symmetrical square matrix $A \equiv [a_{ij}]$, having a_{ij}

in its ith row and jth column, is associated with the quadratic form (3): it is symmetrical because $a_{ij} = a_{ji}$, and so $A = A'$, the transpose of A.

In the sequel it will frequently be necessary to bear in mind that the matrix associated with any quadratic form is a symmetrical matrix.

We also associate with the bilinear form (2) the matrix $[a_{ij}]$ of m rows and n columns; and with the linear form (1) we associate the single-row matrix $[a_{i1},..., a_{in}]$. More generally, we associate with the m linear forms

$$\sum_{j=1}^{n} a_{ij} x_j \quad (i = 1,..., m) \tag{4}$$

the matrix $[a_{ij}]$ of m rows and n columns.

2.4. Notation. We denote the associated matrix $[a_{ij}]$ of any one of (2), (3), or (4) by the single letter A. We may then conveniently abbreviate

the bilinear form (2) to $A(x,y)$,
the quadratic form (3) to $A(x,x)$,
and the m linear forms (4) to Ax.

The first two of these are merely shorthand notations; the third, though it also can be so regarded, is better envisaged as the product of the matrix A by the matrix x, a single-*column* matrix having $x_1,..., x_n$ as elements: the matrix product Ax, which has as many rows as A and as many columns as x, is then a single-column matrix of m rows having the m linear forms as its elements.

2.5. Matrix expressions for quadratic and bilinear forms. As in § 2.4, let x denote a single-column matrix with elements $x_1,..., x_n$: then x', the transpose of x, is a single-row matrix with elements $x_1,..., x_n$.

Let A denote the matrix of the quadratic form $\sum\sum a_{rs} x_r x_s$. Then Ax is the single-column matrix having

$$a_{r1} x_1 + ... + a_{rn} x_n$$

as the element in its rth row, and $x'Ax$ is a matrix of one row (the number of rows of x') and one column (the number of columns of Ax), the single element being

$$\sum_{r=1}^{n} x_r(a_{r1}x_1+...+a_{rn}x_n) = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}x_r x_s.$$

Thus the quadratic form is represented by the single-element matrix $x'Ax$.

Similarly, when x and y are single-column matrices having $x_1,..., x_n$ and $y_1,..., y_n$ as row elements, the bilinear form (2) is represented by the single-element matrix $x'Ay$.

2.6. DEFINITION 2. *The* DISCRIMINANT *of the quadratic form*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \quad (a_{ij} = a_{ji})$$

is the determinant of its coefficients, namely $|a_{ij}|$.

3. Linear transformations

3.1. The set of equations

$$X_i = \sum_{j=1}^{n} a_{ij}x_j \quad (i = 1,..., n), \tag{1}$$

wherein the a_{ij} are given constants and the x_j are variables, is said to be a linear transformation connecting the variables x_j and the variables X_i. When the a_{ij} are constants in a given field $\mathfrak{F}$ we say that the transformation has coefficients in $\mathfrak{F}$; when the a_{ij} are real numbers we say that the transformation is real.

DEFINITION 3. *The determinant* $|a_{ij}|$, *whose elements are the coefficients* a_{ij} *of the transformation* (1), *is called the* MODULUS OF THE TRANSFORMATION.

DEFINITION 4. *A transformation is said to be* NON-SINGULAR *when its modulus is not zero, and is said to be* SINGULAR *when its modulus is zero.*

We sometimes speak of (1) as a transformation from the x_i to the X_i; or, briefly, $x \rightarrow X$.

3.2. The transformation (1) is most conveniently written as

$$X = Ax, \tag{2}$$

a matrix equation in which X and x denote single-column matrices with elements $X_1,..., X_n$ and $x_1,..., x_n$ respectively,

A denotes the matrix (a_{ij}), and Ax denotes the product of the two matrices A and x.

When A is a non-singular matrix, it has a reciprocal A^{-1} (Chap. VII, § 3) and

$$A^{-1}X = A^{-1}Ax = x, \tag{3}$$

which expresses x directly in terms of X. Also $(A^{-1})^{-1} = A$, and so, with a non-singular transformation, it is immaterial whether it be given in the form $x \to X$ or in the form $X \to x$. Moreover, when $X = Ax$, given any X whatsoever, there is one and only one corresponding x, and it is given by $x = A^{-1}X$.

When A is a singular matrix, there are X for which no corresponding x can be defined. For in such a case, r, the rank of the matrix A, is less than n, and we can select r rows of A and express every row as a sum of multiples of these rows (Theorem 31). Thus, the rows being suitably numbered, (1) gives relations

$$X_k = \sum_{i=1}^{r} l_{ki} X_i \quad (k = r+1,..., n), \tag{4}$$

wherein the l_{ki} are constants, and so the set $X_1,..., X_n$ is limited to such sets as will satisfy (4): a set of X that does not satisfy (4) will give no corresponding set of x. For example, in the linear transformation

$$X = 2x+3y, \qquad Y = 4x+6y,$$

which has the singular matrix $\begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}$, the pair X, Y must satisfy the relation $2X = Y$; for any pair X, Y that does not satisfy this relation, there is no corresponding pair x, y.

3.3. The product of two transformations. Let x, y, z be '*numbers*' (Chap. VI, § 1), each with n components, and let all suffixes be understood to run from 1 to n. Use the summation convention. Then the transformation

$$x_i = a_{ij} y_j \tag{1}$$

may be written as a matrix equation $x = Ay$ (§ 3.2), and the transformation

$$y_j = b_{jk} z_k \tag{2}$$

may be written as a matrix equation $y = Bz$.

If in (1) we substitute for y in terms of z, we obtain

$$x_i = a_{ij} b_{jk} z_k, \tag{3}$$

which is a transformation whose matrix is the product AB (compare Chap. VI, § 4). Thus the result of two successive transformations $x = Ay$ and $y = Bz$ may be written as

$$x = ABz.$$

Since AB is the product of the two matrices, we adopt a similar nomenclature for the transformations themselves.

DEFINITION 5. *The transformation $x = ABz$ is called the* PRODUCT OF THE TWO TRANSFORMATIONS $x = Ay$, $y = Bz$. [*It is the result of the successive transformations $x = Ay$, $y = Bz$.*]

The modulus of the transformation (3) is the determinant $|a_{ij} b_{jk}|$, that is, the product of the two determinants $|A|$ and $|B|$. Similarly, if we have three transformations in succession,

$$x = Ay, \qquad y = Bz, \qquad z = Cu,$$

the resulting transformation is $x = ABCu$ and the modulus of this transformation is the product of the three moduli $|A|$, $|B|$, and $|C|$; and so for any finite number of transformations.

4. Transformations of quadratic and bilinear forms

4.1. Consider a given transformation

$$x_i = \sum_{k=1}^{n} b_{ik} X_k \quad (i = 1,\ldots, n) \tag{1}$$

and a given quadratic form

$$A(x,x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} = a_{ji}). \tag{2}$$

When we substitute in (2) the values of $x_1,\ldots, x_n$ given by (1), we obtain a quadratic expression in $X_1,\ldots, X_n$. This quadratic expression is said to be a transformation of (2) by (1), or the result of applying (1) to the form (2).

THEOREM 37. *If a transformation $x = BX$ is applied to a quadratic form $A(x,x)$, the result is a quadratic form $C(X,X)$ whose matrix C is given by*

$$C = B'AB,$$

where B' is the transpose of B.

First proof—a proof that depends entirely on matrix theory.

The form $A(x,x)$ is given by the matrix product $x'Ax$ (compare § 2.5). By hypothesis, $x = BX$, and this implies $x' = X'B'$ (Theorem 23). Hence, by the associative law of multiplication,

$$\begin{aligned} x'Ax &= X'B'ABX \\ &= X'CX, \end{aligned} \tag{3}$$

where C denotes the matrix $B'AB$.

Moreover, (3) is not merely a quadratic expression in $X_1,..., X_n$, but is, in fact, a quadratic *form* having $c_{ij} = c_{ji}$. To prove this we observe that, when $C = B'AB$,

$$\begin{aligned} C' &= B'A'B \quad \text{(Theorem 23, Corollary)} \\ &= B'AB \quad (\text{since } A = A') \\ &= C; \end{aligned}$$

that is to say, $c_{ij} = c_{ji}$.

Second proof—a proof that depends mainly on the use of the summation convention.

Let all suffixes run from 1 to n and use the summation convention throughout. Then the quadratic form is

$$a_{ij}x_i x_j \quad (a_{ij} = a_{ji}). \tag{4}$$

The transformation $$x_i = b_{ik}X_k \tag{5}$$

expresses (4) in the form

$$c_{kl}X_k X_l. \tag{6}$$

We may calculate the values of the c's by carrying out the actual substitutions for the x in terms of the X. We have

$$\begin{aligned} a_{ij}x_i x_j &= a_{ij}b_{ik}X_k b_{jl}X_l \\ &= b_{ik}a_{ij}b_{jl}X_k X_l, \end{aligned}$$

so that $$c_{kl} = b_{ik}a_{ij}b_{jl} = b'_{ki}a_{ij}b_{jl},$$

where $b'_{ki} = b_{ik}$ is the element in the kth row and ith column of B'. Hence the matrix C of (6) is equal to the matrix product $B'AB$.

Moreover, $c_{kl} = c_{lk}$; for, since $a_{ij} = a_{ji}$, we have

$$c_{kl} = b_{ik}a_{ij}b_{jl} = b_{ik}a_{ji}b_{jl},$$

and, since both i and j are dummies,

$$b_{ik}a_{ji}b_{jl} = b_{jk}a_{ij}b_{il},$$

which is the same sum as $b_{il}a_{ij}b_{jk}$ or c_{lk}. Thus (6) is a quadratic form, having $c_{kl} = c_{lk}$.

4.2. THEOREM 38. *If transformations $x = CX$, $y = BY$ are applied to a bilinear form $A(x,y)$, the result is a bilinear form $D(X,Y)$ whose matrix D is given by*

$$D = C'AB,$$

where C' is the transpose of C.

First proof. As in the first proof of Theorem 37, we write the bilinear form as a matrix product, namely $x'Ay$. It follows at once, since $x' = X'C'$, that

$$x'Ay = X'C'ABY = X'DY,$$

where $D = C'AB$.

4.3. *Second proof.* This proof is given chiefly to show how much detail is avoided by the use of matrices: it will also serve as a proof in the most elementary form obtainable.

The bilinear form

$$A(x,y) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_i y_j \tag{1}$$

may be written as

$$A(x,y) = \sum_{i=1}^{m} l_i x_i, \qquad l_i = \sum_{j=1}^{n} a_{ij}y_j.$$

Transform the y's by putting

$$y_j = \sum_{k=1}^{n} b_{jk}Y_k \quad (j = 1,..., n),$$

so that

$$l_i = \sum_{j=1}^{n} a_{ij}\sum_{k=1}^{n} b_{jk}Y_k,$$

and

$$A(x,y) = \sum_{i=1}^{m}\sum_{k=1}^{n}\Big(\sum_{j=1}^{n} a_{ij}b_{jk}\Big)x_i Y_k. \tag{2}$$

The matrix of this bilinear form in x and Y has for its element in the ith row and kth column

$$\sum_{j=1}^{n} a_{ij}b_{jk}.$$

This matrix is therefore AB, where A is $[a_{ik}]$ and B is $[b_{ik}]$.

Again, transform the x's in (1) by putting

$$x_i = \sum_{k=1}^{m} c_{ik} X_k \quad (i = 1,..., m), \tag{3}$$

and leave the y's unchanged. Then

$$\begin{aligned} A(x, y) &= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j \\ &= \sum_{j=1}^{n} y_j \sum_{i=1}^{m} a_{ij} \sum_{k=1}^{m} c_{ik} X_k \\ &= \sum_{k=1}^{m} \sum_{j=1}^{n} \Big(\sum_{i=1}^{m} c_{ik} a_{ij}\Big) X_k y_j. \end{aligned} \tag{4}$$

The sum $\sum_{i=1}^{m} c_{ik} a_{ij}$ is the inner product of the ith row of

$$\begin{bmatrix} c_{11} & . & . & . & c_{m1} \\ . & . & . & . & . \\ c_{1m} & . & . & . & c_{mm} \end{bmatrix} \tag{5}$$

by the jth column of

$$\begin{bmatrix} a_{11} & . & . & . & a_{1n} \\ . & . & . & . & . \\ a_{m1} & . & . & . & a_{mn} \end{bmatrix}. \tag{6}$$

Now (5) is the transpose of

$$\begin{bmatrix} c_{11} & . & . & . & c_{1m} \\ . & . & . & . & . \\ c_{m1} & . & . & . & c_{mm} \end{bmatrix},$$

that is, of C, the matrix of the transformation (3), while (6) is the matrix A of the bilinear form (1). Hence the matrix of the bilinear form (4) is $C'A$.

The theorem follows on applying the two transformations in succession.

4.4. The discriminant of a quadratic form.

THEOREM 39. *Let a quadratic form*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} = a_{ji}) \tag{1}$$

be transformed by a linear transformation

$$x_i = \sum_{k=1}^{n} l_{ik} X_k \quad (i = 1,..., n)$$

whose modulus, that is, the determinant $|l_{ik}|$, *is equal to* M; *let the resulting quadratic form be*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}X_iX_j \quad (c_{ij}=c_{ji}). \tag{2}$$

Then the discriminant of (2) *is* M^2 *times the discriminant of* (1); *in symbols*

$$|c_{ij}| = M^2|a_{ij}|. \tag{3}$$

The content of this theorem is usefully abbreviated to the statement 'WHEN A QUADRATIC FORM IS CHANGED TO NEW VARIABLES, THE DISCRIMINANT IS MULTIPLIED BY THE SQUARE OF THE MODULUS OF THE TRANSFORMATION'.

The result is an immediate corollary of Theorem 37. By that theorem, the matrix of the form (2) is $C = L'AL$, where L is the matrix of the transformation. The discriminant of (2), that is, the determinant $|C|$, is given (Chap. VI, § 10) by the product of the three determinants $|L'|$, $|A|$, and $|L|$.

But $|L| = M$, and since the value of a determinant is unaltered when rows and columns are interchanged, $|L'|$ is also equal to M. Hence

$$|C| = M^2|A|.$$

This theorem is of fundamental importance; it will be used many times in the chapters that follow.

5. Hermitian forms

5.1. In its most common interpretation a HERMITIAN BILINEAR FORM is given by

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j, \tag{1}$$

wherein the coefficients a_{ij} belong to the field of complex numbers and are such that

$$a_{ji} = \bar{a}_{ij}.$$

The bar denotes, as in the theory of the complex variable, the conjugate complex; that is, if $z = \alpha+i\beta$, where α and β are real, then $\bar{z} = \alpha-i\beta$.

The matrix A of the coefficients of the form (1) satisfies the condition

$$A' = \bar{A}. \tag{2}$$

Any matrix A that satisfies (2) is said to be a HERMITIAN MATRIX.

5.2. A form such as

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i \bar{x}_j \quad (a_{ji} = \bar{a}_{ij})$$

is said to be a HERMITIAN FORM.

The theory of these forms is very similar to that of ordinary bilinear and quadratic forms. Theorems concerning Hermitian forms appear as examples at the end of this and later chapters.

6. Cogredient and contragredient sets

6.1. When two sets of variables $x_1,..., x_n$ and $y_1,..., y_n$ are related to two other sets $X_1,..., X_n$ and $Y_1,..., Y_n$ by the same transformation, say

$$x = AX, \qquad y = AY,$$

then the two sets x and y (equally, X and Y) are said to be cogredient sets of variables.

If a set $z_1,..., z_n$ is related to a set $Z_1,..., Z_n$ by a transformation whose matrix is the reciprocal of the transpose of A, that is,

$$z = (A')^{-1}Z, \quad \text{or} \quad Z = A'z,$$

then the sets x and z (equally, X and Z) are said to be contragredient sets of variables.

Examples of cogredient sets readily occur. A transformation in plane analytical geometry from one triangle of reference to another is of the type

$$\left.\begin{aligned} x &= l_1 X + m_1 Y + n_1 Z, \\ y &= l_2 X + m_2 Y + n_2 Z, \\ z &= l_3 X + m_3 Y + n_3 Z. \end{aligned}\right\} \tag{1}$$

The sets of variables (x_1, y_1, z_1) and (x_2, y_2, z_2), regarded as the coordinates of two distinct points, are cogredient sets: the coordinates of the two points referred to the new triangle of reference are (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2), and each is obtained by putting in the appropriate suffix in the equations (1).

Analytical geometry also furnishes an important example of contragredient sets. Let

$$lx+my+nz = 0 \tag{2}$$

be regarded as the equation of a given line α in a system of homogeneous point-coordinates (x, y, z) with respect to a given triangle of reference. Then the line (tangential) coordinates of α are (l, m, n). Take a new triangle of reference and suppose that (1) is the transformation for the point-coordinates. Then (2) becomes

$$LX+MY+NZ = 0,$$

where

$$\left.\begin{aligned} L &= l_1 l+l_2 m+l_3 n, \\ M &= m_1 l+m_2 m+m_3 n, \\ N &= n_1 l+n_2 m+n_3 n. \end{aligned}\right\} \tag{3}$$

The matrix of the coefficients of l, m, n in (3) is the transpose of the matrix of the coefficients of X, Y, Z in (1). Hence point-coordinates (x, y, z) and line-coordinates (l, m, n) are contragredient variables when the triangle of reference is changed. [Notice that (1) is $X, Y, Z \to x, y, z$ whilst (3) is $l, m, n \to L, M, N$.]

The notation of the foregoing becomes more compact when we consider n dimensions, a point x with coordinates $(x_1,..., x_n)$, and a 'flat' l with coordinates $(l_1,..., l_n)$. The transformation $x = AX$ of the point-coordinates entails the transformation $L = A'l$ of the tangential coordinates.

6.2. Another example of contragredient sets, one that contains the previous example as a particular case, is provided by differential operators. Let $F(x_1,...,x_n)$ be expressed as a function of $X_1,..., X_n$ by means of the transformation $x = AX$, where A denotes the matrix $[a_{ij}]$. Then

$$\frac{\partial F}{\partial X_i} = \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\frac{\partial x_j}{\partial X_i} = \sum_{j=1}^{n} a_{ji}\frac{\partial F}{\partial x_j}.$$

That is to say, in matrix form, if $x = AX$,

$$\frac{\partial}{\partial X} = A'\frac{\partial}{\partial x}.$$

Accordingly, the x_i and the $\partial/\partial x_i$ form contragredient sets.

7. The characteristic equation of a transformation

7.1. Consider the transformation $X = Ax$ or, in full,

$$X_i = \sum_{j=1}^{n} a_{ij} x_j \quad (i = 1,..., n). \tag{1}$$

Is it possible to assign values to the variables x so that $X_i = \lambda x_i$ $(i = 1,..., n)$, where λ is independent of i? If such a result is to hold, we must have

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j \quad (i = 1,..., n), \tag{2}$$

and this demands (Theorem 11), when one of $x_1,..., x_n$ differs from zero, that λ be a root of the equation

$$A(\lambda) \equiv \begin{vmatrix} a_{11}-\lambda & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22}-\lambda & . & . & a_{2n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nn}-\lambda \end{vmatrix} = 0. \tag{3}$$

This equation is called the CHARACTERISTIC EQUATION of the transformation (1); any root of the equation is called a CHARACTERISTIC NUMBER or a LATENT ROOT of the transformation.

If λ is a characteristic number, there is a set of numbers $x_1,..., x_n$, not all zero (Theorem 11), that satisfy equations (2). Let $\lambda = \lambda_1$, a characteristic number. If the determinant $A(\lambda_1)$ is of rank $(n-1)$, there is a unique corresponding set of ratios, say

$$x_1^{(1)} : x_2^{(1)} : ... : x_n^{(1)},$$

that satisfies (2). If the rank of the determinant $A(\lambda_1)$ is $n-2$ or less, the set of ratios is not unique (Chap. VIII, § 5).

A set of ratios that satisfies (2) when λ is equal to a characteristic number λ_r, is called a POLE CORRESPONDING TO λ_r.

7.2. If (x_1, x_2, x_3) and (X_1, X_2, X_3) are homogeneous coordinates referred to a given triangle of reference in a plane, then (1), with $|A| \neq 0$, may be regarded as a method of generating a one-to-one correspondence between the variable point (x_1, x_2, x_3) and the variable point (X_1, X_2, X_3). A pole of the transformation is then a point which corresponds to itself.

We shall not elaborate the geometrical implications† of such transformations; a few examples are given on the next page.

EXAMPLES X

1. Prove that all numbers of the form $a+b\surd 3$, where a and b are integers (or zero), constitute a ring; and that all numbers of the form $a+b\surd 3$, where a and b are the ratios of integers (or zero) constitute a field.

2. Express $ax^2+2hxy+by^2$ as a quadratic form (in accordance with the definition of § 2.1) and show that its discriminant is $ab-h^2$ (§ 2.6).

3. Write down the transformation which is the product of the two transformations

$$\left.\begin{aligned} x &= l_1\xi+m_1\eta+n_1\zeta \\ y &= l_2\xi+m_2\eta+n_2\zeta \\ z &= l_3\xi+m_3\eta+n_3\zeta \end{aligned}\right\}, \qquad \left.\begin{aligned} \xi &= \lambda_1 X+\mu_1 Y+\nu_1 Z \\ \eta &= \lambda_2 X+\mu_2 Y+\nu_2 Z \\ \zeta &= \lambda_3 X+\mu_3 Y+\nu_3 Z \end{aligned}\right\}.$$

4. Prove that, in solid geometry, if $\mathbf{i}_1, \mathbf{j}_1, \mathbf{k}_1$ and $\mathbf{i}_2, \mathbf{j}_2, \mathbf{k}_2$ are two unit frames of reference for cartesian coordinates and if, for $r = 1, 2$,

$$\mathbf{i}_r \wedge \mathbf{j}_r = \mathbf{k}_r, \qquad \mathbf{j}_r \wedge \mathbf{k}_r = \mathbf{i}_r, \qquad \mathbf{k}_r \wedge \mathbf{i}_r = \mathbf{j}_r,$$

then the transformation of coordinates has unit modulus.

[Omit if the vector notation is not known.]

5. Verify Theorem 37, when the original quadratic form is

$$ax^2+2hxy+by^2$$

and the transformation is $x = l_1X+m_1Y$, $y = l_2X+m_2Y$, by actual substitution on the one hand and the evaluation of the matrix product ($B'AB$ of the theorem) on the other.

6. Verify Theorem 38, by the method of Example 5, when the bilinear form is

$$a_{11}x_1y_1+a_{12}x_1y_2+a_{21}x_2y_1+a_{22}x_2y_2.$$

7. The homogeneous coordinates of the points D, E, F referred to ABC as triangle of reference are (x_1, y_1, z_1), (x_2, y_2, z_2), and (x_3, y_3, z_3). Prove that, if (l, m, n) are line-coordinates referred to ABC and (L, M, N) are line-coordinates referred to DEF, then

$$\Delta l = X_1L+X_2M+X_3N, \quad \text{etc.},$$

where Δ is the determinant $|x_1\,y_2\,z_3|$ and $X_r,\ldots$ are the co-factors of $x_r,\ldots$ in Δ.

HINT. First obtain the transformation $x, y, z \rightarrow X, Y, Z$ and then use § 6.1.

---

† Cf. R. M. Winger, *An Introduction to Projective Geometry* (New York, 1922), or, for homogeneous coordinates in three dimensions, G. Darboux, *Principes de géométrie analytique* (Gauthier-Villars, Paris, 1917).

8. In the transformation $X = Ax$, new variables Y and y, defined by

$$y = Bx, \qquad Y = BX \qquad (|B| \neq 0),$$

are introduced. Prove that the transformation $Y \rightarrow y$ is given by $Y = Cy$, where $C = BAB^{-1}$.

9. The transformations $x = BX$, $y = \bar{B}Y$ change the Hermitian† $a_{ij}x_i y_j$ (with $A' = \bar{A}$) into $c_{ij}X_i Y_j$. Prove that

$$C = B'A\bar{B}, \qquad C' = \bar{B}'A'B = \bar{C}.$$

10. Prove that the transformation $x = BX$, together with its conjugate complex $\bar{x} = \bar{B}\bar{X}$, changes $a_{ij}x_i\bar{x}_j$ $(A' = \bar{A})$ into $c_{ij}X_i\bar{X}_j$ $(C' = \bar{C})$, where $C = B'A\bar{B}$.

11. The elements of the matrices A and B belong to a given field $\mathfrak{F}$. Prove that, if the quadratic form $A(x,x)$ is transformed by the non-singular transformation $x = BX$ (or by $X = Bx$) into $C(X,X)$, then every coefficient of C belongs to $\mathfrak{F}$.

12. Prove that, if each x_r denotes a complex variable and a_{rs} are complex numbers such that $a_{rs} = \bar{a}_{sr}$ $(r = 1,..., n;\ s = 1,..., n)$, the Hermitian form $a_{rs}x_r\bar{x}_s$ is a real number.

[*Examples* 13–16 *are geometrical in character.*]

13. Prove that, if the points in a plane are transformed by the scheme

$$x' = a_1x+b_1y+c_1z, \qquad y' = a_2x+b_2y+c_2z,$$
$$z' = a_3x+b_3y+c_3z,$$

then every straight line transforms into a straight line. Prove also that there are in general three points that are transformed into themselves and three lines that are transformed into themselves.

Find the conditions to be satisfied by the coefficients in order that every point on a given line may be transformed into itself.

14. Show that the transformation of Example 13 can, in general, by a change of triangle of reference be reduced to the form

$$X' = \alpha X, \qquad Y' = \beta Y, \qquad Z' = \gamma Z.$$

Hence, or otherwise, show that the transformation is a homology (plane perspective, collineation) if a value of λ can be found which will make

$$\begin{vmatrix} a_1-\lambda & b_1 & c_1 \\ a_2 & b_2-\lambda & c_2 \\ a_3 & b_3 & c_3-\lambda \end{vmatrix}$$

of rank one.

HINT. When the determinant is of rank one there is a line l such that every point of it corresponds to itself. In such a case any two corresponding lines must intersect on l.

---

† The summation convention is employed in Examples 9, 10, and 12.

15. Obtain the equations giving, in two dimensions, the homology (collineation, plane perspective) in which the point (x_1', y_1', z_1') is the centre and the line (l_1', l_2', l_3') the axis of the homology.

16. Points in a plane are transformed by the scheme

$$\begin{aligned} x' &= \rho x+\alpha(\lambda x+\mu y+\nu z), \\ y' &= \rho y+\beta(\lambda x+\mu y+\nu z), \\ z' &= \rho z+\gamma(\lambda x+\mu y+\nu z). \end{aligned}$$

Find the points and lines that are transformed into themselves.

Show also that the transformation is involutory (i.e. two applications of it restore a figure to its original position) if and only if

$$2\rho+\alpha\lambda+\beta\mu+\gamma\nu = 0.$$

[Remember that (kx', ky', kz') is the same point as (x', y', z').]

CHAPTER XI

THE POSITIVE-DEFINITE FORM

1. Definite real forms

1.1. DEFINITION 6. *The real quadratic form*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \quad (a_{ij} = a_{ji})$$

is said to be POSITIVE-DEFINITE *if it is positive for every set of real values of* $x_1,..., x_n$ *other than the set* $x_1 = x_2 = ... = x_n = 0$. *It is said to be* NEGATIVE-DEFINITE *if it is negative for every set of real values of* $x_1,..., x_n$ *other than* $x_1 = x_2 = ... = x_n = 0$.

For example, $3x_1^2+2x_2^2$ is positive-definite, while $3x_1^2-2x_2^2$ is not positive-definite; the first is positive for every pair of real values of x_1 and x_2 except the pair $x_1 = 0$, $x_2 = 0$; the second is positive when $x_1 = x_2 = 1$, but it is negative when $x_1 = 1$ and $x_2 = 2$, and is zero when $x_1 = \sqrt{2}$ and $x_2 = \sqrt{3}$.

An example of a real form that can never be negative but is not positive-definite, according to the definition, is given by

$$(3x_1-2x_2)^2+4x_3^2.$$

This quadratic form is zero when $x_1 = 2$, $x_2 = 3$, and $x_3 = 0$, and so it does not come within the scope of Definition 6. The point of excluding such a form from the definition is that, whereas it appears as a function of three variables, x_1, x_2, x_3, it is essentially a function of two variables, namely,

$$X_1 = 3x_1-2x_2, \qquad X_2 = x_3.$$

A positive-definite form in one set of variables is still a positive-definite form when expressed in a new set of variables, provided only that the two sets are connected by a real non-singular transformation. This we now prove.

THEOREM 40. *A real positive-definite form in the n variables* $x_1,..., x_n$ *is a positive-definite form in the n variables* $X_1,..., X_n$ *provided that the two sets of variables are connected by a real, non-singular transformation.*

Let the positive-definite form be $B(x,x)$, and let the real non-singular transformation be $X = Ax$. Then $x = A^{-1}X$. Let $B(x,x)$ become $C(X,X)$ when expressed in terms of X.

Since $B(x,x)$ is positive-definite, $C(X,X)$ is positive for every X save that which corresponds to $x = 0$. But $X = Ax$, where A is a non-singular matrix, and so $X = 0$ if and only if $x = 0$. Hence $C(X,X)$ is positive for every X other than $X = 0$.

NOTE. The equation $x = 0$ is a matrix equation, x being a single-column matrix whose elements are $x_1,..., x_n$.

1.2. The most obvious type of positive-definite form in n variables is

$$a_{11}x_1^2+...+a_{nn}x_n^2 \qquad (a_{rr} > 0).$$

We now show that every positive-definite form in n variables is a transformation of this obvious type.

THEOREM 41. *Every real positive-definite form, $B(x, x)$, can be transformed by a real transformation of unit modulus into a form*

$$c_{11}X_1^2+...+c_{nn}X_n^2$$

wherein each c_{rr} is positive.

The manipulations that follow are typical of others that occur in later work. Here we give them in detail; later we shall refer back to this section and omit as many of the details as clarity permits.

Let the given positive-definite form be

$$B(x,x) \equiv \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_i x_j \quad (b_{ij} = b_{ji})$$

$$\equiv \sum_{r=1}^{n} b_{rr}x_r^2+2\sum_{r<s}^{n} b_{rs}x_r x_s. \tag{1}$$

The terms of (1) that involve x_1 are

$$b_{11}x_1^2+2b_{12}x_1x_2+...+2b_{1n}x_1x_n. \tag{2}$$

Moreover, since (1) is positive-definite, it is positive when $x_1 = 1$ and $x_2 = ... = x_n = 0$; hence $b_{11} > 0$. Accordingly, the terms (2) may be written as[†]

$$b_{11}\left(x_1+\frac{b_{12}}{b_{11}}x_2+...+\frac{b_{1n}}{b_{11}}x_n\right)^2-\frac{1}{b_{11}}(b_{12}x_2+...+b_{1n}x_n)^2,$$

[†] It is essential to this step that b_{11} is not zero.

and so, if

$$\left.\begin{aligned} X_1 &= x_1+\frac{b_{12}}{b_{11}}x_2+\ldots+\frac{b_{1n}}{b_{11}}x_n, \\ X_r &= x_r \quad (r = 2,\ldots, n), \end{aligned}\right\} \tag{3}$$

we have

$$B(x,x) = b_{11}X_1^2+\text{ a quadratic form in } X_2,\ldots, X_n$$

$$= b_{11}X_1^2+\sum_{r=2}^{n}\sum_{s=2}^{n}\beta_{rs}X_rX_s, \tag{4}$$

say.

The transformation (3) is non-singular, the determinant of the transformation being

$$\begin{vmatrix} 1 & b_{12}/b_{11} & . & . & b_{1n}/b_{11} \\ 0 & 1 & . & . & 0 \\ . & . & . & . & . \\ 0 & 0 & . & . & 1 \end{vmatrix} = 1.$$

Hence (Theorem 40) the form (4) in X is positive-definite and $\beta_{22} > 0$ (put $X_2 = 1$, $X_r = 0$ when $r \neq 2$).

Working as before, we have

$$B(x,x) = b_{11}X_1^2+\beta_{22}\left(X_2+\frac{\beta_{23}}{\beta_{22}}X_3+\ldots+\frac{\beta_{2n}}{\beta_{22}}X_n\right)^2+$$

$$+\text{ a quadratic form in } X_3,\ldots, X_n.$$

Let this quadratic form in $X_3,\ldots, X_n$ be $\sum\sum\gamma_{rs}X_rX_s$; then, on writing

$$\left.\begin{aligned} Y_1 &= X_1, \\ Y_2 &= X_2+\frac{\beta_{23}}{\beta_{22}}X_3+\ldots+\frac{\beta_{2n}}{\beta_{22}}X_n, \\ Y_r &= X_r \quad (r = 3,\ldots, n), \end{aligned}\right\} \tag{5}$$

we have

$$B(x,x) = b_{11}Y_1^2+\beta_{22}Y_2^2+\sum_{r=3}^{n}\sum_{s=3}^{n}\gamma_{rs}Y_rY_s. \tag{6}$$

The transformation (5) is of unit modulus and so, by Theorem 40 applied to the form (4), form (6) is positive-definite and $\gamma_{33} > 0$ (put $Y_3 = 1$, $Y_r = 0$ when $r \neq 3$).

Working as before, we obtain as our next form

$$b_{11}Z_1^2+\beta_{22}Z_2^2+\gamma_{33}Z_3^2+\text{ a quadratic form in } Z_4,\ldots, Z_n, \tag{7}$$

wherein b_{11}, β_{22}, γ_{33} are positive. As a preparation for the next step, the coefficient of Z_4^2 can be shown to be positive by proving that (7) is a positive-definite form. Moreover, we have

$$Z_1 = Y_1 = X_1 = x_1+\frac{b_{12}}{b_{11}}x_2+\ldots+\frac{b_{1n}}{b_{11}}x_n,$$

$$Z_2 = Y_2 = x_2+\frac{\beta_{23}}{\beta_{22}}x_3+\ldots+\frac{\beta_{2n}}{\beta_{22}}x_n,$$

$$Z_3 = x_3+\frac{\gamma_{34}}{\gamma_{33}}x_4+\ldots+\frac{\gamma_{3n}}{\gamma_{33}}x_n,$$

$$Z_r = Y_r = X_r = x_r \quad (r > 3).$$

Proceeding step by step in this way, we finally obtain

$$B(x,x) = b_{11}\xi_1^2+\beta_{22}\xi_2^2+\gamma_{33}\xi_3^2+\ldots+\kappa_{nn}\xi_n^2, \tag{8}$$

wherein $b_{11},\ldots, \kappa_{nn}$ are positive, and

$$\left.\begin{array}{rl}
\xi_1 = & x_1+\dfrac{b_{12}}{b_{11}}x_2+\ldots+\dfrac{b_{1n}}{b_{11}}x_n, \\
\xi_2 = & \qquad\qquad x_2+\ldots+\dfrac{\beta_{2n}}{\beta_{22}}x_n, \\
\cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
\xi_n = & \qquad\qquad\qquad\qquad\qquad x_n.
\end{array}\right\} \tag{9}$$

We have thus transformed $B(x,x)$ into (8), which is of the form required by the theorem; moreover, the transformation $x \to \xi$ is (9), which is of unit modulus.

2. Necessary and sufficient conditions for a positive-definite form

2.1. THEOREM 42. *A set of necessary and sufficient conditions that the real quadratic form*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \quad (a_{ij} = a_{ji}) \tag{1}$$

be positive-definite is

$$a_{11} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \qquad \ldots, \qquad \begin{vmatrix} a_{11} & \cdot & \cdot & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \cdot & \cdot & a_{nn} \end{vmatrix} > 0.$$

If the form is positive-definite, then, as in § 1.2, there is a transformation

$$\begin{aligned} X_1 &= x_1+\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n, \\ X_2 &= \qquad\quad x_2+\ldots+\frac{\beta_{2n}}{\beta_{22}}x_n, \\ &\ldots\ldots\ldots\ldots \\ X_n &= \qquad\qquad\qquad\quad x_n, \end{aligned}$$

of unit modulus, whereby the form is transformed into

$$a_{11}X_1^2+\beta_{22}X_2^2+\ldots+\kappa_{nn}X_n^2, \tag{2}$$

and in this form a_{11}, β_{22},..., κ_{nn} are all positive. The discriminant of the form (2) is $a_{11}\beta_{22}\ldots\kappa_{nn}$. Hence, by Theorem 39,

$$\begin{vmatrix} a_{11} & . & . & a_{1n} \\ . & . & . & . \\ a_{n1} & . & . & a_{nn} \end{vmatrix} = a_{11}\beta_{22}\ldots\kappa_{nn},$$

and so is positive.

Now consider (1) when $x_n \equiv 0$. By the previous argument applied to a form in the $n-1$ variables x_1,..., x_{n-1},

$$\begin{vmatrix} a_{11} & . & . & a_{1,n-1} \\ . & . & . & . \\ a_{n-1,1} & . & . & a_{n-1,n-1} \end{vmatrix} = a_{11}\beta_{22}\ldots \text{ to } n-1 \text{ terms}$$

and so is positive.

Similarly, on putting $x_n \equiv 0$ and $x_{n-1} \equiv 0$,

$$\begin{vmatrix} a_{11} & . & . & a_{1,n-2} \\ . & . & . & . \\ a_{n-2,1} & . & . & a_{n-2,n-2} \end{vmatrix} = a_{11}\beta_{22}\ldots \text{ to } n-2 \text{ terms},$$

and so on. Hence the given set of conditions is necessary.

Conversely, if the set of conditions holds, then, in the first place, $a_{11} > 0$ and we may write, as in § 1.2,

$$\begin{aligned} A(x,x) &= a_{11}\left(x_1+\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n\right)^2+ \\ &\qquad\qquad\qquad + \text{a quadratic form in } x_2,\ldots, x_n \\ &= a_{11}X_1^2+\beta_{22}X_2^2+\ldots+\beta_{nn}X_n^2+2\beta_{23}X_2X_3+\ldots, \end{aligned} \tag{3}$$

say, where

$$X_1 = x_1+\frac{a_{12}}{a_{11}}x_2+\dots+\frac{a_{1n}}{a_{11}}x_n,$$

$$X_k = x_k \quad (k > 1).$$

Before we can proceed we must prove that β_{22} is positive. Consider the form and transformation when $x_k \equiv 0$ for $k > 2$. The discriminant of the form (3) is then $a_{11}\beta_{22}$, the modulus of the transformation is unity, and the discriminant of

$$a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2$$

is $a_{11}a_{22}-a_{12}^2$. Hence, by Theorem 39 applied to a form in the two variables x_1 and x_2 only,

$$a_{11}\beta_{22} = a_{11}a_{22}-a_{12}^2,$$

and so is positive by hypothesis. Hence β_{22} is positive, and we may write (3) as

$$a_{11}Y_1^2+\beta_{22}Y_2^2+\text{ a quadratic form in } Y_3,\dots, Y_n,$$

where

$$Y_1 = X_1,$$

$$Y_2 = X_2+\frac{\beta_{23}}{\beta_{22}}X_3+\dots+\frac{\beta_{2n}}{\beta_{22}}X_n,$$

$$Y_k = X_k \quad (k > 2).$$

That is, we may write

$$A(x,x) = a_{11}Y_1^2+\beta_{22}Y_2^2+\gamma_{33}Y_3^2+\dots+2\gamma_{34}Y_3Y_4+\dots, \tag{4}$$

where a_{11} and β_{22} are positive.

Consider the forms (3), (4) and the transformation from X to Y when $x_k \equiv 0$ for $k > 3$. The discriminant of (4) is $a_{11}\beta_{22}\gamma_{33}$, the modulus of the transformation is unity, and the discriminant of (3) is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \tag{5}$$

which is positive by hypothesis. Hence, by Theorem 39 applied to a form in the three variables x_1, x_2, x_3 only, the product $a_{11}\beta_{22}\gamma_{33}$ is equal to the determinant (5), and so is positive. Accordingly, γ_{33} is positive.

We may proceed thus, step by step, to the result that, when the set of conditions is satisfied, we may write $A(x,x)$ in the form

$$\alpha_{11}X_1^2+\beta_{22}X_2^2+...+\kappa_{nn}X_n^2, \tag{6}$$

wherein α_{11}, β_{22},..., κ_{nn} are positive and

$$\left.\begin{aligned} X_1 &= x_1+\frac{\alpha_{12}}{\alpha_{11}}x_2+...+\frac{\alpha_{1n}}{\alpha_{11}}x_n, \\ X_2 &= \qquad\quad x_2+...+\frac{\beta_{2n}}{\beta_{22}}x_n, \\ &. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ X_n &= \qquad\qquad\qquad\quad x_n. \end{aligned}\right\} \tag{7}$$

The form (6) is positive-definite in X_1,..., X_n and, since the transformation (7) is non-singular, $A(x,x)$ is positive-definite in x_1,..., x_n (Theorem 40).

2.2. We have considered the variables in the order x_1, x_2,..., x_n and we have begun with a_{11}. We might have considered the variables in the order x_n, x_{n-1},..., x_1 and begun with a_{nn}. We should then have obtained a different set of necessary and sufficient conditions, namely,

$$a_{nn} > 0, \qquad \begin{vmatrix} a_{n,n} & a_{n,n-1} \\ a_{n-1,n} & a_{n-1,n-1} \end{vmatrix} > 0, \qquad ....$$

Equally, any permutation of the order of the variables will give rise to a set of necessary and sufficient conditions.

2.3. The form $A(x,x)$ is negative-definite if the form $\{-A(x,x)\}$ is positive-definite. Accordingly, the form $A(x,x)$ is negative-definite if and only if

$$a_{11} < 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \qquad ....$$

EXAMPLES XI

1. Prove that each of the quadratic forms

(i) $6x^2+35y^2+11z^2+34yz$,

(ii) $6x^2+49y^2+51z^2-82yz+20zx-4xy$

is positive-definite, but that

(iii) $4x^2+9y^2+2z^2+8yz+6zx+6xy$

is not.

2. Prove that if

$$F \equiv \sum_{r=1}^{3} \sum_{s=1}^{3} a_{rs} x_r x_s \quad (a_{rs} = a_{sr})$$

is a positive-definite form, then

$$a_{11} F = (a_{11} x_1 + a_{12} x_2 + a_{13} x_3)^2 + K(x_2, x_3),$$

where K is a positive-definite form in x_2, x_3 whose discriminant is a_{11} times the discriminant of F.

HINT. Use the transformation $X_1 = a_{11} x_1 + a_{12} x_2 + a_{13} x_3$, $X_2 = x_2$, $X_3 = x_3$ and Theorem 39.

3. Prove that the discriminant of the quadratic form in x, y, z

$$F = (a_1 x + b_1 y + c_1 z)^2 + (a_2 x + b_2 y + c_2 z)^2$$

is the square (by rows) of a determinant whose columns are

$$a_1, b_1, c_1; \ a_2, b_2, c_2; \ 0, 0, 0.$$

4. *Extension of Example* 3. Prove that a sum of squares of r distinct linear forms in n variables is a quadratic form in n variables of zero discriminant whenever $r < n$.

5. *Harder.* Prove also that the rank of such a discriminant is, in general, equal to r: and that the exception to the general rule arises when the r distinct forms are not linearly independent.

6. Prove that the discriminant of

$$F = \sum_{r=1}^{n} \left(\sum_{s=1}^{n} a_{sr} x_s \right)^2$$

is not zero unless the forms

$$\sum_{s=1}^{n} a_{sr} x_s \quad (r = 1, \ldots, n)$$

are linearly dependent.

HINT. Compare Examples 3 and 4.

7. Prove that the discriminant of the quadratic form

$$\sum_{r \neq s} (x_r - x_s)^2 \quad (r, s = 1, \ldots, n)$$

is of rank $n-1$.

8. If $f(x_1, \ldots, x_n)$ is a function of n variables, f_i denotes $\partial f / \partial x_i$, and f_{ij} denotes $\partial^2 f / \partial x_i \partial x_j$ evaluated at $x_r = \alpha_r$ $(r = 1, \ldots, n)$, prove that f has a minimum at $x_r = \alpha_r$ provided that $\sum \sum f_{ij} \xi_i \xi_j$ is a positive-definite form and each f_i is zero. Write down conditions that $f(x, y)$ may be a minimum at the point $x = \alpha$, $y = \beta$.

9. By Example 12, p. 133, a Hermitian form has a real value for every set of values of the variables. It is said to be positive-definite when this value is positive for every set of values of the variables other than $x_1 = \ldots = x_n = 0$. Prove Theorem 40 for a positive-definite Hermitian form and any non-singular transformation.

10. Prove that a positive-definite Hermitian form can be transformed by a transformation of unit modulus into a form

$$c_{11}X_1\bar{X}_1+...+c_{nn}X_n\bar{X}_n$$

wherein each c_{rr} is positive.

HINT. Compare the proof of Theorem 41.

11. The determinant $|A| \equiv |a_{ij}|$ is called the discriminant of the Hermitian form $\sum\sum a_{ij}x_i\bar{x}_j$. Remembering that a Hermitian form is characterized by the matrix equation $A' = \bar{A}$, it follows, by considering the determinant equation

$$|\bar{A}| = |A'| = |A|,$$

that the discriminant is real.

Prove the analogue of Theorem 42 for Hermitian forms.

12. Prove, by analogy with Example 3, that the discriminant of the Hermitian form

$$H \equiv X\bar{X}+Y\bar{Y},$$

where $X = a_1x+b_1y+c_1z,\qquad Y = a_2x+b_2y+c_2z,$

is zero.

Obtain the corresponding analogues of Examples 4, 5, and 6.

CHAPTER XII

THE CHARACTERISTIC EQUATION AND CANONICAL FORMS

1. The λ equation of two quadratic forms

1.1. We have seen that the discriminant of a quadratic form $A(x,x)$ is multiplied by M^2 when the variables x are changed to variables X by a transformation of modulus M (Theorem 39).

If $A(x,x)$ and $C(x,x)$ are any two distinct quadratic forms in n variables and λ is an arbitrary parameter, the discriminant of the form $A(x,x)-\lambda C(x,x)$ is likewise multiplied by M^2 when the variables are submitted to a transformation of modulus M. Hence, if a transformation

$$x_i = \sum_{j=1}^{n} b_{ij} X_j, \qquad |b_{ij}| = M,$$

changes $\sum a_{rs} x_r x_s$, $\sum c_{rs} x_r x_s$ into $\sum \alpha_{rs} X_r X_s$, $\sum \gamma_{rs} X_r X_s$, then

$$M^2 \begin{vmatrix} a_{11}-\lambda c_{11} & . & . & a_{1n}-\lambda c_{1n} \\ . & . & . & . \\ a_{n1}-\lambda c_{n1} & . & . & a_{nn}-\lambda c_{nn} \end{vmatrix} = \begin{vmatrix} \alpha_{11}-\lambda\gamma_{11} & . & . & \alpha_{1n}-\lambda\gamma_{1n} \\ . & . & . & . \\ \alpha_{n1}-\lambda\gamma_{n1} & . & . & \alpha_{nn}-\lambda\gamma_{nn} \end{vmatrix}.$$

The equation $|A-\lambda C| = 0$, or, in full,

$$\begin{vmatrix} a_{11}-\lambda c_{11} & . & . & a_{1n}-\lambda c_{1n} \\ . & . & . & . \\ a_{n1}-\lambda c_{n1} & . & . & a_{nn}-\lambda c_{nn} \end{vmatrix} = 0,$$

is called THE λ EQUATION OF THE TWO FORMS.

What we have said above may be summarized in the theorem:

THEOREM 43. *The roots of the λ equation of any two quadratic forms in n variables are unaltered by a non-singular linear transformation.*

The coefficient of each power λ^r $(r = 0, 1,..., n)$ is multiplied by the square of the modulus of the transformation.

1.2. The λ equation of $A(x,x)$ and the form $x_1^2+...+x_n^2$ is called THE CHARACTERISTIC EQUATION of A. In full, the equation is

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22}-\lambda & . & . & a_{2n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nn}-\lambda \end{vmatrix} = 0.$$

The roots of this equation are called the LATENT ROOTS of the matrix A. The equation itself may be denoted by $|A-\lambda I| = 0$, where I is the unit matrix of order n.

The term independent of λ is the determinant $|a_{ik}|$, so that the characteristic equation has no zero root when $|a_{ik}| \neq 0$. When $|a_{ik}| = 0$, so that the rank of the matrix $[a_{ik}]$ is $r < n$, there is at least one zero latent root; as we shall prove later, there are then $n-r$ zero roots.

2. The reality of the latent roots

2.1. THEOREM 44. *If $[c_{ik}]$ is the matrix of a positive-definite form $C(x,x)$ and $[a_{ik}]$ is any symmetrical matrix with real elements, all the roots of the λ equation*

$$\begin{vmatrix} a_{11}-c_{11}\lambda & a_{12}-c_{12}\lambda & . & . & a_{1n}-c_{1n}\lambda \\ a_{21}-c_{21}\lambda & a_{22}-c_{22}\lambda & . & . & a_{2n}-c_{2n}\lambda \\ . & . & . & . & . \\ a_{n1}-c_{n1}\lambda & a_{n2}-c_{n2}\lambda & . & . & a_{nn}-c_{nn}\lambda \end{vmatrix} = 0$$

are real.

Let λ be any root of the equation. Then the determinant vanishes and (Theorem 11) there are numbers $Z_1, Z_2,..., Z_n$, not all zero, such that

$$\sum_{s=1}^{n} (a_{rs}-c_{rs}\lambda)Z_s = 0 \quad (r = 1,..., n);$$

that is,
$$\lambda \sum_{s=1}^{n} c_{rs} Z_s = \sum_{s=1}^{n} a_{rs} Z_s \quad (r = 1,..., n). \tag{1}$$

Multiply each equation (1) by $\bar{Z}_r$, the conjugate complex† of Z_r, and add the results. If $Z_r = X_r+iY_r$, where X_r and Y_r are real, we obtain terms of two distinct types, namely,

$$Z_r\bar{Z}_r = (X_r+iY_r)(X_r-iY_r) = X_r^2+Y_r^2,$$

$$\begin{aligned} Z_r\bar{Z}_s+\bar{Z}_r Z_s &= (X_r+iY_r)(X_s-iY_s)+(X_r-iY_r)(X_s+iY_s) \\ &= 2(X_r X_s+Y_r Y_s). \end{aligned}$$

---

† When $Z = X+iY$, $\bar{Z} = X-iY$; $i = \sqrt{(-1)}$, X and Y real.

Hence the result of the operation is

$$\lambda\Big\{\sum_{r=1}^{n} c_{rr}(X_r^2+Y_r^2)+2\sum_{r<s}^{n} c_{rs}(X_r X_s+Y_r Y_s)\Big\}$$
$$= \Big\{\sum_{r=1}^{n} a_{rr}(X_r^2+Y_r^2)+2\sum_{r<s}^{n} a_{rs}(X_r X_s+Y_r Y_s)\Big\},$$

or, on using an obvious notation,

$$\lambda\{C(X,X)+C(Y,Y)\} = A(X,X)+A(Y,Y). \qquad (2)$$

Since, by hypothesis, $C(x,x)$ is positive-definite and since the numbers $Z_r = X_r+iY_r$ $(r = 1,..., n)$ are not all zero, the coefficient of λ in equation (2) is positive. Moreover, since each a_{ik} is real, the right-hand side of (2) is real and hence λ must be real.

NOTE. If the coefficient of λ in (2) were zero, then (2) would tell us nothing about the reality of λ. It is to preclude this that we require $C(x,x)$ to be positive-definite.

COROLLARY. *If both $A(x,x)$ and $C(x,x)$ are positive-definite forms, every root of the given equation is positive.*

2.2. When $c_{rr} = 1$ and $c_{rs} = 0$ $(r \neq s)$, the form $C(x,x)$ is $x_1^2+x_2^2+...+x_n^2$ and is positive-definite. Thus Theorem 44 contains, as a special case, the following theorem, one that has a variety of applications in different parts of mathematics.

THEOREM 45. *When $[a_{ik}]$ is any symmetrical matrix with real elements, every root λ of the equation $|A-\lambda I| = 0$, that is, of*

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22}-\lambda & . & . & a_{2n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{nn}-\lambda \end{vmatrix} = 0 \quad (a_{rs} = a_{sr}),$$

is real.

When $[a_{ik}]$ is the matrix of a positive-definite form, every root is positive.

3. Canonical forms

3.1. In this section we prove that, if $A = [a_{ik}]$ is a square matrix of order n and rank r, then there is a non-singular transformation from variables $x_1,..., x_n$ to variables $X_1,..., X_n$ that changes $A(x,x)$ into

$$g_1 X_1^2+...+g_r X_r^2, \qquad (3)$$

where $g_1,\ldots, g_r$ are distinct from zero. We show, further, that if the a_{ik} are elements in a particular field F (such as the field of real numbers), then so are the coefficients of the transformation and the resulting g_i. We call (3) a CANONICAL FORM.

The first step in the proof is to show that any given quadratic form can be transformed into one whose coefficient of x_1^2 is not zero. This is done in §§ 3.2 and 3.3.

3.2. Elementary transformations.

We shall have occasion to use two particular types of transformation.

TYPE I. The transformation

$$x_1 = X_r, \quad x_r = X_1, \quad x_s = X_s \quad (s \neq 1, r)$$

is non-singular; its modulus is -1, as is seen by writing the determinant in full and interchanging the first and rth columns. Moreover, each coefficient of the transformation is either 1 or 0, and so belongs to every field of numbers F.

If in the quadratic form $\sum\sum a_{rs}x_r x_s$ one of the numbers $a_{11},\ldots, a_{nn}$ is not zero, then either $a_{11} \neq 0$ or a suitable transformation of type I will change the quadratic form into $\sum\sum b_{rs}X_r X_s$, wherein $b_{11} \neq 0$.

If, in the quadratic form $\sum\sum a_{rs}x_r x_s$, all the numbers $a_{11},\ldots, a_{nn}$ are zero, but one number a_{rs} $(r \neq s)$ is not zero, then a suitable transformation of type I will change the form into $\sum\sum b_{rs}X_r X_s$, wherein every b_{rr} is zero but one of the numbers $b_{12},\ldots, b_{1n}$ is not zero.

TYPE II. The transformation in n variables

$$x_1 = X_1+X_s, \quad x_s = X_1-X_s, \quad x_t = X_t \quad (t \neq s, 1)$$

is non-singular. Its modulus is the determinant which

(i) in the first row, has 1 in the first and sth columns and 0 elsewhere,

(ii) in the sth row, has 1 in the first, -1 in the sth, and 0 in every other column,

(iii) in the tth row $(t \neq s, 1)$ has 1 in the principal diagonal position and 0 elsewhere.

The value of such a determinant is ± 2. Moreover, each element of the determinant is 1, 0, or -1 and so belongs to every field of numbers F.

If, in the quadratic form $\sum\sum b_{rs}x_r x_s$, all the numbers $b_{11},\ldots,$ b_{nn} are zero, but one b_{1s} is not zero, then a suitable transformation of type II will express the form as $\sum\sum c_{rs}X_r X_s$, wherein c_{11} is not zero.

3.3. The first step towards the canonical form.

The foregoing elementary transformations enable us to transform every quadratic form into one in which a_{11} is not zero. For consider any given form

$$A(x,x)\equiv\sum_{r=1}^{n}\sum_{s=1}^{n}a_{rs}x_r x_s \quad (a_{rs}=a_{sr}). \tag{1}$$

If one a_{rr} is not zero, a transformation of type I changes (1) into a form $B(X,X)$ in which b_{11} is not zero.

If every a_{rr} is zero but at least one a_{rs} is not zero, a transformation of type I followed by one of type II changes (1) into a form $C(Y,Y)$ in which c_{11} is not zero. The product of these two transformations changes (1) directly into $C(Y,Y)$ and the modulus of this transformation is ± 2.

We summarize these results in a theorem.

THEOREM 46. *Every quadratic form $A(x,x)$, with coefficients in a given field F, and having one a_{rs} not zero, can be transformed by a non-singular transformation with coefficients in F into a form $B(X,X)$ whose coefficient b_{11} is not zero.*

3.4. Proof of the main theorem.

DEFINITION 7. *The rank of a quadratic form is defined to be the rank of the matrix of its coefficients.*

THEOREM 47. *A quadratic form in n variables and of rank r, with coefficients in a given field F, can be transformed by a non-singular transformation, with coefficients in F, into the form*

$$\alpha_1 X_1^2+\ldots+\alpha_r X_r^2, \tag{1}$$

where $\alpha_1,\ldots,$ α_r are numbers in F and no one of them is equal to zero.

Let the form be $\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_i x_j$; and let A denote the matrix

$[a_{ij}]$. If every a_{ij} is zero, the rank of A is zero and there is nothing to prove.

If one a_{ij} is not zero (Theorem 46), there is a non-singular transformation with coefficients in F that changes $A(x,x)$ into

$$\alpha_1 X_1^2+2\sum_{i=2}^{n}\alpha_{1i}X_1X_i+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}X_iX_j,$$

where $\alpha_1 \neq 0$. This may be written as

$$\alpha_1\left(X_1+\frac{\alpha_{12}}{\alpha_1}X_2+\ldots+\frac{\alpha_{1n}}{\alpha_1}X_n\right)^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}X_iX_j,$$

and the non-singular transformation, with coefficients in F,

$$Y_1 = X_1+\frac{\alpha_{12}}{\alpha_1}X_2+\ldots+\frac{\alpha_{1n}}{\alpha_1}X_n,$$

$$Y_i = X_i \quad (i = 2,\ldots, n),$$

enables us to write $A(x,x)$ as

$$\alpha_1 Y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}Y_iY_j. \tag{2}$$

Moreover, the transformation direct from x to Y is the product of the separate transformations employed; hence it is non-singular and has its coefficients in F, and every b is in F.

If every b_{ij} in (2) is zero, then (2) reduces to the form (1); the question of rank we defer until the end.

If one b_{ij} is not zero, we may change the form $\sum\sum b_{ij}Y_iY_j$ in $n-1$ variables in the same way as we have just changed the original form in n variables. We may thus show that there is a non-singular transformation

$$Z_i = \sum_{j=2}^{n} l_{ij}Y_j \quad (i = 2,\ldots, n), \tag{3}$$

with coefficients in F, which enables us to write

$$\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}Y_iY_j = \alpha_2 Z_2^2+\sum_{i=3}^{n}\sum_{j=3}^{n}c_{ij}Z_iZ_j,$$

where $\alpha_2 \neq 0$. The equations (3), together with $Y_1 = Z_1$, constitute a non-singular transformation of the n variables $Y_1,\ldots, Y_n$ into $Z_1,\ldots, Z_n$. Hence there is a non-singular transformation, the product of all the transformations so far employed, which has coefficients in F and changes $A(x,x)$ into

$$\alpha_1 Z_1^2+\alpha_2 Z_2^2+\sum_{i=3}^{n}\sum_{j=3}^{n} c_{ij} Z_i Z_j, \tag{4}$$

wherein $\alpha_1 \neq 0$ and $\alpha_2 \neq 0$, and every c is in F.

On proceeding in this way, one of two things must happen. Either we arrive at a form

$$\alpha_1 X_1^2+...+\alpha_k X_k^2+\sum_{i=k+1}^{n}\sum_{j=k+1}^{n} d_{ij} X_i X_j, \tag{5}$$

wherein $k < n$, $\alpha_1 \neq 0,..., \alpha_k \neq 0$, but every $d_{ij} = 0$, in which case (5) reduces to

$$\alpha_1 X_1^2+...+\alpha_k X_k^2 \quad (k < n);$$

or we arrive after n steps at a form

$$\alpha_1 X_1^2+...+\alpha_n X_n^2.$$

In either circumstance we arrive at a final form

$$\alpha_1 X_1^2+...+\alpha_k X_k^2 \quad (k \leqslant n) \tag{6}$$

by a product of transformations each of which is non-singular and has its coefficients in the given field F.

It remains to prove that the number k in (6) is equal to r, the rank of the matrix A. Let B denote the transformation whereby we pass from $A(x,x)$ to the form

$$\alpha_1 X_1^2+...+\alpha_k X_k^2+\sum_{i=k+1}^{n} 0.X_i^2. \tag{7}$$

Then the matrix of (7) is (Theorem 37) $B'AB$. Since B is non-singular, the matrix $B'AB$ has the same rank as A (Theorem 34), that is, r. But the matrix of the quadratic form (7) consists of $\alpha_1,..., \alpha_k$ in the first k places of the principal diagonal and zero elsewhere, so that its rank is k. Hence $k = r$.

COROLLARY. *If $A(x,x)$ is a quadratic form in n variables, with its coefficients in a given field F, and if its discriminant is zero, it can be transformed by a non-singular transformation, with coefficients in F, into a form (in $n-1$ variables at most)*

$$\alpha_1 X_1^2+...+\alpha_{n-1} X_{n-1}^2,$$

where $\alpha_1,..., \alpha_{n-1}$ are numbers in F.

This corollary is, of course, merely a partial statement of the theorem itself; since the discriminant of $A(x,x)$ is zero, the rank of the form is $n-1$ or less. We state the corollary with a view to its immediate use in the next section, wherein F is the field of real numbers.

4. The simultaneous reduction of two real quadratic forms

THEOREM 48. *Let $A(x,x)$, $C(x,x)$ be two real quadratic forms in n variables and let $C(x,x)$ be positive-definite. Then there is a real, non-singular transformation that expresses the two forms as*

$$\lambda_1 X_1^2+\ldots+\lambda_n X_n^2, \qquad X_1^2+\ldots+X_n^2,$$

where $\lambda_1,\ldots,\lambda_n$ are the roots of $|A-\lambda C| = 0$ and are all real.

The roots of $|A-\lambda C| = 0$ are all real (Theorem 44). Let λ_1 be any one root. Then $A(x,x)-\lambda_1 C(x,x)$ is a real quadratic form whose discriminant is zero and so (Theorem 47, Corollary) there is a real non-singular transformation from $x_1,\ldots, x_n$ to $Y_1,\ldots, Y_n$ such that

$$A(x,x)-\lambda_1 C(x,x) = \alpha_2 Y_2^2+\ldots+\alpha_n Y_n^2,$$

where the α's are real numbers. Let this same transformation, when applied to $C(x,x)$, give

$$C(x,x) = \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} Y_i Y_j. \tag{1}$$

Then we have, for an arbitrary λ,

$$\begin{aligned} A(x,x)-\lambda C(x,x) &= A(x,x)-\lambda_1 C(x,x)+(\lambda_1-\lambda)C(x,x) \\ &= \sum_{i=2}^{n} \alpha_i Y_i^2+(\lambda_1-\lambda)\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} Y_i Y_j. \end{aligned} \tag{2}$$

Since $C(x,x)$ is positive-definite in the variables x, it is also positive-definite in the variables Y (Theorem 40). Hence γ_{11} is positive and we may use the transformation

$$Z_1 = Y_1+\frac{\gamma_{12}}{\gamma_{11}}Y_2+\ldots+\frac{\gamma_{1n}}{\gamma_{11}}Y_n,$$
$$Z_s = Y_s \quad (s = 2,\ldots, n).$$

This enables us to write (2) in the form (Chap. XI, § 1.2)

$$A(x,x)-\lambda C(x,x) = \phi(Z_2,\ldots, Z_n)+(\lambda_1-\lambda)\{\gamma_{11} Z_1^2+\psi(Z_2,\ldots, Z_n)\},$$

where ϕ and ψ are real quadratic forms in $Z_2,\ldots, Z_n$.

Since this holds for an arbitrary value of λ, we have

$$\left.\begin{aligned} A(x,x) &= \lambda_1 \gamma_{11} Z_1^2 + \phi + \lambda_1 \psi = \lambda_1 \gamma_{11} Z_1^2 + \theta, \\ C(x,x) &= \quad \gamma_{11} Z_1^2 + \psi, \end{aligned}\right\} \tag{3}$$

where θ and ψ denote quadratic forms in $Z_2,\ldots, Z_n$.

This is the first step towards the forms we wish to establish.

Before we proceed to the next step we observe two things:

(i) ψ is a positive-definite form in the $n-1$ variables $Z_2,\ldots, Z_n$; for

$$\gamma_{11} Z_1^2 + \psi(Z_2,\ldots, Z_n)$$

is obtained by a non-singular transformation from the form (1), which is positive-definite in $Y_1,\ldots, Y_n$.

(ii) The roots of $|\theta - \lambda\psi| = 0$, together with $\lambda = \lambda_1$, account for all the roots of $|A - \lambda C| = 0$; for the forms (3) derive from $A(x,x)$ and $C(x,x)$ by a non-singular transformation and so (Theorem 43) the roots of

$$|(\lambda_1 \gamma_{11} - \lambda \gamma_{11}) Z_1^2 + \theta - \lambda \psi| = 0,$$

that is, of

$$\begin{vmatrix} \gamma_{11}(\lambda_1 - \lambda) & 0 & . & . & 0 \\ 0 & \theta_{22} - \lambda\psi_{22} & . & . & \theta_{2n} - \lambda\psi_{2n} \\ . & . & . & . & . \\ 0 & \theta_{n2} - \lambda\psi_{n2} & . & . & \theta_{nn} - \lambda\psi_{nn} \end{vmatrix} = 0,$$

are the roots of $|A - \lambda C| = 0$. If λ_1 is a repeated root of $|A - \lambda C| = 0$, then λ_1 is also a root of $|\theta - \lambda\psi| = 0$.

Thus we may, by using the first step with θ and ψ in place of A and C, reduce θ, ψ to forms

$$\left.\begin{aligned} \theta &= \lambda_2 \alpha_2 U_2^2 + \theta'(U_3,\ldots, U_n), \\ \psi &= \quad \alpha_2 U_2^2 + \phi'(U_3,\ldots, U_n), \end{aligned}\right\} \tag{4}$$

where α_2 is positive, λ_2 is a root of $|A - \lambda C|$, and the transformation between $Z_2,\ldots, Z_n$ and $U_2,\ldots, U_n$ is real and non-singular. When we adjoin the equation $U_1 = Z_1$ we have a real, non-singular transformation between $Z_1,\ldots, Z_n$ and $U_1,\ldots, U_n$. Hence there is a real non-singular transformation (the product of all transformations so far used) from $x_1,\ldots, x_n$ to $U_1,\ldots, U_n$ such that

$$\left.\begin{aligned} A(x,x) &= \lambda_1 \alpha_1 U_1^2 + \lambda_2 \alpha_2 U_2^2 + f(U_3,\ldots, U_n), \\ C(x,x) &= \quad \alpha_1 U_1^2 + \quad \alpha_2 U_2^2 + F(U_3,\ldots, U_n), \end{aligned}\right\} \tag{5}$$

where α_1 and α_2 are positive; λ_1 and λ_2 are roots, not necessarily distinct, of $|A-\lambda C|$; f and F are real quadratic forms. As at the first stage, we can show that F is positive-definite and that the roots of $|f-\lambda F| = 0$ together with λ_1 and λ_2 account for all the roots of $|A-\lambda C| = 0$.

Thus we may proceed step by step to the reduction of $A(x,x)$ and $C(x,x)$ by real non-singular transformations to the two forms

$$A(x,x) = \lambda_1 \alpha_1 Y_1^2+\lambda_2 \alpha_2 Y_2^2+...+\lambda_n \alpha_n Y_n^2,$$
$$C(x,x) = \alpha_1 Y_1^2+ \alpha_2 Y_2^2+...+ \alpha_n Y_n^2,$$

wherein each α_i is positive and $\lambda_1,..., \lambda_n$ account for all the roots of $|A-\lambda C| = 0$.

Finally, the real transformation

$$X_r = \sqrt{\alpha_r}.Y_r$$

gives the required result, namely,

$$A(x,x) = \lambda_1 X_1^2+...+\lambda_n X_n^2, \qquad C(x,x) = X_1^2+...+X_n^2.$$

5. Orthogonal transformations

If, in Theorem 48, the positive-definite form is $x_1^2+...+x_n^2$, the transformation envisaged by the theorem transforms $x_1^2+...+x_n^2$ into $X_1^2+...+X_n^2$. Such a transformation is called an ORTHOGONAL TRANSFORMATION. We shall examine such transformations in Chapter XIII; we shall see that they are necessarily non-singular. Meanwhile, we note an important theorem.

THEOREM 49. *A real quadratic form $A(x, x)$ in n variables can be reduced by a real orthogonal transformation to the form*

$$\lambda_1 X_1^2+...+\lambda_n X_n^2,$$

where $\lambda_1,..., \lambda_n$ account for all the roots of $|A-\lambda I| = 0$. Moreover, all the roots are real.

The proof consists in writing $x_1^2+...+x_n^2$ for $C(x,x)$ in Theorem 48. [See also Chapter XV, § 3.]

6. The number of non-zero latent roots

If B is the matrix of the orthogonal transformation whereby $A(x,x)$ is reduced to the form

$$\lambda_1 X_1^2+...+\lambda_n X_n^2, \tag{1}$$

then (Theorem 37) $B'AB$ is the matrix of the form (1) and this has $\lambda_1,..., \lambda_n$ in its leading diagonal and zero elsewhere. Its rank is the number of λ's that are not zero. But since B is non-singular, the rank of $B'AB$ is equal to the rank of A. Hence *the rank of A is equal to the number of non-zero roots of the characteristic equation* $|A-\lambda I| = 0$.

7. The signature of a quadratic form

7.1. As we proved in Theorem 47 (§3.4), a real quadratic form of rank r can be transformed by a real non-singular transformation into the form

$$\alpha_1 X_1^2+...+\alpha_r X_r^2, \tag{1}$$

wherein $\alpha_1,..., \alpha_r$ are real and not zero.

As a glance at §3.3 will show, there are, in general, many different ways of effecting such a reduction; even at the first step we have a wide choice as to which non-zero a_{rs} we select to become the non-zero b_{11} or c_{11}, as the case may be.

The theorem we shall now prove establishes the fact that, starting from the one given form $A(x,x)$, the number of positive α's and the number of negative α's in (1) is independent of the method of reduction.

THEOREM 50. *If a given real quadratic form of rank r is reduced by two real, non-singular transformations, B_1 and B_2 say, to the forms*

$$\alpha_1 X_1^2+...+\alpha_r X_r^2, \tag{2}$$

$$\beta_1 Y_1^2+...+\beta_r Y_r^2, \tag{3}$$

the number of positive α's is equal to the number of positive β's and the number of negative α's is equal to the number of negative β's.

Let n be the number of variables $x_1,..., x_n$ in the initial quadratic form; let μ be the number of positive α's and ν the number of positive β's. Let the variables X, Y be so numbered that the positive α's and β's come first. Then, since (2) and (3) are transformations of the same initial form, we have

$$\alpha_1 X_1^2+...+\alpha_\mu X_\mu^2-|\alpha_{\mu+1}|X_{\mu+1}^2-...-|\alpha_r|X_r^2$$
$$\equiv \beta_1 Y_1^2+...+\beta_\nu Y_\nu^2-|\beta_{\nu+1}|Y_{\nu+1}^2-...-|\beta_r|Y_r^2. \tag{4}$$

Now suppose, contrary to the theorem, that $\mu > \nu$. Then the $n+\nu-\mu$ equations

$$Y_1 = 0, \quad ..., \quad Y_\nu = 0, \qquad X_{\mu+1} = 0, \quad ..., \quad X_n = 0 \tag{5}$$

are homogeneous equations in the n variables† $x_1,..., x_n$. There are less equations than there are variables and so (Chap. VIII, § 5) the equations have a solution $x_1 = \xi_1,..., x_n = \xi_n$ in which $\xi_1,..., \xi_n$ are not all zero.

Let X'_r, Y'_r be the values of X_r, Y_r when $x = \xi$. Then, from (4) and (5),

$$\alpha_1 X_1'^2+...+\alpha_\mu X_\mu'^2 = -|\beta_{\nu+1}|Y_{\nu+1}'^2-...-|\beta_r|Y_r'^2,$$

which is impossible unless each X' and Y' is zero.

Hence either we have a contradiction or

$$\left.\begin{aligned} X'_1 = 0, \quad ..., \quad X'_\mu = 0, \\ X'_{\mu+1} = 0, \quad ..., \quad X'_n = 0. \end{aligned}\right\} \tag{6}$$

and, from (5),

But (6) means that the n equations

$$X_1 = 0, \quad ..., \quad X_n = 0, \tag{7}$$

say

$$\sum_{k=1}^{n} l_{ik} x_k = 0 \quad (i = 1,..., n)$$

in full, have a solution $x_r = \xi_r$ in which $\xi_1,..., \xi_n$ are not all zero; this, in turn, means that the determinant $|l_{ik}| = 0$, which is a contradiction of the hypothesis that the transformation

$$X_i = \sum_{k=1}^{n} l_{ik} x_k \quad (i = 1,..., n)$$

is non-singular.

Hence the assumption that $\mu > \nu$ leads to a contradiction. Similarly, the assumption that $\nu > \mu$ leads to a contradiction. Accordingly, $\mu = \nu$ and the theorem is proved.

7.2. One of the ways of reducing a form of rank r is by the orthogonal transformation of Theorem 49. This gives the form

$$\lambda_1 X_1^2+...+\lambda_r X_r^2,$$

where $\lambda_1,..., \lambda_r$ are the non-zero roots of the characteristic equation.

Hence the number of positive α's, or β's, in Theorem 50 is the number of positive latent roots of the form.

---

† Each X and Y is a linear form in $x_1,..., x_n$.

7.3. We have proved that, associated with every quadratic form are two numbers P and N, the number of positive and of negative coefficients in any canonical form. The sum of P and N is the rank of the form.

DEFINITION 8. *The number* $P-N$ *is called the* SIGNATURE *of the form.*

7.4. We conclude with a theorem which states that any two quadratic forms having the same rank and signature are, in a certain sense, EQUIVALENT.

THEOREM 51. *Let* $A_1(x,x)$, $A_2(y,y)$ *be two real quadratic forms having the same rank* r *and the same signature* s. *Then there is a real non-singular transformation* $x = By$ *that transforms* $A_1(x,x)$ *into* $A_2(y,y)$.

When $A_1(x,x)$ is reduced to its canonical form it becomes

$$\alpha_1 X_1^2+\ldots+\alpha_\mu X_\mu^2-\beta_{\mu+1} X_{\mu+1}^2-\ldots-\beta_r X_r^2, \tag{1}$$

where the α's and β's are positive, where $\mu = \frac{1}{2}(s+r)$, and where the transformation from x to X is real and non-singular.

The real transformation

$$\xi_1 = X_1\sqrt{\alpha_1},\quad \ldots,\quad \xi_\mu = X_\mu\sqrt{\alpha_\mu},$$
$$\xi_{\mu+1} = X_{\mu+1}\sqrt{\beta_{\mu+1}},\quad \ldots;\quad \xi_r = X_r\sqrt{\beta_r}$$

changes (1) to $\quad \xi_1^2+\ldots+\xi_\mu^2-\xi_{\mu+1}^2-\ldots-\xi_r^2. \qquad (2)$

There is, then, a real non-singular transformation, say $x = C_1\xi$, that changes $A_1(x,x)$ into (2).

Equally, there is a real non-singular transformation, say $y = C_2\xi$, that changes $A_2(y,y)$ into (2). Or, on considering the reciprocal process, (2) is changed into $A_2(y,y)$ by the transformation $\xi = C_2^{-1}y$. Hence

$$x = C_1 C_2^{-1}y$$

changes $A_1(x,x)$ into $A_2(y,y)$.

EXAMPLES XII

1. Two forms $A(x,x)$ and $C(x,x)$ have

$$a_{rs} = c_{rs} \quad \text{when } r = 1,\ldots, k \text{ and } s = 1,\ldots, n.$$

Prove that the λ equation of the two forms has k roots equal to unity.

2. Write down the λ equation of the two forms $ax^2+2hxy+by^2$ and $a'x^2+2h'xy+b'y^2$ and prove, by elementary methods, that the roots are real when $a > 0$ and $ab-h^2 > 0$.

HINT. Let

$$ax^2+2hxy+by^2 = a(x-\alpha_1 y)(x-\beta_1 y),$$
$$a'x^2+2h'xy+b'y^2 = a'(x-\alpha_2 y)(x-\beta_2 y).$$

The condition for real roots in λ becomes

$$(ab'+a'b-2hh')^2-4(ab-h^2)(a'b'-h'^2) \geqslant 0,$$

which can be written as

$$a^2a'^2(\alpha_1-\alpha_2)(\beta_1-\beta_2)(\alpha_1-\beta_2)(\beta_1-\alpha_2) \geqslant 0.$$

When $ab-h^2 > 0$, α_1 and β_1 are conjugate complexes and the result can be proved by writing $\alpha_1 = \gamma+i\delta$, $\beta_1 = \gamma-i\delta$.

In fact, the λ roots are real save when α_1, β_1, α_2, β_2 are real and the roots with suffix 1 separate those with suffix 2.

3. Prove that the latent roots of the matrix A, where

$$A(x,x) = 6x_1^2+35x_2^2+11x_3^2+34x_2x_3,$$

are all positive.

4. Prove that when

$$A(x,x) = 4x_1^2+9x_2^2+2x_3^2+8x_2x_3+6x_3x_1+6x_1x_2$$

two latent roots are positive and one is negative.

5. $$A \equiv \sum\sum a_{rs}x_r x_s; \qquad X_r = \tfrac{1}{2}\partial A/\partial x_r;$$

$$A_k = \begin{vmatrix} a_{11} & . & . & a_{1k} & X_1 \\ . & . & . & . & . \\ a_{k1} & . & . & a_{kk} & X_k \\ X_1 & . & . & X_k & A \end{vmatrix}, \qquad \xi_k = \begin{vmatrix} a_{11} & . & . & a_{1k} \\ . & . & . & . \\ a_{k-1,1} & . & . & a_{k-1,k} \\ X_1 & . & . & X_k \end{vmatrix}$$

By subtracting from the last row multiplies of the other rows and then, for A_k, by subtracting from the last column multiples of the other columns, prove that A_k and ξ_k are independent of the variables $x_1,\dots, x_k$ when $k < n$. Prove also that $A_n \equiv 0$.

6. With the notation of Example 5, and with

$$D_k = \begin{vmatrix} a_{11} & . & . & a_{1k} \\ . & . & . & . \\ a_{k1} & . & . & a_{kk} \end{vmatrix},$$

show, by means of Theorem 18 applied to A_k, that

$$\begin{vmatrix} A_{k-1} & \xi_k \\ \xi_k & D_k \end{vmatrix} = D_{k-1}A_k.$$

7. Use the result of Example 6 and the result $A_n = 0$ of Example 5 to prove that, when no D_k is zero,

$$A = \sum_{k=1}^{n} \frac{\xi_k^2}{D_{k-1}D_k}.$$

Consider what happens when the rank, r, of A is less than n and $D_1,..., D_r$ are all distinct from zero.

✓8. Transform $4x_2x_3+2x_3x_1+6x_1x_2$ into a form $\sum\sum a_{rs}X_rX_s$ with $a_{11} \neq 0$, and hence (by the method of Chap. XI, § 1.2) find a canonical form and the signature of the form.

9. Show, by means of Example 4, that

$$4x_1^2+9x_2^2+2x_3^2+8x_2x_3+6x_3x_1+6x_1x_2$$

is of rank 3 and signature 1. Verify Theorem 50 for any two independent reductions of this form to a canonical form.

10. Prove that a quadratic form is the product of two linear factors if and only if its rank does not exceed 2.

HINT. Use Theorem 47.

13. Prove that the discriminant of a Hermitian form $A(x,\bar{x})$, when it is transformed to new variables by means of transformations $x = BX$, $\bar{x} = \bar{B}\bar{X}$, is multiplied by $|B|\times|\bar{B}|$.

Deduce the analogue of Theorems 43 and 44 for Hermitian forms.

12. Prove that, if A is a Hermitian matrix, then all the roots of $|A-\lambda I| = 0$ are real.

HINT. Compare Theorem 45 and use

$$C(x,\bar{x}) = \sum_{r=1}^{n} x_r\bar{x}_r.$$

13. Prove the analogues of Theorems 46 and 47 for Hermitian forms.

14. $A(x,\bar{x})$, $C(x,\bar{x})$ are Hermitian forms, of which C is positive-definite. Prove that there is a non-singular transformation that expresses the two forms as

$$\lambda_1X_1\bar{X}_1+...+\lambda_nX_n\bar{X}_n, \qquad X_1\bar{X}_1+...+X_n\bar{X}_n,$$

where $\lambda_1,...,\lambda_n$ are the roots of $|A-\lambda C| = 0$ and are all real.

15. *An example of some importance in analytical dynamics.*† Show that when a quadratic form in $m+n$ variables,

$$2T = \sum a_{rs}x_rx_s \quad (a_{rs} = a_{sr}),$$

is expressed in terms of $\xi_1,..., \xi_m, x_{m+1},..., x_{m+n}$, where $\xi_r = \partial T/\partial x_r$, there are no terms involving the product of a ξ by an x.

Solution. The result is easily proved by careful manipulation. Use the summation convention: let the range of r and s be $1,..., m$; let the range of u and t be $m+1,..., m+n$. Then $2T$ may be expressed as

$$2T = a_{rs}x_rx_s+2a_{ru}x_ry_u+a_{ut}y_uy_t,$$

where, as an additional distinguishing mark, we have written y instead of x whenever the suffix exceeds m.

We are to express T in terms of the y's and new variables ξ_s given by

$$\xi_s = a_{sr}x_r+a_{su}y_u \quad (s = 1,..., m).$$

---

† Cf. Lamb, *Higher Mechanics*, § 77: The Routhian function. I owe this example to Mr. J. Hodgkinson.

Multiply by A_{sr}, the co-factor of a_{sr} in $\Delta = |a_{rs}|$, a determinant of order m, and add: we get
$$\Delta x_r = A_{sr}\xi_s - A_{sr}a_{su}y_u.$$

Now
$$\begin{aligned} 2T &= x_r(a_{rs}x_s + 2a_{ru}y_u) + a_{ut}y_u y_t \\ &= x_r(\xi_r + a_{ru}y_u) + a_{ut}y_u y_t, \\ 2\Delta T &= (A_{sr}\xi_s - A_{sr}a_{su}y_u)(\xi_r + a_{ru}y_u) + a_{ut}y_u y_t\Delta \\ &= A_{sr}\xi_s\xi_r + y_u(a_{ru}A_{sr}\xi_s - a_{su}A_{sr}\xi_r) + \phi(y,y), \end{aligned}$$
where $\phi(y,y)$ denotes a quadratic form in the y's only.

But, since $a_{rs} = a_{sr}$, we also have $A_{sr} = A_{rs}$. Thus the term multiplying y_u in the above may be written as
$$a_{ru}A_{rs}\xi_s - a_{su}A_{sr}\xi_r,$$
and this is zero, since both r and s are dummy suffixes. Hence, $2T$ may be written in the form
$$2T = (A_{sr}/\Delta)\xi_s\xi_r + b_{ut}y_u y_t,$$
where r, s run from 1 to m and u, t run from $m+1$ to $m+n$.

CHAPTER XIII

ORTHOGONAL TRANSFORMATIONS

1. Definition and elementary properties

1.1. We recall the definition of the previous chapter.

DEFINITION 9. *A transformation* $x = AX$ *that transforms* $x_1^2+...+x_n^2$ *into* $X_1^2+...+X_n^2$ *is called an* ORTHOGONAL TRANSFORMATION. *The matrix A is called an* ORTHOGONAL MATRIX.

The best-known example of such a transformation occurs in analytical geometry. When (x, y, z) are the coordinates of a point P referred to rectangular axes Ox, Oy, Oz and (X, Y, Z) are its coordinates referred to rectangular axes OX, OY, OZ, whose direction-cosines with regard to the former axes are (l_1, m_1, n_1), (l_2, m_2, n_2), (l_3, m_3, n_3), the two sets of coordinates are connected by the equations

$$\begin{aligned} x &= l_1 X+l_2 Y+l_3 Z, \\ y &= m_1 X+m_2 Y+m_3 Z, \\ z &= n_1 X+n_2 Y+n_3 Z. \end{aligned}$$

Moreover, $\qquad x^2+y^2+z^2 = X^2+Y^2+Z^2 = OP^2.$

1.2. The matrix of an orthogonal transformation must have some special property. This is readily obtained.

If $A \equiv [a_{rs}]$ is an orthogonal matrix, and $x = AX$,

$$\sum_{r=1}^{n} X_r^2 = \sum_{r=1}^{n} x_r^2 = \sum_{r=1}^{n} (a_{r1} X_1+...+a_{rn} X_n)^2$$

for every set of values of the variables X_r. Hence

$$\left.\begin{aligned} a_{1s}^2+...+a_{ns}^2 &= 1 \quad (s = 1,..., n), \\ a_{1s} a_{1t}+...+a_{ns} a_{nt} &= 0 \quad (s \neq t). \end{aligned}\right\} \tag{1}$$

These are the relations that mark an orthogonal matrix: they are equivalent to the matrix equation

$$A'A = I, \tag{2}$$

as is seen by forming the matrix product $A'A$.

1.3. When (2) is satisfied A' is *the* reciprocal of A (Theorem 25), so that AA' is also equal to the unit matrix, and from this fact follow the relations (by writing AA' in full)

$$\left.\begin{aligned} a_{s1}^2+\ldots+a_{sn}^2 &= 1 \quad (s = 1,\ldots, n), \\ a_{s1}a_{t1}+\ldots+a_{sn}a_{tn} &= 0 \quad (s \neq t). \end{aligned}\right\} \tag{3}$$

When $n = 3$ and the transformation is the change of axes noted in § 1.1, the relations (1) and (3) take the well-known forms

$$l_1^2+l_2^2+l_3^2 = 1, \qquad l_1 m_1+l_2 m_2+l_3 m_3 = 0,$$
$$l_1^2+m_1^2+n_1^2 = 1, \qquad l_1 l_2+m_1 m_2+n_1 n_2 = 0,$$

and so on.

1.4. We now give four theorems that embody important properties of an orthogonal matrix.

Theorem 52. *A necessary and sufficient condition for a square matrix A to be orthogonal is $AA' = I$.*

This theorem follows at once from the work of §§ 1.2, 1.3.

Corollary. *Every orthogonal transformation is non-singular.*

Theorem 53. *The product of two orthogonal transformations is an orthogonal transformation.*

Let $x = AX$, $X = BY$ be orthogonal transformations. Then

$$AA' = I, \qquad BB' = I.$$

Hence
$$\begin{aligned} (AB)(AB)' &= ABB'A' \quad \text{(Theorem 23)} \\ &= AIA' \\ &= AA' = I, \end{aligned}$$

and the theorem is proved.

Theorem 54. *The modulus of an orthogonal transformation is either $+1$ or -1.*

If $AA' = I$ and $|A|$ is the determinant of the matrix A, then $|A|.|A'| = 1$. But $|A'| = |A|$, and hence $|A|^2 = 1$.

Theorem 55. *If λ is a latent root of an orthogonal transformation, then so is $1/\lambda$.*

Let A be an orthogonal matrix; then $AA' = I$, and so

$$A' = A^{-1}. \tag{4}$$

By Theorem 36, Corollary, the characteristic (latent) roots of A' are the reciprocals of those of A. But the characteristic equation of A' is a determinant which, on interchanging its rows and columns, becomes the characteristic equation of A.

Hence the n latent roots of A are the reciprocals of the n latent roots of A, and the theorem follows. (An alternative proof is given in Example 6, p. 168.)

2. The standard form of an orthogonal matrix

2.1. In order that A may be an orthogonal matrix the equations (1) of § 1.2 must be satisfied. There are

$$n+{}^nC_2 = n+\frac{n(n-1)}{2} = \frac{n(n+1)}{2}$$

of these equations. There are n^2 elements in A. We may therefore expect† that the number of independent constants necessary to define A completely will be

$$n^2-\frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

If the general orthogonal matrix of order n is to be expressed in terms of some other type of matrix, we must look for a matrix that has $\frac{1}{2}n(n-1)$ independent elements. Such a matrix is the skew-symmetric matrix of order n; that is,

$$[b_{ik}], \qquad b_{ik} = -b_{ki}.$$

For example, when $n = 3$,

$$\begin{bmatrix} 0 & b_{12} & b_{13} \\ -b_{12} & 0 & b_{23} \\ -b_{13} & -b_{23} & 0 \end{bmatrix}$$

has $1+2 = 3$ independent elements, namely, those lying above the leading diagonal. The number of such elements in the general case is

$$1+2+...+(n-1) = \tfrac{1}{2}n(n-1).$$

---

† The method of counting constants indicates what results to *expect*: it rarely *proves* those results, and in the crude form we have used above it certainly proves nothing.

THEOREM 56. *If S is a skew-symmetric matrix of order n, then, provided that $I+S$ is non-singular,*

$$O = (I-S)(I+S)^{-1}$$

is an orthogonal matrix of order n.

Since S is skew-symmetric,

$$S' = -S, \qquad (I-S)' = I+S, \qquad (I+S)' = I-S;$$

and when $I+S$ is non-singular it has a reciprocal $(I+S)^{-1}$. Hence the non-singular matrix O being defined by

$$O = (I-S)(I+S)^{-1},$$

we have

$$O' = \{(I+S)^{-1}\}'(I-S)' \quad \text{(Theorem 23)}$$

$$= (I-S)^{-1}(I+S) \quad \text{(Theorem 27)}$$

and $$OO' = (I-S)(I+S)^{-1}(I-S)^{-1}(I+S). \tag{5}$$

Now $I-S$ and $I+S$ are commutative,† and, by hypothesis, $I+S$ is non-singular. Hence‡

$$(I-S)(I+S)^{-1} = (I+S)^{-1}(I-S),$$

and, from (5),

$$OO' = (I+S)^{-1}(I-S)(I-S)^{-1}(I+S) = (I+S)^{-1}(I+S) = I.$$

Hence O is an orthogonal matrix (Theorem 52).

NOTE. When the elements of S are real, $I+S$ cannot be singular. This fact, proved as a lemma in § 2.2, was well known to Cayley, who discovered Theorem 56.

2.2. LEMMA. *If S is a real skew-symmetric matrix, $I+S$ is non-singular.*

Consider the determinant Δ obtained by writing down S and replacing the zeros of the principal diagonal by x; e.g., with $n = 3$,

$$\Delta = \begin{vmatrix} x & a & b \\ -a & x & c \\ -b & -c & x \end{vmatrix}, \qquad S = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

---

† $(I-S)(I+S) = I^2+IS-SI-S^2 = I^2-S^2 = (I+S)(I-S)$ $[SI = IS]$.

‡ If $AQ = QA$ and Q is non-singular, then $AQ^{-1} = Q^{-1}A$.

For $$Q^{-1}AQ = Q^{-1}QA = A$$

and so $$AQ^{-1} = (Q^{-1}AQ)Q^{-1} = Q^{-1}AQQ^{-1} = Q^{-1}A.$$

The differential coefficient of Δ with respect to x is the sum of n determinants (Chap. I, § 7), each of which is equal to a skew-symmetric determinant of order $n-1$ when $x = 0$. Thus, when $n = 3$,

$$\frac{d\Delta}{dx} = \begin{vmatrix} 1 & 0 & 0 \\ -a & x & c \\ -b & -c & x \end{vmatrix} + \begin{vmatrix} x & a & b \\ 0 & 1 & 0 \\ -b & -c & x \end{vmatrix} + \begin{vmatrix} x & a & b \\ -a & x & c \\ 0 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} x & c \\ -c & x \end{vmatrix} + \begin{vmatrix} x & b \\ -b & x \end{vmatrix} + \begin{vmatrix} x & a \\ -a & x \end{vmatrix},$$

and the latter become skew-symmetric determinants of order 2 when $x = 0$.

Differentiating again, in the general case, we see that $d^2\Delta/dx^2$ is the sum of $n(n-1)$ determinants each of which is equal to a skew-symmetric determinant of order $n-2$ when $x = 0$; and so on.

By Maclaurin's theorem, we then have

$$\Delta = \Delta_0 + x\sum_1 + x^2\sum_2 + \ldots + x^n, \tag{6}$$

where Δ_0 is a skew-symmetric determinant of order n, $\sum_1$ a sum of skew-symmetric determinants of order $n-1$, and so on.

But a skew-symmetric determinant of odd order is equal to zero and one of even order is a perfect square (Theorems 19, 21). Hence

$$\left.\begin{array}{ll} n \textit{ even} & \Delta = P_0 + x^2P_2 + \ldots + x^n, \\ n \textit{ odd} & \Delta = xP_1 + x^3P_3 + \ldots + x^n, \end{array}\right\} \tag{7}$$

where P_0, P_1,... are either squares or the sums of squares, and so P_0, P_1,... are, in general, positive and, though they may be zero in special cases, they cannot be negative.

The lemma follows on putting $x = 1$.

2.3. THEOREM 57. *Every real orthogonal matrix A can be expressed in the form*

$$J(I-S)(I+S)^{-1},$$

where S is a skew-symmetric matrix and J is a matrix having ± 1 in each diagonal place and zero elsewhere.

As a preliminary to the proof we establish a lemma that is true for all square matrices, orthogonal or not.

Lemma. *Given a square matrix A it is possible to choose a matrix J, having ± 1 in each diagonal place and zero elsewhere, so that -1 is not a latent root of the matrix JA.*

Multiplication by J merely changes the signs of the elements of a matrix row by row; e.g.

$$\begin{bmatrix} 1 & . & . \\ . & -1 & . \\ . & . & -1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ -a_{21} & -a_{22} & -a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{bmatrix}.$$

Bearing this in mind, we see that, either the lemma is true or, for every possible combination of row signs, we must have

$$\begin{vmatrix} \pm a_{11}+1 & \pm a_{12} & . & . & \pm a_{1n} \\ \pm a_{21} & \pm a_{22}+1 & . & . & \pm a_{2n} \\ . & . & . & . & . \\ \pm a_{n1} & \pm a_{2n} & . & . & \pm a_{nn}+1 \end{vmatrix} = 0. \tag{8}$$

But we can show that (8) is impossible. Suppose it is true. Then, adding the two forms of (8), (i) with *plus* in the first row, (ii) with *minus* in the first row, we obtain

$$\begin{vmatrix} \pm a_{22}+1 & . & . & . & \pm a_{2n} \\ . & . & . & . & . \\ \pm a_{2n} & . & . & . & \pm a_{nn}+1 \end{vmatrix} = 0 \tag{9}$$

for every possible combination of row signs. We can proceed by a like argument, reducing the order of the determinant by unity at each step, until we arrive at $\pm a_{nn}+1 = 0$: but it is impossible to have both $+a_{nn}+1 = 0$ and $-a_{nn}+1 = 0$. Hence (8) cannot be true.

2.4. *Proof of Theorem* 57. Let A be a given orthogonal matrix with real elements. If A has a latent root -1, let $J_1 A \equiv A_1$ be a matrix whose latent roots are all different from -1, where J_1 is of the same type as the J of the lemma.

Now J_1 is non-singular and its reciprocal J_1^{-1} is also a matrix having ± 1 in the diagonal places and zero elsewhere, so that

$$A = J_1^{-1} J_1 A = J_1^{-1} A_1 = J A_1,$$

where $J = J_1^{-1}$ and is of the type required by Theorem 57. Moreover, A_1, being derived from A by a change of signs of certain rows, is orthogonal and we have chosen it so that the matrix A_1+I is non-singular. Hence it remains to prove that when A_1 is a real orthogonal matrix such that A_1+I is non-singular, we can choose a skew-symmetric matrix S so that

$$A_1 = (I-S)(I+S)^{-1}.$$

To do this, let

$$S = (I-A_1)(I+A_1)^{-1}. \tag{10}$$

The matrices on the right of (10) are commutative† and we may write, without ambiguity,

$$S = \frac{I-A_1}{I+A_1}, \qquad S' = \frac{I-A_1'}{I+A_1'}. \tag{11}$$

Further, $A_1 A_1' = A_1' A_1 = I$, since A_1 is orthogonal. Thus A_1 and A_1' are commutative and we may work with A_1, A_1', and I as though they were ordinary numbers and so obtain, on using the relation $A_1 A_1' = I$,

$$S+S' = \frac{I-A_1}{I+A_1}+\frac{I-A_1'}{I+A_1'}$$

$$= \frac{2I-2A_1 A_1'}{I+A_1+A_1'+A_1 A_1'} = 0.$$

Hence, when S is defined by (10), we have $S = -S'$; that is, S is a skew-symmetric matrix. Moreover, from (10),

$$S+SA_1 = I-A_1$$

and so, since $I+S$ is non-singular‡ and $I-S$, $I+S$ are commutative, we have

$$A_1 = (I-S)(I+S)^{-1}, \tag{12}$$

the order|| of the two matrices on the right of (12) being immaterial.

We have thus proved the theorem.

2.4. Combining Theorems 56 and 57 in so far as they relate to matrices with real elements we see that

'*If S is a real skew-symmetric matrix, then*

$$A = J(I-S)(I+S)^{-1}$$

---

† Compare the footnote on p. 163.

‡ Compare § 2.2.

|| Compare the footnote on p. 163.

is a real orthogonal matrix, and every real orthogonal matrix can be so written.'

2.5. THEOREM 58. *The latent roots of a real orthogonal matrix are of unit modulus.*

Let $[a_{rs}]$ be a real orthogonal matrix. Then (Chap. X, § 7), if λ is a latent root of the matrix, the equations

$$\lambda x_r = a_{r1}\,x_1+\ldots+a_{rn}\,x_n \quad (r = 1,\ldots, n) \tag{13}$$

have a solution $x_1,\ldots, x_n$ other than $x_1 = \ldots = x_n = 0$.

These x are not necessarily real, but since a_{rs} is real we also have, on taking the conjugate complex of (13),

$$\bar{\lambda}\bar{x}_r = a_{r1}\,\bar{x}_1+\ldots+a_{rn}\,\bar{x}_n \quad (r = 1,\ldots, n).$$

By using the orthogonal relations (1) of § 1, we have

$$\lambda\bar{\lambda}\sum_{r=1}^{n} x_r\,\bar{x}_r = \sum_{r=1}^{n} x_r\,\bar{x}_r.$$

But not all of $x_1,\ldots, x_n$ are zero, and therefore $\sum x_r\,\bar{x}_r > 0$. Hence $\lambda\bar{\lambda} = 1$, which proves the theorem.

EXAMPLES XIII

1. Prove that the matrix

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

is orthogonal. Find its latent roots and verify that Theorems 54, 55, and 58 hold for this matrix.

2. Prove that the matrix

$$\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

is orthogonal.

3. Prove that when

$$S = \begin{bmatrix} 0 & -c & -b \\ c & 0 & -a \\ b & a & 0 \end{bmatrix},$$

$(I-S)(I+S)^{-1}$ is

$$\begin{bmatrix} \dfrac{1+a^2-b^2-c^2}{1+a^2+b^2+c^2} & \dfrac{2(c-ab)}{1+a^2+b^2+c^2} & \dfrac{2(ac+b)}{1+a^2+b^2+c^2} \\ \dfrac{-2(c+ab)}{1+a^2+b^2+c^2} & \dfrac{1-a^2+b^2-c^2}{1+a^2+b^2+c^2} & \dfrac{2(a-bc)}{1+a^2+b^2+c^2} \\ \dfrac{2(ac-b)}{1+a^2+b^2+c^2} & \dfrac{-2(a+bc)}{1+a^2+b^2+c^2} & \dfrac{1-a^2-b^2+c^2}{1+a^2+b^2+c^2} \end{bmatrix}.$$

Hence find the general orthogonal transformation of three variables.[†]

4. A given symmetric matrix is denoted by A; $X'AX$ is the quadratic form associated with A; S is a skew-symmetric matrix such that $I+SA$ is non-singular. Prove that, when $SA = AS$ and

$$X = \frac{I-SA}{I+SA}Y,$$

$X'AX = Y'AY$.

5. (*Harder.*)[‡] Prove that, when A is symmetric, S skew-symmetric, and $R = (A+S)^{-1}(A-S)$,

$$R'(A+S)R = A+S, \qquad R'(A-S)R = A-S;$$

$$R'AR = A, \qquad R'SR = S.$$

Prove also that, when $X = RY$,

$$X'AX = Y'AY.$$

6. Prove Theorem 55 by the following method. Let A be an orthogonal matrix. Multiply the determinant $|A-\lambda I|$ by $|A|$ and, in the resulting determinant, put $\lambda' = 1/\lambda$.

7. A transformation $x = UX$, $\bar{x} = \bar{U}\bar{X}$ which makes

$$\sum x_r \bar{x}_r = \sum X_r \bar{X}_r$$

is called a unitary transformation. Prove that the mark of a unitary transformation is the matrix equation $U\bar{U}' = I$.

8. Prove that the product of two unitary transformations is itself unitary.

9. Prove that the modulus M of a unitary transformation satisfies the equation $M\bar{M} = I$.

10. Prove that each latent root of a unitary transformation is of the form $e^{i\alpha}$, where α is real.

[†] Compare Lamb, *Higher Mechanics*, chapter i, examples 20 and 21, where the transformation is obtained from kinematical considerations.

[‡] These results are proved in Turnbull, *Theory of Determinants, Matrices, and Invariants*.

CHAPTER XIV

INVARIANTS AND COVARIANTS

1. Introduction

1.1. A detailed study of invariants and covariants is not possible in a single chapter of a small book. Such a study requires a complete book, and books devoted exclusively to that study already exist. All we attempt here is to introduce the ideas and to develop them sufficiently for the reader to be able to employ them, be it in algebra or in analytical geometry.

1.2. We have already encountered certain invariants. In Theorem 39 we proved that when the variables of a quadratic form are changed by a linear transformation, the discriminant of the form is multiplied by the square of the modulus of the transformation. Multiplication by a power of the modulus, not necessarily the square, as a result of a linear transformation is the mark of what is called an 'invariant'. Strictly speaking, the word should mean something that does not change at all; it is, in fact, applied to anything whose only change after a linear transformation of the variables is multiplication by a power of the modulus of the transformation. Anything that does not change at all after a linear transformation of the variables is called an 'ABSOLUTE INVARIANT'.

1.3. Definition of an algebraic form. Before we can give a precise definition of 'invariant' we must explain certain technical terms that arise.

DEFINITION 10. *A sum of terms, each of degree k in the n variables x, y,..., t,*

$$\sum a_{\alpha,\beta,\ldots,\lambda}\frac{k!}{\alpha!\,\beta!\ldots\lambda!}x^{\alpha}y^{\beta}\ldots t^{\lambda}, \tag{1}$$

wherein the a's are arbitrary constants and the sum is taken over all integer or zero sets of values of α, β,..., λ which satisfy the conditions

$$0 \leqslant \alpha \leqslant k, \quad \ldots, \quad 0 \leqslant \lambda \leqslant k, \quad \alpha+\beta+\ldots+\lambda = k, \tag{2}$$

is called an ALGEBRAIC FORM *of degree k in the variables x, y,..., t.*

For instance, $ax^2+2hxy+by^2$ is an algebraic form of degree 2 in x and y, and

$$\sum_{r=0}^{n} a_r \frac{n!}{r!\,(n-r)!} x^r y^{n-r} \tag{3}$$

is an algebraic form of degree n in x and y.

The multinomial coefficients $k!/\alpha!\ldots\lambda!$ in (1) and the binomial coefficients $n!/r!\,(n-r)!$ in (3) are not essential, but they lead to considerable simplifications in the resulting theory of the forms.

1.4. Notations. We shall use $F(a, x)$ and similar notations, such as $\phi(b, X)$, to denote an algebraic form; in the notation $F(a, x)$ the single a symbolizes the various constants in (1) and (3) and the single x symbolizes the variables. If we wish to mark the degree k and the number of variables n, we shall use $F(a, x)_n^k$.

An alternative notation is

$$(a_0, a_1,\ldots, a_n)(x, y)^n, \tag{4}$$

which is used to denote the form (3), and

$$(a_0,\ldots)(x,\ldots, t)^k, \quad \text{or} \quad (a_0,\ldots)(x_1,\ldots, x_n)^k,$$

which is used to denote the form (1). In this notation the index marks the degree of the form, while the number of variables is either shown explicitly, as in (4), or is inferred from the context.

Clarendon type, such as $\mathbf{x}$ or $\mathbf{X}$, will be used to denote single-column matrices with n rows, the elements in the rows of $\mathbf{x}$ or $\mathbf{X}$ being the variables of whatever algebraic forms are under discussion.

The standard linear transformation from variables $x_1,\ldots, x_n$ to variables $X_1,\ldots, X_n$, namely,

$$x_r = l_{r1} X_1 + \ldots + l_{rn} X_n \quad (r = 1,\ldots, n), \tag{5}$$

will be denoted by

$$\mathbf{x} = M\mathbf{X}, \tag{6}$$

M denoting the matrix of the coefficients l_{rs} in (5). As in previous chapters, $|M|$ will denote the determinant whose elements are the elements of the matrix M.

On making the substitutions (5) in a form

$$F(a,x) \equiv (a_0,\ldots)(x_1,\ldots,x_n)^k, \tag{7}$$

we obtain a form of degree k in $X_1,\ldots, X_n$: this form we denote by

$$G(A,X) \equiv (A_0,\ldots)(X_1,\ldots,X_n)^k. \tag{8}$$

The constants A in (8) depend on the constants a in (7) and on the coefficients l_{rs}. By its mode of derivation,

$$G(A,X) \equiv F(a,x)$$

for all values of the variables $\mathbf{x}$. For example, let

$$F(a,x) = a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2,$$

$$x_1 = l_{11}X_1+l_{12}X_2, \qquad x_2 = l_{21}X_1+l_{22}X_2;$$

then

$$F(a,x) \equiv G(A,X) = A_{11}X_1^2+2A_{12}X_1X_2+A_{22}X_2^2,$$

where
$$\begin{aligned} A_{11} &= a_{11}l_{11}^2+2a_{12}l_{11}l_{21}+a_{22}l_{21}^2, \\ A_{12} &= a_{11}l_{11}l_{12}+a_{12}(l_{11}l_{22}+l_{21}l_{12})+a_{22}l_{21}l_{22}, \\ A_{22} &= a_{11}l_{12}^2+2a_{12}l_{12}l_{22}+a_{22}l_{22}^2. \end{aligned}$$

In the sequel, the last thing we shall wish to do will be to calculate the actual expressions for the A's in terms of the a's: it will be sufficient for us to reflect that they could be calculated if necessary and to remember, at times, that the A's are linear in the a's.

1.5. Definition of an invariant.

DEFINITION 11. *A function of the coefficients a of the algebraic form $F(a,x)$ is said to be an invariant of the form if, whatever the matrix M of* (6) *may be, the same function of the coefficients A of the form $G(A,X)$ is equal to the original function* (*of the coefficients a*) *multiplied by a power of the determinant $|M|$, the power of $|M|$ in question being independent of M.*

For instance, in the example of § 1.4,

$$A_{11}A_{22}-A_{12}^2 = |M|^2(a_{11}a_{22}-a_{12}^2),$$

a result that may be proved either by laborious calculation or by an appeal to Theorem 39. Hence, in accordance with Definition 11, $a_{11}a_{22}-a_{12}^2$ is an invariant of the algebraic form $a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2$.

Again, we may consider not merely one form $F(a,x)$ and its transform $G(A,X)$, but several forms $F_1(a,x),\ldots, F_r(a,x)$ and their transforms $G_1(A,X),\ldots, G_r(A,X)$: as before, we write $G_r(A,X)$ for the result of substituting for x in terms of X in $F_r(a,x)$, the substitution being given by $\mathbf{x} = M\mathbf{X}$.

Definition 12. *A function of the coefficients a of a number of algebraic forms $F_r(a,x)$ is said to be an invariant (sometimes a joint-invariant) of the forms if, whatever the matrix M of the transformation $\mathbf{x} = M\mathbf{X}$, the same function of the coefficients A of the resulting forms $G_r(A,X)$ is equal to the original function (of the coefficients a) multiplied by a certain power of the determinant $|M|$.*

For example, if

$$F_1(a,x) = a_1x+b_1y, \qquad F_2(a,x) = a_2x+b_2y, \tag{9}$$

and the transformation $\mathbf{x} = M\mathbf{X}$ is, in full,

$$x = \alpha_1 X+\beta_1 Y, \qquad y = \alpha_2 X+\beta_2 Y, \tag{10}$$

so that

$$F_1(a,x) = G_1(A,X) = (a_1\alpha_1+b_1\alpha_2)X+(a_1\beta_1+b_1\beta_2)Y,$$

$$F_2(a,x) = G_2(A,X) = (a_2\alpha_1+b_2\alpha_2)X+(a_2\beta_1+b_2\beta_2)Y,$$

we see that

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1+b_1\alpha_2 & a_1\beta_1+b_1\beta_2 \\ a_2\alpha_1+b_2\alpha_2 & a_2\beta_1+b_2\beta_2 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \times \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix}. \tag{11}$$

Hence, in accordance with Definition 12, $a_1b_2-a_2b_1$ is a joint-invariant of the two forms a_1x+b_1y, a_2x+b_2y.

This example is but a simple case of the rule for forming the 'product' of two transformations: if we think of (9) as the transformation from variables F_1 and F_2 to variables x and y, and (10) as the transformation from x and y to X and Y, then (11) is merely a statement of the result proved in § 3.3 of Chapter X.

1.6. Covariants. An invariant is a function of the coefficients only. Certain functions which depend both on the coefficients and on the variables of a form $F(a,x)$, or of a number of forms $F_r(a,x)$, share with invariants the property of being unaltered, save for multiplication by a power of $|M|$, when the variables are changed by a substitution $\mathbf{x} = M\mathbf{X}$. Such functions are called COVARIANTS.

For example, let u and v be forms in the variables x and y. In these, make the substitutions

$$x = l_1 X + m_1 Y, \qquad y = l_2 X + m_2 Y,$$

so that u, v are expressed in terms of X and Y. Then we have, by the rules of differential calculus,

$$\frac{\partial u}{\partial X} = l_1 \frac{\partial u}{\partial x} + l_2 \frac{\partial u}{\partial y}, \qquad \frac{\partial v}{\partial X} = l_1 \frac{\partial v}{\partial x} + l_2 \frac{\partial v}{\partial y},$$

$$\frac{\partial u}{\partial Y} = m_1 \frac{\partial u}{\partial x} + m_2 \frac{\partial u}{\partial y}, \qquad \frac{\partial v}{\partial Y} = m_1 \frac{\partial v}{\partial x} + m_2 \frac{\partial v}{\partial y}.$$

The rules for multiplying determinants show at once that

$$\begin{vmatrix} \frac{\partial u}{\partial X} & \frac{\partial u}{\partial Y} \\ \frac{\partial v}{\partial X} & \frac{\partial v}{\partial Y} \end{vmatrix} = \begin{vmatrix} l_1 & l_2 \\ m_1 & m_2 \end{vmatrix} \times \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}. \tag{12}$$

The import of (12) is best seen if we write it in full. Let

$$\left.\begin{aligned} u &= a_0 x^n + n a_1 x^{n-1} y + \ldots + a_n y^n, \\ v &= b_0 x^n + n b_1 x^{n-1} y + \ldots + b_n y^n, \end{aligned}\right\} \tag{13}$$

and, when expressed in terms of X, Y,

$$u = A_0 X^n + n A_1 X^{n-1} Y + \ldots + A_n Y^n,$$
$$v = B_0 X^n + n B_1 X^{n-1} Y + \ldots + B_n Y^n.$$

Then (12) asserts that

$$\begin{vmatrix} A_0 X^{n-1} + \ldots + A_{n-1} Y^{n-1} & A_1 X^{n-1} + \ldots + A_n Y^{n-1} \\ B_0 X^{n-1} + \ldots + B_{n-1} Y^{n-1} & B_1 X^{n-1} + \ldots + B_n Y^{n-1} \end{vmatrix}$$
$$= \begin{vmatrix} l_1 & l_2 \\ m_1 & m_2 \end{vmatrix} \times \begin{vmatrix} a_0 x^{n-1} + \ldots + a_{n-1} y^{n-1} & a_1 x^{n-1} + \ldots + a_n y^{n-1} \\ b_0 x^{n-1} + \ldots + b_{n-1} y^{n-1} & b_1 x^{n-1} + \ldots + b_n y^{n-1} \end{vmatrix}.$$

The function

$$\begin{vmatrix} a_0 x^{n-1} + \ldots + a_{n-1} y^{n-1} & a_1 x^{n-1} + \ldots + a_n y^{n-1} \\ b_0 x^{n-1} + \ldots + b_{n-1} y^{n-1} & b_1 x^{n-1} + \ldots + b_n y^{n-1} \end{vmatrix}, \tag{14}$$

depending on the constants a_r, b_r of the forms u, v and on the variables x, y, is, apart from the factor $l_1 m_2 - l_2 m_1$, unaltered when the constants a_r, b_r are replaced by the constants A_r, B_r and the variables x, y replaced by the variables X, Y. Accordingly, we say that (14) is a covariant of the forms u, v.

1.7. Note on the definitions. When an invariant is defined as

> 'a function of the coefficients a, of a form $F(a,x)$, which is equal to the same function of the coefficients A, of the transform $G(A,X)$, *multiplied by a factor that depends only on the constants of the transformation*',

it can be proved that the factor in question *must* be a power of the modulus of the transformation.

In the present, elementary, treatment of the subject we have left aside the possibility of the factor being other than a power of $|M|$. It is, however, a point of some interest to note that the wider definition can be adopted with the same ultimate restriction on the nature of the factor, namely, to be a power of the modulus.

2. Examples of invariants and covariants

2.1. Jacobians. Let u, v,..., w be n forms in the n variables x, y,..., z. The determinant

$$\begin{vmatrix} u_x & v_x & . & . & w_x \\ u_y & v_y & . & . & w_y \\ . & . & . & . & . \\ u_z & v_z & . & . & w_z \end{vmatrix},$$

where u_x, u_y,... denote $\partial u/\partial x$, $\partial u/\partial y$,..., is called the Jacobian of the n forms. It is usually written as

$$\frac{\partial(u,v,...,w)}{\partial(x,y,...,z)}.$$

It is, as an extension of the argument of § 1.6 will show, a covariant of the forms u, v,..., w, and if $\mathbf{x} = M\mathbf{X}$,

$$\frac{\partial(u,v,...,w)}{\partial(X,Y,...,Z)} = |M|\frac{\partial(u,v,...,w)}{\partial(x,y,...,z)}. \tag{1}$$

2.2. Hessians. Let u be a form in the variables x, y,..., z and let u_{xx}, u_{xy},... denote $\partial^2 u/\partial x^2$, $\partial^2 u/\partial x\partial y$,.... Then the determinant

$$\begin{vmatrix} u_{xx} & u_{xy} & . & . & u_{xz} \\ u_{yx} & u_{yy} & . & . & u_{yz} \\ . & . & . & . & . \\ u_{zx} & u_{zy} & . & . & u_{zz} \end{vmatrix} \tag{2}$$

is called the Hessian of u and is a covariant of u; in fact, if we symbolize (2) as $H(u;x)$, we may prove that

$$H(u;X) = |M|^2 H(u;x). \tag{3}$$

For, by considering the Jacobian of the forms u_x, u_y,..., and calling them u_1, u_2,..., we find that § 2.1 gives

$$\frac{\partial(u_1, u_2,...)}{\partial(X, Y,...)} = |M| \frac{\partial(u_1, u_2,...)}{\partial(x, y,...)}. \tag{4}$$

But
$$\frac{\partial u_1}{\partial X} = \frac{\partial}{\partial X} \frac{\partial u}{\partial x},$$

and if
$$\begin{aligned} x &= l_1 X + l_2 Y + ..., \\ y &= m_1 X + m_2 Y + ..., \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned}$$

then
$$\left.\begin{aligned} \frac{\partial}{\partial X} &= l_1 \frac{\partial}{\partial x} + m_1 \frac{\partial}{\partial y} + ..., \\ \frac{\partial}{\partial Y} &= l_2 \frac{\partial}{\partial x} + m_2 \frac{\partial}{\partial y} + ..., \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned}\right\} \tag{5}$$

Hence
$$\begin{vmatrix} u_{XX} & u_{XY} & \cdot & \cdot & u_{XZ} \\ u_{YX} & u_{YY} & \cdot & \cdot & u_{YZ} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u_{ZX} & u_{ZY} & \cdot & \cdot & u_{ZZ} \end{vmatrix}$$

$$= \begin{vmatrix} l_1 & m_1 & \cdot & \cdot & \cdot \\ l_2 & m_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ l_n & m_n & \cdot & \cdot & \cdot \end{vmatrix} \times \begin{vmatrix} u_{xX} & u_{yX} & \cdot & \cdot & \cdot \\ u_{xY} & u_{yY} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u_{xZ} & u_{yZ} & \cdot & \cdot & \cdot \end{vmatrix},$$

as may be seen by multiplying the last two determinants by rows and applying (5). That is to say,

$$H(u;X) = |M| \frac{\partial(u_1, u_2,...)}{\partial(X, Y,...)},$$

which, on using (4), yields

$$H(u;X) = |M|^2 H(u;x).$$

2.3. Eliminant of linear forms. When the forms u, v,..., w of § 2.1 are linear in the variables we obtain the result that the determinant $|a_{rs}|$ of the linear forms

$$a_{r1}x_1+...+a_{rn}x_n \quad (r = 1,..., n)$$

is an invariant. It is sometimes called the ELIMINANT of the forms. As we have seen in Theorem 11, the vanishing of the eliminant is a necessary and sufficient condition for the equations

$$a_{r1}x_1+...+a_{rn}x_n = 0 \quad (r = 1,..., n)$$

to have a solution other than $x_1 = ... = x_n = 0$.

2.4. Discriminants of quadratic forms. When the form u of § 2.2 is a quadratic form, its Hessian is independent of the variables and so is an invariant. When

$$u = \sum a_{rs}x_r x_s \quad (a_{rs} = a_{sr}), \tag{6}$$

we have

$$\partial^2 u/\partial x_r \partial x_s = 2a_{rs},$$

so that the Hessian of (6) is, apart from a power of 2, the determinant $|a_{rs}|$, namely, the discriminant of the quadratic form.

This provides an alternative proof of Theorem 39.

2.5. Invariants and covariants of a binary cubic. A form in two variables is usually called a binary form : thus the general binary cubic is

$$a_0x^3+3a_1x^2y+3a_2xy^2+a_3y^3. \tag{7}$$

We can write down one covariant and deduce from it a second covariant and one invariant. The Hessian of (7) is (§ 2.2) a covariant; it is, apart from numerical factors,

$$\begin{vmatrix} a_0x+a_1y & a_1x+a_2y \\ a_1x+a_2y & a_2x+a_3y \end{vmatrix},$$

i.e.

$$(a_0a_2-a_1^2)x^2+(a_0a_3-a_1a_2)xy+(a_1a_3-a_2^2)y^2. \tag{8}$$

The discriminant of (8) is an invariant of (8): we may expect to find (we prove a general theorem later, § 3.6) that it is an invariant of (7) also.† The discriminant is

$$\begin{vmatrix} a_0a_2-a_1^2 & \frac{1}{2}(a_0a_3-a_1a_2) \\ \frac{1}{2}(a_0a_3-a_1a_2) & a_1a_3-a_2^2 \end{vmatrix}. \tag{9}$$

---

† The reader will clarify his ideas if he writes down the precise meaning of the phrases 'is an invariant of (7)', 'is an invariant of (8)'.

Again, the Jacobian of (7) and (8), which we may expect† to be a covariant of (7) itself, is, apart from numerical factors,

$$\begin{vmatrix} a_0 x^2+2a_1 xy+a_2 y^2 & a_1 x^2+2a_2 xy+a_3 y^2 \\ 2(a_0 a_2-a_1^2)x+(a_0 a_3-a_1 a_2)y & (a_0 a_3-a_1 a_2)x+2(a_1 a_3-a_2^2)y \end{vmatrix}$$

i.e. $(a_0^2 a_3-3a_0 a_1 a_2+2a_1^3)x^3+3(a_0 a_1 a_3-2a_0 a_2^2+a_1^2 a_2)x^2y+$

$$+3(2a_1^2 a_3-a_0 a_2 a_3-a_1 a_2^2)xy^2+(3a_1 a_2 a_3-a_0 a_3^2-2a_2^3)y^3. \quad (10)$$

At a later stage we shall show that an algebraical relation connects (7), (8), (9), and (10). Meanwhile we note an interesting (and in the advanced theory, an important) fact concerning the coefficients of the covariants (8) and (10).

If we write

$$c_0 x^2+c_1 xy+\tfrac{1}{2}c_2 y^2 \quad \text{for a quadratic covariant,}$$

$$c_0 x^3+c_1 x^2y+\frac{1}{2}c_2 xy^2+\frac{1}{2.3}c_3 y^3 \quad \text{for a cubic covariant,}$$

and so on for covariants of higher degree, the coefficients c_r are not a disordered set of numbers, as a first glance at (10) would suggest, but are given in terms of c_0 by means of the formula

$$c_{r+1} = \left(pa_1\frac{\partial}{\partial a_0}+(p-1)a_2\frac{\partial}{\partial a_1}+...+a_p\frac{\partial}{\partial a_{p-1}}\right)c_r, \quad (11)$$

where p is the degree of the form u. Thus, when we start with the cubic (7), for which $p = 3$, and take the covariant (8), so that $c_0 = a_0 a_2-a_1^2$, we find that

$$3a_1\frac{\partial c_0}{\partial a_0}+2a_2\frac{\partial c_0}{\partial a_1}+a_3\frac{\partial c_0}{\partial a_2} = a_0 a_3-a_1 a_2 = c_1$$

and $$3a_1\frac{\partial c_1}{\partial a_0}+2a_2\frac{\partial c_1}{\partial a_1}+a_3\frac{\partial c_1}{\partial a_2} = 2(a_1 a_3-a_2^2) = c_2$$

and the whole covariant (8) is thus derived, by differentiations from its leading term $(a_0 a_2-a_1^2)x^2$.

Exercise. Prove that when (10) is written as

$$c_0 x^3+c_1 x^2y+\tfrac{1}{2}c_2 xy^2+\tfrac{1}{6}c_3 y^3,$$

$$c_{r+1} = 3a_1\frac{\partial c_r}{\partial a_0}+2a_2\frac{\partial c_r}{\partial a_1}+a_3\frac{\partial c_r}{\partial a_2} \qquad (r = 0, 1, 2).$$

---

† Compare §§ 3.6, 3.7.

2.6. A method of forming joint-invariants.

Let

$$u = F(a,x) = (a_0,\ldots,a_m)(x_1,\ldots,x_n)^k, \tag{12}$$

and suppose we know that a certain polynomial function of the a's, say

$$\phi(a_0,\ldots,a_m),$$

is an invariant of $F(a,x)$. That is to say, when we substitute $\mathbf{x} = M\mathbf{X}$ in (12) we obtain a form

$$u = G(A,X) = (A_0,\ldots,A_m)(X_1,\ldots,X_n)^k, \tag{13}$$

and we suppose that the a of (12) and the A of (13) satisfy the relation

$$\phi(A_0,\ldots,A_m) = |M|^s\phi(a_0,\ldots,a_m), \tag{14}$$

where s is some fixed constant, independent of the matrix M.

Now (12) typifies any form of degree k in the n variables $x_1,\ldots, x_n$, and (14) may be regarded as a statement concerning the coefficients of any such form. Thus, if

$$v = F(a',x) = (a'_0,\ldots,a'_m)(x_1,\ldots,x_n)^k,$$

is transformed by $\mathbf{x} = M\mathbf{X}$ into

$$v = G(A',X) = (A'_0,\ldots,A'_m)(X_1,\ldots,X_n)^k,$$

the coefficients a', A' satisfy the relation

$$\phi(A'_0,\ldots,A'_m) = |M|^s\phi(a'_0,\ldots,a'_m). \tag{15}$$

Equally, when λ is an arbitrary constant, $u+\lambda v$ is a form of degree k in the variables $x_1,\ldots, x_n$. It may be written as

$$(a_0+\lambda a'_0,\ldots, a_m+\lambda a'_m)(x_1,\ldots, x_n)^k$$

and, after the transformation $\mathbf{x} = M\mathbf{X}$, it becomes

$$(A_0,\ldots, A_m)(X_1,\ldots, X_n)^k+\lambda(A'_0,\ldots, A'_m)(X_1,\ldots, X_n)^k,$$

that is,

$$(A_0+\lambda A'_0,\ldots, A_m+\lambda A'_m)(X_1,\ldots, X_n)^k.$$

Just as (15) may be regarded as being a mere change of notation in (14), so, by considering the invariant ϕ of the form $u+\lambda v$, we have

$$\phi(A_0+\lambda A'_0,\ldots, A_m+\lambda A'_m) = |M|^s\phi(a_0+\lambda a'_0,\ldots, a_m+\lambda a'_m), \tag{16}$$

which is a mere change of notation in (14).

Now each side of (16) is a polynomial in λ whose coefficients are functions of the A, a, and $|M|$; moreover, (16) is true for all values of λ and so the coefficients of λ^0, λ, λ^2,... on each side of (16) are equal. Hence the coefficient of each power of λ in the expansion of

$$\phi(a_0+\lambda a_0',\dots,\, a_m+\lambda a_m')$$

is a joint-invariant of u and v; for if this coefficient is

$$\phi_r(a_0,\dots,\, a_m,\, a_0',\dots,\, a_m'),$$

then (16) gives

$$\phi_r(A_0,\dots,\, A_m,\, A_0',\dots,\, A_m') = |M|^s\phi_r(a_0,\dots,\, a_m,\, a_0',\dots,\, a_m'). \quad (17)$$

The rules of the differential calculus enable us to formulate the coefficients of the powers of λ quite simply. Write

$$b_0 = a_0+\lambda a_0', \qquad \dots, \qquad b_m = a_m+\lambda a_m'.$$

Then

$$\frac{d}{d\lambda}\phi(a_0+\lambda a_0',\dots,\, a_m+\lambda a_m')$$
$$= \frac{\partial\phi}{\partial b_0}\frac{db_0}{d\lambda}+\dots+\frac{\partial\phi}{\partial b_m}\frac{db_m}{d\lambda}$$
$$= a_0'\frac{\partial\phi}{\partial b_0}+\dots+a_m'\frac{\partial\phi}{\partial b_m}, \quad (18)$$

and the value of (18) when $\lambda = 0$ is given by

$$\left(a_0'\frac{\partial}{\partial a_0}+\dots+a_m'\frac{\partial}{\partial a_m}\right)\phi(a_0,\dots,\, a_m). \quad (19)$$

Hence, by Maclaurin's theorem,

$$\phi(a_0+\lambda a_0',\dots,\, a_m+\lambda a_m')$$
$$= \phi(a_0,\dots,\, a_m)+\lambda\left(a_0'\frac{\partial}{\partial a_0}+\dots+a_m'\frac{\partial}{\partial a_m}\right)\phi(a_0,\dots,\, a_m)+$$
$$+\frac{\lambda^2}{2!}\left(a_0'\frac{\partial}{\partial a_0}+\dots+a_m'\frac{\partial}{\partial a_m}\right)^2\phi(a_0,\dots,\, a_m)+\dots,$$

the expansion terminating after a certain point since

$$\phi(a_0+\lambda a_0',\dots,\, a_m+\lambda a_m')$$

is a polynomial in λ. Hence, remembering the result (17), we have

$$\left(A_0'\frac{\partial}{\partial A_0}+\dots+A_m'\frac{\partial}{\partial A_m}\right)^r\phi(A_0,\dots,\, A_m)$$
$$= |M|^s\left(a_0'\frac{\partial}{\partial a_0}+\dots+a_m'\frac{\partial}{\partial a_m}\right)^r\phi(a_0,\dots,\, a_m). \quad (20)$$

EXAMPLES XIV A

Examples 1–4 can be worked by straightforward algebra and without appeal to any general theorem. In these examples the transformation is taken to be
$$x = l_1X+m_1Y, \qquad y = l_2X+m_2Y,$$
and so $|M|$ is equal to $l_1m_2-l_2m_1$; forms $ax+by$, $ax^2+2bxy+cy^2$ are transformed into $AX+BY$, $AX^2+2BXY+CY^2$; and so for other forms.

1. Prove that $ab'-a'b$ is an invariant (joint-invariant) of the forms $ax+by, a'x+b'y$.

Ans. $AB'-A'B = |M|(ab'-a'b)$.

2. Prove that $ab'^2-2ba'b'+ca'^2$ is an invariant of the forms
$$ax^2+2bxy+cy^2, \qquad a'x+b'y.$$
Ans. $AB'^2-2BA'B'+CA'^2 = |M|^2(ab'^2-2ba'b'+ca'^2)$.

3. Prove that $ab'+a'b-2hh'$ is an invariant of the two quadratic forms $ax^2+2hxy+by^2, a'x^2+2h'xy+b'y^2$.

Ans. $AB'+A'B-2HH' = |M|^2(ab'+a'b-2hh')$.

4. Prove that $b'(ax+by)-a'(bx+cy)$ is a covariant of the two forms $ax^2+2bxy+cy^2, a'x+b'y$.

Ans. $B'(AX+BY)-A'(BX+CY) = |M|\{b'(ax+by)-a'(bx+cy)\}$.

The remaining examples are not intended to be proved by sheer substitution.

5. Prove the result of Example 4 by considering the Jacobian of the two forms $ax^2+2bxy+cy^2$, $a'x+b'y$.

6. Prove that $(ab'-a'b)x^2+(ac'-a'c)xy+(bc'-b'c)y^2$ is the Jacobian of the forms $ax^2+2bxy+cy^2, a'x^2+2b'xy+c'y^2$.

7. Prove the result of Example 1 by considering the Jacobian of the forms $ax+by, a'x+b'y$.

8. Prove that $\sum(b_1c_2-b_2c_1)(ax+hy+gz)$ is the Jacobian (and so a covariant) of the three forms
$$ax^2+by^2+cz^2+2fyz+2gzx+2hxy, \qquad a_1x+b_1y+c_1z, \qquad a_2x+b_2y+c_2z.$$

Examples 9–12 are exercises on § 2.6.

9. Prove the result of Example 3 by first showing that $ab-h^2$ is an invariant of $ax^2+2hxy+by^2$.

10. Prove the result of Example 2 by considering the joint-invariant of $ax^2+2hxy+by^2$ and $(a'x+b'y)^2$.

11. Prove that $abc+2fgh-af^2-bg^2-ch^2$, i.e. the determinant
$$\Delta \equiv \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix},$$
is an invariant of the quadratic form $ax^2+by^2+cz^2+2fyz+2gzx+2hxy$, and find two joint-invariants of this form and of the form $a'x^2+...+2h'xy$.

NOTE. These joint-invariants, usually denoted by Θ, Θ',

$$\Theta = a'A+b'B+c'C+2f'F+2g'G+2h'H,$$
$$\Theta' = aA'+bB'+cC'+2fF'+2gG'+2hH',$$

where A, A',... denote co-factors of a, a',... in the discriminants Δ, Δ', are of some importance in analytical geometry. Compare Somerville, *Analytical Conics*, chapter xx.

12. Find a joint-invariant of the quadratic form

$$ax^2+by^2+cz^2+2fyz+2gzx+2hxy$$

and the linear form $lx+my+nz$. What is the geometrical significance of this joint-invariant?

Prove that the second joint-invariant, indicated by Θ' of Example 11, is identically zero.

13. Prove that

$$\begin{vmatrix} l^2 & lm & m^2 \\ 2ll' & lm'+l'm & 2mm' \\ l'^2 & l'm' & m'^2 \end{vmatrix} = (lm'-l'm)^3.$$

14. Prove that

$$\begin{vmatrix} \frac{\partial^4 u}{\partial x^4} & \frac{\partial^4 u}{\partial x^3 \partial y} & \frac{\partial^4 u}{\partial x^2 \partial y^2} \\ \frac{\partial^4 u}{\partial x^3 \partial y} & \frac{\partial^4 u}{\partial x^2 \partial y^2} & \frac{\partial^4 u}{\partial x \partial y^3} \\ \frac{\partial^4 u}{\partial x^2 \partial y^2} & \frac{\partial^4 u}{\partial x \partial y^3} & \frac{\partial^4 u}{\partial y^4} \end{vmatrix}$$

is a covariant of a form u whose degree exceeds 4 and an invariant of a form u of degree 4.

HINT. Multiplication by the determinant of Example 13 gives a determinant whose first row is $[x = lX+mY,\ y = l'X+m'Y]$

$$\frac{\partial^4 u}{\partial X^2 \partial x^2} \quad \frac{\partial^4 u}{\partial X \partial Y \partial x^2} \quad \frac{\partial^4 u}{\partial Y^2 \partial x^2}.$$

Compare § 2.2: here $\left(l\frac{\partial}{\partial x}+l'\frac{\partial}{\partial y}\right)^2 = \frac{\partial^2}{\partial X^2}$, and so on.

A multiplication of the determinant just obtained by that of Example 13 will give a determinant whose first row is

$$\frac{\partial^4 u}{\partial X^4} \quad \frac{\partial^4 u}{\partial X^3 \partial Y} \quad \frac{\partial^4 u}{\partial X^2 \partial Y^2}.$$

Hence $\mathbf{x} = M\mathbf{X}$ multiplies the initial determinant by $|M|^6$.

15. Prove that $ace+2bcd-ad^2-b^2e-c^3$, i.e.

$$\begin{vmatrix} a & b & c \\ b & c & d \\ c & d & e \end{vmatrix}$$

is an invariant of $(a,b,c,d,e,)(x,y)^4$.

16. Prove that

$$\begin{vmatrix} u_{xx} & u_{xy} & u_{yy} \\ v_{xx} & v_{xy} & v_{yy} \\ w_{xx} & w_{xy} & w_{yy} \end{vmatrix}$$

is a covariant of the forms u, v, w.

HINT. The multiplying factor is $|M|^3$.

3. Properties of invariants

3.1. Invariants are homogeneous forms in the coefficients.

Let $I(a_0,\dots, a_m)$ be an invariant of

$$u \equiv (a_0,\dots, a_m)(x_1,\dots, x_n)^k.$$

Let the transformation $\mathbf{x} = M\mathbf{X}$ change u into

$$(A_0,\dots, A_m)(X_1,\dots, X_n)^k.$$

Then, whatever M may be,

$$I(A_0,\dots, A_m) = |M|^s I(a_0,\dots, a_m), \tag{1}$$

the index s being independent of M.

Consider the transformation

$$x_1 = \lambda X_1, \quad x_2 = \lambda X_2, \quad \dots, \quad x_n = \lambda X_n,$$

wherein $|M| = \lambda^n$.

Since u is homogeneous and of degree k in $x_1,\dots, x_n$, the values of $A_0,\dots, A_m$ are, for this particular transformation,

$$a_0\lambda^k, \quad \dots, \quad a_m\lambda^k.$$

Hence (1) becomes

$$I(a_0\lambda^k,\dots, a_m\lambda^k) = \lambda^{ns} I(a_0,\dots, a_m). \tag{2}$$

That is to say, I is a homogeneous† function of its arguments $a_0,\dots, a_m$ and, if its degree is q, $\lambda^{kq} = \lambda^{ns}$, a result which proves, further, that s is given in terms of k, q, and n by

$$s = kq/n.$$

3.2. Weight: binary forms.

In the binary form

$$a_0 x^k + k a_1 x^{k-1} y + \dots + a_k y^k$$

the suffix of each coefficient is called its WEIGHT and the weight of a product of coefficients is defined to be the sum of

---

† The result is a well-known one. If the reader does not know the result, a proof can be obtained by assuming that I is the sum of I_1, I_2,..., each homogeneous and of degrees q_1, q_2,..., and then showing that the assumption contradicts (2).

the weights of its constituent factors. Thus $a_0^\lambda a_1^\mu a_2^\nu...$ is of weight $\mu+2\nu+...$.

A polynomial form in the coefficients is said to be ISOBARIC when each term has the same weight. Thus $a_0 a_2-a_1^2$ is isobaric, for $a_0 a_2$ is of weight $0+2$, and a_1^2 is of weight 2×1; we refer to the whole expression $a_0 a_2-a_1^2$ as being of weight 2, the phrase implying that each term is of that weight.

3.3. Invariants of binary forms are isobaric. Let $I(a_0,..., a_k)$ be an invariant of

$$u \equiv \sum_{r=0}^{k} \frac{k!}{r!\,(k-r)!} a_r x^{k-r} y^r, \tag{3}$$

and let I be a polynomial of degree q in $a_0,..., a_k$. Let the transformation $\mathbf{x} = M\mathbf{X}$ change u into

$$\sum_{r=0}^{k} \frac{k!}{r!\,(k-r)!} A_r X^{k-r} Y^r. \tag{4}$$

Then, by § 3.1 (there being now two variables, so that $n = 2$),

$$I(A_0,..., A_k) = |M|^{\frac{1}{2}kq} I(a_0,..., a_k) \tag{5}$$

whatever the matrix M may be.

Consider the transformation

$$x = X, \qquad y = \lambda Y,$$

for which $|M| = \lambda$. The values of $A_0,..., A_k$ are, for this particular transformation,

$$a_0, \quad a_1\lambda, \quad ..., \quad a_r\lambda^r, \quad ..., \quad a_k\lambda^k.$$

That is, *the power of λ associated with each coefficient A is equal to the weight of the coefficient.* Thus, for this particular transformation, the left-hand side of (5) is a polynomial in λ such that the coefficient of a power λ^w is a function of the a's of weight w. But the right-hand side of (5) is merely

$$\lambda^{\frac{1}{2}kq} I(a_0,..., a_k), \tag{6}$$

since, for this particular transformation, $|M| = \lambda$. Hence the only power of λ that can occur on the left-hand side of (5) is $\lambda^{\frac{1}{2}kq}$ and its coefficient must be a function of the a's of weight $\frac{1}{2}kq$. Moreover, since the left-hand and the right-hand side of

(5) are identical, this coefficient of $\lambda^{\frac{1}{2}kq}$ on the left of (5) must be $I(a_0,..., a_k)$.

Hence $I(a_0,..., a_k)$ is of weight $\frac{1}{2}kq$.

3.4. Notice that the weight of a polynomial in the a's must be an integer, so that we have, when $I(a_0,..., a_k)$ is a *polynomial* invariant of a binary form,

$$I(A_0,..., A_k) = |M|^s I(a_0,..., a_k), \tag{7}$$
$$s = \tfrac{1}{2}kq,$$

and $\quad \frac{1}{2}kq =$ weight of a polynomial = an integer.

Accordingly, the s of (7) must be an integer.

3.5. Weight: forms in more than two variables. In the form (§ 1.3)

$$\sum a_{\alpha,\beta,...,\lambda}\frac{k!}{\alpha!\,\beta!\,...\,\lambda!}x_1^\alpha x_2^\beta ... x_n^\lambda \tag{8}$$

the suffix λ (corresponding to the power of x_n in the term) is called the weight of the coefficient $a_{\alpha,\beta,...,\lambda}$ and the weight of a product of coefficients is defined to be the sum of the weights of its constituent factors.

Let $I(a)$ denote a polynomial invariant of (8) of degree q in the a's. Let the transformation $\mathbf{x} = MX$ change (8) into

$$\sum A_{\alpha,\beta,...,\lambda}\frac{k!}{\alpha!\,\beta!\,...\,\lambda!}X_1^\alpha X_2^\beta ... X_n^\lambda. \tag{9}$$

Then, by § 3.1, $\quad I(A) = |M|^{kq/n} I(a) \quad$ (10)

whatever the matrix M may be.

Consider the particular transformation

$$x_1 = X_1, \quad ..., \quad x_{n-1} = X_{n-1}, \quad x_n = \lambda X_n,$$

for which $|M| = \lambda$. Then, by a repetition of the argument of § 3.3, we can prove that $I(a)$ must be isobaric and of weight kq/n. Moreover, since the weight is necessarily an integer, kq/n must be an integer.

EXAMPLES XIV B

Examples 2–4 are the extensions to covariants of results already proved for invariants. The proofs of the latter need only slight modification.

1. If $C(a,x)$ is a covariant of a given form u, and $C(a,x)$ is a sum of algebraic forms of different degrees, say

$$C(a,x) = C_1(a,x)^{\varpi_1}+...+C_r(a,x)^{\varpi_r},$$

where the index ϖ denotes degree in the variables, then each separate term $C(a,x)^{\varpi}$ is itself a covariant.

HINT. Since $C(a,x)$ is a covariant,

$$C(A,X) = |M|^s C(a,x).$$

Put $x_1 t,..., x_n t$ for $x_1,..., x_n$ and, consequently, $X_1 t,..., X_n t$ for $X_1,..., X_n$. The result is

$$\sum_r C_r(A,X)^{\varpi_r} t^{\varpi_r} = |M|^s \sum_r C_r(a,x)^{\varpi_r} t^{\varpi_r}.$$

Each side is a polynomial in t and we may equate coefficients of like powers of t.

2. A covariant of degree ϖ in the variables is homogeneous in the coefficients.

HINT. The particular transformation

$$x_1 = \lambda X_1, \quad ..., \quad x_n = \lambda X_n$$

gives, as in § 3.1, $$C(a\lambda^k, X)^{\varpi} = \lambda^{ns} C(a,x)^{\varpi},$$

i.e. $$\lambda^{-\varpi} C(a\lambda^k, x)^{\varpi} = \lambda^{ns} C(a,x)^{\varpi}$$

This proves the required result and, further, shows that, if C is of degree q in the coefficients a,

$$\lambda^{kq} = \lambda^{ns}\lambda^{\varpi}$$

or $$kq = ns+\varpi.$$

3. If in a covariant of a binary form in x, y we consider x to have weight unity and y to have weight zero (x^2 of weight 2, etc.), then a covariant $C(a,x)^{\varpi}$, of degree q in the coefficients a, is isobaric and of weight $\frac{1}{2}(kq+\varpi)$.

HINT. Consider the particular transformation, $x = X$, $y = \lambda Y$ and follow the line of argument of § 3.3.

4. If in a covariant of a form in $x_1,..., x_n$ we consider $x_1,..., x_{n-1}$ to have unit weight and x_n to have zero weight ($x_1 x_2 x_n$ of weight 2, etc.), then a covariant $C(a,x)^{\varpi}$, of degree q in the coefficients a, is isobaric and of weight $\{kq+(n-1)\varpi\}/n$.

HINT. Compare § 3.5.

3.6. Invariants of a covariant. Let $u(a,x)_n^k$ be a given form of degree k in the n variables $x_1,..., x_n$. Let $C(a,x)^{\varpi}$ be a covariant of u, of degree ϖ in the variables and of degree q in the coefficients a. The coefficients in $C(a,x)^{\varpi}$ are not the actual a's of u but homogeneous functions of them.

The transformation $\mathbf{x} = M\mathbf{X}$ changes $u(a,x)$ into $u(A,X)$ and, since C is a covariant, there is an s such that

$$C(A,X)^{\varpi} = |M|^{s}C(a,x)^{\varpi}.$$

Thus the transformation changes $C(a,x)^{\varpi}$ into $|M|^{-s}C(A,X)^{\varpi}$. But C is a homogeneous function of degree q in the a's, and so also in the A's. Hence

$$C(a,x)^{\varpi} = |M|^{-s}C(A,X)^{\varpi} = C(A|M|^{-s/q},X)^{\varpi}. \qquad (11)$$

Now suppose that $I(b)$ is an invariant of $v(b,x)$, a form of degree ϖ in n variables. That is, if the coefficients are indicated by B when v is expressed in terms of X, there is a t for which

$$I(B) = |M|^{t}I(b). \qquad (12)$$

Take $v(b,x)$ to be the covariant $C(a,x)^{\varpi}$; then (11) shows that the corresponding B are the coefficients of the terms in $C(A|M|^{-s/q},X)$.

If $I(b)$ is of degree r in the b, then when expressed in terms of a it is $I_1(a)$, a function homogeneous and of degree rq in the coefficients a. Moreover, (12) gives

$$I_1(A|M|^{-s/q}) =.|M|^{t}I_1(a),$$

or, on using the fact that I_1 is homogeneous and of degree rq in the a,

$$|M|^{-rs}I_1(A) \doteq |M|^{t}I_1(a).$$

That is, $I_1(a)$ is an invariant of the original form $u(a,x)$.

The generalities of the foregoing work are not easy to follow. The reader should study them in relation to the particular example which we now give; it is taken from § 2.5.

$$u(a,x)^3 \equiv a_0x^3+3a_1x^2y+3a_2xy^2+a_3y^3$$

has a covariant

$$C(a,x)^2 \equiv (a_0a_2-a_1^2)x^2+(a_0a_3-a_1a_2)xy+(a_1a_3-a_2^2)y^2.$$

Being the Hessian of u, this covariant has a multiplying factor $|M|^2$; that is to say,

$$(A_0A_2-A_1^2)X^2+... = |M|^2\{(a_0a_2-a_1^2)x^2+...\},$$

or

$$(a_0a_2-a_1^2)x^2+... = \left\{\frac{A_0}{|M|}\frac{A_2}{|M|}-\left(\frac{A_1}{|M|}\right)^2\right\}X^2+..., \qquad (11\,\text{a})$$

which corresponds to (11) above.

But if $\mathbf{x} = M\mathrm{X}$ changes the quadratic form

$$b_0 x^2 + 2b_1 xy + b_2 y^2$$

into $$B_0 X^2 + 2B_1 XY + B_2 Y^2,$$

then $$B_0 B_2 - B_1^2 = |M|^2(b_0 b_2 - b_1^2). \tag{12a}$$

Take for the b's the coefficients of (11 a) and we see that

$$|M|^{-4}\{(A_0 A_2 - A_1^2)(A_1 A_3 - A_2^2) - \tfrac{1}{4}(A_0 A_3 - A_1 A_2)^2\}$$
$$= |M|^2\{(a_0 a_2 - a_1^2)(a_1 a_3 - a_2^2) - \tfrac{1}{4}(a_0 a_3 - a_1 a_2)^2\},$$

which proves that (9) of § 2.5 is an invariant of the cubic $a_0 x^3 + 3a_1 x^2 y + 3a_2 xy^2 + a_3 y^3$.

3.7. Covariants of a covariant. A slight extension of the argument of § 3.6, using a covariant $C(b, x)^{\varpi'}$ of a form of degree ϖ in n variables where § 3.6 uses $I(b)$, will prove that $C(b, x)^{\varpi'}$ gives rise to a covariant of the original form $u(a, x)$.

3.8. Irreducible invariants and covariants. If $I(a)$ is an invariant of a form $u(a, x)_n^k$, it is immediately obvious, from the definition of an invariant, that the square, cube,... of $I(a)$ are also invariants of u. Thus there is an unlimited number of invariants; and so for covariants.

On the other hand, there is, for a given form u, only a finite number of invariants which cannot be expressed rationally and integrally in terms of invariants of equal or lower degree. Equally, there is only a finite number of covariants which cannot be expressed rationally and integrally in terms of invariants and covariants of equal or lower degree in the coefficients of u. Invariants or covariants which cannot be so expressed are called irreducible. The theorem may then be stated

'*The number of irreducible covariants and invariants of a given form is finite.*'

This theorem and its extension to the joint covariants and invariants of any given system of forms is sometimes called the Gordan-Hilbert theorem. Its proof is beyond the scope of the present book.

Even among the irreducible covariants and invariants there may be an algebraical relation of such a kind that no one of its terms can be expressed as a rational and integral function of the rest. For example, a cubic u has three irreducible covariants, itself, a covariant G, and a covariant H; it has one irreducible invariant Δ: there is a relation connecting these four, namely,

$$\Delta u^2 = G^2+4H^3.$$

We shall prove this relation in § 4.3.

4. Canonical forms

4.1. Binary cubics. Readers will be familiar with the device of considering $ax^2+2hxy+by^2$ in the form $aX^2+b_1Y^2$, obtained by writing

$$ax^2+2hxy+by^2 = a\left(x+\frac{h}{a}y\right)^2+\left(b-\frac{h^2}{a}\right)y^2$$

and making the substitution $X = x+(h/a)y$, $Y = y$. The general quadratic form was considered in Theorem 47, p. 148.

We now show that the cubic

$$a_0x^3+3a_1x^2y+3a_2xy^2+a_3y^3 \tag{1}$$

may, in general, be written as pX^3+qY^3; the exceptions are cubics that contain a squared factor.

There are several ways of proving this: the method that follows is, perhaps, the most direct. We pose the problem

'Is it possible to find p, q, α, β so that (1) is identically equal to

$$p(x+\alpha y)^3+q(x+\beta y)^3\,?\text{'} \tag{2}$$

It is possible if, having chosen α and β, we can then choose p and q to satisfy the four equations

$$\left.\begin{aligned} p+q &= a_0, & p\alpha+q\beta &= a_1, \\ p\alpha^2+q\beta^2 &= a_2, & p\alpha^3+q\beta^3 &= a_3. \end{aligned}\right\} \tag{3}$$

For general values of a_0, a_1, a_2, a_3 these four equations cannot be consistent if $\alpha = \beta$. We shall proceed, at first, on the assumption $\alpha \neq \beta$.

The third equation of (3) follows from the first two if we can choose P, Q so that, simultaneously,

$$a_2+Pa_1+Qa_0 = 0, \quad \alpha^2+P\alpha+Q = 0, \quad \beta^2+P\beta+Q = 0.$$

The fourth equation of (3) follows from the second and third if P, Q also satisfy

$$a_3+Pa_2+Qa_1 = 0, \quad \alpha(\alpha^2+P\alpha+Q) = 0, \quad \beta(\beta^2+P\beta+Q) = 0.$$

That is, the third and fourth equations are linear consequences of the first two if α, β are the roots of

$$t^2+Pt+Q = 0,$$

where P, Q are determined from the two equations

$$a_2+Pa_1+Qa_0 = 0,$$
$$a_3+Pa_2+Qa_1 = 0.$$

This means that α, β are the roots of

$$(a_0 a_2-a_1^2)t^2+(a_1 a_2-a_0 a_3)t+(a_1 a_3-a_2^2) = 0$$

and, provided they are distinct, p, q can then be determined from the first two equations of (3).

Thus (1) can be expressed in the form (2) provided that

$$(a_1 a_2-a_0 a_3)^2-4(a_0 a_2-a_1^2)(a_1 a_3-a_2^2) \tag{4}$$

is not zero, this condition being necessary to ensure that α is not equal to β.

We may readily see what cubics are excluded by the provision that (4) is not zero. If (4) is zero, the two quadratics

$$a_0 x^2+2a_1 xy+a_2 y^2,$$
$$a_1 x^2+2a_2 xy+a_3 y^2$$

have a common linear factor;[†] hence

$$a_0 x^2+2a_1 xy+a_2 y^2, \qquad a_0 x^3+3a_1 x^2y+3a_2 xy^2+a_3 y^3$$

have a common linear factor, and so the latter has a repeated linear factor. Such a cubic may be written as X^2Y, where X and Y are linear in x and y.

[†] The reader will more readily recognize the argument in the form

$$a_0x^2+2a_1x+a_2 = 0, \qquad a_1x^2+2a_2x+a_3 = 0$$

have a common root; hence

$$a_0x^2+2a_1x+a_2 = 0, \qquad a_0x^3+3a_1x^2+3a_2x+a_3 = 0$$

have a common root, and so the latter has a repeated root. Every root common to $F(x) = 0$ and $F'(x) = 0$ is a repeated root of the former.

4.2. Binary quartics. The quartic

$$(a_0,\dots,a_4)(x,y)^4 \tag{5}$$

has four linear factors. By combining these in pairs we may write (5) as the product of two quadratic factors, say

$$a_0'x^2+2h'xy+b'y^2, \qquad a''x^2+2h''xy+b''y^2. \tag{6}$$

Let us, at first, preclude quartics that have a squared factor, linear in x and y. Then, by a linear transformation, not necessarily real, we may write the factors (6) in the forms†

$$X^2+Y^2, \qquad \alpha X^2+2\beta XY+\gamma Y^2.$$

By applying Theorem 48, without insisting on the transformation being real, we can write these in the forms

$$X_1^2+X_2^2, \qquad \lambda_1 X_1^2+\lambda_2 X_2^2. \tag{7}$$

Moreover, since we are precluding quartics that have a squared linear factor, neither λ_1 nor λ_2 is zero, and the quartic may be written as

$$\lambda_1 X_1^4+(\lambda_1+\lambda_2)X_1^2X_2^2+\lambda_2 X_2^4,$$

or, on making a final substitution

$$X = \lambda_1^{\frac{1}{4}}X_1, \qquad Y = \lambda_2^{\frac{1}{4}}X_2,$$

we may write the quartic as

$$X^4+6mX^2Y^2+Y^4. \tag{8}$$

This is the canonical form for the general quartic. A quartic having a squared factor, hitherto excluded from our discussion, may be written as $X^2(\lambda_1 X^2+\lambda_2 Y^2)$.

4.3. Application of canonical forms. Relations among the invariants and covariants of a form are fairly easy to detect when the canonical form is considered. For example, the cubic

$$u = a_0x^3+3a_1x^2y+3a_2xy^2+a_3y^3$$

is known (§ 2.5) to have

a covariant $H = (a_0a_2-a_1^2)x^2+\dots$ [(8) of § 2.5],

a covariant $G = (a_0^2a_3-3a_0a_1a_2+2a_1^3)x^3+\dots$ [(10) of § 2.5],

and an invariant

$$\Delta = (a_0a_3-a_1a_2)^2-4(a_0a_2-a_1^2)(a_1a_3-a_2^2).$$

---

† We have merely to identify the two distinct factors of $a'x^2+2h'xy+b'y^2$ with $X+iY$ and $X-iY$.

Unless $\Delta = 0$ the cubic can, by the transformation $X = x+\alpha y$, $Y = x+\beta y$ of § 4.1, be expressed as

$$u = pX^3+qY^3.$$

In these variables the covariants H and G become

$$H_1 = pqXY, \qquad G_1 = p^2qX^3-pq^2Y^3,$$

while $$\Delta_1 = p^2q^2.$$

It is at once obvious that

$$\Delta_1 u^2 = G_1^2+4H_1^3.$$

Now, $|M|$ being the modulus of the transformation $\mathbf{x} = M\mathbf{X}$,

$$H_1 = |M|^2H, \qquad G_1 = |M|^3G, \qquad \Delta_1 = |M|^6\Delta.$$

The factor of H is $|M|^2$ because H is a Hessian. The factor $|M|^3$ of G is most easily determined by considering the transformation $x = \lambda X$, $y = \lambda Y$, which gives $A_0 = a_0\lambda^3,\ldots, A_3 = a_3\lambda^3$, so that

$$\begin{aligned}(A_0^2A_3-3A_0A_1A_2+2A_1^3)X^3+\ldots\\ &= \lambda^9\{(a_0^2a_3-3a_0a_1a_2+2a_1^3)x^3\lambda^{-3}+\ldots\}\\ &= \lambda^6\{(a_0^2a_3-3a_0a_1a_2+2a_1^3)x^3+\ldots\},\end{aligned}$$

and the multiplying factor is λ^6, or $|M|^3$.

The factor $|M|^6$ of Δ may be determined in the same way.

Accordingly, we have proved that, unless $\Delta = 0$,

$$\Delta u^2 = G^2+4H^3. \tag{9}$$

If $\Delta = 0$, the cubic can be written as X^2Y, a form for which both G and H are zero.

5. Geometrical invariants

5.1. Projective invariants. Let a point P, having homogeneous coordinates x, y, z (say areal, or trilinear, to be definite) referred to a triangle ABC in a plane π, be projected into the point P' in the plane π'. Let $A'B'C'$ be the projection on π' of ABC and let x', y', z' be the homogeneous coordinates of P' referred to $A'B'C'$. Then, as many books on analytical geometry prove (in one form or another), there are constants l, m, n such that

$$x = lx', \qquad y = my', \qquad z = nz'.$$

Let X, Y, Z be the homogeneous coordinates of P' referred to a triangle, in π', other than $A'B'C'$. Then there are relations

$$x' = \lambda_1 X+\mu_1 Y+\nu_1 Z,$$
$$y' = \lambda_2 X+\mu_2 Y+\nu_2 Z,$$
$$z' = \lambda_3 X+\mu_3 Y+\nu_3 Z.$$

Thus the projection of a figure in π on to a plane π' gives rise to a transformation of the type

$$\left.\begin{aligned} x &= l_1 X+m_1 Y+n_1 Z, \\ y &= l_2 X+m_2 Y+n_2 Z, \\ z &= l_3 X+m_3 Y+n_3 Z, \end{aligned}\right\} \qquad (1)$$

wherein, since $x = 0$, $y = 0$, $z = 0$ are not concurrent lines, the determinant $(l_1\, m_2\, n_3)$ is not zero.

Thus projection leads to the type of transformation we have been considering in the earlier sections of the chapter. Geometrical properties of figures that are unaltered by projection (projective properties) may be expected to correspond to invariants or covariants of algebraic forms and, conversely, any invariant or covariant of algebraic forms may be expected to correspond to some projective property of a geometrical figure.

The binary transformation

$$x = l_1 X+m_1 Y, \qquad y = l_2 X+m_2 Y, \qquad (2)$$

may be considered as the form taken by (1) when only lines through the vertex C of the original triangle of reference are in question; for such an equation as $ax^2+2hxy+by^2 = 0$ corresponds to a pair of lines through C; it becomes

$$AX^2+2HXY+BY^2 = 0,$$

say, after transformation by (2), which is a pair of lines through a vertex of the triangle of reference in the projected figure.

We shall not attempt any systematic development of the geometrical approach to invariant theory: we give merely a few isolated examples.

The cross-ratio of a pencil of four lines is unaltered by projection: the condition that the two pairs of lines

$$ax^2+2hxy+by^2, \qquad a'x^2+2h'xy+b'y^2$$

should form a harmonic pencil is $ab'+a'b-2hh' = 0$: this represents a property that is unaltered by projection and, as we should expect, $ab'+a'b-2hh'$ is an invariant of the two algebraic forms. Again, we may expect the cross-ratio of the four lines

$$ax^4+4bx^3+6cx^2y^2+4dxy^3+ey^4 = 0$$

to indicate some of the invariants of the quartic. The condition that the four lines form a harmonic pencil is

$$J \equiv ace+2bcd-ad^2-b^2e-c^3 = 0$$

and, as we have seen (Example 15, p. 181), J is an invariant of the quartic. The condition that the four lines form an equiharmonic pencil, i.e. that the first and fourth of the cross-ratios

$$\rho, \quad \frac{1}{\rho}, \quad 1-\rho, \quad \frac{1}{1-\rho}, \quad 1-\frac{1}{\rho}, \quad \frac{\rho}{\rho-1}$$

(the six values arising from the permutations of the order of the lines) are equal, is

$$I \equiv ae-4bd+3c^2 = 0.$$

Moreover, I is an invariant of the quartic.

Again, in geometry, the Hessian of a given curve is the locus of a point whose polar conic with respect to the curve is a pair of lines: the locus meets the curve only at the inflexions and multiple points of the curve. Now proper conics project into proper conics, line-pairs into line-pairs, points of inflexion into points of inflexion, and multiple points into multiple points. Thus all the geometrical marks of a Hessian are unaltered by projection and, as we should expect in such circumstances, the Hessian proves to be a covariant of an algebraic form. Similarly, the Jacobian of three curves is the locus of a point whose polar lines with respect to the three curves are concurrent; thus its geometrical definition turns upon projective properties of the curves and, as we should expect, the Jacobian of any three algebraic forms is a covariant.

In view of their close connexion with the geometry of projection, the invariants and covariants we have hitherto been considering are sometimes called projective invariants and covariants.

5.2. Metrical invariants. The orthogonal transformation

$$\left.\begin{aligned} x &= l_1 X+m_1 Y+n_1 Z, \\ y &= l_2 X+m_2 Y+n_2 Z, \\ z &= l_3 X+m_3 Y+n_3 Z, \end{aligned}\right\} \tag{3}$$

wherein (l_1, m_1, n_1), (l_2, m_2, n_2), and (l_3, m_3, n_3) are the direction-cosines of three mutually perpendicular lines, is a particular type of linear transformation. It leaves unchanged certain functions of the coefficients of

$$ax^2+by^2+cz^2+2fyz+2gzx+2hxy \tag{4}$$

which are not invariant under the general linear transformation. If (3) transforms (4) into

$$a_1 X^2+b_1 Y^2+c_1 Z^2+2f_1 YZ+2g_1 ZX+2h_1 XY, \tag{5}$$

we have, since (3) also transforms $x^2+y^2+z^2$ into $X^2+Y^2+Z^2$, the values of λ for which (4) $-$ $\lambda(x^2+y^2+z^2)$ is a pair of linear factors are the values of λ for which (5) $-$ $\lambda(X^2+Y^2+Z^2)$ is a pair of linear factors. These values of λ are given by

$$\begin{vmatrix} a-\lambda & h & g \\ h & b-\lambda & f \\ g & f & c-\lambda \end{vmatrix} = 0, \qquad \begin{vmatrix} a_1-\lambda & h_1 & g_1 \\ h_1 & b_1-\lambda & f_1 \\ g_1 & f_1 & c_1-\lambda \end{vmatrix} = 0,$$

respectively. By equating the coefficients of powers of λ in the two equations we obtain

$$\begin{aligned} a+b+c &= a_1+b_1+c_1, \\ A+B+C &= A_1+B_1+C_1 \quad (A \equiv bc-f^2, \text{ etc.}), \\ \Delta &= \Delta_1, \end{aligned}$$

where Δ is the discriminant of the form (4). That is to say, $a+b+c$, $bc+ca+ab-f^2-g^2-h^2$ are invariant under orthogonal transformation; they are not invariant under the general linear transformation. On the other hand, the discriminant Δ is an invariant of the form for all linear transformations.

The orthogonal transformation (3) is equivalent to a change from one set of rectangular axes of reference to another set of rectangular axes. It leaves unaltered all properties of a geometrical figure that depend solely on measurement. For example,

$$\frac{\lambda\lambda_1+\mu\mu_1+\nu\nu_1}{\sqrt{\{(\lambda_1^2+\mu_1^2+\nu_1^2)(\lambda^2+\mu^2+\nu^2)\}}} \tag{6}$$

is an invariant of the linear forms

$$\lambda x+\mu y+\nu z, \qquad \lambda_1 x+\mu_1 y+\nu_1 z$$

under any orthogonal transformation: for (6) is the cosine of the angle between the two planes. Such examples have but little algebraic interest.

6. Arithmetic and other invariants

6.1. There are certain integers associated with algebraic forms (or curves) that are obviously unchanged when the variables undergo the transformation $\mathbf{x} = M\mathbf{X}$. Such are the degree of the curve, the number of intersections of a curve with its Hessian, and so on. One of the less obvious† of these arithmetic invariants, as they are called, is the rank of the matrix formed by the coordinates of a number of points.

Let the components of $\mathbf{x}$ be $x, y,..., z$, n in number. Let m points (or particular values of $\mathbf{x}$) be given; say

$$(x_1,..., z_1), \quad ..., \quad (x_m,..., z_m).$$

Then the rank of the matrix

$$\begin{bmatrix} x_1 & . & . & z_1 \\ . & . & . & . \\ x_m & . & . & z_m \end{bmatrix}$$

is unaltered by any non-singular linear transformation

$$\mathbf{x} = M\mathbf{X}.$$

Suppose the matrix is of rank r ($< m$). Then (Theorem 31) we can choose r rows and express the others as sums of multiples of these r rows. For convenience of writing, suppose all rows after the rth can be expressed as sums of multiples of the first r rows. Then there are constants $\lambda_{1k},..., \lambda_{rk}$ such that, for any letter t of $x,..., z$,

$$t_{r+k} = \lambda_{1k} t_1+...+\lambda_{rk} t_r. \tag{1}$$

---

† This also is obvious to anyone with some knowledge of n-dimensional geometry.

The transformation $\mathbf{x} = M\mathbf{X}$ has an inverse $\mathbf{X} = M^{-1}\mathbf{x}$ or, in full, say

$$X = c_{11}x+...+c_{1n}z,$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$Z = c_{n1}x+...+c_{nn}z.$$

Putting in suffixes for particular points

$$(X_1,..., Z_1), \quad ..., \quad (X_m,..., Z_m),$$

and supposing that t is the jth letter of x, y,..., z, we have

$$T_{r+k}-\sum_{q=1}^{r}\lambda_{qk}T_k$$
$$= (c_{j1}x_{r+k}+...+c_{jn}z_{r+k})-\sum_{q=1}^{r}\lambda_{qk}(c_{j1}x_q+...+c_{jn}z_q).$$

By (1), the coefficient of each of c_{j1},..., c_{jn} is zero and hence (1) implies

$$T_{r+k} = \lambda_{1k}T_1+...+\lambda_{rk}T_r. \tag{2}$$

Conversely, as we see by using $\mathbf{x} = M\mathbf{X}$ where in the foregoing we have used $\mathbf{X} = M^{-1}\mathbf{x}$, (2) implies (1). By Theorems 31 and 32, it follows that the ranks of the two matrices

$$\begin{bmatrix} x_1 & . & . & z_1 \\ . & . & . & . \\ x_m & . & . & z_m \end{bmatrix}, \quad \begin{bmatrix} X_1 & . & . & Z_1 \\ .. & . & . & . \\ X_m & . & . & Z_m \end{bmatrix}$$

are equal.

6.2. Transvectants. We conclude with an application to binary forms of invariants which are derived from the consideration of particular values of the variables.

Let (x_1, y_1), (x_2, y_2) be two cogredient pairs of variables (x, y) each subject to the transformation

$$x = lX+mY, \qquad y = l'X+m'Y, \tag{3}$$

wherein $|M| = lm'-l'm$. Then

$$\frac{\partial}{\partial X} = l\frac{\partial}{\partial x}+l'\frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial Y} = m\frac{\partial}{\partial x}+m'\frac{\partial}{\partial y},$$

the transformation being contragredient to (3) (compare Chap. X, § 6.2).

If now the operators $\frac{\partial}{\partial X_1}, \frac{\partial}{\partial Y_1}, \frac{\partial}{\partial X_2}, \frac{\partial}{\partial Y_2}$ operate on a function of X_1, Y_1, X_2, Y_2 and these are taken as independent variables, we have

$$\frac{\partial}{\partial X_1}\frac{\partial}{\partial Y_2}-\frac{\partial}{\partial Y_1}\frac{\partial}{\partial X_2}$$

$$=\left(l\frac{\partial}{\partial x_1}+l'\frac{\partial}{\partial y_1}\right)\left(m\frac{\partial}{\partial x_2}+m'\frac{\partial}{\partial y_2}\right)-$$

$$-\left(m\frac{\partial}{\partial x_1}+m'\frac{\partial}{\partial y_1}\right)\left(l\frac{\partial}{\partial x_2}+l'\frac{\partial}{\partial y_2}\right)$$

$$=(lm'-l'm)\left(\frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2}-\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\right). \tag{4}$$

Hence (4) is an invariant operator.

Let u_1, v_1 be binary forms u, v when we put $x = x_1$, $y = y_1$; u_2, v_2 the same forms u, v when we put $x = x_2$, $y = y_2$; U_1, V_1 and U_2, V_2 the corresponding forms when u, v are expressed in terms of X, Y and the particular values X_1, Y_1 and X_2, Y_2 are introduced. Then, operating on the product $U_1 V_2$, we have

$$\left(\frac{\partial}{\partial X_1}\frac{\partial}{\partial Y_2}-\frac{\partial}{\partial Y_1}\frac{\partial}{\partial X_2}\right)(U_1 V_2)=|M|\left(\frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2}-\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\right)(u_1 v_2),$$

$$\left(\frac{\partial}{\partial X_1}\frac{\partial}{\partial Y_2}-\frac{\partial}{\partial Y_1}\frac{\partial}{\partial X_2}\right)^2(U_1 V_2)=|M|^2\left(\frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2}-\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\right)^2(u_1 v_2),$$

and so on. That is, for any integer r,

$$\frac{\partial^r U_1}{\partial X_1^r}\frac{\partial^r V_2}{\partial Y_2^r}-r\frac{\partial^r U_1}{\partial X_1^{r-1}\partial Y_1}\frac{\partial^r V_2}{\partial X_2\partial Y_2^{r-1}}+\dots$$

$$=|M|^r\left(\frac{\partial^r u_1}{\partial x_1^r}\frac{\partial^r v_2}{\partial y_2^r}-r\frac{\partial^r u_1}{\partial x_1^{r-1}\partial y_1}\frac{\partial^r v_2}{\partial x_2\partial y_2^{r-1}}+\dots\right).$$

These results are true for all pairs (x_1, y_1) and (x_2, y_2). We may, then, replace both pairs by (x, y) and so obtain the theorem that

$$\frac{\partial^r u}{\partial x^r}\frac{\partial^r v}{\partial y^r}-r\frac{\partial^r u}{\partial x^{r-1}\partial y}\frac{\partial^r v}{\partial x\partial y^{r-1}}+\dots \tag{5}$$

is a covariant (possibly an invariant) of the two forms u and v. It is called the rth transvectant of u and v.

Finally, when we take $v \equiv u$ in (5), we obtain covariants of the single form u.

This method can be made the basis of a systematic attack on the invariants and covariants of a given set of forms.

7. Further reading

The present chapter has done no more than sketch some of the elementary results from which the theory of invariants and covariants starts. The following books deal more fully with the subject:

E. B. ELLIOTT, *An Introduction to the Algebra of Quantics* (Oxford, 1895).
GRACE and YOUNG, *The Algebra of Invariants* (Cambridge, 1902).
WEITZENBÖCK, *Invariantentheorie* (Gröningen, 1923).
H. W. TURNBULL, *The Theory of Determinants, Matrices, and Invariants* (Glasgow, 1928).

EXAMPLES XIV C

1. Write down the most general function of the coefficients in $a_0x^3+3a_1x^2y+3a_2xy^2+a_3y^3$ which is (i) homogeneous and of degree 2 in the coefficients, isobaric, and of weight 3; (ii) homogeneous and of degree 3 in the coefficients, isobaric, and of weight 4.

HINT. (i) Three is the sum of 3 and 0 or of 2 and 1; the only terms possible are numerical multiples of a_0a_3 and a_1a_2.

Ans. $\alpha a_0a_3+\beta a_1a_2$, where α, β are numerical constants.

2. What is the weight of the invariant

$$I \equiv ae-4bd+3c^2 \text{ of the quartic } (a,b,c,d,e)(x,y)^4\,?$$

HINT. Rewrite the quartic as $(a_0,\ldots,a_4)(x,y)^4$.

3. What is the weight of the discriminant $|a_{rs}|$ of the quadratic form

$$a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2+2a_{23}x_2x_3+2a_{31}x_3x_1\,?$$

HINT. Rewrite the form as

$$\sum \frac{2!}{\alpha!\beta!\gamma!}a_{\alpha,\beta,\gamma}x_1^\alpha x_2^\beta x_3^\gamma,$$

summed for $\alpha+\beta+\gamma = 2$. Thus $a_{11} = a_{2,0,0}, a_{23} = a_{0,1,1}$.

4. Write down the Jacobian of $ax^2+2hxy+by^2$, $a'x^2+2h'xy+b'y^2$ and deduce from it (§ 3.6) that

$$(ab'-a'b)^2+4(ah'-a'h)(bh'-b'h)$$

is a joint-invariant of the two forms.

5. Verify the theorem 'When I is an invariant of the form $(a_0,\ldots,a_n)(x,y)^n$,

$$\left(a_0\frac{\partial}{\partial a_1}+2a_1\frac{\partial}{\partial a_2}+\ldots+na_{n-1}\frac{\partial}{\partial a_n}\right)I = 0,$$

$$\left(na_1\frac{\partial}{\partial a_0}+(n-1)a_2\frac{\partial}{\partial a_1}+\ldots+a_n\frac{\partial}{\partial a_{n-1}}\right)I = 0,\text{'}$$

in so far as it relates to (i) the discriminant of a quadratic form, (ii) the invariant Δ (§ 4.3) of a cubic, (iii) the invariants I, J (§ 5.1) of a quartic.

6. Show that the general cubic $ax^3+3bx^2y+3cxy^2+dy^3$ can be reduced to the canonical form AX^3+DY^3 by the substitution $X = px+qy$, $Y = p'x+q'y$, where the Hessian of the cubic is

$$(ac-b^2)x^2+(ad-bc)xy+(bd-c^2)y^2 \equiv (px+qy)(p'x+q'y).$$

HINT. The Hessian of $AX^3+3BX^2Y+3CXY^2+DY^3$ reduces to XY: i.e. $AC = B^2$, $BD = C^2$. Prove that, if $A \neq 0$, either $B = C = 0$ or the form is a cube; if $A = 0$, then $B = C = 0$ and the Hessian is identically zero.

7. Find the equation of the double lines of the involution determined by the two line-pairs

$$ax^2+2hxy+by^2 = 0, \qquad a'x^2+2h'xy+b'y^2 = 0$$

and prove that it corresponds to a covariant of the two forms.

HINT. If the line-pair is $a_1x^2+2h_1xy+b_1y^2 = 0$, then, by the usual condition for harmonic conjugates,

$$ab_1+a_1b-2hh_1 = 0, \qquad a'b_1+a_1b'-2h'h_1 = 0.$$

8. If two conics S, S' are such that a triangle can be inscribed in S' and circumscribed to S, then the invariants (Example 11, p. 180) $\Delta, \Theta, \Theta', \Delta'$ satisfy the relation $\Theta^2 = 4\Delta\Theta'$ independently of the choice of triangle of reference.

HINT. Consider the equations of the conics in the forms

$$\textstyle\sum \equiv 2Fmn+2Gml+2Hlm = 0,$$
$$S' \equiv 2f'yz+2g'zx+2h'xy = 0.$$

The general result follows by linear transformation.

9. Prove that the rank of the matrix of the coefficients a_{rs} of m linear forms

$$a_{r1}x_1+...+a_{rn}x_n \quad (r = 1,...,m)$$

in n variables is an arithmetic invariant.

10. Prove that if (x_r, y_r, z_r) are transformed into (X_r, Y_r, Z_r) by $\mathbf{x} = M\mathbf{X}$, then the determinant $|x_1\, y_2\, z_3|$ is an invariant.

11. Prove that the nth transvectant of two forms

$$(a_0,..., a_n)(x, y)^n, \qquad (b_0,..., b_n)(x, y)^n$$

is linear in the coefficients of each form and is a joint-invariant of the two forms. It is called the lineo-linear invariant.

Ans. $a_0b_n-na_1b_{n-1}+\frac{1}{2}n(n-1)a_2b_{n-2}+...$.

12. Prove, by Theorems 34 and 37, that the rank of a quadratic form is unaltered by any non-singular linear transformation of the variables.

CHAPTER XV

LATENT VECTORS

1. Introduction

1.1. Vectors in three dimensions. In three-dimensional geometry, with rectangular axes $O\xi_1$, $O\xi_2$, and $O\xi_3$, a vector $\boldsymbol{\xi} = OP$ has components

$$(\xi_1, \xi_2, \xi_3),$$

these being the coordinates of P with respect to the axes. The length of the vector $\boldsymbol{\xi}$ is

$$|\boldsymbol{\xi}| = \sqrt{(\xi_1^2+\xi_2^2+\xi_3^2)} \tag{1}$$

and the direction-cosines of the vector are

$$\frac{\xi_1}{|\boldsymbol{\xi}|}, \quad \frac{\xi_2}{|\boldsymbol{\xi}|}, \quad \frac{\xi_3}{|\boldsymbol{\xi}|}.$$

Two vectors, $\boldsymbol{\xi}$ with components (ξ_1, ξ_2, ξ_3) and $\boldsymbol{\eta}$ with components (η_1, η_2, η_3), are orthogonal if

$$\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3 = 0. \tag{2}$$

Finally, if $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ have components

$$(1,0,0), \quad (0,1,0), \quad \text{and} \quad (0,0,1),$$

a vector $\boldsymbol{\xi}$ with components (ξ_1, ξ_2, ξ_3) may be written as

$$\boldsymbol{\xi} = \xi_1\mathbf{e}_1+\xi_2\mathbf{e}_2+\xi_3\mathbf{e}_3.$$

Also, each of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is of unit length, by (1), and the three vectors are, by (2), mutually orthogonal. They are, of course, unit vectors along the axes.

1.2. Single-column matrices. A purely algebraical presentation of these details is available to us if we use a single-column matrix to represent a vector.

A vector OP is represented by a single-column matrix $\boldsymbol{\xi}$ whose elements are ξ_1, ξ_2, ξ_3. The length of the vector is defined to be

$$\sqrt{(\xi_1^2+\xi_2^2+\xi_3^2)}. \tag{1}$$

The condition for two vectors to be orthogonal now takes a purely matrix form. The transpose $\boldsymbol{\xi}'$ of $\boldsymbol{\xi}$ is a single-row matrix with elements ξ_1, ξ_2, ξ_3 and

$$\boldsymbol{\xi}'\boldsymbol{\eta} = [\xi_1 \quad \xi_2 \quad \xi_3] \times \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix}$$

is a single-element matrix

$$[\xi_1 \eta_1 + \xi_2 \eta_2 + \xi_3 \eta_3].$$

The condition for $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ to be orthogonal is, in matrix notation,

$$\boldsymbol{\xi}'\boldsymbol{\eta} = 0. \tag{2}$$

The matrices $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are defined by

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};$$

they are mutually orthogonal and

$$\boldsymbol{\xi} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \xi_3 \mathbf{e}_3.$$

The change from three dimensions to n dimensions is immediate. With n variables, a vector $\boldsymbol{\xi}$ is a single-column matrix with elements

$$\xi_1, \xi_2, \ldots, \xi_n.$$

The length of the vector is DEFINED to be

$$|\boldsymbol{\xi}| = \sqrt{(\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2)}, \tag{3}$$

and when $|\boldsymbol{\xi}| = 1$ the vector is said to be a unit vector.

Two vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are said to be orthogonal when

$$\xi_1 \eta_1 + \xi_2 \eta_2 + \ldots + \xi_n \eta_n = 0, \tag{4}$$

which, in matrix notation, may be expressed in either of the two forms†

$$\boldsymbol{\xi}'\boldsymbol{\eta} = 0, \qquad \boldsymbol{\eta}'\boldsymbol{\xi} = 0. \tag{5}$$

In the same notation, when $\boldsymbol{\xi}$ is a unit vector

$$\boldsymbol{\xi}'\boldsymbol{\xi} = 1.$$

Here, and later, 1 denotes the unit single-element matrix.

Finally, we define the unit vector $\mathbf{e}_r$ as the single-column matrix which has unity in the rth row and zero elsewhere. With

---

† Note the order of multiplication; we can form the product AB only when the number of columns of A is equal to the number of rows of B (cf. p. 73) and hence we cannot form the products $\boldsymbol{\xi}\boldsymbol{\eta}'$ or $\boldsymbol{\eta}\boldsymbol{\xi}'$.

this notation the n vectors $\mathbf{e}_r$ $(r = 1,..., n)$ are mutually orthogonal and

$$\boldsymbol{\xi} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + ... + \xi_n \mathbf{e}_n.$$

1.3. Orthogonal matrices. In this subsection we use $\mathbf{x}_r$ to denote a vector, or single-column matrix, with elements

$$x_{1r}, x_{2r},..., x_{nr}.$$

THEOREM 59. *Let* $\mathbf{x}_1, \mathbf{x}_2,..., \mathbf{x}_n$ *be mutually orthogonal unit vectors. Then the square matrix* X *whose columns are the* n *vectors* $\mathbf{x}_1, \mathbf{x}_2,..., \mathbf{x}_n$ *is an orthogonal matrix.*

Proof. Let X' be the transpose of X. Then the product $X'X$ is

$$\begin{bmatrix} x_{11} & x_{21} & . & . & . & x_{n1} \\ . & . & . & . & . & . \\ x_{1r} & x_{2r} & . & . & . & x_{nr} \\ . & . & . & . & . & . \\ x_{1n} & x_{2n} & . & . & . & x_{nn} \end{bmatrix} \times \begin{bmatrix} x_{11} & . & . & . & x_{1s} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2s} & . & . & . & x_{2n} \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ x_{n1} & . & . & . & x_{ns} & . & . & . & x_{nn} \end{bmatrix}.$$

The element in the rth row and sth column of the product is, on using the summation convention,

$$x_{\alpha r} x_{\alpha s}.$$

When $s = r$ this is unity, since $\mathbf{x}_r$ is a unit vector; and when $s \neq r$ this is zero, since the vectors $\mathbf{x}_r$ and $\mathbf{x}_s$ are orthogonal. Hence

$$X'X = I,$$

the unit matrix of order n, and so X is an orthogonal matrix.

Aliter. (The same proof in different words.)

X is a matrix whose columns are

$$\mathbf{x}_1,..., \mathbf{x}_s,..., \mathbf{x}_n,$$

while X' is a matrix whose rows are

$$\mathbf{x}'_1,..., \mathbf{x}'_r,..., \mathbf{x}'_n.$$

The element in the rth row and sth column of the product $X'X$ is $\mathbf{x}'_r \mathbf{x}_s$ (strictly the numerical value in this single-element matrix).

Let $r \neq s$; then $\mathbf{x}'_r \mathbf{x}_s = 0$, since $\mathbf{x}_r$ and $\mathbf{x}_s$ are orthogonal.

Let $r = s$; then $\mathbf{x}'_r \mathbf{x}_s = \mathbf{x}'_s \mathbf{x}_s = 1$, since $\mathbf{x}_s$ is a unit vector

Hence $X'X = I$ and X is an orthogonal matrix.

COROLLARY. *In the transformation* $\boldsymbol{\eta} = X'\boldsymbol{\xi}$, or $\boldsymbol{\xi} = X\boldsymbol{\eta}$, $[X' = X^{-1}]$ *let* $\boldsymbol{\eta} = \mathbf{y}_r$ *when* $\boldsymbol{\xi} = \mathbf{x}_r$. *Then* $\mathbf{y}_r = \mathbf{e}_r$.

Proof. Let Y be the square matrix whose columns are $\mathbf{y}_1,\ldots, \mathbf{y}_n$. Then, since $\mathbf{y}_r = X'\mathbf{x}_r$ and the columns of X are $\mathbf{x}_1,\ldots, \mathbf{x}_n$,

$$Y = X'X = I.$$

Since the columns of I are $\mathbf{e}_1,\ldots, \mathbf{e}_n$, the result is established.

1.4. Linear dependence. As in § 1.3, $\mathbf{x}_r$ denotes a vector with elements

$$x_{1r}, x_{2r},\ldots, x_{nr}.$$

The vectors $\mathbf{x}_1,\ldots, \mathbf{x}_m$ are linearly dependent if there are numbers $l_1,\ldots, l_m$, not all zero, for which

$$l_1\,\mathbf{x}_1+\ldots+l_m\,\mathbf{x}_m = 0. \tag{1}$$

If (1) is true only when $l_1,\ldots, l_m$ are all zero, the vectors are linearly independent.

LEMMA 1. *Of any set of vectors at most* n *are linearly independent.*

Proof. Let there be given $n+k$ vectors. Let A be a matrix having these vectors as columns. The rank r of A cannot exceed n and, by Theorem 31 (with columns for rows) we can select r columns of A and express the others as sums of multiples of the selected r columns.

Aliter. Let $\mathbf{x}_1,\ldots, \mathbf{x}_n$ be any given set of linearly independent vectors; say

$$\mathbf{x}_s = \sum_{r=1}^{n} x_{rs}\,\mathbf{e}_r \quad (s = 1,\ldots, n). \tag{2}$$

The determinant $|X| = |x_{rs}| \neq 0$, since the given vectors are not linearly dependent. Let X_{rs} be the cofactor of x_{rs} in X; then, from (2),

$$\sum_{s=1}^{n} X_{rs}\,\mathbf{x}_s = X\mathbf{e}_r \quad (r = 1,\ldots, n). \tag{3}$$

Any vector $\mathbf{x}_p$ $(p > n)$ must be of the form

$$\mathbf{x}_p = \sum_{r=1}^{n} x_{rp}\,\mathbf{e}_r,$$

and therefore, by (3), must be a sum of multiples of the vectors $\mathbf{x}_1,\ldots, \mathbf{x}_n$.

LEMMA 2. *Any n mutually orthogonal unit vectors are linearly independent.*

Proof. Let $\mathbf{x}_1,\dots,\mathbf{x}_n$ be mutually orthogonal unit vectors. Suppose there are numerical multiples k_r for which

$$k_1\mathbf{x}_1+\dots+k_n\mathbf{x}_n = 0. \tag{4}$$

Pre-multiply by $\mathbf{x}'_s$, where s is any one of $1,\dots,n$. Then

$$k_s\mathbf{x}'_s\mathbf{x}_s = 0,$$

and hence, since $\mathbf{x}'_s\mathbf{x}_s = 1$, $k_s = 0$.

Hence (4) is true only when

$$k_1 = \dots = k_n = 0.$$

Aliter. By Theorem 59, $X'X = I$ and hence the determinant $|X| = \pm 1$. The columns $\mathbf{x}_1,\dots,\mathbf{x}_n$ of X are therefore linearly independent.

2. Latent vectors

2.1. Definition. We begin by proving a result indicated, but not worked to its logical conclusion, in an earlier chapter.†

THEOREM 60. *Let A be a given square matrix of n rows and columns and λ a numerical constant. The matrix equation*

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{1}$$

in which $\mathbf{x}$ is a single-column matrix of n elements, has a solution with at least one element of $\mathbf{x}$ not zero if and only if λ is a root of the equation

$$|A-\lambda I| = 0. \tag{2}$$

Proof. If the elements of $\mathbf{x}$ are $\xi_1, \xi_2,\dots, \xi_n$ the matrix equation (1) is equivalent to the n equations

$$\sum_{j=1}^{n} a_{ij}\xi_j = \lambda\xi_i \quad (i = 1,\dots,n).$$

These linear equations in the n variables $\xi_1,\dots,\xi_n$ have a non-zero solution if and only if λ is a root of the equation [Theorem 11]

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & \cdot & a_{nn}-\lambda \end{vmatrix} = 0, \tag{3}$$

and (3) is merely (2) written in full.

---

† Chapter X, § 7.

DEFINITION. *The roots in* λ *of* $|A-\lambda I| = 0$ *are the latent roots of the matrix* A; *when* λ *is a latent root, a non-zero vector* $\mathbf{x}$ *satisfying*

$$A\mathbf{x} = \lambda\mathbf{x}$$

is a LATENT VECTOR *of* A *corresponding to the root* λ.

2.2. THEOREM 61. *Let* $\mathbf{x}_r$ *and* $\mathbf{x}_s$ *be latent vectors that correspond to two distinct latent roots* λ_r *and* λ_s *of a matrix* A. *Then*

(i) $\mathbf{x}_r$ *and* $\mathbf{x}_s$ *are always linearly independent, and*
(ii) *when* A *is symmetrical,* $\mathbf{x}_r$ *and* $\mathbf{x}_s$ *are orthogonal.*

Proof. By hypothesis,

$$A\mathbf{x}_r = \lambda_r\,\mathbf{x}_r, \qquad A\mathbf{x}_s = \lambda_s\,\mathbf{x}_s. \tag{1}$$

(i) Let k_r, k_s be numbers for which

$$k_r\,\mathbf{x}_r + k_s\,\mathbf{x}_s = 0. \tag{2}$$

Then, since $A\mathbf{x}_s = \lambda_s\,\mathbf{x}_s$, (2) gives

$$Ak_r\,\mathbf{x}_r = -\lambda_s(k_s\,\mathbf{x}_s) = \lambda_s(k_r\,\mathbf{x}_r),$$

and so, by (1), $\qquad k_r\lambda_r\,\mathbf{x}_r = k_r\lambda_s\,\mathbf{x}_r.$

That is to say, when (2) holds,

$$k_r(\lambda_r-\lambda_s)\mathbf{x}_r = 0.$$

By hypothesis $\mathbf{x}_r$ is a non-zero vector and $\lambda_r \neq \lambda_s$. Hence $k_r = 0$ and (2) reduces to $k_s\,\mathbf{x}_s = 0$. But, by hypothesis, $\mathbf{x}_s$ is a non-zero vector and therefore $k_s = 0$.

Hence (2) is true only if $k_r = k_s = 0$; accordingly, $\mathbf{x}_r$ and $\mathbf{x}_s$ are linearly independent vectors.

(ii) Let $A' = A$. Then, by the first equation in (1),

$$\mathbf{x}_s' A\mathbf{x}_r = \mathbf{x}_s'\lambda_r\,\mathbf{x}_r. \tag{3}$$

The transpose of the second equation in (1) gives

$$\mathbf{x}_s' A' = \lambda_s\,\mathbf{x}_s',$$

and from this, since $A' = A$,

$$\mathbf{x}_s' A\,\mathbf{x}_r = \lambda_s\,\mathbf{x}_s'\,\mathbf{x}_r. \tag{4}$$

From (3) and (4), $\qquad (\lambda_r-\lambda_s)\mathbf{x}_s'\,\mathbf{x}_r = 0,$

and so, since the numerical multiplier $\lambda_r-\lambda_s \neq 0$, $\mathbf{x}_s'\,\mathbf{x}_r = 0$.

Hence $\mathbf{x}_r$ and $\mathbf{x}_s$ are orthogonal vectors.

3. Application to quadratic forms

We have already seen that (cf. Theorem 49, p. 153):

A real quadratic form $\boldsymbol{\xi}'A\boldsymbol{\xi}$ *in the n variables* $\xi_1,\ldots,\xi_n$ *can be reduced by a real orthogonal transformation*

$$\boldsymbol{\xi} = X\boldsymbol{\eta}$$

to the form $$\lambda_1 \eta_1^2+\lambda_2 \eta_2^2+\ldots+\lambda_n \eta_n^2,$$

wherein $\lambda_1,\ldots,\lambda_n$ *are the n roots of* $|A-\lambda I| = 0$.

We now show that the latent vectors of A provide the columns of the transforming matrix X.

3.1. When A has n distinct latent roots.

THEOREM 62. *Let A be a real symmetrical matrix having n distinct latent roots* $\lambda_1,\ldots,\lambda_n$. *Then there are n distinct real unit latent vectors* $\mathbf{x}_1,\ldots,\mathbf{x}_n$ *corresponding to these roots. If X is the square matrix whose columns are* $\mathbf{x}_1,\ldots,\mathbf{x}_n$, *the transformation*

$$\boldsymbol{\xi} = X\boldsymbol{\eta}$$

from variables $\xi_1,\ldots,\xi_n$ *to variables* $\eta_1,\ldots,\eta_n$ *is a real orthogonal transformation and*

$$\boldsymbol{\xi}'A\boldsymbol{\xi} = \lambda_1 \eta_1^2+\lambda_2 \eta_2^2+\ldots+\lambda_n \eta_n^2.$$

Proof. The roots $\lambda_1,\ldots,\lambda_n$ are necessarily real (Theorem 45) and so the elements of a latent vector $\mathbf{x}_r$ satisfying

$$A\mathbf{x}_r = \lambda_r \mathbf{x}_r \tag{1}$$

can be found in terms of real numbers† and, if the length of any one such vector is k, the vector $k^{-1}\mathbf{x}_r$ is a real unit vector satisfying (1). Hence there are n real unit latent vectors $\mathbf{x}_1,\ldots,\mathbf{x}_n$, and the matrix X having these vectors in its n columns has real numbers as its elements.

Again, by Theorem 61, $\mathbf{x}_r$ is orthogonal to $\mathbf{x}_s$ when $r \neq s$ and, since each is a unit vector,

$$\mathbf{x}_r'\mathbf{x}_s = \delta_{rs}, \tag{2}$$

where $\delta_{rs} = 0$ when $r \neq s$ and $\delta_{rr} = 1$ $(r = 1,\ldots,n)$. As in the

---

† When $\mathbf{x}_r$ is a solution and α is any number, real or complex, $\alpha\mathbf{x}_r$ is also a solution. For our present purposes, we leave aside all complex values of α.

proof of Theorem 59, the element in the rth row and sth column of the product $X'X$ is δ_{rs}; whence

$$X'X = I$$

and X is an orthogonal matrix.

The transformation $\boldsymbol{\xi} = X\boldsymbol{\eta}$ gives

$$\boldsymbol{\xi}'A\boldsymbol{\xi} = \boldsymbol{\eta}'X'AX\boldsymbol{\eta} \tag{3}$$

and the discriminant of the form in the variables $\eta_1,\ldots, \eta_n$ is the matrix $X'AX$. Now the columns of AX are $A\mathbf{x}_1,\ldots, A\mathbf{x}_n$, and these, from (1), are $\lambda_1 \mathbf{x}_1,\ldots, \lambda_n \mathbf{x}_n$. Hence

(i) the rows of X' are $\mathbf{x}_1',\ldots, \mathbf{x}_n'$,

(ii) the columns of AX are $\lambda_1 \mathbf{x}_1,\ldots, \lambda_n \mathbf{x}_n$ and the element in the rth row and sth column of $X'AX$ is the numerical value of

$$\mathbf{x}_r' \lambda_s \mathbf{x}_s = \lambda_s \mathbf{x}_r' \mathbf{x}_s = \lambda_s \delta_{rs},$$

by (2). Thus $X'AX$ has $\lambda_1,\ldots, \lambda_n$ as its elements in the principal diagonal and zero elsewhere. The form $\boldsymbol{\eta}'(X'AX)\boldsymbol{\eta}$ is therefore

$$\lambda_1 \eta_1^2 + \lambda_2 \eta_2^2 + \ldots + \lambda_n \eta_n^2.$$

3.2. When A has repeated latent roots. It is in fact true that, whether $|A-\lambda I| = 0$ has repeated roots or has all its roots distinct, *there are always n mutually orthogonal real unit latent vectors* $\mathbf{x}_1,\ldots, \mathbf{x}_n$ *of the* SYMMETRICAL *matrix A and, X being the matrix with these vectors as columns, the transformation* $\boldsymbol{\xi} = X\boldsymbol{\eta}$ *is a real orthogonal transformation that gives*

$$\boldsymbol{\xi}'A\boldsymbol{\xi} = \lambda_1 \eta_1^2 + \ldots + \lambda_n \eta_n^2,$$

wherein $\lambda_1,\ldots, \lambda_n$ *are the n roots (some of them possibly equal) of the characteristic equation* $|A-\lambda I| = 0$.

The proof will not be given here. The fundamental difficulty is to prove[†] that, when λ_1 (say) is a k-ple root of $|A-\lambda I| = 0$, there are k linearly independent latent vectors corresponding to λ_1. The setting up of a system of n mutually orthogonal unit

---

[†] *Quart. J. Math.* (Oxford) 18 (1947) 183–5 gives a proof by J. A. Todd; another treatment is given, ibid., pp. 186–92, by W. L. Ferrar. Numerical examples are easily dealt with by actually finding the vectors; see Examples XV, 11 and 14.

vectors is then effected by Schmidt's orthogonalization process† or by the equivalent process illustrated below.

Let $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ be three given real linearly independent vectors. Choose k_1 so that $k_1\mathbf{y}_1$ is a unit vector and let $\mathbf{x}_1 = k_1\mathbf{y}_1$. Choose the constant α so that

$$\mathbf{x}_1'(\mathbf{y}_2+\alpha\mathbf{x}_1) = 0; \tag{1}$$

that is, since $\mathbf{x}_1'\mathbf{x}_1 = 1$, $\alpha = -\mathbf{x}_1'\mathbf{y}_2$.‡

Since $\mathbf{y}_2$ and $\mathbf{x}_1$ are linearly independent, the vector $\mathbf{y}_2+\alpha\mathbf{x}_1$ is not zero and has a non-zero length, l_2 say. Put

$$\mathbf{x}_2 = l_2^{-1}(\mathbf{y}_2+\alpha\mathbf{x}_1).$$

Then $\mathbf{x}_2$ is a unit vector and, by (1), it is orthogonal to $\mathbf{x}_1$.

Now determine β and γ so that

$$\mathbf{x}_1'(\mathbf{y}_3+\beta\mathbf{x}_1+\gamma\mathbf{x}_2) = 0 \tag{2}$$

and

$$\mathbf{x}_2'(\mathbf{y}_3+\beta\mathbf{x}_1+\gamma\mathbf{x}_2) = 0. \tag{3}$$

That is, since $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal,

$$\beta = -\mathbf{x}_1'\mathbf{y}_3, \qquad \gamma = -\mathbf{x}_2'\mathbf{y}_3.$$

The vector $\mathbf{y}_3+\beta\mathbf{x}_1+\gamma\mathbf{x}_2 \neq 0$, since the vectors $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ are linearly independent; it has a non-zero length l_3 and with

$$\mathbf{x}_3 = l_3^{-1}(\mathbf{y}_3+\beta\mathbf{x}_1+\gamma\mathbf{x}_2)$$

we have a unit vector which, by (2) and (3), is orthogonal to $\mathbf{x}_1$ and $\mathbf{x}_2$. Thus the three $\mathbf{x}$ vectors are mutually orthogonal unit vectors.

Moreover, if $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ are latent vectors corresponding to the latent root λ of a matrix A, so that

$$A\mathbf{y}_1 = \lambda\mathbf{y}_1, \qquad A\mathbf{y}_2 = \lambda\mathbf{y}_2, \qquad A\mathbf{y}_3 = \lambda\mathbf{y}_3,$$

then also

$$A\mathbf{x}_1 = \lambda\mathbf{x}_1, \qquad A\mathbf{x}_2 = \lambda\mathbf{x}_2, \qquad A\mathbf{x}_3 = \lambda\mathbf{x}_3$$

and $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ are mutually orthogonal unit latent vectors corresponding to the root λ.

---

† W. L. Ferrar, *Finite Matrices*, Theorem 29, p. 139.

‡ More precisely, α is the numerical value of the element in the single-entry matrix $-\mathbf{x}_1'\mathbf{y}_2$. Both $\mathbf{x}_1'\mathbf{x}_1$ and $\mathbf{x}_1'\mathbf{y}_2$ are single-entry matrices.

4. Collineation

Let $\boldsymbol{\xi}$ be a single-column matrix with elements $\xi_1,\ldots,\xi_n$ and so for other letters. Let these elements be the current coordinates of a point in a system of n homogeneous coordinates and let A be a given square matrix of order n. Then the matrix relation

$$\mathbf{y} = A\mathbf{x} \tag{1}$$

expresses a relation between the variable point $\mathbf{x}$ and the variable point $\mathbf{y}$.

Now let the coordinate system be changed from $\boldsymbol{\xi}$ to $\boldsymbol{\eta}$ by means of a transformation

$$\boldsymbol{\xi} = T\boldsymbol{\eta},$$

where T is a non-singular square matrix. The new coordinates, $\mathbf{X}$ and $\mathbf{Y}$, of the points $\mathbf{x}$ and $\mathbf{y}$ are then given by

$$\mathbf{x} = T\mathbf{X}, \qquad \mathbf{y} = T\mathbf{Y}.$$

In the new coordinate system the relation (1) is expressed by

$$TY = ATX$$

or, since T is non-singular, by

$$Y = T^{-1}ATX. \tag{2}$$

The effect of replacing A in (1) by a matrix of the type $T^{-1}AT$ amounts to considering the same geometrical relation expressed in a different coordinate system. The study of such replacements, A by $T^{-1}AT$, is important in projective geometry. Here we shall prove only one theorem, the analogue of Theorem 62.

4.1. When the latent roots of A are all distinct.

THEOREM 63. *Let the square matrix A, of order n, have n distinct latent roots $\lambda_1,\ldots,\lambda_n$ and let $\mathbf{t}_1,\ldots,\mathbf{t}_n$ be corresponding latent vectors. Let T be the square matrix having these vectors as columns. Then*†

(i) *T is non-singular,*

(ii) $T^{-1}AT = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$.

Proof. (i) We prove that T is non-singular by showing that $\mathbf{t}_1,\ldots,\mathbf{t}_n$ are linearly independent.

---

† The NOTATION $\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$ indicates a matrix whose elements are $\lambda_1,\ldots,\lambda_n$ in the principal diagonal and zero everywhere else.

Let $k_1,\ldots,k_n$ be numbers for which

$$k_1\mathbf{t}_1+k_2\mathbf{t}_2+\ldots+k_n\mathbf{t}_n = 0. \tag{1}$$

Then, since $A\mathbf{t}_n = \lambda_n\mathbf{t}_n$,

$$A(k_1\mathbf{t}_1+\ldots+k_{n-1}\mathbf{t}_{n-1}) = \lambda_n(k_1\mathbf{t}_1+\ldots+k_{n-1}\mathbf{t}_{n-1}).$$

That is, since $A\mathbf{t}_r = \lambda_r\mathbf{t}_r$ for $r = 1,\ldots,n-1$,

$$k_1(\lambda_1-\lambda_n)\mathbf{t}_1+\ldots+k_{n-1}(\lambda_{n-1}-\lambda_n)\mathbf{t}_{n-1} = 0.$$

Since $\lambda_1,\ldots,\lambda_n$ are all different, this is

$$c_1k_1\mathbf{t}_1+\ldots+c_{n-1}k_{n-1}\mathbf{t}_{n-1} = 0, \tag{2}$$

wherein all the numbers $c_1,\ldots,c_{n-1}$ are non-zero.

A repetition of the same argument with $n-1$ for n gives

$$d_1k_1\mathbf{t}_1+\ldots+d_{n-2}k_{n-2}\mathbf{t}_{n-2} = 0,$$

wherein $d_1,\ldots,d_{n-2}$ are non-zero. Further repetitions lead, step by step, to

$$\alpha_1k_1\mathbf{t}_1 = 0, \qquad \alpha_1 \neq 0.$$

By hypothesis, $\mathbf{t}_1$ is a non-zero vector and therefore $k_1 = 0$.

We can repeat the same argument to show that a linear relation

$$k_2\mathbf{t}_2+\ldots+k_n\mathbf{t}_n = 0$$

implies $k_2 = 0$; and so on until we obtain the result that (1) holds only if

$$k_1 = k_2 = \ldots = k_n = 0.$$

This proves that $\mathbf{t}_1,\ldots,\mathbf{t}_n$ are linearly independent and, these being the columns of T, the rank of T is n and T is non-singular.

(ii) Moreover, the columns of AT are

$$\lambda_1\mathbf{t}_1,\ldots,\lambda_n\mathbf{t}_n.$$

Hence[†] $$AT = T\times\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$$

and it at once follows that

$$T^{-1}AT = \operatorname{diag}(\lambda_1,\ldots,\lambda_n).$$

4.2. When the latent roots are not all distinct. If A is not symmetrical, it does not follow that a k-ple root of $|A-\lambda I| = 0$ gives rise to k linearly independent latent vectors.

[†] If necessary, work out the matrix product

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

If it so happens that A has n linearly independent latent vectors $\mathbf{t}_1,\ldots,\mathbf{t}_n$ and T has these vectors as columns, then, as in the proof of the theorem,

$$T^{-1}AT = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$$

whether the λ's are distinct or not.

On the other hand, not all square matrices of order n have n linearly independent latent vectors. We illustrate by simple examples.

(i) Let
$$A = \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix}.$$

The characteristic equation is $(\alpha-\lambda)^2 = 0$ and a latent vector $\mathbf{x}$, with components x_1 and x_2, must satisfy $A\mathbf{x} = \alpha\mathbf{x}$; that is,

$$\alpha x_1+x_2 = \alpha x_1, \qquad \alpha x_2 = \alpha x_2.$$

From the first of these, $x_2 = 0$ and the only non-zero vectors to satisfy $A\mathbf{x} = \alpha\mathbf{x}$ are numerical multiples of

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

(ii) Let
$$A = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}.$$

A latent vector $\mathbf{x}$ must again satisfy $A\mathbf{x} = \alpha\mathbf{x}$. This now requires merely

$$\alpha x_1 = \alpha x_1, \qquad \alpha x_2 = \alpha x_2.$$

The two linearly independent vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and, indeed, all vectors of the form $x_1\mathbf{e}_1+x_2\mathbf{e}_2$ are latent vectors of A.

5. Commutative matrices and latent vectors

Two matrices A and B may or may not commute; they may or may not have common latent vectors. We conclude this chapter with two relatively simple theorems that connect the two possibilities. Throughout we take A and B to be square matrices of order n.

5.1. Let A and B have n linearly independent latent vectors in common, say $\mathbf{t}_1,\ldots,\mathbf{t}_n$. Let T be the square matrix having these vectors as columns.

Then $\mathbf{t}_1,\ldots,\mathbf{t}_n$ are latent vectors of A corresponding to latent roots $\lambda_1,\ldots,\lambda_n$, say, of A; they are latent vectors of B corresponding to latent roots $\mu_1,\ldots,\mu_n$, say, of B. As in § 4.1 (ii), p. 210,

$$\text{the columns of } AT \text{ are } \lambda_r\,\mathbf{t}_r \quad (r = 1,\ldots,n),$$

$$\text{the columns of } BT \text{ are } \mu_r\,\mathbf{t}_r \quad (r = 1,\ldots,n),$$

and, with the notation

$$L = \operatorname{diag}(\lambda_1,\ldots,\lambda_n), \qquad M = \operatorname{diag}(\mu_1,\ldots,\mu_n),$$

$$AT = TL, \qquad BT = TM.$$

That is $$T^{-1}AT = L, \qquad T^{-1}BT = M.$$

It follows that

$$T^{-1}ATT^{-1}BT = LM = ML = T^{-1}BTT^{-1}AT,$$

so that $$T^{-1}ABT = T^{-1}BAT$$

and, on pre-multiplying by T and post-multiplying by T^{-1},

$$AB = BA.$$

We have accordingly proved

THEOREM 64. *Two matrices of order n with n linearly independent latent vectors in common are commutative.*

5.2. Now suppose that $AB = BA$. Let $\mathbf{t}$ be a latent vector of A corresponding to a latent root λ of A. Then $A\mathbf{t} = \lambda\mathbf{t}$ and

$$AB\mathbf{t} = BA\mathbf{t} = B\lambda\mathbf{t} = \lambda B\mathbf{t}. \tag{1}$$

When B is non-singular $B\mathbf{t} \neq 0$ and is a latent vector of A corresponding to λ. [$B\mathbf{t} = 0$ would imply $\mathbf{t} = B^{-1}B\mathbf{t} = 0$.]

If every latent vector of A corresponding to λ is a multiple of $\mathbf{t}$, then $B\mathbf{t}$ is a multiple of $\mathbf{t}$, say

$$B\mathbf{t} = k\mathbf{t},$$

and $\mathbf{t}$ is a latent vector of B corresponding to the latent root k of B.

If there are m, but not more than m,[†] linearly independent latent vectors of A corresponding to the latent root λ, say

$$\mathbf{t}_1,\ldots,\mathbf{t}_m,$$

[†] By § 1.4, $m \leqslant n$.

then, since $$A\mathbf{t}_r = \lambda\mathbf{t}_r \quad (r = 1,..., m),$$
provided that $k_1,..., k_m$ are not all zero,
$$k_1\mathbf{t}_1 + ... + k_m\mathbf{t}_m \tag{2}$$
is a latent vector of A corresponding to λ. By our hypothesis that there are not more than m such vectors which are linearly independent, every latent vector of A corresponding to λ is of the form (2).

As in (1), $$AB\mathbf{t}_r = \lambda B\mathbf{t}_r \quad (r = 1,..., m)$$
so that $B\mathbf{t}_r$ is a latent vector of A corresponding to λ; it is therefore of the form (2). Hence there are constants k_{sr} such that
$$B\mathbf{t}_r = \sum_{s=1}^{m} k_{sr}\mathbf{t}_s \quad (r = 1,..., m).$$
For any constants $l_1,..., l_m$,
$$B\sum_{r=1}^{m} l_r\mathbf{t}_r = \sum_{r=1}^{m} l_r \sum_{s=1}^{m} k_{sr}\mathbf{t}_s = \sum_{s=1}^{m}\Big(\sum_{r=1}^{m} l_r k_{sr}\Big)\mathbf{t}_s.$$
Let θ be a latent root of the matrix
$$K = \begin{bmatrix} k_{11} & . & . & . & k_{1m} \\ . & . & . & . & . \\ k_{m1} & . & . & . & k_{mm} \end{bmatrix}$$
Then there are numbers $l_1,..., l_m$ (not all zero) for which
$$\sum_{r=1}^{m} l_r k_{sr} = \theta l_s \quad (s = 1,..., m)$$
and, with this choice of the l_r,
$$B\sum_{r=1}^{m} l_r\mathbf{t}_r = \theta\sum_{s=1}^{m} l_s\mathbf{t}_s.$$

Hence θ is a latent root of B and $\sum l_r\mathbf{t}_r$ is a latent vector of B corresponding to θ; it is also a latent vector of A corresponding to λ. We have thus proved†

THEOREM 65. *Let A, B be square matrices of order n; let $|B| \neq 0$ and let $AB = BA$. Then to each distinct latent root λ of A corresponds at least one latent vector of A which is also a latent vector of B.*

---

† For a comprehensive treatment of commutative matrices and their common latent vectors see S. N. Afriat, *Quart. J. Math.* (2) 5 (1954) 82–85.

EXAMPLES XV

1. Referred to rectangular axes $Oxyz$ the direction cosines of OX, OY, OZ are (l_1, m_1, n_1), (l_2, m_2, n_2), (l_3, m_3, n_3). The vectors OX, OY, OZ are mutually orthogonal. Show that

$$\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix}$$

is an orthogonal matrix.

2. Four vectors have elements

$$(a,a,a,a), \quad (b,-b,b,-b), \quad (c,0,-c,0), \quad (0,d,0,-d).$$

Verify that they are mutually orthogonal and find positive values of a, b, c, d that make these vectors columns of an orthogonal matrix.

3. The vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal and T is an orthogonal matrix. Prove that $\mathbf{X} = T\mathbf{x}$ and $\mathbf{Y} = T\mathbf{y}$ are orthogonal vectors.

4. The vectors $\mathbf{a}$ and $\mathbf{b}$ have components $(1,2,3,4)$ and $(2,3,4,5)$. Find k_1 so that

$$\mathbf{b}_1 = \mathbf{b}+k_1\mathbf{a}$$

is orthogonal to $\mathbf{a}$.

Given $\mathbf{c} = (1,0,0,1)$ and $\mathbf{d} = (0,0,1,0)$, find constants l_1, m_1, l_2, m_2, n_2 for which, when $\mathbf{c}_1 = \mathbf{c}+l_1\mathbf{b}_1+m_1\mathbf{a}$, $\mathbf{d}_1 = \mathbf{d}+l_2\mathbf{c}_1+m_2\mathbf{b}_1+n_2\mathbf{a}$, the four vectors $\mathbf{a}$, $\mathbf{b}_1$, $\mathbf{c}_1$, $\mathbf{d}_1$ are mutually orthogonal.

5. The columns of a square matrix A are four vectors $\mathbf{a}_1,\ldots,\mathbf{a}_4$ and

$$A'A = \operatorname{diag}(\alpha_1^2, \alpha_2^2, \alpha_3^2, \alpha_4^2).$$

Show that the matrix whose columns are

$$\alpha_1^{-1}\mathbf{a}_1,\ldots,\alpha_4^{-1}\mathbf{a}_4$$

is an orthogonal matrix.

6. Prove that the vectors with elements

$$(1,-2,3), \quad (0,1,-2), \quad (0,0,1)$$

are latent vectors of the matrix whose columns are

$$(1,2,-2), \quad (0,2,2), \quad (0,0,3).$$

7. Find the latent roots and latent vectors of the matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 2 & 3 \end{bmatrix},$$

showing (i) that the first has only two linearly independent latent vectors; (ii) that the second has three linearly independent latent vectors $\mathbf{l}_1 = (1,0,1)$, $\mathbf{l}_2 = (0,1,-1)$, $\mathbf{l}_3 = (0,0,1)$ and that $k_1\mathbf{l}_1+k_2\mathbf{l}_2$ is a latent vector for any values of the constants k_1 and k_2.

8. Prove that, if $\mathbf{x}$ is a latent vector of A corresponding to a latent root λ of A and $C = TAT^{-1}$, then

$$CT\mathbf{x} = TA\mathbf{x} = \lambda T\mathbf{x}$$

and $T\mathbf{x}$ is a latent vector of C corresponding to a latent root λ of C. Show also that, if $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, so also are $T\mathbf{x}$ and $T\mathbf{y}$.

9. *Find* three unit latent vectors of the symmetrical matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 1 & -3 \end{bmatrix}$$

in the form

$$(1,0,0), \qquad \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \qquad \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

and deduce the orthogonal transformation whereby

$$2x^2-3y^2-3z^2+2yz = 2X^2-2Y^2-4Z^2.$$

10. (*Harder*†) Find an orthogonal transformation from variables x, y, z to variables X, Y, Z whereby

$$2yz+2zx+2xy = 2X^2-Y^2-Z^2.$$

11. Prove that the latent roots of the discriminant of the quadratic form

$$2yz+2zx+2xy$$

are -1, -1, 2.

Prove that, corresponding to $\lambda = -1$,

(i) (x,y,z) is a latent vector whenever $x+y+z = 0$;

(ii) $(1,-1,0)$ and $(1,0,-1)$ are linearly independent and that $(1,-1,0)$ and $(1+k,-k,-1)$ are orthogonal when $k = -\frac{1}{2}$;

(iii) $$\mathbf{a} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \qquad \mathbf{b} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$$

are orthogonal unit latent vectors.

Prove that a unit latent vector corresponding to $\lambda = 2$ is

$$\mathbf{c} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

Verify that when T is the matrix having $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as columns and $\mathbf{x} = T\mathbf{X}$, where $\mathbf{x} = (x,y,z)$ and $\mathbf{X} = (X,Y,Z)$,

$$2yz+2zx+2xy = -X^2-Y^2+2Z^2.$$

12. Prove that $x = Z$, $y = Y$, $z = X$ is an orthogonal transformation.

With n variables $x_1,\ldots,x_n$, show that

$$x_1 = X_\alpha, \quad x_2 = X_\beta, \quad \ldots, \quad x_n = X_\kappa,$$

where $\alpha,\beta,\ldots,\kappa$ is a permutation of $1,2,\ldots,n$, is an orthogonal transformation.

13. Find an orthogonal transformation $\mathbf{x} = T\mathbf{X}$ which gives

$$2yz+2zx = (Y^2-Z^2)\sqrt{2}.$$

14. When $r = 1,2,3$, $s = 2,3,4$ and

$$\mathbf{x}'A\mathbf{x} = 2\sum_{r<s} x_r x_s,$$

---

† Example 11 is the same sum broken down into a step-by-step solution.

prove that the latent roots of A are -1, -1, -1, 3 and that

$$(1,0,0,-1),\quad (0,1,-1,0),\quad (1,-1,-1,1),\quad (1,1,1,1)$$

are mutually orthogonal latent vectors of A.

Find the matrix T whose columns are unit vectors parallel to the above (i.e. numerical multiples of them) and show that, when $\mathbf{x} = T\mathbf{X}$,

$$\mathbf{x}'A\mathbf{x} = -X_1^2-X_2^2-X_3^2+3X_4^2.$$

15. Show that there is an orthogonal transformation[†] $\mathbf{x} = T\mathbf{X}$ whereby

$$2x_1x_3+2x_1x_4-4x_2x_3+4x_2x_4 = \surd 2(X_1^2+2X_2^2-2X_3^2-X_4^2).$$

16. A square matrix A, of order n, has n mutually orthogonal unit latent vectors. Prove that A is symmetrical.

17. Find the latent vectors of the matrices

$$\begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},\qquad \begin{bmatrix} -2 & -1 & -5 \\ 1 & 2 & 1 \\ 3 & 1 & 6 \end{bmatrix}.$$

18. The matrix A has latent vectors

$$(1,0,0),\quad (1,1,0),\quad (1,2,3)$$

corresponding to the latent roots $\lambda = 1, 0, -1$. The matrix B has the same latent vectors corresponding to the latent roots $\mu = 1, 2, 3$. Find the elements of A and B and verify that $AB = BA$.

19. The square matrices A and B are of order n; A has n distinct latent roots $\lambda_1,\ldots,\lambda_n$ and B has non-zero latent roots $\mu_1,\ldots,\mu_n$, not necessarily distinct; and $AB = BA$. Prove that there is a matrix T for which

$$T^{-1}AT = \operatorname{diag}(\lambda_1,\ldots,\lambda_n),\qquad T^{-1}BT = \operatorname{diag}(\mu_1,\ldots,\mu_n).$$

HINTS AND ANSWERS

2. $a = b = \frac{1}{2}$, $c = d = 1/\surd 2$; these give *unit* vectors.

3. Work out $X'Y$.

4. $k_1 = -\frac{4}{3}$; $l_1 = -\frac{1}{2}$, $m_1 = -\frac{1}{6}$; $l_2 = \frac{1}{2}$, $m_2 = 0$, $n_2 = -\frac{1}{10}$; cf. § 3.2.

6. $\lambda = 1, 2, 3$.

7. (i) $\lambda = 1, 1, 3$; vectors $(0, 1, -1)$, $(0, 0, 1)$.

9. $\lambda = 2, -2, -4$; use Theorem 62.

12. $\mathbf{x} = T\mathbf{X}$ gives $\mathbf{x}'I\mathbf{x} = \mathbf{X}'T'IT\mathbf{X} = \mathbf{X}'T'T\mathbf{X}$. The given transformation gives $\sum x^2 = \sum X^2$; therefore $T'T = I$.

13. $\lambda = 0, \pm\surd 2$; unit latent vectors are

$$\left(\frac{1}{\surd 2}, -\frac{1}{\surd 2}, 0\right),\qquad \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\surd 2}\right),\qquad \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{\surd 2}\right).$$

Use Theorem 62.

[†] The actual transformation is not asked for: too tedious to be worth the pains.

16. Let $\mathbf{t}_1,\ldots,\mathbf{t}_n$ be the vectors and T have the $\mathbf{t}$'s as columns. As in § 4.1, $T^{-1}AT = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. But $T^{-1} = T'$ and so

$$T'AT = \operatorname{diag}(\lambda_1,\ldots,\lambda_n).$$

A diagonal matrix is its own transpose and so

$$T'AT = (T'AT)' = T'A'T \qquad \text{and} \qquad A = A'.$$

17. (i) $\lambda = -1, 3, 4$; vectors $(3, -1, 0)$, $(1, 1, -4)$, $(2, 1, 0)$.
(ii) $\lambda = 1, 2, 3$; vectors $(2, -1, -1)$, $(1, 1, -1)$, $(1.0, -1)$.

18. $$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & -\frac{2}{3} \\ 0 & 0 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & \frac{2}{3} \\ 0 & 0 & 3 \end{bmatrix}.$$

19. A has n linearly independent latent vectors (Theorem 63), which are also latent vectors of B (Theorem 65). Finish as in proof of Theorem 64.

INDEX

Numbers refer to pages. Note the entries under 'Definitions'.

Absolute invariant 169.
Adjoint, adjugate 55, 85.
Algebraic form 169.
Alternant 22.
Array 50.

Barnard and Child 61.
Bilinear form 120.
  as matrix product 122.
Binary cubic 188.
Binary form 176, 190.
Binary quartic 190.

Canonical form 146, 151, 188.
Characteristic equation:
  of matrix 110.
  of quadratic form 144.
  of transformation 131.
Circulant 23.
Co-factor 28.
Cogredient 129.
Collineation 209.
Commutative matrices 211.
Complementary minor 56.
Contragredient 129.
Covariant:
  definition 172.
  of covariant 187.
  of cubic 176.

Definitions:
  adjoint matrix 85.
  adjugate determinant 55.
  alternant 22.
  bilinear form 120.
  characteristic equation 110.
  circulant 23.
  covariant 172.
  determinant 8.
  discriminant 122.
  invariant 171, 172.
  isobaric form 183.
  latent root 110.
  linear form 119.
  matrix 69.
  Pfaffian 61.
  quadratic form 120.
  rank of matrix 94.
  reciprocal matrix 85.
  unit matrix 77.
  weight 182.
Determinant:
  adjugate 55.
  co-factors 28.
  definition 7, 8, 25.
  differentiation 25.
  expansions 12, 36, 39.
  Hessian 174.
  Jacobian 174.
  Jacobi's theorem 55.
  minors (q.v.) 28.
  notation 21.
  product 41.
  skew-symmetrical 59.
  symmetrical 59.
Division law 77, 89.

Eliminant 176.
Elliott, E. B. 198.
Equations:
  characteristic 110, 131, 144.
  consistent 31.
  homogeneous 32.
  λ equation of forms 144.
Expansion of determinant 12, 36.

Field, number 2, 94, 119, 122.
Fundamental solutions 101.

Grace and Young 198.

Hermitian form 3, 128, 142, 143, 158.
Hermitian matrix 129.
Hessian 174.

Invariant:
  arithmetic 195.
  definition 171.
  further reading on 198.
  geometrical 191.
  joint 178.
  lineo-linear 199.
  metrical 194.
  of covariant 185.
  of cubic 176.
  of quartic 193.
  properties of 182.

Jacobians 174, 193.
Jacobi's theorem 55.

Kronecker delta 84.

Laplace 36.
Latent roots 110, 131, 144, 145, 151, 153, 206, 207.
    vectors 204.
Linear dependence 94, 203, 204.
    equations 30, 98.
    form 119.

Linear substitution 67.
        transformation 122.

Matrix:
  addition 69.
  adjoint, adjugate 85.
  characteristic equation 110.
  commutative 211.
  definition 69.
  determinant of 78.
  division law 77, 89.
  elementary transformation 113.
  equivalent 113.
  Hermitian 129.
  latent roots of 110.
  multiplication 73.
  orthogonal 160, 202.
  rank 93.
  rank of product 109.
  reciprocal 85.
  singular 85.
  transpose 83.
  unit 77.
Minors 28.
  of zero determinant 33.
  complementary 56.
Modulus 122.
Muir, Sir T. 61.
Multiples, sum of 94.

Notation:
  algebraic forms 170.
  bilinear forms 121, 122.
  determinants 21.
  matrices 69.
  minors 28, 29.
  'number' 67.
  substitutions 67.

Orthogonal matrix 160, 202
            transformation 153.
            vectors 201, 205.

Permutations 24.
Pfaffian 61.
Positive-definite form 135.

Quadratic form:
  as matrix product 122.
  canonical form 146, 151, 206.
  definition 120.
  discriminant 122, 127.
  equivalent 156.
  positive-definite 135.
  rank 148.
  signature 154.

Rank
  of matrix 93.
  of quadratic form 148.
Reciprocal matrix 85.
Reversal law
  for reciprocal 88.
  for transpose 83.
Rings, number 1.

Salmon, G. 61.
Schmidt's orthogonalization process 208.
Sign of term 11, 29, 37, 39.
Signature 154.
Substitutions, linear 67.
Summation convention 41.
Sum of four squares 49.

Transformation:
  characteristic equation 131.
  elementary 147.
  latent roots 131.
  modulus 122.
  non-singular 122.
  of quadratic form 124, 206.
  orthogonal, 153, 160, 206.
  pole of 131.
  product 123.
Transpose 83.
Transvectant 196.
Turnbull, H. W. 198.

Vectors:
  as matrices 200.
  latent 204.
  orthogonal 201, 205.
  unit 201.

Weight 182, 184.
Weitzenbock, R. 198.
Winger, R. M. 132.