BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 10, Number 1, January 1984

# THE INVARIANT THEORY OF BINARY FORMS

BY JOSEPH P. S. KUNG[1] AND GIAN-CARLO ROTA[2]

*Dedicated to Mark Kac on his seventieth birthday*

## Table of Contents

**1. Introduction.** Like the Arabian phoenix rising out of its ashes, the theory of invariants, pronounced dead at the turn of the century, is once again at the forefront of mathematics. During its long eclipse, the language of modern algebra was developed, a sharp tool now at long last being applied to the very

Received by the editors May 26, 1983.

1980 *Mathematics Subject Classification.* Primary 05-02, 13-02, 14-02, 15-02.

[1] Supported by a North Texas State University Faculty Research Grant.

[2] Supported by NSF Grant MCS 8104855.

purpose for which it was intended. More recently, the artillery of combinatorics began to be aimed at the problems of invariant theory bequeathed to us by the nineteenth century, abandoned at the time because of insufficient combinatorial thrust.

Even the seemingly polished and well-developed chapters of classical invariant theory, such as the theory of binary forms, now reveal themselves on closer inspection to be sadly deficient in content and proof. The grey area between the known and the unknown, between the solidly established result and the likely conjecture, casts a shade of uncertainty behind which the open problems still hide themselves from the reader born several generations later.

Thus, further progress on the fascinating trip that is invariant theory, a great idea foreshadowed in the work of Boole and gradually formulated by Cayley, Sylvester, Clebsch, Gordan, Capelli, MacMahon, Hilbert, Young, Study, and others, must begin with a presentation, complete with proofs and up-to-date algorithms, of those results that lie within the range of present day mathematical methods. The purpose of the present work is to give such a presentation of the central chapter of classical invariant theory.

The basic results of the theory of invariants of binary forms are developed here by constructive methods. Our objective is to enable the reader to eventually appreciate the computations of the nineteenth century invariant theorists, or at times, to make such computations superfluous.

In the exposition, we are guided by the following criteria. Our language and notation are, wherever possible, patterned after nineteenth century usage. It might have been easier to adopt instead one of the many—too many, perhaps—equivalent languages that have been taking turns in the annual Paris display of mathematical fashion. One could, for example, rephrase the results in the language of representations of GL(2) over certain tensor spaces, or as the study of moduli parameterizing certain algebraic varieties. However, in the attempt to reach as wide an audience as possible, we chose to describe and make rigorous the original notation and follow it as closely as possible. Thus, what is perhaps the main novelty of the present work is a rigorous and yet manageable account of the umbral or symbolic calculus, what Hermann Weyl called "the great war-horse of nineteenth century invariant theory". To be sure, rigor can be injected into the umbral method by simply citing one of the bromides of multilinear algebra, such as "every tensor is the sum of decomposable tensors" (as did Weyl, for example). What is not as easily accomplished is to combine rigor with the suppleness of the bracket notation, so that the computations of covariants and their syzygies can measure up to the artisanship of the past century. Our rigorization of the umbral method by operators and linear functionals obeying a crucial multiplicative property was suggested by earlier work by one of us, and the idea can be traced back to his 1964 paper on enumerating the partitions of a set.

A similar salvage operation could not, unfortunately, be carried out on the proofs. Most of the proofs presented here are new. Among the techniques, the major innovation is an explicit algorithm (Algorithm 4.1) for expressing in terms of the roots a covariant expressed in umbral notation. This algorithm leads to an alternative representation of covariants as symmetrizations of difference in the roots. The formal similarity between this representation and the umbral representation yields a one-line proof of Hermite's reciprocity law.

The discussion of apolarity is also simplified by this algorithm which enables us to immediately infer the significance in terms of the roots of the vanishing of the basic covariants in the theory of apolarity. As typical applications, we give a new proof of Sylvester's theorem (including exceptional cases) and complete lists of canonical forms for the cubic and the quintic. Our task is made easier by the introduction of homogenized roots, which allows the exceptional cases to be handled without undue commotion. We have also computed the umbral representation for several covariants of apolarity theory using a device for converting into umbral form covariants given in the form of determinants (Lemmas 5.3 and 5.4). This device was not exploited in the nineteenth century because of lack of rigor in the umbral method; if nothing else, its simplicity justifies our proposed formalization of the umbral method.

We give two proofs of the finiteness theorem. Both use the representation of covariants as symmetrizations of differences in the roots. Neither, however, uses Hilbert's basis theorem. The first is based on the circular straightening algorithm. The idea of this proof goes back to Kempe, although considerable retouching is necessary to apply it to covariants. This proof yields an explicit construction of a generating set of covariants and the method is made clear—we hope—by our discussion of the cubic. The second proof is a little known proof due to Hilbert. This uses a lemma due to Gordan which was the combinatorial mainstay of nineteenth century invariant theory. Gordan's lemma is given a short proof here using a combinatorial property of partially ordered sets.

It may not be amiss to recount the history of the new proofs given here of the First and Second Fundamental Theorems. Alfred Young invented, in 1900, his celebrated tableaux—to be followed by standard tableaux in 1928—as a method for computing covariants of forms. The main difficulty of the umbral method is that covariants which look quite different umbrally may reveal themselves to be identical upon permutations and substitutions of equivalent letters and applications of the syzygy. This is, in fact, the gist of the Second Fundamental Theorem (which is given a new formulation (Lemma 3.4) using the notion of symmetrization of letters). Young saw that this difficulty could be obviated in part by a decomposition of the group of permutations of equivalent letters into what are nowadays called irreducible representations. However, it did not dawn on him—nor to anyone after him—that standard tableaux would be the ideal method of proving the First Fundamental Theorem, thereby getting rid once and for all of the ponderous Cayley omega operator or the nonconstructive device of the Reynolds operator. Our proof of the First Fundamental Theorem, besides being constructive, gives a simple algebraic approach to the averaging procedures that must be used at some point in every proof of the finiteness theorem, as Hurwitz was the first to perceive.

Although our presentation is limited to binary forms over a field (not assumed to be algebraically closed, except in the section on apolarity) of characteristic zero, we have taken pains in selecting those proofs which, whenever possible, work over a field of arbitrary characteristic, or which point clearly to the step where such an extension fails. The extension of the umbral method to arbitrary characteristic must remain in the realm of speculation—not surprisingly, since the right concept of covariant in positive characteristic is yet to come.

Similarly, the proofs have been chosen to generalize, whenever possible, to forms in several variables, commutative, anticommutative, or not, corresponding to symmetric tensors, antisymmetric tensors, and tensors. Umbral notation and the proofs of the two fundamental theorems carry over without change. For general tensors, the language of double tableaux—what we have called elsewhere the letter place algebra—must be used. What fails in more than two variables is the expression of invariants in terms of the roots, and hence the present proofs of the finiteness theorem do not generalize, as far as one can see. In fact, the notion of a covariant ramifies in several variables into several kinds of concomitants, and the various kinds of apolarity never seem to have been fully explored.

In closing, we remark that, to this day, the covariants of no nontrivial form (except for conics) in three or more variables have been fully classified, not even those of the ternary quartic which persuaded Emmy Noether to quit invariant theory. We surmise that only a deeper combinatorial understanding of the umbral method will lead to a complete list of covariants. This work has been written with the purpose of stimulating such understanding.

## 2. Umbral notation.

2.1 *Binary forms and their covariants.* The theory of invariants of binary forms is concerned with properties of homogeneous polynomials in two variables which are independent of the choice of coordinates.

More specifically, we shall deal throughout this paper with binary forms. A *binary form $f(x, y)$ of degree $n$ in the variables $x$ and $y$* is a homogeneous polynomial of degree $n$ in $x$ and $y$. Thus,

$$\begin{aligned} f(x, y) &= \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k} \\ &= a_n x^n + \binom{n}{1} a_{n-1} x^{n-1} y + \cdots + \binom{n}{n-1} a_1 x y^{n-1} + a_0 y^n. \end{aligned}$$

The numbers $a_k$ are called the *coefficients* of $f(x, y)$ and are assumed to belong to a field of characteristic zero. A *linear change of variables* $(c_{ij})$ is a transformation of the variables $x$ and $y$ given by

$$x = c_{11}\bar{x} + c_{12}\bar{y}, \qquad y = c_{21}\bar{x} + c_{22}\bar{y} \tag{2.1}$$

such that the determinant of the entries, $c_{11}c_{22} - c_{12}c_{21}$, is nonzero. Under a linear change of variables (2.1), the binary form $f(x, y)$ is transformed into another binary form $\bar{f}(\bar{x}, \bar{y})$ in the new variables $\bar{x}$ and $\bar{y}$ defined by

$$\bar{f}(\bar{x}, \bar{y}) = \sum_{k=0}^{n} \binom{n}{k} a_k (c_{11}\bar{x} + c_{12}\bar{y})^k (c_{21}\bar{x} + c_{22}\bar{y})^{n-k}. \tag{2.2}$$

After expanding and regrouping terms, we obtain a binary form

$$\bar{f}(\bar{x}, \bar{y}) = \sum_{k=0}^{n} \binom{n}{k} \bar{a}_k \bar{x}^k \bar{y}^{n-k}$$

in the variables $\bar{x}$ and $\bar{y}$ whose coefficients $\bar{a}_k$ are polynomials in $a_i$ and $c_{ij}$. The precise expressions of these polynomials need not concern us for the moment and will be derived shortly.

Let $g$ be a nonnegative integer. A nonconstant polynomial $I(A_0, A_1, \ldots, A_n, X, Y)$ in the variables $A_0, A_1, \ldots, A_n$, $X$ and $Y$ is said to be a *covariant of index* $g$ of binary forms of degree $n$ if for all binary forms $f(x, y)$ of degree $n$ and all linear changes of variables, the following identity holds:

$$I(\bar{a}_0, \bar{a}_1, \ldots, \bar{a}_n, \bar{x}, \bar{y}) = (c_{11}c_{22} - c_{21}c_{12})^g I(a_0, a_1, \ldots, a_n, x, y).$$

A covariant in which the variables $X$ and $Y$ do not occur is said to be an *invariant.*

Our objective is to determine as explicitly as possible all the covariants of binary forms.

For many purposes it is indispensable to consider several binary forms simultaneously. A nonconstant polynomial

$$I(A_{10}, A_{11}, \ldots, A_{1n(1)}, A_{20}, \ldots, A_{2n(2)}, \ldots, A_{r0}, \ldots, A_{rn(r)}, X, Y)$$

in the variables $A_{ij}$, $X$ and $Y$ is said to be a *joint covariant of index* $g$ of $r$-tuples of binary forms $f_i(x, y)$ of degree $n(i)$ if for all linear changes of variables $(c_{ij})$ and for all $r$-tuple of binary forms

$$f_i(x, y) = \sum_{k=0}^{n(i)} \binom{n(i)}{k} a_{ik} x^k y^{n(i)-k}, \qquad i = 1, 2, \ldots, r,$$

the following identity holds:

$$\begin{aligned} &I(\bar{a}_{10}, \bar{a}_{11}, \ldots, \bar{a}_{1n(1)}, \ldots, \bar{a}_{r0}, \bar{a}_{r1}, \ldots, \bar{a}_{rn(r)}, \bar{x}, \bar{y}) \\ &\quad = (c_{11}c_{22} - c_{12}c_{21})^g I(a_{10}, a_{11}, \ldots, a_{1n(1)}, \ldots, a_{r0}, a_{r1}, \ldots, a_{rn(r)}, x, y). \end{aligned}$$

A *joint invariant* of $f_1(x, y), \ldots, f_r(x, y)$ is a joint covariant in which the variables $X$ and $Y$ do not occur.

A covariant $I(A_0, A_1, \ldots, A_n, X, Y)$ is said to be *homogeneous* if it is homogeneous both as a polynomial in the variables $A_0, A_1, \ldots, A_n$ and as a polynomial in the variables $X$ and $Y$. Every covariant can be written as a linear combination of homogeneous covariants. If $I$ is a homogeneous covariant, the *degree* of $I$ is the (total) degree of $I$ as a polynomial in $A_0, A_1, \ldots, A_n$, while the *order* of $I$ is the (total) degree of $I$ as a polynomial in $X$ and $Y$.

2.2 *The umbral calculus.* The simplest binary form is an $n$th power of a linear form, namely, one of the form

$$f(x, y) = (\alpha_1 x + \alpha_2 y)^n.$$

The device we are about to describe permits us to reduce computations with binary forms to this special case.

Let $\mathcal{A} = \{\alpha, \beta, \ldots, \omega, u\}$ be an alphabet consisting of an infinite supply of Greek letters followed by the single Roman letter $u$. The letters in $\mathcal{A}$ are called *umbral letters*. To each Greek letter $\alpha$ and the Roman letter $u$, we associate two variables $\alpha_1$ and $\alpha_2$. Thus, we have the variables $\alpha_1$, $\alpha_2$, $\beta_1$, $\beta_2, \ldots, \omega_1$, $\omega_2$, $u_1$, $u_2$. The ring of all polynomials in these variables is an (infinite-dimensional) vector space called the *umbral space* $\mathcal{U}$.

The *umbral operator* $U$ for the space of binary forms of degree $n$ is the linear operator defined from the umbral space $\mathcal{U}$ to the space $\mathcal{P}$ of polynomials in the variables $A_0$, $A_1, \ldots, A_n$, $X$ and $Y$ defined in the following way. We denote the action of the operator $U$ on a polynomial $P(\alpha_1, \alpha_2, \ldots)$ in $\mathcal{U}$ by $\langle U | P(\alpha_1, \alpha_2, \ldots)\rangle$ and set:

$\langle U | \alpha_1^k \alpha_2^{n-k}\rangle = A_k$ for any Greek umbral letter $\alpha$;

$\langle U | \alpha_1^j \alpha_2^k\rangle = 0$ if $j + k \neq n$ and $\alpha$ is any Greek umbral letter;

$\langle U | u_1^k\rangle = (-Y)^k$;

$\langle U | u_2^k\rangle = X^k$;

$\langle U | \alpha_1^i \alpha_2^j \beta_1^k \beta_2^l \cdots u_1^p u_2^q\rangle = \langle U | \alpha_1^i \alpha_2^j\rangle\langle U | \beta_1^k \beta_2^l\rangle \cdots \langle U | u_1^p\rangle\langle U | u_2^q\rangle$.

The final rule is called the *multiplicative rule*. These rules uniquely define, by linearity, the umbral operator $U$ on the umbral space $\mathcal{U}$. If

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k}$$

is a binary form of degee $n$, we define the *umbral linear functional $U(f)$ associated with $f(x, y)$* to be the linear functional on $\mathcal{U}$ obtained by evaluating the umbral operator at

$$A_0 = a_0, \quad A_1 = a_1, \ldots, \quad A_n = a_n, \quad X = x, \quad Y = y.$$

Every polynomial $I(A_0, A_1, \ldots, A_n, X, Y)$ can be written as $\langle U | Q(\alpha_1, \alpha_2, \ldots)\rangle$ for some polynomial $Q(\alpha_1, \alpha_2, \ldots)$ in the umbral space; the polynomial $Q$ is called *an umbral representation* for the polynomial $I$, and $I$ is called *the umbral evaluation* of $Q$. To see this, it suffices to consider the case when $I$ is a monomial. But if $I$ is the monomial $A_0^{d_0} A_1^{d_1} \cdots A_n^{d_n} X^{e_1} Y^{e_2}$, then a simple computation shows that $I$ equals

$$\begin{array}{ccc} \Big\langle U | \alpha_1^0 \alpha_2^n \cdots \gamma_1^0 \gamma_2^n & \delta_1^1 \delta_2^{n-1} \cdots \varepsilon_1^1 \varepsilon_2^{n-1} & \cdots (-u_1^{e_2}) u_2^{e_2} \Big\rangle, \\ \leftarrow d_0 \text{ times} \rightarrow & \leftarrow \quad d_1 \text{ times} \quad \rightarrow & \ldots \end{array}$$

where the umbral letters $\alpha, \ldots, \gamma$, $\delta, \ldots, \varepsilon, \ldots$ are distinct. In general, the umbral representation of a polynomial $I$ is far from unique.

The umbral notation is easily extended to several binary forms. Briefly, the umbral operator $U$ for the space of $r$-tuples of binary forms $f_1(x, y), \ldots, f_r(x, y)$ of degree $n(1), \ldots, n(r)$ is defined as follows. Partition the set of Greek umbral letters into $r$ mutually disjoint infinite subsets $\mathcal{A}_i$ and assign the Greek letters in the $i$th subset $\mathcal{A}_i$ to the $i$th form $f_i(x, y)$. If two letters are assigned to the same form $f_i(x, y)$, they are said to be *equivalent*. The umbral operator $U$ is the linear operator defined from the umbral space $\mathcal{U}$ to the space of polynomials in the variables $A_{10}$, $A_{11}, \ldots, A_{1n(1)}, \ldots, A_{r0}$, $A_{r1}, \ldots, A_{rn(r)}$, $X$ and $Y$ by the rules:

$\langle U | \alpha_1^k \alpha_2^{n(i)-k}\rangle = A_{ik}$ if $\alpha$ is in $\mathcal{A}_i$;

$\langle U | \alpha_1^j \alpha_2^k\rangle = 0$ if $\alpha$ is in $\mathcal{A}_i$ and $j + k \neq n(i)$;

$\langle U|u_1^k\rangle = (-Y)^k;$
$\langle U|u_2^k\rangle = X^k,$
and the multiplicative rule, extended here to hold even if the umbral letters are assigned to different forms. Whenever possible, our discussion will be carried out for a single binary form, the extension to several binary forms being mostly a matter of notation.

2.3 *Changes of variables.* We shall now introduce a notation for changes of variables which, combined with the umbral notation, leads to the classification of covariants of binary forms in terms of their umbral representations.

Let $(c_{ij})$ be a (linear) change of variables. We set

$$\begin{aligned} c_2 &= \phantom{-}c_{11}, & d_2 &= \phantom{-}c_{12}, \\ c_1 &= -c_{21}, & d_1 &= -c_{22}. \end{aligned}$$

We shall denote a change of variables written in this way by $(c, d)$. If $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are two vectors (of dimension two), we define the *bracket* $[u\quad v]$ by

$$[u\quad v] = u_1v_2 - u_2v_1 = \det\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}.$$

The bracket notation gives a simple way of computing the umbral representation of any polynomial $I(\bar{a}_0,\ldots,\bar{a}_n, \bar{x}, \bar{y})$ in terms of the umbral representation of $I(a_0,\ldots,a_n, x, y)$ and various brackets involving the vectors $c = (c_1, c_2)$ and $d = (d_1, d_2)$. Specifically, we have

PROPOSITION 2.1. *Let*

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k}$$

*be a binary form of degree n and let $\bar{f}(\bar{x}, \bar{y})$ be the binary form obtained from $f(x, y)$ by the change of variables $(c, d)$. Let I be a polynomial in $\mathcal{P}$ and let*

$$I(a_0, a_1,\ldots,a_n, x, y) = \langle U(f)|P(\alpha_1, \alpha_2, \beta_1, \beta_2,\ldots,u_1, u_2)\rangle$$

*be an umbral representation of I. Then*

$$\begin{aligned} I(\bar{a}_0, \bar{a}_1,\ldots,\bar{a}_n, \bar{x}, \bar{y}) &= \langle U(\bar{f})|P(\alpha_1, \alpha_2, \beta_1, \beta_2,\ldots,u_1, u_2)\rangle \\ &= \langle U(f)|P([\alpha\quad c],[\alpha\quad d],[\beta\quad c],[\beta\quad d], \\ &\qquad \ldots,[u\quad c]/[c\quad d],[u\quad d]/[c\quad d])\rangle. \end{aligned}$$

PROOF. As umbral operators are linear and obey the multiplicative rule, the proof is reduced to verifying the following identities:

A. For any Greek umbral letter $\alpha$,

$$\langle U(f)|[\alpha\quad c]^j[\alpha\quad d]^k\rangle = \langle U(\bar{f})|\alpha_1^j\alpha_2^k\rangle;$$

B. $$\langle U(f)|[u\quad c]/[c\quad d]\rangle = \langle U(\bar{f})|u_1\rangle = -\bar{y};$$

C. $$\langle U(f)|[u\quad d]/[c\quad d]\rangle = \langle U(\bar{f})|u_2\rangle = \bar{x}.$$

We shall prove these identities in reverse order.

Inverting the matrix of the change of variables $(c, d)$, we obtain

$$\bar{x} = (-d_1 x - d_2 y)/[c \quad d], \qquad \bar{y} = (c_1 x + c_2 y)/[c \quad d].$$

Umbrally, this is

$$\langle U(f) \,|\, [u \quad d]/[c \quad d] \rangle = \bar{x}, \qquad \langle U(f) \,|\, [u \quad c]/[c \quad d] \rangle = -\bar{y}.$$

This verifies identities B and C.

To verify the first identity, let $\alpha$ be any Greek umbral letter. Using the umbral representation

$$f(x, y) = \left\langle U(f) \,|\, (\alpha_1 u_2 - \alpha_2 u_1)^n \right\rangle = \left\langle U(f) \,|\, [\alpha \quad u]^n \right\rangle$$

and the fact that, by definition,

$$\bar{f}(\bar{x}, \bar{y}) = f(x, y),$$

we have

$$\bar{f}(\bar{x}, \bar{y}) = \left\langle U(f) \,|\, [\alpha \quad u]^n \right\rangle.$$

By the determinantal identity

$$[\alpha \quad u][c \quad u] = \det\begin{pmatrix} [\alpha \quad c] & [u \quad c] \\ [\alpha \quad d] & [u \quad d] \end{pmatrix},$$

we obtain

$$\begin{aligned} \bar{f}(\bar{x}, \bar{y}) &= \left\langle U(f) \,|\, \det\begin{pmatrix} [\alpha \quad c] & [u \quad c]/[c \quad d] \\ [\alpha \quad d] & [u \quad d]/[c \quad d] \end{pmatrix}^n \right\rangle \\ &= \left\langle U(f) \,|\, \sum_{k=0}^{n} \binom{n}{k} [\alpha \quad c]^k [\alpha \quad d]^{n-k} \left( -\frac{[u \quad c]}{[c \quad d]} \right)^{n-k} \left( \frac{[u \quad d]}{[c \quad d]} \right)^k \right\rangle \\ &= \sum_{k=0}^{n} \binom{n}{k} \left\langle U(f) \,|\, [\alpha \quad c]^k [\alpha \quad d]^{n-k} \right\rangle \bar{x}^k \bar{y}^{n-k}. \end{aligned}$$

Equating like powers of $\bar{x}$ and $\bar{y}$, we conclude that

$$\left\langle U(f) \,|\, [\alpha \quad c]^k [\alpha \quad d]^{n-k} \right\rangle = \bar{a}_k = \left\langle U(\bar{f}) \,|\, \alpha_1^k \alpha_2^{n-k} \right\rangle.$$

Finally, for $j + k \neq n$,

$$\left\langle U(f) \,|\, [\alpha \quad c]^j [\alpha \quad d]^k \right\rangle = 0$$

since any monomial in the expansion of $[\alpha \quad c]^j [\alpha \quad d]^k$ is of the form $\alpha_1^p \alpha_2^q$, where $p + q = j + k$ and $p + q \neq n$. Hence, in these cases,

$$\left\langle U(f) \,|\, [\alpha \quad c]^j [\alpha \quad d]^k \right\rangle = \left\langle U(\bar{f}) \,|\, \alpha_1^j \alpha_2^k \right\rangle = 0.$$

This completes the proof. □

EXAMPLE. From Proposition (2.1) we obtain the explicit expression of the coefficient $\bar{a}_k$ in terms of $a_k$ and the entries $c_{11}, c_{12}, c_{21}, c_{22}$ of the change of variables matrix $(c_{ij})$:

$$\begin{aligned} \bar{a}_k &= \left\langle U(f) \,|\, (\alpha_1 c_{11} + \alpha_2 c_{21})^k (\alpha_1 c_{12} + \alpha_2 c_{22})^{n-k} \right\rangle \\ &= \sum_{m=0}^{n} \left( \sum_{i=m-n+k}^{\min(m,k)} \binom{k}{i} \binom{n-k}{m-i} c_{11}^i c_{12}^{m-i} c_{21}^{k-i} c_{22}^{n-k-m+i} \right) a_m. \end{aligned}$$

No further use will be made of this formula.

## 3. The fundamental theorems.

3.1 *Bracket polynomials.* The formula given in Proposition 2.1 for computing the effect of a change of variables on the umbral representation suggests a subspace of the umbral space whose umbral evaluations are obviously covariants. This is the subspace $\mathcal{B}$ of bracket polynomials.

Recall that a bracket $[\alpha \quad \beta]$ or $[\alpha \quad u]$ is defined by

$$[\alpha \quad \beta] = \alpha_1\beta_2 - \alpha_2\beta_1, \qquad [\alpha \quad u] = \alpha_1 u_2 - \alpha_2 u_1.$$

A *bracket monomial* $M$ in the umbral space $\mathcal{U}$ is a nonconstant polynomial in $\mathcal{U}$ which can be written as a product of brackets: that is,

$$M = [\alpha \quad \beta][\alpha \quad \delta] \cdots [\omega \quad u]$$

for brackets $[\alpha \quad \beta]$, $[\alpha \quad \delta], \dots, [\omega \quad u]$. In particular, a bracket monomial is never a monomial in the variables $\alpha_1$, $\alpha_2, \dots, u_1$, $u_2$. The *index* of a bracket monomial $M$ is the number of brackets in $M$ containing *only* Greek letters, its *order* is the number of brackets containing the Roman letter $u$, and its *height* is the total number of brackets in $M$.

A *bracket polynomial* is a linear combination of bracket monomials. The bracket polynomials form a subspace $\mathcal{B}$ of the umbral space $\mathcal{U}$. The bracket monomials of index $g$ span a subspace $\mathcal{B}_g$ of the space of bracket polynomials. The bracket polynomials in $\mathcal{B}_g$, which are linear combinations of bracket monomials all of the same index $g$, are called *bracket polynomials of index g.*

THEOREM 3.1 (THE FIRST FUNDAMENTAL THEOREM): PART I. *The umbral evaluation $\langle U \mid P \rangle$ of a bracket polynomial $P$ of index $g$ is a covariant of index $g$.*

PROOF. It suffices to prove this for bracket monomials. Let $M$ be a bracket monomial of index $g$. Then, by Proposition 2.1 and the determinantal identities

$$\det\begin{pmatrix} [\alpha \quad c] & [\beta \quad c] \\ [\alpha \quad d] & [\beta \quad d] \end{pmatrix} = [c \quad d][\alpha \quad \beta],$$

$$\det\begin{pmatrix} [\alpha \quad c] & [u \quad c]/[c \quad d] \\ [\alpha \quad d] & [u \quad d]/[c \quad d] \end{pmatrix} = [\alpha \quad u],$$

we have, for any binary form $f(x, y)$ and any change of variables $(c, d)$,

$$\langle U(\bar{f}) \mid M \rangle = \langle U(f) \mid [c \quad d]^g M \rangle = [c \quad d]^g \langle U(f) \mid M \rangle.$$

Hence, $\langle U \mid M \rangle$ is a covariant of index $g$. □

Remarkably, the converse of this theorem is also true. To prove this, we require two results: the straightening algorithm and the second fundamental theorem of the umbral notation. These results will occupy the next two sections.

3.2 *The straightening algorithm.* In order to prove the converse of Theorem 3.1, we must first engage in a detailed study of the combinatorics of bracket polynomials. The straightening algorithm, which we now describe, can be viewed as the central combinatorial algorithm in invariant theory.

Let $\mathcal{A} = \{\alpha, \beta, \gamma, \dots\}$ be an alphabet linearly ordered in such a way that $\alpha < \beta < \gamma < \cdots$. Let $M$ be a bracket monomial. Thus, $M$ is a product, say,

$[\alpha\quad\beta][\alpha\quad\gamma]\cdots[\delta\varepsilon]$ of $h$ brackets. We rewrite $M$ as a *tableau* (of *height* $h$)

$$\begin{bmatrix} \alpha & \beta \\ \alpha & \gamma \\ \vdots & \\ \delta & \varepsilon \end{bmatrix} = [\alpha\quad\beta][\alpha\quad\gamma]\cdots[\delta\quad\varepsilon].$$

We call such a tableau *standard* if the letters in each row are increasing from left to right, and the letters in each column are nondecreasing from top down. A bracket monomial is *standard* if, by permuting the brackets and replacing a bracket $[\alpha\quad\beta]$ by $-[\beta\quad\alpha]$, it can be written as a standard tableau.

THEOREM 3.2. *The standard bracket monomials form a basis for the vector space of bracket polynomials.*

The theorem follows from three lemmas.

LEMMA 3.1 (THE SYZYGY). *Let $\alpha, \beta, \gamma, \delta$ be letters in the alphabet $\mathcal{Q}$ with $\alpha < \beta < \gamma < \delta$. Then*

$$\begin{bmatrix} \alpha & \delta \\ \beta & \gamma \end{bmatrix} = -\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} + \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}.$$

PROOF. This is equivalent to the determinantal identity

$$\det\begin{pmatrix} [\alpha\quad\beta] & [\alpha\quad\gamma] \\ [\delta\quad\beta] & [\delta\quad\gamma] \end{pmatrix} = \det\begin{pmatrix} \alpha_1 & \alpha_2 \\ \delta_1 & \delta_2 \end{pmatrix}\det\begin{pmatrix} \beta_2 & \gamma_2 \\ -\beta_1 & -\gamma_1 \end{pmatrix}. \qquad \square$$

LEMMA 3.2. *Any bracket monomial can be written as a linear combination with integer coefficients of standard bracket monomials.*

PROOF. We first impose a total ordering on the collection of all tableaux of a given height. If $M$ is the tableau

$$\begin{bmatrix} \alpha & \beta \\ \alpha & \gamma \\ \vdots & \\ \delta & \varepsilon \end{bmatrix},$$

we associate with $M$ the *row sequence* of letters $\alpha\beta\alpha\gamma\cdots\delta\varepsilon$, obtained by writing out the tableau along a straight line. We say that $M > N$ if the row sequence of $M$ is lexicographically greater than the row sequence of $N$.

Rewrite a bracket monomial as a tableau $M$ such that the rows are increasing and the *first* column is nondecreasing. Suppose the resulting tableau is not standard and, going down the second column, the first violation of standardness occurs between the $i$th and $(i+1)$st rows. These two rows must be of the form

$$\begin{bmatrix} \alpha & \delta \\ \beta & \gamma \end{bmatrix}$$

where $\alpha$, $\beta$, $\gamma$, $\delta$ are letters such that $\alpha < \beta < \gamma < \delta$. Using the syzygy, we have

$$M = \begin{bmatrix} & \vdots & \\ \alpha & & \delta \\ \beta & & \gamma \\ & \vdots & \end{bmatrix} = - \begin{bmatrix} & \vdots & \\ \alpha & & \beta \\ \gamma & & \delta \\ & \vdots & \end{bmatrix} + \begin{bmatrix} & \vdots & \\ \alpha & & \gamma \\ \beta & & \delta \\ & \vdots & \end{bmatrix}.$$

Observe now that the row sequences of both tableaux on the right are *strictly* lexicographically smaller than the row sequence of $M$. If these two tableaux are nonstandard, we repeat the procedure on each. As there are only a finite number of tableaux of the same height as $M$, this process must terminate, yielding an expression of $M$ as a linear combination with integer coefficients of standard bracket monomials. □

The process described in the proof specifies an algorithm, called the *straightening algorithm*, for writing a bracket polynomial as a linear combination with integer coefficients of standard bracket monomials.

LEMMA 3.3. *The standard bracket monomials form a linearly independent set in the bracket subspace* $\mathcal{B}$.

PROOF. Among all the nontrivial linear dependence relations between standard bracket monomials, choose one (with nonzero scalars $c_k$),

$$\sum_{k=1}^{m} c_k M_k = 0,$$

in which (a) the number of distinct letters in the bracket monomials $M_k$ is as small as possible, and (b) subject to (a), the maximum height of the bracket monomials $M_k$ is as small as possible. Let $\delta$ and $\varepsilon$ be the two largest letters occurring in this linear relation. By (b), the bracket $[\delta \quad \varepsilon]$ cannot be a common factor of all the bracket monomials $M_k$. Hence, on setting $\delta$ equal to $\varepsilon$, not all the bracket monomials become zero. Moreover, since $\delta$ and $\varepsilon$ are the two largest letters, those bracket monomials which remain nonzero also remain standard. We thus obtain a nontrivial dependence relation with fewer distinct letters, contradicting our initial choice. □

Theorem 3.2 can be regarded as a rigorous reformulation of the following nineteenth century heuristic.

COROLLARY 3.1. *Every algebraic relation between bracket polynomials is deducible from the syzygy and antisymmetry.*

3.3 *The Second Fundamental Theorem.* In this section we answer the following question: Let $P$ be a polynomial in the umbral space $\mathcal{U}$ and suppose $\langle U | P \rangle = 0$. What can be said about $P$? In addition to playing a crucial rôle in the proof of the converse of Theorem 3.1, the answer to this question also gives a simple criterion for deciding when two polynomials in the umbral space have equal umbral evaluations.

Let $P$ be a polynomial in the umbral space and $\mathcal{C}$ the set of Greek umbral letters occurring (nontrivially) in $P$. The polynomial $P$ is said to be *irredundant* (*for binary forms of degree* $n$) if for every monomial $N$ in $P$ and every Greek umbral letter $\alpha$ in $\mathcal{C}$, the total degree of the variables $\alpha_1$ and $\alpha_2$ is $n$. It is evident from §2.2 that every homogeneous polynomial $I(A_0,\ldots,A_n, X, Y)$ can be represented umbrally by an irredundant polynomial.

Our main result is a characterization of the irredundant polynomials $P$ in $\mathcal{U}$ whose umbral evaluation $\langle U|P\rangle$ is the identically zero polynomial. To state this result we need the notion of symmetrization. Let $P$ be an irredundant polynomial in $\mathcal{U}$, $\mathcal{C}$ the set of Greek umbral letters occurring in $P$, and $d$ the cardinality of $\mathcal{C}$. If $\pi$ is a permutation of $\mathcal{C}$, the polynomial $\pi(P)$ is defined to be the polynomial obtained from $P$ by replacing each letter $\gamma$ in $P$ by its image $\pi(\gamma)$ under the permutation $\pi$. The *symmetrization* $S(P)$ of the polynomial $P$ is the irredundant polynomial defined by

$$S(P) = \frac{1}{d!}\sum_{\pi}\pi(P),$$

where the summation is over all permutations $\pi$ of $\mathcal{C}$. If $P = S(P)$, we say that $P$ is a *symmetrized* polynomial.

LEMMA 3.4 (THE SYMMETRIZATION CONDITION). *Let $P$ be an irredundant polynomial in $\mathcal{U}$. Then $\langle U|P\rangle$ is identically zero if and only if the symmetrization $S(P)$ is the identically zero polynomial in $\mathcal{U}$.*

PROOF. By definition of the umbral operator $U$,

$$\langle U|P\rangle = \langle U|S(P)\rangle.$$

Hence, if $S(P) \equiv 0$, then $\langle U|P\rangle \equiv 0$.

To prove the converse, let $\mathcal{C}$ be the set of Greek umbral letters in $P$ and let $F(x, y)$ be the binary form defined by

$$F(x, y) = \sum_{\gamma\in\mathcal{C}} \lambda_\gamma(\gamma_1 x + \gamma_2 y)^n,$$

where $\lambda_\gamma$, $\gamma \in \mathcal{C}$, are new variables. The coefficient of $x^i y^{n-i}$ in the form $F(x, y)$ equals

$$\sum_{\gamma\in\mathcal{C}} \lambda_\gamma \gamma_1^i \gamma_2^{n-i}.$$

Therefore, by definition of the umbral functional $U(F)$,

$$\langle U(F)|\alpha_1^i\alpha_2^{n-i}\rangle = \sum_{\gamma\in\mathcal{C}} \lambda_\gamma\gamma_1^i\gamma_2^{n-i},$$

and, by the multiplicative rule,

$$\left\langle U(F)|\prod_{\alpha\in\mathcal{C}} \alpha_1^{e(\alpha)}\alpha_2^{n-e(\alpha)}\right\rangle = \prod_{\alpha\in\mathcal{C}}\left(\sum_{\gamma\in\mathcal{C}} \lambda_\gamma\gamma_1^{e(\alpha)}\gamma_2^{n-e(\alpha)}\right).$$

Now expand the right-hand side as a polynomial in the variables $\lambda_\gamma$. The multilinear monomial $\Pi_{\gamma\in\mathcal{C}}\lambda_\gamma$ in this expansion is obtained by choosing one summand $\lambda_\gamma\gamma_1^{e(\alpha)}\gamma_2^{n-e(\alpha)}$ from each sum in such a way that no two summands from two distinct sums have the same umbral letter $\gamma$, taking their product, and summing over all products obtained in this way. Thus, the coefficient of $\Pi_{\gamma \in \mathcal{C}} \lambda_\gamma$ equals

$$\sum_{\pi} \prod_{\alpha \in \mathcal{C}} \pi(\alpha)_1^{e(\alpha)} \pi(\alpha)_2^{n-e(\alpha)},$$

where the summation ranges over all permutations $\pi$ of $\mathcal{C}$: that is, it equals the symmetrization

$$S\Big(\prod_{\alpha \in \mathcal{C}} \alpha_1^{e(\alpha)} \alpha_2^{n-e(\alpha)}\Big).$$

Now, as $P$ is an irredundant polynomial, it is a linear combination of monomials of the form

$$\Big(\prod_{\alpha \in \mathcal{C}} \alpha_1^{e(\alpha)} \alpha_2^{n-e(\alpha)}\Big) u_1^i u_2^j.$$

Therefore, although $\langle U(F) | P \rangle$ cannot be easily computed, we know nevertheless that the coefficient of $\Pi_{\gamma \in \mathcal{C}} \lambda_\gamma$ in $\langle U(F) | P \rangle$, considered as a polynomial in $\lambda_\gamma$, is the symmetrization $d!S(P)$, since symmetrization is a linear operator. We conclude that if $\langle U | P \rangle$ is identically zero, then $d!S(P)$, and hence $S(P)$, is also identically zero. □

Example. Let $U$ be the umbral operator for binary cubics. Consider the irredundant bracket polynomial $[\alpha \quad \beta]^3$. As

$$\big\langle U | [\alpha \quad \beta]^3 \big\rangle = -\big\langle U | [\beta \quad \alpha]^3 \big\rangle = -\big\langle U | [\alpha \quad \beta]^3 \big\rangle,$$

the umbral evaluation of $[\alpha \quad \beta]^3$ is identically zero. As the lemma predicts, its symmetrization $\frac{1}{2}([\alpha \quad \beta]^3 + [\beta \quad \alpha]^3)$ is also identically zero. □

From the preceding lemma, we obtain

Theorem 3.3 (The Second Fundamental Theorem). *Let $U$ be the umbral operator for binary forms of degree $n$ and let $P$ and $Q$ be polynomials in the umbral space $\mathfrak{U}$ such that $\langle U | P \rangle = \langle U | Q \rangle$. Then $P$ can be obtained from $Q$ by a sequence of operations of the following four types*:

I. *an application of the $k$-algebra axioms in the polynomial algebra $k[\alpha_1, \alpha_2, \beta_1, \beta_2, \ldots, u_1, u_2]$*;

II. *adding a scalar multiple of a* redundant *monomial, that is a monomial in $\mathfrak{U}$ containing a Greek umbral letter $\gamma$ for which the total degree of $\gamma_1$ and $\gamma_2$ is not equal to $n$ or zero*:

III. *replacing any monomial $M$ by $M'$, where $M'$ is obtained from $M$ by replacing the variables $\alpha_1$ and $\alpha_2$ for some Greek umbral letter $\alpha$ occurring in $M$ by the variables $\delta_1$ and $\delta_2$, where $\delta$ is an umbral letter* not *occurring in $M$*;

(IV). *replacing any monomial $M$ by $\pi(M)$, where $\pi$ is a permutation of the set of umbral letters occurring in $M$.*

Proof. By applying operations of type II, we can assume $P$ and $Q$ contain no redundant monomials. We next write $P$ as a sum, $P = P_1 + P_2 + \cdots + P_r$, where $P_i$ is the linear combination with the same coefficients as in $P$ of all monomials in $P$ containing $i$ distinct Greek umbral letters. Similarly, write

$Q = Q_1 + Q_2 + \cdots + Q_s$. Since $\langle U|P\rangle = \langle U|Q\rangle$, and the degree (as a polynomial in $k[\alpha_1, \alpha_2, \ldots, u_1, u_2]$) of $\langle U|M\rangle$, where $M$ is a monomial, equals the number of distinct Greek umbral letters in $M$, we have, for all $i$,

$$\langle U|P_i\rangle = \langle U|Q_i\rangle.$$

Thus, it suffices to prove the theorem for polynomials $P$ and $Q$ such that every monomial in $P$ and $Q$ has the same number, $d$ say, of distinct Greek umbral letters. By applying operations of type III, we can assume that the set of distinct Greek umbral letters in $P$ and the set of distinct Greek umbral letters in $Q$ are the same and have cardinality $d$: that is to say, $P$ and $Q$ are irredundant polynomials formed with the same set of Greek umbral letters.

The proof can now be completed by observing that, as $\langle U|P\rangle = \langle U|Q\rangle$, $\langle U|P - Q\rangle = 0$ and hence, by Lemma 3.4, the symmetrization $S(P - Q) = S(P) - S(Q) = 0$. Thus, $S(P) = S(Q)$. But $P$ can be obtained from $S(P)$ and $S(Q)$ obtained from $Q$ by operations of types I and IV. □

3.4 *The First Fundamental Theorem.* We are now ready to prove the basic result underlying the use of umbral notation in invariant theory.

THEOREM 3.1 (THE FIRST FUNDAMENTAL THEOREM): PART II. *Let $I$ be a covariant of index $g$ of binary forms of degree $n$. Then there exists a bracket polynomial $P$ of index $g$ such that $I = \langle U|P\rangle$.*

PROOF. It suffices to prove the theorem for a homogeneous covariant $I$ of degree $d$, order $t$, and index $g$. Let

$$I = \langle U|P(\alpha_1, \alpha_2, \beta_1, \beta_2, \ldots, u_1, u_2)\rangle$$

be an irredundant umbral representation of $I$. As the umbral evaluation of a polynomial $P$ and its symmetrization $S(P)$ are equal, we can assume that the polynomial $P$ is a symmetrized polynomial. As $I$ is a covariant, we have, for any binary form

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k}$$

of degree $n$ and any change of variables $(c, d)$,

$$\begin{aligned}[c\ \ d]^g I(a_0, \ldots, a_n, x, y) &= I(\bar{a}_0, \ldots, \bar{a}_n, \bar{x}, \bar{y}) \\ &= \langle U(f)|P([\alpha\ \ c], [\alpha\ \ d], [\beta\ \ c], [\beta\ \ d], \ldots, \\ &\qquad [u\ \ c]/[c\ \ d], [u\ \ d]/[c\ \ d])\rangle.\end{aligned}$$

On multiplying both sides by $[c\ \ d]^t$, we obtain

$$(3.1)\quad \begin{aligned}&[c\ \ d]^{g+t} I(a_0, \ldots, a_n, x, y) \\ &= \langle U(f)|P([\alpha\ \ c], [\alpha\ \ d], [\beta\ \ c], [\beta\ \ d], \ldots, [u\ \ c], [u\ \ d])\rangle\end{aligned}.$$

Observe that the polynomial $P([\alpha\ \ c], [\alpha\ \ d], \ldots, [u\ \ c], [u\ \ d])$ remains irredundant and symmetrized.

The identity (3.1) holds for all scalar values $c_1$, $c_2$, $d_1$, $d_2$ for which $c_1 d_2 - c_2 d_1 \neq 0$. Therefore, it holds as a polynomial identity in the variables

$c_1, c_2, d_1, d_2$. Using this fact, we shall prove that

$$P([\alpha\ \ c],[\alpha\ \ d],\ldots,[u\ \ c],[u\ \ d]) = [c\ \ d]^{g+t}Q(\alpha_1,\alpha_2,\ldots,u_1,u_2),$$

where $Q$ is a *bracket* polynomial of index $g$ in the letters $\alpha, \beta,\ldots,u$, *not* containing any of the variables $c_1, c_2, d_1, d_2$. We shall then be able to cancel the factor $[c\ \ d]^{g+t}$ from both sides of identity (3.1), thus obtaining an umbral representation of $I$ by a bracket polynomial.

Let $\mathcal{A}^+$ be the alphabet $\{c, d, \alpha, \beta,\ldots,\omega, u\}$ linearly ordered in such a way that $c < d < \alpha < \beta < \cdots < \omega < u$, and let $\mathcal{B}^+$ be the space of bracket polynomials formed with the letters in $\mathcal{A}^+$. Applying the straightening algorithm to $P([\alpha\ \ c],[\alpha\ \ d],\ldots,[u\ \ c],[u\ \ d])$, considered as a bracket polynomial in $\mathcal{B}^+$, we can write $P([\alpha\ \ c],[\alpha\ \ d],\ldots,[u\ \ c],[u\ \ d])$ as a linear combination $\Sigma b_k M_k$, with $b_k \neq 0$, of distinct standard bracket monomials in $\mathcal{B}^+$.

LEMMA 3.5. *The polynomial $P(\alpha_1, \alpha_2,\ldots,u_1, u_2)$ may be so chosen that the letter $c$, as well as the letter $d$, occurs exactly $g + t$ times in each of the standard monomials $M_k$.*

PROOF. Let $A$ be a new variable. Using the fact that $[\alpha\ \ (Ac)] = A[\alpha\ \ c]$ for any letter $\alpha$, we obtain, on replacing $c_1$ and $c_2$ by $Ac_1$ and $Ac_2$ in (3.1),

$$A^{g+t}[c\ \ d]^{g+t}I(a_0,\ldots,a_n, x, y) = \left\langle U(f)\,|\,\sum b_k A^{c(k)}M_k\right\rangle,$$

where $c(k)$ is the number of occurrences of $c$ in the bracket monomial $M_k$. Equating coefficients of $A^{g+t}$, we obtain

$$[c\ \ d]^{g+t}I(a_0,\ldots,a_n, x, y) = \left\langle U(f)\,|\,\sum{}' b_k M_k\right\rangle,$$

where the prime on the summation indicates that those bracket monomials with $c(k) \neq g + t$ are omitted. We can now replace $P(\alpha_1, \alpha_2,\ldots,u_1, u_2)$ with the polynomial obtained from $\Sigma' b_k M_k$ by setting $c_1 = 1$, $c_2 = 0$, $d_1 = 0$, $d_2 = 1$. □

Consider now a bracket monomial $M_k$ in the expansion of

$$P([\alpha\ \ c],[\alpha\ \ d],\ldots,[u\ \ c],[u\ \ d])$$

as a linear combination of standard bracket monomials. It is of the form

$$[c\ \ d]^{l(k)}\begin{bmatrix} c & \alpha \\ & \vdots \\ c & \beta \\ d & \gamma \\ & \vdots \end{bmatrix},$$

where $l(k)$ is the number of brackets $[c\ \ d]$ occurring in $M_k$. Let $l$ be the minimum of these integers $l(k)$. By Lemma 3.5, $l \leq g + t$.

If $g + t = l$, then $M_k = [c\ \ d]^{g+t}M_k'$, where (by Lemma 3.5 again) there are no further occurrences of the letters $c$ and $d$ in the bracket monomial $M_k'$. We can thus cancel $[c\ \ d]^{g+t}$ from both sides of (3.1).

Now suppose $g + t > l$. Writing $M_k = [c \quad d]^l M_k'$, we can cancel $[c \quad d]^l$ from both sides of (3.1) to obtain

$$(3.2) \qquad [c \quad d]^{g+t-l} I(a_0, \ldots, a_n, x, y) = \left\langle U(f) \,|\, \sum b_k M_k' \right\rangle,$$

where $g + t - l > 0$ and there is a standard bracket monomial $M_j'$ which does not contain the bracket $[c \quad d]$. The identity (3.2) holds for all $c_1$, $c_2$, $d_1$, $d_2$ in the infinite field $k$ such that $[c \quad d] \neq 0$; thus, it holds as a polynomial identity in the variables $c_1$, $c_2$, $d_1$, $d_2$. We can therefore set $c_1 = d_1$ and $c_2 = d_2$. This yields the identity

$$\left\langle U(f) \,|\, \sum b_k \hat{M}_k \right\rangle = 0,$$

where $\hat{M}_k$ is the bracket monomial obtained from $M_k'$ by setting $c = d$. Note that, as $c$ and $d$ precede all the other letters in the linear ordering, the bracket monomials which do not vanish after $d$ is set equal to $c$ remain standard. Moreover, $\Sigma b_k \hat{M}_k$ is still symmetrized. We can thus appeal to Lemma 3.4 to conclude that $\Sigma b_k \hat{M}_k$ is identically zero.

As remarked earlier, there is a standard bracket monomial $M_j'$ which does not contain the bracket $[c \quad d]$. As $\Sigma b_k \hat{M}_k$ is zero and the standard bracket monomials are linearly independent, there exists a subset $E$ of indices with $j \in E$ such that $\Sigma_{k \in E} b_k \hat{M}_k = 0$, and for all $k$ in $E$, $M_k' \neq M_j'$ but $\hat{M}_k = \hat{M}_j$. As $b_j \neq 0$, there is an index $m$ not equal to $j$ in $E$. As $\hat{M}_j = \hat{M}_m$ and $\hat{M}_j$ is standard, we have

$$\hat{M}_j = \hat{M}_m = \begin{bmatrix} c & * \\ c & * \\ & \vdots \\ c & * \\ * & * \\ & \vdots \\ * & * \end{bmatrix}$$

where $c$ occurs as the first letter in the first $2(g + t - l)$ rows and an asterisk stands for a Greek or Roman umbral letter. However, as $M_j'$ and $M_m'$ differ from $\hat{M}_j$ and $\hat{M}_m$ *only* in that $c$ is set equal to $d$, we have, by Lemma 3.5,

$$M_j' = M_m' = \begin{bmatrix} c & * \\ & \vdots \\ c & * \\ d & * \\ & \vdots \\ d & * \\ * & * \\ & \vdots \\ * & * \end{bmatrix},$$

where $c$ occurs as the first letter in the first $g+t-l$ rows and $d$ occurs as the first letter in the next $g+t-l$ rows. This contradicts the assumption that $M_j'$ and $M_m'$ are different.

This completes the proof of the first fundamental theorem. □

Our proof gives an algorithm for expressing any covariant $I$ as the umbral evaluation of a linear combination of standard bracket monomials. We shall illustrate this algorithm with the discriminant of binary quadratics.

Example (The Discriminant). Let $D$ be the discriminant of binary quadratics, that is,

$$D = A_0A_2 - A_1^2.$$

Then an umbral representation of $D$ is $\langle U \,|\, P(\alpha_1, \alpha_2, \beta_1, \beta_2)\rangle$, where

$$P(\alpha_1, \alpha_2, \beta_1, \beta_2) = \alpha_2^2\beta_1^2 - \alpha_1\alpha_2\beta_1\beta_2.$$

Symmetrizing $P$, we obtain

$$\frac{1}{2}\left(\alpha_2^2\beta_1^2 + \alpha_1^2\beta_2^2 - 2\alpha_1\alpha_2\beta_1\beta_2\right).$$

Replacing $\alpha_1$ by $[\alpha \quad c]$, $\alpha_2$ by $[\alpha \quad d]$,..., we obtain the following linear combination of (nonstandard) bracket monomials in the letters $\alpha$, $\beta$, $c$ and $d$:

$$\frac{1}{2}\left(\begin{bmatrix} \alpha & d \\ \alpha & d \\ \beta & c \\ \beta & c \end{bmatrix} + \begin{bmatrix} \alpha & c \\ \alpha & c \\ \beta & d \\ \beta & d \end{bmatrix} - 2\begin{bmatrix} \alpha & c \\ \alpha & d \\ \beta & c \\ \beta & d \end{bmatrix}\right).$$

Applying the straightening algorithm to each bracket monomial and adding the results, we obtain

$$\frac{1}{2}\begin{bmatrix} c & d \\ c & d \\ \alpha & \beta \\ \alpha & \beta \end{bmatrix}.$$

Thus, an umbral representation for the discriminant $D$ in terms of bracket polynomials is

$$D = \left\langle U \,|\, \tfrac{1}{2}[\alpha \quad \beta]^2 \right\rangle. \quad \square$$

Let $I$ be a homogeneous covariant of binary forms of degree $n$, of degree $d$, order $t$, and index $g$. Let $I = \langle U \,|\, \Sigma b_k M_k\rangle$ be an irredundant umbral representation of $U$ by a linear combination of bracket monomials. Then for every bracket monomial $M_k$, the number of brackets in $M_k$ containing only Greek letters is equal to $g$, the index of $I$, and the number of brackets in $M_k$ containing the Roman letter $u$ is equal to $t$, the order of $I$. Thus, every bracket monomial has the same number $h$ of brackets, where $h = g + t$. Since each bracket contains two letters, each Greek umbral letter occurs $n$ times, and the Roman letter $u$ occurs $t$ times, the numbers $n$, $m$, $t$, $h$, and $g$ satisfy the relations

$$2h = dn + t, \qquad 2g + t = dn.$$

In particular, the index $g$ of a homogeneous covariant of binary forms of degree $n$ can be deduced from its degree $d$ and order $t$ by the equation

$$g = \tfrac{1}{2}(dn - t).$$

The two fundamental theorems and their proofs extend immediately to several binary forms. An important consequence is the following result which implies that in most situations, it suffices to consider only joint *in*variants.

COROLLARY 3.2. *Let $I(A_0,\ldots,A_n, X, Y)$ be a homogeneous polynomial of degree $d$ and order $l$. Then $I(A_0,\ldots,A_n, X, Y)$ is a covariant of index $g$ of binary forms of degree $n$ if and only if the polynomial $I(A_0,\ldots,A_n, S, -T)$, obtained by setting $X = S$ and $Y = -T$, is a joint invariant of index $g + l$ of binary forms of degree $n$ and linear forms $tx + sy$.*

PROOF. If $\alpha$ is an umbral letter belonging to the linear form $tx + sy$, then $\langle U|\alpha_1\rangle = T$ and $\langle U|\alpha_2\rangle = S$. Since $\langle U|u_1\rangle = -Y$ and $\langle U|u_2\rangle = X$, if $P$ is an irredundant polynomial in $\mathfrak{A}$ such that $\langle U|P\rangle = I(A_0,\ldots,A_n, X, Y)$, where $I$ is a homogeneous polynomial of degree $d$ and order $l$, then

$$\langle U|P'\rangle = I(A_0,\ldots,A_n, S, -T),$$

where $P'$ is the irredundant polynomial obtained by replacing the factor $u_1^i u_2^j$ in each monomial of $P$ by $\alpha_1 \cdots \beta_1\gamma_2 \cdots \delta_2$, where $\alpha,\ldots,\beta$ are the first $i$ letters and $\gamma,\ldots,\delta$ are the last $j$ letters in a set $\{\alpha,\ldots,\delta\}$ of $l$ Greek umbral letters belonging to the linear form $tx + sy$. In particular, let $I(A_0,\ldots,A_n, X, Y)$ be a homogeneous covariant and $\langle U|P\rangle$ an irredundant umbral representation of $I$ by a bracket monomial $P$. Then $P'$ is also a bracket polynomial and, hence, $\langle U|P'\rangle$ is a joint invariant. This argument can be reversed. Finally, observe that if $\alpha$ and $\beta$ are Greek umbral letters belonging to the linear form $tx + sy$, then $\langle U|[\alpha\ \ \beta]\rangle$ equals zero and, hence, the index of a joint invariant is at least the degree $l$ of that joint invariant in the coefficients of the linear form. □

The first fundamental theorem, together with Lemma 3.4, yields immediately an umbral representation for algebraic relations (or syzygies) between covariants in terms of bracket polynomials.

COROLLARY 3.3. *Every relation*

$$\sum_j b_j I_{j1} I_{j2} \cdots I_{jm(j)} = 0$$

*between covariants $I_{jk}$ of binary forms of degree $n$ can be written in the form*

$$\sum_j b_j \langle U|M_{j1}\rangle\langle U|M_{j2}\rangle \cdots \langle U|M_{jm(j)}\rangle = 0,$$

*where $I_{jk} = \langle U|M_{jk}\rangle$, $M_{jk}$ is a bracket polynomial, and*

$$P = \sum_j b_j M_{j1} M_{j2} \cdots M_{jm(j)}$$

*is a bracket polynomial in the umbral space whose symmetrization is zero.*

The First and Second Fundamental Theorems can be summarized by the following theorem.

THEOREM 3.4. *Let $\mathcal{G}[n, d, t]$ be the (finite-dimensional) vector space of homogeneous covariants of binary forms of degree $n$, with degree $d$, order $t$, and index $g$, where $g = \frac{1}{2}(dn - t)$. Let $\mathcal{U}[n, d, t]$ be the subspace of the space $\mathcal{B}$ of bracket polynomials spanned by bracket monomials of height $g + t$ formed with $d$ distinct Greek umbral letters each occurring $n$ times and the Roman letter $u$ occurring $t$ times. Finally, let $\mathcal{U}^S[n, d, t]$ be the space of all symmetrized polynomials in $\mathcal{U}[n, d, t]$. Then*

$$\mathcal{U}^S[n, d, t] \simeq \mathcal{G}[n, d, t],$$

*the isomorphism being given by the umbral operator $U$ for binary forms of degree $n$.*

## 4. Covariants in terms of the roots.

4.1 *Homogenized roots.* When the leading coefficient $a_n$ is nonzero, the remaining coefficients $a_0, a_1, \dots, a_{n-1}$ of the binary form

$$\begin{aligned} f(x, y) &= \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k} \\ &= a_n(x - \lambda_1 y)(x - \lambda_2 y) \cdots (x - \lambda_n y) \end{aligned}$$

can be written in terms of the leading coefficient $a_n$ and the roots $\lambda_1, \lambda_2, \dots, \lambda_n$ of the polynomial $f(x, 1)$. Indeed, the coefficient $a_{n-k}$ is a multiple of the $k$th elementary symmetric function $e_k(\lambda_1, \dots, \lambda_n)$ of the roots. More precisely,

$$(4.1) \qquad \binom{n}{k} a_{n-k} = (-1)^k a_n e_k(\lambda_1, \dots, \lambda_n) = (-1)^k a_n \sum \lambda_{i_1} \cdots \lambda_{i_k},$$

the sum ranging over all $k$-element subsets of $\{1, 2, \dots, n\}$. We note an alternative version of (4.1):

$$(4.2) \qquad a_{n-k} = (-1)^k \frac{a_n}{n!} \sum_{\pi} \lambda_{\pi(1)} \cdots \lambda_{\pi(k)},$$

where the sum ranges over all permutations $\pi$ of $\{1, 2, \dots, n\}$. (The equivalence of (4.1) and (4.2) follows from the fact that for any given $k$-element subset $S$ of $\{1, 2, \dots, n\}$, there exist $k!(n - k)!$ permutations such that $\pi(S) = S$.)

Now let $k[a_n, \lambda_1, \dots, \lambda_n, x, y]$ be the ring of all polynomials in the variables $a_n, \lambda_1, \dots, \lambda_n, x$ and $y$. This ring will be called the *algebra of roots*. The substitutions

$$\begin{aligned} &A_n \leftarrow a_n, \\ &A_{n-k} \leftarrow (-1)^k \frac{a_n}{n!} \sum_{\pi} \lambda_{\pi(1)} \cdots \lambda_{\pi(k)}, \\ &X \leftarrow x, \quad Y \leftarrow y, \end{aligned}$$

define an algebra homomorphism $\mathbf{r}$ from the algebra $k[A_0, \dots, A_n, X, Y]$ of all polynomials in the variables $A_0, A_1, \dots, A_n$, $X$ and $Y$ to the algebra of roots. If $I(A_0, \dots, A_n, X, Y)$ is a polynomial in $A_i$, $X$ and $Y$, the image of $I$ under $\mathbf{r}$ is called the *representation of $I$ in terms of the roots*.

For studying covariants it is often useful to consider a more symmetrical representation in terms of homogenized roots. Let the binary form $f(x, y)$ be written (not uniquely, of course) as a product of $n$ linear forms:

$$f(x, y) = (\mu_1 x - \nu_1 y)(\mu_2 x - \nu_2 y) \cdots (\mu_n x - \nu_n y).$$

The coefficients $\mu_i$, $\nu_i$ of the linear forms are called the *homogenized roots* of $f(x, y)$. On expanding and equating coefficients, we obtain

$$\begin{aligned} (4.3) \qquad a_{n-k} &= \frac{(-1)^k}{n!} \sum_{\pi} \nu_{\pi(1)} \cdots \nu_{\pi(k)} \mu_{\pi(k+1)} \cdots \mu_{\pi(n)} \\ &= \frac{(-1)^k}{n!} \mu_1 \cdots \mu_n e_k\left(\frac{\nu_1}{\mu_1}, \dots, \frac{\nu_n}{\mu_n}\right), \end{aligned}$$

the sum ranging over all permutations $\pi$ of $\{1, 2, \dots, n\}$.

In analogy with the roots, we define the *algebra of homogenized roots* to be the ring $k[\mu_1, \dots, \mu_n, \nu_1, \dots, \nu_n, x, y]$ of polynomials in the variables $\mu_1, \dots, \mu_n$, $\nu_1, \dots, \nu_n$, $x$ and $y$. The substitutions

$$\begin{aligned} &A_{n-k} \leftarrow \frac{(-1)^k}{n!} \sum_{\pi} \nu_{\pi(1)} \cdots \nu_{\pi(k)} \mu_{\pi(k+1)} \cdots \mu_{\pi(n)}, \\ &X \leftarrow x, \quad Y \leftarrow y, \end{aligned}$$

define an algebra homomorphism $\mathbf{h}$ from the algebra $k[A_i, X, Y]$ to the algebra of homogenized roots. If $I$ is a polynomial in $A_i$, $X$ and $Y$, the image of $I$ under $\mathbf{h}$ is called the *representation of $I$ in terms of homogenized roots*.

When is a polynomial in homogenized roots expressible in terms of the coefficients $A_0, \dots, A_n$? The answer is contained in our next result, which is a homogenized version of the fundamental theorem of symmetric functions. To state this result, we need the following definitions. Let

$$M = \mu_1^{a_1} \mu_2^{a_2} \cdots \mu_n^{a_n} \nu_1^{b_1} \nu_2^{b_2} \cdots \nu_n^{b_n} x^{c_1} y^{c_2}$$

be a monomial in the homogenized roots. For $i \in \{1, \dots, n\}$, the *multiplicity $m_i$* of $i$ in $M$ is defined by

$$m_i = a_i + b_i.$$

A monomial $M$ is said to be *regular of degree $d$* if

$$m_1 = m_2 = \cdots = m_n = d.$$

A polynomial $R(\mu_i, \nu_i, x, y)$ in the algebra of homogenized roots is said to be *regular of degree $d$* if every monomial in $R$ is regular of the same degree $d$.

PROPOSITION 4.1. *Let $R(\mu_i, \nu_i, x, y)$ be a polynomial in the algebra of homogenized roots. Then $R$ is expressible as a polynomial in the* homogenized symmetric functions

$$a_k(\mu_i, \nu_i) = \frac{1}{n!} \sum_{\pi} \nu_{\pi(1)} \cdots \nu_{\pi(k)} \mu_{\pi(k+1)} \cdots \mu_{\pi(n)}$$

(*with coefficients in the algebra $k[x, y]$ of polynomials in the variables $x$ and $y$*) *if and only if $R$ is regular and $R$ is jointly symmetric in $\mu_i$ and $\nu_i$* (*that is, for all permutations $\pi$ of $\{1, 2, \dots, n\}$, $R(\mu_i, \nu_i, x, y) = R(\mu_{\pi(i)}, \nu_{\pi(i)}, x, y)$.*

PROOF. As $a_k(\mu_i, \nu_i)$ is jointly symmetric and regular, and these properties are preserved under multiplication, one implication is immediate.

To prove the converse, let $R$ be regular of degree $d$. Then $R$ can be rewritten as

$$R(\mu_i, \nu_i, x, y) = (\mu_1 \cdots \mu_n)^d \hat{R}(\nu_i/\mu_i, x, y),$$

where $\hat{R}$ is a symmetric polynomial in the variables $\lambda_i = \nu_i/\mu_i$. By the fundamental theorem of symmetric functions, we can write $\hat{R}$ as a polynomial $Q$ (with coefficients in $k[x, y]$) in the elementary symmetric functions $e_k(\nu_i/\mu_i)$. Multiplying $Q$ by $(\mu_1 \cdots \mu_n)^d$ and distributing factors of $\mu_1 \cdots \mu_n$ among the elementary symmetric functions $e_k(\nu_i/\mu_i)$, we obtain an expression of $R$ in terms of the homogenized symmetric functions. □

All definitions and results in this section extend immediately to several binary forms.

4.2 *Tableaux in terms of roots.* Let $P$ be a bracket polynomial in the umbral space $\mathfrak{U}$. Applying the umbral operator $U$ and then the homomorphism $\mathbf{h}$ to $P$, we obtain a polynomial in the algebra of homogenized roots. Our main result in this section is a constructive description of the composite function $\mathbf{h} \circ U$. This description is in the form of an algorithm for translating the umbral representation of a covariant $I$ of the form $f(x, y)$ into a representation of $I$ in terms of the homogenized roots of $f(x, y)$.

Let $T$ be a bracket monomial in $\mathfrak{U}$ and let $U$ be the umbral operator for binary forms of degree $n$. The element $\mathbf{h}(\langle U | T \rangle)$ in the algebra of homogenized roots, which is also the representation of the covariant $\langle U | T \rangle$ in terms of homogenized roots, can be obtained by the following algorithm.

ALGORITHM 4.1. *Let $\mathcal{A}$ be the set of all Greek umbral letters occurring in $T$ and let $d$ be the cardinality of $\mathcal{A}$. We shall assume that every letter in $\mathcal{A}$ occurs exactly $n$ times in $T$. If not,* $\mathbf{h}(\langle U | T \rangle) = 0$ *and the algorithm terminates.*

*Step* 1. Let the brackets in $T$ be written as a tableau in some fixed order:

$$T = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \\ & \vdots \\ \omega & u \end{bmatrix}. \tag{4.4}$$

From this particular expression of $T$, construct a new tableau as follows. Let $\alpha$ be a Greek umbral letter in $\mathcal{A}$. Going down the first column and then the second column of the tableau in (4.4), replace the first occurrence of $\alpha$ by the integer 1, the second by 2,..., and the $n$th by $n$. Repeat this for every Greek umbral letter, thus obtaining a tableau $\hat{T}$ whose entries are either integers in $\{1, 2, \ldots, n\}$ or the Roman letter $u$. Putting the tableaux $T$ and $\hat{T}$ side by side, we obtain the *double tableau*

$$\left[\begin{array}{cc|cc} \alpha & \beta & i & j \\ \gamma & \delta & k & l \\ \vdots & & \vdots & \\ \omega & u & p & u \end{array}\right].$$

*Step* 2. Next let $\Phi$ be a function from $\mathcal{Q}$ to the set of all permutations on $\{1, 2, \ldots, n\}$. The double tableau $T[\Phi]$ is defined by

$$T[\Phi] = \left[\begin{array}{cc|cc} \alpha & \beta & \Phi(\alpha, i) & \Phi(\beta, j) \\ \gamma & \delta & \Phi(\gamma, k) & \Phi(\delta, l) \\ \vdots & & \vdots & \vdots \\ \omega & u & \Phi(\omega, p) & u \end{array}\right],$$

where $\phi(\alpha, i)$ is the image of the integer $i$ under the permutation $\Phi(\alpha)$.

*Step* 3. Let $T[\Phi]$ be the double tableau

$$\left[\begin{array}{cc|cc} \alpha & \beta & i' & j' \\ \gamma & \delta & k' & l' \\ \vdots & & \vdots & \\ \omega & u & p' & u \end{array}\right].$$

To each row in $T[\Phi]$, assign a polynomial in the homogenized roots according to the following two rules: for $i$, $j \in \{1, 2, \ldots, n\}$ and $\alpha$, $\beta \in \mathcal{Q}$,

$$(4.5) \quad [\alpha \quad \beta | i \quad j] \leftarrow (\mu_i\nu_j - \nu_i\mu_j), \qquad [\alpha \quad u | i \quad u] \leftarrow (\mu_i x - \nu_i y).$$

The polynomials $(\mu_i\nu_j - \nu_i\mu_j)$ and $(\mu_i x - \nu_i y)$ are called *differences*.

(This assignment can be visualized as a substitution by interpreting

$$[\alpha \quad \beta | i \quad j] \quad \text{and} \quad [\alpha \quad u | i \quad u]$$

as the determinants

$$\det\begin{pmatrix} (\alpha|i)_1 & (\beta|j)_1 \\ (\alpha|i)_2 & (\beta|j)_2 \end{pmatrix} \quad \text{and} \quad \det\begin{pmatrix} (\alpha|i)_1 & u_1 \\ (\alpha|i)_2 & u_2 \end{pmatrix}.$$

where $(\alpha|i)_1$, $(\alpha|i)_2, \ldots$ are new variables, and making the substitutions $(\alpha|i)_1 \leftarrow \mu_i$, $(\alpha|i)_2 \leftarrow \nu_i$, $u_1 \leftarrow y$, and $u_2 \leftarrow x$.)

Taking the product of all the differences, we obtain the polynomial $T^h[\Phi]$, given explicitly as follows:

$$T^h[\Phi] = (\mu_{i'}\nu_{j'} - \nu_{i'}\mu_{j'})(\mu_{k'}\nu_{l'} - \nu_{k'}\mu_{l'}) \cdots (\mu_{p'}x - \nu_{p'}y).$$

*Step* 4. Set

$$\mathbf{h}(\langle U | T \rangle) = \frac{(-1)^g}{(n!)^d} \sum_{\Phi} T^h[\Phi],$$

where the sum ranges over all functions $\Phi$ from the set of Greek umbral letters $\mathcal{Q}$ to the set $\Omega_n$ of permutations of $\{1, 2, \ldots, n\}$ and $g$, the index of the covariant $\langle U | T \rangle$, equals the number of brackets in $T$ not containing the Roman letter $u$. □

Example. Let $D$ be the discriminant of a binary quadratic form. Then $D$ is an invariant and is given by

$$D = A_0A_2 - A_1^2.$$

An umbral representation for $D$ is $\frac{1}{2}[\alpha \quad \beta]^2$. Following the above algorithm, we obtain four double tableaux:

$$\frac{1}{2}\left[\begin{array}{cc|cc} \alpha & \beta & 1 & 1 \\ \alpha & \beta & 2 & 2 \end{array}\right], \quad \frac{1}{2}\left[\begin{array}{cc|cc} \alpha & \beta & 1 & 2 \\ \alpha & \beta & 2 & 1 \end{array}\right],$$
$$\frac{1}{2}\left[\begin{array}{cc|cc} \alpha & \beta & 2 & 1 \\ \alpha & \beta & 1 & 2 \end{array}\right], \quad \frac{1}{2}\left[\begin{array}{cc|cc} \alpha & \beta & 2 & 2 \\ \alpha & \beta & 1 & 1 \end{array}\right].$$

On substituting according to the rules (4.5), we obtain for each of the tableaux,

$$0, \quad \tfrac{1}{2}(\mu_1\nu_2 - \nu_1\mu_2)(\mu_2\nu_1 - \nu_2\mu_1), \quad \tfrac{1}{2}(\mu_2\nu_1 - \nu_2\mu_1)(\mu_1\nu_2 - \nu_1\mu_2), \quad 0.$$

Thus, in terms of homogenized roots, the discriminant has the representation

$$D = -\tfrac{1}{4}(\mu_1\nu_2 - \nu_1\mu_2)^2.$$

THEOREM 4.1. *Algorithm* 4.1 *computes the representation of the covariant* $\langle U | T \rangle$ *in terms of the homogenized roots.*

PROOF. Let $m$ be the number of rows in $T$ and label the rows in $T$ with the integers $\{1,2,\ldots,m\}$. For $\gamma$ a Greek umbral letter occurring in $T$ and $Z$ a subset of $\{1,\ldots,m\}$, let $E_1(\gamma, Z)$ and $\bar{E}_1(\gamma, Z)$ be the subsets of row labels defined by

$$E_1(\gamma, Z) = \{i\colon i \in Z \text{ and } \gamma \text{ is the first letter in the } i\text{th row}\},$$
$$\bar{E}_1(\gamma, Z) = \{i\colon i \notin Z \text{ and } \gamma \text{ is the second letter in the } i\text{th row}\}.$$

Similarly, let

$$E_2(\gamma, Z) = \{i\colon i \in Z \text{ and } \gamma \text{ is the second letter in the } i\text{th row}\},$$
$$\bar{E}_2(\gamma, Z) = \{i\colon i \notin Z \text{ and } \gamma \text{ is the first letter in the } i\text{th row}\}.$$

Let

$$e_1(\gamma, Z) = |E_1(\gamma, Z)| + |\bar{E}_1(\gamma, Z)|$$

and

$$e_2(\gamma, Z) = |E_2(\gamma, Z)| + |\bar{E}_2(\gamma, Z)|.$$

By the assumption that the Greek umbral letter $\gamma$ occurs exactly $n$ times in $T$, we have

$$e_1(\gamma, Z) + e_2(\gamma, Z) = n. \tag{4.6}$$

For the Roman letter $u$, the sets $E_i(u, Z)$, $\bar{E}_i(u, Z)$ and the numbers $e_i(u, Z)$ are defined similarly.

LEMMA 4.1. *As a polynomial in the variables* $\gamma_1$, $\gamma_2$, $\gamma \in \mathcal{Q}$, *and* $u_1$, $u_2$, *the tableau $T$ can be expanded as*

$$T = \sum_{Z} (-1)^{m-|Z|} M(Z), \tag{4.7}$$

*where*

$$M(Z) = \left( \prod_{\gamma \in \mathfrak{A}} \gamma_1^{e_1(\gamma, Z)} \gamma_2^{e_2(\gamma, Z)} \right) u_1^{e_1(u,Z)} u_2^{e_2(u,Z)},$$

*and the sum is over all subsets $Z$ of the set of row labels $\{1, 2, \ldots, m\}$.*

PROOF. Recall the algebraic identity

$$\prod_{i=1}^{m} (A_i - B_i) = \sum_{Z} (-1)^{m-|Z|} \prod_{i \in Z} A_i \prod_{j \notin Z} B_j, \tag{4.8}$$

where the sum is over all subsets $Z \subseteq \{1, \ldots, m\}$. As a polynomial in $\gamma_1$, $\gamma_2, \ldots, u_1$, $u_2$, the tableau $T$ is a product of type (4.8) where the factor $A_i - B_i$ is the result of expanding the bracket in the $i$th row of $T$. Suppose this bracket is $[\gamma \quad \delta]$. Then

$$A_i = \gamma_1 \delta_2 \quad \text{and} \quad B_i = \gamma_2 \delta_1.$$

Hence, in (4.8),

$$\prod_{i \in Z} A_i = \left( \prod_{\gamma} \gamma_1^{p(\gamma)} \gamma_2^{q(\gamma)} \right) u_1^{p(u)} u_2^{q(u)},$$

where the second product ranges over all $\gamma$ in $\mathfrak{A}$, and for any Greek umbral letter $\gamma$ (and similarly for the Roman umbral letter $u$), $p(\gamma) = |E_1(\gamma, Z)|$ is the number of times $\gamma$ is the first entry in a row labelled by an integer in $Z$, and $q(\gamma) = |E_2(\gamma, Z)|$ is the number of times $\gamma$ is the second entry in a row labelled by an integer in $Z$. Similarly,

$$\prod_{j \notin Z} B_j = \left( \prod_{\gamma} \gamma_1^{r(\gamma)} \gamma_2^{s(\gamma)} \right) u_1^{r(u)} u_2^{s(u)},$$

where $r(\gamma) = |\overline{E}_1(\gamma, Z)|$ and $s(\gamma) = |\overline{E}_2(\gamma, Z)|$. Since for any umbral letter $\gamma$,

$$p(\gamma) + r(\gamma) = e_1(\gamma, Z) \quad \text{and} \quad q(\gamma) + s(\gamma) = e_2(\gamma, Z),$$

we have

$$\prod_{i \in Z} A_i \prod_{j \notin Z} B_j = \left( \prod_{\gamma \in \mathfrak{A}} \gamma_1^{e_1(\gamma, Z)} \gamma_2^{e_2(\gamma, Z)} \right) u_1^{e_1(u,Z)} u_2^{e_2(u,Z)}.$$

This concludes the proof of the lemma. □

Returning to the proof of the theorem, apply the umbral operator $U$ to both sides of (4.7) to obtain

$$\begin{aligned} \langle U | T \rangle &= \sum_{Z} (-1)^{m-|Z|} \langle U | M(Z) \rangle \\ &= \sum_{Z} (-1)^{m-|Z|} \left\langle U \middle| \left( \prod_{\gamma \in \mathfrak{A}} \gamma_1^{e_1(\gamma, Z)} \gamma_2^{e_2(\gamma, Z)} \right) u_1^{e_1(u,Z)} u_2^{e_2(u,Z)} \right\rangle. \end{aligned}$$

Using the definition and the multiplicative property of the umbral operator $U$, we obtain

$$\begin{aligned}\langle U|M(Z)\rangle &= \Big(\prod_{\gamma\in\mathcal{A}} A_{e_1(\gamma,Z)}\Big)x^{e_2(u,Z)}(-y)^{e_1(u,Z)}\\ &= A_{e_1(\alpha,Z)}A_{e_1(\beta,Z)}\cdots A_{e_1(\delta,Z)}x^{e_2(u,Z)}(-y)^{e_1(u,Z)},\end{aligned}$$

where $\alpha, \beta,\ldots,\omega$ are all the umbral letters in $\mathcal{A}$. To compute $\mathbf{h}(\langle U|M(Z)\rangle)$, we replace each coefficient $A_k$ by the corresponding symmetric function in the homogenized roots. This yields

$$\begin{aligned}\mathbf{h}(\langle U|M(Z)\rangle) = \frac{1}{(n!)^d}&\prod_\gamma(-1)^{e_2(\gamma,Z)}\\ &\times\Big[\sum_\pi \nu_{\pi(1)}\cdots\nu_{\pi(e_2(\alpha,z))}\mu_{\pi(e_2(\alpha,Z)+1)}\cdots\mu_{\pi(n)}\Big]\\ &\times\Big[\sum_\sigma \nu_{\sigma(1)}\cdots\nu_{\sigma(e_2(\beta,Z))}\mu_{\sigma(e_2(\beta,Z)+1)}\cdots\mu_{\sigma(n)}\Big]\\ &\cdots x^{e_2(u,Z)}(-y)^{e_1(u,Z)}.\end{aligned}$$

Bringing the sums out of the products and collecting the negative signs together, we obtain

$$\begin{aligned}\text{(4.9)}\quad \mathbf{h}(\langle U|M(Z)\rangle) &= (1/n!)^d(-1)^{e_2(\alpha,Z)+e_2(\beta,Z)+\cdots+e_2(\omega,Z)+e_1(u,Z)}\\ &\times\sum_{\pi,\sigma,\ldots}\nu_{\pi(1)}\cdots\nu_{\pi(e_2(\alpha,Z))}\mu_{\pi(e_2(\alpha,Z)+1)}\cdots\mu_{\pi(n)}\\ &\times\nu_{\sigma(1)}\cdots\nu_{\sigma(e_2(\beta,Z))}\mu_{\sigma(e_2(\beta,Z)+1)}\cdots\mu_{\sigma(n)}\cdots x^{e_2(u,Z)}y^{e_1(u,Z)},\end{aligned}$$

the sum ranging over all $d$-tuples $\pi, \sigma,\ldots$ of permutations of $\{1,\ldots,m\}$ indexed by the Greek umbral letters in $\mathcal{A}$.

LEMMA 4.2.

$$e_2(\alpha, Z) + e_2(\beta, Z) + \cdots + e_2(\omega, Z) + e_2(u, Z) = m,$$

*where $m$ is the number of brackets in $T$.*

PROOF. As the set $E_2(\gamma, Z)$ are pairwise disjoint, it suffices to prove

$$E_2(\alpha, Z) \cup E_2(\beta, Z) \cup \cdots \cup E_2(\omega, Z) \cup E_2(u, Z) = \{1,2,\ldots,m\}.$$

To show this, let $i$ be an integer in $\{1,\ldots,m\}$. If $i \in Z$, then $i \in E_2(\gamma, Z)$, where $\gamma$ is the second letter in the $i$th row. If $i \notin Z$, then $i \in E_2(\gamma', Z)$, where $\gamma'$ is the first letter in the $i$th row. Hence, $i$ is in the union $E_2(\alpha, Z) \cup \cdots \cup E_2(u, Z)$. □

Using the lemma, we conclude that

$$\begin{aligned}&e_2(\alpha, Z) + e_2(\beta, Z) + \cdots + e_2(\omega, Z) + e_1(u, Z)\\ &\qquad\equiv m - (e_1(u, Z) + e_2(u, Z)) \equiv m - p \pmod 2,\end{aligned}$$

where $p$ is the number of occurrences of $u$ in $T$. Observing that $m - p$ is simply the index $g$ of the covariant $\langle U \,|\, T \rangle$ (see §3.4), we obtain

$$(-1)^{e_2(\alpha,Z)+e_2(\beta,Z)+\cdots+e_2(\omega,Z)+e_1(u,Z)} = (-1)^g.$$

Substituting this into (4.9) and summing over all subsets $Z$ in $\{1,2,\ldots,m\}$, we obtain

$$\text{(4.10)} \quad \mathbf{h}(\langle U \,|\, T \rangle) = \frac{(-1)^g}{(n!)^d} \sum_Z (-1)^{m-|Z|} \times \sum_{\pi,\sigma,\ldots} \nu_{\pi(1)} \cdots \nu_{\pi(e_2(\alpha,Z))} \mu_{\pi(e_2(\alpha,Z))+1)} \cdots \mu_{\pi(n)} \cdots x^{e_2(u,Z)} y^{e_1(u,Z)}.$$

To finish the proof, we will show that the polynomial in (4.10) equals the polynomial

$$\text{(4.11)} \qquad \frac{(-1)^g}{(n!)^d} \sum_\Phi T^h[\Phi],$$

computed by Algorithm 4.1. To this end, consider the tableau $\hat{T}$ of integers constructed in Step 1 of Algorithm 4.1. For $Z \subseteq \{1,\ldots,m\}$, let $D_1(\gamma, Z)$ be the subset of $\{1,\ldots,n\}$ defined by

$$D_1(\gamma, Z) = \{j : j \text{ is the first entry in the } i\text{th row of } \hat{T} \text{ for some } i \in E_1(\gamma, Z)\};$$

Thus, $D_1(\gamma, Z)$ can be constructed by first listing all the rows of $\hat{T}$ whose row labels are in $E_1(\gamma, Z)$ and then extracting the first entry from each row. The sets $\bar{D}_1(\gamma, Z)$, $D_2(\gamma, Z)$, and $\bar{D}_2(\gamma, Z)$ are defined analogously. Note that

$$|D_1(\gamma, Z) \cup \bar{D}_1(\gamma, Z)| = e_1(\gamma, Z)$$

and

$$|D_2(\gamma, Z) \cup \bar{D}_2(\gamma, Z)| = e_2(\gamma, Z).$$

LEMMA 4.3. *Let $\Phi$ be a function from $\mathcal{Q}$ to the set of permutations on $\{1,\ldots,n\}$. As a polynomial in $\mu_i$, $\nu_i$, $x$ and $y$,*

$$\text{(4.12)} T^h[\Phi] = \sum_Z (-1)^{m-|Z|} \left[ \prod_{\gamma \in \mathcal{Q}} \prod_{i \in D_1(\gamma, Z) \cup \bar{D}_1(\gamma, Z)} \mu_{\Phi(\gamma,i)} \times \prod_{j \in D_2(\gamma, Z) \cup \bar{D}_2(\gamma, Z)} \nu_{\Phi(\gamma,j)} \right] x^{e_2(u,Z)} y^{e_1(u,Z)},$$

*and the sum ranges over all subsets $Z$ of $\{1,\ldots,m\}$.*

PROOF. As in Lemma 4.1, we use the algebraic identity (4.8). As a polynomial in $\mu_i$, $\nu_i$, $x$ and $y$, $T^h[\Phi]$ is a product

$$\prod_{i=1}^m (A_i - B_i) = \sum_Z (-1)^{m-|Z|} \prod_{i \in Z} A_i \prod_{j \notin Z} B_j,$$

where the factor $A_i - B_i$ is the result of substituting for the double bracket in the $i$th row of $T[\Phi]$ according to the rules (4.5). Suppose the double bracket in the $i$th row is $[\gamma \quad \delta \,|\, \Phi(\gamma, p)\, \Phi(\delta, q)]$. Then

$$A_i = \mu_{\Phi(\gamma,p)}\nu_{\Phi(\delta,q)} \quad \text{and} \quad B_i = \mu_{\Phi(\delta,q)}\nu_{\Phi(\gamma,q)}.$$

Hence, as in the proof of Lemma 4.1,

$$\prod_{i\in Z} A_i \prod_{j\notin Z} B_j = \left[ \prod_{\gamma\in\mathcal{A}} \prod_{i\in D_1(\gamma,Z)\cup\bar{D}_1(\gamma,Z)} \mu_{\Phi(\gamma,i)} \right.$$
$$\left. \times \prod_{j\in D_2(\gamma,Z)\cup\bar{D}_2(\gamma,Z)} \nu_{\Phi(\gamma,j)} \right] x^{e_2(u,Z)} y^{e_1(u,Z)}. \quad \square$$

Substituting (4.12) into (4.11) and changing the order of summation, we obtain

$$(4.13) \quad \frac{(-1)^g}{(n!)^d} \sum_{\Phi} T^h[\Phi] = \frac{(-1)^g}{(n!)^d} \sum_{Z} (-1)^{m-|Z|}$$
$$\times \sum_{\Phi} \left[ \prod_{,\gamma\in\mathcal{A}} \prod_{i\in D_1(\gamma,Z)\cup\bar{D}_1(\gamma,Z)} \mu_{\Phi(\gamma,i)} \right.$$
$$\left. \times \prod_{j\in D_2(\gamma,Z)\cup\bar{D}_2(\gamma,Z)} \nu_{\Phi(\gamma,j)} \right] x^{e_2(u,Z)} y^{e_1(u,Z)}.$$

It remains to show that the polynomials on the right-hand side of (4.10) and (4.13) are equal. Let $Z$ be a subset of $\{1,\ldots,m\}$ and let $\Psi$ be a function from $\mathcal{A}$ to the set $\Omega_n$ of permutations on $\{1,\ldots,n\}$ such that for all $\gamma \in \mathcal{A}$, $\Psi(\gamma)$ sends the subset

$$\begin{aligned}
&\{1,2,\ldots,|E_2(\gamma,Z)|\} \quad \text{to} \quad D_2(\gamma,Z),\\
&\{|E_2(\gamma,Z)|+1,\ldots,e_2(\gamma,Z)\} \quad \text{to} \quad \bar{D}_2(\gamma,Z),\\
&\{e_2(\gamma,Z)+1,\ldots,e_2(\gamma,Z)+|E_1(\gamma,Z)|\} \quad \text{to} \quad D_1(\gamma,Z),\\
&\{e_2(\gamma,Z)+|E_1(\gamma,Z)|+1,\ldots,n\} \quad \text{to} \quad \bar{D}_1(\gamma,Z).
\end{aligned}$$

If $\Phi\colon \mathcal{A} \to \Omega_n$, then let $\Phi'$ be the function given by $\Phi'(\gamma) = \Phi(\gamma) \circ \Psi(\gamma)$, the binary operation $\circ$ being composition of permutations. As $\Phi$ ranges over all functions from $\mathcal{A}$ to $\Omega_n$, so does $\Phi'$. Hence, the inner sum in (4.13) can be rewritten as

$$\sum_{\Phi} \left[ \prod_{\gamma\in\mathcal{A}} \nu_{\Phi(\gamma,1)} \cdots \nu_{\Phi(\gamma,e_2(\gamma,Z))} \mu_{\Phi(\gamma,e_2(\gamma,Z)+1)} \cdots \mu_{\Phi(n)} \right] x^{e_2(u,Z)} y^{e_1(u,Z)}.$$

On writing a function $\Phi\colon \mathcal{A} \to \Omega_n$ as a $d$-tuple $(\pi, \sigma, \ldots)$ where $\pi = \Phi(\alpha)$, $\sigma = \Phi(\beta), \ldots,$ we obtain the inner sum in (4.10).

This completes the proof of the theorem. □

Algorithm 4.1 can be extended without major changes to find the representation of a joint covariant $\langle U | T \rangle$ of several binary forms $f_1(x, y), \ldots, f_r(x, y)$ in terms of the homogenized roots of $f_1(x, y), \ldots, f_r(x, y)$.

Let $f_i(x, y)$ be of degree $n(i)$ and let $\mu_1^{(i)},\ldots,\mu_{n(i)}^{(i)}$, $\nu_1^{(i)},\ldots,\nu_{n(i)}^{(i)}$ be the homogenized roots of $f_i(x, y)$.

ALGORITHM 4.2. *Let $\mathcal{Q}_i$ be the set of Greek umbral letters belonging to the form $f_i(x, y)$ appearing in $T$ and let $d_i$ be the cardinality of $\mathcal{Q}_i$. We shall assume that every letter in $\mathcal{Q}_i$ occurs exactly $n(i)$ times. If not,* $\mathbf{h}(\langle U|T\rangle) = 0$ *and the algorithm terminates.*

*Step* 1. Construct the tableau $\hat{T}$ by repeating Step 1 in Algorithm 4.1 for each set $\mathcal{Q}_i$ of Greek umbral letters.

*Step* 2. Let $\Phi$ be a function from $\mathcal{Q}_1 \cup \mathcal{Q}_2 \cup \cdots \cup \mathcal{Q}_r$ to the set of permutations such that if $\alpha \in \mathcal{Q}_i$, then $\Phi(\alpha)$ is a permutation of the set $\{1, 2,\ldots, n(i)\}$. The double tableau $T[\Phi]$ is defined as in Algorithm 4.1.

*Step* 3. Let $T[\Phi]$ be as given in Algorithm 4.1. To each row in $T[\Phi]$, assign a polynomial in the homogenized roots according to the following two rules:

If $\alpha \in \mathcal{Q}_p$ and $\beta \in \mathcal{Q}_q$, then

$$(4.14) \qquad \begin{aligned} [\alpha \quad \beta | i \quad j] &\leftarrow \mu_i^{(p)}\nu_j^{(q)} - \nu_i^{(p)}\mu_j^{(q)}, \\ [\alpha \quad u | i \quad u] &\leftarrow \mu_i^{(p)}x - \nu_j^{(p)}y. \end{aligned}$$

The polynomial $T^h[\Phi]$ is obtained by taking the product of all the difference so obtained.

*Step* 4. Set

$$\mathbf{h}(\langle U|T\rangle) = \left((-1)^g \Big/ \prod_{i=1}^{r} (n(i)!)^{d_i}\right) \sum_{\Phi} T^h[\Phi],$$

where $g$ is the index of the covariant $\langle U|T\rangle$ and the sum ranges over all functions $\Phi$ of the type defined in Step 2.

For many applications it is more convenient to work with the roots rather than the homogenized roots. Algorithm 4.2 can be easily modified to obtain the representation of $\langle U|T\rangle$ in terms of the roots.

Let $f_i(x, y)$ be of degree $n(i)$. Let $\lambda_1^{(i)},\ldots,\lambda_{n(i)}^{(i)}$ be the roots of $f_i(x, 1)$, and let $A_{n(i)}$ be the leading coefficient of $f_i(x, y)$.

ALGORITHM 4.3. *To obtain the representation of $\langle U|T\rangle$ in terms of the roots, proceed as in Algorithm* 4.2 *with the following modifications.*

I. *In Step* 3 *use the following assignments instead of* (4.14): *if $\alpha \in \mathcal{Q}_p$ and $\beta \in \mathcal{Q}_q$,*

$$[\alpha \quad \beta | i \quad j] \leftarrow \lambda_j^{(q)} - \lambda_i^{(p)}, \qquad [\alpha \quad u | i \quad u] \leftarrow x - \lambda_i^{(p)}y$$

*The expressions $\lambda_j^{(p)} - \lambda_i^{(p)}$ and $x - \lambda_i^{(p)}y$ are also called* differences. *The polynomial $T^r[\Phi]$ is obtained by taking the product of all the differences so obtained.*

II. *In Step* 4 *set*

$$\mathbf{r}(\langle U|T\rangle) = (-1)^g \prod_{i=1}^{r} \left(\frac{A_{n(i)}}{n(i)!}\right)^{d_i} \sum_{\Phi} T^r[\Phi].$$

We end this section with an example.

EXAMPLE (TRANSVECTANTS). The $k$th *transvectant* of two binary forms $f(x, y)$ and $g(x, y)$ of degree $n$ and $m$, $n \geqslant m \geqslant k$, is the joint covariant defined umbrally by

$$\{f, g\}^k = \left\langle U(f, g) \,|\, [\alpha \quad \beta]^k [\alpha \quad u]^{n-k} [\beta \quad u]^{m-k} \right\rangle,$$

where $\alpha$ is an umbral letter of $f$ and $\beta$ an umbral letter of $g$. The first transvectant is the Jacobian of $f$ and $g$; the last, or $(n - m)$th transvectant is the apolar covariant of $f$ and $g$ (see §5).

Applying Algorithm 4.2, the expression of the $k$th transvectant in terms of the homogenized roots $\mu_1, \ldots, \mu_n$, $\nu_1, \ldots, \nu_n$ of $f(x, y)$ and $\xi_1, \ldots, \xi_m$, $\eta_1, \ldots, \eta_m$ of $g(x, y)$ is given by

$$\begin{aligned}\{f, g\}^k = (-1)^k \sum_{\pi, \sigma} &\left(\mu_{\pi(1)}\eta_{\sigma(1)} - \nu_{\pi(1)}\xi_{\sigma(1)}\right) \cdots \left(\mu_{\pi(k)}\eta_{\sigma(k)} - \nu_{\pi(k)}\xi_{\sigma(k)}\right) \\ &\times \left(\mu_{\pi(k+1)}x - \nu_{\pi(k+1)}y\right) \cdots \left(\mu_{\pi(n)}x - \nu_{\pi(n)}y\right) \\ &\times \left(\xi_{\sigma(k+1)}x - \eta_{\sigma(k+1)}y\right) \cdots \left(\xi_{\sigma(m)}x - \eta_{\sigma(m)}y\right).\end{aligned}$$

4.3 *Covariants and differences.* In the previous section, we showed that the symmetric functions of homogenized roots which represent covariants can be expressed as polynomials in the differences of homogenized roots. The converse is also true, provided that certain simple numerical constraints are satisfied.

Let $\mu_1, \ldots, \mu_n$, $\nu_1, \ldots, \nu_n$ be the homogenized roots of a binary form of degree $n$. Recall that a *difference* of homogenized roots is a polynomial of the form $\mu_i\nu_j - \mu_j\nu_i$ or $\mu_i x - \nu_i y$. A *difference monomial* $N$ is a product of differences. If $i$ is an integer in $\{1, 2, \ldots, n\}$, the *multiplicity* $m_i$ of $i$ in the difference monomial $N$ is the number of differences in $N$ containing the variable $\mu_i$ (which equals the number of differences in $N$ containing $\nu_i$). The *order* of $N$ is the number of differences containing the variable $x$. The *index* of $N$ is the number of differences *not* containing the variable $x$ or $y$. A difference monomial $N$ is said to be *regular of degree* $d$ if the multiplicities of $i$ are equal to $d$, that is, if $m_1 = m_2 = \cdots = m_n = d$. If

$$N = (\mu_i\nu_j - \mu_j\nu_i)(\mu_k\nu_l - \mu_l\nu_k) \cdots (\mu_p x - \nu_p y)(\mu_q x - \nu_q y) \cdots$$

and $\pi$ is a permutation of $\{1, \ldots, n\}$, the difference monomial $\pi(N)$ is defined by

$$\pi(N) = \left(\mu_{\pi(i)}\nu_{\pi(j)} - \mu_{\pi(j)}\nu_{\pi(i)}\right) \cdots \left(\mu_{\pi(p)}x - \nu_{\pi(p)}y\right) \cdots.$$

A *symmetric difference term* (*of index* $g$) is a polynomial of the form $\Sigma_\pi \pi(N)$, where $N$ is a *regular* difference monomial of index $g$ and the sum is over all permutations $\pi$ of $\{1, \ldots, n\}$. By definition, a symmetric difference term is a jointly symmetric in the variables $\mu_i$ and $\nu_i$.

THEOREM 4.2. *Let $R$ be a polynomial in the algebra of homogenized roots $k[\mu_1, \ldots, \mu_n, \nu_1, \ldots, \nu_n, x, y]$. Then $R$ is the representation in terms of homogenized roots of a covariant $I$ of index $g$ of binary forms of degree $n$ if and only if $R$ is a linear combination of symmetric difference terms, all of the same index $g$.*

Proof. Let $I$ be a covariant. We will show that $\mathbf{h}(I)$, the representation of $I$ in terms of homogenized roots, is a linear combination of symmetric difference terms. By linearity it suffices to prove the assertion in the case when $I = \langle U \mid T \rangle$, where $T$ is a bracket monomial in the umbral space. Let $\mathcal{Q}$ be the set of Greek umbral letters in $T$, $d$ the cardinality of $\mathcal{Q}$, and $g$ the index of $I$. Using Algorithm 4.1, we obtain

$$\mathbf{h}(\langle U \mid T \rangle) = \frac{(-1)^g}{(n!)^d} \sum_{\Phi} T^h[\Phi], \tag{4.15}$$

where the sum ranges over all functions $\Phi$ from $\mathcal{Q}$ to $\Omega_n$, the set of all permutations on $\{1,\dots,n\}$, that is, over all $d$-tuples $(\Phi(\alpha)$: $\alpha \in \mathcal{Q})$ in $\Omega_n \times \Omega_n \times \cdots \times \Omega_n$, the $d$-fold direct product of $\Omega_n$. Consider $\Omega_n \times \cdots \times \Omega_n$ as the $d$-fold direct product of the symmetric group $\Omega_n$. Let $\Delta$ be the subgroup consisting of all $d$-tuples of the form $(\pi, \pi, \dots)$ where $\pi \in \Omega_n$, and let $\mathcal{C}$ be a set of right coset representatives of $\Delta$ in $\Omega_n \times \cdots \times \Omega_n$. Then the sum (4.15) can be broken down into smaller sums:

$$\frac{(-1)^g}{(n!)^d} \sum_{(\sigma,\tau,\dots)\in\mathcal{C}} \sum_{(\pi,\pi,\dots)\in\Delta} T^h[\pi\sigma, \pi\tau, \dots],$$

where the outer sum ranges over all $d$-tuples of permutations in $\mathcal{C}$ and the inner sum ranges over all $d$-tuples of permutations in the subgroup $\Delta$. Now let $(\sigma, \tau, \dots)$ be an $d$-tuple of permutations in $\mathcal{C}$. The polynomial $T^h[\sigma, \tau, \dots]$ is, by construction, a regular difference monomial (of degree $d$, the number of distinct umbral letters in $T$). Hence, the inner sum can be rewritten as a symmetric difference term:

$$\sum_{\pi\in\Omega_n} T^h[\pi\sigma, \pi\tau, \dots] = \sum_{\pi\in\Omega_n} \pi\left(T^h[\sigma, \tau, \dots]\right).$$

Thus, $\mathbf{h}(\langle U \mid T \rangle)$ is a linear combination of symmetric difference terms. This proves the implication.

To prove the converse, it suffices to prove that a symmetric difference term $R$ is the representation in terms of homogenized roots of a covariant $I$. First, observe that when $R$ is expanded as a polynomial in the variable $\mu_i$, $\nu_i$, $x$ and $y$, every one of its monomials is regular of the same degree. Thus, $R$ is regular as a polynomial in $\mu_i$ and $\nu_i$. As $R$ is jointly symmetric in $\mu_i$ and $\nu_i$, by Proposition 4.1, there exists a polynomial $I(a_0,\dots,a_n, x, y)$ such that

$$R = I\left(a_0(\mu_i, \nu_j), \dots, a_n(\mu_i, \nu_j), x, y\right).$$

It remains to show that $I(A_0,\dots,A_n, X, Y)$ is covariant.

Lemma 4.4. *Let $(c, d)$ be a linear change of variables from $x$ and $y$ to $\bar{x}$ and $\bar{y}$. Then*

$$\bar{\mu}_i\bar{\nu}_j - \bar{\mu}_j\bar{\nu}_i = [c \quad d](\mu_i\nu_j - \mu_j\nu_i),$$

$$\bar{\mu}_i\bar{x} - \bar{\nu}_i\bar{y} = \mu_i x - \nu_i y.$$

Proof. Under the linear change of variables $(c, d)$, the linear form $\mu_i x - \nu_i y$ is transformed into $\bar{\mu}_i\bar{x} - \bar{\nu}_i\bar{y}$. This proves the second identity. Equating coefficients of $\bar{x}$ and $\bar{y}$, we have

$$\bar{\mu}_i = c_2\mu_i + c_1\nu_i, \qquad \bar{\nu}_i = -d_2\mu_i - d_1\nu_i.$$

Observing that the difference $\mu_i\nu_j - \mu_j\nu_i$ can be rewritten as a determinant

$$\mu_i\nu_j - \mu_j\nu_i = \det\begin{pmatrix} \mu_i & \mu_j \\ \nu_i & \nu_j \end{pmatrix},$$

we conclude that

$$\bar{\mu}_i\bar{\nu}_j - \bar{\mu}_j\bar{\nu}_i = \det\begin{pmatrix} c_2 & c_1 \\ -d_2 & -d_1 \end{pmatrix} \det\begin{pmatrix} \mu_i & \mu_j \\ \nu_i & \nu_j \end{pmatrix} = [c \quad d](\mu_i\nu_j - \mu_j\nu_i). \qquad \square$$

To finish the proof, let $g$ be the number of differences of the type $\mu_i\nu_j - \mu_j\nu_i$ in a monomial in $R$. Using the lemma, we infer that

$$R(\bar{\mu}_i, \bar{\nu}_j, \bar{x}, \bar{y}) = [c \quad d]^g R(\mu_i, \nu_j, x, y).$$

Thus, for any binary form $f(x, y) = \sum_{k=0}^{n}\binom{n}{k}a_k x^k y^{n-k}$ with homogenized roots $\mu_i$, $\nu_i$, and any change of variables $(c, d)$,

$$\begin{aligned} I(\bar{a}_0,\ldots,\bar{a}_n, \bar{x}, \bar{y}) &= I\big(a_0(\bar{\mu}_i, \bar{\nu}_j),\ldots,a_n(\bar{\mu}_i, \bar{\nu}_j), \bar{x}, \bar{y}\big) \\ &= R(\bar{\mu}_i, \bar{\nu}_j, \bar{x}, \bar{y}) = [c \quad d]^g R(\mu_i, \nu_j, x, y) \\ &= [c \quad d]^g I(a_0,\ldots,a_n, x, y). \end{aligned}$$

Hence, $I(A_0,\ldots,A_n, X, Y)$ is a covariant of binary forms of degree $n$. □

Symmetric difference terms can be written more elegantly by using bracket notation. Briefly, set

$$[i \quad j] = \mu_i\nu_j - \mu_j\nu_i, \qquad [i \quad u] = \mu_i x - \mu_i y$$

and let $\mathcal{V}$ be the subalgebra generated by these brackets in the algebra of homogeneous roots. Define the *symmetrization operator* $S$ on $\mathcal{V}$ by setting

$$\begin{aligned} &\langle S | [i \quad j][k \quad l]\cdots[p \quad u]\rangle \\ &\quad = \sum_{\pi}[\pi(i) \quad \pi(j)][\pi(k) \quad \pi(l)]\cdots[\pi(p) \quad u] \end{aligned}$$

on bracket monomials and extending by linearity. Thus, in this notation, a symmetric difference term is the image of a bracket monomial under the symmetrization operator.

This description yields another representation of the space $\mathcal{I}[n, d, t]$ of covariants of degree $d$ and order $t$ of binary forms of degree $n$.

Proposition 4.2. *Let $\mathcal{V}^S[n, d, t]$ be the space of all symmetrized bracket monomials formed with the n integers $\{1, 2,\ldots,n\}$ each occurring d times and the Roman letter u occurring t times. Then*

$$\mathcal{I}[n, d, t] \simeq \mathcal{V}^S[n, d, t],$$

*the isomorphism being given by restricting the homomorphism* $\mathbf{h}$ *to* $\mathcal{I}[n, d, t]$.

Proof. Observe that if $I$ is a covariant and $\mathbf{h}(I)$ is zero, then $I$ must be identically zero. □

By Theorem 4.2 the symmetrization of a difference monomial $M$ represents a covariant if and only if $M$ is regular. This condition, stated in terms of brackets, yields

PROPOSITION 4.3. *Let $M$ be a nonconstant bracket monomial in $\mathcal{V}$. Then $\langle S\,|\,M\rangle$ is the representation in terms of homogeneous roots of a covariant $I$ of degree $d$ and order $t$ of binary forms of degree $n$ if and only if the following conditions are satisfied: Let*

$$m_{ij} = \textit{number of occurrences of the bracket } [i \quad j] \textit{ or } [j \quad i] \textit{ in } M,$$

$$t_i = \textit{number of occurrences of the bracket } [i \quad u] \textit{ or } [u \quad i] \textit{ in } M.$$

*Then*:

A. *For all $i$ and $j$, $m_{ij} = m_{ji}$ and $m_{ii} = 0$.*

B. *For all $i$,*

$$t_i + m_{i1} + m_{i2} + m_{i3} + \cdots + m_{in} = d.$$

C. $t_1 + t_2 + \cdots + t_n = t$.

D. *The sum of all the $t_i$'s and $m_{ij}$'s is even: that is, there exists a positive integer $h$ such that*

$$\sum_{i=1}^{n} t_i + \sum_{i,j=1}^{n} m_{ij} = 2h.$$

*Conversely, any covariant can be represented in terms of homogenized roots as a linear combination of symmetric difference terms $\langle S\,|\,M\rangle$, where $M$ satisfies the above conditions.*

PROOF. This result follows readily from Theorem 4.2, on observing that:

A. As $M$ is nonzero, the number $m_{ii}$ of occurrences of the bracket $[i \quad i]$ must be zero. Furthermore, as the definition of $m_{ij}$ is symmetric in $i$ and $j$, $m_{ij} = m_{ji}$.

B. $t_i + m_{i1} + m_{i2} + \cdots + m_{in}$ is the number of occurrences of the variable $\mu_i$ (or the number of occurrences of $\nu_i$) in the bracket monomial $M$. By regularity, these numbers are all equal, say, to $d$. When $\langle S\,|\,M\rangle$ is written as a polynomial $I$ in the homogenized symmetric functions $a_0(\mu_i, \nu_j), \dots, a_n(\mu_i, \nu_j)$, each variable $\mu_i$ belongs to exactly one homogenized symmetric function $a_k(\mu_i, \nu_j)$. Hence, $d$ equals the total degree of the variables $A_0, \dots, A_n$ in $I$.

C. $t_1 + t_2 + \cdots + t_n$ equals the total order of the variables $x$ and $y$ in $\langle S\,|\,M\rangle$ and, hence, equals the order of $I$.

D. The sum of all the $t_i$'s and $m_{ij}$'s equals the total number of integers from $\{1, 2, \dots, n\}$ or letters $u$ (counted according to their multiplicity) occurring in the bracket monomial $M$. As each bracket contains two entries, this number equals $2h$, where $h$ is the number of brackets in $M$. □

EXAMPLE. Let $\langle S\,|\,M\rangle$ be a symmetric difference term representing an *in*variant of the binary cubic. Then the nonzero entries $m_{ij}$, $i \neq j$, $i, j = 1, 2, 3$, satisfy the following diophantine equations:

$$m_{ij} = m_{ji},$$

$$m_{12} + m_{13} = m_{21} + m_{23} = m_{31} + m_{32} = d,$$

$$m_{12} + m_{13} + m_{21} + m_{23} + m_{31} + m_{32} = 2h.$$

Solving these equations, we obtain

$$m_{12} = m_{13} = m_{23}.$$

Thus,

$$M = ([1 \quad 2][1 \quad 3][2 \quad 3])^k$$

for some positive integer $k$. As

$$[\pi(1) \quad \pi(2)][\pi(1) \quad \pi(3)][\pi(2) \quad \pi(3)] = \operatorname{sgn}\pi[1 \quad 2][1 \quad 3][2 \quad 3],$$

$\langle S \mid M \rangle$ equals zero if $k$ is odd. Thus,

$$\langle S \mid M \rangle = ([1 \quad 2][1 \quad 3][2 \quad 3])^k,$$

where $k$ is even, are the only nonzero symmetric difference terms representing invariants. For $k = 2$,

$$([1 \quad 2][1 \quad 3][2 \quad 3])^2 = ((\mu_1\nu_2 - \mu_2\nu_1)(\mu_1\nu_3 - \mu_3\nu_1)(\mu_2\nu_3 - \mu_3\nu_2))^2,$$

a constant multiple of the discriminant of the cubic. We conclude that every nonzero invariant of the binary cubic is a constant multiple of a power of the discriminant. □

By suitably extending the notion of regularity, the results in this section generalize to several binary forms. Let $f_1(x, y), \ldots, f_r(x, y)$ be binary forms of degrees $n(1), \ldots, n(r)$, and let $\mu_i^{(k)}$, $\nu_i^{(k)}$, $i = 1, 2, \ldots, n(k)$, be the homogenized roots of the $k$th binary form $f_k(x, y)$. A *difference* is a polynomial in the algebra $k[\mu_i^{(k)}, \nu_i^{(k)}, x, y]$ of homogenized roots of the form

$$\mu_i^{(k)}\nu_j^{(l)} - \mu_j^{(l)}\nu_i^{(k)} \quad \text{or} \quad \mu_i^{(k)}x - \nu_i^{(k)}y,$$

and a *difference monomial* is a product of differences. The *multiplicity* $m_i^{(k)}$ *of* $i$ *relative to the kth binary form* in a difference monomial $N$ is the number of differences in $N$ containing the variable $\mu_i^{(k)}$. A difference monomial $N$ is said to be *regular of degree* $(d_1, \ldots, d_r)$ if the multiplicities of $i$ relative to the $k$th binary form are all equal to $d_k$, that is, if $m_1^{(k)} = m_2^{(k)} = \cdots = m_{n(k)}^{(k)} = d_k$. With this definition of regularity, both Theorem 4.2 and Proposition 4.3 extend readily to several binary forms.

4.4. *Hermite's reciprocity law.* The bracket notation for differences, combined with the umbral notation, yields a transparent proof of Hermite's reciprocity law.

THEOREM 4.3 (HERMITE'S RECIPROCITY LAW). *Let $c(n, d, t)$ be the dimension of the vector space of covariants of degree d and order t of binary forms of degree n. Then*

$$c(n, d, t) = c(d, n, t).$$

PROOF. Representing the covariants of degree $d$ and order $t$ of binary forms of degree $n$ umbrally, we have, by Theorem 3.4,

$$c(n, d, t) = \dim \mathfrak{U}^S[n, d, t],$$

where $\mathcal{U}^S[n, d, t]$ is the vector space spanned by all symmetrized bracket monomials formed with $d$ distinct Greek umbral letters $\alpha, \beta, \ldots, \delta$ each occurring $n$ times and the Roman letter $u$ occurring $t$ times. On the other hand, representing the covariants by homogenized roots, we have, by Proposition 4.2,

$$c(d, n, t) = \dim \mathcal{V}^S[d, n, t],$$

where $\mathcal{V}^S[d, n, t]$ is the vector space spanned by all symmetrized brackets monomials formed with $d$ integers $\{1, 2, \ldots, d\}$ each occurring $n$ times and the Roman letter $u$ occurring $t$ times. These two vector spaces are isomorphic on identifying the first umbral letter $\alpha$ with the integer 1, the second letter $\beta$ with $2, \ldots,$ and the $d$th letter $\delta$ with $d$. □

## 5. Apolarity.

5.1 *The apolar covariant.* By making suitable changes of variables, a binary form may sometimes be brought to a simpler form. For example, a binary form may be written as the $n$th power of a linear form, or it may be written as the sum of $k$, but no fewer than $k$, powers of linear forms. Such properties of a binary form are independent of the choice of coordinates, and we therefore expect them to be expressible by the vanishing or nonvanishing of covariants. We shall now see how such covariants may be constructed.

Consider two binary forms

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k} = \left\langle U(f, g) \mid [\alpha \quad u]^n \right\rangle,$$

$$g(x, y) = \sum_{k=0}^{m} \binom{m}{k} b_k x^k y^{m-k} = \left\langle U(f, g) \mid [\beta \quad u]^m \right\rangle,$$

where $f(x, y)$ is of degree $n$, $g(x, y)$ is of degree $m$, $m \leqslant n$, $\alpha$ is an umbral letter for $f$, and $\beta$ is an umbral letter for $g$. Their apolar covariant $\{f, g\}$ is the binary form of degree $n - m$ defined umbrally by

$$\{f, g\} = \left\langle U(f, g) \mid [\alpha \quad \beta]^m [\alpha \quad u]^{n-m} \right\rangle.$$

LEMMA 5.1. *Let $\mathcal{F}_n$ be the vector space of all binary forms of degree $n$. Then the apolar covariant $\{f, g\}$ is a bilinear mapping from $\mathcal{F}_n \times \mathcal{F}_m$ to $\mathcal{F}_{n-m}$ which is jointly covariant in $f$ and $g$. Conversely, any jointly covariant bilinear mapping from $\mathcal{F}_n \times \mathcal{F}_m$ to $\mathcal{F}_{n-m}$ is a constant multiple of $\{f, g\}$.*

PROOF. The lemma follows from two easy observations: (a) A joint covariant $I$ of $f$ and $g$ is bilinear in $f$ and $g$ if and only if $I$ can be represented umbrally by a bracket polynomial in which every bracket monomial contains exactly one umbral letter belonging to $f$ and exactly one umbral letter belonging to $g$. (b) The only standard tableau with $n$ occurrences of $\alpha$, $m$ occurrences of $\beta$, and $n - m$ occurrences of $u$, where $\alpha < \beta < u$, is $[\alpha \quad \beta]^m [\alpha \quad u]^{n-m}$. □

In the special case when $n = m$, the apolar covariant $\{f, g\}$ has the explicit expression

$$\{f, g\} = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} a_k b_{n-k}.$$

This is a scalar and is called the *lineo-linear invariant*. In general, the apolar covariant is given explicitly by

$$(5.1)\quad \{f, g\} = \sum_{l=0}^{n-m} \binom{n-m}{l} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} b_{m-k} a_{k+l} x^l y^{m-n-l}.$$

Let

$$f(x, y) = a(\mu_1 x - \nu_1 y)(\mu_2 x - \nu_2 y) \cdots (\mu_n x - \nu_n y),$$
$$g(x, y) = b(\xi_1 x - \eta_1 y)(\xi_2 x - \eta_2 y) \cdots (\xi_m x - \eta_m y)$$

be factorizations of $f(x, y)$ and $g(x, y)$ into linear forms. We say that two linear forms $\mu_i x - \nu_i y$ and $\mu_j x - \nu_j y$ in the factorization of $f(x, y)$ are *distinct* if for all constants $c$, $\mu_i x - \nu_i y \neq c(\mu_j x - \nu_j y)$. From Algorithm 4.2, we immediately obtain the expression of the apolar covariant $\{f, g\}$ in terms of the coefficients $\mu_i$, $\nu_i$ and $\xi_i$, $\eta_i$ of the linear forms occurring in the above factorizations, namely,

$$(5.2)\ \{f, g\} = \frac{(-1)^{n-m} ab}{m!n!} \sum_{\pi, \sigma} \left(\mu_{\pi(1)} \eta_{\sigma(1)} - \nu_{\pi(1)} \xi_{\sigma(1)}\right) \cdots \left(\mu_{\pi(m)} \eta_{\sigma(m)} - \nu_{\pi(m)} \xi_{\sigma(m)}\right) \times \left(\mu_{\pi(m+1)} x - \nu_{\pi(m+1)} y\right) \cdots \left(\mu_{\pi(n)} x - \nu_{\pi(n)} y\right),$$

where the sum ranges over all permutations $\pi$ of $\{1, 2, \ldots, n\}$ and $\sigma$ of $\{1, 2, \ldots, m\}$.

A useful fact about the apolar covariant is

LEMMA 5.2. *Let $f(x, y)$ be a form of degree $n$, $g(x, y)$ a form of degree $m_1$, $h(x, y)$ a form of degree $m_2$, with $m_1 + m_2 \leqslant n$. Then*

$$\{f, gh\} = \{\{f, g\}, h\}.$$

The *proof* is an easy computation using the umbral representation of $\{f, g\}$.

As a corollary, we obtain

COROLLARY 5.1. *Under the same hypotheses as Lemma 5.2, if $\{f, g\} = 0$, then $\{f, gh\} = 0$.*

5.2 *Forms apolar to a given form.* Two binary forms $f(x, y)$ and $g(x, y)$ are said to be *apolar* if their apolar covariant $\{f, g\}$ is the identically zero form. The basic question about apolarity which we will study in this section is: Given a form of degree $s$ and a positive integer $t$, what is the dimension of the vector space of all forms of degree $t$ apolar to the given form? The answer turns out to be different depending on whether (A) $t \geqslant s$ or (B) $t < s$.

The answer for case (A) is given by

PROPOSITION 5.1. *Let $g(x, y)$ be a nonzero form of degree $m$, and let $n$ be a positive integer such that $n \geqslant m$. Then the dimension of the vector space of all forms of degree $n$ apolar to $g(x, y)$ equals $m$. More explicitly, if*

$$g(x, y) = a(\mu_1 x - \nu_1 y)^{m_1} (\mu_2 x - \nu_2 y)^{m_2} \cdots (\mu_p x - \nu_p y)^{m_p}$$

*is a factorization of $g(x, y)$ into distinct linear forms, then the binary forms*

$$(\mu_i x - \nu_i y)^{n-m_i+1} x^j y^{m_i-j-1}, \qquad i = 1,\dots,p, \quad j = 0, 1,\dots,m_i - 1,$$

*form a basis for the vector space of all forms of degree n apolar to $g(x, y)$.*

PROOF. Let

$$g(x, y) = \sum_{k=0}^{m} \binom{m}{k} b_k x^k y^{m-k}.$$

Then the condition $\{f, g\} \equiv 0$ yields $n - m + 1$ linear equations which have to be satisfied by the coefficients $a_k$ of $f(x, y)$ if $f$ is apolar to $g$:

(5.3)

$$\begin{aligned}
&\binom{m}{0} b_m a_0 - \binom{m}{1} b_{m-1} a_1 + \binom{m}{2} b_{m-2} a_2 - \cdots \pm \binom{m}{m} b_0 a_m &&= 0\\
&\qquad \binom{m}{0} b_m a_1 - \binom{m}{1} b_{m-1} a_2 + \quad \cdots \quad \pm \binom{m}{m} b_0 a_{m+1} &&= 0\\
&\qquad \vdots\\
&\qquad \binom{m}{0} b_m a_{n-m} - \binom{m}{1} b_{m-1} a_{n-m+1} + \cdots \quad \pm \binom{m}{m} b_0 a_n &&= 0.
\end{aligned}$$

As $g(x, y)$ is nonzero, these linear equations are linearly independent and hence determine a subspace of the vector space of all binary forms of degree $n$ of dimension exactly $(n + 1) - (n - m + 1) = m$.

To prove the second assertion let $\mu_i x - \nu_i y$ be a linear form occurring with multiplicity $m_i$ in the factorization of $g(x, y)$. Let

$$h(x, y) = (\mu_i x - \nu_i y)^{n-m_i+1} x^j y^{m_i-j-1}. \tag{5.4}$$

Using (5.2), we have

$$\begin{aligned}
\{h, g\} = \frac{(-1)^{n-m} a}{m!n!} \sum_{\pi, \sigma} &\left(\kappa_{\pi(1)}\eta_{\sigma(1)} - \lambda_{\pi(1)}\xi_{\sigma(1)}\right)\\
&\cdots \left(\kappa_{\pi(m)}\eta_{\sigma(m)} - \lambda_{\pi(m)}\xi_{\sigma(m)}\right)\\
&\times \left(\kappa_{\pi(m+1)}x - \lambda_{\pi(m+1)}y\right) \cdots \left(\kappa_{\pi(n)}x - \lambda_{\pi(n)}y\right),
\end{aligned}$$

where $\kappa_i$, $\lambda_i$ are the coefficients of the linear forms in the factorization (5.3) of $h(x, y)$ and $\xi_i$, $\eta_i$ are the coefficients of the linear forms in the factorization of $g(x, y)$ given in the statement of the proposition. Observe that the coefficients $\mu_i$, $\nu_i$ occur with multiplicity $n - m_i + 1$ among the $\kappa_i$, $\lambda_i$, and with multiplicity $m_i$ among the $\xi_i$, $\eta_i$. As $m_i + (n - m_i + 1) = n + 1 > n$, and there are exactly $n$ factors in each summand in the above sum, there must, by the pigeonhole principle, be a factor of the form $(\mu_i\nu_i - \nu_i\mu_i)$ in each summand. Hence $\{h, g\} \equiv 0$ and $h(x, y)$ is apolar to $g(x, y)$.

Finally, consider the forms

$$(\mu_i x - \nu_i y)^{n-m_i+1} x^j y^{m_i-m-1}, \qquad i = 1,\dots,p, \quad j = 0, 1,\dots,m_i - 1,$$

where $(\mu_1 x - \nu_1 y),\dots,(\mu_p x - \nu_p y)$ are all the distinct linear forms occurring with multiplicity $m_1,\dots,m_p$ in a factorization of $g(x, y)$ into linear forms. There are $m_1 + m_2 + \cdots + m_p = m$ such forms and they are all apolar to $g(x, y)$. To finish the proof it remains to observe that they are linearly independent. The proof of this fact is elementary. □

The answer for case (B) cannot, in general, be given explicitly. However, the following partial answer often suffices.

PROPOSITION 5.2. *Let $f(x, y)$ be a binary form of degree $n$ and let $m$ be a positive integer such that $m \leqslant n$. Then the subspace of all binary forms of degree $m$ apolar to $f(x, y)$ has dimension greater than or equal to $2m - n$.*

PROOF. Given

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k},$$

the condition $\{f, g\} \equiv 0$ yields $n - m + 1$ linear equations on the coefficients $b_i$ of any form $g(x, y)$ of degree $m$ apolar to $f(x, y)$:

$$\sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} a_{k+l} b_{m-k} = 0, \qquad l = 0, 1, \dots, n - m. \tag{5.5}$$

These linear equations may be dependent. Thus, the dimension of the subspace of all binary forms of degree $m$ apolar to $f(x, y)$ has dimension at least $m + 1 - (n - m + 1) = 2m - n$. □

Analyzing the proof of the previous proposition, we obtain a somewhat more useful result.

COROLLARY 5.2. *Under the same hypotheses as Proposition 5.2, the space of binary forms of degree $m$ apolar to $f(x, y)$ has dimension $m - r + 1$, where $r$ is the rank of the system* (5.5) *of linear equations.*

5.3 *Sylvester's theorem.* We now have all the tools in hand to prove Sylvester's theorem on canonical forms for binary forms of odd degree. Sylvester's theorem states that, in general, a binary form of odd degree $n = 2j + 1$ can be written as a linear combination of $j + 1$ $n$th powers of linear forms. Thus, it gives "generically" a canonical form for binary forms of odd degree.

THEOREM 5.1 (SYLVESTER). *Let $f(x, y)$ be a binary form of odd degree $n = 2j + 1$. Then there exists a nonzero form $g(x, y)$ of degree $m = j + 1$ apolar to $f(x, y)$. If, in addition, there exists one such form $g(x, y)$ with $m$ distinct linear factors $\mu_1 x - \nu_1 y,\dots,\mu_m x - \nu_m y$, then there exist scalars $c_i$ such that*

$$f(x, y) = \sum_{i=1}^{m} c_i (\mu_i x - \nu_i y)^n.$$

PROOF. By Proposition 5.2 the dimension of the space of all forms of degree $m$ apolar to $f(x, y)$ is at least

$$2m - n = 2(j + 1) - (2j + 1) = 1$$

and is thus nonzero. If there exists one such form $g(x, y)$ with $m$ distinct linear factors, then the second assertion follows from Proposition 5.1. □

By using the full power of Proposition 5.2, we can sharpen Sylvester's theorem so as to omit the qualification "generically".

THEOREM 5.2. *Let $f(x, y)$ be a binary form of odd degree $n = 2j + 1$ and let $g(x, y)$ be a nonzero form of degree $m = j + 1$ apolar to $f(x, y)$. If*

$$g(x, y) = (\mu_1 x - \nu_1 y)^{m_1}(\mu_2 x - \nu_2 y)^{m_2} \cdots (\mu_p x - \nu_p y)^{m_p}$$

*is a factorization of $g(x, y)$ into $p$ distinct linear forms, then there exist forms $h_i(x, y)$ of degree $m_i - 1$ such that*

$$f(x, y) = \sum_{i=1}^{p} h_i(x, y)(\mu_i x - \nu_i y)^{n-m_i+1}.$$

When the system (5.5) of linear equations is of rank $m$, then the nonzero form $g(x, y)$ in Sylvester's theorem is determined up to a constant multiple. When this is the case, the form $g(x, y)$ is in fact a covariant, classically denoted by $J$, of binary forms of degree $n$. Our next result provides an explicit umbral representation for $J$.

LEMMA 5.3. *Let $f(x, y) = \sum_{k=0}^{n}\binom{n}{k} a_k x^k y^{n-k}$ be a binary form of odd degree $n = 2j + 1$ and $m = j + 1$. Let $J$ be the covariant given umbrally by*

$$J = \left\langle U| \prod_{\delta<\varepsilon} [\delta \quad \varepsilon]^2 \prod_{\delta} [\delta \quad u] \right\rangle,$$

*where $\{\alpha, \beta, \dots, \omega\}$ is a set of $m$ linearly ordered umbral letters of $f(x, y)$, the first product is over all pairs $(\delta, \varepsilon)$ of umbral letters such that $\delta < \varepsilon$, and the second product is over all letters $\delta$. Then $J(a_0, a_1, \dots, a_n, x, y)$ is apolar to $f(x, y)$, and if $J(a_0, \dots, a_n, x, y) \not\equiv 0$, every form of degree $m$ apolar to $f(x, y)$ is a constant multiple of $J(a_0, \dots, a_n, x, y)$.*

PROOF. Let $g(x, y) = \sum_{k=0}^{m}\binom{m}{k} b_k x^k y^{m-k}$ be a form of degree $m$ apolar to $f(x, y)$. Then the coefficients $b_0, b_1, \dots, b_m$ of $g(x, y)$ satisfy the following system of linear equations (see Proposition 5.2):

(5.6)

$$\begin{array}{llllll}
(-1)^m a_0 b_m & + (-1)^{m-1}\binom{m}{1} a_1 b_{m-1} & + (-1)^{m-2}\binom{m}{2} a_2 b_{m-2} & + \cdots + & a_m b_0 & = 0 \\
(-1)^m a_1 b_m & + (-1)^{m-1}\binom{m}{1} a_2 b_{m-1} & + (-1)^{m-2}\binom{m}{2} a_3 b_{m-2} & + \cdots + & a_{m+1} b_0 & = 0 \\
\vdots & & & & & \\
(-1)^m a_{m-1} b_m & + (-1)^{m-1}\binom{m}{1} a_m b_{m-1} & + \cdots & & + a_{2m-1} b_0 & = 0.
\end{array}$$

Solving this for $b_0, \ldots, b_m$ using Cramer's rule and regrouping the terms into a determinant, we obtain

$$(5.7)\quad g(x, y) = (-1)^m \begin{vmatrix} a_0 & -\binom{m}{1}a_1 & \binom{m}{2}a_2 & \cdots & a_{m-1} \\ a_1 & -\binom{m}{1}a_2 & \binom{m}{2}a_3 & & a_m \\ a_2 & -\binom{m}{1}a_3 & \binom{m}{2}a_4 & & a_{m+1} \\ \vdots & & & & \\ a_{m-1} & -\binom{m}{1}a_m & \cdots & & \\ x^m & \binom{m}{1}x^{m-1}y & \binom{m}{2}x^{m-2}y^2 & \cdots & y^m \end{vmatrix}$$

We can express $g(x, y)$ umbrally by

$$g(x, y) = \left[ \prod_{k=1}^{n} (-1)^{m-k} \binom{m}{k} \right]$$

$$\times \left\langle U(f) \,\middle|\, \begin{vmatrix} \alpha_2^n & \alpha_1\alpha_2^{n-1} & \alpha_1^2\alpha_2^{n-2} & \cdots & \alpha_1^m\alpha_2^{m-1} \\ \beta_1\beta_2^{n-1} & \beta_1^2\beta_2^{n-2} & \beta_1^3\beta_2^{n-3} & & \beta_1^{m+1}\beta_2^{m-2} \\ \gamma_1^2\gamma_2^{n-2} & \gamma_1^3\gamma_2^{n-3} & \cdots & & \\ \vdots & & & & \\ \omega_1^{m-1}\omega_2^m & \omega_1^m\omega_2^{m-1} & \omega_1^{m+1}\omega_2^{m-2} & & \\ u_2^m & u_2^{m-1}u_1 & u_2^{m-2}u_1^2 & \cdots & u_1^m \end{vmatrix} \right\rangle$$

where $\alpha, \beta, \ldots, \omega$ are $m$ Greek umbral letters of $f(x, y)$ linearly ordered in such a way that $\alpha < \beta < \cdots < \omega$. Let $\Delta$ denote the determinant inside the umbral expression. Factoring out the first entry in each row, we obtain

$$\Delta = \alpha_2^n\beta_1\beta_2^{n-1}\gamma_1^2\gamma_2^{n-2} \cdots \omega_1^{m-1}\omega_2^m u_2^m$$

$$\times \begin{vmatrix} 1 & \alpha_1/\alpha_2 & (\alpha_1/\alpha_2)^2 & \cdots & (\alpha_1/\alpha_2)^{m-1} \\ 1 & \beta_1/\beta_2 & (\beta_1/\beta_2)^2 & & (\beta_1/\beta_2)^{m-1} \\ 1 & \gamma_1/\gamma_2 & \cdots & & \\ \vdots & & & & \\ 1 & \omega_1/\omega_2 & \cdots & & (\omega_1/\omega_2)^{m-1} \\ 1 & u_1/u_2 & \cdots & & (u_1/u_2)^{m-1} \end{vmatrix}.$$

Apart from a factor, the determinant Δ is a Vandermonde determinant and we obtain

$$\Delta = \alpha_2^n\beta_1\beta_2^{n-1}\gamma_1^2\gamma_2^{n-2}\cdots\omega_1^{m-1}\omega_2^m u_2^m \prod_{\delta<\varepsilon}\left(\frac{\delta_1}{\delta_2}-\frac{\varepsilon_1}{\varepsilon_2}\right)\prod_{\delta}\left(\frac{\delta_1}{\delta_2}-\frac{u_1}{u_2}\right).$$

Since

$$\frac{\delta_1}{\delta_2}-\frac{\varepsilon_1}{\varepsilon_2}=\frac{\delta_1\varepsilon_2-\delta_2\varepsilon_1}{\delta_2\varepsilon_2}=\frac{[\delta\quad\varepsilon]}{\delta_2\varepsilon_2},$$

and there are $m$ terms in the products containing a given Greek umbral letter, we obtain

$$\Delta = \alpha_2^{m-1}\beta_1\beta_2^{m-2}\gamma_1^2\gamma_2^{m-3}\cdots\omega_1^{m-1}\prod_{\delta<\varepsilon}[\delta\quad\varepsilon]\prod_{\delta}[\delta\quad u].$$

The determinant Δ is not yet a bracket polynomial, but on symmetrizing the letters (see §3.3, although full details will be be given here), we will obtain a bracket monomial.

Let $\pi$ be a permutation of the set $\{\alpha, \beta, \dots, \omega\}$ of Greek umbral letters. The permutation $\pi$ acts on Δ by permuting the letters in Δ, that is,

$$\pi(\Delta) = \pi(\alpha)_2^{m-1}\pi(\beta)_1\pi(\beta)_2^{m-2}\cdots\pi(\omega)_1^{m-1}\prod_{\delta<\varepsilon}[\pi(\delta)\quad\pi(\varepsilon)]\prod_{\delta}[\pi(\delta)\quad u].$$

As the product

$$\prod_{\delta<\varepsilon}[\delta\quad\varepsilon]\prod_{\delta}[\delta\quad u]$$

is alternating, we have

$$\pi(\Delta) = \operatorname{sgn}(\pi)\pi(\alpha)_2^{m-1}\pi(\beta)_1\pi(\beta)_2^{m-2}\cdots\pi(\omega)_1^{m-1}\prod_{\delta<\varepsilon}[\delta\quad\varepsilon]\prod_{\delta}[\delta\quad u].$$

Observing that $\alpha, \beta, \dots, \omega$ are equivalent umbral letters, we have

$$\langle U(f)\,|\,\pi(\Delta)\rangle = \langle U(f)\,|\,\Delta\rangle.$$

Averaging over all permutations, we have

$$\begin{aligned} n!\langle U(f)\,|\,\Delta\rangle &= \sum_{\pi}\langle U(f)\,|\,\pi(\Delta)\rangle \\ &= \Big\langle U(f)\,\Big|\,\sum_{\pi}\operatorname{sgn}(\pi)\pi(\alpha)_2^{m-1}\pi(\beta)_1\pi(\beta)_2^{m-2} \\ &\qquad\cdots\pi(\omega)_1^{m-1}\prod_{\delta<\varepsilon}[\delta\quad\varepsilon]\prod_{\delta}[\delta\quad u]\Big\rangle. \end{aligned}$$

But

$$\sum_{\pi} \operatorname{sgn}(\pi)\pi(\alpha)_2^{m-1}\pi(\beta)_1\pi(\beta)_2^{m-2}\cdots\pi(\omega)_1^{m-1}$$

$$= \begin{vmatrix} \alpha_2^{m-1} & \alpha_1\alpha_2^{m-2} & \cdots & \alpha_1^{m-1} \\ \beta_2^{m-1} & \beta_1\beta_2^{m-2} & \cdots & \beta_1^{m-1} \\ \vdots & & & \\ \omega_2^{m-1} & \omega_1\omega_2^{m-2} & \cdots & \omega_1^{m-1} \end{vmatrix}$$

$$= \prod_{\delta<\varepsilon} [\delta \quad \varepsilon].$$

We conclude that

$$\left\langle U(f) \mid \prod_{\delta<\varepsilon} [\delta \quad \varepsilon]^2 \prod_{\delta} [\delta \quad u] \right\rangle$$

is a constant multiple of $g(x, y)$ and, hence, is apolar to $f(x, y)$. Further, if $J(a_0, a_1, \ldots, a_n, x, y) \not\equiv 0$, then one of its coefficients is nonzero and the system (5.6) of linear equations is linearly independent. From this we conclude that the dimension of the space of all binary forms of degree $m$ apolar to $f(x, y)$ is exactly one, or every form of degree $m$ apolar to $f(x, y)$ is a constant multiple of $J(a_0, a_1, \ldots, a_n, x, y)$. □

As an illustration of the relation between apolarity and canonical forms, we derive the canonical forms for the binary quintic. Let

$$f(x, y) = \sum_{k=0}^{5} \binom{5}{k} a_k x^k y^{5-k}, \qquad a_5 \neq 0,$$

be a binary quintic with nonzero leading coefficient. The space of all binary cubics

$$g(x, y) = \sum_{k=0}^{3} \binom{3}{k} b_k x^k y^{3-k}$$

apolar to $f(x, y)$ can be found by solving the following simultaneous linear equations for the unknowns $b_0$, $b_1$, $b_2$, and $b_3$:

$$\text{(5.8)} \qquad \begin{aligned} -a_0b_3 + 3a_1b_2 - 3a_2b_1 + a_3b_0 &= 0 \\ -a_1b_3 + 3a_2b_2 - 3a_3b_1 + a_4b_0 &= 0 \\ -a_2b_3 + 3a_3b_2 - 3a_4b_1 + a_5b_0 &= 0. \end{aligned}$$

If this system of linear equations is linearly independent, then all the cubic forms apolar to $f(x, y)$ are a constant multiple of $J(a_0, \ldots, a_5, x, y)$, where $J$ is the covariant

$$\langle U \mid [\alpha \quad \beta]^2[\beta \quad \gamma]^2[\gamma \quad \alpha]^2[\alpha \quad u][\beta \quad u][\gamma \quad u] \rangle.$$

If $J(x, y)$ has three distinct linear factors, $\mu_1 x - \nu_1 y$, $\mu_2 x - \nu_2 y$, and $\mu_3 x - \nu_3 y$, then we have

A. $$f(x, y) = a(\mu_1 x - \nu_1 y)^5 + b(\mu_2 x - \nu_2 y)^5 + c(\mu_3 x - \nu_3 y)^5.$$

If

$$J(x, y) = (\mu_1 x - \nu_1 y)^2(\mu_2 x - \nu_2 y),$$

then we have

B. $$f(x, y) = (ax + by)(\mu_1 x - \nu_1 y)^4 + c(\mu_2 x + \nu_2 y)^5.$$

Finally, if

$$J(x, y) = (\mu x - \nu y)^3,$$

C. $$f(x, y) = (ax^2 + bx + cy^2)(\mu x - \nu y)^3.$$

Now suppose the system (5.8) of linear equations is not linearly independent. This is the case if and only if $J(x, y)$ is identically zero. If (5.8) has rank 2, then we can set $b_3 = 0$ and solve for $b_2$, $b_1$, and $b_0$ to obtain a nonzero quadratic $Q(x, y)$, uniquely determined up to a constant multiple, apolar to $f(x, y)$. If $Q(x, y)$ has two distinct linear factors, $\mu_1 x - \nu_1 y$ and $\mu_2 x - \nu_2 y$, then we have

D. $$f(x, y) = a(\mu_1 x - \nu_1 y)^5 + b(\mu_2 x - \nu_2 y)^5.$$

If

$$Q(x, y) = (\mu x - \nu y)^2,$$

then we have

E. $$f(x, y) = (ax + by)(\mu x - \nu y)^4.$$

Finally, (5.8) may have rank 1. We can then set $b_3 = b_2 = 0$ and solve for $b_1$ and $b_0$ to obtain a nonzero linear form $l(x, y) = \mu x - \nu y$ apolar to $f(x, y)$. In this case, we have

F. $$f(x, y) = a(\mu x - \nu y)^5.$$

This completes the classification of the canonical forms of the binary quintic.

There is no analogue of Sylvester's theorem for binary forms of even degree in general. However, under certain conditions, we can obtain a similar canonical form.

Let

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k}$$

be a binary form of even degree $n = 2j$. The *catalecticant* of $f(x, y)$ is defined to be the determinant

$$\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_j \\ a_1 & a_2 & a_3 & & a_{j+1} \\ a_2 & a_3 & & & \\ \vdots & & & & \\ a_j & a_{j+1} & & \cdots & a_{2j} \end{vmatrix}.$$

LEMMA 5.4. *The catalecticant has the umbral representation*

$$\left\langle U(f) \,\middle|\, \frac{1}{(j+1)!} \prod_{\gamma<\delta} [\gamma \quad \delta]^2 \right\rangle,$$

*where $\mathcal{C} = \{\alpha, \beta, \ldots, \varepsilon\}$ is a set of $j + 1$ linearly ordered umbral letters of $f(x, y)$ and the product is over all pairs $(\gamma, \delta)$ of umbral letters such that $\gamma < \delta$.*

The *proof* is similar to that of Lemma 5.3.

As examples, the catalecticant for the quartic has the umbral representation

$$\langle U(f) | [\alpha \quad \beta]^2 [\alpha \quad \gamma]^2 [\beta \quad \gamma]^2 \rangle,$$

and the catalecticant for the sextic has the umbral representation

$$\langle U(f) | [\alpha \quad \beta]^2 [\alpha \quad \gamma]^2 [\alpha \quad \delta]^2 [\beta \quad \gamma]^2 [\beta \quad \delta]^2 [\gamma \quad \delta]^2 \rangle.$$

THEOREM 5.3. *Let $f(x, y)$ be a binary form of even degree $n = 2j$. Then there exists a nonzero form $g(x, y)$ of degree $j$ apolar to $f(x, y)$ if and only if the catalecticant of $f(x, y)$ is zero. Further, if there exists one such form $g(x, y)$ with $j$ distinct linear factors, $\mu_1 x - \nu_1 y, \ldots, \mu_j x - \nu_j y$, then*

$$f(x, y) = \sum_{i=0}^{j} c_i (\mu_i x - \nu_i y)^n.$$

PROOF. Let

$$f(x, y) = \sum_{k=0}^{n} \binom{n}{k} a_k x^k y^{n-k}$$

be a form of even degree $n = 2j$. The rank $r$ of the system (5.5) of linear equations

$$\sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} a_{k+l} b_{m-k} = 0, \qquad l = 0, 1, \ldots, j,$$

is strictly less than $j + 1$ if and only if the catalecticant is zero. But $r < j + 1$ if and only if the space of all forms of degree $j$ apolar to $f(x, y)$ has dimension at least one. This proves the first assertion. The second assertion follows from Proposition 5.1. □

5.4 *The Hessian and the cubic.* In this section we shall apply invariant theoretic reasoning to the cubic. To this end, we first consider a covariant, the Hessian, which is of independent interest.

The *Hessian* $H(x, y)$ of a binary form $f(x, y)$ of degree $n$ is defined umbrally by

$$H(x, y) = \tfrac{1}{2}n^2(n-1)^2\left\langle U(f) \,|\, [\alpha \;\; \beta]^2[\alpha \;\; u]^{n-2}[\beta \;\; u]^{n-2}\right\rangle,$$

where $\alpha$ and $\beta$ are umbral letters of $f(x, y)$. A more effective way to compute the Hessian is given by

LEMMA 5.5.

$$H(x, y) = \det\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\, \partial y} \\ \dfrac{\partial^2 f}{\partial x\, \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$

PROOF. To show this, we note the following commutation relations:

$$\frac{\partial}{\partial x}U(f) = U(f)\frac{\partial}{\partial u_2}, \qquad \frac{\partial}{\partial y}U(f) = -U(f)\frac{\partial}{\partial u_1}.$$

Applying this to $f(x, y) = \langle U(f) \,|\, [\gamma \;\; u]^n\rangle$, $\gamma$ an umbral letter of $f$, we obtain

$$\frac{\partial^2 f}{\partial x^2} = \left\langle U(f) \,\middle|\, \frac{\partial^2}{\partial u_2^2}[\gamma \;\; u]^n\right\rangle = \left\langle U(f) \,\middle|\, n(n-1)\gamma_1^2[\gamma \;\; u]^{n-2}\right\rangle,$$

$$\frac{\partial^2 f}{\partial x\, \partial y} = \left\langle U(f) \,\middle|\, \frac{\partial^2}{\partial u_2\, \partial u_1}[\gamma \;\; u]^n\right\rangle = \left\langle U(f) \,\middle|\, n(n-1)\gamma_1\gamma_2[\gamma \;\; u]^{n-2}\right\rangle,$$

$$\frac{\partial^2 f}{\partial y^2} = \left\langle U(f) \,\middle|\, \frac{\partial^2}{\partial u_1^2}[\gamma \;\; u]^n\right\rangle = \left\langle U(f) \,\middle|\, n(n-1)\gamma_2^2[\gamma \;\; u]^{n-2}\right\rangle.$$

Thus,

$$\det\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\, \partial y} \\ \dfrac{\partial^2 f}{\partial x\, \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

$$= \det\begin{pmatrix} \dfrac{\partial^2}{\partial x^2}\left\langle U(f) \,\middle|\, [\alpha \;\; u]^n\right\rangle & \dfrac{\partial^2}{\partial x\, \partial y}\left\langle U(f) \,\middle|\, [\beta \;\; u]^n\right\rangle \\ \dfrac{\partial^2}{\partial x\, \partial y}\left\langle U(f) \,\middle|\, [\alpha \;\; u]^n\right\rangle & \dfrac{\partial^2}{\partial y^2}\left\langle U(f) \,\middle|\, [\beta \;\; u]^n\right\rangle \end{pmatrix}$$

$$= n^2(n-1)^2\left\langle U(f) \,\middle|\, \left(\alpha_1^2\beta_2^2 - \alpha_1\alpha_2\beta_1\beta_2\right)[\alpha \;\; u]^{n-2}[\beta \;\; u]^{n-2}\right\rangle.$$

As $\alpha$ and $\beta$ are equivalent umbral letters, we can replace $\alpha_1^2\beta_2^2[\alpha \;\; u]^{n-2}$ $[\beta \;\; u]^{n-2}$ in the above expression by $\frac{1}{2}(\alpha_1^2\beta_2^2 + \alpha_2^2\beta_1^2)[\alpha \;\; \mathrm{u}]^{n-2}[\beta \;\; u]^{n-2}$. The proof can now be completed by observing that

$$\tfrac{1}{2}\left(\alpha_1^2\beta_2^2 + \alpha_2^2\beta_1^2\right) - \alpha_1\alpha_2\beta_1\beta_2 = \tfrac{1}{2}[\alpha \;\; \beta]^2. \qquad \square$$

More useful theoretically is the following expansion of the Hessian in terms of the homogenized roots of $f(x, y)$:

LEMMA 5.6. *The expression of the Hessian in terms of homogenized roots $\mu_i$, $\nu_i$ is given by*

$$H(x, y) = \frac{1}{2((n-2)!)^2} \sum_{\pi, \sigma} (\mu_{\pi(1)}\nu_{\sigma(1)} - \mu_{\sigma(1)}\nu_{\pi(1)}) (\mu_{\pi(2)}\nu_{\sigma(2)} - \mu_{\sigma(2)}\nu_{\pi(2)})$$
$$\times \prod_{i=3}^{n} (\mu_{\pi(i)}x - \nu_{\pi(i)}y)(\mu_{\sigma(i)}x - \nu_{\sigma(i)}y).$$

PROOF. This can be obtained immediately from the umbral representation and Algorithm 4.1. □

The importance of the Hessian in the theory of canonical forms is due to the following property: if the Hessian of a binary form vanishes identically, then the binary form has the simplest possible canonical form. More precisely, we have

PROPOSITION 5.3. *The Hessian of the binary form $f(x, y)$ of degree $n$ vanishes identically if and only if the form is the nth power of linear form.*

PROOF. If $f(x, y)$ is the $n$th power of a linear form, the homogenized roots of $f(x, y)$ are all equal. Using Lemma 5.4, we conclude that $H(x, y) \equiv 0$.

Now suppose $H(x, y) \equiv 0$. From the umbral representation or Lemma 5.5, we obtain the following equations:

$$\begin{aligned} &a_n a_{n-2} - a_{n-1}^2 = 0, \\ &a_{n-3}a_n - a_{n-2}a_{n-1} = 0, \\ &(n-3)a_n a_{n-4} - (n-1)a_{n-2}^2 + 2a_{n-1}a_{n-3} = 0, \\ &\quad\vdots \end{aligned}$$

Suppose first that $a_n \neq 0$. Setting $a_n = a$, $a_{n-1} = a\lambda$, and applying these equations one by one, we obtain

$$a_{n-2} = a\lambda^2, \quad a_{n-3} = a\lambda^3, \quad \dots, \quad a_0 = a\lambda^n.$$

This implies

$$f(x, y) = a(x + \lambda y)^n.$$

If $a_n = 0$, a similar argument shows that $f(x, y) = ay^n$. □

We now turn our attention to obtaining the canonical forms of a binary cubic $f(x, y)$ with nonzero leading coefficient. If the Hessian of $f(x, y)$ is identically zero, then $f(x, y)$ is the cube of a linear form and

A. $$f(x, y) = (\mu x - \nu y)^3.$$

In the case of the cubic, the Hessian plays another rôle. It is also a multiple of the covariant

$$J = \langle U(f) | [\alpha \quad \beta]^2 [\alpha \quad u][\beta \quad u] \rangle.$$

Thus, by Lemma 5.3, $H(x, y)$ is apolar to $f(x, y)$.

We shall now assume $H(x, y) \not\equiv 0$. If $H(x, y) = (\mu x - \nu y)^2$, then, by Proposition 5.1,

B. $$f(x, y) = (ax + by)(\mu x - \nu y)^2.$$

In particular, $f(x, y)$ has the same repeated linear factor as $H(x, y)$. If $H(x, y)$ has two distinct linear factors $\mu_1 x - \nu_1 y$ and $\mu_2 x - \nu_2 y$, then by Proposition 5.1,

C. $$f(x, y) = a(\mu_1 x - \nu_1 y)^3 + b(\mu_2 x - \nu_2 y)^3.$$

This completes the classification of the canonical forms of the binary cubic.

This classification offers a procedure for solving a cubic polynomial $p(x)$ by radicals. Let $f(x, y)$ be the homogenization $y^3 p(x/y)$ of $p(x)$. Find the Hessian $H(x, y)$ of $f(x, y)$ by the formula in Lemma 5.5. If $H(x, y) \equiv 0$, then $p(x)$ is a perfect cube and the triple root can easily be found. If $H(x, 1)$ is a quadratic, its roots $\lambda_1$ and $\lambda_2$ can be found explicitly by the quadratic formula. If $\lambda_1 = \lambda_2$, then $\lambda_1$ is also a double root of $p(x)$ and the remaining root can be found by division. If $\lambda_1 \neq \lambda_2$, then

$$p(x) = a(x - \lambda_1)^3 + b(x - \lambda_2)^3$$

where

$$a = p(\lambda_2)/(\lambda_2 - \lambda_1)^3 \quad \text{and} \quad b = p(\lambda_1)/(\lambda_1 - \lambda_2)^3.$$

Once in this form, the roots of $p(x)$ can easily be obtained by extracting cube roots. If $H(x, 1)$ is linear with root $\lambda$, then

$$p(x) = a(x - \lambda)^3 + b.$$

The numbers $a$ and $b$ are given by

$$a = (p(\mu) - p(\lambda))/(\mu - \lambda)^3 \quad \text{where } \mu \neq \lambda, b = p(\lambda),$$

and the roots of $p(x)$ can be obtained by extracting cube roots.

Among all the methods for solving a cubic equation by radicals, the present one, which is closest in spirit to the method described by Mark Kac in his first published paper, is easiest to apply and remember.

## 6. The finiteness theorem.

6.1 *Generating sets of covariants.* We shall now consider the central problem of both classical and modern invariant theory: Does there exist a finite generating set for the set of covariants?

A set $\mathbb{S}$ of covariants of binary forms of degree $n$ is said to be a *generating set* if for every covariant $I$, there exists a polynomial $P(X_1, \ldots, X_s)$ such that $I = P(C_1, \ldots, C_s)$, where $C_1, \ldots, C_s$ are covariants in $\mathbb{S}$. The central result of the invariant theory of binary forms is

THEOREM 6.1 (THE FINITENESS THEOREM). *There exists a finite generating set for the covariants of binary forms of degree n.*

We shall present two constructive proofs of the finiteness theorem. The first, which occupies §§6.2–6.4, is based on the idea of circular straightening. The second, which occupies §6.5 relies on a combinatorial lemma of Gordan.

6.2 *Circular straightening.* In order to prove the finiteness theorem, we first describe another basis, the basis of cyclically standard bracket monomials, for the space $\mathfrak{B}$ of bracket polynomials. This basis has combinatorial properties similar to the basis of standard bracket monomials and is useful in other contexts.

Let $\mathcal{Q} = \{\alpha, \beta, \gamma, \delta, \dots\}$ be an alphabet. A *cyclic order* $\Gamma$ on the alphabet $\Gamma$ is a relation, denoted $\alpha \Rightarrow \beta$, satisfying: for every letter $\beta$ in $\mathcal{Q}$ there exists a unique $\alpha$ such that $\alpha \Rightarrow \beta$ and a unique $\gamma$ such that $\beta \Rightarrow \gamma$. The letter $\alpha$ is called the *predecessor* of $\beta$ and the letter $\gamma$ is called the *successor* of $\beta$. A cyclic order $\Gamma$ can be visualized as a directed graph, also denoted by $\Gamma$, on the vertex set $\mathcal{Q}$ such that there is a directed edge from $\alpha$ to $\beta$ if and only if $\alpha \Rightarrow \beta$. This directed graph is a directed cycle and there is a unique simple path (that is, a path without any repeated vertices) *from* any vertex $\alpha$ *to* any other vertex $\delta$. We say that $\beta$ is *between* $\alpha$ and $\delta$ and write $\alpha \to \beta \to \delta$ if $\beta$ is a vertex distinct from $\alpha$ and $\delta$ on the unique simple path from $\alpha$ to $\delta$. If $\mathcal{Q}'$ is a subset of $\mathcal{Q}$, the restriction of $\Gamma$ to $\mathcal{Q}'$ is the cyclic order defined by: $\alpha \Rightarrow \beta$ if every letter between $\alpha$ and $\beta$ is not in $\mathcal{Q}'$.

Now let $\mathfrak{U}$ be the umbral space formed with the alphabet $\mathcal{Q}$ and let $\mathfrak{B}$ be the space of bracket monomials. Let $M$ be a bracket monomial in $\mathfrak{B}$. Two brackets, $[\alpha \quad \gamma]$, $[\beta \quad \delta]$, in $M$ are said to form a *crossing pair* if $\alpha \to \beta \to \gamma \to \delta$. This may be visualized as follows: let the letters in $\mathcal{Q}$ be placed, according to their cyclic order, on a circle in the plane and represent every bracket $[\alpha \quad \beta]$ by a straight line segment between the points $\alpha$ and $\beta$. Then two brackets cross if and only if their line segments have a point of intersection inside the circle. We say that a bracket monomial is *cyclically standard* if it is nonzero and no two brackets in $M$ form a crossing pair.

THEOREM 6.2. *The cyclically standard bracket monomials form a basis for the space $\mathfrak{B}$ of bracket polynomials.*

The proof consists of the following two lemmas.

LEMMA 6.1. *Every bracket monomial can be written as a linear combination with integer coefficients of cyclically standard bracket monomials.*

PROOF. Let $M$ be a bracket monomial and let $\mathcal{C}$ be a list (with suitable multiplicity) of the crossing pairs of brackets in $M$. The length $|\mathcal{C}|$ of $\mathcal{C}$ is called the *crossing number* of $M$. Suppose that $M$ is not cyclically standard. Let $[\alpha \quad \gamma]$, $[\beta \quad \delta]$ be a crossing pair of brackets in $M$ and write $M = [\alpha \quad \gamma]$ $[\beta \quad \delta]M'$. By the syzygy (Lemma 3.1),

$$M = [\alpha \quad \gamma][\beta \quad \delta]M' = [\alpha \quad \beta][\gamma \quad \delta]M' + [\alpha \quad \delta][\beta \quad \gamma]M'.$$

We claim that both bracket monomials on the right-hand side have crossing numbers strictly smaller than $|\mathcal{C}|$. To see this, let $[\xi \quad \eta]$, $[\zeta \quad \omega]$ be a crossing pair of brackets in $[\alpha \quad \beta][\gamma \quad \delta]M'$. If $[\xi \quad \eta]$ and $[\zeta \quad \omega]$ are both in the submonomial $M'$, then the pair $[\xi \quad \eta]$, $[\zeta \quad \omega]$ is also in $\mathcal{C}$. If $[\xi \quad \eta] = [\alpha \quad \beta]$, then we have $\alpha \to \zeta \to \beta \to \omega$. If $\omega$ is between $\beta$ and $\delta$, then $\zeta \to \beta \to \omega \to \delta$ and $[\zeta \quad \omega]$, $[\beta \quad \delta]$ is a crossing pair of brackets in $\mathcal{C}$. Similarly, if $\omega = \delta$ or is between $\delta$ and $\alpha$, then $\zeta \to \gamma \to \omega \to \alpha$ and $[\zeta \quad \omega]$, $[\alpha \quad \gamma]$ is a crossing pair of brackets in $\mathcal{C}$. A similar argument can be applied if $[\xi \quad \eta] = [\gamma \quad \delta]$. Hence, to every crossing pair of brackets in the bracket monomial $[\alpha \quad \beta][\gamma \quad \delta]M'$ is associated, in a one-to-one manner, a crossing pair of brackets in $\mathcal{C}$. However, the pair $[\alpha \quad \gamma][\beta \quad \delta]$ in $\mathcal{C}$ is *not* associated with any crossing pair in $[\alpha \quad \beta][\gamma \quad \delta]M'$. Hence, the crossing number of $[\alpha \quad \beta][\gamma \quad \delta]M'$ is strictly smaller than $|\mathcal{C}|$, the crossing number of $M$. Similarly, the crossing number of $[\alpha \quad \delta][\beta \quad \gamma]M'$ is strictly smaller than the crossing number of $M$.

Iterating this procedure, we can write the bracket monomial $M$ as a linear combination (with integer coefficients) of bracket monomials whose crossing numbers are zero, that is, cyclically standard bracket monomials. □

LEMMA 6.2. *The cyclically standard bracket monomials form a linearly independent set.*

PROOF. Suppose not. From the set of all nontrivial linear dependence relations between cyclically standard bracket monomials, choose one, $\Sigma_{k=1}^{m} c_k M_k = 0$, in which (a) $c_k \neq 0$ for all $k$, (b) the number of distinct letters is as small as possible, and (c) subject to (b), the maximum number of brackets in a monomial $M_k$ occurring in the linear relation is as small as possible. Let $\mathcal{Q}'$ be the set of all letters occurring in the linear relation cyclically ordered by the restriction of the cyclic order on $\mathcal{Q}$. Let $\delta$ and $\varepsilon$ be two letters in $\mathcal{Q}'$ such that $\delta \Rightarrow \varepsilon$ in the restricted order. By condition (c), $[\delta \quad \varepsilon]$ is not a common factor of all the bracket monomials $M_k$. Thus, on setting $\delta$ equal to $\varepsilon$, not all the bracket monomials $M_k$ vanish. By our choice of $\delta$ and $\varepsilon$, those bracket monomials which remain nonzero also remain cyclically standard. We thus obtain a nontrivial linear relation with a strictly smaller number of distinct letters, contradicting our initial choice. □

For our purposes, the most important property of cyclically standard bracket monomials is that they have outer segments. If $\alpha$ and $\varepsilon$ are letters in $\mathcal{Q}$, the *segment* $(\alpha, \varepsilon)$ is the set of all letters (strictly) between $\alpha$ and $\varepsilon$: that is,

$$(\alpha, \varepsilon) = \{\gamma \colon \alpha \to \gamma \to \varepsilon\}.$$

Note that the segments $(\alpha, \varepsilon)$ and $(\varepsilon, \alpha)$ are distinct: indeed, $\mathcal{Q} = (\alpha, \varepsilon) \cup (\varepsilon, \alpha) \cup \{\alpha, \varepsilon\}$. Let $M$ be a bracket monomial. A bracket $[\gamma \quad \delta]$ is said to be *diagonal* if neither $\gamma \Rightarrow \delta$ nor $\delta \Rightarrow \gamma$. Now let $\mathcal{Q}_M$ be the set of letters occurring in $M$ cyclically ordered by the restriction of the cyclic order on $\mathcal{Q}$. A nonempty segment $(\alpha, \varepsilon)$ in $\mathcal{Q}_M$ is said to be an *outer segment* of $M$ if for all $\gamma$ in $(\alpha, \varepsilon)$, there are no diagonal brackets in $M$ containing $\gamma$. An outer segment is said to be *maximal* if it is not strictly contained in any outer segment.

PROPOSITION 6.1. *Let $M$ be a cyclically standard bracket monomial. Then either $M$ has no diagonal brackets or there exist (at least) two maximal outer segments of $M$.*

PROOF. We proceed by induction on $|\mathcal{Q}_M|$, the number of distinct letters occurring in $M$. If $M$ has no diagonal brackets, we are done. Now, if $[\alpha \quad \varepsilon]$ is a bracket, the distance of $[\alpha \quad \varepsilon]$ is the length of the shortest *un*directed path between $\alpha$ and $\varepsilon$. Among all the diagonal brackets in $M$, choose one, $\pm[\alpha \quad \varepsilon]$, for which the distance is at a minimum and is attained by the directed path from $\alpha$ to $\varepsilon$. As $M$ has no crossing pair of brackets, the segment $(\alpha, \varepsilon)$ is a maximal outer segment.

To find the second maximal outer segment, observe that the brackets in $M$ can be partitioned into three blocks: the brackets containing only letters from $\{\alpha, \varepsilon\} \cup (\alpha, \varepsilon)$, the brackets equal to $[\alpha \quad \varepsilon]$, and the brackets containing only letters from $\{\alpha, \varepsilon\} \cup (\varepsilon, \alpha)$. Consider the submonomial $M'$ of $M$ consisting of all the brackets in $M$ from the second and third blocks. As $|\mathcal{A}_{M'}| < |\mathcal{A}_M|$, by induction, $M'$ has no diagonal brackets or $M'$ has two disjoint maximal outer segments. In the first case, $(\varepsilon, \alpha)$ is a maximal outer segment. In the second case, one of the outer segments does not contain the subset $\{\alpha, \varepsilon\}$ and is also an outer segment of $M$. □

6.3 *Elemental bracket monomials.* We shall now consider a space $\mathcal{B}$ of bracket polynomials formed with the alphabet $\mathcal{A} = \{\alpha, \beta, \ldots, \varepsilon, u\}$ consisting of a finite number of Greek letters and the single Roman letter $u$. As in §4.3, a bracket monomial $M$ is said to be *regular* of degee $d$ if for every Greek umbral letter $\alpha$ in $\mathcal{A}$, the number of occurrences of $\alpha$ in $M$ equals $d$; the number of occurrences of $u$ need not equal $d$ and is called the *order* of $M$. An *elemental* bracket monomial is either a regular bracket monomial of degree one or a regular bracket monomial of degree two which is not the product of two regular bracket monomials of degree one. For example, if $\mathcal{A} = \{\alpha, \beta, \gamma, \delta, \varepsilon, u\}$, then $[\alpha \quad \beta][\gamma \quad \delta][\varepsilon \quad u]$, $[\alpha \quad \beta][\gamma \quad u][\delta \quad u][\varepsilon \quad u]$, and $[\alpha \quad \beta][\beta \quad \gamma][\gamma \quad \delta]$ $[\delta \quad \varepsilon][\varepsilon \quad \alpha]$ are all elemental bracket monomials.

The main result in this section is that the elemental bracket monomials form a generating set for the set of regular bracket polynomials.

PROPOSITION 6.2 (KEMPE'S LEMMA). *Every regular bracket monomial formed with the alphabet $\mathcal{A} = \{\alpha, \beta, \ldots, \varepsilon, u\}$ can be written as a linear combination with integer coefficients of products of elemental bracket monomials.*

PROOF. We proceed by induction on $|\mathcal{A}| - 1$, the number of Greek letters in $\mathcal{A}$. To do so we need to strengthen the induction hypothesis slightly. Let $\mathcal{G}$ be another alphabet of Greek letters, and let $e$: $\mathcal{G} \rightarrow \mathcal{A}$ be a function whose image is $\{\alpha, \beta, \ldots, \varepsilon\}$, the subset of Greek letters in $\mathcal{A}$. Two letters $\xi$ and $\eta$ are said to be *equivalent* (*to* $\gamma$) if $e(\xi) = e(\eta) = \gamma$. A bracket monomial $M$ formed with letters from $\mathcal{G} \cup \{u\}$ is said to be a bracket monomial *with equivalent letters*. If $M$ is a bracket monomial with equivalent letters, we extend our terminology by saying that $M$ is *regular*, *cyclically standard*, etc. if the bracket monomial formed from $M$ by replacing each Greek letter $\xi$ by $e(\xi)$ is regular, cyclically standard, etc. As all our results are proved by exhibiting a constructive algorithm, the version for bracket monomials is equivalent to the version for bracket monomials with equivalent letters.

By definition of an elemental bracket monomial, our assertion is true if $|\mathcal{A}| - 1 = 1$ or 2. By examining cases, it is also true for $|\mathcal{A}| - 1 = 3$. We will now assume there exists an algorithm for writing any bracket monomial with equivalent letters as a linear combination with integer coefficients of products of elemental bracket monomials with equivalent letters if $\mathcal{A}$ has $n - 1$ or fewer Greek letters.

Let $M$ be a regular bracket monomial on the alphabet $\mathscr{Q}$ consisting of $n$ Greek letters and the Roman letter $u$. By Theorem 6.2 it suffices to prove our assertion for $M$ a cyclically standard bracket monomial. If $M$ is cyclically standard, by Proposition 6.1 there are two maximal outer segments in $M$ and, hence, there exists a Greek letter $\beta$ in an outer segment. There are now two possible cases: first, there exists such a Greek letter $\beta$ such that both its predecessor $\alpha$ and its successor $\gamma$ are Greek letters, or, second, for all such Greek letters, either the predecessor or the successor is the Roman letter $u$.

To deal with the first case, we first prove the following lemma.

Lemma 6.3. *The bracket monomial $M$ can be written as a linear combination with integer coefficients of bracket monomials $N$ (which may not be cyclically standard) such that $\beta$ is still in an outer segment and $[\alpha \quad \gamma]$ does not appear as a bracket in $N$: that is,*

$$N = [\alpha \quad \beta]^{m-k}[\beta \quad \gamma]^{k} N',$$

*where $N'$ is a bracket monomial not containing the letter $\beta$ or the bracket $[\alpha \quad \gamma]$.*

Proof. Let $M$ be a bracket monomial of degree $d$ and order $t$ containing $r$ brackets equal to $[\alpha \quad \gamma]$. A simple counting argument shows that there are $r + \frac{1}{2}t + \frac{1}{2}d(n-4)$ brackets in $M$ not containing $\alpha$, $\beta$, or $\gamma$. As $n \geqslant 4$, there are at least $r$ such brackets in $M$.

Now, if $[\delta \quad \varepsilon]$ is a bracket not containing $\alpha$, $\beta$, or $\gamma$, we can write $M$ as a linear combination of bracket monomials (which are not necessarily cyclically standard) containing $r-1$ brackets equal to $[\alpha \quad \gamma]$ by using the syzygy

$$[\alpha \quad \gamma][\delta \quad \varepsilon] = [\alpha \quad \varepsilon][\delta \quad \gamma] - [\alpha \quad \delta][\varepsilon \quad \gamma].$$

As there are at least $r$ brackets of the form $[\delta \quad \varepsilon]$, we can continue this process till none of the brackets $[\alpha \quad \gamma]$ remains in any of the bracket monomials. □

It now suffices to prove our assertion for a monomial of the form

$$N = [\alpha \quad \beta]^{m-k}[\beta \quad \gamma]^{k} N',$$

where $N'$ contains no bracket equal to $[\alpha \quad \gamma]$. The total number of occurrences of $\alpha$ and $\gamma$ in $N'$ is exactly $d$, the degree of $N$. Thus, if $\alpha$ and $\gamma$ are defined to be equivalent to a new letter $\zeta$, $N'$ is a regular bracket monomial with letters equivalent to the alphabet $(\mathscr{Q} - \{\alpha, \beta, \gamma\}) \cup \{\zeta\}$ with $n-1$ Greek letters. By induction we can write $N'$ as a linear combination

$$N' = \sum_i b_i E_{i1} E_{i2} \cdots E_{ik(i)}$$

of products of elemental bracket monomials $E_{ij}$ with equivalent letters. Multiplying by $[\alpha \quad \beta]^{d-k}[\beta \quad \gamma]^{k}$, we obtain

$$N = \sum_i b_i [\alpha \quad \beta]^{d-k}[\beta \quad \gamma]^{k} E_{i1} E_{i2} \cdots E_{ik(i)}.$$

We shall next distribute the brackets $[\alpha \quad \beta]$ and $[\beta \quad \gamma]$ among the elemental bracket monomials $E_{ij}$ with equivalent letters to obtain elemental bracket monomials $E_{ij}$ with letters from $\mathcal{Q}$. This distribution is done according to the following scheme:

I. If $E_{ij}$ is of degree two and contains two occurrences of $\alpha$ (or $\gamma$), then set $\hat{E}_{ij} = [\beta \quad \gamma]^2 E_{ij}$ (or $[\alpha \quad \beta]^2 E_{ij}$).

II. If $E_{ij}$ is of degree two and contains one occurrence each of $\alpha$ and $\gamma$, then set $\hat{E}_{ij} = [\alpha \quad \beta][\beta \quad \gamma] E_{ij}$.

III. If $E_{ij}$ is of degree one and contains one occurrence of $\alpha$ (or $\gamma$), then set $\hat{E}_{ij} = [\beta \quad \gamma] E_{ij}$ (or $[\alpha \quad \beta] E_{ij}$).

A simple counting argument shows that the number of brackets $[\alpha \quad \beta]$ and $[\beta \quad \gamma]$ matches up with the number of different types of elemental bracket monomials. Thus,

$$N = \sum_i b_i \hat{E}_{i1} \hat{E}_{i2} \cdots \hat{E}_{ik(i)}.$$

This completes our proof for the first case.

In the second case, every Greek letter $\beta$ in an outer segment has one of its predecessors or successors equal to the Roman letter $u$. This is possible only in the case when there are exactly two maximal outer segments and they are of the form $(\alpha, u)$ and $(u, \delta)$ and contain a single Greek letter. Thus, we have $\alpha \to \beta \to u \to \gamma \to \delta$ in the cyclic order, where $(\alpha, u) = \{\beta\}$ and $(u, \delta) = \{\gamma\}$.

LEMMA 6.4. *Under these conditions,* $[\zeta \quad u]$ *is a bracket in* $M$ *for every Greek letter* $\zeta$ *in* $\mathcal{Q}$.

PROOF. Suppose $[\zeta \quad u]$ is not a bracket in $M$. Then $\zeta \neq \beta$ or $\gamma$. As $\zeta$ is not in an outer segment, there exists a Greek letter $\eta$ such that $[\zeta \quad \eta]$ is diagonal. Choose $\eta$ such that the distance is smallest. Then, as in Proposition 6.1, we can conclude that one of the segments $(\zeta, \eta)$ or $(\eta, \zeta)$ is an outer segment, contrary to our assumptions. □

By the lemma,

$$M = \Big( \prod_{\zeta \in \mathcal{Q}} [\zeta \quad u] \Big) M'.$$

As $\prod_{\zeta \in \mathcal{Q}} [\zeta \quad u]$ is elemental of degree one and $M'$ is regular of degree one less than the degree of $M$, we can repeat our entire argument using the smaller bracket monomial $M'$.

This completes the proof of Proposition 6.2. □

6.4 *Reduction of degree.* To prove the finiteness theorem, recall from §4.3 that every covariant of binary forms of degree $n$ can be expressed as a linear combination of symmetric difference terms $\langle S | M \rangle$, where $M$ is a regular bracket monomial formed with the alphabet $\{1, 2, \ldots, n, u\}$. Regarding the integers $1, 2, \ldots, n$ as Greek letters, we can use Kempe's lemma to conclude that every regular bracket monomial $M$ is a linear combination of products of elemental bracket monomials: that is, the finite set $\{E_1, \ldots, E_m\}$ of elemental bracket monomials is a generating set for the set of regular bracket monomials. However, it is not true (except when $n = 1$) that the set $\{\langle S | E_1 \rangle, \ldots, \langle S | E_m \rangle\}$ is a generating set for the set of symmetric difference terms and we need to take a larger (but still finite) set constructed from the elemental bracket monomials. This construction is given in the following lemma.

LEMMA 6.5 (HILBERT). *Let $r = n!$ and let $\{E_1, \dots, E_m\}$ be a generating set for the set of regular bracket monomials on the alphabet $\{1, 2, \dots, n, u\}$. Then the set of symmetric difference terms*

$$\langle S \,|\, E_1^{e_1} E_2^{e_2} \cdots E_m^{e_m} \rangle, \qquad 0 \leqslant e_i \leqslant r - 1, \quad e_i \neq 0 \text{ for some } i,$$
$$\langle S \,|\, E_i^r \rangle, \qquad 1 \leqslant i \leqslant m,$$

*is a generating set for the set of symmetric difference terms.*

PROOF. Let $E_i$ be a bracket monomial in the generating set and let $\pi(E_i)$ be the bracket monomial obtained from $E_i$ by replacing each integer $h$ in $E_i$ by its image $\pi(h)$ under the permutation $\pi$ on the set $\{1, 2, \dots, n\}$. Let $a_1(\pi(E_i))$, $a_2(\pi(E_i)), \dots, a_r(\pi(E_i))$ be the elementary symmetric functions of the $r$ bracket monomials $\pi(E_i)$, $\pi \in \Omega_n$. As

$$\prod_{\pi \in \Omega_n} (S - \pi(E_i)) = S^r - a_1(\pi(E_i)) S^{r-1}$$
$$+ a_2(\pi(E_i)) S^{r-2} + \cdots \pm a_r(\pi(E_i)),$$

we have

$$E_i^r = a_1(\pi(E_i)) E_i^{r-1} - a_2(\pi(E_i)) E_i^{r-2} + \cdots \pm a_r(\pi(E_i)).$$

More generally, for $k \geqslant r$, we have

$$E_i^k = a_1(\pi(E_i)) E_i^{k-1} - a_2(\pi(E_i)) E_i^{k-2} + \cdots \pm a_r(\pi(E_i)) E_i^{k-r}.$$

Now observe that the elementary symmetric functions $a_j(\pi(E_i))$ are invariant under permutations of the integers $\{1, 2, \dots, n\}$ inside the bracket monomials $\pi(E_i)$. Thus, if $M$ is a bracket monomial,

$$\langle S \,|\, a_j(\pi(E_i)) M \rangle = \sum_{\pi \in \Omega_n} a_j(\pi(E_i)) \pi(M) = a_j(\pi(E_i)) \langle S \,|\, M \rangle.$$

From this, we obtain, if $k \geqslant r$,

$$\langle S \,|\, E_1^{e_1} \cdots E_i^k \cdots E_m^{e_m} \rangle = a_1(\pi(E_i)) \langle S \,|\, E_1^{e_1} \cdots E_i^{k-1} \cdots E_m^{e_m} \rangle$$
$$- \cdots \pm a_r(\pi(E_i)) \langle S \,|\, E_1^{e_1} \cdots E_i^{k-r} \cdots E_m^{e_m} \rangle.$$

The degree of $E_i$ in every term on the right-hand side is strictly smaller than $k$. Therefore, by iterating this process, we can write any symmetric difference term

$$\langle S \,|\, M \rangle = \left\langle S \,\middle|\, \sum_i b_i E_1^{e_{i1}} E_2^{e_{i2}} \cdots E_m^{e_{im}} \right\rangle$$

as a linear combination of products of $a_j(\pi(E_i))$ and $\langle S \,|\, E_1^{e_1} \cdots E_m^{e_m} \rangle$, where $0 \leqslant e_i \leqslant r - 1$ and $e_i \neq 0$ for some $i$.

To complete the proof, recall that the elementary symmetric functions $a_j(\pi(E_i))$ can be written in terms of the power sum symmetric functions $h_j(\pi(E_i))$, where

$$h_j(\pi(E_i)) = \sum_{\pi \in \Omega_n} \pi(E_i)^j = \langle S \,|\, E_i^j \rangle.$$

Hence, the set consisting of the symmetric difference terms

$$\langle S \,|\, E_1^{e_1} E_2^{e_2} \cdots E_m^{e_m} \rangle, \qquad 0 \leqslant e_i \leqslant r-1, \quad e_i \neq 0 \textit{ for some } i,$$

and

$$\langle S \,|\, E_i^r \rangle, \qquad i = 1, 2, \ldots, m,$$

is a generating set for the set of symmetric difference terms. □

Applying this lemma to the set $\{E_1, \ldots, E_m\}$ of elemental bracket monomials, we obtain the following, more explicit, version of the finiteness theorem.

THEOREM 6.3. *Let $\{E_1, \ldots, E_m\}$ be the set of elemental bracket monomials formed with the alphabet $\{1, 2, \ldots, n, u\}$. The set of covariants whose representations in terms of the homogenized roots are given by*

$$\langle S \,|\, E_1^{e_1} E_2^{e_2} \cdots E_m^{e_m} \rangle, \qquad 0 \leqslant e_i \leqslant n! - 1, \quad e_i \neq 0 \textit{ for some } i,$$
$$\langle S \,|\, E_i^{n!} \rangle, \qquad 0 \leqslant i \leqslant m,$$

*is a finite generating set for the set of covariants of binary forms of degree n.*

6.4 *Gordan's lemma.* Our second proof of the finiteness theorem is similar in structure to the first proof. Once again, we begin by finding a generating set for the set of regular bracket monomials. The combinatorial tool for doing this is a lemma of Gordan.

Consider the system

$$\begin{gathered} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1m}X_m = 0 \\ \vdots \\ a_{k1}X_1 + a_{k2}X_2 + \cdots + a_{km}X_m = 0 \end{gathered}$$

of linear equations in the variables $X_i$ where the coefficients $a_{ij}$ are (positive or negative) integers. Let $N$ be the set of *nonnegative* integers and consider the set $\mathcal{N}$ of solutions $\mathbf{s} = (s_i)$ in $N^m$. As the system is linear, $\mathcal{N}$ contains the zero solution and is closed under componentwise addition.

PROPOSITION 6.3 (GORDAN'S LEMMA). *There exists a finite set $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p\}$ of solutions such that if $\mathbf{s}$ is a solution, then*

$$\mathbf{s} = \sum_{j=1}^{p} c_j \mathbf{b}_j,$$

*where $c_j$ are nonnegative integers. Such a finite set is called a* basis *of solutions.*

PROOF. We need a preliminary combinatorial result concerning the order structure of $N^m$. The set $N$ of nonnegative integers is a totally ordered set under the usual order relation $\leqslant$ of less than or equal to. The $m$-fold product $N^m$ can be given a partial order by: $(s_i) \leqslant (t_i)$ if $s_i \leqslant t_i$ for every index $i$. Two elements $(s_i)$ and $(t_i)$ are *comparable* if either $(s_i) \leqslant (t_i)$ or $(s_i) \geqslant (t_i)$. An *antichain* is a subset of $N^m$ in which no two elements are comparable.

LEMMA 6.6. *Let $(\mathbf{s}_k)_{k=1}^{\infty}$ be an infinite sequence of elements in $N^m$. Then there exist indices i and j such that $i < j$ and $\mathbf{s}_i \leqslant \mathbf{s}_j$. In particular, $N^m$ has no infinite antichains.*

PROOF. We proceed by induction on $m$. The lemma is certainly true when $m = 0$. Suppose that it is true for $N^{m-1}$. Consider $N^m$ as the product $N^{m-1} \times N$ and let $((\mathbf{a}_k, s_k))_{k=1}^{\infty}$ be an infinite sequence of elements in $N^{m-1} \times N$.

We first show that there exists an infinite nondecreasing subsequence $(s_{k'})$ in the sequence $(s_k)$ of nonnegative integers. Suppose first that the set $\{s_k$: $1 \leqslant k < \infty\}$ of elements in the sequence $(s_k)$ is finite. Then there exists an element $s$ such that $s_{k'} = s$ for infinitely many $k'$. The subsequence $(s_{k'})$ is constant, hence nondecreasing. Now consider the case when the set $\{s_k\}$ is infinite. Let $S_1$ be the set $\{s_j$: $j > 1$ and $s_j \geqslant s_1\}$. As $\{s_k\}$ is infinite, $S_1$ is nonempty and so there exists $j_1$ such that $1 < j_1$ and $s_1 \leqslant s_{j_1}$. Repeating this argument, we obtain an infinite nondecreasing subsequence $(s_{j_k})$.

Finally, let $(s_{k'})$ be a nondecreasing subsequence of $(s_k)$. Consider the infinite sequence $(\mathbf{a}_{k'})$ in $N^{m-1}$. By induction, there exist indices $i'$ and $j'$ such that $i' < j'$ and $\mathbf{a}_{i'} \leqslant \mathbf{a}_{j'}$. For the same pair of indices $(\mathbf{a}_{i'}, s_{i'}) \leqslant (\mathbf{a}_{j'}, s_{j'})$. □

Consider the set $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p\}$ in $\mathfrak{N}$ of *minimal nonzero solutions*, that is, the set $\{\mathbf{b}_j\}$ of all solutions in $\mathfrak{N}$ satisfying: $\mathbf{b}_j > \mathbf{0}$ and there exists no solution $\mathbf{b}'$ such that $\mathbf{b}_j > \mathbf{b}' > \mathbf{0}$. This set $\{\mathbf{b}_j\}$ is an antichain in $N^m$ and is therefore finite by the preceding lemma. It remains to show that every solution $\mathbf{s}$ is a linear combination of $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p$.

Let $\mathbf{s}$ be a solution in $\mathfrak{N}$. If $\mathbf{s} \neq \mathbf{0}$, then $\mathbf{s} \geqslant \mathbf{b}_j$ for some $j$. The vector $\mathbf{s}' = \mathbf{s} - \mathbf{b}_j$ is still in $N^m$ and by linearity is a solution. Further, $\Sigma s_i > \Sigma s_i'$. Thus, if we iterate this process, we must arrive at the zero solution after a finite number of steps. Thus, $\mathbf{s} - \Sigma \mathbf{c}_j \mathbf{b}_j = \mathbf{0}$ for some nonnegative integers $c_j$, as desired. □

Now consider the system of diophantine equations in the unknowns $m_{ij}$, $t_i$, $d$, $t$, $h$, where $i, j = 1, 2, \ldots, n$ from Proposition 4.3:

$$\begin{aligned} & m_{ij} = m_{ji}, \quad m_{ii} = 0, \\ & t_i + m_{i1} + m_{i2} + \cdots + m_{in} = d \quad \text{for } i = 1, 2, \ldots, n, \\ (6.1) \qquad & t_1 + t_2 + \cdots + t_n = t, \\ & \sum_{i=1}^{n} t_i + \sum_{i,j=1}^{n} m_{ij} = 2h. \end{aligned}$$

If $\mathbf{s}$ is a solution to the system (6.1), then, by Proposition 4.3, the bracket monomial $M(\mathbf{s})$, defined by

$$M(\mathbf{s}) = \left( \prod_{i,j=1}^{n} [i \quad j]^{m_{ij}} \right) \left( \prod_{i=1}^{n} [i \quad u]^{t_i} \right),$$

is a regular bracket monomial, and conversely. Now let $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ be a basis of solutions of (6.1). As

$$M(\mathbf{s} + \mathbf{s}') = M(\mathbf{s}) M(\mathbf{s}'),$$

every regular bracket monomial is a product of bracket monomials of the form $M(\mathbf{b}_i)$: thus, $\{M(\mathbf{b}_1), \ldots, M(\mathbf{b}_p)\}$ is a generating set for the set of regular bracket monomials. The proof of the finiteness theorem can now be completed as in §6.4 by using Lemma 6.5.

The finiteness theorem holds for joint convariants of several binary forms. The second proof generalizes immediately. However, Kempe's lemma does not hold in general for several binary forms.

6.6 *The binary cubic.* We end by computing explicitly a generating set for the covariants of the binary cubic.

We begin by listing the elemental bracket monomials formed with the alphabet $\{1,2,3,u\}$, grouped according to their order:

*order* 0: $[1\ \ 2][2\ \ 3][3\ \ 1]$;

*order* 1: $[1\ \ 2][3\ \ u]$, $[1\ \ 3][2\ \ u]$, $[2\ \ 3][1\ \ u]$;

*order* 3: $[1\ \ u][2\ \ u][3\ \ u]$.

(Note: There are no elemental bracket monomials of order 2 since the only regular bracket monomials of order 2 are of the form $[1\ \ 2][2\ \ 3][3\ \ u][1\ \ u]$, and they are products of two elemental bracket monomials of order 1. All regular bracket monomials of order greater than 3 are not elemental.) Since

$$[2\ \ 3][1\ \ u]=[1\ \ 3][2\ \ u]-[1\ \ 2][3\ \ u],$$

we can take as a generating set for the regular bracket monomials

$$A=[1\ \ 2][2\ \ 3][3\ \ 1],\qquad B=[1\ \ 2][3\ \ u],$$
$$C=[1\ \ 3][2\ \ u],\qquad D=[1\ \ u][2\ \ u][3\ \ u].$$

Thus, by Lemma 6.5, a generating set for the covariants of the cubic is given by

$$\begin{gathered}\left\langle S\,|\,A^aB^bC^cD^d\right\rangle,\qquad 0\leqslant a,b,c,d\leqslant 5,\\ \left\langle S\,|\,A^6\right\rangle,\ \left\langle S\,|\,B^6\right\rangle,\ \left\langle S\,|\,C^6\right\rangle,\ \left\langle S\,|\,D^6\right\rangle.\end{gathered}\tag{6.2}$$

This generating set is highly redundant and may be reduced considerably.

PROPOSITION 6.4. *A generating set for the covariants of the binary cubic consists of* $\Delta=\langle S\,|\,A^2\rangle$, $f=\langle S\,|\,D\rangle$, $-H=\langle S\,|\,B^2\rangle$, *and* $T=\langle S\,|\,B^2C\rangle$, *where* $\Delta$, *the discriminant*, $f$, *the form itself*, $H$ *the Hessian*, *and* $T$, *the Jacobian of the form and the Hessian*, *are given umbrally by*

$$\begin{aligned}\Delta&=\tfrac{27}{2}\left\langle U\,|\,[\alpha\ \ \beta]^2[\alpha\ \ \gamma][\beta\ \ \delta][\gamma\ \ \delta]^2\right\rangle,\\ f&=\left\langle U\,|\,[\alpha\ \ u]^3\right\rangle,\\ H&=18\left\langle U\,|\,[\alpha\ \ \beta]^2[\alpha\ \ u][\beta\ \ u]\right\rangle,\\ T&=108\left\langle U\,|\,[\alpha\ \ \beta]^2[\alpha\ \ \gamma][\beta\ \ u][\gamma\ \ u]^2\right\rangle.\end{aligned}$$

The proof consists of showing that all the covariants in the list (6.2) can be expressed in terms of $\Delta$, $f$, $H$, and $T$.

Observe first that the elemental bracket monomial $D$ is invariant under permutation of the integers $\{1,2,3\}$. Hence, for any bracket monomial $M$,

$$\langle S\,|\,DM\rangle=\sum_{\pi\in\Omega_n}D\pi(M)=D\langle S\,|\,M\rangle.$$

Thus, except for $\langle S\,|\,D\rangle$ itself, all symmetric difference terms $\langle S\,|\,M\rangle$, where $M$ contains $D$ as a factor, can be deleted from (6.2).

Similarly, $A^2$ is invariant under permutations of the integers $\{1,2,3\}$, and except for $\langle S|A^2\rangle$, every symmetric difference term $\langle S|M\rangle$, where $M$ contains $A^2$ as a factor, can be deleted from (6.2).

The remainder of the proof consists of somewhat tedious computations. We first consider symmetric difference terms of the form $\langle S|B^bC^c\rangle$. These yield two covariants not already obtained, namely, $\langle S|B^2\rangle = -H$ and $\langle S|B^2C\rangle = T$. Further computations show that the other symmetric difference terms are either zero (examples of such symmetric difference terms are $\langle S|B\rangle$, $\langle S|B^3\rangle$, $\langle S|B^5\rangle$, $\langle S|B^4C\rangle$, etc.) or yield covariants expressible in terms of $\Delta$, $f$, $H$, or $T$ (examples are: $\langle S|BC\rangle = -\frac{1}{2}H$, $\langle S|B^4\rangle = \frac{1}{8}H^2$, $\langle S|B^6\rangle = \frac{1}{32}H^3 + 4\Delta f^2$, $\langle S|C^2B\rangle = T$, etc.). To finish the proof, we consider symmetric difference terms of the form $\langle S|AB^bC^c\rangle$. Computations show that these are either zero (examples are $\langle S|A\rangle = \langle S|AB\rangle = \langle S|AC\rangle = 0$) or yield covariants expressible in terms of $\Delta$, $f$, $H$ or $T$ (an example is $\langle S|AB^2C\rangle = \frac{3}{2}\Delta^2 f$). This completes the proof of Proposition 6.4. □

**7. Further work.** We have chosen to list only a few of the open problems in the invariant theory of binary forms. The selection is short and disregards the problems arising in the applications of invariant theory to number theory, algebraic geometry, computational complexity, and other fields. We have also limited the selection to problems which could have been formulated in the last century, though they seldom were.

1. Gram's theorem, as it is somewhat optimistically called, asserts that in certain cases, the vanishing of covariants is equivalent to a projective property. A clear statement of this widely held doctrine has not been given, yet the correspondence between geometric properties and the vanishing of covariants is admittedly the *raison d'être* of invariant theory. What is needed is a formulation of the first order logic of binary forms and an algorithm coding sentences of such a first order logic into the vanishing of sets of covariants.

2. The vanishing of some covariants expresses properties of binary forms which can be stated without invoking the underlying projective geometry. For example, the vanishing of the Hessian indicates that the form is a power of a linear form; the vanishing of other covariants, such as the catalecticant, obtained by apolarity have similar meanings. As a further example, the vanishing of the Jacobian of two forms indicates that the two forms are algebraically dependent. On the other hand, the vanishing of the Jacobian of a cubic form and its Hessian cannot be given a "meaning" without using the vanishing of certain cross-ratios, and a rather stilted meaning at that (see Gurevich, p. 270). Is there a criterion for distinguishing the two kinds of covariants?

3. There is a class of covariants, namely covariants of a form of degree $n$ that are obtainable by setting to zero certain coefficients of a form of degree $n + 1$, which can be regarded as covariants of a single form of infinite degree. Such "stable" covariants were classically known as perpetuants. There is strong evidence in the work of Cayley, Grace, MacMahon, and Stroh that perpetuants and their syzygies can be completely classified. This area is in a particularly sorry state. MacMahon's method of partitions is at variance with Grace's use

of tableaux and with Cayley's differential operators, or hyperdeterminants as he called them. Perpetuants may in fact provide the answer to the previous problem, and it may well turn out that the explicit computation of a generating set of covariants for a form of degree $n$ other than perpetuants will be a sterile exercise.

4. Among the transvectants, only the first (the Jacobian) and the last (the apolar covariant) have been interpreted. Do the intermediate transvectants have a "meaning"?

5. Little work has been done on the significance of the vanishing of covariants in the real or $p$-adic fields. Sylvester expressed Sturm's theorem in terms of real invariants, but it is hard to find other work in the same spirit. Inequalities between covariants preserved under the group of linear changes of variables with positive determinant seem never to have been investigated, even though such inequalities are essential in the study of the distribution of roots in the complex plane. Similarly, P. Cohen's decision procedure for $p$-adic fields can be invariantly expressed.

6. Another interesting group of covariants, expressed in terms of the roots of a form of degree $n$, is given by the cumulants. Cumulants originated in statistics. Keeping to the simplest case, let

$$s_k = \sum_{i=1}^{n} \frac{\lambda_i^k}{n}$$

and write

$$\sum_{k=1}^{\infty} \frac{s_k t^k}{k!} = \exp\left( c_1 t + \frac{c_2 t^2}{2!} + \cdots \right).$$

It can then be verified that for $k > 1$, the expressions $c_k$ are invariants of the translation group. We conjecture that the $c_k$'s and their associated multilinearized joint invariants provide a basis for all "interpretable" invariants. An explicit interpretation of the $c_k$'s probably exists, but it has not been stated.

7. The expression of covariants in terms of determinants of partial derivatives of the form is as yet poorly understood. Cayley's hyperdeterminant notation can be used as an alternative to umbral notation—in fact, it is a thinly disguised equivalent—but does not give, for example, even the expression of the Hessian as a determinant of second partial derivatives. What is missing is an algorithm for obtaining one expression in terms of the other.

8. Some connections between covariants (especially transvectants) of binary forms and ordinary differential operators were investigated by F. Klein (see Grace and Young, Appendix II). It would be worthwhile to re-examine Klein's work in the light of differential algebra.

9. Gordan's method of transvectants for his proof of the finiteness theorem was based on an ingenious method of substitutions of brackets into brackets. After Hilbert's work, Gordan's ideas were abandoned. However, Gordan's method remains the most effective one. Further insight into the explicit generation of covariants will require a systematization of Gordan's brackets of brackets (plethysms) and a concomitant deepening of the straightening algorithm.

10. On a smaller scale, the umbral representation of discriminants and resultants can be of use in the study of these covariants, particularly for forms in more than two variables. However, not much that is explicit is known at present, even for binary forms.

11. Kempe's lemma (see §6.3) does not extend immediately to several binary forms. Is there a similar result for several binary forms?

12. A generalization of circular straightening to higher dimension would be of the utmost interest, since it would yield, among other returns, another explicit construction of the representations of the symmetric group.

13. Apolarity is the study of invariant bilinear forms on tensor spaces. For forms in more than two variables, there are several notions of apolarity corresponding to various symmetry classes of tensors. No systematic classification has ever been attempted of such invariant bilinear forms. The closest analogue to the apolar covariant for binary forms is an apolar covariant defined for forms in several variables which are products of linear forms. An analogue of Sylvester's theorem can be proved for such forms.

14. Our proof of Hermite's reciprocity law in §4.4 yields an explicit isomorphism **i** between the space $\mathcal{U}^S[n, d, t]$ of symmetrized bracket polynomials and the space $\mathcal{V}^S[d, n, t]$ of symmetrized difference polynomials. Combining this isomorphism with the homomorphism **h** sending the umbral representation of a covariant into its representation in terms of homogenized roots as follows,

$$\mathcal{U}^S[n, d, t] \xrightarrow{\mathbf{h}} \mathcal{V}^S[n, d, t] \xrightarrow{\mathbf{i}} \mathcal{U}^S[d, n, t] \xrightarrow{\mathbf{h}} \mathcal{V}^S[d, n, t] \xrightarrow{\mathbf{i}} \mathcal{U}^S[n, d, t],$$

we obtain a linear map from the space of covariants into itself. There is some evidence to suggest that this is the identity; if so, Algorithm 4.1 can also be applied to obtain the umbral representation of a covariant from its representation in terms of homogenized roots.

## References

*Books*

A. Cayley, *Collected mathematical papers*, Vols. 1–12, Cambridge Univ. Press, Cambridge, 1889–1897.

L. E. Dickson, *Algebraic invariants*, Wiley, New York, 1914.

J. A. Dieudonné and J. B. Carrell, *Invariant theory, old and new*, Academic Press, New York and London, 1971.

E. B. Elliot, *An introduction to the algebra of quantics*, 2nd ed., Oxford Univ. Press, Oxford, 1913.

P. Gordan, *Vorlesungen über Invariantentheorie* (G. Kershensteiner, ed.), Vols. I, II, Teubner, Leipzig, 1885, 1887.

J. H. Grace and A. Young, *The algebra of invariants*, Cambridge Univ. Press, Cambridge, 1903.

G. B. Gurevich, *Foundations of the theory of algebraic invariants*, Noordhoff, Groningen, 1964.

C. Hermite, *Oeuvres* (E. Picard, ed.), Vols. I–IV, Gauthier-Villars, Paris, 1905–1917.

D. Hilbert, *Gesammelte Abhandlungen*, Vol. II, Springer-Verlag, Berlin, 1933.

J. P. S. Kung (Editor), *Young tableaux in combinatorics, invariant theory, and algebra*, Academic Press, New York and London, 1982.

P. A. MacMahon, *Collected papers* (G. Andrews, ed.), Vol. II, M. I. T. Press, Cambridge, 1984.

W. F. Meyer, *Bericht über den gegenwartigen Stand der Invariantentheorie*, Jahresber. Deutsch. Math.-Verein. **1** (1890–1891).

D. Mumford, *Geometric invariant theory*, Springer-Verlag, Berlin and New York, 1965.

I. Schur, *Vorlesungen über Invariantentheorie* (H. Grunsky, ed.), Springer-Verlag, Berlin and New York, 1968.

T. A. Springer, *Invariant theory*, Lecture Notes in Math., vol. 585, Springer-Verlag, Berlin and New York, 1977.

E. Study, *Methoden zur Theorie der ternären Formen*, Teubner, Leipzig, 1889 (reprinted with an introduction by G.-C. Rota, Springer-Verlag, Berlin and New York, 1982).

J. J. Sylvester, *Collected mathematical papers*, Vols. I–IV, Cambridge Univ. Press, New York, 1904–1912.

H. W. Turnbull, *The theory of determinants, matrices, and invariants*, Blackie, London, 1931.

H. Weyl, *The classical groups, their invariants and representations*, 2nd ed., Princeton Univ. Press, Princeton, N. J., 1946.

______, *Gesammelte Abhandlungen* (K. Chandrasekharan, ed.), Vols. 1–4, Springer-Verlag, Berlin and New York, 1968.

A. Young, *Collected papers*, Univ. of Toronto Press, Toronto and Buffalo, 1977.

*Papers*

P. J. Cohen, *Decision procedures for real and p-adic fields*, Comm. Pure Appl. Math. **22** (1969), 131–152.

J. Désarménien, J. P. S. Kung and G.-C. Rota, *Invariant theory, Young bitableaux, and combinatorics*, Adv. in Math. **27** (1978), 63–92.

P. Doubilet, G.-C. Rota and J. Stein, *On the foundations of combinatorial theory*. IX: *Combinatorial methods in invariant theory*, Stud. Appl. Math. **53** (1974), 185–216.

C. Hermite, *Sur la théorie, des fonctions homogènes à deux indéterminées*, Cambridge and Dublin Math. J., 1854 (= *Oeuvres*, Vol. I, pp. 298–349).

D. Hilbert, *Über die Endlichkeit des Invariantensystems für binären Grundformen*, Math. Ann. **33** (1889), 223–226 (= *Gesammelte Abhandlungen*, Vol. II, pp. 162–164).

M. Katz, *O Nowym Sposobie rozwiazywanie rownan*, Stopnia trzeciego; Mlody Matematyk, Nos. 4–5 (April/May 1931), 69–71.

A. B. Kempe, *On regular difference terms*, Proc. London Math. Soc. **25** (1894), 343–359.

G.-C. Rota, *The number of partitions of a set*, Amer. Math. Monthly **71** (1964), 498–504.

J. J. Sylvester, *An essay on canonical forms, supplement to a sketch of a memoir on elimination, transformation and canonical forms*, George Bell, Fleet Street, 1851 (= *Collected papers*, Vol. I, Paper 34).

______, *On a remarkable discovery in the theory of canonical forms and of hyperdeterminants*, Phil. Mag. **2** (1851), 391–410 (= *Collected papers*, Vol. I, Paper 41).

DEPARTMENT OF MATHEMATICS, NORTH TEXAS STATE UNIVERSITY, DENTON, TEXAS 76203

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139