TIGHT BINDING BOOK

UNIVERSAL
LIBRARY

FROM
DETERMINANT
TO
TENSOR

BY

W. F. SHEPPARD, Sc.D., LL.M.

FORMERLY FELLOW OF TRINITY COLLEGE, CAMBRIDGE

OXFORD

AT THE CLARENDON PRESS

1923

OXFORD UNIVERSITY PRESS
London Edinburgh Glasgow Copenhagen
New York Toronto Melbourne Cape Town
Bombay Calcutta Madras Shanghai
HUMPHREY MILFORD
Publisher to the University

Printed in England

PREFACE

THE tensor calculus used in the mathematical treatment of relativity, and concisely explained by Professor A. S. Eddington in his 'Report on the Relativity Theory of Gravitation', is, like the various kinds of vector calculus, a system of condensed notation which not only conduces to economy in the writing of symbols, but, what is more important, enables spatial and physical relationships to be grasped as a whole without having to be built up from a number of components which really represent views from different parts of space. Three-dimensional geometry or physics is troublesome enough: the addition of a fourth dimension made the need of a condensed notation imperative.

Professor Eddington has recently pointed out that the tensor notation and methods can be applied, with happy results, to other and more elementary classes of problems than those for which they were originally devised; and this book is an attempt to put his somewhat compressed exposition into a form in which it may appeal to a larger circle of readers. The book, therefore, is not intended as an introduction to the mathematical theory of relativity—though I hope it may be of some use for that purpose—but rather as an exercise in the elementary application of methods which, apart from any practical use, possess a special beauty of their own.

The new notation is not introduced until the fifth chapter. The properties of determinants, which serve as the starting-point for the application of the notation, are familiar to the mathematician; but, as I hope the book may be read by some who are not entirely at ease with determinants, I have commenced with four chapters on the elementary theory of the subject. I make no apology for doing this, instead of referring the reader to the ordinary text-book on algebra. The text-book treatment is not always stimulating; the reasons for the various stages are not necessarily clear to the student; and attempt at simplicity sometimes leads to loss of rigidity in proof. In such a subject it is necessary to take the reader into one's confidence; and this earlier part may in this respect be found helpful to some, teachers or students, to whom the later part makes at first a less strong appeal.

I have added a chapter on some applications to the theory of statistics, to which the tensor calculus seems specially suitable. The basis of this portion, so far as method is concerned, is a short paper by Professor Eddington, mentioned at the end of the chapter. This is one only of many possible applications.

What I have called double sets will be recognized by the advanced student as matrices; and many of the propositions will be found to be familiar. But the tensor calculus may fairly claim that, in bringing into close relation various branches of mathematical study, previously regarded as distinct, it gives them a new life.

I have to thank Professor Eddington for looking at my manuscript and making some corrections and suggestions.

5 *June* 1923. W. F. S.

CONTENTS

ERRATA

PAGE 34, line 4 from bottom, for 'the 0's' read 'the first m 0's'.

PAGE 76, line 2, for 'of X's' read 'of the X's'.

Sheppard *Tensor Calculus*

INTRODUCTION

As this is a comparatively new subject to most readers, it may be as well to explain briefly what it is about.

A vector in (say) 3 dimensions is a directed quantity, determined as regards both direction and magnitude by its components, which are magnitudes measured along three definite axes. These axes being supposed to have been fixed beforehand, we can take them in some definite order; and a vector $\mathfrak{A}$ is then determined by a set of 3 quantities, which we may call

$$A_1 \quad A_2 \quad A_3.$$

Algebraically, the idea of a vector can be extended to something which is determined by a set of m quantities

$$A_1 \quad A_2 \quad A_3 \ldots A_m,$$

where m may have any value.

A determinant (say)

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

is the algebraical sum of all the products that can be formed in a certain way according to a certain rule of signs from the set of quantities

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

Each 'element' of this set has its position fixed in the set by the numbers of horizontal and of vertical steps required to reach it from the initial element a. Thus the set is extended in two directions, while the set which determines a vector is extended in one direction only. This applies to a determinant with any number m^2 of elements.

In the same way we might have a set extended in 3 directions, the symbols being written along the edges of a cube and along lines inside the cube or on its faces; and we could, in theory, increase the number of 'directions' to 4 or more, by proper convention as to the order in which the elements are to be taken. On the other hand, a single quantity—what in the language of vectors is called a scalar—may be regarded as a set not extended in any direction.

The tensor calculus, using the word 'tensor' in its broad sense, deals with all these different kinds of sets, in relation to sets of variables by which we can regard axes of reference as being determined. In the narrower sense in which the word is used in reference to the theory of relativity, only sets which satisfy certain conditions are called tensors.

In this book I have treated the tensor calculus as arising out of the use of determinants. Chapters I–IV deal with the elementary theory of determinants, so far as it is required for our purpose. (The student who is familiar with determinants can skip these chapters.) In Chapter V the tensor notation is introduced in successive steps, with explanatory remarks. These latter are in small print, not because they are less important, but in order not to break the continuity of the chapter as a whole. In Chapter VI these explanatory paragraphs (or parts of them) are brought together and amplified so as to give a general idea of the elementary properties of sets. Chapters VII and VIII deal with some developments of the subject in its general aspect. Chapter IX shows the application of the methods to certain problems in the theory of statistics and of error; this can be omitted by any one who wishes to pass on to Chapter X, which deals very briefly with the tensor in its more limited sense, as applied to the theory of relativity.

DETERMINANTS

I. ORIGIN OF DETERMINANTS

I. 1. **Solution of simultaneous equations.**—Determinants ordinarily arise out of the solution of simultaneous equations. Suppose we have two equations

$$\left.\begin{aligned} 5x+2y &= 19 \\ 4x+3y &= 18 \end{aligned}\right\}.$$

Then, if we used only elementary methods, we could multiply the first by 3 and the second by 2, which would give

$$\left.\begin{aligned} 15x+6y &= 57 \\ 8x+6y &= 36 \end{aligned}\right\};$$

and thence, by subtracting, we should have

$$7x = 21, \quad x = 3,$$

whence either equation would give

$$y = 2.$$

Similarly, if we had three equations

$$\left.\begin{aligned} 2x+5y+3z &= 4 \\ x-3y-2z &= -1 \\ -5x-4y+z &= 7 \end{aligned}\right\},$$

we could, by eliminating z between the first and the second and between the second and the third, obtain

$$\left.\begin{aligned} 7x+y &= 5 \\ -9x-11y &= 13 \end{aligned}\right\};$$

whence, proceeding as before, we should obtain

$$x = 1, \quad y = -2, \quad z = 4.$$

This process of successive elimination is tedious, especially when there are more than two unknowns; and it is found better to obtain a formula for the general solution and apply it to the numerical values of the particular case.

Since such an equation as $x-3y-2z = -1$ can be written in the form $(+1)x+(-3)y+(-2)z = (-1)$, we can use positive signs throughout, it being understood that the quantity represented by any symbol may be either positive or negative.

I. 2. **Formula for solution.**—(i) For completeness we begin with one unknown. The equation

$$a_1x = k_1$$

gives

$$x = \frac{k_1}{a_1}.$$

(ii) For two unknowns the equations may be written

$$\left.\begin{aligned} a_1x+b_1y &= k_1 \\ a_2x+b_2y &= k_2 \end{aligned}\right\}.$$

Multiplying the first equation by b_2 and the second by b_1, and subtracting, we get

$$(a_1b_2-a_2b_1)x = k_1b_2-k_2b_1,$$

whence

$$x = \frac{k_1b_2-k_2b_1}{a_1b_2-a_2b_1}.$$

Similarly, interchanging a's and b's,

$$y = \frac{k_1a_2-k_2a_1}{b_1a_2-b_2a_1}$$

$$= \frac{a_1k_2-a_2k_1}{a_1b_2-a_2b_1}.$$

It should be noticed that the expressions for x and for y have the same denominator, and that the numerators are obtained from the denominator by replacing the a's in the one case, and the b's in the other, by k's.

(iii) Next take the case of three unknowns. Let the equations be

$$\left.\begin{aligned} a_1x+b_1y+c_1z &= k_1 \\ a_2x+b_2y+c_2z &= k_2 \\ a_3x+b_3y+c_3z &= k_3 \end{aligned}\right\}.$$

Eliminating z from the first two equations, we obtain

$$(a_1c_2-a_2c_1)x+(b_1c_2-b_2c_1)y = k_1c_2-k_2c_1.$$

Similarly from the second and third equations

$$(a_2c_3-a_3c_2)x+(b_2c_3-b_3c_2)y = k_2c_3-k_3c_2.$$

Then, eliminating y from these equations, we get

$$x = \frac{(k_1c_2-k_2c_1)(b_2c_3-b_3c_2)-(k_2c_3-k_3c_2)(b_1c_2-b_2c_1)}{(a_1c_2-a_2c_1)(b_2c_3-b_3c_2)-(a_2c_3-a_3c_2)(b_1c_2-b_2c_1)}.$$

As before, the numerator is got from the denominator by replacing a's by k's, and we therefore need only consider the denominator. Multiplying out, it becomes

$$a_1b_2c_2c_3-a_1b_3c_2^2-a_2b_2c_1c_3+a_2b_3c_1c_2-a_2b_1c_2c_3+a_2b_2c_1c_3 \\ +a_3b_1c_2^2-a_3b_2c_1c_2$$

$$=c_2(a_1b_2c_3-a_1b_3c_2-a_2b_1c_3+a_2b_3c_1+a_3b_1c_2-a_3b_2c_1).$$

Hence, replacing the a's by k's for the numerator,

$$x = \frac{k_1b_2c_3-k_1b_3c_2-k_2b_1c_3+k_2b_3c_1+k_3b_1c_2-k_3b_2c_1}{a_1b_2c_3-a_1b_3c_2-a_2b_1c_3+a_2b_3c_1+a_3b_1c_2-a_3b_2c_1}.$$

Corresponding expressions can be obtained for y and for z.

I. 3. **General problem.**—(i) We might proceed in the same way for equations in four or more unknowns. But this would mean that each case would have to be considered separately; and not only should we fail to get a general formula, but the algebraical work would soon become practically impossible. We therefore alter our tactics.

We write down the general equations involving m unknowns $x, y, z, \ldots w$

$$\left.\begin{array}{r} a_1x+b_1y+\ c_1z+\ldots+f_1w=k_1 \\ a_2x+b_2y+\ c_2z+\ldots+f_2w=k_2 \\ \vdots \\ a_mx+b_my+c_mz+\ldots+f_mw=k_m \end{array}\right\}, \quad \text{(I. 3. A)}$$

guess at a solution, and then verify that this solution does actually satisfy the equations.

(ii) The values of $x, y, z \ldots w$ as found from these equations will be in the form of fractions. We will consider first the denominators. Putting together the results obtained in § 2, for the cases of $m = 1, 2, 3$, we find that the successive denominators, which we will call $D^{(1)}$, $D^{(2)}$, $D^{(3)}$, are

$$\left.\begin{array}{l} D^{(1)} = a_1 \\ D^{(2)} = a_1b_2 - a_2b_1 \\ D^{(3)} = a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2b_3c_1 + a_3b_1c_2 \\ \qquad\qquad - a_3b_2c_1 \end{array}\right\}. \quad \text{(I. 3. 1)}$$

We want to obtain an expression $D^{(m)}$, which we should guess to be the common denominator in the solution of the equations (I. 3. A), and of which $D^{(1)}$, $D^{(2)}$, $D^{(3)}$ shall be the particular cases for $m = 1, 2, 3$.

(iii) The three D's in (I. 3. 1) have a general similarity, which enables us to obtain a formula for $D^{(m)}$. It will be seen that both in $D^{(2)}$ and in $D^{(3)}$ some of the terms have sign + and some have sign −. We will see first how the terms are constructed, and then consider the question of sign.

I. 4. **Construction of terms.**—(i) For each of the three D's, for which the values of m are 1, 2, 3 respectively, each term is the product of m factors, which are the coefficients in the equations in § 2. In writing down these coefficients, it is convenient to keep them in the relative positions in which they occur in the equations. Thus we get

For $D^{(1)}$	For $D^{(2)}$	For $D^{(3)}$
a_1	$a_1 \quad b_1$	$a_1 \quad b_1 \quad c_1$
	$a_2 \quad b_2$	$a_2 \quad b_2 \quad c_2$
		$a_3 \quad b_3 \quad c_3$

In each case we have a **set** of quantities arranged in the form of a square. The individual quantities are called the **elements** of the set; the quantities in a vertical line constitute a **column**, and the quantities in a horizontal line constitute a **row**. The columns are numbered from the left, and the rows from the top. The diagonal drawn from the top left-hand corner—i. e. the diagonal through a_1—is called the **leading diagonal.**

(ii) Each term contains m factors, which are taken from the set in such a way that one (only) shall come from each column and that one (only) shall come from each row. Also $D^{(m)}$ contains every term which can be constructed in this way. Take, for example, $D^{(3)}$. Since one factor is to come from each column, the factors are an a, a b, and a c. The a can be either a_1 or a_2 or a_3, i. e. it can be taken in three ways; when one of these three a's has been taken, one row has been used up, and the b can only be taken in two ways; and, when one of the two b's has been taken, the c can similarly only be taken in one way. There are therefore $3 \,.\, 2 \,.\, 1 = 6$ possible combinations of factors; and this is the number of terms in $D^{(3)}$.

(iii) Another way of stating the thing is that, if we keep to a fixed order $a\ b\ c$ of the factors in each term, the suffixes of the factors are the numbers 1 or 1 2 or 1 2 3, arranged in different ways, and there is one term for each of the possible arrangements.

(iv) We conclude that $D^{(m)}$ contains terms, each of which is constructed by taking m factors from the set of $m \times m = m^2$ quantities

$$\begin{matrix} a_1 & b_1 & c_1 \dots f_1 \\ a_2 & b_2 & c_2 \dots f_2 \\ \vdots & \vdots & \vdots \quad \vdots \\ a_m & b_m & c_m \dots f_m \end{matrix}$$

in such a way that there shall be one factor (only) taken from each column and one (only) from each row; there being one term for each of the $m\ (m-1) \dots 1 \equiv m!$ ways in which this can be done. Or, which comes to the same thing, that the terms are made up of factors $a\ b\ c \dots f$ with suffixes $1\ 2\ 3 \dots m$ arranged in different ways, there being one term for each of the $m!$ possible different arrangements.

I. 5. **Rule of signs.**—(i) It will be seen that, in the D's after $D^{(1)}$, half of the terms are positive and half negative, and that in each case the term found from the elements in the leading diagonal—namely a_1 or $a_1 b_2$ or $a_1 b_2 c_3$—is positive. We should therefore expect that half of the terms in $D^{(m)}$ would be positive and half negative, the term $a_1\ b_2\ c_3 \dots f_m$—which we call the **leading term** and usually place first—being positive. The difficult question is that of signs. In the case of (say) $m = 5$, how are we to know whether such a term as $a_3 b_5 c_2 d_4 e_1$ is to have the sign $+$ or $-$?

(ii) The sign of a term must, if the letters $a\ b\ c \dots f$ are kept in their original order, depend on the arrangement of the suffixes, i. e. on the extent to which this has departed from the initial arrangement $1\ 2\ 3 \dots m$. Now any arrangement such as 35241 can be got from the initial arrangement 12345 by a series of interchanges of adjacent figures. We must fix a definite order in which these interchanges are to be made. We therefore say that each figure is to be moved in turn, beginning with that which ultimately comes first, then that which ultimately comes second, and so on. Thus in this particular case the successive stages would (repeating for each group of interchanges the arrangement from which we start) be 12345, 13245, 31245; 1245, 1254, 1524, 5124; 124, 214; 14, 41; 1: a total of seven interchanges. Now let us look at the signs in $D^{(2)}$ and $D^{(3)}$. In $D^{(2)}$ the arrangement 21 is obtained from 12 by 1 interchange, and the sign for 21 is $-$. In $D^{(3)}$ the signs of the successive terms, and the stages by which the final arrangements of suffixes are obtained, are as follows, the numbers of interchanges being added in heavy type:

+	123	**0**	+	123, 213, 231	**2**
−	123, 132 . .	**1**	+	123, 132, 312	**2**
−	123, 213 . .	**1**	−	123, 132, 312, 321 . .	**3**

It will be seen that both for $D^{(2)}$ and for $D^{(3)}$ the sign is $-$ or $+$ according as the number of interchanges is odd or even. We therefore adopt this as our rule; in the case of $a_3b_5c_2d_4e_1$, for instance, seven interchanges are necessary, and the sign is therefore $-$.

(iii) In order to find the sign of any given term by the above rule, it would be necessary to perform all the interchanges. A shorter method is to look at the term as it stands and to consider the *reversals of order* in it; i.e. taking the suffixes of the term in pairs in every possible way without altering their order, to see in how many cases the numbers are in the reverse of their order in the leading term, i.e. are in descending instead of ascending order. The term $a_3b_5c_2d_4e_1$, for instance, gives the following pairs, those in which the order is reversed being printed in heavier type:—35, **32**, 34, **31**, **52**, **54**, **51**, 24, **21**, **41**.

It is easily seen that each interchange, of the kind described in (ii) above, produces one reversal of order; for, while we are shifting one number, such as the 5 in the second group of arrangements there shown, the relative order of the other numbers remains unaltered. It follows that the number of reversals of order is the same as the number of interchanges of this kind; and therefore *the sign of a term will be − or + according as the number of reversals of order is odd or even.*

(iv) *The interchange of any two suffixes in a term changes the sign of the term.*

[Let the two suffixes be ϕ and ψ; ϕ coming before ψ in the term in question, but not necessarily being before it in numerical order.

(1) First let ϕ and ψ be adjacent. Then the interchange of ϕ and ψ increases or decreases the number of reversals of order by 1, and therefore changes the sign of the term.

(2) Next suppose that there are x suffixes between ϕ and ψ. Then we can move ψ in front of ϕ by $x + 1$ interchanges with the adjacent term, and then move ϕ into the original position of ψ by x interchanges. This is a total of $2x + 1$ interchanges, each of which in succession makes a change of sign: the total result is to change the sign of the term.]

(v) We have so far assumed that the factors of a term are arranged in the original order of the letters $a\ b\ c\ d \ldots$. Now suppose that the order of the factors is altered in any way. How does this affect the rule of signs?

The alteration of order can be brought about by a series of interchanges of factors. Suppose there is an interchange of a_ϕ and b_ψ. Then, by (iv), the number of reversals of order of suffixes is altered by an odd number, but the number of reversals of order of letters is also altered by an odd number; and therefore, if we consider the sum of the numbers of reversals of order of letters and of suffixes, this sum either is not altered or is altered by an even number. It follows that, *if the factors of a term have been shifted about so that the letters a b c ... are not in their original order, the sign of the term depends on the sum of the numbers of reversals of order of letters and of suffixes respectively,* being − or + according as this sum is odd or even. For example, in $d_4 b_5 c_2 a_3 e_1$ there are five reversals of order of the letters and eight of the suffixes, so that the sign is −.

II. PROPERTIES OF DETERMINANTS

II. 1. **Definition of determinant.**—We can now combine the results obtained in I. 4 and I. 5. We suppose that we are dealing with a set of $m \times m = m^2$ quantities, which we can arrange in the form of a square, thus (the quantities being denoted by crosses):

$$\begin{matrix} \times & \times & \times \ldots \times \\ \times & \times & \times \ldots \times \\ \vdots & \vdots & \vdots \quad \vdots \\ \times & \times & \times \ldots \times \end{matrix}$$

Then the expression which we have to consider is the algebraical sum of a number of terms, of which some are taken positively and some negatively. Each term (apart from sign) is the product of m elements of the set, taken in such a way that one element (only) shall come from each column and that one element (only) shall come from each row; and there are $m!$ terms, corresponding to the $m!$ different ways in which this can be done. The leading term is the term containing the elements in the leading diagonal of the square, and has sign $+$. The signs of the other terms are to be found by replacing the elements of the set by $a_1, a_2, \ldots b_1$, etc., arranged as a *key set*:

$$\begin{matrix} a_1 & b_1 & c_1 \ldots f_1 \\ a_2 & b_2 & c_2 \ldots f_2 \\ \vdots & \vdots & \vdots \quad \vdots \\ a_m & b_m & c_m \ldots f_m \end{matrix}$$

The sign of a term is then $-$ or $+$ according as the sum of the numbers of reversals of order of the letters and of the suffixes, as compared with the leading term $a_1 b_2 c_3 \ldots f_m$, is odd or even.

The algebraical sum of the terms so obtained, namely $a_1 b_2 c_3 \ldots f_m - a_2 b_1 c_3 \ldots f_m +$ etc., is called the **determinant** of the set, and will be denoted by D. It should however be observed that it is the determinant of the set *as so arranged*; with different arrangements of the elements of a set, still keeping them in a square, we may obtain different determinants.

We can therefore define the determinant as *the algebraical sum of terms of the form* $a_p b_q c_r \ldots$, *where* $p\ q\ r \ldots$ *are the numbers* 1 2 3 . . . m *arranged in some order, there being a term for each of the* $m!$ *possible orders, and the sign prefixed to the term being* + *for the natural order* 1 2 3 . . . m *and* − *or* + *for other orders according as the number of reversals of natural order is odd or even.*

The symbol for the determinant is constructed by placing single vertical lines before and after the set; thus

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

means the determinant $a_1 b_2 c_3 -$ etc., which we have called $D^{(3)}$.

The terminology is the same as is given in I. 4 and I. 5 for a set. The quantities between the vertical lines are the **elements** of the determinant. Those in a vertical line are a **column**; those in a horizontal line are a **row**. The **leading diagonal** is the diagonal drawn from the top left-hand corner; and the **leading term** is that containing the elements through which the leading diagonal passes. The leading term, as already stated, is taken positively.

If the symbol for a determinant contains m columns and m rows, the determinant is said to be of the mth **order**.

II. 2. Elementary properties.—(i) From the mode of construction it follows that each element of the determinant appears in $(m-1)!$ out of the $m!$ terms; and there are no two terms having more than $m-2$ factors alike.

(ii) If each element of a column or of a row is 0, the determinant is $=0$. [For each term contains one of these elements as a factor.]

II. 3. Properties depending on the rule of signs.—

(i) *The value of a determinant is not altered by making the columns rows and the rows columns*; e.g., for $m=4$,

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

[Let D be the determinant, and R the new determinant obtained by making columns rows and rows columns in the symbol for D. Then, apart from sign, D and R obviously have the same $m!$ terms. We have therefore only to consider signs. The two determinants have the same leading term, which is positive in both. Let t be any other term of D, say $a_3b_4c_2d_1\ldots$. Then t also occurs in R, but, since the terms of R must be constructed according to the system prescribed in our definition of a determinant, the factors of t in R will be arranged in the numerical order of the suffixes, namely $d_1c_2a_3b_4\ldots$. The sign of t in D depends on the number of reversals of order in the suffixes 3 4 2 1 . . ., and the sign of t in R depends on the number of reversals of order in the letters $d\,c\,a\,b\ldots$. But each of these numbers is the sum of the numbers of reversals of order of letters and of suffixes as compared with the original orders $a\,b\,c\,d\ldots$ and 1 2 3 4 . . .; and, by I. 5 (v), these sums are either both odd or both even. It follows that the sign of is the same in both determinants. This is true for each term of D or R; and the two determinants are therefore equal.]

If two determinants correspond so that the columns of one are the rows of the other, each determinant is said to be the **transposed** of the other

(ii) It follows from (i) that any statement as to columns or rows is equally true for rows or columns. We shall indicate this, for conciseness, by 'column [row]' or 'row [column]'.

(iii) *If any two columns [rows] of a determinant are interchanged, the absolute magnitude of the determinant remains unaltered, but its sign is changed.*

[Suppose, e. g., that we interchange the b's and the e's. Let ϕ and ψ be any two suffixes. Then, in the original determinant, corresponding to any term which contains b_ϕ and e_ψ, there is another term exactly similar except that the factors are b_ψ and e_ϕ; and these two terms, by I. 5 (iv), are of opposite sign. The effect of interchanging the b's and the e's is that the two terms are interchanged, i. e. the sign of each is changed. This applies to every such pair of terms.]

(iv) *If any two columns [rows] of a determinant are identical, the determinant is* $= 0$.

[We can see this in either of two ways.

(1) Consider a pair of terms such as are mentioned in (iii). The one contains b_ϕ and e_ψ; the other is exactly similar, except that it contains b_ψ and e_ϕ; and the two terms have opposite signs. If $b_\phi = e_\phi$ and $b_\psi = e_\psi$, the two terms cancel. The whole determinant is made up of such pairs.

(2) More briefly, suppose we interchange the two columns which are identical. Then the determinant remains unaltered. But, by (iii), its sign is changed. This can only be the case if the determinant is 0.]

(v) Since columns and rows may be interchanged, a determinant is sometimes represented by its leading diagonal alone, if this indicates a system for insertion of the remaining elements. The notation is

$$|a_1 b_2 c_3 \dots f_m| \equiv \begin{vmatrix} a_1 & b_1 & c_1 & \dots f_1 \\ a_2 & b_2 & c_2 & \dots f_2 \\ a_3 & b_3 & c_3 & \dots f_3 \\ \vdots & \vdots & \vdots & \vdots \\ a & b & c & f \end{vmatrix};$$

it is then immaterial, so far as the value of the determinant is concerned, whether we enter the a's as a column or as a row. But it should be mentioned that the relative arrangement of columns and rows is of importance later on, when we come to consider properties of sets of quantities.

II. 4. **Cofactors and minors.**—(i) In the complete expression for D, each term contains one a, which is either a_1 or a_2 or a_3 etc. We can therefore group the terms according to the a's they contain. In the case of $D^{(3)}$, for instance,

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1$$

$$= a_1(b_2c_3 - b_3c_2) + a_2(-b_1c_3 + b_3c_1) + a_3(b_1c_2 - b_2c_1).$$

Suppose that the terms of D are grouped in this way; and let the resulting coefficients of $a_1\, a_2\, a_3 \ldots a_m$ be denoted by $A_1\, A_2\, A_3 \ldots A_m$. Then

$$D = a_1A_1 + a_2A_2 + a_3A_3 + \ldots + a_mA_m.$$

Similarly, if we group the terms according to the b's or c's etc. they contain and denote the coefficients of the b's or c's etc. by $B_1\, B_2\, B_3 \ldots B_m$ or $C_1\, C_2\, C_3 \ldots C_m$ etc., we shall have

$$D = b_1B_1 + b_2B_2 + b_3B_3 + \ldots + b_mB_m,$$
$$D = c_1C_1 + c_2C_2 + c_3C_3 + \ldots + c_mC_m,$$
etc.

The A's, B's, etc., are called the **cofactors** of the corresponding elements of the determinant; thus the cofactor of b_3 is B_3, where b_3B_3 is the sum of all the terms which contain b_3.

(ii) The terms which contain the leading element a_1 are obtained from the leading term $a_1b_2c_3 \ldots f_m$ by altering the suffixes, and prefixing the proper sign to the term, in the manner already described, with the proviso that the factor a_1 remains unaltered. But this process will give us the products, by a_1, of the terms so constructed from a leading term $b_2 c_3 \ldots f_m$. In other words, the cofactor of a_1 is the determinant

$$\begin{vmatrix} b_2 & c_2 & \cdots f_2 \\ b_3 & c_3 & \cdots f_3 \\ \vdots & \vdots & \vdots \\ b_m & c_m & \cdots f_m \end{vmatrix}.$$

In $D^{(3)}$, for example, it is

$$b_2 c_3 - b_3 c_2 = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}.$$

(iii) The determinant which is obtained from D by striking out the column and the row which contain any element of the determinant is called the **minor** of that element in the determinant.

We see from (ii) that

$$A_1 \equiv \text{cofactor of } a_1 = \text{minor of } a_1.$$

We might show in the same way that the cofactor of any other element, say c_4, is equal to the minor of that element, with the sign $-$ or $+$ prefixed according to the position of the element in the determinant: but it is simpler to find the cofactor by bringing the element into the position of a_1. Let the element be in the qth column and the rth row. Then we can make it the leading element, without altering the order of the other columns or rows, by means of $q-1$ interchanges of its column with an adjoining column and $r-1$ interchanges of its row with an adjoining row. Each of these interchanges, by II. 3 (iii), multiplies the determinant by -1; and the total result is to multiply by -1 or by $+1$ according as $q+r-2$ is odd or even. Having got the element into the position of the leading element, we strike out the first column and the first row; the result, apart from the prefixed sign, is still to give the minor of the element, since the relative positions of the other columns and rows are unaltered. Hence *the cofactor of any element is equal to its minor with the sign − or + prefixed according as the number of steps from the leading element to this element is odd or even*; it being understood that each step is either horizontally from one column to the next or vertically from one row to the next. For example,

$$\begin{aligned} A_3 &= +\text{minor of } a_3, \\ C_4 &= -\text{minor of } c_4, \\ &\text{etc.} \end{aligned}$$

(iv) We have found in (i) that

$$\left.\begin{aligned} a_1A_1+a_2A_2+a_3A_3+\ldots+a_mA_m &= D \\ b_1B_1+b_2B_2+b_3B_3+\ldots+b_mB_m &= D \\ \text{etc.} \end{aligned}\right\}. \qquad \textbf{(II. 4. 1)}$$

We have now to find the value of such sums as

$$\begin{aligned} &a_1B_1+a_2B_2+a_3B_3+\ldots+a_mB_m, \\ &b_1A_1\,+b_2A_2\,+b_3A_3\,+\ldots+b_mA_m, \\ &c_1A_1\,+c_2A_2\,+c_3A_3\,+\ldots+c_mA_m, \\ &\text{etc.} \end{aligned}$$

Let us take the second and third of these as examples, but replace the b's or the c's by θ's. Then we want to find the value of

$$\theta_1A_1+\theta_2A_2+\theta_3A_3+\ldots+\theta_mA_m.$$

Now we see from (II. 4. 1) that this is the value of the determinant

$$\begin{vmatrix} \theta_1 & b_1 & c_1 & \cdots f_1 \\ \theta_2 & b_2 & c_2 & \cdots f_2 \\ \theta_3 & b_3 & c_3 & \cdots f_3 \\ \vdots & \vdots & \vdots & \vdots \\ \theta_m & b_m & c_m & \cdots f_m \end{vmatrix};$$

for the cofactors of $\theta_1\, \theta_2\, \theta_3 \ldots \theta_m$ in this determinant are the same as the cofactors of $a_1\, a_2\, a_3 \ldots a_m$ in D, i.e. are $A_1\, A_2\, A_3 \ldots A_m$. Let us replace θ throughout this determinant by any letter, other than a, occurring in D, e.g. by c. Then the determinant becomes

$$\begin{vmatrix} c_1 & b_1 & c_1 \ldots f_1 \\ c_2 & b_2 & c_2 \ldots f_2 \\ c_3 & b_3 & c_3 \ldots f_3 \\ \vdots & \vdots & \vdots \quad \vdots \\ c_m & b_m & c_m \ldots f_m \end{vmatrix}.$$

But this is a determinant which has two columns identical, and its value, by § 3 (iv), is 0. Hence

$$\left.\begin{array}{ll} & c_1 A_1 + c_2 A_2 + c_3 A_3 + \ldots + c_m A_m = 0 \\ \text{Similarly} & b_1 A_1 + b_2 A_2 + b_3 A_3 + \ldots + b_m A_m = 0 \\ & a_1 B_1 + a_2 B_2 + a_3 B_3 + \ldots + a_m B_m = 0 \\ & \text{etc.} \end{array}\right\} . \quad \textbf{(II. 4. 2)}$$

(v) Now let us interchange the columns with the rows, so that the determinant becomes

$$\begin{vmatrix} a_1 & a_2 & a_3 \ldots a_m \\ b_1 & b_2 & b_3 \ldots b_m \\ c_1 & c_2 & c_3 \ldots c_m \\ \vdots & \vdots & \vdots \quad \vdots \\ f_1 & f_2 & f_3 \ldots f_m \end{vmatrix}.$$

Then, by § 3 (i), the value of the new determinant is the same as that of the old, i.e. is D. Also the minor of any element ϵ in the new determinant is the same as its minor in the old determinant, but with columns and rows interchanged, so that its value is unaltered; and the number of steps from a_1 to ϵ is the same in both determinants. The cofactor of ϵ in the new determinant is therefore equal to its cofactor in the old determinant. Hence by applying

(II. 4. 1) and (II. 4. 2) to the new determinant we get new sets of relations, namely

$$\left.\begin{array}{l} a_1 A_1 + b_1 B_1 + c_1 C_1 + \ldots + f_1 F_1 = D \\ a_2 A_2 + b_2 B_2 + c_2 C_2 + \ldots + f_2 F_2 = D \\ \quad \text{etc.} \end{array}\right\} \qquad \textbf{(II. 4. 3)}$$

and

$$\left.\begin{array}{l} a_3 A_1 + b_3 B_1 + c_3 C_1 + \ldots + f_3 F_1 = 0 \\ a_2 A_1 + b_2 B_1 + c_2 C_1 + \ldots + f_2 F_1 = 0 \\ \quad \text{etc.} \end{array}\right\}. \qquad \textbf{(II. 4. 4)}$$

(vi) If all the elements in a column [row], except one, are 0, the determinant is equal to the product of that one by its cofactor.

III. SOLUTION OF SIMULTANEOUS EQUATIONS

III. 1. Statement of previous results.—We have next to consider the solution of the simultaneous equations (I. 3. A). Before we do this, it will be convenient to express in determinant form the results obtained in I. 2. These results are as follows:

(1) If $a_1 x = k_1$, then* $x = |k_1| \div |a_1|$.

(2) If
$$\left.\begin{aligned} a_1 x + b_1 y &= k_1 \\ a_2 x + b_2 y &= k_2 \end{aligned}\right\},$$
then
$$x = \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad y = \begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

(3) If
$$\left.\begin{aligned} a_1 x + b_1 y + c_1 z &= k_1 \\ a_2 x + b_2 y + c_2 z &= k_2 \\ a_3 x + b_3 y + c_3 z &= k_3 \end{aligned}\right\},$$
then
$$x = \begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix} \div D^{(3)}, \quad y = \begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix} \div D^{(3)},$$
$$z = \begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix} \div D^{(3)},$$
where
$$D^{(3)} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

* Here, as elsewhere, vertical lines denote a determinant, not 'absolute value'.

III. 2. **General solution.**—The general equations of which we require a solution are those set out in (I. 3. A), namely:

$$\left.\begin{aligned} a_1x + b_1y + c_1z + \dots + f_1w &= k_1 \\ a_2x + b_2y + c_2z + \dots + f_2w &= k_2 \\ &\vdots \\ a_mx + b_my + c_mz + \dots + f_mw &= k_m \end{aligned}\right\}. \qquad \textbf{(III. 2. A)}$$

The form of the solution is suggested by the results given above. Multiplying the successive equations by $A_1, A_2, \dots A_m$, and adding, we have

$$\begin{aligned} &(a_1A_1 + a_2A_2 + \dots + a_mA_m)x \\ &+ (b_1A_1 + b_2A_2 + \dots + b_mA_m)y \\ &+ (c_1A_1 + c_2A_2 + \dots + c_mA_m)z \\ &+ \dots \\ &+ (f_1A_1 + f_2A_2 + \dots + f_mA_m)w = k_1A_1 + k_2A_2 + \dots + k_mA_m. \end{aligned}$$

By (II. 4. 1) and (II. 4. 2) the coefficient of x is equal to D, and those of $y, z, \dots w$ are equal to 0. Also the expression on the right-hand side is what D would become if we replaced the a's by k's. Hence

$$\left.\begin{aligned} x &= (k_1A_1 + k_2A_2 + \dots + k_mA_m) \div D \\ &= \begin{vmatrix} k_1 & b_1 & c_1 & \dots f_1 \\ k_2 & b_2 & c_2 & \dots f_2 \\ \vdots & \vdots & \vdots & \vdots \\ k_m & b_m & c_m & \dots f_m \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 & c_1 & \dots f_1 \\ a_2 & b_2 & c_2 & \dots f_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_m & b_m & c_m & \dots f_m \end{vmatrix}. \\ &\text{Similarly} \\ y &= \begin{vmatrix} a_1 & k_1 & c_1 & \dots f_1 \\ a_2 & k_2 & c_2 & \dots f_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_m & k_m & c_m & \dots f_m \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 & c_1 & \dots f_1 \\ a_2 & b_2 & c_2 & \dots f_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_m & b_m & c_m & \dots f_m \end{vmatrix}, \\ &\text{and so on.} \end{aligned}\right\} \qquad \textbf{(III. 2. 1)}$$

If, for verification, we substitute these values in the original equations, it will be found that the relations (II. 4. 3) and (II. 4. 4) come into play.

IV. PROPERTIES OF DETERMINANTS
(*continued*)

IV. 1. **Sum of determinants.**—If two determinants are identical except as regards one column [row], their sum is a similar determinant in which the elements of that column [row] are the sums of corresponding elements in the two determinants. [For example

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+d_1 & b_1 & c_1 \\ a_2+d_2 & b_2 & c_2 \\ a_3+d_3 & b_3 & c_3 \end{vmatrix},$$

since it is $= (a_1A_1 + a_2A_2 + a_3A_3) + (d_1A_1 + d_2A_2 + d_3A_3)$
$= (a_1 + d_1)A_1 + (a_2 + d_2)A_2 + (a_3 + d_3)A_3$.]

IV. 2. **Multiplication of determinant by a single factor.**—If each element of a column [row] is multiplied by the same factor, the determinant is multiplied by that factor. [For example

$$\begin{vmatrix} \lambda a_1 & b_1 & c_1 \\ \lambda a_2 & b_2 & c_2 \\ \lambda a_3 & b_3 & c_3 \end{vmatrix} = \lambda a_1A_1 + \lambda a_2A_2 + \lambda a_3A_3 = \lambda \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.]$$

IV. 3. **Alteration of column or row.**—If the elements of a column [row] are multiplied by a single factor and added to the corresponding elements of another column [row], the value of the determinant is not altered. [For example

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

IV. 4. Calculation of determinant.—(i) There are two main methods for calculation of a numerical determinant.

(*a*) When m is small, we can use the formula

$$D = a_1 A_1 + a_2 A_2 + \ldots + a_m A_m,$$

repeating the process as often as may be necessary. Thus, if

$$D \equiv \begin{vmatrix} 3 & -2 & 7 \\ 5 & 1 & -3 \\ 4 & 6 & 1 \end{vmatrix},$$

then $D = 3\begin{vmatrix} 1 & -3 \\ 6 & 1 \end{vmatrix} - 5\begin{vmatrix} -2 & 7 \\ 6 & 1 \end{vmatrix} + 4\begin{vmatrix} -2 & 7 \\ 1 & -3 \end{vmatrix}$

$$= 3 . 19 - 5(-44) + 4(-1) = 273.$$

(*b*) When m is large, we can reduce the determinant to one of order $m-1$ by means of § 3 and II. 4 (vi). Applying this method to the above example, we could multiply the first row by $\frac{5}{3}$ and subtract from the second, and also multiply it by $\frac{4}{3}$ and subtract from the third. To avoid fractions, we multiply D twice by 3. Then

$$9D = \begin{vmatrix} 3 & -2 & 7 \\ 15 & 3 & -9 \\ 12 & 18 & 3 \end{vmatrix} = \begin{vmatrix} 3 & -2 & 7 \\ 0 & 13 & -44 \\ 0 & 26 & -25 \end{vmatrix}$$

$$= 3 \times \begin{vmatrix} 13 & -44 \\ 26 & -25 \end{vmatrix} = 3(-325 + 1144) = 3.819.$$

$\therefore \quad D = 273.$

(ii) For algebraical determinants various devices have to be used. An important determinant is

$$D \equiv \begin{vmatrix} a^{m-1} & b^{m-1} & c^{m-1} \ldots e^{m-1} & f^{m-1} \\ a^{m-2} & b^{m-2} & c^{m-2} \ldots e^{m-2} & f^{m-2} \\ \vdots & \vdots & \vdots \quad\quad \vdots & \vdots \\ a & b & c \quad \ldots e & f \\ 1 & 1 & 1 \quad \ldots 1 & 1 \end{vmatrix}.$$

This is $=0$ if $a=b$ or if $a=c$, etc. Hence it contains $a-b$, $a-c, \ldots a-e$, $a-f$ as factors. Similarly it contains $b-c, \ldots b-e$, $b-f$; and so on. Looking to the leading term, it will be seen that there can be no other factors; i.e.

$$D=(a-b)(a-c)\ldots(a-e)(a-f).(b-c)\ldots(b-e)(b-f)\ldots(e-f).$$

[Example.—Hence prove that

$$\begin{vmatrix} b^{m-1} & c^{m-1} \ldots\ldots & e^{m-1} & f^{m-1} \\ b^{m-2} & c^{m-2} \ldots\ldots & e^{m-2} & f^{m-2} \\ \vdots & \vdots & \vdots & \vdots \\ b^{r+1} & c^{r+1} \ldots\ldots & e^{r+1} & f^{r+1} \\ b^{r-1} & c^{r-1} \ldots\ldots & e^{r-1} & f^{r-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 \ldots\ldots & 1 & 1 \end{vmatrix} = (-)^{m-r+1} \frac{D}{\phi(a)} \times \text{coefficient of } a^r \text{ in } \phi(a),$$

where

$$\phi(a) \equiv (a-b)(a-c)\ldots(a-e)(a-f).]$$

IV. 5. Product of determinants.—(i) The product of a determinant of order m and a determinant of order n can be expressed as a determinant of order $m+n$ by placing the leading diagonals in line and filling in with 0's. For example

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} d_4 & e_4 \\ d_5 & e_5 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & 0 \\ a_3 & b_3 & c_3 & 0 & 0 \\ 0 & 0 & 0 & d_4 & e_4 \\ 0 & 0 & 0 & d_5 & e_5 \end{vmatrix}.$$

[For the only terms of this latter determinant which are not 0 are those for which the first three factors (collectively) are taken from the first three columns and rows and the next two factors (collectively) from the last two columns and rows; and in each such case the first three factors form a term of the first determinant and the next two factors form a term of the second determinant. Thus all the terms of the product of the one expanded determinant by the other are accounted for, and there are no others.]

(ii) We have now to show that the product of two determinants, each of order m, can be expressed as a determinant of order m. To obtain the general formula, it will be sufficient to take a particular case, e.g.

$$D \equiv \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \quad E \equiv \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix},$$

provided that in our reasoning we retain m as the order of each determinant.

By means of the first sentence of (i) we can write down DE as a determinant of order $2m$, i.e.

$$DE = \begin{vmatrix} a_1 & b_1 & c_1 & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & 0 & 0 \\ a_3 & b_3 & c_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_1 & \beta_1 & \gamma_1 \\ 0 & 0 & 0 & \alpha_2 & \beta_2 & \gamma_2 \\ 0 & 0 & 0 & \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix}.$$

Two of the quarters of this determinant contain 0's only; and it will be seen, from the method of forming those terms of the determinant which do not contain 0 as a factor, that we can fill in either of these quarters in any way we like, provided we leave the other quarter alone. Also, in order to reduce the determinant from order $2m$ to order m, we ought to get m 1's in the leading diagonal. We therefore shift the last m columns to be the first m, and then replace the 0's in the new leading diagonal by 1's. The first process involves m^2 interchanges: we can avoid change of sign of the determinant, in the case where m is odd, by changing the signs of the first m rows (before inserting the 1's), whether m is even or odd. Then, inserting the 1's, we get

$$DE = \begin{vmatrix} 1 & 0 & 0 & -a_1 & -b_1 & -c_1 \\ 0 & 1 & 0 & -a_2 & -b_2 & -c_2 \\ 0 & 0 & 1 & -a_3 & -b_3 & -c_3 \\ \alpha_1 & \beta_1 & \gamma_1 & 0 & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 & 0 & 0 & 0 \\ \alpha_3 & \beta_3 & \gamma_3 & 0 & 0 & 0 \end{vmatrix}.$$

We now reduce each of the elements in the right-hand top quarter to 0 by means of § 3; i.e. we add a_1 times the 1st column to the $(m+1)$th (in this case the 4th), thus getting

$$\begin{vmatrix} 1 & 0 & 0 & 0 & -b_1 & -c_1 \\ 0 & 1 & 0 & -a_2 & -b_2 & -c_2 \\ 0 & 0 & 1 & -a_3 & -b_3 & -c_3 \\ \alpha_1 & \beta_1 & \gamma_1 & a_1\alpha_1 & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 & a_1\alpha_2 & 0 & 0 \\ \alpha_3 & \beta_3 & \gamma_3 & a_1\alpha_3 & 0 & 0 \end{vmatrix},$$

then do the same with a_2 times the 2nd column and a_3 times the 3rd..., and then deal in the same way with the $(m+2)$th and $(m+3)$th... columns. We get finally

$$DE = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \alpha_1 & \beta_1 & \gamma_1 & a_1\alpha_1+a_2\beta_1+a_3\gamma_1 & b_1\alpha_1+b_2\beta_1+b_3\gamma_1 & c_1\alpha_1+c_2\beta_1+c_3\gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 & a_1\alpha_2+a_2\beta_2+a_3\gamma_2 & b_1\alpha_2+b_2\beta_2+b_3\gamma_2 & c_1\alpha_2+c_2\beta_2+c_3\gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 & a_1\alpha_3+a_2\beta_3+a_3\gamma_3 & b_1\alpha_3+b_2\beta_3+b_3\gamma_3 & c_1\alpha_3+c_2\beta_3+c_3\gamma_3 \end{vmatrix}$$

By the same reasoning as that employed at the beginning of this paragraph, we can replace each of the elements in the left-hand bottom quarter of the above determinant by 0. Hence, if the determinant formed by the elements in the lower right-hand quarter of the above is called P, we have, by (i),

$$DE = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \times P = P;$$

i.e.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1\alpha_1 + a_2\beta_1 + a_3\gamma_1 & b_1\alpha_1 + b_2\beta_1 + b_3\gamma_1 & c_1\alpha_1 + c_2\beta_1 + c_3\gamma_1 \\ a_1\alpha_2 + a_2\beta_2 + a_3\gamma_2 & b_1\alpha_2 + b_2\beta_2 + b_3\gamma_2 & c_1\alpha_2 + c_2\beta_2 + c_3\gamma_2 \\ a_1\alpha_3 + a_2\beta_3 + a_3\gamma_3 & b_1\alpha_3 + b_2\beta_3 + b_3\gamma_3 & c_1\alpha_3 + c_2\beta_3 + c_3\gamma_3 \end{vmatrix}. \quad \textbf{(IV.5.1)}$$

The reasoning is quite general, and the product of two determinants of any order can be written down from the above.

By interchanging columns and rows in one or other or both of the original determinants we get three other expressions. All four expressions, of course, are equal when expanded: we shall take the above to be the standard form. It is to be noticed that DE or $D \times E$ means the product of D and E, in this order; by reversing the order of multiplication we get four other forms, but these are only the 'transposed' of the previous four. The eight forms are given, in the new notation (see V. 6 (vi)), in the Appendix (p. 122). They are based on the principle that in the standard form the element in the qth column and rth row of the product is formed in a particular way from the qth column of the first determinant and the rth row of the second.

IV. 6. The adjoint determinant.—(i) Let D' denote the determinant whose elements are the cofactors of the corresponding elements in the original determinant, i.e.

$$D' \equiv \begin{vmatrix} A_1 & B_1 & C_1 \dots F_1 \\ A_2 & B_2 & C_2 \dots F_2 \\ \vdots & \vdots & \vdots \quad \vdots \\ A_m & B_m & C_m \dots F_m \end{vmatrix}. \qquad \textbf{(IV. 6. A)}$$

If we interchange columns and rows in D', and then express the product DD' as in (IV. 5. 1), we find that the elements in the principal diagonal of the product are $a_1A_1 + a_2A_2 + \dots a_m A_m$, $b_1B_1 + b_2B_2 + \dots + b_mB_m$, etc., each of which, by (II. 4. 1), is $= D$, while the other elements are $a_1B_1 + a_2B_2 + \dots + a_mB_m$, $b_1A_1 + b_2A_2 + \dots + b_mA_m$, etc., each of which, by (II. 4. 2), is $= 0$. Hence

$$DD' = \begin{vmatrix} D & 0 & 0 \dots 0 \\ 0 & D & 0 \dots 0 \\ \vdots & \vdots & \vdots \quad \vdots \\ 0 & 0 & 0 \dots D \end{vmatrix} = D^m, \qquad \textbf{(IV. 6. 1)}$$

and therefore

$$D' = D^{m-1}. \qquad \textbf{(IV. 6. 2)}$$

The determinant D' was formerly called the *reciprocal*, but is now more usually called the **adjoint**, of D. It is not the true reciprocal of D, since the product of the two is not 1 but D^m (cf. V. 2).

(ii) Let the cofactors of $A_1, B_1, C_1, \dots F_1$ in D' be denoted by $\alpha_1, \beta_1, \gamma_1, \dots \zeta_1$. Then, applying (II. 4. 3) and (II. 4. 4) to D', we have

$$\left.\begin{aligned} A_1\alpha_1 + B_1\beta_1 + C_1\gamma_1 + \dots + F_1\zeta_1 &= D' \\ A_2\alpha_1 + B_2\beta_1 + C_2\gamma_1 + \dots + F_2\zeta_1 &= 0 \\ &\;\vdots \\ A_m\alpha_1 + B_m\beta_1 + C_m\gamma_1 + \dots + F_m\zeta_1 &= 0 \end{aligned}\right\}.$$

We can regard these as equations for determining $\alpha_1, \beta_1, \gamma_1, \dots \zeta_1$. Comparing them with

$$\left.\begin{aligned} A_1a_1 + B_1b_1 + C_1c_1 + \dots + F_1f_1 &= D \\ A_2a_1 + B_2b_1 + C_2c_1 + \dots + F_2f_1 &= 0 \\ &\;\vdots \\ A_ma_1 + B_mb_1 + C_mc_1 + \dots + F_mf_1 &= 0 \end{aligned}\right\},$$

which are obtained from (II. 4. 3) and (II. 4. 4), we see that the solution is

$$a_1 = \frac{D'}{D} a_1 = D^{m-2} a_1, \quad \beta_1 = \frac{D'}{D} b_1 = D^{m-2} b_1, \quad \text{etc.}$$

Hence *the cofactor of any element of the adjoint determinant is equal to the corresponding element of the original determinant, multiplied by the ratio of the adjoint determinant to the original determinant.*

V. THE TENSOR NOTATION *

V. 1. **Main properties of determinant.**—(i) The notation so far used is the ordinary one for elementary work. For higher work we reduce the number of letters and make a more liberal use of suffixes. We shall replace $x, y, z \ldots w$ by $X_1, X_2, X_3, \ldots X_m$; $a_1, b_1, c_1, \ldots f_1$ by $d_{11}, d_{21}, d_{31}, \ldots d_{m1}$; and so on. Also, it being understood that the values assignable to each of the letters q and r are 1, 2, 3, ... m, we can use $|d_{qr}|$ and $|d_{rq}|$ to denote the determinants whose elements in the qth column and rth row are respectively † d_{qr} and d_{rq}, i.e.

$$\left.\begin{aligned} |d_{qr}| &\equiv \begin{vmatrix} d_{11} & d_{21} & d_{31} \ldots d_{m1} \\ d_{12} & d_{22} & d_{32} \ldots d_{m2} \\ \vdots & \vdots & \vdots \quad\; \vdots \\ d_{1m} & d_{2m} & d_{3m} \ldots d_{mm} \end{vmatrix} \\ |d_{rq}| &\equiv \begin{vmatrix} d_{11} & d_{12} & d_{13} \ldots d_{1m} \\ d_{21} & d_{22} & d_{23} \ldots d_{2m} \\ \vdots & \vdots & \vdots \quad\; \vdots \\ d_{m1} & d_{m2} & d_{m3} \ldots d_{mm} \end{vmatrix} \end{aligned}\right\}, \qquad \textbf{(V. 1. A)}$$

so that the interchangeability of columns and rows gives

$$|d_{qr}| = |d_{rq}|. \qquad \textbf{(V. 1. 1)}$$

(ii) The k's in (III. 2. A) were supposed to be known

* The paragraphs in small print may be omitted on first reading, but should be read before Chapter VI is taken.

† It is more usual to have the suffixes the other way round; i.e. to use 11, 12, 13,... $1m$ as suffixes for the first row. But my arrangement seems to follow more naturally from the $a_1 b_1 \ldots$ notation, and I also find that it fits in better with the subsequent work.

quantities, so that the equations served to determine $x, y, z, \ldots w$. We shall have to consider the relations between the set of quantities which we have denoted by $x, y, z, \ldots w$ and those which we have denoted by $k_1, k_2, k_3, \ldots k_m$. We therefore, in altering $x, y, z, \ldots w$ to $X_1, X_2, X_3, \ldots X_m$, also alter $k_1, k_2, k_3, \ldots k_m$ to $Y_1, Y_2, Y_3, \ldots Y_m$.

(iii) The main properties which we have to consider are set out below. Definitions are marked with capital letters: propositions with arabic numbers.

$$D \equiv | d_{qr} | \equiv \begin{vmatrix} d_{11} & d_{21} & d_{31} \ldots d_{m1} \\ d_{12} & d_{22} & d_{32} \ldots d_{m2} \\ \vdots & \vdots & \vdots \quad\;\; \vdots \\ d_{1m} & d_{2m} & d_{3m} \ldots d_{mm} \end{vmatrix} \quad . \quad \text{(A)}$$

$$D_{ps} \equiv \text{cofactor of } d_{ps} \text{ in } D \quad . \quad . \quad . \quad \text{(B)}$$

$$\binom{p = 1, 2, \ldots m;}{q = 1, 2, \ldots m} \; D_{p1} d_{q1} + D_{p2} d_{q2} + \ldots + D_{pm} d_{qm} = \begin{cases} D \text{ if } q = p \\ 0 \text{ if } q \neq p \end{cases} \quad \text{(1)}$$

$$D' \equiv | D_{qr} | \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(C)}$$

$$DD' = D^m \quad . \quad . \quad . \quad . \quad . \quad \text{(2)}$$

$$\text{cofactor of } D_{ps} \text{ in } D' = D^{m-2} d_{ps} \quad . \quad . \quad \text{(3)}$$

If
$$\left.\begin{aligned} d_{11}X_1 + d_{21}X_2 + d_{31}X_3 + \ldots + d_{m1}X_m &= Y_1 \\ d_{12}X_1 + d_{22}X_2 + d_{32}X_3 + \ldots + d_{m2}X_m &= Y_2 \\ &\;\vdots \\ d_{1m}X_1 + d_{2m}X_2 + d_{3m}X_3 + \ldots + d_{mm}X_m &= Y_m \end{aligned}\right\},$$
then
$$(p = 1, 2, \ldots m)$$
$$X_p = \frac{D_{p1}Y_1 + D_{p2}Y_2 + D_{p3}Y_3 + \ldots + D_{pm}Y_m}{D}. \quad \text{(4)}$$

(iv) As a preliminary, to make the statements more concise, we may—though this is not essential—introduce the ordinary notation of summation. Thus (1) can be written

$$\begin{pmatrix} p = 1, 2, \dots m; \\ q = 1, 2, \dots m \end{pmatrix} \sum_{s=1}^{m} D_{ps} d_{qs} = \begin{Bmatrix} D \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}; \quad . \quad (1)$$

and (4) can be written

$$\left.\begin{aligned} &\text{If } (s = 1, 2, \dots m) \sum_{p=1}^{m} d_{ps} X_p = Y_s, \\ &\text{then } (p = 1, 2, \dots m) X_p = \sum_{s=1}^{m} \frac{D_{ps} Y_s}{D}. \end{aligned}\right\} \quad . \quad . \quad (4)$$

(v) The tensor notation involves five steps, which are set out in §§ 2–4, 6, 8 below. The reader will find it helpful to copy out the statement in (iii), modified as in (iv), and make the successive alterations which are now to be described.

V. 2. **Reciprocal determinant.**—The first step (which will be found in modern text-books) relates to (1) and (2). The determinant D' has sometimes been called the reciprocal of D. But, as has already been pointed out in IV. 6 (i), it is not a true reciprocal, since the product of the two determinants is not 1 but D^m. Since, however, D' contains m columns, we see that if we form a new determinant by dividing each element of it by D the product of this new determinant and D will be 1. We therefore write (1) in the form

$$\begin{pmatrix} p = 1, 2, \dots m; \\ q = 1, 2, \dots m \end{pmatrix} \sum_{s=1}^{m} \frac{D_{ps}}{D} d_{qs} = \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}; \quad . \quad (1)$$

and form a new determinant D'' defined by

$$D'' \equiv | D_{qr}/D |. \quad . \quad . \quad . \quad . \quad . \quad . \quad (C)$$

This new determinant will be equal to $D'/D^m = 1/D$, so that the product of the two determinants will be **1**; and the cofactor of D_{ps}/D in the new determinant will be $d_{ps}/D = d_{ps}D''$. Hence the two determinants D and D'' are so related that their product is 1 and that the cofactor of any element of either determinant, divided by the determinant, is equal to the corresponding element of the other determinant. We can therefore call each determinant the **reciprocal** of the other.

V. 3. Elements of reciprocal determinant.—The next step is to have a single symbol for

$$(\text{cofactor of } d_{ps} \text{ in } D) \div D.$$

We have already used D_{ps} for the cofactor of d_{ps}; and it is inconvenient to introduce a new letter in place of d or D. We therefore denote the above expression by

$$d^{ps}.$$

We accordingly, in our statement, replace (B), (1), (C), (2), (3), and the second line of (4), by

$$d^{ps} \equiv (\text{cofactor of } d_{ps} \text{ in } D) \div D \quad . \quad . \quad . \quad (\text{B})$$

$$\begin{pmatrix} p = 1, 2, \ldots m; \\ q = 1, 2, \ldots m \end{pmatrix} \sum_{s=1}^{m} d^{ps} d_{qs} = \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix} \quad . \quad . \quad (1)$$

$$D'' \equiv | d^{qr} | \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{C})$$

$$DD'' = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$$d_{ps} = (\text{cofactor of } d^{ps} \text{ in } D'') \div D'' \quad . \quad . \quad (3)$$

$$(p = 1, 2, \ldots m)\ X_p = \sum_{s=1}^{m} d^{ps} Y_s \quad . \quad [\text{2nd line of}] \quad (4)$$

The parallelism of (B) and (3), and of the two lines of (4), should be noted.

V. 4. Set-notation.—(i) The next step is to abbreviate (4), as altered. This contains two statements—a hypothesis and a conclusion; they are similar in form, so that we need only consider the first one, namely

$$(s = 1, 2, \dots m) \sum_{p=1}^{m} d_{ps} X_p = Y_s.$$

(ii) The expression on the left-hand side of this statement has a definite value for each value of s; we can denote these values by

$$E_1\ E_2\ E_3 \dots E_m.$$

The statement then takes the form

$$(s = 1, 2, \dots m) E_s = Y_s.$$

We cannot merely omit the '$(s = 1, 2, \dots m)$' from this, without leaving it doubtful whether we are speaking of some particular s or of each s. We get over the difficulty by omitting the '$(s = 1, 2, \dots m)$' and replacing the s in '$E_s = Y_s$' by a *Greek* letter. The convention then is that a statement of the form

$$E_\sigma = Y_\sigma$$

means that E_s is equal to Y_s for each of the values of s, i.e. that

$$E_1 = Y_1,\ E_2 = Y_2,\ E_3 = Y_3, \dots E_m = Y_m;$$

it being understood that the values 1, 2, 3, ... m which are to be given to a Greek letter have been settled beforehand and remain the same throughout our work.

(iii) Applying this convention to the two statements in (4), it becomes:

$$\text{If } \sum_{p=1}^{m} d_{p\sigma} X_p = Y_\sigma, \text{ then } X_\lambda = \sum_{s=1}^{m} d^{\lambda s} Y_s \quad . \quad . \quad . \quad (4)$$

It is immaterial what Greek letter we use in either of these statements, provided the letter is the same on both sides. We could have used the same letter in the two statements, but in the particular case it is better to have the letters different.*

We have rather spoilt the symmetry of (4), but we will put this right in § 6.

(iv) The statement in (1) is a statement as to the value of a certain expression for all values of p and all values of q; and, so far as the left-hand side is concerned, we could extend the above principle by using two Greek letters. But the right-hand side presents difficulties; and we must therefore leave this over for later consideration (§ 8).

V. 5. **Principles of set-notation.**—It is desirable at this stage to consider the principles underlying the notation which we have just adopted.

(i) Take first the case of a **single set** of m quantities or **elements**; i. e. an aggregate of m quantities which fall into a certain linear arrangement. We denote these by, say,

$$A_1 \ A_2 \ A_3 \ldots A_m.$$

We have settled that a statement such as

$$A_\lambda = E_\lambda$$

is a comprehensive way of saying that

$$A_1 = E_1, A_2 = E_2, A_3 = E_3, \ldots A_m = E_m.$$

Thus we use A_p etc. when we are referring to a particular member of the set, and we use A_λ etc. when we are making a statement with regard to each member of the set in turn. We may also want to speak of the set as a whole. It will be found that no confusion arises from using A_λ in this sense also. We can therefore say that

$$A_\lambda \equiv (A_1 \ A_2 \ A_3 \ldots A_m);$$

the brackets being used in order to show that we are considering the set as a whole. In this sense we might regard the statement $A_\lambda = E_\lambda$ as meaning that the set A_λ as a whole is equal to the set

* I have as far as possible used λ, μ, ρ, σ in this chapter to correspond to p, q, r, s, reserving ν for product-sums (§ 6). Later on it is better to have no fixed rule, beyond that laid down in § 5 (iii).

E_λ as a whole, this equality of the wholes implying equality of the parts. We shall have to take this step later (Chapter VI): for the present it will be sufficient to regard the statement $A_\lambda = E_\lambda$ as merely an abbreviated way of saying that $A_1 = E_1$, $A_2 = E_2$, etc.

(ii) The interpretation of A_λ as the set as a whole recalls the idea of a vector as the resultant of a number of components. But the operations which we shall have to perform with single sets do not follow exactly the same laws as those which govern operations with vectors as ordinarily understood (see note to (vi) below), so that the analogy must not be pushed too far.

(iii) Next take the case of a **double set** of m^2 quantities, i.e. a set consisting of m single sets, each containing m elements. If we denote the elements of the qth single set by $F_{q1}, F_{q2}, \ldots F_{qm}$, this single set as a whole can be called $F_{q\rho}$, and the complete double set can be called $F_{\mu\rho}$. We think of the elements as arranged in a square, the columns of which are the single sets: by regrouping (e.g. $F_{1r}, F_{2r}, \ldots F_{mr}$) we get the rows of the square. We have already, in § 1, adopted the convention that in d_{qr} or d_{rq} the q represents the column and the r the row; and similarly we shall say that in $F_{\mu\rho}$ or $F_{\rho\mu}$ etc. the first letter, according to alphabetical order, means the column and the second the row. Hence

$$F_{\mu\rho} \equiv \begin{pmatrix} F_{11} & F_{21} & F_{31} \ldots F_{m1} \\ F_{12} & F_{22} & F_{32} \ldots F_{m2} \\ \vdots & \vdots & \vdots \quad\quad \vdots \\ F_{1m} & F_{2m} & F_{3m} \ldots F_{mm} \end{pmatrix}, \quad F_{\rho\mu} \equiv \begin{pmatrix} F_{11} & F_{12} & F_{13} \ldots F_{1m} \\ F_{21} & F_{22} & F_{23} \ldots F_{2m} \\ \vdots & \vdots & \vdots \quad\quad \vdots \\ F_{m1} & F_{m2} & F_{m3} \ldots F_{mm} \end{pmatrix},$$

the brackets being inserted, as before, in order to show that the set of quantities is in each case to be regarded as a whole. Then the statements

$$F_{\mu\rho} = G_{\mu\rho}, \quad F_{\mu\rho} = H_{\rho\mu},$$

mean respectively that $F_{qr} = G_{qr}$, and that $F_{qr} = H_{rq}$, for every value of q taken with every value of r.

A double set is **symmetrical** if it is not altered by interchanging columns and rows.

(iv) A particular form of double set is obtained by multiplying together every element of one single set and every element of another single set (of the same number of elements). If these two sets are B_μ and C_ρ (in this order), the representative element of the resulting double set will be $B_q\, C_r$, so that the double set can be represented by $B_\mu\, C_\rho$ or $C_\rho\, B_\mu$. This double set is called the **product** of the two single sets. It should be noted that we must not write it as '$B_\mu\, C_\mu$' or as '$B_\rho\, C_\rho$'; partly because this would not define the particular arrangement of the elements of the double set, and partly because we shall presently have to give a special meaning to these latter expressions.

(v) In addition to double sets and single sets, we have to use single quantities, such as a or k. Any such quantity is called a **scalar**. It need not be a constant: it may, as will be seen later, be a definite function of the elements of one or more sets.

(vi) We shall for the present be dealing only with expressions which, interpreted according to the laws of ordinary algebra, are obtained from scalars, single sets, and double sets by addition, subtraction, and multiplication. The rule of interpretation is the same that we adopted in (iv) for $B_\mu\, C_\rho$: we replace the Greek letters $\lambda\ \mu\ \rho \ldots$ by $p\ q\ r \ldots$ and take the total expression to be the set obtained by giving to each of the quantities $p\ q\ r \ldots$ each of the values $1\ 2\ 3 \ldots m$. For example:

(*a*) kA_ρ means the single set whose elements are

$$kA_1,\ kA_2, \ldots kA_m;$$

(*b*) $A_\lambda \pm bB_\lambda$ means the single set whose elements are *

$$A_1 \pm bB_1,\ A_2 \pm bB_2, \ldots A_m \pm bB_m;$$

(*c*) $A_{\mu\rho} - aF_\rho\, G_\mu$ means the double set whose element in the qth column and the rth row is $A_{qr} - aF_r\, G_q$.

It is obvious that this system of interpretation is in accordance with the laws of ordinary algebra; for instance

$$k(A_\mu + B_\mu) = kA_\mu + kB_\mu = kB_\mu + kA_\mu,$$
$$F_\lambda (G_\rho - H_\rho) = F_\lambda\, G_\rho - F_\lambda H_\rho,$$

and so on.

We are further restricted, in the case of expressions containing

* It will be seen from (vi) (*a*) and (*b*) and from (iv) that single sets follow the same rule as ordinary vectors as regards multiplication by a scalar, addition, and subtraction, but not as regards multiplication together.

more than one term, to (1) scalar expressions, (2) single sets arising as sums ('sum' of course including 'difference') of expressions which contain the same letter, e.g. $aA_\lambda + bB_\lambda + \ldots$, (3) double sets arising as sums of expressions which contain the same pair of letters, e.g. $aA_{\mu\nu} + bB_\nu C_\mu$. We do not therefore have to consider such an expression as $A_\lambda + B_\sigma$, which is really a double set, not a single set.

(vii) The suffixes which we have so far attached to a symbol have usually, in accordance with the regular practice in algebra and with the ordinary meaning of the word, been placed below the line: the exception being the use of d^{ps} to mean (cofactor of d_{ps} in D) $\div D$. This latter system of having upper suffixes as well as lower suffixes will sometimes be found convenient. We may, for instance, want to denote a single set by A^λ; and in that case $A^1, A^2, \ldots$ would be members of the set. Where there is any risk of confusion, we shall not use the ordinary indices of algebra at all; thus the square of A_p will be $A_p A_p$, not $A_p{}^2$.

V. 6. **Product-sum notation.**—(i) Our next simplification consists in dropping the sign of summation in (1) and (4). But, since merely to drop it and to replace, say, $\sum_{s=1}^{m} d^{ps} d_{qs}$ by $d^{ps} d_{qs}$ would be misleading, we use a special notation. The number of alphabets at our disposal is limited: and it will be found not only that we can use Greek letters for this purpose without risk of error, but that there are actual advantages in doing so.

(ii) The rule we adopt is that, when an expression of the form $B_p C_p$ has to be summed for the values $1, 2, \ldots m$ of p, we denote the result by replacing p by a Greek letter in both places; and, conversely, the meaning of such an expression as $B_\nu C_\nu$ is

$$B_\nu C_\nu \equiv B_1 C_1 + B_2 C_2 + \ldots + B_m C_m. \qquad \textbf{(V. 6. A)}$$

(iii) The pair of ν's in $B_\nu C_\nu$ could be replaced by a pair of any other identical Greek letters; e.g.

$$B_\nu C_\nu = B_\rho C_\rho. \qquad \text{(V. 6. 1)}$$

The ν (or ρ) is for this reason called a **dummy**. We can think of the sum represented by $B_\nu C_\nu$ as the result of linking the elements of B_ν with the corresponding elements of C_ν; we can therefore describe a Greek letter which occurs twice as a **linked suffix**, and one which occurs once only as a **free suffix.**

(iv) The rule in (ii) applies if either or both of the expressions B_p and C_p has a free suffix as well as the p; e.g. $A_\nu B_{\mu\nu}$ means $A_1 B_{\mu 1} + A_2 B_{\mu 2} + \dots + A_m B_{\mu m}$, and $A_{\lambda\nu} B_{\rho\nu}$ means $A_{\lambda 1} B_{\rho 1} + A_{\lambda 2} B_{\rho 2} + \dots + A_{\lambda m} B_{\rho m}$, which is a double set whose typical element is $A_{q\nu} B_{r\nu}$.

(v) We can also have successive summations expressed in the same way. Thus

$$B_\lambda C_{\lambda\mu} D_{\mu\rho} E_{\sigma\rho}$$

involves summations with regard to λ, with regard to μ, and with regard to ρ. It is easy to show that these summations can be made in any order: e.g. we can take $C_{\lambda\mu}$ and $D_{\mu\rho}$ together as if the λ and ρ were free, and then bring in B_λ and $E_{\sigma\rho}$.

(vi) As an example of the brevity effected by this notation we may take the expression for the product of determinants. Even for so small a value of m as 3, the expression obtained in IV. 5 (ii) for the product of two determinants is formidable. We can condense it by introducing Σ's, but the result is clumsy. In the new notation it will be found that the method of IV. 5 (ii) gives

$$|a_{qr}| \times |b_{qr}| = |a_{q\lambda} b_{\lambda r}|. \qquad \text{(V. 6. 2)}$$

The process can be repeated: e.g.

$$| a_{qr} | \times | b_{qr} | \times | c_{qr} | \times | d_{qr} | = | a_{q\lambda} b_{\lambda\mu} c_{\mu\nu} d_{\nu r} |. \qquad \text{(V. 6. 3)}$$

(vii) The result of applying the product-sum notation to the statement in § 1 (iii) is that (1) and (4) become (p and s in (4) being replaced by λ and σ respectively)

$$\binom{p = 1, 2, \dots m;}{q = 1, 2, \dots m} d^{p\sigma} d_{q\sigma} = \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}. \quad . \quad (1)$$

$$\text{If } d_{\lambda\sigma} X_\lambda = Y_\sigma, \text{ then } X_\lambda = d^{\lambda\sigma} Y_\sigma. \quad . \quad . \quad (4)$$

V. 7. **Inner products of sets.**—(i) The quantity $B_\nu C_\nu$ behaves in many respects like an algebraical product. We call it the **inner product** of B_ν and C_ν, to distinguish it from an ordinary or **outer** product such as $B_\lambda C_\rho$ (§ 5 (iv)). The inner product of B_ν and C_ν is the sum of the elements in the leading diagonal of the outer product of B_λ and C_ρ.

(ii) In the same way $A_\nu B_{\mu\nu}$ is the inner product of A_ν and $B_{\mu\nu}$, and $A_{\lambda\nu} B_{\nu\rho}$ is the inner product of $A_{\lambda\nu}$ and $B_{\nu\rho}$.

(iii) The process of forming an inner product, as above, may be called **inner multiplication.**

V. 8. **Unit-set notation.**—(i) We have finally to consider the form of (1), which is a statement that

$$\binom{p = 1, 2, \dots m;}{q = 1, 2, \dots m} d^{p\sigma} d_{q\sigma} = \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}.$$

So far as the left-hand side is concerned, this is a statement as to the values of the elements of a double set

$$d^{\lambda\sigma} d_{\mu\sigma}.$$

As regards the right-hand side, however, the statement falls into two; first that $d^{p\sigma} d_{p\sigma} = 1$, and next that $d^{p\sigma} d_{q\sigma} = 0$ if p and q are different. We want to replace these by a single statement.

(ii) We do this by converting the statement into one as to the equality of two double sets. For this purpose we construct a set whose typical element, in the qth column and rth row, is 1 if q and r are the same and 0 if they are different; a set, in other words, which has 1 for each element of its leading diagonal and 0 everywhere else. If we call this set *

$$|^{\lambda}_{\mu},$$

then our definition of $|^{p}_{q}$ is that

$$|^{p}_{q} \equiv \text{the function of } p \text{ and } q \text{ which}$$

$$\text{is} = \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}. \qquad \text{(V. 8. A)}$$

We can therefore write (1) in the form

$$\begin{pmatrix} p = 1, 2, \ldots m; \\ q = 1, 2, \ldots m \end{pmatrix} d^{p\sigma} d_{q\sigma} = |^{p}_{q};$$

or, in the set-notation,

$$d^{\lambda\sigma} d_{\mu\sigma} = |^{\lambda}_{\mu}. \qquad \ldots \ldots \qquad (1)$$

V. 9. **Properties of the unit set.**—(i) We have defined $|^{\lambda}_{\mu}$ as the set whose typical element is

$$|^{p}_{q} \equiv \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}. \qquad \text{(V. 9. A)}$$

Hence each element of the leading diagonal of $|^{\lambda}_{\mu}$ is 1, and the other elements are all 0; in other words

$$|^{\lambda}_{\mu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 & 0 \ldots 0 \\ 0 & 0 & 1 & 0 \ldots 0 \\ \vdots & \vdots & \vdots & \vdots \quad \vdots \\ 0 & 0 & 0 & 0 \ldots 1 \end{pmatrix}. \qquad \text{(V. 9. B)}$$

* The usual symbol, adopted by Einstein, is δ^{λ}_{μ}; J. E. Wright ('Invariants of quadratic differential forms') uses $\eta_{\lambda\mu}$. Neither of these seems sufficiently distinctive; and δ already has a considerable number of other uses. I have therefore altered the symbol to $|^{\lambda}_{\mu}$ ('unit $\lambda\mu$'), as an experiment.

This set (with any pair of letters) will be called the **unit set.** The following are its chief properties.

(ii) The set is symmetrical, i. e.

$$|^{\lambda}_{\mu} = |^{\mu}_{\lambda}. \qquad \text{(V. 9. 1)}$$

(iii) The determinant of the set is

$$||^{q}_{r}| = \begin{vmatrix} 1 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 \ldots 0 \\ 0 & 0 & 1 \ldots 0 \\ \vdots & \vdots & \vdots \quad \vdots \\ 0 & 0 & 0 \ldots 1 \end{vmatrix} = 1. \qquad \text{(V. 9. 2)}$$

(iv) Also, if we multiply together this determinant and any other determinant, in either order, it will be found that we merely reproduce the latter, i. e.

$$||^{q}_{r}| \times |a_{qr}| = |a_{qr}| \times ||^{q}_{r}| = |a_{qr}|. \qquad \text{(V. 9. 3)}$$

(v) The special importance of the set, or of any column or row of it, lies in its effect when combined with another set to form a product-sum. It will be found that, t having any one of the values 1, 2, 3, ... m,

$$|^{t}_{\mu} A_{\mu} = |^{\mu}_{t} A_{\mu} = A_{t}, \quad |^{t}_{\mu} A_{\mu\nu} = |^{\mu}_{t} A_{\mu\nu} = A_{t\nu}, \qquad \text{(V. 9. 4)}$$

$$|^{\lambda}_{\mu} A_{\mu} = |^{\mu}_{\lambda} A_{\mu} = A_{\lambda}, \quad |^{\lambda}_{\mu} A_{\mu\nu} = |^{\mu}_{\lambda} A_{\mu\nu} = A_{\lambda\nu}. \qquad \text{(V. 9. 5)}$$

[For example, take $t = 3$ in the first part of (V. 9. 4). Then

$$\begin{aligned} |^{3}_{\mu} A_{\mu} &= |^{3}_{1} A_{1} + |^{3}_{2} A_{2} + |^{3}_{3} A_{3} + |^{3}_{4} A_{4} + \ldots \\ &= 0 . A_{1} + 0 . A_{2} + 1 . A_{3} + 0 . A_{4} + \ldots \\ &= A_{3}.] \end{aligned}$$

Thus the effect of inner multiplication by $|^{\lambda}_{\mu}$ of a single or double set which contains μ (or λ) is to alter the latter to λ (or μ).

V. 10. Determinant properties.

—(i) Before we write down the final results, there is another small change which we shall find it convenient to make. In the statement $d^{\lambda\sigma} d_{\mu\sigma} = |^{\lambda}_{\mu}$, obtained in § 8 (ii), the linked suffixes are an upper σ and a lower σ, which cancel one another; and the free suffixes λ and μ are in the same respective positions on the two sides. It is desirable that, whenever possible, these two conditions should exist. In each of the statements $d_{\lambda\sigma} X_\lambda = Y_\sigma$ and $X_\lambda = d^{\lambda\sigma} Y_\sigma$, given at the end of § 6, one of the conditions exists but the other does not. We make them both exist by replacing X_λ, throughout, by

$$X^\lambda \equiv (X^1 X^2 \dots X^m),$$

as explained in § 5 (vii).

(ii) Our statement, after carrying out the alterations indicated in §§ 2–4, 6 and 8, and in (i) above, becomes—

Notation.	$D \equiv \mid d_{qr} \mid.$	(**V. 10. A**)
	$d^{ps} \equiv (\text{cofactor of } d_{ps} \text{ in } D) \div D.$	(**V. 10. B**)
	$D'' \equiv \mid d^{qr} \mid.$	(**V. 10. C**)
Properties.	$d^{\lambda\sigma} d_{\mu\sigma} = 1^\lambda_\mu.$	(**V. 10. 1**)
	$DD'' = 1.$	(**V. 10. 2**)
	$d_{ps} = (\text{cofactor of } d^{ps} \text{ in } D'') \div D''.$	(**V. 10. 3**)
	If $Y_\sigma = d_{\lambda\sigma} X^\lambda$, then $X^\lambda = d^{\lambda\sigma} Y_\sigma.$	(**V. 10. 4**)

To these we may add the formula (V. 6. 2) for multiplication of determinants, namely

$$\mid a_{qr} \mid \times \mid b_{qr} \mid = \mid a_{q\lambda} b_{\lambda r} \mid. \qquad \textbf{(V. 10. 5)}$$

V. 11. **Example of method.**—To illustrate the methods that we are now able to use, let us verify (V. 10. 4) by means of (V. 10. 1). It is given that

$$Y_\sigma = d_{\lambda\sigma} X^\lambda.$$

To find the value of $d^{\lambda\sigma} Y_\sigma$, it will not do to replace Y_σ by the above value as it stands, since we should then have three λ's. We must first replace λ in the expression for I_σ by some other suffix, say μ. We then have, by (V. 10. 1) and (V. 9. 5),

$$d^{\lambda\sigma} Y_\sigma = d^{\lambda\sigma} d_{\mu\sigma} X^\mu = |^\lambda_\mu X^\mu = X^\lambda,$$

which is what we wanted to prove. The reader will find it instructive, for comparison, to write out the proof in the ordinary notation.

In the above proof we have proceeded from $d^{\lambda\sigma}(d_{\mu\sigma} X^\mu)$ to $(d^{\lambda\sigma} d_{\mu\sigma}) X^\mu$. It has already been pointed out, in § 6 (v), that summations in a case of this kind can be made in any order.

SETS

VI. SETS OF QUANTITIES

1. 1. Introductory.—(i) It may have been noticed that in V. 11, in deducing (V. 10. 4) from (V. 10. 1) and (V. 9. 5), we made no direct use of determinant properties: the only indirect use being in the relations between elements and their cofactors, from which (V. 10. 1) was derived. But we can dispense even with this indirect use. In the equations $d_{\lambda\sigma} X^{\lambda} = Y_{\sigma}$ the values of d_{λ} are supposed to be known; and we can treat the ... nt in (V. 10. 1), namely

$$d^{\lambda\sigma} d_{\mu\sigma} = |^{\lambda}_{\mu},$$

a set of equations giving the values of $d^{\lambda\sigma}$ in terms of those of $d_{\lambda\sigma}$. If, for instance, $m = 20$, the set $d_{\lambda\sigma}$ contains 400 elements, and (V. 10. 1) is a condensed statement of the 400 equations (each with 20 terms on one side) which give the 400 values of $d^{\lambda\sigma}$. Thus for $p = 2$ we should have

$$\left.\begin{array}{l} d_{11} d^{21} + d_{12} d^{22} + d_{13} d^{23} + \ldots + d_{1m} d^{2m} = 0 \\ d_{21} d^{21} + d_{22} d^{22} + d_{23} d^{23} + \ldots + d_{2m} d^{2m} = 1 \\ d_{31} d^{21} + d_{32} d^{22} + d_{33} d^{23} + \ldots + d_{3m} d^{2m} = 0 \\ \qquad\qquad\qquad \vdots \\ d_{m1} d^{21} + d_{m2} d^{22} + d_{m3} d^{23} + \ldots + d_{mm} d^{2m} = 0 \end{array}\right\},$$

which give the values of $d^{21}, d^{22}, d^{23} \ldots d^{2m}$, i.e. of $d^{2\sigma}$. Similarly for $d^{1\sigma}, d^{3\sigma}$, etc.

(ii) We have, in fact, arrived at a position similar to that reached at the end of the first chapter. We started with the problem of solving a set of simultaneous equations, and arrived at a probable solution, involving what we called determinants. To verify the solution, we had to investigate the properties of determinants. The determinant thus took the leading place, its applicability to the solving of equations being one only of its properties. A determinant of order m is based on a set of m^2 quantities, which for convènience of reference are thought of as arranged in a square, the determinant being expressed by enclosing the set of symbols of the quantities between vertical lines: and we have reached the stage at which the set of quantities becomes the important thing, its existence as the basis of a determinant being one only of its properties.

(iii) These properties we have now to consider. The following sections of this chapter are mainly a restatement, with obvious modifications and extensions, of results obtained in the preceding chapter.

VI. 2. Single sets.—(i) We may have a **single set**

$$A_\rho \equiv (A_1 \quad A_2 \quad A_3 \dots A_m).$$

The separate quantities $A_1 A_2 \dots A_m$ comprised in the set are called its **elements**. The **order** of the set is the number of elements comprised in it. The typical statement with regard to such a set is of the form

$$A_\rho = E_\rho.$$

This, in the first instance, we regarded merely as a short way of saying that

$$A_1 = E_1, \quad A_2 = E_2, \dots A_m = E_m;$$

but we must now think of it as a statement that the two sets A_ρ and E_ρ, each taken as a whole, are equal, this equality of the wholes implying the equality of corresponding elements. The analogy of a vector may help us here.

The statement that two vectors are equal implies that the components are equal, each to each; but what we really think of is not the separate equalities of the components but the equality, in all respects, of the vectors.

(ii) Single sets behave like ordinary vectors as regards addition and subtraction and multiplication by a scalar (§ 4 (iii)), but not as regards multiplication of one single set by another.

VI. 3. Double sets.—(i) We may have a **double set** of order m—i.e. comprising m^2 elements—

$$A_{\mu\rho} \equiv \begin{pmatrix} A_{11} & A_{21} & A_{31} \dots A_{m1} \\ A_{12} & A_{22} & A_{32} \dots A_{m2} \\ \vdots & \vdots & \vdots \quad\quad \vdots \\ A_{1m} & A_{2m} & A_{3m} \dots A_{mm} \end{pmatrix}.$$

Here the quantity in the qth column and rth row is A_{qr}. We adopt the convention that the first Greek letter (in alphabetical order) represents the column and the second the row, so that

$$A_{\rho\mu} \equiv \begin{pmatrix} A_{11} & A_{12} & A_{13} \dots A_{1m} \\ A_{21} & A_{22} & A_{23} \dots A_{2m} \\ \vdots & \vdots & \vdots \quad\quad \vdots \\ A_{m1} & A_{m2} & A_{m3} \dots A_{mm} \end{pmatrix},$$

the representative element of which is A_{rq}. The sets $A_{\mu\rho}$ and $A_{\rho\mu}$ are called the **transposed** of each other.

(ii) The determinant

$$| A_{qr} |$$

is called the determinant of the set* $A_{\mu\rho}$, and similarly $| A_{rq} |$ is the determinant of $A_{\rho\mu}$.

* The set is usually called the *matrix* of the determinant. It is a singularly inappropriate name, as the symbol of the set is the inner part of that of the determinant, not something which surrounds it. The set is really the substance or *core* of the determinant.

(iii) The brackets in which the elements of A_ρ, $A_{\mu\rho}$, $A_{\rho\mu}$ have been placed are not essential,* and have been introduced partly to help the eye and partly to indicate that the sets are being considered as a whole.

(iv) A double set is **symmetrical** if columns and rows can be interchanged without altering it. Thus, if $A_{\mu\rho}$ is symmetrical, then $A_{\mu\rho}=A_{\rho\mu}$; and conversely.

VI. 4. Sets generally.—(i) We describe a single set as being of **rank**† 1, and a double set as being of rank 2.

(ii) Similarly a set of rank 3 of order m is made up of m double sets of order m; and so on. Thus we might represent a set of rank 3 by $A_{\mu\nu\rho}$. There would have to be a convention as to the order of the symbols, so as to distinguish $A_{\mu\nu\rho}$ from $A_{\nu\mu\rho}$ etc. Where, however, the set is symmetrical, so that $A_{\mu\nu\rho}=A_{\nu\mu\rho}=$etc., this difficulty does not arise.

(iii) The set of rank 0 is a single quantity or **scalar**.

(iv) To denote a set generally, without reference to its suffixes, we use a Gothic letter such as $\mathfrak{A}$ or $\mathfrak{B}$.

* An alternative method, in the case of a double set or matrix, is to enclose the symbols between two pairs of vertical lines, so as to distinguish the set from the determinant, which has two single lines. It is not a satisfactory symbolism from our point of view, as it would seem to suggest that the set is more restricted than the determinant, whereas what we are aiming at is to free the set from the bonds of the determinant.

† I have been doubtful as to the appropriate word. In reference to tensors Einstein uses *Rang*, Hilbert *Ordnung*, Eddington *rank*, de Sitter *order*. The objection to either of the last two is that there is already a settled meaning for *order* as regards a determinant, and (though this is not so important) for *rank* as regards a matrix (see note to § 3 (ii)). It would seem reasonable to describe a set containing m^f elements (i. e. composed of m sets each containing m^{f-1} elements) as being of *degree f*. I have, however, felt bound to keep to Eddington's use of *rank*.

VI. 5. Sums and products of sets.—(i) If two or more sets, of the same rank and the same order, have *the same* suffixes, we add (or subtract) them by adding (or subtracting) corresponding elements. Thus

$$A_\lambda + B_\lambda \equiv (A_1 + B_1 \quad A_2 + B_2 \quad A_3 + B_3 \ldots A_m + B_m);$$

and similarly for sets of higher rank.

(ii) We multiply a set, of whatever rank, by a scalar when we multiply every element of the set by the scalar; e.g.

$$kA_{\mu\rho} \equiv \begin{pmatrix} kA_{11} & kA_{21} & kA_{31} & \ldots kA_{m1} \\ kA_{12} & kA_{22} & kA_{32} & \ldots kA_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ kA_{1m} & kA_{2m} & kA_{3m} & \ldots kA_{mm} \end{pmatrix}.$$

Thus the determinant of $kA_{\mu\rho}$ is not k times, but k^m times, the determinant of $A_{\mu\rho}$.

(iii) If $\mathfrak{A}$ and $\mathfrak{B}$ are two sets, with *different* suffixes, o ranks f and g respectively (f and g not being necessarily different), their **product** $\mathfrak{A}\,\mathfrak{B}$ is the set of rank $f+g$ obtained by multiplying every element of one by every element of the other. (Here, as elsewhere, we assume that all the sets we are considering are of the same order.) Thus the product of two single sets A_ρ and B_σ is the double set $A_\rho B_\sigma$ obtained by giving to ρ and σ separately each of the values 1 to m. A product obtained in this way is sometimes called an **outer** product, to distinguish it from an 'inner' product as defined in § 6 below.

VI. 6. Inner product.—(i) When a suffix occurs twice in an expression such as $A_{\nu\nu}$ or $B_\nu C_\nu$, or, more generally, in any single expression or product, e.g. $A_{\lambda\nu\nu\rho\ldots}$ or $B_{\nu\rho\ldots}C_{\nu\sigma\ldots}$ (where the letters may be in any order), this means that the expression is to be summed for the values 1, 2...m of the suffix; e.g.

$$\left.\begin{aligned} B_\nu C_\nu &\equiv B_1 C_1 + B_2 C_2 + \ldots + B_m C_m \\ B_\nu D_{\lambda\nu} &\equiv B_1 D_{\lambda 1} + B_2 D_{\lambda 2} + \ldots + B_m D_{\lambda m} \\ &\text{etc.} \end{aligned}\right\}. \qquad \textbf{(VI. 6. A)}$$

Where a letter occurs twice in this way, each of the two letters is **linked.** Where a letter occurs once only, it is **free.** The linked suffixes are called **dummy**, as they can be replaced by a pair of any other suffixes not already occurring in the expression.

(ii) In the particular case where the expression to be summed is of the form $B_\nu C_\nu$, the result is called the **inner product** of B_ν and C_ν, or of B_λ and C_ρ, etc. It is immaterial what suffixes we use in this latter description, since they have to be replaced by one and the same suffix.

(iii) From a pair of double sets $A_{\mu\rho}$ and $B_{\mu\rho}$, or $A_{\lambda\sigma}$ and $B_{\mu\rho}$, we can by a single product-summation form several different double sets $A_{\mu\nu}\,B_{\nu\rho}$, $A_{\lambda\rho}\,B_{\mu\rho}$, $A_{\lambda\mu}\,B_{\lambda\rho}$, etc. There is also the scalar quantity $A_{\mu\rho}\,B_{\mu\rho}$, formed by two summations, which can be simultaneous or successive; if successive, the first is a product-summation giving us the double set $A_{\lambda\rho}\,B_{\mu\rho}$ or $A_{\mu\rho}\,B_{\mu\sigma}$. Strictly speaking, this scalar quantity $A_{\mu\rho}\,B_{\mu\rho}$ is *the* inner product of $A_{\mu\rho}$ and $B_{\mu\rho}$. But it occurs less frequently than the double sets obtained by a single summation, and it is therefore more convenient to call one of these latter the inner product. We shall call $A_{\mu\rho}\,B_{\mu\rho}$ the **complete inner product** of $A_{\mu\rho}$ and $B_{\mu\rho}$. By analogy with the expression found in V. 6 (vi) for the product of two determinants, we define the **inner product*** of two sets $A_{\mu\rho}$ and $B_{\mu\rho}$—or, more generally, of

* This is what, in the case of matrices, is called the 'product'. The true product of two sets $A_{\mu\sigma}$ and $B_{\lambda\rho}$ is a set of rank 4.

If this use of 'inner product' seemed likely to lead to confusion with the 'complete inner product', we could use a different phrase, such as 'interproduct'. It should be noted that the intertwo sets $A_{\mu\sigma}$ and $B_{\lambda\rho}$—, in each of which the letters are in their proper alphabetical order, as the set $A_{\mu\nu}B_{\nu\rho}$; the linked suffix in this latter expression being chosen so as to be (alphabetically) intermediate between the two free suffixes. This is equivalent to saying, as regards this case, that *the inner product of two double sets is the double set whose element in the* q*th column and* r*th row is the inner product of the* q*th column of the first set and the* r*th row of the second set*; and we apply this rule to all cases. The simplest way of applying it is to use the result for $A_{\mu\rho}$ and $B_{\mu\rho}$ and alter the order of the suffixes where necessary. Suppose, for instance, that we want the inner product of $A_{\mu\rho}$ and $B_{\rho\mu}$; then by writing $B_{\rho\mu}\equiv F_{\mu\rho}$ we see that the inner product is $A_{\mu\nu}F_{\nu\rho}=A_{\mu\nu}B_{\rho\nu}$. It should be noticed that in all cases the inner product depends on the relative position of the original sets; thus the inner product of $B_{\mu\rho}$ and $A_{\mu\rho}$ is not $A_{\mu\nu}B_{\nu\rho}$ but $B_{\mu\nu}A_{\nu\rho}$.

(iv) There are four main forms of inner product constructed in accordance with (iii); and four others, which are really repetitions, can be obtained by interchanging the two sets. Denoting the inner product of $\mathfrak{A}$ and $\mathfrak{B}$ (in this order) by $\mathfrak{A}\times\mathfrak{B}$ (cf. IV. 5 (ii) as to product of two determinants), the forms are as follows:

$$A_{\mu\rho}\times B_{\mu\rho}=A_{\mu\nu}B_{\nu\rho} \quad (1) \qquad B_{\mu\rho}\times A_{\mu\rho}=B_{\mu\nu}A_{\nu\rho}=A_{\nu\rho}B_{\mu\nu} \quad (5)$$
$$A_{\mu\rho}\times B_{\rho\mu}=A_{\mu\nu}B_{\rho\nu} \quad (2) \qquad B_{\rho\mu}\times A_{\mu\rho}=B_{\nu\mu}A_{\nu\rho}=A_{\nu\rho}B_{\nu\mu} \quad (6)$$
$$A_{\rho\mu}\times B_{\mu\rho}=A_{\nu\mu}B_{\nu\rho} \quad (3) \qquad B_{\mu\rho}\times A_{\rho\mu}=B_{\mu\nu}A_{\rho\nu}=A_{\rho\nu}B_{\mu\nu} \quad (7)$$
$$A_{\rho\mu}\times B_{\rho\mu}=A_{\nu\mu}B_{\rho\nu} \quad (4) \qquad B_{\rho\mu}\times A_{\rho\mu}=B_{\nu\mu}A_{\rho\nu}=A_{\rho\nu}B_{\nu\mu} \quad (8)$$

(v) The transposed of an inner product such as $A_{\mu\nu}B_{\nu\rho}$ is found in the usual way (V. 5 (iii)) by interchanging the free suffixes μ and ρ. By comparison of (1) with (8), (2) with (7), etc., it will be seen that *the transposed of the inner product of two double sets is the*

---

mediate product-sum, for $A_{\mu\rho}B_{\mu\rho}$, is not $A_{\mu\nu}B_{\nu\rho}$ but either $A_{\lambda\rho}B_{\mu\rho}$ or $A_{\mu\rho}B_{\mu\sigma}$. There are various reasons for taking the former, rather than one of the two latter, as 'the' inner product.

inner product of the transposed sets, in reverse order; e.g. the transposed of the inner product of $A_{\rho\mu}$ and $B_{\rho\mu}$ is the inner product of $B_{\mu\rho}$ and $A_{\mu\rho}$.

(vi) It will be seen by comparison with IV. 5 (ii) and the Appendix that the rule for construction of the inner product of two double sets is exactly the same as that for construction of the product of two determinants; so that the determinant of the inner product of two double sets—whether we call them (say) $A_{\mu\rho}$ and $B_{\rho\mu}$ or $A_{\mu\sigma}$ and $B_{\rho\lambda}$—is equal to the product of the determinants of the two sets.

VI. 7. The unit set.—(i) The **unit set***

$$|_{\rho}^{\mu}$$

is defined as the set whose typical term is

$$|_{r}^{q} \equiv \begin{Bmatrix} 1 \text{ if } r = q \\ 0 \text{ if } r \neq q \end{Bmatrix}, \qquad \textbf{(VI. 7. A)}$$

so that

$$|_{\rho}^{\mu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 & 0 \ldots 0 \\ 0 & 0 & 1 & 0 \ldots 0 \\ \vdots & \vdots & \vdots & \vdots \quad \vdots \\ 0 & 0 & 0 & 0 \ldots 1 \end{pmatrix}. \qquad \textbf{(VI. 7. 1)}$$

(ii) From the definition it follows that the set is symmetrical, i.e.

$$|_{\rho}^{\mu} = |_{\mu}^{\rho}, \qquad \textbf{(VI. 7. 2)}$$

and that

$$\left| \, |_{r}^{q} \right| = 1. \qquad \textbf{(VI. 7. 3)}$$

(iii) The special property of this set is that, if A_{μ} is any set (possibly containing other suffixes $\sigma\ \tau \ldots$), then

$$|_{\mu}^{\nu} A_{\mu} = |_{\nu}^{\mu} A_{\mu} = A_{\nu}, \qquad \textbf{(VI. 7. 4)}$$

so that the inner product of the unit set and any other set

* This is by analogy with the 'unit matrix'.

is the same as the latter but with the suffix changed. In other words, the unit set acts, for inner multiplication, as a *substitution-operator*.

(iv) In particular, the inner product of two unit sets is a unit set, i.e.

$$|^{\mu}_{\nu}|^{\nu}_{\rho} = |^{\mu}_{\rho}. \qquad \text{(VI. 7. 5)}$$

VI. 8. Inverse double sets.

—(i) We take any double set

$$A_{\mu\rho},$$

and we say that there is another set

$$A^{\rho\mu}$$

connected with it by the condition that the inner product of the former and the latter is a unit set, i.e. (see §§ 6 (iii) and 7 (i)) that

$$A_{\mu\nu} A^{\rho\nu} = |^{\rho}_{\mu}. \qquad \text{(VI. 8. 1)}$$

This represents m^2 equations, which are sufficient to determine the m^2 values of A^{rq} when those of A_{qr} are known. The set $A^{\rho\mu}$, as defined by the above condition, is called the **inverse** of the set $A_{\mu\rho}$. We shall keep to this notation, so that (VI. 8. 1) will always hold, however we alter the letters A, μ, ν, ρ.

(ii) The above is subject to one condition. If we write down the equations which determine the elements in, say, the second row of $A^{\rho\mu}$, namely

$$\left.\begin{aligned} A_{11}A^{21} + A_{12}A^{22} + A_{13}A^{23} + \ldots + A_{1m}A^{2m} &= 0 \\ A_{21}A^{21} + A_{22}A^{22} + A_{23}A^{23} + \ldots + A_{2m}A^{2m} &= 1 \\ A_{31}A^{21} + A_{32}A^{22} + A_{33}A^{23} + \ldots + A_{3m}A^{2m} &= 0 \\ &\vdots \\ A_{m1}A^{21} + A_{m2}A^{22} + A_{m3}A^{23} + \ldots + A_{mm}A^{2m} &= 0 \end{aligned}\right\},$$

we see that in order that there may be a solution it is necessary that we should have $|A_{rq}| \neq 0$, which is the same thing as

$$|A_{qr}| \neq 0. \qquad \text{(VI. 8. 2)}$$

This applies also to the other rows. It is a sufficient as well as a necessary condition for the existence of $A^{\rho\mu}$.

(iii) Taking it that $|A_{qr}| \neq 0$, we have, by (V. 6. 2) and (VI. 8. 1) and (VI. 7. 3),

$$|A^{qr}| \times |A_{rq}| = |A^{q\nu}A_{r\nu}| = ||{}^q_r| = 1. \qquad \text{(VI. 8. 3)}$$

It follows that

$$|A^{qr}| \neq 0, \quad |A^{rq}| \neq 0. \qquad \text{(VI. 8. 4)}$$

(iv) The statement (VI. 8. 1) is a statement as to the m^2 relations obtained by taking each value of ρ with each value of μ. It is therefore equally true to say, by interchanging μ and ρ, that

$$A^{\mu\nu} A_{\rho\nu} = |{}^\mu_\rho. \qquad \text{(VI. 8. 5)}$$

The expression on the left-hand side is (§ 6 (iii)) the inner product of $A^{\mu\rho}$ and $A_{\rho\mu}$; and these are the transposed of $A^{\rho\mu}$ and $A_{\mu\rho}$ respectively. Hence, if $\mathfrak{B}$ is the inverse of $\mathfrak{A}$, the transposed of $\mathfrak{A}$ is the inverse of the transposed of $\mathfrak{B}$.

(v) The relation in (VI. 8. 1) is a relation connecting columns of the original set and rows of the inverse set. There is a similar relation connecting rows and columns. For (VI. 8. 1) gives

$$A^{\mu\lambda} A_{\mu\nu} A^{\rho\nu} = |{}^\rho_\mu A^{\mu\lambda} = A^{\rho\lambda} = |{}^\lambda_\nu A^{\rho\nu};$$

whence, as will be shown in § 9 (v), it follows that

$$A^{\mu\lambda} A_{\mu\nu} = |{}^\lambda_\nu, \qquad \text{(VI. 8. 6)}$$

and hence also

$$A_{\mu\lambda} A^{\mu\nu} = |{}^\nu_\lambda. \qquad \text{(VI. 8. 7)}$$

The expression on the left-hand side of (VI. 8. 6) is the inner product of $A^{\nu\lambda}$ and $A_{\lambda\nu}$, which is the equivalent of that of $A^{\rho\mu}$ and $A_{\mu\rho}$; so that, by the definition in (i), the latter is the inverse of the former. Hence, if $\mathfrak{B}$ is the inverse of $\mathfrak{A}$, then $\mathfrak{A}$ is the inverse of $\mathfrak{B}$. The relation (VI. 8. 7) is similar to (VI. 8. 5), and shows that the transposed of $\mathfrak{B}$ is the inverse of the transposed of $\mathfrak{A}$.

VI. 9. **Reciprocation.**—(i) Suppose there are two single sets X^λ and Y_λ connected by the relation

$$Y_\sigma = A_{\lambda\sigma} X^\lambda.$$

Then, as in V. 11, we have

$$A^{\lambda\sigma} Y_\sigma = A^{\lambda\sigma} A_{\mu\sigma} X^\mu = |^{\lambda}_{\mu} X^\mu = X^\lambda;$$

i. e.

$$\text{If} \quad Y_\sigma = A_{\lambda\sigma} X^\lambda, \quad \text{then} \quad X^\lambda = A^{\lambda\sigma} Y_\sigma. \qquad \text{(VI. 9. 1)}$$

Similarly by taking X^λ to be each single set, in turn, of a set of second or higher rank, with Y_λ to correspond, we find that

$$\text{If } Y_{\sigma\tau\ldots} = A_{\lambda\sigma} X^\lambda_{\tau\ldots}, \text{ then } X^\lambda_{\tau\ldots} = A^{\lambda\sigma} Y_{\sigma\tau\ldots}. \qquad \text{(VI. 9. 2)}$$

(ii) Thus the operation represented by $A_{\lambda\sigma}$ is annulled by the operation $A^{\lambda\sigma}$; and conversely. The sets $A_{\lambda\sigma}$ and $A^{\lambda\sigma}$ will be said to be **reciprocal** to one another: and the process adopted in (VI. 9. 1) and (VI. 9. 2)—which we shall have to use very frequently—will be called **reciprocation.**

(iii) We see from § 8 that the reciprocal of a set is the transposed of the inverse of the set, and conversely. If a set is symmetrical, its inverse and its reciprocal are identical.

(iv) As an example of the application of (VI. 9. 1), suppose that

$$A_{\lambda\sigma} X^\lambda = 0$$

for all values of σ. Then, by reciprocation,

$$X^\lambda = A^{\lambda\sigma}\,0 = 0,$$

provided that $|A_{qr}| \neq 0$. (This is practically the same thing as saying that, if the inner products of X^λ by m independent single sets are all 0, then X^λ is 0.)

(v) Similarly, suppose that

$$A_{\lambda\nu}\,C^{\nu\sigma} = A_{\lambda\nu}\,D^{\nu\sigma}$$

identically, i.e. for all values of λ and σ, and that $|A_{qr}|$ is not $= 0$. Then, by reciprocation,

$$C^{\nu\sigma} = A^{\lambda\nu}\,A_{\lambda\rho}\,D^{\rho\sigma} = |^{\nu}_{\rho}\,D^{\rho\sigma} = D^{\nu\sigma}.$$

Thus we can divide both sides of the equation by $A_{\lambda\nu}$. This supplies the missing step in § 8 (v).

(vi) The two definitions, and the proposition, used in the establishment of (VI. 9. 1) are

$$|^{q}_{r} = \begin{cases} 1 \text{ if } r = q \\ 0 \text{ if } r \neq q \end{cases}, \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(A)}$$

$$|^{\lambda}_{\mu}\,A^{\mu} = A^{\lambda}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$A_{\mu\sigma}A^{\lambda\sigma} = |^{\lambda}_{\mu}; \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(B)}$$

and from these we deduce (VI. 9. 1). We could have altered the order in various ways. For instance, we could have defined $|^{\lambda}_{\mu}$ by (1); thence, by giving λ its successive values, and equating coefficients, we should have got (A). Also we might have defined $A^{\lambda\sigma}$ by (VI. 9. 1), instead of by (B), as the coefficients of Y_σ when the equations $Y_\sigma = A_{\lambda\sigma}X^\lambda$ are solved for X^λ. This would give $X^\lambda = A^{\lambda\sigma}\,A_{\mu\sigma}X^\mu$. Then, if we defined $|^{\lambda}_{\mu}$ by (1), we should have $A^{\lambda\sigma}A_{\mu\sigma} = |^{\lambda}_{\mu}$; or, if we defined $|^{\lambda}_{\mu}$ by (B), we should have $X^\lambda = |^{\lambda}_{\mu}\,X^\mu$, which is (1).

VI. 10. Continued inner products.—(i) We can construct continued inner products without ambiguity, provided we adhere to the rule laid down in § 6 (iii). Suppose, for instance, that we want the inner product of $A_{\lambda\sigma}$, $B_{\lambda\sigma}$, and $C_{\sigma\lambda}$. That of $A_{\lambda\sigma}$ and $B_{\lambda\sigma}$, according to the rule, is $A_{\lambda\mu}B_{\mu\sigma}$. If we call this $F_{\lambda\sigma}$, then, by (2) of § 6 (iv), the inner product of $F_{\lambda\sigma}$ and $C_{\sigma\lambda}$ is $F_{\lambda\nu}C_{\sigma\nu}$; i.e. the inner product of $A_{\lambda\sigma}$, $B_{\lambda\sigma}$, and $C_{\sigma\lambda}$ is $A_{\lambda\mu}B_{\mu\nu}C_{\sigma\nu}$. Similarly that of $A_{\lambda\sigma}$, $B_{\lambda\sigma}$, $C_{\sigma\lambda}$, and $D_{\lambda\sigma}$ is $A_{\lambda\mu}B_{\mu\nu}C_{\rho\nu}D_{\rho\sigma}$.

(ii) On the other hand, the inner product of $A_{\sigma\lambda}$, $B_{\sigma\lambda}$, $C_{\lambda\sigma}$, $D_{\sigma\lambda}$ is not $A_{\sigma\rho}B_{\rho\nu}C_{\mu\nu}D_{\mu\lambda}$. For, by (4) of § 6 (iv), that of $A_{\sigma\lambda}$ and $B_{\sigma\lambda}$ is $B_{\sigma\mu}A_{\mu\lambda}$. Calling this $G_{\sigma\lambda}$, the inner product of $G_{\sigma\lambda}$ and $C_{\lambda\sigma}$ is, by (3) of § 6 (iv), $C_{\nu\sigma}G_{\nu\lambda} = C_{\nu\sigma}B_{\nu\mu}A_{\mu\lambda}$. Similarly that of $A_{\sigma\lambda}$, $B_{\sigma\lambda}$, $C_{\lambda\sigma}$, $D_{\sigma\lambda}$ is $D_{\sigma\rho}C_{\nu\rho}B_{\nu\mu}A_{\mu\lambda}$.

(iii) The transposed of the inner product of any number of double sets is the inner product of the transposed sets, in reverse order; e.g. the transposed of the inner product of $A_{\lambda\sigma}$, $B_{\lambda\sigma}$, $C_{\sigma\lambda}$, $D_{\lambda\sigma}$ is the inner product of $D_{\sigma\lambda}$, $C_{\lambda\sigma}$, $B_{\sigma\lambda}$, $A_{\sigma\lambda}$. [For we have shown, in § 6 (v), that this is true for the inner product of two sets; and thence it follows, by induction, for any number of sets.]

(iv) The inverse of the inner product of any number of double sets is the inner product of the inverse sets, in reverse order; e.g. the inverse of the inner product of $B_{a\sigma}$, $C_{a\sigma}$, $D_{a\sigma}$, $E_{a\sigma}$ is the inner product of $E^{\sigma a}$, $D^{\sigma a}$, $C^{\sigma a}$, $B^{\sigma a}$, i.e.

If $A_{a\sigma} \equiv B_{a\beta}C_{\beta\gamma}D_{\gamma\delta}E_{\delta\sigma}$, then $A^{\sigma a} = B^{\sigma\delta}C^{\delta\gamma}D^{\gamma\beta}E^{\beta a}$. **(VI.10.1)**

[Denote this latter expression (right-hand side) by $F^{\sigma a}$, and alter $\beta\ \gamma\ \delta$ in it to $\mu\ \nu\ \rho$. Then the inner product of $A_{a\sigma}$ and $F^{\sigma a}$ is

$$\begin{aligned} A_{a\lambda}F^{\sigma\lambda} &= B_{a\beta}B^{\sigma\rho}C_{\beta\gamma}C^{\rho\nu}D_{\gamma\delta}D^{\nu\mu}E_{\delta\lambda}E^{\mu\lambda} \\ &= B_{a\beta}B^{\sigma\rho}C_{\beta\gamma}C^{\rho\nu}D_{\gamma\delta}D^{\nu\mu}\big|^{\mu}_{\delta} \\ &= B_{a\beta}B^{\sigma\rho}C_{\beta\gamma}C^{\rho\nu}\big|^{\nu}_{\gamma} \\ &= B_{a\beta}B^{\sigma\rho}\big|^{\rho}_{\beta} \\ &= \big|^{\sigma}_{a}. \end{aligned}$$

Hence, by reciprocation,

$$F^{\sigma\lambda} = A^{a\lambda}\big|^{\sigma}_{a} = A^{\sigma\lambda},$$

so that

$$F^{\sigma a} = A^{\sigma a}.]$$

(v) It follows from (iv) and (iii), since $A^{a\sigma}$ is the transposed of $A^{\sigma a}$, that the reciprocal of the inner product of any number of double sets is the inner product of the reciprocal sets; e.g.

If $A_{a\sigma} \equiv B_{a\beta}C_{\beta\gamma}D_{\gamma\delta}E_{\delta\sigma}$, then $A^{a\sigma} = B^{a\beta}C^{\beta\gamma}D^{\gamma\delta}E^{\delta\sigma}$. **(VI. 10. 2)**

VI. 11. **Partial sets.**—(i) When we are dealing with a set

$$A_\lambda \equiv (A_1\ A_2\ A_3 \ldots A_m)$$

we sometimes want to consider the separate or mutual relations of groups of the A's. The simplest case is when the set divides into two groups, one group consisting of k elements, which we shall take to be the first k, and the other group consisting of the other $m-k$ elements. If we use suffixes $\alpha\beta\gamma\ldots$ in reference to the first group, and $\phi\chi\psi\ldots$ in reference to the second, reserving $\lambda\mu\nu\ldots$ for the set as a whole, we may treat the two groups as **partial single sets** of orders k and $m-k$ respectively, and write

$$A_\alpha \equiv (A_1\ A_2 \ldots A_k),\ A_\phi \equiv (A_{k+1}\ A_{k+2} \ldots A_m).$$

(ii) In the same way a double set $A_{\mu\rho}$ may fall into four groups by division by two lines cutting off k columns and k rows respectively. We could denote these groups by

$$\left\{ \begin{matrix} A_{\alpha\gamma} & A_{\phi\gamma} \\ A_{\alpha\psi} & A_{\phi\psi} \end{matrix} \right\},$$

ϕ being regarded as coming before γ. The groups $A_{\alpha\gamma}$ and $A_{\phi\psi}$ would be **partial double sets** of orders k and $m-k$ respectively. The groups $A_{\alpha\psi}$ and $A_{\phi\gamma}$ would each have different numbers of columns and of rows, and therefore would not be double sets; but this would usually not matter, as we should be specially concerned with $A_{\alpha\gamma}$ and $A_{\phi\psi}$. The important point to notice is that, if we take $A_{\alpha\gamma}$, say, as a partial set and construct the inverse set $A^{\gamma\alpha}$ or the reciprocal set $A^{\alpha\gamma}$, the set so constructed will not in general be the same as the set made up of the corresponding elements of the set inverse or reciprocal to $A_{\mu\rho}$. The inverse set $A^{\gamma\alpha}$, for instance, is given by

$$A_{\alpha\beta}A^{\gamma\beta} = |\,^\gamma_\alpha,$$

with summations made only from 1 to k instead of from 1 to m. To avoid mistake, we may write it $(A^{\gamma\alpha})_k$. Similarly the set inverse to the partial set $A_{\phi\psi}$ may be written $[A^{\psi\phi}]_{m-k}$.

(iii) If, however, all the elements in the portions $A_{\alpha\psi}$ and $A_{\phi\gamma}$ are 0, so that the set $A_{\mu\rho}$ practically consists only of the two partial sets $A_{\alpha\gamma}$ and $A_{\phi\psi}$, this is also the case for the complete reciprocal set $A^{\mu\rho}$; all the elements in $A^{\alpha\psi}$ and $A^{\phi\gamma}$ are 0, and the elements of $A^{\alpha\gamma}$ and $A^{\phi\psi}$ are just the same whether they are regarded as obtained from the complete set $A_{\mu\rho}$ or from the partial sets $A_{\alpha\gamma}$ and $A_{\phi\psi}$.

(iv) If we had two or more sets, single or double, divided in the manner described above, we could take portions from different sets to form new sets. If, for instance, we had divided A_λ into A_α and A_ϕ, and B_λ into B_α and B_ϕ (orders again k and $m-k$), we could construct a new set consisting of A_α and B_ϕ.

(v) It would, of course, be incorrect to describe this new set as being $A_\alpha + B_\phi$, or the old set as being $A_\alpha + A_\phi$: for we cannot add together two sets of different order. We could, however, look at the matter in another way. Consider the two sets

$$(A_1 \quad A_2 \ldots A_k \quad 0 \quad 0 \quad \ldots 0 \ ),$$
$$(0 \quad 0 \ldots 0 \quad A_{k+1} \quad A_{k+2} \ldots A_m).$$

The sum of these, if we regard each as having the suffix λ, is A_λ; and in this sense, if we denote the two sets by A_α and A_ϕ, and regard α and ϕ as connoting λ, we could say that

$$A_\lambda = A_\alpha + A_\phi.$$

If we compare A_λ with a vector, we see that A_α and A_ϕ correspond to the projections of A_λ on a 'plane', i.e. a surface of the first degree, passing through the first k axes, and on a 'plane' through the last $m-k$ axes, respectively.

The extreme form, if we made further divisions, would be that in which A_λ was split up into m component single sets, each having $m-1$ 0's in it. It would only be in this sense that we could describe the set as being the sum of its m components.

VII. RELATED SETS OF VARIABLES

VII. 1. **Variable sets.**—(i) In the earlier chapters we considered the manner in which determinants arose in solving a set of equations of the form

$$(s = 1, 2, \ldots m)\, d_{1s}X^1 + d_{2s}X^2 + \ldots + d_{ms}X^m = Y_s; \quad (1)$$

and in the chapter preceding this we have considered the general aspects of the system under which we express these equations and their solution in the form

$$Y_\sigma = d_{\lambda\sigma}X^\lambda, \quad X^\lambda = d^{\lambda\sigma}Y_\sigma. \quad (2)$$

According to the definitions we gave to the notation, X^λ and Y_σ are each used in different senses in the two places where they occur in (2): X^λ means 'the elements of X^λ' on its first occurrence and 'each element of X^λ' on its second occurrence; and similarly for Y_σ, but in the reverse order. We have, however, by this time practically reached the stage of treating a set as a whole, so that we can now regard (2) as a pair of statements, one of which gives an expression for the set Y_σ in terms of the set X^λ, while the other expresses X^λ in terms of Y_σ. The form of either expression determines the nature of the relation between the two sets.

(ii) The special features of the particular case were that the X's were unknown quantities which we wanted to find, that the d's were coefficients, more or less accidental, and that the Y's were known quantities arising from the application of these coefficients to the X's; and, more important, that this was merely an isolated set of equations, for which we had no further use when we had found the X's.

(iii) The relations with which we have to deal in this and subsequent chapters are of a different nature. We have a set $\mathfrak{A}$ and a set $\mathfrak{B}$, each consisting of a number (the same for both) of elements, which we will call the A's and the B's. In each set the elements need not be all of the same kind. Each of the A's is a variable; i.e. it either has, or can (as in the theory of statistics or of error) be regarded as having, a very large number of actual or possible values. These variables, the values of which are algebraically independent,* together constitute the variable set $\mathfrak{A}$. In the same way the B's are variables, and constitute another variable set $\mathfrak{B}$. But the two sets of variables are not independent of each other: they are connected by certain relations, by means of which the B's are known if the A's are known, and conversely. Thus the B's are functions of the A's, in the ordinary sense of the word, and the A's are functions of the B's. In this case we say that $\mathfrak{B}$ is a function of $\mathfrak{A}$, and $\mathfrak{A}$ a function of $\mathfrak{B}$. But we must not only say it, but think it; i.e. we must treat the functional relations of the A's and the B's rather as interpreting the nature of the functionality of $\mathfrak{A}$ and $\mathfrak{B}$ than as actually constituting this functionality.

(iv) In the particular case we have been considering, $\mathfrak{A}$ and $\mathfrak{B}$ were the single sets X^λ and Y_λ, and the relation between them was *linear*; i.e. the Y's were linear functions of the X's, and the X's were therefore linear functions of

* By algebraical independence of m quantities we mean that each may have any of its values, whatever the values of the other $m-1$ may be. This does not imply statistical independence, which is a different thing.

the Y's. In such a case we say that Y_λ is a **linear function** of X^λ, and it follows that X^λ is a linear function of Y_λ.

(v) In dealing with the theory of the subject, as distinct from its applications, we are concerned not with the actual values of elements of sets but with the relations between the sets. Thus in the case of the linear relation $Y_\sigma = d_{\lambda\sigma}X^\lambda$, where the variable single set Y_σ is expressed in terms of the variable single set X^λ and the fixed double set $d_{\lambda\sigma}$, the elements of $d_{\lambda\sigma}$ form a kind of framework

$$\left.\begin{array}{l} \wedge = d_{11} \wedge + d_{21} \wedge + d_{31} \wedge + \dots + d_{m1} \wedge \\ \wedge = d_{12} \wedge + d_{22} \wedge + d_{32} \wedge + \dots + d_{m2} \wedge \\ \vdots \\ \wedge = d_{1m} \wedge + d_{2m} \wedge + d_{3m} \wedge + \dots + d_{mm} \wedge \end{array}\right\}$$

into which the values of the X's and the Y's can be fitted; and what we are really investigating are the properties and mutual relations of such frameworks. In the present chapter we shall consider certain simple relations between two such frameworks, namely relations between the linear relation of one pair of sets and the linear relation of another pair of sets.

VII. 2. **Direct proportion of single sets.**—(i) If a quantity Z is a linear function of m X's, which we will call $X^1, X^2, \dots X^m$, it is of the form

$$Z = h_1X^1 + h_2X^2 + \dots + h_mX^m = h_\lambda X^\lambda. \qquad (1)$$

This is the simplest form of statement of a linear relation. Suppose, for instance, that Z is the 3rd difference of the X's, formed in the usual way, i.e.

$$Z = \Delta\Delta\Delta X^1.$$

This is equivalent to $Z = X^4 - 3X^3 + 3X^2 - X^1$, so that

$$h_1 = -1,\ h_2 = 3,\ h_3 = -3,\ h_4 = 1,\ h_5 = h_6 = \dots = h_m = 0.$$

The h's having these values, the X's may alter, but (1) will always give the 3rd difference.

(ii) Now suppose that there is another set A^λ, and that C is the same function of the A's that Z is of the X's; e.g., as in the above example, that it is the 3rd difference. Then the h's are the same, so that

$$C = h_\lambda A^\lambda. \tag{2}$$

We can write (1) in the form

$$Z/X^\lambda = h_\lambda,$$

on the understanding that a suffix in a denominator is linked with a similar suffix on the other side and implies an inner multiplication. Similarly we can write (2) as

$$C/A^\lambda = h_\lambda.$$

Equating the two values of h_λ, we have

$$\frac{C}{A^\lambda} = \frac{Z}{X^\lambda} \tag{3}$$

as our way of stating that C is the same linear function of the A's that Z is of the X's.

(iii) Next suppose that a set Y_ρ * is a linear function of the set X^λ, so that each of the Y's is a linear function of the X's. Then we can take Z of (1) to be each of the Y's in turn: but the sets of h's will be different, so that the relation will be of the form

$$Y_\rho = d_{\mu\rho} X^\mu.$$

If there is also a set B_ρ, each element of which is the same

* The set might be called either Y^ρ or Y_ρ; we choose the latter as giving a convenient symbol $d_{\mu\rho}$ for the coefficient of X^μ (see V. 10 (i)).

linear function of the A's that the corresponding element of Y_ρ is of X's, then

$$B_\rho = d_{\mu\rho} A^\mu,$$

the two sets of d's being the same. We therefore have *

$$\frac{B_\rho}{A^\mu} = \frac{Y_\rho}{X^\mu}$$

as a statement of the fact that the B's have the same linear relations to the A's that the Y's have to the X's. This would be the case, for instance, if the Y's were the successive differences of the X's, and the B's were those of the A's according to the same system.

In view of the variety of ways in which we are able to deal with sets according to algebraical laws, it is perhaps permissible to describe this as a case of **direct proportion**, and to say that B_ρ bears the same **ratio** to A^μ that Y_ρ bears to X^μ. This 'ratio', here denoted by $d_{\mu\rho}$, is really the operator that is required to convert A^μ into B_ρ or X^μ into Y_ρ.

(iv) If the linear relation of the B's to the A's is the same as that of the Y's to the X's, then that of the A's to the B's is the same as that of the X's to the Y's (*or*, if the ratio of B_ρ to A^μ is the same as that of Y_ρ to X^μ, then the ratio of A^μ to B_ρ is the same as that of X^μ to Y_ρ); i.e.

$$\text{If } \frac{B_\rho}{A^\mu} = \frac{Y_\rho}{X^\mu}, \text{ then } \frac{A^\mu}{B_\rho} = \frac{X^\mu}{Y_\rho}. \qquad \textbf{(VII. 2. 1)}$$

[Let $B_\rho = d_{\mu\rho} A^\mu$. Then $A^\mu = d^{\mu\rho} B_\rho$. Similarly $X^\mu = d^{\mu\rho} Y_\rho$. Therefore $A^\mu/B_\rho = X^\mu/Y_\rho$.]

* This expression B_ρ/A^μ must not be confused with B_ρ/A^μ as the double set whose typical element is B_r/A^q. The limitation in V. 5 (vi) excludes double sets of this kind from consideration.

(v) The above sets of relations can be expressed by the diagram in Fig. 1. The crosses may be taken as representing either coefficients of the X's in the values of the Y's and of the A's in the values of the B's; or—in virtue of (VII. 2. 1)—coefficients of the Y's in the values of the X's and of the B's in the values of the A's.

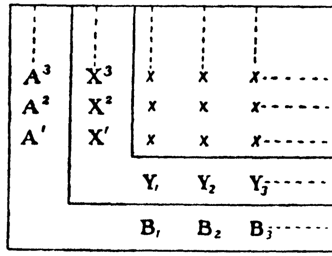

FIG. 1.

(vi) Ratios of the kind considered above can be combined according to the laws of ordinary algebra; e.g.

$$\frac{B_\nu}{A^\mu}\cdot\frac{C^\sigma}{B_\nu}=\frac{C^\sigma}{B_\nu}\cdot\frac{B_\nu}{A^\mu}=\frac{C^\sigma}{A^\mu}. \qquad \textbf{(VII. 2. 2)}$$

The expression on the left-hand side is, of course, an inner product. A special ratio is

$$\frac{A^\nu}{A^\mu}=|^{\nu}_{\mu}=|^{\mu}_{\nu}=\frac{A^\mu}{A^\nu}\cdot \qquad \textbf{(VII. 2. 3)}$$

Example. If $B_\rho/A^\mu = Y_\rho/X^\mu$, prove that $B_\rho/A^\mu.\ X^\mu/Y_\sigma = |^{\sigma}_{\rho}$.

VII. 3. Reciprocal proportion of single sets.—

(i) The other important class of cases is that in which the linear relation (or ratio) of B_ρ to A^μ is the *reciprocal* of that of Y_ρ to X^μ, i.e. in which, B_ρ and A^μ being altered to B^ρ and A_μ,

$$Y_\rho = k_{\mu\rho}X^\mu,\quad B^\rho = k^{\mu\rho}A_\mu.$$

This gives

$$X^\mu = k^{\mu\rho}Y_\rho,$$

so that

$$\frac{B^\rho}{A_\mu}=\frac{X^\mu}{Y_\rho}.$$

We can call this a case of **reciprocal proportion.**

(ii) If the ratio of B^ρ to A_μ is the reciprocal of that of Y_ρ to X^μ, then the ratio of Y_ρ to X^μ is the reciprocal of that of B^ρ to A_μ; i. e.

$$\text{If } \frac{B^\rho}{A_\mu} = \frac{X^\mu}{Y_\rho}, \text{ then } \frac{A_\mu}{B^\rho} = \frac{Y_\rho}{X^\mu}. \qquad \textbf{(VII. 3. 1)}$$

[Let $B^\rho = k^{\mu\rho} A_\mu$. Then $A_\mu = k_{\mu\rho} B^\rho$. Similarly $Y_\rho = k_{\mu\rho} X^\mu$. Therefore $A_\mu / B^\rho = Y_\rho / X^\mu$.]

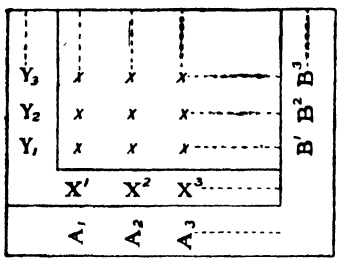

FIG. 2.

(iii) The sets of relations can be expressed by a diagram as in Fig. 2, where the crosses represent coefficients of the Y's in the X's and of the A's in the B's, or of the X's in the Y's and of the B's in the A's.

(iv) The inner products of the reciprocally corresponding sets are equal; i. e.

$$\text{If } \frac{B^\rho}{A_\mu} = \frac{X^\mu}{Y_\rho}, \text{ then } B^\rho Y_\rho = A_\mu X^\mu. \qquad \textbf{(VII. 3. 2)}$$

[Let $B^\rho = k^{\mu\rho} A_\mu$. Then $X^\mu = k^{\mu\rho} Y_\rho$; and therefore $Y_\rho = k_{\mu\rho} X^\mu$. Hence $B^\rho Y_\rho = k^{\mu\rho} A_\mu k_{\nu\rho} X^\nu = k^{\mu\rho} k_{\nu\rho} A_\mu X^\nu = |^\mu_\nu A_\mu X^\nu = A_\mu X^\mu$.]

(v) An interesting case is that in which the Y's are the successive differences of the X's. It will be found that in this case the B's are linear functions of successive sums, and therefore of successive moments, of the A's. In the ordinary system, for instance, which is such as to give $Y_1 = X^1$, $Y_2 = \Delta X^1 = X^2 - X^1$, $Y_3 = \Delta\Delta X^1 = X^3 - 2X^2 + X^1, \ldots,$ we have $X^1 = Y_1$, $X^2 = Y_1 + Y_2$, $X^3 = Y_1 + 2Y_2 + Y_3, \ldots$; and these give $B^1 = -\Sigma A_1$, $B^2 = +\Sigma\Sigma A_2$, $B^3 = -\Sigma\Sigma\Sigma A_3, \ldots,$ the constants in the sums being chosen so that $\Sigma A_{m+1} = 0$, $\Sigma\Sigma A_{m+1} = 0$, $\Sigma\Sigma\Sigma A_{m+1} = 0, \ldots$

VII. 4. Cogredience and contragredience.—(i) In §§ 2 and 3 we have not assumed the existence of any relation between A^λ and X^λ or between B_λ and Y_λ. Where a relation does exist, the important cases are those of *cogredience* and *contragredience*. We start with a set X^λ, and a set A^λ which is derived in a definite way from X^λ, in other words is a function of X^λ. We then take Y_λ to be any linear function of X^λ, and B_λ to be derived from Y_λ in the same way that A^λ is derived from X^λ. Then B_λ is some function of A^λ. Of the cases in which this is a linear function, we are concerned with two special classes:—

(1) If B_ρ/A^μ is always $= Y_\rho/X^\mu$, then A^λ and X^λ are said to be **cogredient.**

(2) If B_ρ/A^μ is always $= X^\mu/Y_\rho$, then A^λ and X^λ are said to be **contragredient.**

Examples of cogredience are given in IX. 6 (x), and of contragredience in VIII. 3 (iv) and IX. 4 (v).

(ii) Instead of saying that A^λ and X^λ are cogredient or contragredient, we might say that A^λ in the one case *varies directly* as X^λ and in the other case *varies reciprocally* as X^λ. When we say that A^λ **varies directly** as X^λ, we mean that, if X^λ is multiplied by any double set involving λ, A^λ is multiplied by the same set: when we say that A^λ **varies reciprocally** as X^λ, we mean that, if X^λ is multiplied by any double set involving λ, A^λ is multiplied by the reciprocal of this set.

(iii) If A^λ and X^λ are $\left\{\begin{matrix}\text{cogredient}\\ \text{contragredient}\end{matrix}\right\}$, and P^λ and A^λ are also $\left\{\begin{matrix}\text{cogredient}\\ \text{contragredient}\end{matrix}\right\}$, then P^λ and X^λ are cogredient; if A^λ and X^λ are $\left\{\begin{matrix}\text{cogredient}\\ \text{contragredient}\end{matrix}\right\}$, but P^λ and A^λ are $\left\{\begin{matrix}\text{contragredient}\\ \text{cogredient}\end{matrix}\right\}$, then P^λ and X^λ are contragredient.

[Suppose, for instance, that A^λ and X^λ are contragredient, and P^λ and A^λ are also contragredient. Then we have relations of the form

$$B_\rho/A^\mu = X^\mu/Y_\rho, \quad Q_\rho/P^\mu = A^\mu/B_\rho,$$

whence, by (VII. 3. 1),

$$Q_\rho/P^\mu = Y_\rho/X^\mu,$$

so that P^λ and X^λ are cogredient.]

VII. 5. Contragredience with linear relation.—

(i) The simplest case of contragredience is that in which the contragredient sets are connected by a linear relation.

(*a*) Suppose that, with the notation of § 3, the relation between A_λ and X^λ is

$$A_\mu = a_{\mu\nu} X^\nu.$$

Then, if Q denotes the inner product (cross-product) in (VII. 3. 2),

$$Q = A_\mu X^\mu = a_{\mu\nu} X^\mu X^\nu = a_{11} X^1 X^1 + (a_{12} + a_{21}) X^1 X^2 + a_{22} X^2 X^2 + (a_{13} + a_{31}) X^1 X^3 + \dots + a_{mm} X^m X^m.$$

Thus Q is a quadratic in $X^1, X^2, X^3, \dots X^m$, i.e. in X^μ. Also, since each of the four sets X^μ, Y_μ, A_μ, B^μ is a linear function of each of the others, Q can be expressed in a good many other ways, e.g. as a quadratic in B^μ, or in the form $k^{\mu\rho} A_\mu Y_\rho$ or $a^{\mu\nu} A_\mu A_\nu$.

(*b*) If the relation between A_μ and X^μ is symmetrical, i.e. if $a_{\mu\rho} = a_{\rho\mu}$, it can be shown that we shall, in addition to (VII. 3. 1), have the further relations

$$\frac{A_\mu}{Y_\rho} = \frac{B^\rho}{X^\mu}, \quad \frac{Y_\rho}{A_\mu} = \frac{X^\mu}{B^\rho}.$$

(ii) Conversely, suppose that we are dealing with a set of quantities or coordinates $X^\mu \equiv (X^1 X^2 \ldots X^m)$, and that we come across an expression

$$Q \equiv a_{\mu\nu} X^\mu X^\nu,$$

where $a_{\mu\nu} = a_{\nu\mu}$. Then we may construct a new set given by

$$X_\mu \equiv a_{\mu\nu} X^\nu,$$

and we shall have

$$Q = X^\mu X_\mu.$$

Now suppose we change the system of coordinates linearly or replace the X's by some linear functions of them. The a's will normally have some definite meaning; and this meaning, though not their actual values, will remain unchanged when the X's are changed. Suppose that, when X^μ becomes Y_μ, X_μ—as based on this meaning of the a's—becomes Y^μ. Then we shall have

$$Y_\mu Y^\mu = X^\mu X_\mu,$$

and also

$$\frac{Y_\mu}{X^\nu} = \frac{X_\nu}{Y^\mu} = \frac{Q}{X^\nu Y^\mu}, \text{ etc.}$$

VII. 6. **Ratios of sets generally.**—(i) The word *ratio* has so far only been used in reference to single sets connected by a linear relation; if the relation between Y_ρ and X^μ is of the form $Y_\rho = d_{\mu\rho} X^\mu$, we call $d_{\mu\rho}$ the ratio of Y_ρ to X^μ, and we call $d^{\mu\rho}$ the reciprocal of this ratio. We can extend the use of the word to sets other than single sets.

(ii) We have already had examples of a ratio which involves a scalar. Thus in § 2 we had relations $Z = h_\lambda X^\lambda$, $C = h_\lambda A^\lambda$, and we said that

$$\frac{C}{A^\lambda} = h_\lambda = \frac{Z}{X^\lambda}.$$

Here we can quite well call h_λ the ratio of C to A^λ or of Z to X^λ;

i.e. it is the ratio of a scalar to a single set. Similarly in § 5 (ii) the statement

$$\frac{Y_\mu}{X^\nu} = \frac{Q}{X^\nu Y^\mu}$$

may be regarded as a statement that the ratio of Q to $X^\nu Y^\mu$ is equal to that of Y_μ to X^ν. In each case the ratio of one set to another is the set by which the latter has to be multiplied in order to obtain the former.

(iii) The more important case is that of the ratio of a variable set of any rank to another variable set of the same rank. If $\mathfrak{A}$ is a variable set of any rank, and $\mathfrak{B}$ is a set which is a function of $\mathfrak{A}$ and is of the same rank, and if the relation between $\mathfrak{B}$ and $\mathfrak{A}$ is of the form

$$\mathfrak{B} = \mathfrak{p}\,\mathfrak{A},$$

where $\mathfrak{p}$ is a constant set whose symbol contains all the suffixes occurring in $\mathfrak{A}$ and $\mathfrak{B}$, then we can call $\mathfrak{p}$ the ratio of $\mathfrak{B}$ to $\mathfrak{A}$, and denote this ratio by $\mathfrak{B}/\mathfrak{A}$.

(iv) In these cases we can continue to speak of the relation of $\mathfrak{B}$ to $\mathfrak{A}$ as linear. Also, by solving the equations, we find that there is a relation of the form

$$\mathfrak{A} = \mathfrak{p}'\mathfrak{B},$$

so that, if $\mathfrak{B}$ is a linear function of $\mathfrak{A}$, then $\mathfrak{A}$ is a linear function of $\mathfrak{B}$. Here, $\mathfrak{p}$ being the ratio of $\mathfrak{B}$ to $\mathfrak{A}$, $\mathfrak{p}'$ is the ratio of $\mathfrak{A}$ to $\mathfrak{B}$, and we can call each ratio the reciprocal of the other.

(v) As an example, suppose that $\mathfrak{A}$ and $\mathfrak{B}$ are of rank 3, and that $\mathfrak{p}$ is the product of three double sets, each of which has inner multiplication with $\mathfrak{A}$. Then the relation might be of the form

$$B^\tau_{\rho\sigma} = a_{\lambda\rho} b_{\mu\sigma} c^{\nu\tau} A^{\lambda\mu}_{\nu}.$$

It is easy to show that in this case

$$A_{\nu}^{\lambda\mu} = a^{\lambda\rho} b^{\mu\sigma} c_{\nu\tau} B_{\rho\sigma}^{\tau}.$$

Thus the reciprocal of the product of the three double sets is the product of their reciprocals.

VII. 7. **Related sets of higher rank.**—With the preceding explanation, there is no difficulty in extending the ideas of equality of ratios, and of related systems of sets, to sets of higher rank.

Suppose, for instance, that we have a set $\mathfrak{A}$ which is a function of three single sets, and a set $\mathfrak{B}$ which is a function of three other single sets. If the three latter sets were functions of the three former, $\mathfrak{B}$ would be a function of $\mathfrak{A}$. The cases analogous to those considered in § 4 would be the cases in which there were linear relations between corresponding sets. Suppose that $X_{\lambda}, Y_{\lambda}, Z_{\lambda}$ are linear functions of U^{λ}, V^{λ}, W^{λ} respectively, e.g.

$$X_{\rho} = a_{\lambda\rho} U^{\lambda}, \quad Y_{\sigma} = b_{\mu\sigma} V^{\mu}, \quad Z_{\tau} = c_{\nu\tau} W^{\nu},$$

that $\mathfrak{A}$ is a certain function (in the most general sense) of U^{λ}, V^{λ}, W^{λ}, and that $\mathfrak{B}$ is the same function of $X_{\lambda}, Y_{\lambda}, Z_{\lambda}$. Then $\mathfrak{B}$ is some function of $\mathfrak{A}$. If we suppose that, when the values of $a_{\lambda\rho}, b_{\mu\sigma}, c_{\nu\tau}$ are made to vary, $\mathfrak{B}$ is always a linear function of $\mathfrak{A}$, and the ratio of $\mathfrak{B}$ to $\mathfrak{A}$ is always compounded of the ratios of X_{λ} to U^{λ}, of Y_{λ} to V^{λ}, and of Z_{λ} to W^{λ}, each taken directly or reciprocally, we get an extension of the cases considered in § 4. Thus we might have

$$B_{\rho\sigma}^{\tau} = a_{\lambda\rho} b_{\mu\sigma} c^{\nu\tau} A_{\nu}^{\lambda\mu},$$

so that

$$\frac{B_{\rho\sigma}^{\tau}}{A_{\nu}^{\lambda\mu}} = a_{\lambda\rho} b_{\mu\sigma} c^{\nu\tau} = \frac{X_{\rho}}{U^{\lambda}} \cdot \frac{Y_{\sigma}}{V^{\mu}} \cdot \frac{W^{\nu}}{Z_{\tau}};$$

each of these latter expressions representing a set of rank 6.

In this particular case we can say that $A_{\nu}^{\lambda\mu}$ is *cogredient* as regards U^{λ} and V^{μ} and *contragredient* as regards W^{ν}; or we can say that it *varies directly* as regards U^{λ} and V^{μ} and *reciprocally* as regards W^{ν}, or that it is *directly proportional* to U^{λ} and V^{μ} and *reciprocally proportional* to W^{ν}.

VIII. DIFFERENTIAL RELATIONS OF SETS

VIII. 1. Derivative of a set.—We have now to consider the cases in which two sets vary together continuously, so that there can be a derivative (differential coefficient) of one with regard to the other, this latter being a single set. The derivative will in all cases be a partial one, since the elements of the single set vary independently.

(i) The simplest case is that of a scalar linear function

$$Z = h_\lambda X^\lambda = h_1 X^1 + h_2 X^2 + \ldots + h_m X^m.$$

Here

$$\frac{\partial Z}{\partial X^p} = h_p.$$

Giving p all values 1 to m, we can write this

$$\frac{\partial Z}{\partial X^\lambda} = h_\lambda = \frac{Z}{X^\lambda};$$

and we can regard h_λ as the derivative of the set Z (which is of rank 0) with regard to the set X^λ.

(ii) Similarly, if

$$Y_\rho = d_{\mu\rho} X^\mu,$$

then *

$$\frac{\partial Y_\rho}{\partial X^\mu} = d_{\mu\rho} = \frac{Y_\rho}{X^\mu}.$$

* It should, however, be noticed that in this statement the sign '=' has not the same meaning in the two places in which it is used. When we say that $\partial Y_\rho / \partial X^\mu = d_{\mu\rho}$, we mean that $\partial Y_r / \partial X^q = d_{qr}$ for all values of q and r: but, when we say that $Y_\rho / X^\mu = d_{\mu\rho}$, we mean that $Y_r = d_{\mu r} X^\mu$ for all values of r.

(iii) A particular case of (ii) is where $Y_\lambda = X^\lambda$. Here $\partial X^r/\partial X^q$ is $= 1$ or 0 according as r is $= q$ or $\neq q$, since the X's vary independently. Hence (see (VII. 2. 3))

$$\frac{\partial X^\rho}{\partial X^\mu} = \frac{X^\rho}{X^\mu} = |^\rho_\mu = |^\mu_\rho = \frac{X^\mu}{X^\rho} = \frac{\partial X^\mu}{\partial X^\rho}. \qquad \textbf{(VIII. 1. 1)}$$

(iv) Taking also sets of higher rank, and not limiting ourselves to linear functions, we see that the derivative of a set with regard to a single set X^λ is a set of rank higher by 1 than that of the original set.

VIII. 2. Derivative of sum or product.

—(i) The derivatives of sums and products of sets follow the ordinary laws of derivatives of sums and products; e.g.

$$\frac{\partial(\mathfrak{B}+\mathfrak{C})}{\partial A^\lambda} = \frac{\partial \mathfrak{B}}{\partial A^\lambda} + \frac{\partial \mathfrak{C}}{\partial A^\lambda}, \qquad \textbf{(VIII. 2. 1)}$$

$$\frac{\partial(\mathfrak{B}\mathfrak{C})}{\partial A^\lambda} = \frac{\partial \mathfrak{B}}{\partial A^\lambda}\mathfrak{C} + \mathfrak{B}\frac{\partial \mathfrak{C}}{\partial A^\lambda}. \qquad \textbf{(VIII. 2. 2)}$$

(ii) As a particular case of this last result, take the scalar quadratic form

$$Q \equiv a_{\mu\nu} X^\mu X^\nu$$

considered in VII. 5. Here, taking (VIII. 1. 1) into account, we have

$$\frac{\partial Q}{\partial X^\mu} = a_{\mu\nu} X^\nu + a_{\mu\nu} X^\mu \frac{\partial X^\nu}{\partial X^\mu} = a_{\mu\nu} X^\nu + a_{\mu\nu} X^\mu |^\nu_\mu = 2 a_{\mu\nu} X^\nu.$$

This can be verified by expressing Q in terms of the X's and finding the partial derivative with regard to X^q in the usual way.

(iii) For an application of this, suppose that

$$a_{\mu\nu} X^\mu X^\nu = b_{\mu\nu} X^\mu X^\nu$$

for all values of the X's. By taking adjoining values of

X^μ, we get the result which is expressed by differentiation, namely

$$2a_{\mu\nu}X^\nu = 2b_{\mu\nu}X^\nu.$$

Differentiating again, or equating coefficients, we find that

$$a_{\mu\nu} = b_{\mu\nu}.$$

VIII. 3. **Derivative of function of a set.**—(i) If z is a function of y, and y is a function of x, then we know that

$$\frac{dz}{dy}\frac{dy}{dx} = \frac{dz}{dx};$$

and, total and partial differential coefficients being in this case identical,

$$\frac{\partial z}{\partial y}\frac{\partial y}{\partial x} = \frac{\partial z}{\partial x}.$$

(ii) Now suppose that B_μ is a function of A^λ, and C^ρ is a function of B_μ. Then, p and r being values of λ and ρ respectively, we know that

$$\frac{\partial C^r}{\partial A^p} = \frac{\partial C^r}{\partial B_1}\frac{\partial B_1}{\partial A^p} + \frac{\partial C^r}{\partial B_2}\frac{\partial B_2}{\partial A^p} + \ldots + \frac{\partial C^r}{\partial B_m}\frac{\partial B_m}{\partial A^p} = \frac{\partial C^r}{\partial B_\mu}\frac{\partial B_\mu}{\partial A^p}.$$

Hence, giving p and r all their values,

$$\frac{\partial C^\rho}{\partial A^\lambda} = \frac{\partial C^\rho}{\partial B_\mu}\frac{\partial B_\mu}{\partial A^\lambda}. \qquad \textbf{(VIII. 3. 1)}$$

The argument applies to sets of higher rank. If, e.g., we are dealing with $C^{\rho\sigma\cdots}$, where the $\sigma\ldots$ relates to aspects independent of B_μ, then

$$\frac{\partial C^{\rho\sigma\cdots}}{\partial A^\lambda} = \frac{\partial C^{\rho\sigma\cdots}}{\partial B_\mu}\frac{\partial B_\mu}{\partial A^\lambda}. \qquad \textbf{(VIII. 3. 2)}$$

(iii) As a particular case (see (VIII. 1. 1)),

$$\frac{\partial C^\rho}{\partial B^\mu}\frac{\partial B^\mu}{\partial C^\sigma} = \frac{\partial C^\rho}{\partial C^\sigma} = |^\rho_\sigma. \qquad \textbf{(VIII. 3. 3)}$$

(iv) Suppose that the relation between B_λ and A^λ is linear, say

$$B_\mu = p_{\lambda\mu} A^\lambda, \quad A^\lambda = p^{\lambda\mu} B_\mu.$$

Then, replacing C^r or C^ρ by $\mathfrak{C}$, we have

$$\frac{\partial \mathfrak{C}}{\partial B_\mu} = \frac{\partial \mathfrak{C}}{\partial A^\lambda} \frac{\partial A^\lambda}{\partial B_\mu} = \frac{\partial \mathfrak{C}}{\partial A^\lambda} p^{\lambda\mu}.$$

Hence

$$\frac{\partial \mathfrak{C}}{\partial B_\mu} \bigg/ \frac{\partial \mathfrak{C}}{\partial A^\lambda} = p^{\lambda\mu} = A^\lambda / B_\mu.$$

Thus A^λ and $\partial \mathfrak{C} / \partial A^\lambda$ are contragredient. We can express this by saying that A^λ and the operator $\partial / \partial A^\lambda$ are contragredient.

(v) The determinant of $\partial B_\mu / \partial A^\lambda$ is the Jacobian of B_μ with regard to A^λ; i.e.

$$\frac{\partial (B_1, B_2, \ldots B_m)}{\partial (A^1, A^2, \ldots A^m)} = \left| \frac{\partial B_r}{\partial A^q} \right|. \qquad \textbf{(VIII. 3. 4)}$$

From (VIII. 3. 1), taken with (V. 10. 5), we have the ordinary formula for the product of two Jacobians:

$$\left| \frac{\partial B_r}{\partial A^q} \right| \times \left| \frac{\partial C^r}{\partial B_q} \right| = \left| \frac{\partial B_\mu}{\partial A^q} \frac{\partial C^r}{\partial B_\mu} \right| = \left| \frac{\partial C^r}{\partial A^q} \right|. \qquad \textbf{(VIII. 3. 5)}$$

VIII. 4. Transformation of quadratic form to sum of squares.

—(i) For an example of a Jacobian, take the case in which a quadratic form is to be expressed as the sum of the squares of linear functions of the variables. Let the quadratic form be

$$Q \equiv a_{\mu\rho} X^\mu X^\rho, \qquad (1)$$

where

$$a_{\mu\rho} = a_{\rho\mu}.$$

Let

$$Y_\sigma \equiv b_{\mu\sigma} X^\mu \qquad (2)$$

be a set of linear functions of the X's, so that

$$X^{\mu} = b^{\mu\sigma} Y_{\sigma}. \tag{3}$$

Suppose that the b's are chosen so that (1) gives

$$Q = Y_1 Y_1 + Y_2 Y_2 + \ldots + Y_m Y_m = Y_{\sigma} Y_{\sigma}. \tag{4}$$

Then, by substitution from (2),

$$Q = b_{\mu\sigma} X^{\mu} . b_{\rho\sigma} X^{\rho} = b_{\mu\sigma} b_{\rho\sigma} X^{\mu} X^{\rho}.$$

Hence, by comparison with (1) (see § 2 (iii)),

$$a_{\mu\rho} = b_{\mu\sigma} b_{\rho\sigma}. \tag{5}$$

The Jacobian which we should usually want to find is that of X^{μ} with regard to Y_{σ}, i. e.

$$J \equiv \frac{\partial(X^1, X^2, \ldots X^m)}{\partial(Y_1, Y_2, \ldots Y_m)} = \left|\frac{\partial X^q}{\partial Y_r}\right|.$$

By (3),

$$\frac{\partial X^{\mu}}{\partial Y_{\sigma}} = b^{\mu\sigma};$$

and therefore

$$J = | b^{qr} |. \tag{6}$$

But (5) gives

$$a^{\mu\rho} = b^{\mu\sigma} b^{\rho\sigma},$$

and therefore

$$| a^{qr} | = | b^{q\sigma} b^{r\sigma} | = | b^{qr} | \times | b^{rq} | = JJ.$$

Hence, combining this with (6),

$$J = | b^{qr} | = \{ | a^{qr} | \}^{\frac{1}{2}} = 1/\{ | a_{qr} | \}^{\frac{1}{2}}. \tag{7}$$

Similarly the Jacobian of Y_{σ} with regard to X^{μ} is

$$J' \equiv \frac{\partial(Y_1, Y_2, \ldots Y_m)}{\partial(X^1, X^2, \ldots X^m)} = \frac{1}{J} = | b_{qr} | = \{ | a_{qr} | \}^{\frac{1}{2}}$$
$$= 1/\{ | a^{qr} | \}^{\frac{1}{2}}. \tag{8}$$

There are a good many different ways of expressing Q as in (4), but they all give the same two Jacobians.

(ii) Hence we easily obtain the value of the multiple integral

$$D \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}Q}\, dX^1 dX^2 \dots dX^m. \tag{9}$$

Using $\left(\int\right)^m$ to denote $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots (m$ times), we have

$$D = \left(\int\right)^m e^{-\frac{1}{2}Q} J dY_1 dY_2 \dots dY_m$$

$$= J\int_{-\infty}^{\infty} e^{-\frac{1}{2}Y_1 Y_1} dY_1 \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}Y_m Y_m} dY_m$$

$$= \{|\, a^{qr} \,|\}^{\frac{1}{2}} (2\pi)^{\frac{1}{2}m}. \tag{10}$$

(iii) We shall require, in the next chapter, the value of

$$N/D,$$

where

$$N \equiv \left(\int\right)^m \tfrac{1}{2} X^q \frac{\partial Q}{\partial X^r} e^{-\frac{1}{2}Q} dX^1 dX^2 \dots dX^m \tag{11}$$

and D is as above. We find the value of N, as we have found that of D, by expressing everything in terms of Y's. By (2) and (4), and (VIII. 3. 1) and § 2 (ii),

$$\frac{\partial Q}{\partial X^r} = \frac{\partial Y_\tau}{\partial X^r}\frac{\partial Q}{\partial Y_\tau} = 2 b_{r\tau} Y_\tau;$$

and therefore, by (3),

$$\tfrac{1}{2} X^q \frac{\partial Q}{\partial X^r} = b^{q\sigma} b_{r\tau} Y_\sigma Y_\tau,$$

$$N = b^{q\sigma} b_{r\tau} \left(\int\right)^m Y_\sigma Y_\tau e^{-\frac{1}{2}Q} dX^1 dX^2 \dots dX^m$$

$$= J b^{q\sigma} b_{r\tau} \left(\int\right)^m Y_\sigma Y_\tau e^{-\frac{1}{2}Q} dY_1 dY_2 \dots dY_m. \tag{12}$$

Let us write

$$M_t \equiv J b^{q\sigma} \left(\int\right)^m Y_\sigma Y_t e^{-\frac{1}{2}Q} dY_1 dY_2 \dots dY_m, \tag{13}$$

so that

$$N = b_{r\tau} M_\tau. \tag{14}$$

Then M_t consists of m terms due to the m values of σ. Since Q is the sum of the squares of the Y's, the only term which counts in the integration is that for which $\sigma = t$. Also we know that

$$\int_{-\infty}^{\infty} Y_t Y_t e^{-\frac{1}{2}Y_t Y_t} dY_t = \int_{-\infty}^{\infty} e^{-\frac{1}{2}Y_t Y_t} dY_t.$$

Hence it follows that

$$\begin{aligned} M_t &= Jb^{qt}\left(\int\right)^m e^{-\frac{1}{2}Q} dY_1 dY_2 \ldots dY_m \\ &= b^{qt} D; \end{aligned}$$

and therefore, by (14),

$$\begin{aligned} N/D &= b_{r\tau} b^{q\tau} \\ &= |^q_r. \end{aligned} \tag{15}$$

IX. EXAMPLES FROM THE THEORY OF STATISTICS *

IX. 1. **Preliminary.**—(i) The special feature of a statistical set

$$X_\lambda \equiv (X_1 X_2 X_3 \dots X_m),$$

of the kind which we have to consider in this chapter, is that each X has one only of a very large number of actual or possible values, which together constitute the field from which the X is drawn; and the fundamental facts with which we are concerned are the relative frequencies of occurrence of the various possible combinations formed by taking an X from each of the m fields. Thus the X's are variables, and the expression for the relative frequency of joint occurrence of a particular set $(X_1 X_2 \dots X_m)$ involves the X's of the set, with certain constants. In a large class of cases the constants depend on the mean values of the X's and the mean squares and products of their deviations from their respective means. It is to these cases that the new notation is specially applicable. It may be that some of the X's are drawn from the same field; we shall proceed as if the fields were all different, but this does not affect the validity of the reasoning.

(ii) We shall only consider two kinds of cases.

(*a*) The first kind of case is where our statistical information relates to a large number of individuals, and $X_1 X_2 X_3 \dots$ are the measures of specified attributes, such as height, head-length, chest-expansion, intelligence, etc.,

* I have dealt with the problems of this chapter in as general terms as possible. The explanations in small print may help to show the statistical student the way in which the problems actually arise.

of any individual. Here the 'variability' of any X has reference to the different values that it takes for different individuals. The frequency of joint occurrence of particular values of $X_1 X_2 X_3 \ldots$ may be a complicated function of these values and of certain constants which are to be determined.

(*b*) The other kind of case is that in which the question is one of 'graduation' or 'fitting'. Here $X_1 X_2 X_3 \ldots$ are observed values of different quantities (e.g. rates of mortality at successive ages), or are the results of observation of one quantity at different times or by different observers. The 'variability' of any particular X lies in the fact that the observed value contains an unknown error; and our treatment is based on the assumption that a relation of a particular kind holds between the true X's.

In cases of this latter kind it should be noticed that the only things which we treat as variables are the errors in the X's. We may take, as the typical case, the observed rates of mortality $X_1 X_2 X_3 \ldots$ at ages $t_1 t_2 t_3 \ldots$ If we denote the true rates by $\xi_1 \xi_2 \xi_3 \ldots$, then the assumption which we make is really an assumption that ξ is a certain function, with constants to be determined, of t. So far as this function is concerned, ξ and t might be called variables. But, for our purpose, they are not variables. We are concerned with the fixed values $t_1 t_2 t_3 \ldots$ and the corresponding fixed, though unknown, values $\xi_1 \xi_2 \xi_3 \ldots$; the real variables are the differences between the observed values $X_1 X_2 X_3 \ldots$ and the true values $\xi_1 \xi_2 \xi_3 \ldots$.

(iii) There are two reasons why the same mathematical methods apply to subjects so different as relativity and statistical theory. One is that the number of X's in a statistical set may be very large: in the second kind of case mentioned in (ii) it may be as many as 20 or 30. The need of a condensed notation is therefore even greater than for relativity, where the number of dimensions does not exceed 4. The other reason is that, as in the relativity theory, we are, to a certain extent in the first kind of case, and very largely in the second kind of case, concerned with sets constructed from the original sets by means of linear relations.

(iv) We denote the mean product of the deviations of X_q and X_r by $(X_q . X_r)$, or, more briefly, by f_{qr}; i.e.

$$f_{qr} \equiv (X_q . X_r) \equiv \text{mean value of } (X_q - \text{mean } X_q)(X_r - \text{mean } X_r).$$

It must be clearly understood that, though $(X_q . X_r)$ depends on q and r, it relates to the complete fields from which X_q and X_r are drawn, and, for any particular values of q and r, is not a variable, like X_q and X_r, but a constant.

(v) Although we have defined $(X_q . X_r)$ as a mean value, our dealings with it ultimately depend on the algebraical laws which it follows. These are, first that it satisfies the ordinary laws of multiplication of two expressions X_q and X_r, i.e. that, c being a constant as regards the X's,

$$(X_q . X_r) = (X_r . X_q), \quad (X_q . (X_r + X_s)) = (X_q . X_r) + (X_q . X_s), \quad (X_q . cX_r) = c(X_q . X_r); \qquad \textbf{(IX. 1. 1)}$$

and next that, if u is any linear function of the X's, then $(u . u)$ is positive unless $u = 0$, i.e. that

$$(u . u) > 0 \quad \text{if} \quad u \neq 0. \qquad \textbf{(IX. 1. 2)}$$

It is clear that $(u . u) = 0$ if $u = 0$; for (IX. 1. 1) gives, by putting $c = 0$,

$$(X_q . 0) = 0, \qquad \textbf{(IX. 1. 3)}$$

whence $(u . 0) = 0$ follows by means of (IX. 1. 1) (cf. (vi) below). These are the only properties we shall use; and our results will therefore be true for any meaning of $(X_q . X_r)$ that satisfies these laws, provided, of course, that

$(X_q . X_r)$ is a constant as regards all the X's. And, conversely, we shall only be dealing with sets for which $(X_q . X_r)$ has a meaning and a value for each value of q with each value of r, and satisfies these laws. It will be assumed that the values of $(X_q . X_r)$ are known; or, at any rate, that our results are final when expressed in terms of these values. Whether we are dealing with mean squares or mean products or not, we can call $(X_q . X_r)$ the (.) of X_q and X_r.

In the first kind of case mentioned in (ii) $(X_q . X_q)$ would usually be the mean square of deviation of X_q from its mean, i. e. would be the square of the standard deviation; and $(X_q . X_r)$ would be the mean product of deviations of X_q and X_r from their respective means, and would therefore, in the case of normal correlation, be equal to the product of the standard deviations of X_q and X_r multiplied by their coefficient of correlation. In the second kind of case $(X_q . X_q)$ is the mean square of error of X_q, and $(X_q . X_r)$ is the mean product of errors of X_q and X_r.

(vi) It follows from (IX. 1. 1) that

$$(a_\mu X_\mu . X_r) = a_\mu (X_\mu . X_r), \quad (a_\mu X_\mu . b_\rho X_\rho) = a_\mu b_\rho (X_\mu . X_\rho). \tag{IX. 1. 4}$$

(vii) If X_λ is a set of the kind considered in this section, then so also is any other set which is a linear function of X_λ. Suppose, for instance, that

$$Y^\mu \equiv b^{\lambda\mu} X_\lambda .$$

Then, by (IX. 1. 4),

$$(Y^q . Y^r) = (b^{\lambda q} X_\lambda . b^{\nu r} X_\nu) = b^{\lambda q} b^{\nu r} (X_\lambda . X_\nu),$$

which has a definite meaning for each value of q with each value of r, and can be shown to satisfy the laws stated in (v).

(viii) Our results are also subject to the condition that none of the determinants of the double sets we have to deal with are 0; i.e. that $|(X_q . X_r)| \neq 0$, whether the range of values of q and r is the whole of the range 1 to m or a part of it only.

IX. 2. **Mean-product set.**—(i) The quantities $(X_q . X_r)$ constitute a symmetrical double set

$$f_{\mu\rho} \equiv f_{\rho\mu} \equiv \begin{Bmatrix} (X_1 . X_1) & (X_1 . X_2) & (X_1 . X_3) \ldots (X_1 . X_m) \\ (X_1 . X_2) & (X_2 . X_2) & (X_2 . X_3) \ldots (X_2 . X_m) \\ \vdots & \vdots & \vdots \qquad \vdots \\ (X_1 . X_m) & (X_2 . X_m) & (X_3 . X_m) \ldots (X_m . X_m) \end{Bmatrix}. \qquad \text{(IX. 2. A)}$$

We call this the **mean-product set.**

(ii) Corresponding to this there is a reciprocal set $f^{\mu\rho} = f^{\rho\mu}$ given by

$$f^{\lambda\rho} f_{\mu\rho} = f^{\lambda\rho} f_{\rho\mu} = f^{\rho\lambda} f_{\mu\rho} = f^{\rho\lambda} f_{\rho\mu} = |^{\lambda}_{\mu}. \qquad \text{(IX. 2. 1)}$$

(iii) If $(X_p . X_s) = 0$, we can for the purpose of this chapter describe X_p and X_s as *statistically independent.* Strictly speaking, this is a loose description, since the complete statistical independence of two variables X_p and X_s would imply a good deal more than that the mean product of their deviations from their respective means should be 0. But we are only concerned, here, with mean squares and mean products.

(iv) The simplest class of cases—from the point of view of algebraical treatment—consists of those cases in which the X's are statistically independent of one another and the mean square of deviation of each X is 1. We can express this by duplicating the set, thus:

$$\begin{matrix} X_1 & X_2 & X_3 \ldots X_m \\ X_1 & X_2 & X_3 \ldots X_m \end{matrix}$$

and saying that the (.) of corresponding elements of the two sets is 1, and that the (.) of elements which do not correspond is 0.

For a set of this kind we have

$$f_{qr} = (X_q \,.\, X_r) = |^q_r,$$

so that the mean-product set is the unit set. It follows that in this class of cases the mean-product set and its reciprocal set are identical.

(v) The next kind of case, in point of simplicity, is that in which the X's are statistically independent of one another but the mean squares of deviation are not all 1. This would be the case, for instance, if the X's were independent observations, not all of the same weight, of a single quantity. For practical purposes a case of this kind can be brought under (iv) by expressing each X in terms of its standard deviation (square root of mean square of deviation) as the unit.

(vi) There is also an important class of cases in which the X's fall into two groups, such that each X in one group is statistically independent of each X in the other group. If, as in VI. **11** (i), we denote the two groups by A_α and A_ϕ, then the property is that

$$(A_\alpha \,.\, A_\phi) = 0.$$

IX. 3. Conjugate sets.—(i) When a set X_λ is not of the simple kind described in § 2 (iv), we shall find it useful to introduce another set X^λ which (1) is a linear function of X_λ and (2) is such that, if we place the sets opposite one another, thus:

$$\begin{matrix} X_1 & X_2 & X_3 \ldots X_m \\ X^1 & X^2 & X^3 \ldots X^m \end{matrix}$$

the (.) of corresponding elements of the two sets is 1, and that of elements which do not correspond is 0. This new set X^λ is said to be **conjugate** to X_λ.

(ii) The second of the above conditions can be written

$$(X^p \,.\, X_q) = |^p_q, \qquad \textbf{(IX. 3. 1)}$$

or

$$(X^\lambda . X_\mu) = |^\lambda_\mu. \qquad \text{(IX. 3. 2)}$$

(iii) Each element X^p of the new set will contain m terms, with m coefficients which have to be determined from the m equations given by

$$(X^p . X_\mu) = |^p_\mu.$$

There are altogether m^2 equations to determine X^λ. By regrouping these according to the values of μ in $(X^\lambda . X_\mu)$, we see that if X^λ is conjugate to X_λ then X_λ is conjugate to X^λ. This is in fact evident from the symmetry of (IX. 3. 2).

(iv) To express X^λ in terms of X_λ, or X_λ in terms of X^λ, let us first take W to be any linear function of X_λ, say

$$W = a^\mu X_\mu. \qquad (1)$$

Then we want to find an expression for a^μ.

As we know the value of $(X^\lambda . X_\mu)$, we take the (.) of W and X^λ. By (IX. 1. 4) and (IX. 3. 2) we find that

$$(W . X^\lambda) = (a^\mu X_\mu . X^\lambda) = a^\mu (X_\mu . X^\lambda) = a^\mu |^\lambda_\mu = a^\lambda,$$

whence

$$a^\mu = (W . X^\mu).$$

Substituting in (1),

$$W = (W . X^\mu) X_\mu. \qquad \text{(IX. 3. 3)}$$

Taking W to be each element of X^λ in turn, we have

$$X^\lambda = (X^\lambda . X^\mu) X_\mu. \qquad \text{(IX. 3. 4)}$$

We do not yet know the values of $(X^\lambda . X^\mu)$. But, if we had started with W as a linear function of X^λ, we should similarly have got

$$W = (W . X_\mu) X^\mu, \qquad \text{(IX. 3. 5)}$$

whence

$$X_\lambda = (X_\lambda . X_\mu) X^\mu. \qquad \text{(IX. 3. 6)}$$

Writing this in the form

$$X_\lambda = f_{\lambda\mu} X^\mu, \qquad \textbf{(IX. 3. 7)}$$

we have, by reciprocation,

$$X^\mu = f^{\lambda\mu} X_\lambda, \qquad \textbf{(IX. 3. 8)}$$

which gives X^μ in terms of X_λ. Further, comparing (IX. 3. 8) with (IX. 3. 4), we see that

$$\text{If } (X_\lambda \,.\, X_\mu) \equiv f_{\lambda\mu}, \text{ then } (X^\lambda \,.\, X^\mu) = f^{\lambda\mu}; \qquad \textbf{(IX. 3. 9)}$$

and, of course, the converse also holds. This result is dependent on the assumption, made in § 1 (viii), that $| f_{qr} |$ is not 0.

(v) We could have obtained (IX. 3. 4) and (IX. 3. 6) in fewer steps by considering the set as a whole instead of element by element. If we assume

$$X^\lambda = a^{\lambda\mu} X_\mu,$$

then we get

$$(X^\lambda \,.\, X^\nu) = (a^{\lambda\mu} X_\mu \,.\, X^\nu) = a^{\lambda\mu} |_\mu^\nu = a^{\lambda\nu},$$

so that

$$a^{\lambda\mu} = (X^\lambda \,.\, X^\mu).$$

This gives (IX. 3. 4); and (IX. 3. 6) can be obtained in the same way.

(vi) We can write (IX. 3. 3) in the form

$$W/X_\mu = (W \,.\, X^\mu); \qquad \textbf{(IX. 3. 10)}$$

and similarly from (IX. 3. 5)

$$W/X^\mu = (W \,.\, X_\mu). \qquad \textbf{(IX. 3. 11)}$$

(vii) If the set X_λ is of the special kind considered in § 2 (iv), i.e. is such that

$$(X_\lambda \,.\, X_\mu) = |_\mu^\lambda,$$

then

$$X_\lambda = (X_\lambda \,.\, X_\mu)\, X^\mu = |_\mu^\lambda X^\mu = X^\lambda,$$

so that the set is identical with its conjugate. The set is said to be **self-conjugate.**

(viii) In a case of the kind mentioned in § 2 (v), where $f_{pq} = 0$ if $q \neq p$, but f_{pp} is not necessarily $= 1$, it may be shown that $f^{pp} = 1/f_{pp}$, and $f^{pq} = 0$ if $q \neq p$, so that $X^p = X_p/f_{pp}$.

(ix) Next consider a case of the kind mentioned in § 2 (vi), where X_λ consists of two portions, the elements in each portion being independent of those in the other portion. As before, we take one portion to consist of the first k elements, and the other of the remaining $m-k$, and we denote the two portions by X_a and X_ϕ. Then the special property is that

$$f_{a\phi} \equiv (X_a \,.\, X_\phi) = 0, \tag{1}$$

whence, as in VI. 11 (iii), it follows that

$$f^{a\phi} = 0. \tag{2}$$

Breaking up the right-hand side of (IX. 3. 7) into two portions, we get, according as λ belongs to the first or to the second portion, the two separate results

$$X_a = f_{a\gamma} X^\gamma, \quad X_\phi = f_{\phi\psi} X^\psi. \tag{3}$$

Similarly, from (IX. 3. 8),

$$X^a = f^{a\gamma} X_\gamma, \quad X^\phi = f^{\phi\psi} X_\psi. \tag{4}$$

In finding $f^{a\gamma}$ and $f^{\phi\psi}$ from $f_{\lambda\mu}$, it is (see VI. 11 (iii)) immaterial whether we take the set $f_{\lambda\mu}$ as a whole or the sets $f_{a\gamma}$ and $f_{\phi\psi}$ separately, so that (4) may equally well be written (see VI. 11 (ii))

$$X^a = (f^{a\gamma})_k X_\gamma, \quad X^\phi = [f^{\phi\psi}]_{m-k} X_\psi. \tag{5}$$

IX. 4. Conjugate sets with linear relations.—

(i) Let Y_ρ be any set which is a linear function of X_μ, and therefore of X^μ; and let Y^ρ be its conjugate set. Then, taking W in (IX. 3. 10) to be each element of Y_ρ in turn, we have

$$Y_\rho/X_\mu = (Y_\rho \,.\, X^\mu). \tag{IX. 4. 1}$$

Similarly

$$Y_\rho/X^\mu = (Y_\rho \,.\, X_\mu), \quad Y^\rho/X_\mu = (Y^\rho \,.\, X^\mu), \quad Y^\rho/X^\mu = (Y^\rho \,.\, X_\mu). \tag{IX. 4. 2}$$

(ii) By combining ratios, we get such results as

$$(E_\sigma \, . \, X^\mu) = E_\sigma/X_\mu = E_\sigma/Y_\rho \, . \, Y_\rho/X_\mu = (E_\sigma \, . \, Y^\rho)\,(Y_\rho \, . \, X^\mu). \tag{IX. 4. 3}$$

(iii) To find Y^ρ, suppose that

$$Y_\rho = b_{\rho\mu} X_\mu. \tag{1}$$

Let the conjugate set be

$$Y^\rho = k_{\rho\mu} X^\mu.$$

Then

$$|^\rho_\sigma = (Y^\rho \, . \, Y_\sigma) = (k_{\rho\mu} X^\mu \, . \, b_{\sigma\nu} X_\nu) = |^\mu_\nu k_{\rho\mu} b_{\sigma\nu} = k_{\rho\mu} b_{\sigma\mu};$$

whence, by reciprocation,

$$k_{\rho\mu} = b^{\sigma\mu} |^\rho_\sigma = b^{\rho\mu}.$$

Thus the conjugate set is

$$Y^\rho = b^{\rho\mu} X^\mu. \tag{2}$$

(iv) Similarly, if

$$Y_\epsilon = b_{\epsilon\nu} c_{\nu\rho} d_{\rho\sigma} X_\sigma,$$

then

$$Y^\epsilon = b^{\epsilon\nu} c^{\nu\rho} d^{\rho\sigma} X^\sigma,$$

etc.; in other words, the conjugate of an inner product is the inner product of the $\left\{\begin{matrix}\text{reciprocals}\\ \text{conjugate}\end{matrix}\right\}$ of the factors. [Let us write

$$C^\epsilon \equiv b^{\epsilon\eta} c^{\eta\theta} d^{\theta\lambda} X^\lambda.$$

Then

$$(C^\epsilon . Y_s) = b^{\epsilon\eta} b_{s\nu} c^{\eta\theta} c_{\nu\rho} d^{\theta\lambda} d_{\rho\sigma} (X^\lambda . X_\sigma)$$
$$= |^\epsilon_s,$$

since $(X^\lambda . X_\sigma) = |^\lambda_\sigma$. Hence

$$C^\epsilon = Y^\epsilon.$$

We might, alternatively, have deduced this from (iii) by means of (VI. 10. 2).]

(v) From (IX. 4. 1)

$$Y_\rho/X_\mu = (Y_\rho \, . \, X^\mu) = (X^\mu \, . \, Y_\rho) = X^\mu/Y^\rho;$$

and therefore, by VII. 3 (iv),

$$Y_\rho Y^\rho = X_\mu X^\mu. \tag{IX. 4. 4}$$

Thus conjugate sets are contragredient (VII. 4); and the inner product of a set and its conjugate is the same for all linearly related sets. If we denote this inner product by Q, then the sets X_μ, X^μ, Y_μ, Y^μ are connected by four relations of the form

$$\frac{X_\mu}{Y^\rho} = \frac{Y_\rho}{X^\mu} = (X_\mu \,.\, Y_\rho) = \frac{Q}{X^\mu Y^\rho}. \qquad \textbf{(IX. 4. 5)}$$

We can express Q in such forms as

$$Q = X_\mu X^\mu = (X_\lambda \,.\, X_\mu)\, X^\lambda X^\mu = (X^\lambda \,.\, X^\mu)\, X_\lambda X_\mu,$$

the last of which, when written out in full, is

$$f^{11} X_1 X_1 + 2 f^{12} X_1 X_2 + f^{22} X_2 X_2 + 2 f^{13} X_1 X_3 + \ldots + f^{mm} X_m X_m,$$

or in more general forms such as

$$Q = (C^\lambda \,.\, D^\mu)\, C_\lambda D_\mu,$$

where C_λ and D_λ are any linear functions of X_λ. It must be remembered that the invariability of Q only applies for the particular values $X_1, X_2, \ldots X_m$. If there were a different set of X's there would (in general) be a different Q.

IX. 5. The frequency-quadratic.

—(i) In most of the cases we are considering, whether of the first or of the second kind mentioned in § 1 (ii), the frequency of joint occurrence of values lying within limits

$$X_1 \pm \tfrac{1}{2} d X_1,\ X_2 \pm \tfrac{1}{2} d X_2, \ldots X_m \pm \tfrac{1}{2} d X_m$$

is proportional to

$$e^{-\frac{1}{2}P} dX_1 dX_2 \ldots dX_m,$$

where, if $x_1, x_2, \ldots x_m$ are the deviations of $X_1, X_2 \ldots X_m$ from their respective means, P is of the form

$$P \equiv a^{11} x_1 x_1 + 2 a^{12} x_1 x_2 + a^{22} x_2 x_2 + 2 a^{13} x_1 x_3 + \ldots + a^{mm} x_m x_m. \quad (1)$$

In our notation this becomes

$$P \equiv a^{\lambda\mu} x_\lambda x_\mu ; \qquad (2)$$

where

$$a^{\lambda\mu} = a^{\mu\lambda}. \qquad (3)$$

(ii) Let us write

$$E^\lambda \equiv a^{\lambda\mu}x_\mu \equiv a^{\lambda 1}x_1 + a^{\lambda 2}x_2 + \ldots + a^{\lambda m}x_m$$
$$= \tfrac{1}{2}\partial P/\partial x_\lambda. \quad (4)$$

Then it can be shown (see VIII. 4 (iii)) that

$$\text{mean value of } E^p x_q = \begin{Bmatrix} 1 \text{ if } q = p \\ 0 \text{ if } q \neq p \end{Bmatrix}. \quad (5)$$

Taking $(B_p . C_q)$ to mean, for these cases, the mean product of B_p and C_q, it follows from (5) that E^λ is conjugate to x_λ, i.e.

$$E^\lambda = x^\lambda. \quad (6)$$

Hence, by (4),

$$x^\lambda = a^{\lambda\mu}x_\mu,$$

and therefore, by (IX. 3. 10), or by comparison with (IX. 3. 4),

$$a^{\lambda\mu} = x^\lambda/x_\mu = (x^\lambda . x^\mu). \quad (7)$$

(iii) Thus the a's in the expression for P given in (1) are the mean squares and mean products of the elements of the conjugate set. Similarly, if we expressed P in terms of the conjugate set, the coefficients would be the mean squares and mean products of the elements of the original set; i.e.

$$\left.\begin{aligned} &= a^{\lambda\mu}x_\lambda x_\mu = x_\lambda x^\lambda = a_{\lambda\mu}x^\lambda x^\mu \\ &= a_{11}x^1x^1 + 2a_{12}x^1x^2 + a_{22}x^2x^2 + 2a_{13}x^1x^3 + \ldots + a_{mm}x^mx^m, \end{aligned}\right\} \quad (8)$$

here

$$a_{\lambda\mu} = (x_\lambda . x_\mu). \quad (9)$$

(iv) Take, for example, the case of two quantities X_1, X_2, whose standard deviations and coefficient of correlation are c_1, c_2, and r. Then it is well known that

$$P = \left(\frac{x_1x_1}{c_1c_1} - 2\frac{rx_1x_2}{c_1c_2} + \frac{x_2x_2}{c_2c_2}\right)\Big/(1-rr).$$

This gives, for the members of the conjugate set,

$$x^1 = \left(\frac{x_1}{c_1} - \frac{rx_2}{c_2}\right) \times \frac{1}{c_1(1-rr)}, \quad x^2 = \left(-\frac{rx_1}{c_1} + \frac{x_2}{c_2}\right) \times \frac{1}{c_2(1-rr)}.$$

It is easily verified that the mean square of x^1 is equal to the coefficient of $x_1 x_1$ in P, and so on.

(v) If we express the x's linearly in terms of a new set y_λ, the value of P will remain unaltered. We can put this differently as follows. Suppose that y_λ is any linear function of x_λ. Let $h_{\lambda\mu} \equiv (y_\lambda \cdot y_\mu)$; and let $h^{\lambda\mu}$ be the reciprocal of $h_{\lambda\mu}$. Then, if we write

$$y^\lambda = h^{\lambda\rho} y_\rho,$$

we shall have

$$y^\lambda y_\lambda = P = x^\lambda x_\lambda.$$

(vi) The $(x_q \cdot x_q)$ or $(x_q \cdot x_r)$ which we have so far been considering is the mean square of x_q, or the mean product of x_q and x_r, without regard to the values that each of the other x's may have; i.e. the mean square or mean product taken for all possible values of these other x's according to their relative frequencies. We may also want to know what happens when some of the x's have definite values ascribed to them and are not allowed to vary from these values. In these cases we follow the principle of VI. 11 (ii). Suppose that all the x's after x_k are fixed. Let the x's up to x_k be denoted by x_a or x_β etc. Then our methods apply to the set of order k formed by these x's. The principle, therefore, is as follows. Suppose we want to study the variation of the k quantities x_a when the $m-k$ quantities x_ϕ are fixed. We first construct the mean-product set $f_{\mu\rho}$ $(=a_{\mu\rho})$ for the m quantities x_a and x_ϕ; then construct the reciprocal set $f^{\mu\rho}$ $(=a^{\mu\rho})$, the elements of which are the coefficients in the terms in P; then take out the partial set $f^{a\gamma}$ corresponding to x_a; and then find the set $(f_{a\gamma})_k$ which is the reciprocal of $f^{a\gamma}$. The result is the mean-product set of x_a when x_ϕ is fixed. (This is a well-known theorem, but is usually expressed in terms of determinants.)

(vii) If the partial sets x_a and x_ϕ in (vi) are independent, so that $(x_a . x_\phi) = 0$, the elements of the mean-product set outside the portions corresponding to $(x_a . x_\gamma)$ and $(x_\phi . x_\psi)$ will all be 0, and (see § 3 (ix)) the values of $(f_{a\gamma})_k$ will be the same whether we construct them from the whole set or from the partial set. This is otherwise obvious; for, if the x's of x_a vary independently of those of x_ϕ, they vary in the same way whether the latter are fixed or vary.

IX. 6. **Criteria for improved values.**—(i) Our next problem, considered in this and the following section, is that of reduction of error, in a case of the second kind mentioned in § 1 (ii). We have a set of quantities

$$D_\lambda \equiv (D_1 \quad D_2 \quad D_3 \ldots D_m)$$

which contain errors; the mean products of error being

$$d_{\lambda\mu} \equiv (D_\lambda . D_\mu). \qquad \textbf{(IX. 6. A)}$$

The whole set D_λ consists of two portions

$$D_a \equiv (D_1 \quad D_2 \ldots D_k), \quad D_\phi \equiv (D_{k+1} \quad D_{k+2} \ldots D_m).$$

All the D's are the results (direct or indirect) of observation; but the true values of the D_ϕ are negligible (within the degree of accuracy to which we are working), and, if U is any one of the D's, or any linear function of them, we can add to it any linear function* of D_ϕ without altering, except to a negligible extent, the true value which it represents. If the sum of U and an indeterminate linear function of D_ϕ is represented by U', the problem of *reduction of error* is to determine this linear function so as to make

* This is, of course, an incomplete statement. We could replace U by any function of U and D_ϕ which would be equal to U if the D_ϕ were all 0. But we are only considering linear functions.

$(U'.U')$ a minimum. The resulting value of U' is called the **improved value** of U, and will be denoted by IU. The elements of D_ϕ are called the **auxiliaries**. We can replace a by $\beta, \gamma, \ldots$, and ϕ by $\chi, \psi, \ldots$, as occasion requires.

(ii) The way in which this problem arises is as follows. We start with a set of observed quantities $X_1, X_2, \ldots X_m$, which correspond to a series of values of some other quantity t at equal or unequal intervals; the X's might, e.g., be rates of mortality at different ages. The X's contain errors; and our fundamental assumption, based on general experience and on inspection of the particular *data*, is that the true values are so nearly of the form (in ordinary notation) $c_0 + c_1 t + c_2 t^2 + \ldots + c_{k-1} t^{k-1}$ that their differences (divided differences if the values of t are at unequal intervals) after the $(k-1)$th are negligible. We may therefore add to each X any linear function of these differences, which are what we are calling $D_{k+1}, D_{k+2}, \ldots D_m$. The problem is to determine the coefficients in this linear function so that the mean square of error of the sum of the X and the linear function shall be a minimum.

(iii) We have first to see what relations hold between the two portions of D_λ and the two portions of its conjugate set when similarly divided. Denoting the conjugate set, as usual, by D^λ, let the two portions be

$$D^a \equiv (D^1 \quad D^2 \ldots D^k), \quad D^\phi \equiv (D^{k+1} \quad D^{k+2} \ldots D^m).$$

Here it is to be observed that D^a is not (in general) the set (order k) conjugate to the set D_a (order k), since each element of it is a linear function of the whole D_λ; and similarly for D^ϕ. Now the condition of conjugacy is

$$(D^p . D_q) = |^p_q.$$

But, if D^p and D_q belong to non-corresponding portions of the two sets, q cannot be equal to p. Hence we get the relations

$$(D^a \,.\, D_\beta) = |^a_\beta \ldots (1), \quad (D^a \,.\, D_\phi) = 0 \ldots (2),$$
$$(D^\phi \,.\, D_a) = 0 \ldots (3), \quad (D^\phi \,.\, D_\chi) = |^\phi_\chi \ldots (4).$$

(iv) First let U be an element of D_ϕ or a linear function of D_ϕ, say $a^\phi D_\phi$. Then its improved value must be $U - U = 0$; i. e.

$$I(a^\phi D_\phi) = 0. \qquad \textbf{(IX. 6. 1)}$$

For this makes $(U' \,.\, U') = (0 \,.\, 0) = 0$, by (IX. 1. 3); and, by (IX. 1. 2), $(U' \,.\, U')$ would be > 0 if U' were not $= 0$. Hence $(U' \,.\, U')$ is a minimum when $U' = 0$.

(v) The next most simple case is that in which U is an element of D^a or a linear function of D^a, say

$$U \equiv a_a D^a. \qquad (1)$$

Let the value of U' be

$$U' = U + u,$$

where

$$u \equiv a^\phi D_\phi. \qquad (2)$$

Then, by (IX. 1. 1),

$$(U' \,.\, U') = (U \,.\, U) + 2(U \,.\, u) + (u \,.\, u).$$

But, by (1) and (2), and by (2) of (iii) above,

$$(U \,.\, u) = a_a a^\phi (D^a \,.\, D_\phi) = 0.$$

Hence $(U' \,.\, U')$ is a minimum when $(u \,.\, u)$ is a minimum; and this, by (IX. 1. 2), is when $u = 0$, so that

$$U' = U.$$

Hence the improved value is the same as the original value; i. e.

$$I(a_a D^a) = a_a D^a. \qquad \textbf{(IX. 6. 2)}$$

(vi) The simplicity of the results obtained in (iv) and (v) suggests that we should in all cases regard U as expressed in terms of D^a and D_ϕ. There is no difficulty about this, in theory, whatever linear function of the D's U may be. If, for instance, U is given as a linear function of D_λ, then we obtain our result by eliminating D_a (k values) between this formula for U and the k equations which give D^a in terms of D_λ, i.e. in terms of D_a and D_ϕ. Suppose then that

$$U = V + W,$$

where

$$V = b_a D^a, \quad W = b^\phi D_\phi.$$

Then U' is formed from U by adding some linear function of D_ϕ, so that

$$U' = V + W',$$

where $V = b_a D^a$ as before, and W' is of the form

$$W' = c^\phi D_\phi.$$

Hence

$$(U' . U') = (V . V) + 2(V . W') + (W' . W').$$

But

$$(V . W') = (b_a D^a . c^\phi D_\phi) = b_a c^\phi (D^a . D_\phi) = 0,$$

by (2) of (iii). Hence

$$(U' . U') = (V . V) + (W' . W').$$

But this is a minimum when $W' = 0$. Hence

$$I(b_a D^a + b^\phi D_\phi) = b_a D^a. \qquad \textbf{(IX. 6. 3)}$$

In other words, if we express U in terms of D^a and D_ϕ, the improved value of U is found by omitting the part involving D_ϕ.

(vii) Since, by (IX. 6. 3), IU is a linear function of D^a, and, by (2) of (iii), $(D^a . D_\phi) = 0$, it follows, by (IX. 1. 4), that

$$(IU . D_\phi) = 0. \qquad \textbf{(IX. 6. 4)}$$

In other words, the (.) of any improved value and each of the auxiliaries is 0.

(viii) It also follows from (IX. 6. 3) that if two quantities differ by a linear function of the auxiliaries they have the same improved value.

(ix) By taking U in (vi) to be each member, in turn, of a set

$$B_\lambda \equiv b_{\lambda a} D^a + b^{\lambda\phi} D_\phi,$$

we find that

$$IB_\lambda = b_{\lambda a} D^a.$$

Also

$$I(k^\lambda B_\lambda) = I(k^\lambda b_{\lambda a} D^a + k^\lambda b^{\lambda\phi} D_\phi) = k^\lambda b_{\lambda a} D^a = k^\lambda IB_\lambda; \tag{IX. 6. 5}$$

i. e. the improved value of any linear function of the B's is the same linear function of their improved values.

(x) Altering k^λ in (IX. 6. 5) to k^λ_μ, and writing $C_\mu \equiv k^\lambda_\mu B_\lambda$, we find that

$$\frac{IC_\mu}{IB_\lambda} = \frac{C_\mu}{B_\lambda}; \tag{IX. 6. 6}$$

i. e. the improved values of two linearly connected sets are related in the same way as the original sets; or, more briefly, a set and its improved values are cogredient.

(xi) Since we know that the improved values of D_ϕ are 0, we have really only to determine those of k other quantities. In view of (IX. 6. 5), we can choose these to be any linear functions of D_a that we like, with or without linear functions of D_ϕ added; and similarly we can replace D_ϕ by any linear functions of D_ϕ: provided, in both cases, that none of the mean-product determinants are 0. The functions so chosen can be called D_a and D_ϕ, so that we need only consider the problem of finding ID_a.

(xii) The result stated in (IX. 6. 2) gives us the extension, to the general case in which the errors of the original observations may have any mean squares and mean products, of the 'method of moments' ordinarily applied to the case of a self-conjugate set (§ 3 (vii)). We take X_λ, as in (ii), to be the original observations, and D_λ to be their differences of successive orders. Then we have found that the improved value of any linear function of D^a is the same as the original value. But, by VII. 3 (v), $D^1, D^2, \ldots D^k$ are successive sums of the elements of X^λ, the set conjugate to X_λ; and the first k moments of X^λ are linear functions of these sums. Hence the improved values of these moments are the same as their original values; and this, by (ix), is the same thing as saying that the moments of the improved values of X^λ are equal to the moments of the original values. Thus the method of moments still applies. But it should be observed that it does not apply to the original set of observations, but to the conjugate set.

As a simple example, suppose that X_λ is a set of independent observations of a single quantity, the mean square of error of X_p being $c_p c_p$. Then (§ 3 (viii)) the conjugate set is

$$\left(\frac{X_1}{c_1 c_1}\ \frac{X_2}{c_2 c_2} \ldots\ldots \frac{X_m}{c_m c_m}\right).$$

As the X's will all have the same improved value, which we will call IX, there is only one moment to be considered, namely, the 0th moment, or sum, of the conjugate set. Hence, equating the sums of original and of improved values,

$$\frac{X_1}{c_1 c_1} + \frac{X_2}{c_2 c_2} + \ldots + \frac{X_m}{c_m c_m} = \frac{IX}{c_1 c_1} + \frac{IX}{c_2 c_2} + \ldots + \frac{IX}{c_m c_m},$$

which gives the familiar result.

IX. 7. Determination of improved values.—

(i) From the results obtained in the preceding section we deduce three methods of finding the improved value of any element of D_a, say D_f.

(1) We can express D_f in terms of D^a and D_ϕ, and then omit the part involving D_ϕ. The result is ID_f.

(2) We can say that ID_f is some linear function of D^a. This linear function has k coefficients to be determined; they are determined by the condition that, if the linear

function is expressed in terms of D_λ, the coefficient of D_f is 1 and those of other elements of D_a are 0. The practical application of this method depends on the circumstances of the particular class of cases.

(3) We can say that ID_f is obtained from D_f by adding a linear function of D_ϕ, which we have called $-b^\phi D_\phi$. This linear function has $m-k$ coefficients to be determined. We have found in (IX. 6. 4) that $(ID_f . D_\phi)=0$; this gives $m-k$ equations, from which the coefficients in question can be determined. Thus *the necessary and sufficient conditions for ID_f are that it differs from D_f by a linear function of D_ϕ and that* $(ID_f . D_\phi) = 0$.

These three methods are exhibited in (ii), (iv), and (v) below, and a fourth method is given in (vii).

(ii) To apply the first method, let us write

$$D_a = e_{a\beta} D^\beta + e^{a\phi} D_\phi.$$

We do not need $e^{a\phi}$, since the only part of D_a that counts for the improved value is $e_{a\beta} D^\beta$; we therefore get rid of $e^{a\phi}$ at once, by means of something whose $(. D_\phi)$ is 0. This, by (2) of § 6 (iii), is D^γ. Taking the $(D^\gamma .)$ of both sides, we have

$$(D^\gamma . D_a) = e_{a\beta}(D^\gamma . D^\beta) = e_{a\beta} d^{\beta\gamma}.$$

Also, by (1) of § 6 (iii),

$$(D^\gamma . D_a) = |^\gamma_a.$$

Hence

$$d^{\beta\gamma} e_{a\beta} = |^\gamma_a.$$

Here a, β, γ relate to the partial set $(A_1 A_2 \ldots A_k)$, and the statement is limited to this set. Dealing only with this set, let us denote the reciprocal of $d^{\beta\gamma}$ by $(d_{\beta\gamma})_k$; this, as pointed out in VI. 11 (ii) (cf. § 5 (vi) of the present chapter) is not ordinarily the same thing as $d_{\beta\gamma}$ as obtained from the whole set of order m. We have then, by reciprocation,

$$e_{\alpha\beta} = (d_{\beta\gamma})_k|_{\alpha}^{\gamma} = (d_{\beta\alpha})_k.$$

Substituting in the expression for D_α, and dropping the $e^{\alpha\phi} D_\phi$ in order to get the improved value, we have

$$ID_\alpha = (d_{\beta\alpha})_k D^\beta. \qquad \text{(IX. 7. 1)}$$

(iii) Although, in the above, we have not needed $e^{\alpha\phi}$, we ought to find its value in order to satisfy ourselves that, as has been stated in § 6 (vi), any linear function of D_λ, say $g_\alpha D_\alpha + g_\phi D_\phi$, can be expressed as a linear function of D^α and D_ϕ; to do this, it is only necessary to prove the proposition for D_α, since the formula for $g_\alpha D_\alpha + g_\phi D_\phi$ will follow at once.

We have written

$$D_\alpha = e_{\alpha\beta} D^\beta + e^{\alpha\phi} D_\phi,$$

and have found $e_{\alpha\beta}$. To find $e^{\alpha\phi}$, we must get rid of the first term; so we again use (2) of § 6 (iii), getting

$$(D_\alpha \,.\, D_\chi) = e^{\alpha\phi} (D_\phi \,.\, D_\chi)$$

or

$$d_{\alpha\chi} = e^{\alpha\phi} d_{\phi\chi}.$$

Hence, by reciprocation,

$$e^{\alpha\phi} = [d^{\phi\chi}]_{m-k} d_{\alpha\chi},$$

where $[d^{\phi\chi}]_{m-k}$ is the reciprocal of $d_{\phi\chi}$ obtained from the partial set $(D_{k+1}\; D_{k+2} \ldots D_m)$. The complete expression for D_α is therefore

$$D_\alpha = (d_{\beta\alpha})_k D^\beta + [d^{\phi\chi}]_{m-k} d_{\alpha\chi} D_\phi. \qquad \text{(IX. 7. 2)}$$

The existence of $(d_{\beta\alpha})_k$ and $[d^{\phi\chi}]_{m-k}$ is dependent on the assumption that the determinant $| d_{qr} |$ formed for D_α, and the determinant $| d_{qr} |$ formed for D_ϕ, are both $\neq 0$ (see § 1 (viii)).

(iv) To use the second method, we might have proceeded as follows. We write

$$ID_\alpha = e_{\alpha\beta} D^\beta.$$

To find $e_{\alpha\beta}$, we express D^β in terms of D_μ, i.e. of D_γ and D_ψ, by means of (IX. 3. 4), and we have

$$\begin{aligned} ID_\alpha &= e_{\alpha\beta} d^{\beta\mu} D_\mu \\ &= e_{\alpha\beta} d^{\beta\gamma} D_\gamma + e_{\alpha\beta} d^{\beta\psi} D_\psi. \end{aligned}$$

From the condition stated in (2) of (i), it follows that

$$e_{\alpha\beta} d^{\beta\gamma} = |^\gamma_\alpha,$$

and therefore, by reciprocation,

$$e_{\alpha\beta} = (d_{\beta\gamma})_k |^\gamma_\alpha = (d_{\beta\alpha})_k.$$

Hence we get the same result as before, namely,

$$ID_\alpha = (d_{\beta\alpha})_k D^\beta.$$

(v) For the third method, we write

$$ID_\alpha = D_\alpha - e^{\alpha\phi} D_\phi,$$

and we have to find $e^{\alpha\phi}$. The condition stated in (3) of (i), namely,

$$(ID_\alpha . D_\chi) = 0,$$

gives

$$d_{\alpha\chi} \equiv (D_\alpha . D_\chi) = e^{\alpha\phi}(D_\phi . D_\chi) = e^{\alpha\phi} d_{\phi\chi}.$$

This is true for all the $m-k$ values of χ. By reciprocation

$$e^{\alpha\phi} = [d^{\phi\chi}]_{m-k} d_{\alpha\chi}.$$

This agrees with (IX. 7. 1) and (IX. 7. 2). As $m-k$ will usually be a good deal greater than k, the method is rather of theoretical than of practical interest.

(vi) The elements which we have found to be important in the above processes are the $m-k$ auxiliaries D_ϕ, whose improved values are all 0, and the k elements D^α of the conjugate set which correspond to the remainder of D_λ. These elements together constitute a set of order m; and we have in fact, in (iii), expressed D_a in terms of this set. As the set is important, it is worth while to see what is its conjugate.

We write

$$E^\lambda \equiv D^a \ \& \ D_\phi,$$

where the '&' means that the elements of the two sets of orders k and $m-k$ are combined to form a set of order m. These two partial sets are statistically independent. It follows, by § 3 (ix), that the set conjugate to E^λ is

$$E_\lambda = (d_{a\beta})_k D^\beta \ \& \ [d^{\phi\chi}]_{m-k} D_\chi.$$

(vii) But ($d_{a\beta}$ and $d_{\beta a}$ being identical) we have already found that

$$ID_a = (d_{a\beta})_k D^\beta.$$

Hence we get a concise formula for finding ID_a. Let the set conjugate to D_a & D_ϕ be D^a & D^ϕ; and let the set conjugate to D^a & D_ϕ be F_a & F^ϕ. Then $F_a = ID_a$.

(viii) Since ID_a is of the form $D_a - e^{a\phi} D_\phi$, and ID_β is of the form $e_{\beta\gamma} D^\gamma$, and $(D^\gamma . D_\phi) = 0$, it follows that

$$(ID_a . ID_\beta) = (D_a . ID_\beta),$$

and similarly

$$(ID_a . ID_\beta) = (ID_a . D_\beta).$$

[NOTE.—This chapter is based on (1) a paper by myself in *Phil. Trans.* (1920), ser. A, vol. 221, pp. 199-237, in which the old notation was used; (2) a paper by Professor Eddington in *Proceedings of the London Mathematical Society*, ser. 2, vol. 20, pp. 213-221, showing how the notation and methods of the tensor calculus can be applied, and making some abbreviations and improvements in the work; and (3) a note by myself, following the above, *ibid.*, pp. 222-224. I have altered the notation a good deal.]

X. TENSORS IN THEORY OF RELATIVITY

X. 1. **Preliminary.**—(i) Tensors are sets* which (1) are functions of a set of co-ordinates $(x_1\ x_2\ x_3 \ldots)$ and (2) are subject to certain conditions of transformation when the co-ordinates are transformed.

(ii) In the theory of relativity there are four co-ordinates $(x_1\ x_2\ x_3\ x_4)$, so that all the sets are of order 4, and any inner multiplication with regard to a suffix μ involves addition of the products for the values 1, 2, 3, 4 of μ. But this fixing of the number of elements in a set does not affect the general reasoning with regard to the sets, and we can continue to treat them as of order m.

(iii) In VII. 4 we started with a set X^λ, and a set A^λ which is a definite function of X^λ, and we supposed A^λ to be changed as the result of X^λ being changed by linear substitution; and the cases we considered were those in which, throughout all such changes, A^λ varies either directly or reciprocally as X^λ. In VII. 7 we extended the inquiry by taking a set $\mathfrak{A}$ to be a function of two or more single sets, and considered cases in which, when these sets are changed by linear substitution, $\mathfrak{A}$ varies directly or reciprocally as each of the sets. For tensors we have tc consider cases in which the primary substitutions are not necessarily linear. If in place of the original set of co-ordinates x_λ we take a new set x'_λ, which is a function but not necessarily a linear function of x_λ, the ratio which

* It must be remembered that the elements of a set are not necessarily numbers, but may be quantities; and that a set as a whole is something different from its elements. What we usually mean by a tensor is some physical phenomenon represented by a set: but no confusion arises if we call the set itself a tensor.

we have now to consider is not the ratio of x'_ρ to x_λ but the ratio of their differentials, i.e. the partial differential coefficient $\partial x'_\rho/\partial x_\lambda$. When the substitution is linear, this is equal * to x'_ρ/x_λ, so that our treatment of the general case is consistent with our previous treatment of the particular case.

We will begin with the single set, and then go on to sets of higher rank.

(iv) In the case of sets of higher rank, we sometimes have to deal not only with inner products but also with *inner sums*. By the **inner sum**, in such an expression as $A^\rho_{\mu\nu\sigma}$, we mean the result obtained by replacing ρ by σ and summing the values for $\sigma = 1, 2, 3, 4$. It will be seen presently (§ 5 (iv)) that, as the result of the particular notation adopted, the inner products or sums have only to be considered when one of the two letters concerned is an upper suffix and the other is a lower suffix.

X. 2. Single sets (vectors).—(i) Beginning with single sets, we start with a pair which we call x_λ (the set of co-ordinates) and A^λ, or x_a and A^a; A^λ or A^a being a function of x_λ or x_a. (The a here, like the λ, denotes a complete set, not, as in Chapter IX, a partial set.) Connected with these, or arising out of them, is a plurality of pairs x'_λ and A'^λ. But our purview is limited to the cases in which the relation between A^λ and A'^λ is linear, and in which, further, this linear relation is of one of the two following kinds:

(*a*) where

$$\frac{A'^\lambda}{A^a} = \frac{\partial x'_\lambda}{\partial x_a}:$$

* The difficulty mentioned in the note to VIII. 1 (ii) does not arise, because $\partial x'_\rho/\partial x_\lambda$ does not occur absolutely, but (directly or reciprocally) as one of the factors in inner multiplication with regard to λ or ρ.

(*b*) —replacing A^λ and A'^λ by A_λ and A'_λ—where

$$\frac{A'_\lambda}{A_a} = \frac{\partial x_a}{\partial x'_\lambda}.$$

In the cases under (*a*) A^λ is said to be a **contravariant vector**; in the cases under (*b*) A_λ is said to be a **covariant vector.** Here 'vector' is used as meaning a single set which is a tensor.

(ii) It will be seen from VIII. 1 (ii) that, if the relation between x_λ and x'_λ is linear, these become respectively

$$\frac{A'^\lambda}{A^a} = \frac{x'_\lambda}{x_a},$$

$$\frac{A'_\lambda}{A_a} = \frac{x_a}{x'_\lambda},$$

so that in these particular classes of cases A^λ is *contravariant* if A^λ and x_λ are *cogredient*, and A_λ is *covariant* if A_λ and x_λ are *contragredient.**

X. 3. Other sets.—(i) For sets of higher rank, we are similarly concerned with pairs of partial derivatives

$$\frac{\partial x'_\lambda}{\partial x_a} \text{ and } \frac{\partial x_a}{\partial x'_\lambda},$$

$$\frac{\partial x'_\mu}{\partial x_\beta} \text{ and } \frac{\partial x_\beta}{\partial x'_\mu},$$

etc.; and a set depending on $x_a, x_\beta, \ldots$ is not a tensor unless, when $x_a x_\beta \ldots$ become $x'_\lambda x'_\mu \ldots$, the 'ratio' of the new value of the set to the old value is the product of these partial derivatives, one from each pair. The particular derivatives are indicated by the position of the letters $a\beta \ldots$ or $\mu\nu \ldots$: these are upper suffixes if, so far as the particular variable is concerned, the relation is of the

* It seems desirable to call attention to these classes of cases, as otherwise the tensor terminology may be found rather confusing.

contravariant type, and lower suffixes if the relation is of the covariant type. Thus for double sets (tensors of the second rank) we should have such relations as

$$\frac{A'^{\lambda\mu}}{A^{\alpha\beta}} = \frac{\partial x'_\lambda}{\partial x_\alpha}\frac{\partial x'_\mu}{\partial x_\beta} \text{ (contravariant tensor),}$$

$$\frac{A'_{\lambda\mu}}{A_{\alpha\beta}} = \frac{\partial x_\alpha}{\partial x'_\lambda}\frac{\partial x_\beta}{\partial x'_\mu} \text{ (covariant tensor),}$$

$$\frac{A'^{\mu}_{\lambda}}{A^{\beta}_{\alpha}} = \frac{\partial x_\alpha}{\partial x'_\lambda}\frac{\partial x'_\mu}{\partial x_\beta} \text{ (mixed tensor);}$$

and for a tensor of higher rank we might have such a relation as

$$\frac{A'^{\rho}_{\lambda\mu\nu}}{A^{\delta}_{\alpha\beta\gamma}} = \frac{\partial x_\alpha}{\partial x'_\lambda}\frac{\partial x_\beta}{\partial x'_\mu}\frac{\partial x_\gamma}{\partial x'_\nu}\frac{\partial x'_\rho}{\partial x_\delta}.$$

Two tensors $\mathfrak{A}$ and $\mathfrak{B}$ are said to be of the same **character** if the ratios $\mathfrak{A}'/\mathfrak{A}$ and $\mathfrak{B}'/\mathfrak{B}$ are of the same form.

(ii) For a scalar function of a set or sets the above condition becomes

$$A' = A,$$

so that a scalar (in the general sense) is not a tensor unless it remains constant for all changes in the system of co-ordinates. Such a function as $A_1 + A_2 + A_3 + A_4$, for instance, would not usually be a tensor. For tensor purposes, therefore, 'scalar' practically means 'invariant'.

X. 4. **Reason for limitation.**—The object of limiting the definition of 'tensor' in this way is to ensure that the result of any number of steps, all of the same kind, produced by successive transformations of co-ordinates, shall be the same as if we had passed in one step, also of the same kind, from the initial set of co-ordinates to the final set. That this is in fact ensured is seen from the properties of the partial derivative of a set. Suppose, for example, that

$$\frac{A'^{\lambda}}{A^{a}} = \frac{\partial x'_{\lambda}}{\partial x_{a}}, \tag{1}$$

and that

$$\frac{A''^{\sigma}}{A'^{\lambda}} = \frac{\partial x''_{\sigma}}{\partial x'_{\lambda}}. \tag{2}$$

Then, by (VII. 2. 2) and (VIII. 3. 1),

$$\frac{A''^{\sigma}}{A^{a}} = \frac{A''^{\sigma}}{A'^{\lambda}} \frac{A'^{\lambda}}{A^{a}} = \frac{\partial x''_{\sigma}}{\partial x'_{\lambda}} \frac{\partial x'_{\lambda}}{\partial x_{a}} = \frac{\partial x''_{\sigma}}{\partial x_{a}}, \tag{3}$$

which is of the same form as (1) and (2).

X. 5. Miscellaneous properties.—The following are some miscellaneous properties which are useful in determining whether a set is a tensor.

(i) The sum (or difference) of two tensors of the same rank and character and the same suffix is a tensor.

[Suppose, for instance, that A^{λ} and B^{λ} are contravariant tensors. Then

$$A'^{\lambda} = \frac{\partial x'_{\lambda}}{\partial x_{a}} A^{a}, \quad B'^{\lambda} = \frac{\partial x'_{\lambda}}{\partial x_{a}} B^{a},$$

and therefore

$$(A'^{\lambda} + B'^{\lambda}) = \frac{\partial x'_{\lambda}}{\partial x_{a}} (A^{a} + B^{a}).]$$

(ii) The product of two tensors is a tensor whose character is the combination of the characters of the two. For example, from

$$A'_{\lambda} = \frac{\partial x_{a}}{\partial x'_{\lambda}} A_{a}, \quad B'^{\mu\nu} = \frac{\partial x'_{\mu}}{\partial x_{\beta}} \frac{\partial x'_{\nu}}{\partial x_{\gamma}} B^{\beta\gamma},$$

we see that

$$(A'_{\lambda} B'^{\mu\nu}) = \frac{\partial x_{a}}{\partial x'_{\lambda}} \frac{\partial x'_{\mu}}{\partial x_{\beta}} \frac{\partial x'_{\nu}}{\partial x_{\gamma}} (A_{a} B^{\beta\gamma}).$$

(iii) An inner product of two tensors, or an inner sum (§ 1 (iv)) of a tensor, taken with regard to suffixes of opposite character, is a tensor.

[Take, for example,

$$A'^{\rho}_{\lambda\mu\nu} = \frac{\partial x_a}{\partial x'_\lambda}\frac{\partial x_\beta}{\partial x_\mu'}\frac{\partial x_\gamma}{\partial x'_\nu}\frac{\partial x'_\rho}{\partial x_\delta}A^{\delta}_{a\beta\gamma}.$$

If we replace ρ by ν we have, by (VIII. 3. 3),

$$\frac{\partial x_\gamma}{\partial x'_\nu}\frac{\partial x'_\nu}{\partial x_\delta}A^{\delta}_{a\beta\gamma} = |^{\gamma}_{\delta} A^{\delta}_{a\beta\gamma} = A^{\gamma}_{a\beta\gamma};$$

and therefore

$$A'^{\nu}_{\lambda\mu\nu} = \frac{\partial x_a}{\partial x'_\lambda}\frac{\partial x_\beta}{\partial x'_\mu}A^{\gamma}_{a\beta\gamma},$$

which satisfies the requirements.]

(iv) If in this last example we had replaced μ by λ, instead of ρ by ν, the expression would have contained

$$\frac{\partial x_a}{\partial x'_\lambda}\frac{\partial x_\beta}{\partial x'_\lambda},$$

which has no general significance.

(v) The derivative of a scalar (§ 3 (ii)) is a covariant vector. Suppose, e.g., that A is a scalar function of x_λ, whose value remains constant (§ 3 (ii)) for all transformations. Then

$$\frac{\partial A'}{\partial x'_\lambda} = \frac{\partial A}{\partial x'_\lambda} = \frac{\partial x_a}{\partial x'_\lambda}\frac{\partial A}{\partial x_a},$$

so that $\partial A/\partial x_\lambda$ is a covariant vector.

(vi) If the inner product of a set $\mathfrak{A}$ by each of $m\ (=4)$ $\left\{\begin{matrix}\text{contravariant}\\ \text{covariant}\end{matrix}\right\}$ vectors is a tensor, then $\mathfrak{A}$ is a tensor, and is $\left\{\begin{matrix}\text{covariant}\\ \text{contravariant}\end{matrix}\right\}$ as regards x_μ, where μ is the linked suffix.

[Take the case in which the vectors are contravariant, and suppose that $\mathfrak{A} \equiv A_{\mu\nu\ldots}$. Let the m vectors be denoted by $B^{1\mu}$, $B^{2\mu}, \ldots B^{m\mu}$, or, collectively, by $B^{\lambda\mu}$; the λ not indicating any tensor character. Then, by hypothesis, $B^{\lambda\mu}$ is a tensor as regards x_μ, and $B^{\lambda\mu}A_{\mu\nu\ldots}$ is a tensor as regards $x_\nu\ldots$, so that

$$B^{\lambda\beta} = \frac{\partial x_\beta}{\partial x'_\mu} B'^{\lambda\mu}, \quad B'^{\lambda\mu} A'_{\mu\nu\ldots} = \frac{\partial x_\gamma}{\partial x'_\nu} \ldots B^{\lambda\beta} A_{\beta\gamma\ldots}.$$

From these we deduce

$$B'^{\lambda\mu} A'_{\mu\nu\ldots} = B'^{\lambda\mu} \frac{\partial x_\beta}{\partial x'_\mu} \frac{\partial x_\gamma}{\partial x'_\nu} \ldots A_{\beta\gamma\ldots},$$

whence, by division by $B'^{\lambda\mu}$ (see VI. 9 (v)),

$$A'_{\mu\nu\ldots} = \frac{\partial x_\beta}{\partial x'_\mu} \frac{\partial x_\gamma}{\partial x'_\nu} \ldots A_{\beta\gamma\ldots}.$$

Hence $A_{\mu\nu\ldots}$ is a tensor and is covariant as regards x_μ. The case in which the vectors are covariant can be dealt with in the same way.]

APPENDIX

PRODUCT OF DETERMINANTS

In IV. 5 we have taken as the standard form for product of two determinants—the order being now reduced from 3 to 2, for economy of printing—

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

In this form, the element in the qth column and rth row of the result is the 'inner product' of the qth column of the first determinant and the rth row of the second. By interchanges of columns and rows, and also by changing the order of multiplication, we get seven other forms, all constructed according to this rule. The eight forms can be set out as follows:

$$\begin{vmatrix} a_1 b_1 \\ a_2 b_2 \end{vmatrix} \times \begin{vmatrix} \alpha_1 \beta_1 \\ \alpha_2 \beta_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + a_2\beta_1 & b_1\alpha_1 + b_2\beta_1 \\ a_1\alpha_2 + a_2\beta_2 & b_1\alpha_2 + b_2\beta_2 \end{vmatrix} \tag{1}$$

$$\begin{vmatrix} a_1 b_1 \\ a_2 b_2 \end{vmatrix} \times \begin{vmatrix} \alpha_1 \alpha_2 \\ \beta_1 \beta_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + a_2\alpha_2 & b_1\alpha_1 + b_2\alpha_2 \\ a_1\beta_1 + a_2\beta_2 & b_1\beta_1 + b_2\beta_2 \end{vmatrix} \tag{2}$$

$$\begin{vmatrix} a_1 a_2 \\ b_1 b_2 \end{vmatrix} \times \begin{vmatrix} \alpha_1 \beta_1 \\ \alpha_2 \beta_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + b_1\beta_1 & a_2\alpha_1 + b_2\beta_1 \\ a_1\alpha_2 + b_1\beta_2 & a_2\alpha_2 + b_2\beta_2 \end{vmatrix} \tag{3}$$

$$\begin{vmatrix} a_1 a_2 \\ b_1 b_2 \end{vmatrix} \times \begin{vmatrix} \alpha_1 \alpha_2 \\ \beta_1 \beta_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + b_1\alpha_2 & a_2\alpha_1 + b_2\alpha_2 \\ a_1\beta_1 + b_1\beta_2 & a_2\beta_1 + b_2\beta_2 \end{vmatrix} \tag{4}$$

$$\begin{vmatrix} \alpha_1 \beta_1 \\ \alpha_2 \beta_2 \end{vmatrix} \times \begin{vmatrix} a_1 b_1 \\ a_2 b_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + b_1\alpha_2 & a_1\beta_1 + b_1\beta_2 \\ a_2\alpha_1 + b_2\alpha_2 & a_2\beta_1 + b_2\beta_2 \end{vmatrix} \tag{5}$$

$$\begin{vmatrix} \alpha_1 \alpha_2 \\ \beta_1 \beta_2 \end{vmatrix} \times \begin{vmatrix} a_1 b_1 \\ a_2 b_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + b_1\beta_1 & a_1\alpha_2 + b_1\beta_2 \\ a_2\alpha_1 + b_2\beta_1 & a_2\alpha_2 + b_2\beta_2 \end{vmatrix} \tag{6}$$

$$\begin{vmatrix} \alpha_1 \beta_1 \\ \alpha_2 \beta_2 \end{vmatrix} \times \begin{vmatrix} a_1 a_2 \\ b_1 b_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + a_2\alpha_2 & a_1\beta_1 + a_2\beta_2 \\ b_1\alpha_1 + b_2\alpha_2 & b_1\beta_1 + b_2\beta_2 \end{vmatrix} \tag{7}$$

$$\begin{vmatrix} \alpha_1 \alpha_2 \\ \beta_1 \beta_2 \end{vmatrix} \times \begin{vmatrix} a_1 a_2 \\ b_1 b_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + a_2\beta_1 & a_1\alpha_2 + a_2\beta_2 \\ b_1\alpha_1 + b_2\beta_1 & b_1\alpha_2 + b_2\beta_2 \end{vmatrix} \tag{8}$$

It will be seen that the last four are the transposed of the first four, but in the reverse order; i. e. the transposed of (1) (2) (3) (4) are (8) (7) (6) (5).

In the double-suffix notation these become, the order of (5)–(8) being reversed:

$\mid d_{qr} \mid \times \mid e_{qr} \mid = \mid d_{q\lambda} e_{\lambda r} \mid$ (1)	$\mid e_{rq} \mid \times \mid d_{rq} \mid = \mid d_{r\lambda} e_{\lambda q} \mid$ (8)
$\mid d_{qr} \mid \times \mid e_{rq} \mid = \mid d_{q\lambda} e_{r\lambda} \mid$ (2)	$\mid e_{qr} \mid \times \mid d_{rq} \mid = \mid d_{r\lambda} e_{q\lambda} \mid$ (7)
$\mid d_{rq} \mid \times \mid e_{qr} \mid = \mid d_{\lambda q} e_{\lambda r} \mid$ (3)	$\mid e_{rq} \mid \times \mid d_{qr} \mid = \mid d_{\lambda r} e_{\lambda q} \mid$ (6)
$\mid d_{rq} \mid \times \mid e_{rq} \mid = \mid d_{\lambda q} e_{r\lambda} \mid$ (4)	$\mid e_{qr} \mid \times \mid d_{qr} \mid = \mid d_{\lambda r} e_{q\lambda} \mid$ (5)

It must be borne in mind that in each of these statements q refers to the column and r to the row; e. g. (6) means that

$$\begin{vmatrix} e_{11} e_{12} \\ e_{21} e_{22} \end{vmatrix} \times \begin{vmatrix} d_{11} d_{21} \\ d_{12} d_{22} \end{vmatrix} = \begin{vmatrix} d_{\lambda 1} e_{\lambda 1} & d_{\lambda 1} e_{\lambda 2} \\ d_{\lambda 2} e_{\lambda 1} & d_{\lambda 2} e_{\lambda 2} \end{vmatrix}.$$

INDEX OF SYMBOLS

GENERAL INDEX

Under any heading, separate entries are usually in the order of occurrence, not in alphabetical order.

PRINTED IN ENGLAND
AT THE OXFORD UNIVERSITY PRESS