CHAPTER I.

INTRODUCTION. SYMBOLICAL NOTATION.

1. If in the expression

$$a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2,$$

we write

$$x_1 = \xi_1 X_1 + \eta_1 X_2,$$

$$x_2 = \xi_2 X_1 + \eta_2 X_2,$$

we obtain a new expression, viz.

$$A_0 X_1^2 + 2A_1 X_1 X_2 + A_2 X_2^2,$$

where

$$A_0 = a_0 \xi_1^2 + 2a_1 \xi_1 \xi_2 + a_2 \xi_2^2,$$

$$A_1 = a_0 \xi_1 \eta_1 + a_1 (\xi_1 \eta_2 + \xi_2 \eta_1) + a_2 \xi_2 \eta_2,$$

$$A_2 = a_0 \eta_1^2 + 2a_1 \eta_1 \eta_2 + a_2 \eta_2^2.$$

It is easy to verify the identity

$$A_0 A_2 - A_1^2 = (a_0 a_2 - a_1^2)(\xi_1 \eta_2 - \xi_2 \eta_1)^2,$$

which shews that the function $A_0 A_2 - A_1^2$ of the coefficients of the transformed expression differs from the same function $a_0 a_2 - a_1^2$ of the coefficients of the original expression by a factor involving only the coefficients contained in the transformation.

2. In the present work we shall give an account of the theory and structure of functions of the coefficients possessing properties analogous to that described above; but before proceeding to generalities we shall give some further examples.

If we transform the two expressions

$$a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2,$$

$$a_0' x_1^2 + 2a_1' x_1 x_2 + a_2' x_2^2,$$

in the same way as before, and they become

$$A_0X_1^2 + 2A_1X_1X_2 + A_2X_2^2,$$

$$A_0'X_1^2 + 2A_1'X_1X_2 + A_2'X_2^2,$$

then it is easy to verify the identity

$$A_0A_2' - 2A_1A_1' + A_0'A_2 = (a_0a_2' - 2a_1a_1' + a_0'a_2)(\xi_1\eta_2 - \xi_2\eta_1)^2.$$

Thus we have here a function of the coefficients of two expressions such that the new value differs from the original value by a factor depending only on the transformation employed

3. As a third example, if the cubic expression

$$a_0x_1^3 + 3a_1x_1^2x_2 + 3a_2x_1x_2^2 + a_3x_2^3$$

become $\quad A_0X_1^3 + 3A_1X_1^2X_2 + 3A_2X_1X_2^2 + A_3X_2^3,$

when we put $\quad x_1 = \xi_1X_1 + \eta_1X_2,$

$$x_2 = \xi_2X_1 + \eta_2X_2,$$

then we have

$$(A_0A_2 - A_1^2)X_1^2 + (A_0A_3 - A_1A_2)X_1X_2 + (A_1A_3 - A_2^2)X_2^2$$
$$= \{(a_0a_2 - a_1^2)x_1^2 + (a_0a_3 - a_1a_2)x_1x_2 + (a_1a_3 - a_2^2)x_2^2\}(\xi_1\eta_2 - \xi_2\eta_1)^2.$$

This identity indicates a property quite similar to that illustrated in the two previous examples, but the function, which is unaltered except for the factor $(\xi_1\eta_2 - \xi_2\eta_1)^2$, now involves the variables as well as the coefficients of the expression from which it is formed.

The result we have written down may be verified directly, but more easily as follows:

Denoting the original expression by f and the transformed expression by F we have to prove that

$$\frac{\partial^2 F}{\partial X_1^2}\frac{\partial^2 F}{\partial X_2^2} - \left(\frac{\partial^2 F}{\partial X_1 \partial X_2}\right)^2 = \left\{\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2\right\}(\xi_1\eta_2 - \xi_2\eta_1)^2.$$

Now $$\frac{\partial F}{\partial X_1} = \frac{\partial F}{\partial x_1}\frac{\partial x_1}{\partial X_1} + \frac{\partial F}{\partial x_2}\frac{\partial x_2}{\partial X_1}$$

$$= \xi_1\frac{\partial F}{\partial x_1} + \xi_2\frac{\partial F}{\partial x_2},$$

and in like manner

$$\frac{\partial^2 F}{\partial X_1^2} = \xi_1^2 \frac{\partial^2 F}{\partial x_1^2} + 2\xi_1\xi_2 \frac{\partial^2 F}{\partial x_1 \partial x_2} + \xi_2^2 \frac{\partial^2 F}{\partial x_2^2}$$

$$\frac{\partial^2 F}{\partial X_1 \partial X_2} = \xi_1\eta_1 \frac{\partial^2 F}{\partial x_1^2} + (\xi_1\eta_2 + \xi_2\eta_1) \frac{\partial^2 F}{\partial x_1 \partial x_2} + \xi_2\eta_2 \frac{\partial^2 F}{\partial x_2^2}$$

$$\frac{\partial^2 F}{\partial X_2^2} = \eta_1^2 \frac{\partial^2 F}{\partial x_1^2} + 2\eta_1\eta_2 \frac{\partial^2 F}{\partial x_1 \partial x_2} + \eta_2^2 \frac{\partial^2 F}{\partial x_2^2}.$$

But these equations are exactly the same as those which express A_0, A_1, A_2 in terms of a_0, a_1, a_2 (§ 1), hence

$$\frac{\partial^2 F}{\partial X_1^2}\frac{\partial^2 F}{\partial X_2^2} - \left(\frac{\partial^2 F}{\partial X_1 \partial X_2}\right)^2 = (\xi_1\eta_2 - \xi_2\eta_1)^2 \left\{\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2\right\}.$$

The expression $\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$ is called the Hessian of f.

4. Let us now explain the phraseology in common use when dealing with questions such as arise in our subject.

Quantics. A rational integral homogeneous algebraic function of any number of variables $x_1, x_2, \dots x_p$, is called a quantic.

The degree in the variables is called the *order* of the quantic, and according as the number of variables is two, three, four ...... we call the quantic binary, ternary, quaternary ......

Thus a binary quantic of order n is a rational integral homogeneous algebraic function of two variables which is of the nth degree in those variables.

Such a quantic might be written

$$a_0 x_1^n + a_1 x_1^{n-1} x_2 + a_2 x_1^{n-2} x_2^2 + \dots + a_n x_2^n,$$

but we shall find it invariably more convenient to write it

$$a_0 x_1^n + \binom{n}{1} a_1 x_1^{n-1} x_2 + \binom{n}{2} a_2 x_1^{n-2} x_2^2 + \dots + a_n x_2^n,$$

i.e. with binomial coefficients prefixed to the various a's.

The former of these expressions is now commonly written

$$(a_0, a_1, a_2, \dots, a_n \between x_1, x_2)^n,$$

and the latter $\quad (a_0, a_1, a_2, \dots, a_n \between x_1, x_2)^n,$

a very convenient notation introduced by Cayley.

The mere consideration of the transformation of the binary form

$$a_0x_1^2 + 2a_1x_1x_2 + a_2x_2^2$$

will be sufficient to convince the reader of the advantage of the introduction of binomial coefficients.

Passing now to the case of any number of variables, we call the quantic a p-ary q-ic when it is homogeneous and of degree q in p variables.

Thus the most general ternary quadratic is written

$$a_{200}x_1^2 + a_{020}x_2^2 + a_{002}x_3^2 + 2a_{110}x_1x_2 + 2a_{101}x_1x_3 + 2a_{011}x_2x_3,$$

and in general the ternary n-ic is written

$$\Sigma \frac{n!}{p!\,q!\,r!} a_{pqr}x_1^p x_2^q x_3^r,$$

where the summation is extended to all values of p, q, r satisfying the equality

$$p + q + r = n.$$

It will be noticed that here we have prefixed multinomial coefficients to the a's.

5. Linear Transformations. The equations

$$x_1 = \xi_1 X_1 + \eta_1 X_2$$
$$x_2 = \xi_2 X_1 + \eta_2 X_2$$

are said to constitute a linear transformation from the variables x_1x_2 to the variables X_1X_2—it is of course implied that the coefficients on the right do not involve either set of variables.

The determinant

$$D = \begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix}$$

is called the determinant of the transformation.

If D vanishes it is evident that x_1 and x_2 are virtually identical, for their ratio is constant, and hence, as the variables are always supposed to be independent, we shall throughout only deal with transformations which have a non-vanishing determinant.

On solving for X_1X_2 we find

$$X_1 = (\eta_2x_1 - \eta_1x_2)/D$$
$$X_2 = (-\xi_2x_1 + \xi_1x_2)/D$$

so that the passage back from the new variables to the old is effected by a linear transformation. This is called the inverse of the original transformation; it is evident at once that its determinant is equal to $\frac{1}{D}$.

6. Let us now regard a linear transformation as an operator, which acting on x_1, x_2 changes them to X_1, X_2, and let us consider the effect of two such operators acting successively.

If the coefficients of the first are

$$\xi_1, \eta_1; \ \xi_2, \eta_2,$$

and those of the second

$$\xi_1', \eta_1'; \ \xi_2', \eta_2',$$

then we have

$$\left.\begin{aligned} x_1 &= \xi_1 X_1 + \eta_1 X_2 \\ x_2 &= \xi_2 X_1 + \eta_2 X_2 \end{aligned}\right\},$$

$$\left.\begin{aligned} X_1 &= \xi_1' X_1' + \eta_1' X_2' \\ X_2 &= \xi_2' X_1' + \eta_2' X_2' \end{aligned}\right\},$$

and the effect of the two operators acting successively is to change from the variables x_1, x_2 to X_1', X_2'.

Now on elimination of X_1, X_2 we find

$$x_1 = (\xi_1\xi_1' + \eta_1\xi_2') X_1' + (\xi_1\eta_1' + \eta_1\eta_2') X_2',$$

$$x_2 = (\xi_2\xi_1' + \eta_2\xi_2') X_1' + (\xi_2\eta_1' + \eta_2\eta_2') X_2'.$$

And accordingly we can pass directly from the original to the final variables by means of a single linear transformation which we shall call Σ.

If we call the two preceding operators S and S' we may write

$$\Sigma = SS'$$

and Σ is called the product or the resultant of S and S'.

It must be carefully noticed that the order of the factors S and S' is essential in considering their product. In our example we supposed that S acted first and then S'. If S' had acted first and then S we should have

$$\Sigma' = S'S$$

and it is manifest that Σ and Σ' are not in general the same.

Since the resultant of two or any number of linear transformations is another such transformation, the whole set of linear transformations obtained by varying the coefficients is said to form a group—a continuous group because the coefficients ξ and η may be supposed to vary continuously.

The determinant of Σ is equal to the product of the determinants of S and S', as follows from the multiplication theorem for determinants.

The product of a transformation and its inverse is a transformation which does not affect the variables, *i.e.* it is

$$\left.\begin{aligned} x_1 &= X_1 \\ x_2 &= X_2 \end{aligned}\right\}$$

which is called the identical operator. The determinant of this is unity, and, as we have pointed out, the product of the determinants of a transformation and its inverse is also unity.

7. The idea of a linear transformation admits of immediate extension to any number of variables $x_1, x_2, \dots x_p$ and now the transformation consists of n equations

$$x_r = \xi_{r1} X_1 + \xi_{r2} X_2 + \dots + \xi_{rp} X_p, \qquad r = 1, 2, \dots p.$$

The determinant D formed with the ξ's for elements is called the determinant of the transformation, and inasmuch as when D vanishes there is a linear homogeneous relation between the x's, we exclude as before all transformations having a vanishing determinant.

If $D \neq 0$ we can solve for the X's in terms of the x's and, as can be easily seen, each X is a linear function of $x_1, x_2, \dots x_p$, so that we have

$$X_r = \eta_{r1} x_1 + \eta_{r2} x_2 + \dots + \eta_{rp} x_p,$$

a linear transformation which is the inverse of the preceding one.

As in the case of two variables, the resultant of two linear transformations S and T is a third linear transformation

$$\Sigma = ST,$$

and on examining the coefficients in Σ it will be seen at once by the multiplication theorem that the determinant of Σ is the product of the determinants of S and T.

8. In the earlier portion of this work we shall deal almost entirely with binary forms, and although we shall be constantly considering linear transformations and their effects, yet the fact that they form a group will not be explicitly used. Our only object, in introducing these elementary properties of groups, is to point out that the connection between invariants and groups is intimate and universal—in other words, that every group has its accompanying invariants and, conversely, every set of invariants belongs to a group.

9. Invariants of Binary Forms. If a binary form f be changed by a linear transformation into a new form F, and a function I of the coefficients of F be equal to the same function of the coefficients of f multiplied by a factor depending solely on the transformation, then I is called an invariant of the binary form f.

Thus for example in § 1 the identity

$$(A_0 A_2 - A_1^2) = (a_0 a_2 - a_1^2)(\xi_1 \eta_2 - \xi_2 \eta_1)^2$$

shews that $a_0 a_2 - a_1^2$ is an invariant of the binary quadratic

$$a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2.$$

An exactly similar definition applies to a joint invariant of several binary forms, *e.g.*

$$a_0 a_2' - 2a_1 a_1' + a_0' a_2$$

is an invariant of the two binary forms

$$a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2,$$

and

$$a_0' x_1^2 + 2a_1' x_1 x_2 + a_2' x_2^2.$$

10. For the present we shall confine our attention to invariants which are rational integral functions of the coefficients. It is easy to see that there is no further loss of generality if we suppose the invariants to be homogeneous in each set of coefficients that they contain.

Thus for example if I be an invariant of a single binary form f which is not homogeneous in the coefficients a we can write I in the form

$$I_1 + I_2 + \dots + I_s,$$

where each element in this sum is homogeneous.

Now by definition we have

$$I(A) = M \times I(a),$$

and therefore

$$I_1(A) + I_2(A) + \dots + I_s(A) = M\{I_1(a) + I_2(a) + \dots + I_s(a)\}.$$

But the A's are linear functions of the a's and M is independent of both, and therefore the only part on the left-hand side which is of the same degree as $I_1(a)$ on the right-hand side is $I_1(A)$;

$$\therefore \quad I_1(A) = M\, I_1(a),$$

that is to say I_1 is an invariant. Hence a non-homogeneous invariant is the sum of several homogeneous invariants.

This result can be at once extended to any number of binary forms.

As an example

$$a_0 a_2 - a_1^2 + a_0 a_2' - 2a_1 a_1' + a_2 a_0'$$

is an invariant of the two binary quadratics

$$(a_0, a_1, a_2 \between x_1, x_2)^2 \text{ and } (a_0', a_1', a_2' \between x_1, x_2)^2,$$

but it is the sum of two expressions

$$a_0 a_2 - a_1^2,$$

and $$a_0 a_2' - 2a_1 a_1' + a_2 a_0'$$

each of which is homogeneous in the two sets of coefficients.

11. Covariants of Binary Forms. If a binary form f is changed into a form F by a linear transformation, and a function C of the coefficients of F and the new variables X_1, X_2 be equal to the same function of the coefficients of f and the old variables x_1, x_2 multiplied by a factor depending only on the transformation, then C is called a covariant of the binary form.

Thus from what we have seen

$$\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$$

is a covariant of the binary cubic f and in fact of any binary form.

An exactly similar definition applies to a joint covariant of several binary forms—as an example the reader will have no difficulty in shewing that the Jacobian

$$\frac{\partial f}{\partial x_1}\frac{\partial \phi}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial \phi}{\partial x_1}$$

of any two forms f and ϕ is a covariant of those forms, the multiplier being $(\xi_1\eta_2 - \xi_2\eta_1)$ the determinant of the transformation.

We shall confine our attention to covariants which are rational integral functions both of the coefficients and the variables, and, as in the case of invariants, there is no difficulty in seeing that there is no further loss of generality in supposing such covariants to be homogeneous in the variables and in each set of coefficients involved. In fact if a covariant be not homogeneous it is the sum of several parts each of which is a covariant and homogeneous.

12. Degree and Order of a Covariant. The degree of a covariant of a single form is its degree in the coefficients of that form—the order is the degree in the variables.

The covariant $\dfrac{\partial^2 f}{\partial x_1^2}\dfrac{\partial^2 f}{\partial x_2^2} - \left(\dfrac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$ of a binary form of order n is of degree two and order $2n - 4$.

A covariant of several binary forms has a definite partial degree in each set of coefficients involved and the order is as before the degree in the variables.

The Jacobian of f and ϕ is of degree one in the coefficients of each of the two forms, and its order is the sum of the orders of f and ϕ diminished by two.

13. Symbolical Notation. In our investigations we shall find it of the utmost value to write the binary quantic

$$a_0 x_1^n + n a_1 x_1^{n-1} x_2 + \ldots + \binom{n}{r} a_r x_1^{n-r} x_2^r + \ldots + a_n x_2^n$$

in the symbolical form

$$(\alpha_1 x_1 + \alpha_2 x_2)^n,$$

so that $\alpha_1^n = a_0,\ \alpha_1^{n-1}\alpha_2 = a_1, \ldots \alpha_1^{n-r}\alpha_2^r = a_r, \ldots \alpha_2^n = a_n.$

This representation is startling at first sight, but consider how the use of it would introduce errors into calculation. They would arise because relations of the type

$$a_0 a_2 = \alpha_1^{2n-2}\alpha_2^2 = a_1^2$$

between the coefficients prevent our binary form from being a general one. Now in representing a function of the coefficients symbolically we allow no symbol such as α to occur more than n times in any one term, so that the possibility of relations giving rise to

$$a_0 a_2 = a_1^2$$

is entirely precluded. In fact to obtain this relation there must be $2n$ α's multiplied together in the representation of the function $a_0 a_2$ or a_1^2, whereas, when we allow no more than n α's to occur in any one term, the $(n+1)$ expressions

$$\alpha_1^n,\ \alpha_1^{n-1}\alpha_2, \dots \alpha_1^{n-r}\alpha_2^r, \dots \alpha_2^n$$

are independent quantities, *i.e.* with these restrictions on the use of our symbols the $(n+1)$ coefficients of the original quantic are not necessarily connected by any relation, and therefore the most general quantic can be represented in the form indicated.

Accordingly in addition to the symbol α we introduce a number of equivalent symbols β, γ, ... so that

$$f = (\alpha_1 x_1 + \alpha_2 x_2)^n = (\beta_1 x_1 + \beta_2 x_2)^n = (\gamma_1 x_1 + \gamma_2 x_2)^n = \dots$$

or as it will invariably be written

$$f = \alpha_x^n = \beta_x^n = \gamma_x^n \dots$$

The symbolical equivalent of $a_0 a_2$ is not

$$\alpha_1^{2n-2}\alpha_2^2,$$

because here there are more than n α's multiplied together.

To represent $a_0 a_2$ we must use two different symbols α, β and then

$$a_0 a_2 = \alpha_1^n \beta_1^{n-2}\beta_2^2,$$

which is of course equivalent to

$$\beta_1^n \alpha_1^{n-2}\alpha_2^2,$$

whereas in the same symbols a_1^2 is represented by $\alpha_1^{n-1}\alpha_2\beta_1^{n-1}\beta_2$.

In general to represent an expression of degree m in the coefficients, we have to use m different symbols of the type α, β, γ, ...

We have said that not more than n α's must be multiplied together in a given term—on the other hand if the expression has an actual as well as a symbolical significance not less than n of these symbols must occur together because only the expressions

$$\alpha_1^n,\ \alpha_1^{n-1}\alpha_2, \dots \alpha_2^n$$

have an actual meaning.