Chapter I

SOME PRELIMINARIES

1. Determinants. Summation convention.

Before entering on the subject of Differential Geometry we may, with advantage, devote a little space to the mention of certain results of algebra and analysis, which will be needed in the following pages, explaining at the same time the notation to be employed.

It is assumed that the reader is familiar with the elementary properties of determinants. If the numbers i, j can take all positive integral values from 1 to n, the n^2 quantities a^i_j may be taken as elements of a determinant of *order* n, viz.

$$a \equiv \begin{vmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \dots & \dots & \dots & \dots \\ a^n_1 & a^n_2 & \dots & a^n_n \end{vmatrix}, \qquad \text{......(1)}$$

which is a homogeneous polynomial of the nth degree in the elements. The superscript i of the symbol a^i_j denotes the row to which the element belongs, and the subscript j indicates the column. The determinant is also frequently denoted briefly by $|a^i_j|$. If $a^i_j = a^j_i$, for all values of i and j, the determinant is *symmetric*; while if $a^i_j = -a^j_i$ it is *skew-symmetric*.

Let A^j_i denote the cofactor of the element a^i_j in the determinant a. It is well known that the sum of the products of the elements of the ith row (or column) by the cofactors of the corresponding elements of the jth row (or column) is equal to a if $i = j$, and to zero if $i \neq j$. Consequently

$$a^i_1 A^1_j + a^i_2 A^2_j + \dots + a^i_n A^n_j = a\delta^i_j,$$

where the symbols δ^i_j are defined by

$$\left.\begin{aligned} \delta^i_j &= 1 \quad \text{if} \quad i = j \\ \text{and} \qquad \delta^i_j &= 0 \quad \text{if} \quad i \neq j \end{aligned}\right\}. \qquad \text{......(2)}$$

These symbols are called the *Kronecker deltas*, and are used constantly throughout these pages. The above equation, and the corresponding one obtained by interchanging rows and columns, may be expressed

$$\sum_{k}^{1,\dots,n} a^i_k A^k_j = a\delta^i_j,$$

and

$$\sum_{k}^{1,\dots,n} a^k_i A^j_k = a\delta^j_i.$$

Following the *summation convention*, due to Einstein, we dispense with the sign of summation and write these simply

$$a^i_k A^k_j = a\delta^i_j, \qquad \dots\dots(3)$$

and

$$a^k_i A^j_k = a\delta^j_i. \qquad \dots\dots(3')$$

In accordance with this summation convention, when the same index appears in any term as a subscript and a superscript, this term stands for the sum of all the terms obtained by giving that index all the values it may take. In (3) or (3′) the index k appears as subscript and superscript in the same term; so that the single term expressed stands for the sum of n terms. The repeated index is called a *dummy* or an *umbral* index, because the value of the expression does not depend on the symbol used for this index. Thus

$$a^i_k A^k_j = a^i_h A^h_j.$$

We may also remark that, in agreement with the summation convention,

$$\delta^i_i = \delta^1_1 + \delta^2_2 + \dots + \delta^n_n = n. \qquad \dots\dots(4)$$

Hence the necessity of writing the first of equations (2) in that form.

The determinant of the n^2 cofactors A^j_i of the elements of (1) is called the *adjoint* of a. We denote it by A. Thus

$$A = |A^j_i|.$$

It is well known that*

$$A = a^{n-1}. \qquad \dots\dots(5)$$

* See, e.g., Bôcher, 1907, 1, p. 33. The references are to the Bibliography at the end of the book.

The rule for forming the *product of two determinants* of the same order may be neatly expressed by means of the summation convention. According to this rule the product of the determinants $|a^i_j|$ and $|b^i_j|$ is the determinant whose elements p^i_j are given by

$$p^i_j = a^i_k b^k_j.$$

Thus

$$|a^i_j| . |b^i_j| = |a^i_k b^k_j|.$$

A second application of this rule shows that

$$|a^i_j| . |b^i_j| . |c^i_j| = |a^i_k b^k_h c^h_j|,$$

and so on.

2. Differentiation of a determinant.

If the elements of the determinant a are functions of the independent variables $x, y, \ldots$, the derivatives of a with respect to these variables are given by formulae of the type

$$\frac{\partial a}{\partial x} = A^j_i \frac{\partial}{\partial x} a^i_j, \qquad \ldots\ldots(6)$$

in which the second member stands for a double sum, the repeated indices i, j each taking all integral values from 1 to n.

To prove this formula we observe that the expansion of the determinant consists of a sum of terms, each of which is a product of n elements. The derivative of this sum is a sum of terms, each of which is the product of $n-1$ elements and the derivative of another element; and the derivative of every element occurs in the sum. If we collect all the terms containing the derivative of the element a^i_j, it is clear from (3) that the coefficient of this derivative is A^j_i. Thus the whole sum, which expresses the derivative of a, is the sum of terms such as

$$A^j_i \frac{\partial}{\partial x} a^i_j,$$

the summation being extended to all the elements of the determinant, that is to say, to all rows and all columns. But this summation is indicated by the repeated indices in the term just written. Hence we have the formula (6).

3. Matrices. Rank of a matrix.

A system of mn quantities, arranged in a rectangular array of m rows and n columns, is called a *matrix*. Let the mn quantities be denoted by a^i_j, i taking the values 1, 2, ..., m and j the values 1, 2, ..., n. Then the matrix is usually expressed in the form

$$\begin{Vmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \dots & \dots & \dots & \dots \\ a^m_1 & a^m_2 & \dots & a^m_n \end{Vmatrix}$$

or, more briefly, $\quad \| a^i_j \| \quad \begin{pmatrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n \end{pmatrix}.$

If $m = n$, the matrix is said to be a *square* matrix of order n; and the determinant $|a^i_j|$ is called the determinant of the square matrix.

By striking out certain rows or columns (or both) of a matrix we obtain other matrices. In particular by doing so we obtain certain square matrices, whose determinants are called the determinants contained by the original matrix. If the matrix consists of m rows and n columns, it contains determinants of all orders from 1 to the smaller of the integers m and n. It frequently happens that all the determinants of orders greater than a certain integer are zero. The *rank* of a matrix is defined as the order of the non-zero determinant of highest order contained by the matrix. Thus, if the rank is r, the matrix contains at least one determinant of order r which is not zero, while all its determinants of order greater than r are zero.

4. Linear equations. Cramer's rule.

Consider the system of n linear equations

$$\left.\begin{array}{c} a^1_1 x^1 + a^1_2 x^2 + \dots + a^1_n x^n = b^1 \\ a^2_1 x^1 + a^2_2 x^2 + \dots + a^2_n x^n = b^2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a^n_1 x^1 + a^n_2 x^2 + \dots + a^n_n x^n = b^n \end{array}\right\} \qquad \dots\dots(7)$$

in the n unknowns $x^1, x^2, \ldots, x^n$, where the superscripts are merely distinguishing indices, having no connection with "powers". The determinant $|a^i_j|$ of the coefficients in the first members is the determinant (1). Its value will be denoted by a; and, as above, A^j_i will denote the cofactor of the element a^i_j.

In virtue of the summation convention we may write the ith of equations (7) in the form

$$a^i_j x^j = b^i. \qquad \ldots\ldots(8)$$

If we multiply this by A^k_i, and sum for all integral values of i from 1 to n, we obtain

$$A^k_i a^i_j x^j = A^k_i b^i,$$

which, in consequence of (3′), is equivalent to

$$a\delta^k_j x^j = A^k_i b^i.$$

Now, in the sum indicated by the first member of this equation, j taking the values 1, 2, ..., n, all the quantities δ^k_j are zero, except that in which j has the value k. Thus the equation reduces to

$$ax^k = A^k_i b^i.$$

Consequently, provided a is not zero, the solution of the system (7) is given by

$$x^k = \frac{A^k_i b^i}{a}. \qquad \ldots\ldots(9)$$

This is *Cramer's rule* for the solution of a system of linear equations.

Suppose next that we have *a system of m equations in n unknowns*,

$$\left.\begin{array}{c} a^1_1 x^1 + a^1_2 x^2 + \ldots + a^1_n x^n = b^1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a^m_1 x^1 + a^m_2 x^2 + \ldots + a^m_n x^n = b^m \end{array}\right\}. \qquad \ldots\ldots(10)$$

The matrix $\quad \| a^i_j \|, \quad \begin{pmatrix} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{pmatrix}$

is called the matrix of the system of equations, while

$$\left\| \begin{array}{ccccc} a^1_1 & a^1_2 & \ldots & a^1_n & b^1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a^m_1 & a^m_2 & \ldots & a^m_n & b^m \end{array} \right\|$$

is called the *augmented matrix*. It can be shown that the necessary and sufficient condition that the system of equations may be consistent is that the matrix of the system have the same rank as the augmented matrix.* If this condition is satisfied, and r is the common rank of the matrices, the values of $n-r$ of the unknowns may be assigned arbitrarily, and those of the other unknowns will then be uniquely determined.

Lastly consider the system of *homogeneous linear equations* obtained from (10) by taking all the quantities b^i equal to zero. The augmented matrix has necessarily the same rank as the matrix of the system of equations, so that the system has one or more solutions. Also, as above, if the rank of the system is r, the values of $n-r$ of the unknowns may be assigned arbitrarily, and those of the others will then be uniquely determined. If $r = n$ there is only one solution, which is obviously

$$x^1 = x^2 = \dots = x^n = 0. \qquad \text{......(11)}$$

In order that there may exist a solution different from (11), the rank of the system of equations must be less than n. In particular, if the number of equations is less than the number of unknowns, the equations always possess solutions other than (11). If $m = n$, a necessary and sufficient condition for a solution different from (11) is that the determinant of the coefficients be zero.

5. Linear transformations.

In problems of algebra or analysis it is frequently convenient to change the variables, taking as new variables certain functions of the original ones. A case of particular importance is that in which the new variables are homogeneous linear polynomials in the original variables. Such a transformation, or change of variables, is called a *homogeneous linear transformation*. If $x^1, x^2, \dots, x^n$ are the original variables and $y^1, y^2, \dots, y^n$ the new ones, the transformation is given by

* Bôcher, 1907, 1, p. 46; or Dickson, 1930, 4, p. 63.

equations of the form

$$\left.\begin{array}{l} y^1 = a^1_1 x^1 + a^1_2 x^2 + \dots + a^1_n x^n \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ y^n = a^n_1 x^1 + a^n_2 x^2 + \dots + a^n_n x^n \end{array}\right\}, \qquad \dots\dots(12)$$

where the coefficients a^i_j are constants. If these are real the transformation is said to be *real*. The matrix $\| a^i_j \|$ is called the matrix of the transformation, and its determinant $a \equiv | a^i_j |$ is the determinant of the transformation. If this determinant is zero the transformation is said to be *singular*; otherwise it is *non-singular*. In accordance with the summation convention the transformation (12) may be expressed briefly

$$y^i = a^i_j x^j, \qquad \dots\dots(12')$$

$$(i, j = 1, 2, \dots, n).$$

Let the transformation be non-singular. Then, solving the equations (12) for the x's in terms of the y's, we have, by Cramer's rule,

$$x^k = \frac{1}{a} A^k_i y^i. \qquad \dots\dots(13)$$

The transformation expressed by (13) is called the *inverse* of (12). Since (12) is non-singular so also is (13); for the determinant of (13) has the value

$$a^{-n} | A^k_i | = a^{-1}$$

in virtue of (5).

6. Functional determinants.

Consider n functions of n variables,

$$y^i(x^1, x^2, \dots, x^n), \quad (i = 1, 2, \dots, n),$$

which are finite and continuous, along with their derivatives, in the field considered. The *Jacobian* or *functional determinant* of the y's with respect to the x's is the determinant, of order n, whose elements are the partial derivatives of the functions with respect to the x's, that is to say, the determinant

$$\begin{vmatrix} \dfrac{\partial y^1}{\partial x^1} & \dfrac{\partial y^1}{\partial x^2} & \cdots & \dfrac{\partial y^1}{\partial x^n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial y^n}{\partial x^1} & \dfrac{\partial y^n}{\partial x^2} & \cdots & \dfrac{\partial y^n}{\partial x^n} \end{vmatrix}.$$

This determinant is also denoted briefly by

$$\frac{\partial(y^1, y^2, \ldots, y^n)}{\partial(x^1, x^2, \ldots, x^n)} \quad \text{or} \quad \left|\frac{\partial y^i}{\partial x^j}\right|,$$

or still more briefly by $\left|\dfrac{\partial y}{\partial x}\right|$.

The n functions y^i are said to be *independent* if their functional determinant does not vanish identically. In this case the equations

$$y^i = y^i(x^1, x^2, \ldots, x^n)$$

are soluble for the x's in terms of the y's.

Let $z^i(y^1, y^2, \ldots, y^n)$ be n independent functions of the y's. Then, by the formula for partial differentiation, we have

$$\frac{\partial z^i}{\partial x^j} = \frac{\partial z^i}{\partial y^1}\frac{\partial y^1}{\partial x^j} + \ldots + \frac{\partial z^i}{\partial y^n}\frac{\partial y^n}{\partial x^j},$$

which, in accordance with the summation convention, may be expressed

$$\frac{\partial z^i}{\partial x^j} = \frac{\partial z^i}{\partial y^k}\frac{\partial y^k}{\partial x^j}, \qquad \ldots\ldots(14)$$

on the understanding that a superscript in the denominator has the force of a subscript, so far as concerns the summation convention. Consequently, by the rule for multiplying determinants, we have the important relation

$$\left|\frac{\partial z}{\partial x}\right| = \left|\frac{\partial z}{\partial y}\right| \cdot \left|\frac{\partial y}{\partial x}\right| \qquad \ldots\ldots(15)$$

connecting the functional determinants.

Since the Jacobian $\left|\frac{\partial y}{\partial x}\right|$ is not zero, we may take as a particular case

$$z^i(y^1, y^2, \ldots, y^n) = x^i.$$

The relations (14) then become

$$\frac{\partial x^i}{\partial y^k}\frac{\partial y^k}{\partial x^j} = \delta^i_j,$$

and (15) reduces to

$$\left|\frac{\partial y}{\partial x}\right| = \frac{1}{\left|\frac{\partial x}{\partial y}\right|}. \qquad \ldots\ldots(16)$$

7. Functional matrices.

When the number m of functions y^i is different from the number n of variables x^j, we consider the functional matrix

$$\left\|\begin{matrix} \frac{\partial y^1}{\partial x^1} & \frac{\partial y^1}{\partial x^2} & \cdots & \frac{\partial y^1}{\partial x^n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y^m}{\partial x^1} & \frac{\partial y^m}{\partial x^2} & \cdots & \frac{\partial y^m}{\partial x^n} \end{matrix}\right\|,$$

which is often denoted more briefly by $\left\|\frac{\partial y}{\partial x}\right\|$. When the rank of this matrix is equal to m, the m functions are said to be *independent*. It follows that, if the number of functions is greater than the number of variables, the functions cannot be independent; while, if $m = n$, the above definition of independence is identical with that given in § 6.

If the rank r of the functional matrix is less than the number m of functions, these are not independent; but there exist $m-r$ relations between them of the type*

$$f(y^1, y^2, \ldots, y^m) = 0$$

involving the y's but not the x's.

* Levi-Civita, 1927, 1, pp. 9–12.

8. Quadratic forms.

A homogeneous polynomial in the variables $x^1, x^2, \ldots, x^n$ is called a *form*, or, more fully, an *n-ary form*, since the number of variables is n. We shall be concerned very largely with forms of the second degree, that is to say, *quadratic forms*. An n-ary quadratic form may be expressed

$$a_{ij}x^i x^j, \qquad (i, j = 1, \ldots, n),$$

the single term denoting a double sum, since each of the indices i, j occurs as superscript and subscript. When the coefficients a_{ij} are all real, the form is said to be *real*. The square matrix of the coefficients, viz.

$$\left\| \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{matrix} \right\|,$$

is called the *matrix* of the form; its rank is the *rank* of the quadratic form, and its determinant is the *discriminant* of the form. If this discriminant vanishes the form is said to be *singular*; otherwise it is non-singular. If its rank is less than n the form must be singular.

If the x's in the above form are subjected to a non-singular linear transformation

$$x^i = a^i_j y^j, \qquad (i, j = 1, \ldots, n),$$

we obtain a form in the variables y^j, which we may denote by

$$b_{ij}y^i y^j.$$

It can be shown that the matrix $\|b_{ij}\|$ of this form has the same rank as the matrix $\|a_{ij}\|$; that is to say, *the rank of a quadratic form is unaltered by a non-singular linear transformation of the variables*.*

Further, any quadratic form in n variables can be reduced by a non-singular linear transformation to the form

$$c_i(x^i)^2, \qquad (i = 1, \ldots, n). \qquad \ldots\ldots(17)$$

* Bôcher, 1907, 1, p. 129.