ETERMINATION OF ALL GENERAL HOMOGENEOUS POLYNOMIALS EXPRESSIBLE AS DETERMINANTS WITH LINEAR ELEMENTS*

BY

LEONARD EUGENE DICKSON

1. The general quadratic forms in three and four variables can be transformed into $x_1 x_2 - x_3^2$ and $x_1 x_2 - x_3 x_4$ respectively, and hence are expressible as determinants of order 2. Since any binary form of degree r is a product of r linear forms, it is expressible as an r-rowed determinant whose elements outside the main diagonal are all zero.

It was proved geometrically by H. Schröter† and more simply by L. Cremona‡ that a sufficiently general cubic surface $f = 0$ is the locus of the intersections of corresponding planes of three projective bundles of planes:

$$\kappa l_{i1} + \lambda l_{i2} + \mu l_{i3} = 0 \qquad (i = 1, 2, 3),$$

where κ, λ, μ are parameters and the l_{ij} are linear homogeneous functions§ of $x_1, \cdots, x_4$. Hence $f = 0$ has the determinantal form $|l_{ij}| = 0$. Taking $x_4 = 0$, we see that a general cubic curve is expressible in determinantal form.

I shall prove that every plane curve is expressible in determinantal form and that, aside from the cases mentioned above, no further general homogeneous polynomial is expressible in determinantal form.

The case of quartic surfaces was discussed erroneously by Jessop.|| His argument would apply equally well to the determinant D whose 16 elements are binary linear forms and show that D can be given a form containing a single parameter, whereas every binary quartic can be expressed in the form D.

A new theory of equivalence of pairs of bilinear forms is given in § 9.

2. THEOREM 1. *When the number of terms in the general form of degree r in n variables $(n > 2)$ exceeds $(n - 2)r^2 + 2$, it is not expressible as a determinant whose elements are linear forms.*

* Presented to the Society at Chicago, December 28, 1920.

† Journal für Mathematik, vol. 62 (1863), p. 265.

‡ *Ibid.*, vol. 68 (1868), p. 79.

§ In case their coefficients are rational we obtain all rational solutions of the Diophantine equation $f = 0$ by solving our three linear equations for the ratios of $x_1, \cdots, x_4$, obtaining cubic functions of κ, λ, μ.

|| *Quartic Surfaces*, 1916, p. 160.

Let D be any r-rowed determinant whose elements are linear homogeneous functions of $x_1, \cdots, x_n$. We may express the matrix M of D in the form $x_1 M_1 + \cdots + x_n M_n$, where each M_i is a matrix whose r^2 elements are constants. Evidently D is at most multiplied by a constant not zero if we interchange any two rows or any two columns, or multiply the elements of any row or column by a constant not zero, or add to the elements of any row (or column) the products of the elements of any other row (or column) by a constant. The effect on M of any succession of such "elementary transformations" is known to be the same as forming the product AMB, where A and B are constant matrices whose determinants are not zero.

If the determinant of M_1 is zero, D lacks x_1^r and will not represent the general form. Hence M_1 has an inverse M_1^{-1} such that $M_1 M_1^{-1}$ is the identity matrix I. Consider the new matrix

$$X = MM_1^{-1} = x_1 I + x_2 X_2 + \cdots + x_n X_n \quad (X_i \equiv M_i M_1^{-1}),$$

whose determinant equals the quotient of D by $|M_1|$ and has unity as the coefficient of x_1^r. The product AXB will likewise have I as the coefficient of x_1 if and only if $A = B^{-1}$. Our next step is therefore to choose matrix B so that $B^{-1} X_2 B$ shall have a canonical form. We first interpret this product. If X_2 is the matrix of the transformation

$$\xi_i' = \sum_{j=1}^{r} \alpha_{ij} \xi_j \qquad (i = 1, \cdots, r), \tag{1}$$

and if we introduce new variables $\eta_1, \cdots, \eta_r$ by means of a transformation

$$\eta_i = \sum_{k=1}^{r} \beta_{ik} \xi_k \qquad (i = 1, \cdots, r),$$

whose matrix is B, transformation (1) becomes a transformation on the η's whose matrix is easily verified* to be $B^{-1} X_2 B$. Naturally we desire that as many as possible of the η's shall be transformed into mere multiples of themselves. Hence we seek functions $\eta = \Sigma \beta_k \xi_k$ such that $\eta' = \lambda \eta$ under transformation (1); the conditions are evidently

$$\sum_{i=1}^{r} \beta_i \alpha_{ij} = \lambda \beta_j \qquad (j = 1, \cdots, r).$$

Thus λ must be a root of the characteristic equation

$$\begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} & \cdots & \alpha_{1r} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rr} - \lambda \end{vmatrix} = 0$$

* Dickson, *Linear Groups*, Teubner, 1901, p. 80.

of transformation (1). When λ is any root, the above conditions are known to have solutions $\beta_1, \cdots, \beta_r$ not all zero, so that $\eta' = \lambda\eta$.

If in the last determinant we replace λ by $-x_1/x_2$ and multiply the elements of each row by x_2, we obtain

$$\begin{vmatrix} x_2\alpha_{11} + x_1 & x_2\alpha_{12} & \cdots & x_2\alpha_{1r} \\ \cdot & \cdot & \cdot & \cdot \\ x_2\alpha_{r1} & x_2\alpha_{r2} & \cdots & x_2\alpha_{rr} + x_1 \end{vmatrix},$$

which is the determinant of matrix $x_2 N_2 + x_1 I$ and hence is the value of D when $x_3 = 0, \cdots, x_n = 0$. The latter must be the general binary form in x_1, x_2, since D is required to represent the general form in $x_1, \cdots, x_n$. Hence the above characteristic equation must have r distinct roots $\lambda_1, \cdots, \lambda_r$.

We thus obtain r linear functions $\eta_1, \cdots, \eta_r$ which (1) multiplies by $\lambda_1, \cdots, \lambda_r$ respectively. A simple artifice* shows that these η's are linearly independent functions of $\xi_1, \cdots, \xi_r$ and hence may be taken as new variables to give the desired canonical form

$$(2) \qquad \eta_i' = \lambda_i \eta_i \qquad (i = 1, \cdots, r).$$

Denote its matrix $B^{-1} N_2 B$ by P_2. Similarly, denote $B^{-1} N_j B$ by P_j. Hence our matrix N (and thus M) has been reduced to

$$P = x_1 I + x_2 P_2 + \cdots + x_n P_n.$$

For $n > 2$, the further normalization of P is to be accomplished by means of a matrix K such that $K^{-1} P_2 K = P_2$. But the only linear transformation commutative with (2), in which $\lambda_1, \cdots, \lambda_r$ are distinct, is seen at once to be

$$(3) \qquad \eta_i' = k_i \eta_i \qquad (i = 1, \cdots, r),$$

where the k's need not be distinct, but each is $\neq 0$. If

$$\eta_i' = \sum_{j=1}^{r} c_{ij} \eta_j \qquad (i = 1, \cdots, r)$$

is the transformation whose matrix is P_3, and if K is the matrix of (3), then, in view of the above interpretation, $K^{-1} P_3 K$ is the matrix of

$$\zeta_i' = \sum_{j=1}^{r} \frac{k_i}{k_j} c_{ij} \zeta_j \qquad (i = 1, \cdots, r;\ \zeta_i = k_i \eta_i).$$

Thus the maximum simplification possible in $P_3, \cdots, P_n$ is to make $r - 1$ non-vanishing elements take the value unity, so that at most $(n-2)r^2 - (r-1)$ parameters appear in their canonical forms. Taking account also of $\lambda_1, \cdots, \lambda_r$ and of the factor initially removed from D, we conclude

* *Ibid.*, p. 222.

that D can be given a form containing at most $(n-2)r^2+2$ parameters. A generalization of Theorem 1 is given in § 10.

3. The number of terms in the general form of degree r in n variables is known to be the binomial coefficient

$$\binom{r+n-1}{r}=\frac{(r+n-1)!}{r!(n-1)!}.$$

This follows by two-fold induction since in

$$\binom{r+n-2}{r-1}+\binom{r+n-2}{r}=\binom{r+n-1}{r},$$

the first symbol therefore enumerates the terms with the factor x_n and the second symbol enumerates the terms lacking x_n.

4. By §§ 2–3, the general form of degree r in n variables $(n>2)$ is not expressible in determinantal form if

$$\binom{r+n-1}{r}>(n-2)r^2+2. \tag{4}$$

If $r=2$, this condition reduces to $(n-3)(n-4)>0$. If $r=3$, it is

$$n^3+3n^2-52n+96>0$$

and holds when $n\geqq 5$. If $r\geqq 4$ and $n\geqq 7$, we have

$$\binom{r+n-1}{n-1}>\frac{(n+3)(n+2)(n+1)n\cdots 7(r+2)(r+1)}{(n-1)!},$$

since we have replaced r by 4 in all but the last two factors. The factors $n-1, n-2, \cdots, 7$ (which are absent if $n=7$) may be cancelled. We get

$$(n+3)(n+2)(n+1)n(r+2)(r+1)/6!,$$

which will exceed $t=(n-2)r^2+2$, since $(n-2)(r+2)(r+1)>t$, provided

$$(n+3)(n+2)(n+1)n/6!\geqq n-2.$$

The latter may be written in the form

$$(n-7)(n^3+13n^2+102n)+1440\geqq 0.$$

It remains to treat the cases $n\leqq 6$. For $n=3$, (4) fails if $r\geqq 2$, since $\binom{r+2}{2}\leqq r^2+2$ for $(r-1)(r-2)\geqq 0$. But for $n=4$, (4) holds if $(r-1)(r-2)(r-3)>0$. For $n=5$ and $n=6$, (4) becomes respectively

$$r^4+10r^3-37r^2+50r-24>0,$$

$$r^5+15r^4+85r^3-255r^2+274r-120>0,$$

each of which evidently holds if $r \geqq 3$. Hence we have

THEOREM 2. *The general form of degree r in n variables, $n > 2$, is not expressible in determinantal form if $r = 2$ or 3, $n > 4$, and if $r \geqq 4$, $n \geqq 4$, and hence unless $n = 3$, r any, or $n = 4$, $r = 2$ or 3.*

5. Since general quadric and cubic surfaces are expressible in determinantal form (§ 1), there remains only the case $n = 3$.

For $n = r = 3$, we may employ the canonical forms,* omitting those with a linear factor (§ 6):

$$x^3 + y^3 + z^3 - mxyz \equiv \begin{vmatrix} x & y & z \\ z & x & ay \\ a^{-1}y & z & x \end{vmatrix}, \qquad a + \frac{1}{a} = m - 1,$$

$$x^3 + y^3 - xyz \equiv \begin{vmatrix} x & y & z \\ 0 & x & y \\ y & 0 & \cdot x \end{vmatrix}, \qquad x^3 + yz^2 \equiv \begin{vmatrix} x & y & 0 \\ 0 & x & z \\ z & 0 & x \end{vmatrix}.$$

I obtained the determinants by inspection.

For $n = 3$, $r = 4$, a general quartic curve can be given the form

$$a^2 b^2 + c^2 d^2 + e^2 f^2 - 2abcd - 2abef - 2cdef = 0,$$

where a and b, c and d, e and f are three pairs of bitangents of a Steiner set.† I find that this function equals the determinant

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

But the actual determination of the bitangents depends upon the solution of an equation of very high order. Also numerous radicals appear in the expressions for m and the canonical variables in the above cubic form. The method next explained is not only general, but employs no irrationals other than the roots $\lambda_1, \cdots, \lambda_r$.

6. The following method enables us to express the equation $f = 0$ of any plane curve of order r as a determinant of order r whose elements are linear functions of x, y, z. It will suffice to prove this for irreducible forms f. For, if $f = f_1 f_2$, where f_i is of degree r_i and is expressible as a determinant of order r_i of matrix M_i, then f equals the determinant of the matrix

$$\begin{pmatrix} M_1 & O \\ O & M_2 \end{pmatrix},$$

where O is a matrix all of whose elements are zero.

* P. Gordan, these Transactions, vol. 1 (1900), p. 402.

† Miller, Blichfeldt and Dickson, *Finite Groups*, 1916, p. 355.

We suppose merely that f has no repeated factor. Then there exists a straight line which cuts the curve in r distinct points.* Take it as the side $z = 0$ of a triangle of reference. Take as the side $y = 0$ any line not meeting $z = 0$ at one of its r intersections with the curve. Then $(1, 0, 0)$ is not on the curve, and the coefficient of x^3 in f may be assumed to be unity. Hence, for $z = 0$, f reduces to a product $X_1 X_2 \cdots X_r$ of r distinct linear functions $X_i = x + \lambda_i y$. Thus

$$f = X_1 X_2 \cdots X_r + \sum_{k=1}^{r} z^k F_k(y, x), \tag{5}$$

where F_k is a binary form of order $r - k$.

We shall prove that every such form f, in which $\lambda_1, \cdots, \lambda_r$ are distinct, can be expressed as a determinant of the type suggested by § 2:

$$\begin{vmatrix} X_1 + c_{11} z & c_{12} z & \cdots & c_{1r} z \\ \cdot & \cdot & \cdot & \cdot \\ c_{r1} z & c_{r2} z & \cdots & X_r + c_{rr} z \end{vmatrix}. \tag{6}$$

There are $\frac{1}{2}(r + 2)(r + 1)$ coefficients in a general ternary form of order r. In (5) we have identified $f(x, y, 0)$ with $X_1 \cdots X_r$, thus fixing $r + 1$ coefficients; there remain $\frac{1}{2}(r^2 + r)$ coefficients. Hence the identification of (6) with (5) involves as many conditions as there are c_{ij} in and below the main diagonal. Accordingly we shall assign simple values to the remaining c_{ij}:

$$c_{i\,i+1} = 1, \qquad c_{ij} = 0 \quad (j > i + 1;\ i, j = 1, \cdots, r). \tag{7}$$

This choice is in accord with the general theory in § 2, where it was shown that, without altering determinant (6), we may assign the value unity to $r - 1$ non-vanishing c's.

We proceed to prove that the c_{ij} $(j \leqq i)$ can be uniquely determined so that determinant (6), subject to (7), becomes identical with any given form (5). Use is made of the known expansion of an "axial" determinant (6). First, the terms linear in z are

$$\sum_{i=1}^{r} c_{ii}\, z X_1 \cdots X_{i-1} X_{i+1} \cdots X_r .$$

This sum will be identical with $zF_1(y, x)$, where F_1 is any given binary form

* For, if every line $z = rx + sy$ cuts $f = 0$ in points two of which coincide,

$$F(x, y) \equiv f(x, y, rx + sy) = 0$$

has a double root for every r, s, whence

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + r \frac{\partial f}{\partial z} = 0, \qquad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} + s \frac{\partial f}{\partial z} = 0$$

for every r, s. Thus every point on $f = 0$ is a singular (multiple) point.

of order $r - 1$, if they are equal for the r values $x = -\lambda_i y$ $(i = 1, \cdots, r)$ for which the X_i vanish. The resulting conditions

$$c_{ii}(\lambda_1 - \lambda_i)\cdots(\lambda_{i-1} - \lambda_i)(\lambda_{i+1} - \lambda_i)\cdots(\lambda_r - \lambda_i) = F_1(1, -\lambda_i)$$

uniquely determine c_{ii} $(i = 1, \cdots, r)$. Next, the terms of (6) quadratic in z are

$$\sum \begin{vmatrix} c_{ii} & c_{ij} \\ c_{ji} & c_{jj} \end{vmatrix} z^2 \frac{X_1 \cdots X_r}{X_i X_j} \qquad (i, j = 1, \cdots, r;\ i < j).$$

This sum is to be identified with $z^2 F_2(y, x)$. If $j > i + 1$, then $c_{ij} = 0$ and the two-rowed determinant equals the previously determined number $c_{ii} c_{jj}$. If $j = i + 1$, then $c_{ij} = 1$ and the diagonal term is known. Transposing all the known terms and combining them with $z^2 F_2$, we are to identify

$$-\sum_{i=1}^{r-1} c_{i+1,\, i} X_1 \cdots X_{i-1} X_{i+2} \cdots X_r$$

with a known binary form of order $r - 2$. It suffices* for this purpose to take in turn $X_1 = 0$, $X_2 = 0$, $\cdots$, $X_{r-1} = 0$. Of the resulting conditions, the first determines c_{21}, the second determines c_{32} in terms of $c_{21}, \ldots$, the last determines $c_{r,\, r-1}$ in terms of $c_{r-1,\, r-2}$, so that they uniquely determine all of the $c_{i+1,\, i}$.

To make the general step of the proof by induction, we assume that the terms of (6) of degrees $1, 2, \cdots, k - 1$ in z have been identified with the terms of (5) of corresponding degrees by the unique determination of the c_{ii}, $c_{i+1,\, i}$, $\cdots$, $c_{i+k-2,\, i}$. The terms of (6) of degree k in z are

$$\sum \begin{vmatrix} c_{i_1 i_1} & c_{i_1 i_2} & \cdots & c_{i_1 i_k} \\ \cdot & \cdot & \cdot & \cdot \\ c_{i_k i_1} & c_{i_k i_2} & \cdots & c_{i_k i_k} \end{vmatrix} z^k \frac{X_1 \cdots X_r}{X_{i_1} \cdots X_{i_k}} \quad \begin{pmatrix} i_1, \cdots, i_k = 1, \cdots, r; \\ i_1 < i_2 < \cdots < i_k \end{pmatrix}.$$

This sum is to be identified with $z^k F_k(y, x)$. If in the determinant just written every element just above the main diagonal is unity, then

$$i_2 = i_1 + 1, \quad i_3 = i_2 + 1 = i_1 + 2, \quad \cdots, \quad i_k = i_{k-1} + 1 = i_1 + k - 1,$$

and the difference between the subscripts of any c other than $c_{i_k i_1}$ is $\leqq k - 2$, so that the c is one of those previously determined, while the minor of the exceptional c equals unity. We shall prove that the expansions of the remaining determinants D involve only previously determined c's. Then after transposing known terms and combining them with $z^k F_k$, we have left only the product of z^k by

$$\sum_{i=1}^{r-k+1} (-1)^{k-1} c_{i+k-1,\, i} \frac{X_1 \cdots X_r}{X_i X_{i+1} \cdots X_{i+k-1}},$$

* Or we may determine $c_{r,\, r-1}$ from $X_r = 0$ and work back half way or all the way.

which is to be identified with a given binary form of order $r - k$. The conditions obtained by taking in turn $X_1 = 0, \cdots, X_{r-k+1} = 0$ determine these c's in turn. Hence the induction is complete.

It remains to prove the statement regarding any determinant D which has at least one zero element c just above the main diagonal. Then, by (7), zero is the value of every element of D which lies in the rectangle bounded by c and the elements above it and the elements to the right of it. Hence D is the product of two principal minors (each having its diagonal elements on the main diagonal of D). In case either minor has a zero element just above its diagonal, it decomposes similarly into a product of principal minors. Hence D is a product of two or more principal minors each of which has either a single element c_{ii} or is a principal minor P all of whose t elements just above the diagonal are equal to unity. If the diagonal elements of such a $(t + 1)$-rowed P are

$$c_{ii} \quad (i = i_l, i_m, i_n, \cdots, i_w),$$

then, by (7),

$$i_m = i_l + 1, \qquad i_n = i_m + 1 = i_l + 2, \qquad \cdots, \qquad i_w = i_l + t,$$

so that the maximum difference of subscripts in any element of P is $i_w - i_l = t$. Since P has fewer rows than D, $t + 1 < k$, and $t \leqq k - 2$. Hence by the hypothesis for the induction, every element of P is among those previously determined.

THEOREM 3. *Every plane curve can be represented by equating to zero a determinant whose elements are linear functions. In particular, any ternary form without a repeated factor can be transformed linearly into mf, where m is a constant and $f(x, y, 0) = X_1 \cdots X_r$, $X_i = x + \lambda_i y$, $\lambda_1, \cdots, \lambda_r$ being distinct. Then f can be expressed in one and but one way as a determinant*

$$\begin{vmatrix} X_1 + c_{11} z & z & 0 & \cdots & 0 \\ c_{21} z & X_2 + c_{22} z & z & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{r1} z & c_{r2} z & c_{r3} z & \cdots & X_r + c_{rr} z \end{vmatrix},$$

in which the elements just above the main diagonal equal z and the remaining elements above the diagonal equal zero.

7. The problem for quadratic forms may be treated with attention to rationality. A determinant of order 2 or 3, whose elements are linear forms in x, y, z, w with coefficients in a given field F (or domain of rationality) evidently vanishes for a set of values, not all zero, in F. Hence a quadric surface which is representable as such a determinant must have a point with coördinates in F, which may be taken to be $(1, 0, 0, 0)$. If the surface is not a cone, the coefficient of x may be taken as the new variable y. Then by

adding to x a suitable linear function of y, z, w, we obtain $xy + Q(z, w)$, where Q is a binary quadratic form with coefficients in F. A determinant representing it can evidently be given the form

$$\begin{vmatrix} x + A & B \\ C & y \end{vmatrix},$$

where A is free of x, while B and C are free of x and y. Thus $A \equiv 0$, and the condition is that Q have linear factors with coefficients in F, so that the tangent plane $y = 0$ at $(1, 0, 0, 0)$ cuts the surface in rational lines. For a cone or conic, we note that a ternary quadratic form which lacks x^2 can be transformed rationally* into $xy + az^2$ or $Q(y, z)$. But every binary form can be expressed rationally in determinantal form:

$$a_0 x^r + a_1 x^{r-1} y + \cdots + a_r y^r = \begin{vmatrix} a_0 x + a_1 y & y & 0 & 0 & \cdots & 0 \\ -a_2 y & x & y & 0 & \cdots & 0 \\ a_3 y & 0 & x & y & \cdots & 0 \\ -a_4 y & 0 & 0 & x & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (-1)^r a_{r-1} y & 0 & 0 & 0 & \cdots & y \\ (-1)^{r+1} a_r y & 0 & 0 & 0 & \cdots & x \end{vmatrix}.$$

THEOREM 4. *Every binary form can be expressed rationally in determinantal form. A rational quadric surface not a cone can be expressed rationally in determinantal form if and only if it has a rational point the tangent plane at which cuts the surface in rational lines. A conic (or quadric cone) can be represented rationally in determinantal form if and only if it has a rational point (not the vertex).*

If we desire to ignore irrationalities, we take F to be the field of all complex numbers. A summary of our results gives

THEOREM 5. *Every binary form, every ternary form, every quaternary quadratic form, and a sufficiently general quaternary cubic form can be expressed in determinantal form. No further general form has this property.*

I have treated elsewhere† the problem of quaternary cubic forms with attention to rationality. The number 20 of coefficients equals the number of disposable parameters in the determinant (§ 2), and the problem depends in general upon the solution of a single algebraic equation. The notations suggested by § 2 are less convenient for this problem than those used in the paper cited.

8. We shall examine briefly the conditions under which a given rational ternary cubic form T is expressible rationally in determinant form. If T

* Dickson, *Algebraic Invariants*, 1914, p. 24.

† American Journal of Mathematics, April, 1921.

vanishes at a rational point, the problem has been treated fully in the paper last cited. In the contrary case, we may assume that the coefficient of x^3 is unity and that the terms in $x^2 y$ and $x^2 z$ are lacking. The matrix of the determinant may be taken to be $x_1 I + x_2 X_2 + zX_3$, where $x_1 = x$, $x_2 = y$. Since T is not zero at a rational point, $|x_1 I + x_2 X_2| \neq 0$ when x_1 and x_2 are rational and not both zero. As shown in § 2, the characteristic equation of X_2 therefore has no rational root, and hence has a single invariant factor $\lambda^3 - \alpha\lambda - \beta$, so that there exists (§ 9) a matrix of rational coefficients which transforms X_2 into P_2 and X_3 into* P_3:

$$P_2 = \begin{pmatrix} 0 & \alpha & \beta \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad P_3 = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}.$$

The determinant of $xI + yP_2 + zP_3$ will be identical with

$$T = x^3 - \alpha xy^2 + \beta y^3 + Bxyz + Cy^2 z + Dxz^2 + Eyz^2 + Fz^3$$

if and only if

$$a + e + k = 0, \qquad b + \alpha d + \beta g + f = -B, \qquad c - \alpha k + \beta d + \beta h = C,$$

$$\begin{vmatrix} a & b \\ d & e \end{vmatrix} + \begin{vmatrix} a & c \\ g & k \end{vmatrix} + \begin{vmatrix} e & f \\ h & k \end{vmatrix} = D,$$

$$-\begin{vmatrix} b & c \\ h & k \end{vmatrix} - \alpha \begin{vmatrix} d & f \\ g & k \end{vmatrix} - \begin{vmatrix} a & c \\ d & f \end{vmatrix} + \beta \begin{vmatrix} d & e \\ g & h \end{vmatrix} = E, \qquad |P_3| = F.$$

Employ the linear equations to express a, b, c in terms of the remaining letters. In D, g and f enter linearly, the latter with the coefficient $d - h$. If $h = d$, the result of substituting the values of g and f given by our equations (D) and (E) into $|P_3| = F$ is of the ninth degree in each d and e and of the eighth degree in k. Since this special case thus involves hopeless difficulties, we take $h = d + m$, $m \neq 0$. Then equations (D), (E) and (F) give

$$f = \frac{g\gamma - \delta}{m}, \qquad \gamma \equiv m\beta - \alpha k + 3\beta d - C,$$

$$\delta \equiv D + e^2 + ek + k^2 - Bd - \alpha d^2,$$

$$\alpha\gamma g^2 + \{\gamma(e + 2k) + m\beta(k - e) - \alpha\delta\}g - m\epsilon - \delta(e + 2k) = 0,$$

$$\epsilon \equiv E - Bk - C(m + 2d) - 2\alpha dk - \alpha mk + 3\beta d^2 + 3\beta md + \beta m^2,$$

* For the normalization of P_3 we have available the most general matrix

$$rI + sP_2 + t\begin{pmatrix} 0 & \beta & 0 \\ 0 & 0 & \beta \\ 1 & 0 & -\alpha \end{pmatrix}$$

commutative with P_2. But it transforms P_3 into such a complicated matrix that normalization does not seem worth while.

$$m^2 F = \begin{vmatrix} -e-k & m(-B-\alpha d-\beta g)-g\gamma+\delta & C-m\beta-2\beta d+\alpha k \\ md & m^2 e & g\gamma-\delta \\ g & m^2+md & k \end{vmatrix}.$$

The resultant of our quadratic and cubic equations in g is readily seen to be of the sixth degree in m, with coefficients involving d, e, k to higher degree. Even in the simple case in which T has a rational point, the problem depends upon a cubic irrationality.

9. Two pairs of bilinear forms with matrices M, N and M', N' with elements in a field F and such that N and N' are non-singular (determinant $\neq 0$) are called equivalent with respect to F if there exist non-singular matrices P and Q with elements in F such that $PMQ = M', PNQ = N'$. Necessary and sufficient conditions for equivalence are known to be the identity of the invariant factors of $M - \lambda N$ with those of $M' - \lambda N'$. The special case $N = N' = I$ shows that matrix (or substitution) M can be transformed within F into $Q^{-1}MQ = M'$ if and only if the invariant factors of the characteristic determinants $|M - \lambda I|$ and $|M' - \lambda I|$ of M and M' are the same. This case finds application in many branches of mathematics.

The last result has been proved independently* of the theory of pairs of bilinear forms. From it we can deduce the latter theory. First, let M, N be equivalent to M', N', so that $MN^{-1} = J, I$ are equivalent to $M'N'^{-1} = J', I$, whence $AJB = J'$, $AIB = I$, $A = B^{-1}$. Thus B transforms J into J' and their characteristic determinants $|MN^{-1} - \lambda I|$ and $|M'N'^{-1} - \lambda I|$ have the same invariant factors. The latter are not altered when the determinants are multiplied by the constants $|N|$ and $|N'|$, respectively. Hence $|M - \lambda N|$ and $|M' - \lambda N'|$ have the same invariant factors. This necessary condition is also a sufficient condition for equivalence of the pairs. For, we then have $B^{-1}JB = J'$, $J \equiv MN^{-1}$, $J' \equiv M'N'^{-1}$, where the elements of B are in F. Then, if λ is arbitrary,

$$B^{-1}(MN^{-1} - \lambda I)B = M'N'^{-1} - \lambda I,$$

$$B^{-1}(M - \lambda N)N^{-1}B = (M' - \lambda N')N'^{-1},$$

$$B^{-1}(M - \lambda N)C = M' - \lambda N', \qquad C = N^{-1}BN'.$$

Thus B^{-1} serves as pre-factor and C as post-factor to convert M into M' and N into N', so that the pairs are equivalent.

The ideas in this paper evidently apply also to the question of the equivalence of n-tuples of bilinear forms, and are being developed by one of my students.

10. A generalization of Theorem 1 is given by

THEOREM 6. *Any r-rowed determinant D whose elements are linear homo-*

* Dickson, these Transactions, vol. 3 (1902), pp. 290–2.

geneous functions of $x_1, \cdots, x_n$ *can be expressed in a canonical form involving not more than* $(n-2)r^2+2$ *parameters when* $n>2$.

The matrix of D is of the form $M = x_1 M_1 + \cdots + x_n M_n$.

Since we may assume that D is not identically zero, there exist constants $c_1, \cdots, c_n$ such that the determinant of $C = c_1 M_1 + \cdots + c_n M_n$ is not zero. Replacing $x_1, \cdots, x_n$ by

$$c_1 x_1 + d_1 x_2 + \cdots + k_1 x_n, \; \cdots, \; c_n x_1 + d_n x_2 + \cdots + k_n x_n,$$

where the $d_i, \cdots, k_i$ are chosen so that the determinant of these linear forms is not zero, we see that M is replaced by $x_1 C + \cdots$. Returning to the initial notations, we may therefore assume that $|M_1| \neq 0$.

Proceeding as in § 2, we reduce M to an equivalent matrix

$$P = x_1 I + x_2 P_2 + \cdots + x_n P_n,$$

in which P_2 is the canonical form of a matrix (or linear substitution) obtained by transformation. We proved our theorem in § 2 for the case in which P_2 is of the special form (2) with $\lambda_1, \cdots, \lambda_r$ distinct. In case equalities occur among these λ's, transformations (3) are included among those commutative with P_2, so that we can accomplish the same (and further) specialization of the parameters in $P_3, \cdots, P_n$ as in § 2. Hence for two reasons the number of parameters in $P_2, \cdots, P_n$ is smaller than before.

Finally, let the canonical P_2 be of its most general form:*

$$(8) \quad \eta_1' = \lambda_1 \eta_1 + \eta_2, \quad \eta_2' = \lambda_1 \eta_2 + \eta_3, \quad \cdots, \quad \eta_{e_1-1}' = \lambda_1 \eta_{e_1-1} + \eta_{e_1}, \quad \eta_{e_1}' = \lambda_1 \eta_{e_1};$$

$$(9) \quad \zeta_1' = \lambda_2 \zeta_1 + \zeta_2, \quad \zeta_2' = \lambda_2 \zeta_2 + \zeta_3, \quad \cdots, \quad \zeta_{e_2-1}' = \lambda_2 \zeta_{e_2-1} + \zeta_{e_2}, \quad \zeta_{e_2}' = \lambda_2 \zeta_{e_2};$$

etc. Now (8) is commutative with†

$$(10) \qquad \eta_i' = k_1 \eta_i + k_2 \eta_{i+1} + \cdots + k_{e_1-i+1} \eta_{e_1} \qquad (i = 1, \cdots, e_1).$$

To verify this fact, we may take $\lambda_1 = 0$ in (8), since we may subtract $\lambda_1 I$ from the matrix of (8); then the product of the modified (8) and (10), in either order, replaces η_i by

$$k_1 \eta_{i+1} + k_2 \eta_{i+2} + \cdots + k_{e_1-i} \eta_{e_1}.$$

Similarly, (9) is commutative with

$$(11) \qquad \zeta_i' = l_1 \zeta_i + l_2 \zeta_{i+1} + \cdots + l_{e_2-i+1} \zeta_{e_2} \qquad (i = 1, \cdots, e_2).$$

* Dickson, *Linear Groups*, p. 223; Bôcher, *Higher Algebra*, p. 293.

† We do not need the fact that every linear transformation on $\eta_1, \cdots, \eta_{e_1}$ which is commutative with (8) is of the form (10). If the λ's are all distinct, every transformation commutative with P_2 is given by (10), (11), etc.

Transformations (10), (11), etc. are available for the normalization of $P_j\ (j > 2)$. Let P_j replace η_{e_1} by

$$t_1 \eta_1 + \cdots + t_{e_1} \eta_{e_1} + f, \tag{12}$$

where f is a linear function of the ζ's, etc. If K is the matrix of the transformation defined by (10), (11), etc., $K^{-1} P_j K$ replaces η_{e_1} by the product of k_1^{-1} by the function by which (10), (11), etc., replace (12), the product being

$$\frac{1}{k_1}\left\{f' + \sum_{i=1}^{e_1} t_i (k_1 \eta_i + k_2 \eta_{i+1} + \cdots + k_{e_1-i+1} \eta_{e_1})\right\}$$

$$= t_1 \eta_1 + \left(t_2 + \frac{k_2}{k_1} t_1\right)\eta_2 + \left(t_3 + \frac{k_2}{k_1} t_2 + \frac{k_3}{k_1} t_1\right)\eta_3 + \cdots$$

$$+ \left(t_{e_1} + \frac{k_2}{k_1} t_{e_1-1} + \cdots + \frac{k_{e_1}}{k_1} t_1\right)\eta_{e_1} + \frac{f'}{k_1},$$

where f' is the function by which (11), etc. replace f. Hence the ratios

$$k_2/k_1, \ \cdots, \ k_{e_1}/k_1, \ l_1/k_1, \ l_2/l_1, \ \cdots, \ l_{e_2}/l_1, \ \cdots$$

are available to specialize as many parameters in P_j. Their number is $e_1 + e_2 + \cdots + e_t - 1$ if $e_1 + \cdots + e_t = r$. In § 2 we could specialize only this number $r - 1$ of parameters in the $P_j\ (j > 2)$. We now have only t distinct λ's instead of the former r. Hence the total number of parameters found in § 2 is the true maximum.

The University of Chicago,
Chicago, Ill.