

whose determinant we call c , and denoting the coefficients of the transformed form by a'_1, a'_2, \dots , we have

$$(1) \quad I(a'_1, a'_2, \dots) \equiv c^\lambda I(a_1, a_2, \dots);$$

an identity which may also be written

$$f_1(a'_1, a'_2, \dots) \cdots f_l(a'_1, a'_2, \dots) \equiv c^\lambda f_1(a_1, a_2, \dots) \cdots f_l(a_1, a_2, \dots).$$

We have here a polynomial in the c 's and a 's which, on the second side of the identity, is resolved into its irreducible factors, since by Theorem 1, §61, the determinant c is irreducible. Hence each factor on the first side is equal to the product of some of the factors on the second. That is

$$(2) \quad f_i(a'_1, a'_2, \dots) \equiv c^{\lambda_i} \phi_i(a_1, a_2, \dots) \quad (i = 1, 2, \dots, l)$$

where the ϕ 's are polynomials.

Now let

$$c_{11} = c_{22} = \dots = c_{nn} = 1,$$

and let all the other c 's be zero. Our transformation becomes the identical transformation, the determinant $c = 1$, and each a' is equal to the corresponding a . The identities (2) therefore reduce to

$$f_i(a_1, a_2, \dots) \equiv \phi_i(a_1, a_2, \dots) \quad (i = 1, 2, \dots, l).$$

Substituting this value of ϕ_i in (2), we see that f_i is an invariant, and our theorem is proved.

The general theorem, now, is the following :

THEOREM 2. *If $I(a_1, a_2, \dots; b_1, b_2, \dots; \dots)$ is an integral rational invariant of the system of forms*

$$f_1(x_1, \dots, x_n; a_1, a_2, \dots)$$

$$f_2(x_1, \dots, x_n; b_1, b_2, \dots)$$

$$\dots \dots \dots$$

$$\dots \dots \dots$$

and is reducible, then all its factors are invariants.

The proof of this theorem is practically identical with that of Theorem 1.

CHAPTER XVII

GENERAL THEOREMS ON INTEGRAL RATIONAL INVARIANTS

77. The Invariance of the Factors of Invariants. Let us consider the general n -ary form of the k th degree which we will represent by $f(x_1, \dots, x_n; a_1, a_2, \dots)$, the x 's being the variables and the a 's the coefficients. By suitably changing the a 's, this symbol may be used to represent *any* such form. Hence, if we subject such a form to a linear transformation, the new form, being n -ary and of the same degree as the old, may be represented by the same functional letter : $f(x'_1, \dots, x'_n; a'_1, a'_2, \dots)$. This new form will evidently be homogeneous and linear in the a 's; that is, each of the a 's is a homogeneous linear polynomial in the a 's. It is also clear that each of the a 's is a homogeneous polynomial of the k th degree in the coefficients of the transformation.

It follows from the very definition of invariants that if we have a number of integral rational relative invariants of a form or system of forms, their product will also be an integral rational relative invariant. It is the converse of this that we wish to prove in this section. We begin by stating this converse in the simple case of a single form.

THEOREM 1. *If $I(a_1, a_2, \dots)$ is an integral rational invariant of the n -ary form*

$$f(x_1, \dots, x_n; a_1, a_2, \dots),$$

and is reducible, then all its factors are invariants.

It will evidently be sufficient to prove that the irreducible factors of I are invariants. Let f_1, f_2, \dots, f_l be the irreducible factors of I .

Subjecting f to the linear transformation

$$\begin{cases} x_1 = c_{11}x'_1 + \dots + c_{1n}x'_n, \\ \dots \dots \dots \\ x_n = c_{n1}x'_1 + \dots + c_{nn}x'_n, \end{cases}$$

EXERCISE

If $I(a_1, a_2, \dots; b_1, b_2, \dots; \dots; y_1, \dots, y_n; z_1, \dots, z_n; \dots)$ is an integral rational covariant of the system of forms

$$\begin{aligned} f_1(x_1, \dots, x_n; a_1, a_2, \dots), \\ f_2(x_1, \dots, x_n; b_1, b_2, \dots), \\ \dots \end{aligned}$$

and the system of points $(y_1, \dots, y_n), (z_1, \dots, z_n), \dots$, and is reducible, then all its factors are covariants (or invariants).

78. A More General Method of Approach to the Subject of Relative Invariants. We have called a polynomial I in the coefficients of an n -ary form f a relative invariant of this form if it has the property of being merely multiplied by a power of the determinant of the transformation when f is subjected to a linear transformation. It is natural to inquire what class of functions I we should obtain if we make the less specific demand that I be multiplied by a polynomial in the coefficients of the transformation. We should expect to get in this way a class of functions more general than the invariants we have so far considered. As a matter of fact, we get precisely the same class of functions, as is shown by the following theorem:

THEOREM. Let I be a polynomial not identically zero in the coefficients (a_1, a_2, \dots) of an n -ary form f , and let (a'_1, a'_2, \dots) be the coefficients of the form obtained by subjecting f to the linear transformation

$$\begin{cases} x_1 = c_{11}x'_1 + \dots + c_{1n}x'_n \\ \dots \\ x_n = c_{n1}x'_1 + \dots + c_{nn}x'_n \end{cases}$$

If $I(a'_1, a'_2, \dots) \equiv \psi(c_{11}, \dots, c_{nn}) I(a_1, a_2, \dots)$,

where ψ is a polynomial in the c 's, and this is an identity in the a 's and c 's, then ψ is a power of the determinant of the transformation.

We will first show that $\psi \neq 0$ when $c \neq 0$. If possible let d_{11}, \dots, d_{nn} be a particular set of values of the c_{ij} 's such that

$$\psi(d_{11}, \dots, d_{nn}) = 0,$$

while

$$\begin{vmatrix} d_{11} & \dots & d_{1n} \\ \dots & \dots & \dots \\ d_{n1} & \dots & d_{nn} \end{vmatrix} \neq 0.$$

Then the transformation

$$\begin{cases} x_1 = d_{11}x'_1 + \dots + d_{1n}x'_n \\ \dots \\ x_n = d_{n1}x'_1 + \dots + d_{nn}x'_n \end{cases}$$

has an inverse

$$\begin{cases} x'_1 = \delta_{11}x_1 + \dots + \delta_{1n}x_n \\ \dots \\ x'_n = \delta_{n1}x_1 + \dots + \delta_{nn}x_n \end{cases}$$

Let us consider a special set of a 's such that $I(a_1, a_2, \dots) \neq 0$. Then

$$I(a'_1, a'_2, \dots) = \psi(d_{11}, \dots, d_{nn}) I(a_1, a_2, \dots) = 0.$$

Now apply the inverse transformation, and we have

$$I(a_1, a_2, \dots) = \psi(\delta_{11}, \dots, \delta_{nn}) I(a'_1, a'_2, \dots) = 0,$$

which is contrary to our hypothesis.

Having thus proved that ψ can vanish only when $c = 0$, let us break up ψ into its irreducible factors,

$$\psi(c_{11}, \dots, c_{nn}) \equiv \psi_1(c_{11}, \dots, c_{nn}) \psi_2(c_{11}, \dots, c_{nn}) \dots \psi_\lambda(c_{11}, \dots, c_{nn}).$$

Since ψ vanishes whenever $\psi_i = 0$, ψ_i can vanish only when $c = 0$. Hence by the theorem for n variables which corresponds to Theorem 7, §76, ψ_i must be a factor of c . But c is irreducible. Hence ψ_i can differ from c only by a constant factor, and we may write

$$\psi \equiv Kc^\lambda.$$

It remains then merely to prove that the constant K has the value 1. For this purpose consider the identity

$$I(a'_1, a'_2, \dots) \equiv Kc^\lambda I(a_1, a_2, \dots),$$

and give to the c_{ij} 's the values which they have in the identical transformation. Then $c = 1$, and the a 's are equal to the corresponding a 's. The last written identity therefore becomes

$$I(a_1, a_2, \dots) \equiv KI(a_1, a_2, \dots);$$

from which we infer that $K = 1$.

EXERCISES

1. Prove that if a polynomial I in the coefficients a_1, a_2, \dots of an n -ary form and the coordinates (y_1, \dots, y_n) of a point has the property of being merely multiplied by a certain rational function ψ of the coefficients of the transformation when the form and the point are subjected to a linear transformation, then ψ is a positive or negative power of the determinant of the transformation, and I is a covariant.

2. Generalize the theorem of this section to the case of invariants of a system of forms.

3. Generalize the theorem of Exercise 1 to the case of a system of forms and a system of points.

4. Prove that every rational invariant of a form or system of forms is the ratio of two integral rational invariants.

5. Generalize the theorem of Exercise 4 to the case of covariants.

79. The Isobaric Character of Invariants and Covariants. In many investigations, and in particular in the study of invariants and covariants, it is desirable to attach a definite *weight* to each of the variables with which we have to deal. To a product of two or more such variables we then attach a weight equal to the sum of the weights of the factors, and this weight is supposed to remain unchanged if the product is multiplied by a constant coefficient. Thus if z_1, z_2, z_3 are regarded as having weights w_1, w_2, w_3 respectively, the term

$$5 z_1 z_2 z_3^2$$

would have the weight $w_1 + w_2 + 2w_3$.

If, then, having thus attached a definite weight to each of the variables, we consider a polynomial, each term of this polynomial will be of a definite weight, and *by the weight of a polynomial we understand the greatest weight of any of its terms whose coefficient is not zero*. If moreover all the terms of a polynomial are of the same weight, the polynomial is said to be *isobaric*.

It may be noticed that, according to this definition, a polynomial which vanishes identically is the only one which has no weight, while a polynomial which reduces to a constant different from zero is of weight zero. Moreover if two polynomials are of weights w_1 and w_2 , their product is of weight $w_1 + w_2$.*

* The conception of degree of a polynomial is merely the special case of the conception of weight in which all the variables are supposed to have weight 1. The conception of being isobaric then reduces to the conception of homogeneity.

We will apply this conception of weight first to the case in which the variables of which we have been speaking are the coefficients a_1, a_2, \dots of the n -ary form

$$f(x_1, \dots, x_n; a_1, a_2, \dots).$$

We shall find it desirable to admit n different determinations of the weights of these a 's; one determination corresponding to each of the variables x_1, \dots, x_n .

DEFINITION 1. If a_i is the coefficient of the term

$$a_i x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}$$

in an n -ary form, we assign to a_i the weights p_1, p_2, \dots, p_n respectively with regard to the variables x_1, x_2, \dots, x_n .

In the case of a binary form,

$$a_0 x_1^k + a_1 x_1^{k-1} x_2 + \dots + a_k x_2^k$$

the subscripts of the coefficients indicate their weights with regard to x_2 , while their weights with regard to x_1 are equal to the differences between the degree of the form and these subscripts.

As a second example, we mention the quadratic form

$$\sum_1^n a_{ij} x_i x_j.$$

Here the weight of any coefficient with regard to one of the variables, say x_j , is equal to the number of times j occurs as a subscript to this coefficient.*

In connection with this subject of weight, the special linear transformation

$$(1) \quad \begin{cases} x_i = x'_i \\ x_j = kx'_j \end{cases} \quad (i \neq j)$$

is useful. If a_i is a coefficient which is of weight λ with regard to x_j , the term in which this coefficient occurs contains x_j^λ , and therefore

$$a'_i = k^\lambda a_i.$$

* For forms of higher degree, a similar notation for the coefficients by means of multiple subscripts might be used. The weight of each coefficient could then be at once read off from the subscripts.

That is

THEOREM 1. *The weight with regard to x_j of a coefficient of an n -ary form is the exponent of the power of k by which this coefficient is multiplied after the special transformation (1).*

From this it follows at once that an isobaric polynomial of weight λ with regard to x_j in the coefficients (a_1, a_2, \dots) of an n -ary form is simply multiplied by k^λ if the form is subjected to the linear transformation (1).

Moreover, the converse of this is also true. For if a'_1, a'_2, \dots are the coefficients of the n -ary form after the transformation (1), and if $\phi(a_1, a_2, \dots)$ is a polynomial which has the property that

$$\phi(a'_1, a'_2, \dots) \equiv k^\lambda \phi(a_1, a_2, \dots),$$

this being an identity in the a 's and also in k , we can infer, as follows, that ϕ is isobaric of weight λ . Let us group the terms of ϕ together according to their weights, thus writing ϕ in the form

$$\phi(a_1, a_2, \dots) \equiv \phi_1(a_1, a_2, \dots) + \phi_2(a_1, a_2, \dots) + \dots$$

where ϕ_1, ϕ_2, \dots are isobaric of weights $\lambda_1, \lambda_2, \dots$. We have then

$$\phi(a'_1, a'_2, \dots) \equiv k^{\lambda_1} \phi_1(a_1, a_2, \dots) + k^{\lambda_2} \phi_2(a_1, a_2, \dots) + \dots$$

But on the other hand

$$\phi(a'_1, a'_2, \dots) \equiv k^\lambda \phi(a_1, a_2, \dots) \equiv k^\lambda \phi_1(a_1, a_2, \dots) + k^\lambda \phi_2(a_1, a_2, \dots) + \dots$$

Comparing the last members of these two identities, we see that

$$\lambda = \lambda_1 = \lambda_2 = \dots$$

as was to be proved. We have thus established the theorem:

THEOREM 2. *A necessary and sufficient condition that a polynomial ϕ in the coefficients of an n -ary form be simply multiplied by k^λ when the form is subjected to transformations of the form (1) is that ϕ be isobaric of weight λ with regard to x_j .*

By means of this theorem we can show that the use of the word *weight* introduced in § 31 is in accord with the definition given in the present section. For an integral rational invariant of an n -ary form which, according to the definition of § 31, is of weight λ will, if

the form is subjected to the transformation (1), be merely multiplied by k^λ and must therefore, according to Theorem 2, be isobaric of weight λ with regard to x_j . That is:

THEOREM 3. *If I is an integral rational invariant of a form f which according to the definition of § 31 is of weight λ , it will also be of weight λ with regard to each of the variables x_j of f according to the definitions of this section, and it will be isobaric with regard to each of these variables.*

As an illustration of this theorem we may mention the discriminant

$$a_0 a_2 - a_1^2$$

of the binary quadratic form

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

which is isobaric of weight 2 both with regard to x_1 and with regard to x_2 .

The reader should consider in the same way the discriminant of the general quadratic form.

All of the considerations of the present section may be extended immediately to the case in which we have to deal, not with a single form, but with a system of forms. We state here merely the theorem which corresponds to Theorem 3.

THEOREM 4. *If I is an integral rational invariant of a system of forms which according to the definition of § 31 is of weight λ , it will also be of weight λ with regard to each of the variables x_j of the system, and it will be isobaric with regard to each of these variables.*

The reader may consider as an illustration of this theorem the resultant of a system of linear forms, and also the invariants obtained in Chapters XII and XIII.

We saw in Theorem 5, § 31, that the weight of an integral rational invariant cannot be negative. This fact now becomes still more evident, since the weight of no coefficient is negative. Moreover, we can now add the following further fact:

THEOREM 5. *An integral rational invariant of a form or system of forms cannot be of weight zero.*

For consider any term of the invariant whose coefficient is not zero. This term involves the product of a number of coefficients of the system of forms. Since none of these coefficients can be of nega-

tive weight, the weight of the term will be at least as great as the weight of any one of them. But any one of them is at least of weight 1 with regard to some one of the variables. Hence the invariant is at least of weight 1 with regard to some one of the variables, and hence with regard to any of the variables.

In order, finally, to be able to extend the considerations of this section to the case of covariants, we must lay down the following additional definition:

DEFINITION 2. *If the sets of variables (y_1, \dots, y_n) , (z_1, \dots, z_n) , ... are cogredient with the variables (x_1, \dots, x_n) of a system of n -ary forms, we will assign to y_j, z_j, \dots the weight -1 with regard to x_j , to all the other y 's, z 's, etc. the weight 0.*

It will be noticed that here too, when we perform the transformation (1), each of the variables is multiplied by a power of k whose exponent is the weight of the variable. It is therefore easy* to extend the considerations of this section to this case, and we thus get the theorem:

THEOREM 6. *If I is an integral rational covariant of a system of forms and a system of points which is of weight λ according to the definition of § 31, it will also be of weight λ with regard to each of the variables of the system, and it will be isobaric with regard to each of these variables.*

As an example of this theorem we note that the polar

$$a_0 y_1 z_1 + a_1 (y_1 z_2 + y_2 z_1) + a_2 y_2 z_2$$

of a binary quadratic form is isobaric of weight zero. The reader may satisfy himself that the same is true of the polar of the general quadratic form.

80. Geometric Properties and the Principle of Homogeneity. It is a familiar fact that many geometric properties of plane curves or surfaces are expressed by the vanishing of an integral rational function of the coefficients of their equations. Take, for instance, the surface

$$(1) \quad f(x, y, z; a_1, a_2, \dots) = 0,$$

* Slight additional care must be taken here on account of the possible presence of terms of negative weight.

where f is a polynomial of the k th degree in the non-homogeneous coordinates x, y, z , and a_1, a_2, \dots are the coefficients of this polynomial; and consider the relation

$$(2) \quad \phi(a_1, a_2, \dots) = 0,$$

where ϕ is a polynomial, which we will assume to be of at least the first degree, in the coefficients a_1, a_2, \dots . By Theorem 3, § 6, there are an infinite number of polynomials of the k th degree in (x, y, z) whose coefficients satisfy the relation (2) and also an infinite number whose coefficients do not satisfy this relation. In other words, all polynomials of the k th degree in (x, y, z) may be divided into two classes, A and B , of which the first is characterized by condition (2) being fulfilled, while the second is characterized by this condition not being fulfilled. We may therefore say that (2) is a necessary and sufficient condition that f have a certain property, namely, the property of belonging in class A .

The simplest examples, however, show that this property of f need not correspond to a geometric property of the surface (1). To illustrate this, let $k=1$, so that we have

$$f \equiv a_1 x + a_2 y + a_3 z + a_4,$$

and consider first the polynomial in the a 's:

$$\phi \equiv a_4.$$

The vanishing of ϕ gives a necessary and sufficient condition that f belong to the class of homogeneous polynomials of the first degree in (x, y, z) , and thus expresses a property of the polynomial. This same condition, $a_4=0$, also expresses a property of the plane $f=0$, namely, the property that it pass through the origin.

Suppose, however, that instead of the function ϕ we take the polynomial

$$\phi_1 \equiv a_4 - 1.$$

The vanishing of this polynomial also gives a necessary and sufficient condition that the polynomial f have a certain property, namely, that its constant term have the value 1. It does not serve to distinguish planes into two classes, since we may write the equation of any plane (except those through the origin) either with the constant term 1 or with the constant term different from 1 by merely multiplying the equation through by a constant.

From the foregoing it will be seen that saying that a surface has a certain *property* amounts to the same thing as saying that it belongs to a certain class of surfaces.*

THEOREM 1. *The equation (2) expresses a necessary and sufficient condition for a geometric property of the surface (1) when, and only when, the polynomial ϕ is homogeneous.*

For if ϕ is non-homogeneous, let us write it in the form

$$\phi \equiv \phi_n + \phi_{n-1} + \dots + \phi_1 + \phi_0,$$

where ϕ_n is a homogeneous polynomial of the n th degree and each of the other ϕ 's which is not identically zero is a homogeneous polynomial of the degree indicated by its subscript. Let a'_1, a'_2, \dots be a set of values of the a 's for which ϕ_n and at least one of the other ϕ_i 's is not zero, and consider the surface

$$(3) \quad f(x, y, z; ca'_1, ca'_2, \dots) = 0.$$

The condition (2) for this surface is

$$c^n \phi_n(a'_1, a'_2, \dots) + c^{n-1} \phi_{n-1}(a'_1, a'_2, \dots) + \dots + c \phi_1(a'_1, a'_2, \dots) + \phi_0(a'_1, a'_2, \dots) = 0.$$

This is an equation of the n th degree in c , and since at least one of the coefficients after the first is different from zero, it will have at least one root $c_1 \neq 0$. On the other hand, we can find a value $c_2 \neq 0$ which is not a root of this equation. Hence the surface (3) satisfies condition (2) if we let $c = c_1$ and does not satisfy it if $c = c_2$. But a change in the value of c merely multiplies the equation (3) by a constant and does not change the surface represented by it. Thus we see that one and the same surface can be regarded both as satisfying and as not satisfying condition (2). In other words, if ϕ is non-homogeneous, (2) does not express a property of the surface (1).

Assume now that ϕ is homogeneous of the n th degree, and consider the class A of polynomials f whose coefficients satisfy equation (2) and the class B whose coefficients do not satisfy this equation. Our theorem will be proved if we can show that we have hereby divided the surfaces (1) into two classes, that is, that if

* This brief explanation must not be regarded as an attempt to define the conception *property*, for no specific class can be defined without the use of some property.

a'_1, a'_2, \dots are the coefficients of a polynomial of class A and a''_1, a''_2, \dots the coefficients of a polynomial of class B , then the two surfaces

$$f(x, y, z; a'_1, a'_2, \dots) = 0,$$

$$f(x, y, z; a''_1, a''_2, \dots) = 0,$$

cannot be the same. If they were the same, the coefficients a'_1, a'_2, \dots would be proportional to a''_1, a''_2, \dots (cf. Theorem 7, Corollary, § 76),

$$a''_1 = ca'_1, a''_2 = ca'_2, \dots$$

and therefore $\phi(a''_1, a''_2, \dots) = c^n \phi(a'_1, a'_2, \dots)$.

But this is impossible since by hypothesis

$$\phi(a'_1, a'_2, \dots) = 0, \quad \phi(a''_1, a''_2, \dots) \neq 0.$$

Thus our theorem is proved.

This theorem admits of generalization in various directions. Suppose first that instead of a single surface (1) we have a system of algebraic surfaces, and that ϕ is a polynomial in the coefficients of all these surfaces. Then precisely the reasoning just used shows that the equation $\phi = 0$ gives a necessary and sufficient condition for a geometric property of this system of surfaces when and only when ϕ is homogeneous in the coefficients of each surface taken separately.

On the other hand, we may use homogeneous coordinates in writing the equations of the surfaces, and the results so far stated will obviously hold without change:

THEOREM 2. *Let*

$$f_1(x, y, z, t; a_1, a_2, \dots), f_2(x, y, z, t; b_1, b_2, \dots), \dots$$

be a system of homogeneous polynomials in (x, y, z, t) whose coefficients are $a_1, a_2, \dots; b_1, b_2, \dots; \dots$; and let

$$\phi(a_1, a_2, \dots; b_1, b_2, \dots; \dots)$$

be a polynomial in the a 's, b 's, etc. Then the equation $\phi = 0$ expresses a necessary and sufficient condition that the system of surfaces

$$f_1 = 0, f_2 = 0, \dots$$

have a geometric property when, and only when, the polynomial ϕ is homogeneous in the a 's alone, also in the b 's alone, etc.

In conclusion we note that all the results of this section can be extended at once to algebraic curves in the plane; or, indeed, to the case of space of any number of dimensions.

EXERCISE

If, in Theorem 2, besides the surfaces $f_1 = 0, f_2 = 0, \dots$ we also have a system of points

$$(x_1, y_1, z_1, t_1), (x_2, y_2, z_2, t_2), \dots$$

and if ϕ is a polynomial not merely of the a 's, b 's, etc., but also of the coördinates of these points, prove that $\phi = 0$ expresses a necessary and sufficient condition that this system of surfaces and points have a geometric property when and only when ϕ is homogeneous in the a 's alone, in the b 's alone, etc., and also in (x_1, y_1, z_1, t_1) alone, in (x_2, y_2, z_2, t_2) alone, etc.

81. Homogeneous Invariants. From the developments of the last section it is clear that the only integral rational invariants which will be of importance in geometrical applications are those which are homogeneous in the coefficients of each of the ground-forms taken separately.* Such invariants we will speak of as *homogeneous invariants*. It will be found that all the invariants which we have met so far are of this kind.

An important relation between the weight and the various degrees connected with a homogeneous invariant is given by the following theorem :

THEOREM 1. *If we have a system of n -ary forms,*

$$(1) \quad \begin{cases} f_1(x_1, \dots, x_n; a_1, a_2, \dots), \\ f_2(x_1, \dots, x_n; b_1, b_2, \dots), \\ \dots \\ \dots \end{cases}$$

of degrees m_1, m_2, \dots respectively, and if

$$I(a_1, a_2, \dots; b_1, b_2, \dots; \dots)$$

* This statement must not be taken too literally. It is true if in the geometrical application in question we consider the variables as homogeneous coördinates and if we have to deal with the loci obtained by equating the ground-forms to zero. While this is the ordinary way in which we interpret invariants geometrically, other interpretations are possible. For instance, instead of interpreting the variables (x, y) as homogeneous coördinates on a line and equating the binary quadratic forms

$$\begin{aligned} f_1 &\equiv a_1x^2 + 2a_2xy + a_3y^2, \\ f_2 &\equiv b_1x^2 + 2b_2xy + b_3y^2, \end{aligned}$$

to zero, thus getting two pairs of points on a line, we may interpret (x, y) as non-homogeneous coördinates in the plane, and consider the two conics $f_1 = 1, f_2 = 1$. With this interpretation, the vanishing of the invariant

$$a_1a_3 - a_2^2 + b_1b_3 - b_2^2,$$

which is not homogeneous in the a 's alone or in the b 's alone, has a geometric meaning

is a homogeneous invariant of this system, of weight λ , and of degree α in the a 's, β in the b 's, etc., then

$$(2) \quad m_1\alpha + m_2\beta + \dots = n\lambda.$$

Subjecting the forms (1) to the linear transformation

$$(3) \quad \begin{cases} x_1 = c_{11}x'_1 + \dots + c_{1n}x'_n, \\ \dots \\ x_n = c_{n1}x'_1 + \dots + c_{nn}x'_n, \end{cases}$$

whose determinant we will denote by c , we get

$$\begin{aligned} &f_1(x'_1, \dots, x'_n; a'_1, a'_2, \dots), \\ &f_2(x'_1, \dots, x'_n; b'_1, b'_2, \dots), \\ &\dots \end{aligned}$$

and, since by hypothesis I is an invariant of weight λ ,

$$(4) \quad I(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) \equiv c^\lambda I(a_1, a_2, \dots; b_1, b_2, \dots; \dots).$$

Every a is a homogeneous polynomial in the c_{ij} 's of degree m_1 , every b of degree m_2 , etc.; and since I is itself homogeneous of degree α in the a 's, β in the b 's, etc., we see that the left-hand side of (4) is a homogeneous polynomial of degree $m_1\alpha + m_2\beta + \dots$ in the c_{ij} 's. Equating this to the degree of the right-hand side of (4) in the c_{ij} 's, which is evidently $n\lambda$, our theorem is proved.

An additional reason for the importance of these homogeneous invariants is that the non-homogeneous integral rational invariants can be built up from them, as is stated in the following theorem:

THEOREM 2. *If an integral rational invariant I of the system (1) be written in the form*

$$I \equiv I_1 + I_2 + \dots + I_k$$

where each of the I_i 's is a polynomial in the a 's, b 's, etc., which is homogeneous in the a 's alone, and also in the b 's alone, etc., and such that the sum of no two I_i 's has this property, then each of the functions

$$I_1, I_2, \dots, I_k$$

is a homogeneous invariant of the system (1).

This theorem follows immediately from the definition of an invariant. For from the identity,

$$I_1(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) + \dots + I_k(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) \\ \equiv c^\lambda [I_1(a_1, a_2, \dots; b_1, b_2, \dots; \dots) \\ + \dots + I_k(a_1, a_2, \dots; b_1, b_2, \dots; \dots)],$$

we infer at once the identities,

$$I_1(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) \equiv c^\lambda I_1(a_1, a_2, \dots; b_1, b_2, \dots; \dots), \\ \dots \\ I_k(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) \equiv c^\lambda I_k(a_1, a_2, \dots; b_1, b_2, \dots; \dots).$$

In the case of a single n -ary form, but in that case only, we have the theorem:

THEOREM 3. *An integral rational invariant of a single n -ary form is always homogeneous.*

Let $f(x_1, \dots, x_n; a_1, a_2, \dots)$

be the ground-form, and let I be the invariant. By Theorem 2 we may write

$$I \equiv I_1 + I_2 + \dots + I_k$$

where I_1, \dots, I_k are homogeneous invariants. Let the degrees of these homogeneous invariants in the a 's be $\alpha_1, \dots, \alpha_k$ respectively. Their weights are all the same as the weight of I , which we will call λ . If, then, we call the degree of f , m , we have, by Theorem 1,

$$m\alpha_1 = n\lambda, m\alpha_2 = n\lambda, \dots, m\alpha_k = n\lambda,$$

from which, since $m > 0$, we infer

$$\alpha_1 = \alpha_2 = \dots = \alpha_k.$$

That is, I_1, \dots, I_k are of the same degree, and I is homogeneous.

THEOREM 4. *If we have a system of n -ary forms f_1, f_2, \dots and a polynomial ϕ in their coefficients, the equation $\phi = 0$ gives a necessary and sufficient condition for a projective property of the system of loci in space of $n - 1$ dimensions,*

$$f_1 = 0, f_2 = 0, \dots,$$

when, and only when, ϕ is a homogeneous invariant of the system of forms f .

If ϕ is a homogeneous invariant, its vanishing gives a necessary and sufficient condition for a geometric property (cf. § 80), and this property must be a projective property since when we subject the loci to a non-singular collineation, ϕ is merely multiplied by a non-vanishing constant.

On the other hand let $\phi = 0$ be a necessary and sufficient condition for a projective property. In order to prove that ϕ is an invariant (it must be homogeneous by § 80) let a_1, a_2, \dots be the coefficients of f_1 ; b_1, b_2, \dots the coefficients of f_2 , etc.; and suppose that the linear transformation,

$$(5) \quad \begin{cases} x_1 = c_{11}x'_1 + \dots + c_{1n}x'_n, \\ \dots \\ x_n = c_{n1}x'_1 + \dots + c_{nn}x'_n, \end{cases}$$

carries over f_1 into f'_1 with coefficients a'_1, a'_2, \dots ; f_2 into f'_2 with coefficients b'_1, b'_2, \dots ; etc. The polynomial ϕ formed for the transformed forms is

$$\phi(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots),$$

and may, since the a 's, b 's, \dots are polynomials in the a 's, b 's, \dots and the c 's, be itself regarded as a polynomial in the a 's, b 's, \dots and the c 's. Looking at it from this point of view, let us resolve it into its irreducible factors,

$$(6) \quad \phi(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) \equiv \phi_1(a_1, a_2, \dots; b_1, b_2, \dots; \dots c_{11}, \dots c_{nn}) \\ \dots \phi_k(a_1, a_2, \dots; b_1, b_2, \dots; \dots c_{11}, \dots c_{nn}).$$

It is clear that at least one of the factors on the right must contain the c 's. Let ϕ_1 be such a factor, and let us arrange it as a polynomial in the c 's whose coefficients are polynomials in the a 's, b 's, etc. Let

$$\psi(a_1, a_2, \dots; b_1, b_2, \dots; \dots)$$

be one of these coefficients which is not identically zero and which is the coefficient of a term in which at least one of the c 's has an exponent greater than zero. We can, now, give to the a 's, b 's, \dots values which we will denote by A 's, B 's, \dots such that neither ϕ nor ψ vanish; and consider a neighborhood N of the point

$$(A_1, A_2, \dots; B_1, B_2, \dots; \dots)$$

throughout which

$$(7) \quad \phi(a_1, a_2, \dots; b_1, b_2, \dots; \dots) \neq 0,$$

$$(8) \quad \psi(a_1, a_2, \dots; b_1, b_2, \dots; \dots) \neq 0.$$

Let us now restrict the a 's, b 's, ... to the neighborhood N and ask ourselves under what circumstances we can have $\phi_1 = 0$. If this equation is fulfilled, we see from (6) that ϕ vanishes for the transformed loci, while, by (7), it does not vanish for the original loci. Since, by hypothesis, the vanishing of ϕ gives a necessary and sufficient condition for a projective property, a transformation (5) which causes ϕ to vanish when it did not vanish before must be a singular transformation. That is, *if the a 's, b 's, ... are in the neighborhood N , whenever ϕ_1 vanishes the determinant c of (5) vanishes.* Moreover, ϕ_1 does vanish for values of the a 's, b 's, ... in N , for if we assign to the a 's, b 's, ... any such values, ϕ_1 becomes a polynomial in the c_{ij} 's, which, by (8), is of at least the first degree, and therefore vanishes for suitably chosen values of the c_{ij} 's. We can therefore apply the theorem for more than three variables analogous to Theorem 8, § 76, and infer that ϕ_1 is a factor of the determinant c ; and consequently, since this determinant is irreducible (Theorem 1, § 61), that ϕ_1 is merely a constant multiple of c .

The reasoning we have just applied to ϕ_1 applies equally to any of the factors on the right of (6) which are of at least the first degree in the c_{ij} 's. Accordingly (6) reduces to the form

$$(9) \quad \phi(a'_1, a'_2, \dots; b'_1, b'_2, \dots; \dots) \equiv c^\lambda \chi(a_1, a_2, \dots; b_1, b_2, \dots; \dots),$$

where χ no longer involves the c_{ij} 's. To determine this polynomial χ , let us assign to the c_{ij} 's the values 0, 1 which reduce (5) to the identical transformation. Then the a 's, b 's, ... reduce to the a 's, b 's ..., while $c = 1$; so that from (9) we see that

$$\phi(a_1, a_2, \dots; b_1, b_2, \dots; \dots) \equiv \chi(a_1, a_2, \dots; b_1, b_2, \dots; \dots).$$

Substituting this value of χ in (9), we see that ϕ is really an invariant.

In order to avoid all misunderstanding, we state here explicitly that if we have two or more polynomials, ϕ_1, ϕ_2, \dots in the coefficients of the forms f_i , the equations $\phi_1 = \phi_2 = \dots = 0$ may be a necessary and sufficient condition for a projective property of the loci $f_i = 0$, even though ϕ_1, ϕ_2, \dots are not invariants. For instance, a necessary and sufficient condition that the two lines

$$a_1x_1 + a_2x_2 + a_3x_3 = 0,$$

$$b_1x_1 + b_2x_2 + b_3x_3 = 0$$

coincide is the vanishing of the three two-rowed determinants of the matrix

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

none of which is an invariant. Or, again, a necessary and sufficient condition that a quadric surface break up into two planes, distinct or coincident, is the vanishing of all the three-rowed determinants of its matrix, and these are not invariants. In this case we can also express the condition in question by the *identical* vanishing of a certain contravariant, namely, the adjoint of the quadratic form; and this — a projective relation expressed by the identical vanishing of a covariant or contravariant — is typical of what we shall usually have when a single equation $\phi = 0$ is not sufficient to express the condition. There are, however, cases where the condition is given by the vanishing of two or more invariants; cf. Exercise 6, § 90.

EXERCISES

1. Prove that if in Theorem 1 our system consists not merely of the ground-forms (1) but also of certain points

$$(y_1, \dots, y_n), (z_1, \dots, z_n), \dots,$$

and we have not an invariant I but a covariant of weight λ , and of degree a in the a 's, β in the b 's, etc., η in the y 's, ζ in the z 's, etc., then

$$m_1a + m_2\beta + \dots = n\lambda + \eta + \zeta + \dots$$

2. Extend Theorem 2 to the case of covariants. Does Theorem 3 admit of such extension?

3. Extend Theorem 4 to the case of covariants.

4. Show that an integral rational invariant of a single binary form of odd degree must be of even degree.

5. Show that the weight of an integral rational invariant of a single binary form can never be smaller than the degree of the form.

6. Express the condition that (a) two lines, and (b) two planes coincide, in the form of the identical vanishing of a covariant or contravariant.

7. Prove that a polynomial in the coefficients of a system of n -ary forms which is homogeneous in the coefficients of each form taken by themselves, and which is unchanged when the forms are subjected to any linear transformation of determinant + 1, is an invariant of the system of forms.

8. Generalize Exercise 7 to the case of covariants.

82. Resultants and Discriminants of Binary Forms. If we interpret (x_1, x_2) as homogeneous coördinates in one dimension, the equations obtained by equating the two binary forms

$$\begin{aligned} f(x_1, x_2) &\equiv a_0 x_1^n + a_1 x_1^{n-1} x_2 + \dots + a_n x_2^n, \\ \phi(x_1, x_2) &\equiv b_0 x_1^m + b_1 x_1^{m-1} x_2 + \dots + b_m x_2^m \end{aligned}$$

to zero represent sets of points on a line. The points given by the equation $f=0$ are the points at which the linear factors of f vanish, and the points corresponding to $\phi=0$ are the points at which the linear factors of ϕ vanish. Since two binary linear forms obviously vanish at the same point when, and only when, these linear forms are proportional, it follows that the loci of the two equations $f=0$, $\phi=0$ have a point in common when, and only when, f and ϕ have a common factor other than a constant. Hence, by § 72, a necessary and sufficient condition that the two loci $f=0$, $\phi=0$ have a point in common is that the resultant R of the binary forms f , ϕ vanish.

The property of these two loci having a point in common is, however, a projective property. Thus, by Theorem 4, § 81,

THEOREM 1. *The resultant of two binary forms is a homogeneous invariant of this pair of forms.*

From the determinant form of R given in § 68 it is clear that R is of degree m in the a 's and of degree n in the b 's. Hence by formula (2), § 81,

$$\lambda = mn.$$

THEOREM 2. *The weight of the resultant of two binary forms of degrees m and n is mn .*

The following geometrical problem will lead us to an important invariant of a single binary form.

Let us resolve the form f , which we assume not to be identically zero, into its linear factors (cf. formula (4), § 65),

$$f(x_1, x_2) \equiv (a'_1 x_1 - a'_1 x_2)(a'_2 x_1 - a'_2 x_2) \dots (a'_n x_1 - a'_n x_2).$$

The equation $f=0$ represents n distinct points provided no two of these linear factors are proportional to each other. If, however, two of these factors are proportional, we say that f has a multiple linear factor, and in this case two or more of the n points represented by the equation $f=0$ coincide. Let us inquire under what conditions this will occur.

Form the partial derivatives :

$$(1) \quad \begin{cases} \frac{\partial f}{\partial x_1} \equiv a''_1 (a'_2 x_1 - a'_2 x_2) \dots (a'_n x_1 - a'_n x_2) \\ \quad \quad \quad + a''_2 (a'_1 x_1 - a'_1 x_2)(a'_3 x_1 - a'_3 x_2) \dots (a'_n x_1 - a'_n x_2) \\ \quad \quad \quad + \dots + a''_n (a'_1 x_1 - a'_1 x_2) \dots (a'_{n-1} x_1 - a'_{n-1} x_2), \\ \frac{\partial f}{\partial x_2} \equiv -a'_1 (a'_2 x_1 - a'_2 x_2) \dots (a'_n x_1 - a'_n x_2) \\ \quad \quad \quad - a'_2 (a'_1 x_1 - a'_1 x_2)(a'_3 x_1 - a'_3 x_2) \dots (a'_n x_1 - a'_n x_2) \\ \quad \quad \quad - \dots - a'_n (a'_1 x_1 - a'_1 x_2) \dots (a'_{n-1} x_1 - a'_{n-1} x_2). \end{cases}$$

From these formulæ we see that any multiple linear factor of f is a factor of both of these partial derivatives.

Conversely, if these partial derivatives have a common linear factor, it must be a factor of f on account of the formula,

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} \equiv n f, *$$

a formula which follows immediately from the expressions,

$$(2) \quad \begin{cases} \frac{\partial f}{\partial x_1} \equiv n a_0 x_1^{n-1} + (n-1) a_1 x_1^{n-2} x_2 + \dots + a_{n-1} x_2^{n-1}, \\ \frac{\partial f}{\partial x_2} \equiv a_1 x_1^{n-1} + 2 a_2 x_1^{n-2} x_2 + \dots + n a_n x_2^{n-1}. \end{cases}$$

But, by (1), no linear factor of f can be a factor of $\partial f / \partial x_1$ unless it is a multiple factor of f . Thus we have proved

THEOREM 3. *A necessary and sufficient condition that f have a multiple linear factor is that the resultant of $\partial f / \partial x_1$ and $\partial f / \partial x_2$ vanish.*

DEFINITION. *The resultant of $\partial f / \partial x_1$ and $\partial f / \partial x_2$ is called the discriminant of f .*

From (2) we see that the discriminant of f may be written as a determinant of order $2n-2$ whose elements, so far as they are not zero, are numerical multiples of the coefficients a_0, a_1, \dots, a_n of f . That is, this discriminant is a polynomial in the a 's. Moreover, its vanishing gives a necessary and sufficient condition that the locus $f=0$ have a projective property (namely, that two points of this locus coincide). Hence, by Theorem 4, § 81, this discriminant is a

* This is merely Euler's Theorem for Homogeneous Functions.

homogeneous invariant, whose degree and weight are readily determined. Thus we get the theorem :

THEOREM 4. *The discriminant of a binary form of the n th degree is a homogeneous invariant of this form of degree $2(n-1)$ and of weight $n(n-1)$.*

A slight modification in the definition of the discriminant is often desirable. Let us write the binary form f , not in the above form where the coefficients are a_0, a_1, \dots, a_n , but, by the introduction of binomial coefficients, in the form

$$f(x_1, x_2) \equiv a_0 x_1^n + n a_1 x_1^{n-1} x_2 + \frac{n(n-1)}{2!} a_2 x_1^{n-2} x_2^2 + \dots + n a_{n-1} x_1 x_2^{n-1} + a_n x_2^n.$$

Then we may write

$$\frac{1}{n} \frac{\partial f}{\partial x_1} \equiv a_0 x_1^{n-1} + (n-1) a_1 x_1^{n-2} x_2 + \frac{(n-1)(n-2)}{2!} a_2 x_1^{n-3} x_2^2 + \dots + a_{n-1} x_2^{n-1},$$

$$\frac{1}{n} \frac{\partial f}{\partial x_2} \equiv a_1 x_1^{n-1} + (n-1) a_2 x_1^{n-2} x_2 + \frac{(n-1)(n-2)}{2!} a_3 x_1^{n-3} x_2^2 + \dots + a_n x_2^{n-1}.$$

We may then define the discriminant of f as the resultant of the two binary forms just written. We thus get for the discriminant a polynomial in the a 's which differs from the discriminant as above defined only by a numerical factor, and for which Theorems 3 and 4 obviously still hold. If this last definition be applied to the case of a binary quadratic form, it will be seen that it leads us precisely to what we called the discriminant of this quadratic form in the earlier chapters of this book.

EXERCISES

1. Prove that the resultant of two binary forms of degrees n and m respectively is irreducible.

[SUGGESTION. When $b_0 = 0$, R is equal to a_0 times the resultant of two binary forms of degrees n and $m-1$ respectively. Show that if this last resultant is irreducible, R is also irreducible, and use the method of induction, starting with the case $n=1, m=1$.]

2. Prove by the methods of this chapter that the bordered determinants of Chapter XII are invariants of weight 2.

3. The following account of Bézout's method of elimination is sometimes given:

If f and ϕ are polynomials in x which are both of degree n , the expression

$$f(x)\phi(y) - \phi(x)f(y)$$

vanishes, independently of y , for a value of x for which both f and ϕ vanish, and is divisible by $x-y$, since it is zero for $x=y$. Hence

$$F(x, y) = \frac{f(x)\phi(y) - \phi(x)f(y)}{x-y}$$

is a polynomial of degree $(n-1)$ in x which vanishes for all values of y when x is a common root of f and ϕ . Arranging F according to the powers of y , we have the expression

$$F = c_{00} + c_{01}x + c_{02}x^2 + \dots + c_{0,n-1}x^{n-1} \\ + y(c_{10} + c_{11}x + c_{12}x^2 + \dots + c_{1,n-1}x^{n-1}) \\ + y^2(c_{20} + c_{21}x + c_{22}x^2 + \dots + c_{2,n-1}x^{n-1}) \\ + \dots \\ + y^{n-1}(c_{n-1,0} + c_{n-1,1}x + c_{n-1,2}x^2 + \dots + c_{n-1,n-1}x^{n-1}).$$

If this function is to vanish independently of y , the coefficient of each power of y must be zero. This gives n equations between which we can eliminate the n quantities, $1, x, x^2, \dots, x^{n-1}$, obtaining the resultant in the form of the determinant,

$$R = \begin{vmatrix} c_{00} & c_{01} & \dots & c_{0,n-1} \\ c_{10} & c_{11} & \dots & c_{1,n-1} \\ \dots & \dots & \dots & \dots \\ c_{n-1,0} & c_{n-1,1} & \dots & c_{n-1,n-1} \end{vmatrix} = 0.$$

With the help of the auxiliary function F we have, in this case, reduced the resultant to a determinant of the n th order, while that obtained by the method of Sylvester was of order $2n$.

Criticise this treatment and make it rigorous, applying it, in particular, to the case of homogeneous variables.

4. If f and ϕ are polynomials in (x, y) of degrees n and m respectively and are relatively prime, prove that the curves $f=0, \phi=0$ cannot have more than mn points of intersection.

[SUGGESTION. Show first that the coordinate axes can be turned in such a way that no two points of intersection have the same abscissa, and that the equations of the two curves are of degrees n and m respectively, after the transformation, in y alone. Then eliminate y between the two equations by Sylvester's dialytic method.]

5. Prove that every integral rational invariant of the binary cubic is a constant multiple of a power of the discriminant.

[SUGGESTION. Show that if the discriminant is not zero, every binary cubic can be reduced by a non-singular linear transformation to the normal form $x_1^3 - x_2^3$. Then as in § 48.]