Europ. J. Combinatorics (1993) **14**, 157–181

Apolarity and Canonical Forms for Homogeneous Polynomials

RICHARD EHRENBORG AND GIAN-CARLO ROTA

Dedicated to Bernt Lindström on his 60th birthday

Kăre Bernt Lindström,

Gratulerar på Din sextioårs dag. Vi hoppas att denna artikel kommer falla Dig i smaken.

Vi börjar med en kort studie i algebraiska matroider, och fortsätter med att bevisa relationen mellan Jacobianen av en mängd algebraiska funktioner och deras algebraiska oberoende. Med detta resulat bevisar vi de två huvudsatserna, som behandlar kanoniska former. Dessa satser reducerar frågan om en form är kanonisk för homogena polynom i q variabler och av grad p till att undersöka om ett homogent linjärt ekvations system har bara den triviala lösningen. Genom att använda apolaritet kan detta linjära system enkelt beskrivas. Till sist ger vi en mångfald av exempel av kanoniska former för homogena polynom.

1. INTRODUCTION

Whereas the combinatorial aspects of linear dependence are in our day well understood, largely through the development of the theory of matroids, the same cannot be said of algebraic dependence of sets of polynomials with coefficients in a field. What once appeared to be a simple way out has actually turned out to be a problem; indeed, on a superficial level, the theory of algebraic dependence is combinatorially identical to the theory of linear dependence [5]; in fact, the MacLane–Steinitz exchange axiom that is now the defining property of a matroid was historically first derived in order to deal with algebraic dependence.

It seems, however, that the full meaning of algebraic dependence is not carried by the exchange property, and that a better understanding of algebraic dependence may be the payoff of a deeper study of the finer properties of algebraic dependence in concrete algebraic and combinatorial contexts. Such was the objective of some work carried out at the beginning of this century by British and German algebraists, on the possible canonical forms of homogeneous polynomials under linear changes of variables. We find it regrettable that this line of work was broken off (perhaps as a consequence of the excess of abstraction that swept commutative algebra at the time), and it is our purpose to give in this paper a complete, self-contained, updated introduction to the theory of canonical forms of polynomials, as it has been known to this day, to the best of our knowledge. It turns out that this theory is intimately related to the understanding of algebraic dependence. We begin by giving a bird's eye view of the contents of the paper.

We shall deal mainly with homogeneous polynomials of degree p in q variables over a field of characteristic zero, which will be taken to be the complex number field without warning (the generalization of the following results to fields of arbitrary characteristic is, in our opinion, a completely open problem of the utmost interest). Such polynomials have received, in the last two-hundred-odd years, a wealth of nomenclature that perhaps surpasses that of any other mathematical notion: they have been called forms, quantics, primals, varieties of dimension $q-2$ in a projective space of dimension $q-1$, symmetric tensors, homogeneous elements of the symmetric algebra of a vector space, etc. We shall usually refer to them as 'forms' or, following an

0195-6698/93/030157 + 25 $08.00/0

old British custom, as 'q-ary p-ics'; in particular, we shall speak of q-ary quadratic, binary p-ics, ternary quartics, etc. We shall try to refer as little as possible to the underlying q-dimensional vector space over which q-ary p-ics are defined as 'tensors'. The problem of canonical forms, stated rather loosely, is the following: How much can one such 'generic' form be 'simplified' by linear changes of variables (that is, under the action of the general linear group in q variables)?

Attractive as such a statement may be, it needs to be made more precise. It is easy to see that the lack of uniqueness of any proposed canonical form renders a precise statement rather difficult. Take the example of a q-ary quadratic form. A classic result states that, generically, such a form can be written as the sum of squares of linear forms; of all canonical forms, surely this one seems to be the soundest. But suppose that q is even; say, $q = 2j$. Then one can show, using our second main theorem, that a quadratic form can be (generically) written in the form $x_1x_2 + x_3x_4 + \cdots + x_{q-1}x_q$ after a suitable linear change of variables: this latter canonical form does away with any prejudices that we may have about canonical forms of polynomials in relation to the eigenvalues of symmetric matrices. A new start needs to be made: in the wake of the work of our predecessors, Grace and Young [3], Lasker [6], Richmond [11] and Wakeford [14], on whose shoulders we lean heavily, we have recognized that the fundamental notion that must play a role in a theory of canonical forms is the notion of apolarity.

Again, because of an over-emphasis on definitions that relied on the symbolic method of classical invariant theory, the notion of apolarity has remained sealed in the well of oblivion. It took us far longer than we are willing to admit to recognize that this notion coincided with another notion, one that is alive and well remembered. It turns out that there is a unique scalar-valued bilinear form $\langle g \,|\, f \rangle$ defined for all q-ary p-ics f and all dual q-ary p-ics g, that is, all forms in q variables and of degree p defined over the dual vector space V^*, that is invariant under the joint action of the general linear group on q-ary p-ics and the contragredient action on dual q-ary p-ics. This bilinear form is called the apolar bilinear form, and two forms f and g are said to be apolar when $\langle g \,|\, f \rangle = 0$.

The classical invariant theorists derived a number of identities satisfied by the apolar bilinear form. Although these identities could be motivated by the symbolic method, and although they could be painfully checked one by one, their authentic source remained mysterious. This mystery is probably the reason why, to this day, the constructive finiteness proofs for the invariants of binary forms given by Gordan and Jordan, which made powerful use of these identities, have not attracted many readers, and even fewer followers.

What took us a longer time to realize than we are willing to admit is that the apolar bilinear form is much more than a bilinear form: it is a pairing of Hopf algebras. More precisely, if $S(V)$ is the symmetric (Hopf) algebra of the vector space V, the graded dual Hopf algebra is, as is well known, the algebra of divided powers $D(V^*)$. There is a natural pairing between these two Hopf algebras with the well known properties ('measuring' in the sense of Sweedler or 'Laplace pairing' in the sense of Grosshans–Rota–Stein):

$$\langle gh \,|\, f \rangle = \sum_f \langle g \,|\, f_{(1)} \rangle \langle h \,|\, f_{(2)} \rangle, \qquad \langle g \,|\, fh \rangle = \sum_g \langle g_{(1)} \,|\, f \rangle \langle g_{(2)} \,|\, h \rangle.$$

What struck us one day is the fact that these identities are precisely two of the identities to be prominently found in Grace and Young's 'algebra of invariants' [3] in the language of the symbolic method. On closer examination, it turned out that the apolar bilinear form is nothing but the pairing of the Hopf algebras $S(V)$ and $D(V^*)$,

restricted to characteristic zero, and that all identities pertaining to apolarity could be proved using the language of Hopf algebras; in fact, the symbolic method itself could be interpreted, generalized and made entirely rigorous in terms of Hopf algebras. Important as it is, a purely Hopf algebraic treatment of apolarity is deferred to a parallel publication which is currently being written. The present paper has deliberately been written in the language of elementary linear algebra, with the objection of maximizing the readership.

One such identity, one that will play an important role below, is the fact that polarization operators are adjoint of multiplication by linear forms (and similarly for higher order forms). This duality allows an invariant definition of the notion of apolarity of forms of different degrees.

What is more remarkable, however, is the connection between apolarity and canonical forms, which we proceed to explain.

It may be significant to begin with a specific problem: For which integers n can a 'generic' q-ary p-ic be expressed as the sum of n pth powers of linear forms? The history of this problem is interesting. For $q=3$ and $p=4$, that is, for the ternary quartic, we have to deal with 15 coefficients. If $n=5$, that is, if we are allowed five linear forms each containing 3 coefficients, we have altogether 15 coefficients; thus, if we rely merely on a count of constants, we are led to the erroneous conclusion that a ternary quartic can be written as the sum of five fourth powers of linear forms. Clebsch was the first to prove that this assertion is false. But the natural explanation lies in the connection between apolarity and algebraic dependence, to wit:

(1) If the desired expression is valid, then every coefficient of a ternary quartic equals a polynomial in the coefficients of the linear forms.

(2) Again, the rank of the Jacobian of this set of polynomial must equal at least 15.

(3) Thus, the verification that desired expression is 'canonical', that is, valid for a dense set of ternary quartics, is reduced to the verification that a certain matrix has rank 15, not a bright prospect in general.

The notion of apolarity allows us to bypass such a verification. Specifically, let the linear forms be $X_1, \ldots, X_5$. Then the first main theorem below states that every ternary quartic cannot be written in the form $X_1^4 + \cdots + X_5^4$ iff for some choice of the coefficients of each of these forms, that is, for some specialization of the coefficients, there does exist a dual ternary quartic which is apolar to each of the forms $X_1^3, \ldots, X_5^3$. A few minutes' computation gives Clebsch's theorem.

The preceding example generalizes to give the first main theorem on apolarity below. We believe this theorem to yield the complete solution of the analog of Waring's problem for forms (that is, given a generic q-ary p-ic, which is the smallest integer n such that it can be expressed as the sum of n pth powers of linear forms?) and we hope to present the complete solution elsewhere. In the present work, we have limited our exposition to all cases thus far considered in the literature, supplemented by a few new cases that caught our fancy.

All these cases are applications of one single general result, stating that a q-ary p-ic $f(\mathbf{x})$ can be written as a sum $(\mathbf{a} \mid \mathbf{x})^p + (\mathbf{b} \mid \mathbf{x})^p + \cdots + (\mathbf{c} \mid \mathbf{x})^p$ of pth powers of linear forms iff the following strange sequence of quantified statements holds: there exist vectors $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ such that the q-ary $(p-1)-$ics $(\mathbf{a}' \mid \mathbf{x})^{p-1}$, $(\mathbf{b}' \mid \mathbf{x})^{p-1}, \ldots, (\mathbf{c}' \mid \mathbf{x})^{p-1}$ have no non-zero dual q-ary p-ic which is apolar to all of them. Thus, the verification that a generic q-ary p-ic can be written as the sum of n powers of linear forms boils down to checking whether a simple system of linear equations only has a trivial solution. Several examples of this verification can be found in the corollaries in Section 4.2 below.

Our second main theorem generalizes the notion of canonical form in a direction

which we believe to be as general as can be. The problem is whether a generic q-ary p-ic can be 'expressed' as a polynomial of 'prescribed' form in terms of forms of lower 'prescribed' degrees. The typical classical case of this situation is the quaternary cubic, or 'cubical surface', which can be 'reduced' to the canonical form $h_1h_2h_3 + h_4h_5h_6$, where the h_i's are linear forms. This classical result has proved useful in establishing the existence of at least 27 straight lines on any cubic surface.

Our second main theorem (Theorem 4.4 below) gives what we believe to be the most general result in this direction, a result which is elementary to apply to many important cases. Our result subsumes all previous work of Rosanes, Meyer, Sylvester, Lasker and Wakeford, and gives as special cases all canonical forms described by Richmond. We have striven for a self-contained presentation that dispenses with all techniques of algebraic geometry. In particular, we have deemed necessary to prefix our treatment with a detailed explanation of the combinatorics that underlies the algebraic proof of the Jacobian condition for the algebraic dependence of a set of polynomials. We are grateful to M. Artin and A. Mattuck for their help in providing a proof of Theorem 2.4, without which our presentation would not be self-contained.

We hope that the present work will stimulate further work along these lines, and we hazard the guess that the method of apolairty will prove to be a powerful tool in invariant theory and computer algebra.

2. Dependence

2.1. *Matroids and algebraic dependence.* We recall some of the basic definition of the theory of matroids. Our objective is to proceed as quickly as possible to the treatment of algebraic dependence of polynomials and more generally of algebraic functions.

Definition 2.1. A closure relation on a set S is a map from power set of S to itself, denoted $\bar{A}$ for $A \subseteq S$, such that:
(1) $A \subseteq \bar{A}$,
(2) $A \subseteq B \Rightarrow \bar{A} \subseteq \bar{B}$ and
(3) $\bar{\bar{A}} = \bar{A}$
for all subsets $A, B \subseteq S$.

A set $A \subseteq S$ is called *closed* if $\bar{A} = A$.

Definition 2.2. A closure relation is of finite type if whenever $s \in \bar{A}$ there exists a finite subset $F \subseteq A$ such that $s \in \bar{F}$.

Definition 2.3. A closure relation satisfies the MacLane–Steinitz exchange property if whenever $p, s \in S$, and $p \in \overline{A \cup \{s\}}$ but $p \notin \bar{A}$, then $s \in \overline{A \cup \{p\}}$.

Definition 2.4. A matroid is a set S together with a closure relation of finite type that satisfies the MacLane–Steinitz exchange property.

Definition 2.5. A subset A of S is called dependent if there exists $p \in A$ such that $p \in \overline{A - \{p\}}$. A set that is not dependent is called independent. A maximal independent set is called a basis. A subset A of S is called spanning if $\bar{A} = S$. The rank of a subset A of S, $r(A)$, is defined to be the cardinality of the largest independent subset of A.

The basic result on matroids is the following:

THEOREM 2.1. *In a matroid all bases are spanning sets. Moreover, all bases have the same finite cardinality. This cardinality is called the rank of the matroid.*

In this paper we shall be concerned with the particular instances of matroids arising in the theory of algebraic dependence. These matroids were considered by Bernt Lindström.

Let $\mathbb{C}(x_1, \ldots, x_q)$ be a transcendental extension of the complex numbers $\mathbb{C}$, and let S be the algebraic closure of the field $\mathbb{C}(x_1, \ldots, x_q)$. Define a closure relation on the set S as follows. For $A \subseteq S$, let A' the smallest field that contains the set A and the complex numbers $\mathbb{C}$, let $\bar{A}$ be the algebraic closure of the field A'.

PROPOSITION 2.1. *The following facts about the algebraic closure hold*:
(1) *The algebraic closure $A \to \bar{A}$ is a closure relation on the set S.*
(2) *$p \in \bar{A}$ iff there exists a polynomial $G(t, t_1, \ldots, t_n) \in \mathbb{C}[t, t_1, \ldots, t_n]$ such that*

$$\partial G/\partial t \not\equiv 0$$

and

$$G(p, a_1, \ldots, a_n) = 0$$

for some $a_1, \ldots, a_n \in A$.
(3) *Algebraic closure is of finite type and satisfies the MacLane–Steinitz exchange property.*

A matroid defined by algebraic closure will be called an *algebraic matroid.*

As a consequence of the above proposition we have the following:

PROPOSITION 2.2. *In an algebraic matroid a set A is dependent if there exists a non-zero polynomial $G(t_1, \ldots, t_n) \in \mathbb{C}[t_1, \ldots, t_n]$ such that*

$$G(a_1, \ldots, a_n) \equiv 0$$

for some distinct elements $a_1, \ldots, a_n \in A$.

We say that such a set A is *algebraically dependent.* If a set A is not algebraically dependent we say that A is *algebraically independent.*

The set $B = \{x_1, \ldots, x_q\}$ spans all of S; that is, $\bar{B} = S$. Moreover, B is an algebraically independent set. Hence B is a basis in the matroid defined on S by algebraic closure. Thus we have the following:

PROPOSITION 2.3. *The algebraic matroid defined by the field $\overline{\mathbb{C}(x_1, \ldots, x_q)}$ has rank q.*

We now give a simple combinatorial proof on the classical property of the Jacobian of algebraic functions.

THEOREM 2.2. *Let $p_1(x_1, \ldots, x_q), \ldots, p_q(x_1, \ldots, x_q) \in \overline{\mathbb{C}(x_1, \ldots, x_q)}$. The algebraic functions $p_1, \ldots, p_q$ are algebraically dependent iff their Jacobian of $p_1, \ldots, p_q$ vanishes; that is, iff*

$$\det(\partial p_i/\partial x_j)_{1 \leq i,j \leq q} \equiv 0.$$

PROOF. Assume that $p_1, \ldots, p_q$ are algebraically dependent. Choose a non-trivial polynomial $G(t_1, \ldots, t_q)$ in the variables $t_1, \ldots, t_q$, of lowest degree among all polynomials such that $G(p_1, \ldots, p_q) = 0$. Let $G_i = \partial G / \partial t_i$. By the chain rule we have

$$0 = \partial G / \partial x_j = \sum_{i=1}^{g} (\partial p_i / \partial x_j) G_i(p_1, \ldots, p_q). \tag{1}$$

We claim that $G_i(p_1, \ldots, p_q)$ is non-zero for some i. Indeed, there is at least one t_i such that $\partial G / \partial t_i \neq 0$. Furthermore, G_i is of strictly lower degree than G. If $G_i(p_1, \ldots, p_q)$ were identically 0, then the assumption that G is lowest degree among all polynomials such that $G(p_1, \ldots, p_q) = 0$ would be violated. Hence $G_i \neq 0$ for some i.

We may write equation (1) as the product of a matrix and a vector, as follows:

$$\mathbf{0} = (\partial p_i / \partial x_j)_{1 \leq i,j \leq q} \cdot (G_i(p_1, \ldots, p_q))_{1 \leq i \leq q}.$$

Since at least one of the G_i is non-zero, the matrix $(\partial p_i / \partial x_j)$ is singular, and therefore its determinant is zero. This completes half the proof.

Now assume that $p_1, \ldots, p_q$ are the algebraically independent. By Theorem 2.1, all bases of a matroid have the same number of elements. Hence the q elements $p_1, \ldots, p_q$, form a basis of the algebraic matroid defined on S, and hence the algebraic closure of the set of these q elements is S. In other words, $\overline{\{p_1, \ldots, p_q\}} = \overline{\{x_1, \ldots, x_q\}}$. Therefore, $x_i \in \overline{\{p_1, \ldots, p_q\}}$, and hence we can find a non-zero polynomial $H_i(t, t_1, \ldots, t_q) \in \mathbb{C}[t, t_1, \ldots, t_q]$ of smallest degree such that

$$H_i(x_i, p_1(x_1, \ldots, x_q), \ldots, p_q(x_1, \ldots, x_q)) \equiv 0 \tag{2}$$

and

$$\partial H_i / \partial t \neq 0.$$

Let

$$H_{i,0} = \partial H_i / \partial t, \qquad H_{i,j} = \partial H_i / \partial t_j.$$

Since the polynomials H_i have been chosen to have the smallest degree with respect to the condition $H_i(x_i, p_1, \ldots, p_q) = 0$, we have

$$H_{i,0}(x_i, p_1(x_1, \ldots, x_q), \ldots, p_q(x_1, \ldots, x_q)) \neq 0.$$

Differentiating equation (2) above with respect to x_k we obtain

$$\delta_{i,k} H_{i,0} + \sum_{j=1}^{q} (\partial p_j / \partial x_k) H_{i,j} = 0.$$

Since $H_{i,0}(x_i, p_1, \ldots, p_q) \neq 0$, we can rewrite the above equation in the form

$$\sum_{j=1}^{q} (\partial p_j / \partial x_k) \cdot (-H_{i,j} / H_{i,0}) = \delta_{i,k}.$$

But this identity shows that the Jacobian matrix has as its inverse the matrix $(-H_{i,j}/H_{i,0})_{1 \leq i,j \leq q}$. We conclude that the Jacobian $\det(\partial p_i / \partial x_j)$ is non-zero. □

THEOREM 2.3. *Let $p_1(x_1, \ldots, x_q), \ldots, p_r(x_1, \ldots, x_q) \in \overline{\mathbb{C}(x_1, \ldots, x_q)}$, where $r \leq q$. Then the algebraic functions $p_1, \ldots, p_r$ are algebraically independent iff the matrix $(\partial p_i / \partial x_j)_{1 \leq i \leq r, 1 \leq j \leq q}$ has full rank.*

PROOF. Assume that $p_1, \ldots, p_r$ are algebraically independent. To the algebraically independent set $\{p_1, \ldots, p_r\}$ of elements of $\overline{\mathbb{C}(x_1, \ldots, x_q)}$ add elements $p_{r+1}, \ldots, p_q$

so that $\{p_1, \ldots, p_q\}$ is a basis of the matroid $\overline{\mathbb{C}(x_1, \ldots, x_q)}$. By Theorem 2.2, the matrix $(\partial p_i/\partial x_j)_{1\leqslant i,j\leqslant q}$ has full rank. Hence the submatrix $(\partial p_i/\partial x_j)_{1\leqslant i\leqslant r, 1\leqslant j\leqslant q}$ has full rank. This proves half the theorem.

Now assume that the matrix $(\partial p_i/\partial x_j)_{1\leqslant i\leqslant r, 1\leqslant j\leqslant q}$ has rank r. Without loss of generality we may assume that the submatrix

$$(\partial p_i/\partial x_j)_{1\leqslant i,j\leqslant r} \tag{3}$$

has rank r; that is, that the $r\times r$ submatrix (3) is nonsingular. Let now $p_i = x_i$ for $i = r+1, \ldots, q$. We have $\partial p_i/\partial x_j = \delta_{i,j}$ for $i \geqslant r+1$, so that

$$\det(\partial p_i/\partial x_j)_{1\leqslant i,j\leqslant q} = \det(\partial p_i/\partial x_j)_{1\leqslant i,j\leqslant r}.$$

Hence the $q\times q$ matrix on the left-hand side is non-singular. By Theorem 2.2, the algebraic functions $p_1, \ldots, p_q$ are algebraically independent. In particular, $p_1, \ldots, p_r$ are algebraically independent. This completes the proof. □

PROPOSITION 2.4. *Let $p_1(x_1, \ldots, x_q), \ldots, p_s(x_1, \ldots, x_q) \in \overline{\mathbb{C}(x_1, \ldots, x_q)}$. Then the algebraic rank of the set $\{p_1, \ldots, p_s\}$ is given by*

$$r(\{p_1, \ldots, p_s\}) = \operatorname{rank}(\partial p_i/\partial x_j)_{1\leqslant i\leqslant s, 1\leqslant j\leqslant q}. \tag{4}$$

Thus the algebraic matroid on the set $\overline{\mathbb{C}(x_1, \ldots, x_q)}$ is a linear matroid over the field $\overline{\mathbb{C}(x_1, \ldots, x_q)}$ by the representation

$$p \to \left(\frac{\partial p}{\partial x_1}, \ldots, \frac{\partial p}{\partial x_q}\right).$$

PROOF. Assume that the algebraic rank of the set $\{p_1, \ldots, p_s\}$ is equal to r. Recall that q is the rank of the algebraic matroid, so $r \leqslant q$. Since the rank of the set $\{p_1, \ldots, p_s\}$ is equal to r we can find an independent subset of size r. Without loss of generality we may assume that the set $\{p_1, \ldots, p_r\}$ is algebraically independent. By Theorem 2.3 we conclude that the matrix

$$(\partial p_i/\partial x_j)_{1\leqslant i\leqslant r, 1\leqslant j\leqslant q} \tag{5}$$

has full rank; thus the rank is r. But matrix (5) is a submatrix of the matrix $(\partial p_i/\partial x_j)_{1\leqslant i\leqslant s, 1\leqslant j\leqslant q}$, and thus this matrix has rank at least r. Hence it follows that the left-hand side is less than or equal to the right-hand side in equation (4).

Assume that the matrix $(\partial p_i/\partial x_j)_{1\leqslant i\leqslant s, 1\leqslant j\leqslant q}$ has rank r. Observe that $r \leqslant q$. Then we can find r independent rows in the above matrix. Without loss of generality we can assume that the rows indexed by $i = 1, \ldots, r$ are independent. Hence the matrix $(\partial p_i/\partial x_i)_{1\leqslant i\leqslant r, 1\leqslant j\leqslant q}$ has full rank. By Theorem 2.3 we have that the set $\{p_1, \ldots, p_r\}$ is algebraically independent. Thus the set $\{p_1, \ldots, p_s\}$ has rank greater or equal than r. Thus the left-hand side is greater than or equal to the right side in equation (4).

By joining these two inequalities the identity (4) follows. □

THEOREM 2.4. *Let $p_1(x_1, \ldots, x_q), \ldots, p_r(x_1, \ldots, x_q) \in \overline{\mathbb{C}(x_1, \ldots, x_q)}$, where $r \leqslant q$. Let $P: \mathbb{C}^q \to \mathbb{C}^r$ be defined by*

$$P(x_1, \ldots, x_q) = (p_1(x_1, \ldots, x_q), \ldots, p_r(x_1, \ldots, x_q)).$$

Then the algebraic functions $p_1, \ldots, p_r$ are algebraically independent iff the range of the map P is dense in $\mathbb{C}^r$.

PROOF. Assume that the range of the map $P(x_1, \ldots, x_q) = (p_1(x_1, \ldots, x_q), \ldots, p_r(x_1, \ldots, x_q))$ is dense. Any non-trivial polynomial $G(t_1, \ldots, t_r) \in \mathbb{C}[t_1, \ldots, t_r]$ such that

$$G(p_1(x_1, \ldots, x_q), \ldots, p_r(x_1, \ldots, x_q)) = 0.$$

vanishes on a dense set in $\mathbb{C}^r$, and hence such a G is the zero polynomial. We conclude that $p_1, \ldots, p_r$ are algebraically independent.

Now assume that the algebraic functions $p_1, \ldots, p_r$ are algebraically independent. To the algebraically independent set $\{p_1, \ldots, p_r\}$ of elements of $\overline{\mathbb{C}(x_1, \ldots, x_q)}$ add elements $p_{r+1}, \ldots, p_q$ so that $\{p_1, \ldots, p_q\}$ is a basis of the matroid $\mathbb{C}(x_1, \ldots, x_q)$. Thus we have that $\mathbb{C}(x_1, \ldots, x_q) = \mathbb{C}(q_1, \ldots, p_q)$; in other words, the two fields $\mathbb{C}(x_1, \ldots, x_q)$ and $\mathbb{C}(p_1, \ldots, p_q)$ have the same algebraic closure.

By a well known result in field theory we can find $\alpha \in \overline{\mathbb{C}(p_1, \ldots, p_q)}$ so that

$$\mathbb{C}(x_1, \ldots, x_q) \subseteq \mathbb{C}(\alpha, p_1, \ldots, p_q).$$

Since $\alpha \in \overline{\mathbb{C}(p_1, \ldots, p_q)}$ it follows that α is the root of an equation with coefficients in $\mathbb{C}(p_1, \ldots, p_q)$. Such an equation will then be of the form

$$\alpha^n + \frac{f_{n-1}(p_1, \ldots, p_q)}{F(p_1, \ldots, p_q)} \alpha^{n-1} + \cdots + \frac{f_0(p_1, \ldots, p_q)}{F(p_1, \ldots, p_q)} = 0 \tag{6}$$

where $f_0(t_1, \ldots, t_q), \ldots, f_{n-1}(t_1, \ldots, t_q), F(t_1, \ldots, t_q) \in \mathbb{C}[t_1, \ldots, t_q]$.

Since $x_i \in \mathbb{C}(x_1, \ldots, x_q) \subseteq \mathbb{C}(\alpha, p_1, \ldots, p_q)$, the element x_i is as a rational function in α and $p_1, \ldots, p_q$. But since α is algebraic over the field $\mathbb{C}(p_1, \ldots, p_q)$, we can write x_i in the form

$$x_i = \frac{g_i(\alpha, p_1, \ldots, p_q)}{G(p_1, \ldots, p_q)}, \tag{7}$$

where $g_i(t, t_1, \ldots, t_q) \in \mathbb{C}[t, t_1, \ldots, t_q]$ and $G(t_1, \ldots, t_q) \in \mathbb{C}[t_1, \ldots, t_q]$. We may assume that the degree of the variable t in the polynomial g_i is at most $n - 1$. Note that the denominators in identity (7) are independent of i.

Let

$$\hat{D} = \{(y_1, \ldots, y_q) \in \mathbb{C}^q : F(y_1, \ldots, y_q) \neq 0, G(y_1, \ldots, y_q) \neq 0\}.$$

Clearly, $\hat{D}$ is a dense subset of $\mathbb{C}^q$. We claim that $\hat{D}$ is a subset of the range of the map

$$\hat{P}(x_1, \ldots, x_q) = (p_1(x_1, \ldots, x_q), \ldots, p_q(x_1, \ldots, x_q)).$$

Since, for $(y_1, \ldots, y_q) \in \hat{D}$, we have $F(y_1, \ldots, y_q) \neq 0$, we can solve in equation (6) for α. For all such roots α of equation (6) and for all values $y_1, \ldots, y_q$, such that $G(y_1, \ldots, y_q) \neq 0$, we can find a value of x_i by equation (7). This proves that such a q-tuple $(y_1, \ldots, y_q)$ lies in the range of the map $\hat{P}$.

Let

$$D = \{(y_1, \ldots, y_r) \in \mathbb{C}^r : (y_1, \ldots, y_q) \in \hat{D} \text{ for some } y_{r+1}, \ldots, y_q \in \mathbb{C}\}.$$

Since D is desne in $\mathbb{C}^r$, and D is contained in the range of the map

$$P(x_1, \ldots, x_q) = (p_1(x_1, \ldots, x_q), \ldots, p_r(x_1, \ldots, x_q)),$$

the result of the theorem follows. □

3. BASIC THEORM OF q-ARY p-ICS

3.1. *The space of polynomials and its dual space.* Let $V = \mathbb{C}[x_1, \ldots, x_q]$. Let V_p be the subspace of V consisting of all homogeneous polynomials of of degree p.

Following British nineteenth-century custom, an element of V_p will be called a q-ary p-ic. A basis for V_p is given by the monomials

$$x_1^{i_1} \cdots x_q^{i_q},$$

where $i_1 + \cdots + i_q = p$ and $i_1, \ldots, i_q$ are non-negative integers. The dimension of V_p is

$$\dim(V_p) = \frac{q \cdot (q+1) \cdots (q+p-1)}{p \cdot (p-1) \cdots 1} = \binom{q+p-1}{p}.$$

This quantity equals the number of ways in which to choose a multiset with p elements from a q-set, and (following Comtet) we will denote it by $\left\langle {q \atop p} \right\rangle$. Thus, $\dim(V_p) = \left\langle {q \atop p} \right\rangle$.

For non-negative integers $i_1, \ldots, i_q$, define the multinomial coefficient as

$$\binom{p}{i_1, \ldots, i_q} = \frac{p!}{i_1! \cdots i_q!}$$

if $i_1 + \cdots + i_q = p$ and 0 otherwise.

A typical element of V_p may written in the form

$$f(x_1, \ldots, x_q) = \sum_{i_1 + \cdots + i_q = p} \binom{p}{i_1 \cdots i_q} a_{i_1, \ldots, i_q} x_1^{i_1} \cdots x_q^{i_q},$$

where we shall tacitly assume that $a_{i_1, \ldots, i_q} = 0$ if $i_1 + \cdots i_q \neq p$.

In order to avoid an excess of indices, we use some standard abbreviations, as follows:

$$\mathbf{x} = (x_1, \ldots, x_q), \qquad f(\mathbf{x}) = f(x_1, \ldots, x_q), \qquad \mathbf{i} = (i_1, \ldots, i_q),$$
$$\mathbf{x}^{\mathbf{i}} = x_1^{i_1} \cdots x_q^{i_q}, \qquad a_{\mathbf{i}} = a_{i_1, \ldots, i_q}, \qquad \mathbf{i}! = i_1! \cdots i_q!, \qquad |\mathbf{i}| = i_1 + \cdots + i_q,$$
$$\binom{p}{\mathbf{i}} = \binom{p}{i_1, \ldots, i_q}.$$

A q-tuple $\mathbf{i} \in \mathbb{N}^q$ is called a multi-index.

In this notation, we may write a q-ary p-ic $f(\mathbf{x})$ in the form

$$f(\mathbf{x}) = \sum_{|\mathbf{i}| = p} \binom{p}{\mathbf{i}} a_{\mathbf{i}} \cdot \mathbf{x}^{\mathbf{i}}.$$

Notice that

$$V = \bigoplus_{p \geq 0} V_p$$

and V is a graded algebra, as is well known.

Similarly, let $V^* = \mathbb{C}[u_1, \ldots, u_q]$ and let V_p^* be the space of all homogeneous polynomials of degree p. An element of V_p^* will be called a dual q-ary p-ic. As before, we have $\dim(V_p^*) = \left\langle {q \atop p} \right\rangle$, so that

$$V^* = \bigoplus_{p \geq 0} V_p^*.$$

A typical element of V_p^* is written as

$$g(\mathbf{u}) = \sum_{\mathbf{j}} \binom{p}{\mathbf{j}} b_{\mathbf{j}} \mathbf{u}^{\mathbf{j}}.$$

From now on, the number of variables will be tacitly assumed to be q.

We define a bilinear form $\langle \cdot \mid \cdot \rangle \colon V^* \times V \to \mathbb{C}$, by setting

$$\langle \mathbf{u}^{\mathbf{j}} \mid \mathbf{x}^{\mathbf{i}} \rangle = \mathbf{i}! \cdot \delta_{\mathbf{ij}}$$

and extending by linearity. This bilinear form is called the *apolar form.*

Given a linear functional $L: V_p \to \mathbb{C}$ there exists an unique element $g(\mathbf{u}) \in V_p^*$ such that $L(f(\mathbf{x})) = \langle g(\mathbf{u}) \,|\, f(\mathbf{x}) \rangle$ for all $f(\mathbf{x}) \in V_p$. Thus, the vector space V_p^* is natural isomorphic to the dual space of the vector space V_p.

For vectors $\mathbf{c} = (c_1, \ldots, c_q)$ and $\mathbf{d} = (d_1, \ldots, d_q)$, denote their scalar product by $(\mathbf{c} \,|\, \mathbf{d}) = c_1 d_1 + \cdots + c_q d_q$. We note that the form $(\mathbf{c} \,|\, \mathbf{x})^p = (c_1 x_1 + \cdots + c_q x_q)^p$ is a q-ary p-ic. We shall call such a q-ary p-ic *the pth power of a linear form.* Similarly, $(\mathbf{c} \,|\, \mathbf{u})^p$ is called *the pth power of a dual linear form.*

The following result is fundamental.

PROPOSITION 3.1. *For every vector* $\mathbf{c}$ *and for every* q*-ary* p*-ic* $f(\mathbf{x}) \in V_p$ *we have*

$$\langle (\mathbf{c} \,|\, \mathbf{u})^p \,|\, f(\mathbf{x}) \rangle = p! \cdot f(\mathbf{c}).$$

Similarly, for every vector $\mathbf{c}$ *and for every dual* q*-ary* p*-ic* $g(\mathbf{u}) \in V_p^*$ *we have*

$$\langle g(\mathbf{u}) \,|\, (\mathbf{c} \,|\, \mathbf{x})^p \rangle = p! \cdot g(\mathbf{c}).$$

PROOF. The proof is by direct verification:

$$\begin{aligned} \langle (\mathbf{c} \,|\, \mathbf{u})^p \,|\, f(\mathbf{x}) \rangle &= \left\langle \sum_{\mathbf{i}} \binom{p}{\mathbf{i}} \mathbf{c}^{\mathbf{i}} \mathbf{u}^{\mathbf{i}} \,\middle|\, \sum_{\mathbf{i}} \binom{p}{\mathbf{i}} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} \right\rangle \\ &= \sum_{\mathbf{i}} \binom{p}{\mathbf{i}}^2 a_{\mathbf{i}} \mathbf{c}^{\mathbf{i}} \mathbf{i}! \\ &= p! \cdot \sum_{\mathbf{i}} \binom{p}{\mathbf{i}} a_{\mathbf{i}} \mathbf{c}^{\mathbf{i}} \\ &= p! \cdot f(\mathbf{c}). \end{aligned}$$

□

3.2. *Invariance of the apolar form.* A $q \times q$ matrix A defines a linear endomorphism on V_p by substitution. In other words, given $f \in V_p$, define $\hat{A}f \in V_p$ by setting

$$\hat{A}(f(x_1, \ldots, x_q)) = f\left(\sum_{k_1=1}^{q} a_{1,k_1} x_{k_1}, \ldots, \sum_{k_q=1}^{q} q_{1,k_q} x_{k_q} \right),$$

or, in vector notation,

$$\hat{A}f(\mathbf{x}) = f(A\mathbf{x}).$$

Much like a matrix, A operates on V_p by linear substitution: we can let the matrix A operate on V_p^*. To this end, define $\hat{A}$ by:

$$\hat{A}g(\mathbf{u}) = g(A\mathbf{u}).$$

The *adjoint map* $(\hat{A})^*: V^* \to V^*$ of the linear map $\hat{A}: V_p \leftarrow V_p$ is defined by

$$\langle (\hat{A})^* g \,|\, f \rangle = \langle g \,|\, \hat{A} f \rangle.$$

PROPOSITION 3.2. *For a* $q \times q$ *matrix* A *we have*

$$(\hat{A})^* = \widehat{(A^*)}.$$

where the first asterisk is the adjoint, and the second asterisk is the transpose of the matrix.

PROOF. It is sufficient to verify this fact for basis elements $\mathbf{u}^{\mathbf{j}}$ and $\mathbf{x}^{\mathbf{i}}$. Now,

$$\begin{aligned}\langle (\hat{A})^* \mathbf{u}^{\mathbf{i}} \mid \mathbf{x}^{\mathbf{i}} \rangle &= \langle \mathbf{u}^{\mathbf{j}} \mid \hat{A}\mathbf{x}^{\mathbf{i}} \rangle \\ &= \langle \mathbf{u}^{\mathbf{j}} \mid (A\mathbf{x})^{\mathbf{i}} \rangle \\ &= \left\langle \mathbf{u}^{\mathbf{j}} \,\middle|\, \left(\sum_{k_1=1}^{q} a_{1,k_1} x_{k_1} \right)^{i_1} \cdots \left(\sum_{k_q=1}^{q} a_{q,k_q} x_{k_q} \right)^{i_q} \right\rangle \\ &= \mathbf{j}! \sum_{M} \prod_{l=1}^{q} \binom{i_l}{m_{l,1}, \ldots, m_{l,q}} \prod_{k=1}^{q} a_{l,k}^{m_{l,k}} \\ &= \mathbf{j}! \cdot \mathbf{i}! \sum_{M} \prod_{1 \leq l,k \leq q} \frac{a_{l,k}^{m_{l,k}}}{m_{l,k}!}.\end{aligned}$$

where M ranges over all $q \times q$ non-negative integer matrices such that

$$\sum_{k=1}^{q} m_{l,k} = i_l, \qquad \sum_{l=1}^{q} m_{l,k} = j_k.$$

Interchanging the roles of $\mathbf{x}^{\mathbf{i}}$ and $\mathbf{u}^{\mathbf{j}}$, and taking the transpose of A, we obtain

$$\begin{aligned}\langle (\hat{A})^* \mathbf{u}^{\mathbf{j}} \mid \mathbf{x}^{\mathbf{i}} \rangle &= \langle (A^* \mathbf{u})^{\mathbf{j}} \mid \mathbf{x}^{\mathbf{i}} \rangle \\ &= \langle \widehat{(A^*)} \mathbf{u}^{\mathbf{j}} \mid \mathbf{x}^{\mathbf{i}} \rangle,\end{aligned}$$

as desired. □

COROLLARY 3.1 (invariance of the apolar form). *For a non-singular $q \times q$ matrix A, for every $f(\mathbf{x}) \in V_p$ and for every $g(\mathbf{u}) \in V_p^*$ we have*

$$\langle \widehat{A^{*-1}} g(\mathbf{u}) \mid \hat{A} f(\mathbf{x}) \rangle = \langle g(\mathbf{u}) \mid f(\mathbf{x}) \rangle.$$

PROOF.

$$\begin{aligned}\langle g(\mathbf{u}) \mid f(\mathbf{x}) \rangle &= \langle g(\mathbf{u}) \mid \widehat{A^{-1} \cdot A} f(\mathbf{x}) \rangle \\ &= \langle g(\mathbf{u}) \mid \widehat{A^{-1}} \cdot \hat{A} f(\mathbf{x}) \rangle \\ &= \langle \widehat{A^{*-1}} g(\mathbf{u}) \mid \hat{A} f(\mathbf{x}) \rangle.\end{aligned}$$

□

Note that the action of A^{*-1} is the *contragredient action* of A.

3.3. *Apolarity and polarization.* Our long-awaited definition is the following:

DEFINITION 3.1. Let $f(\mathbf{x})$ be a form of degree r, and let $g(\mathbf{u})$ be a dual form of degree p. If $r \leq p$ then $f(\mathbf{x})$ is apolar to $g(\mathbf{u})$ whenever

$$\langle g(\mathbf{u}) \mid h(\mathbf{x}) \cdot f(\mathbf{x}) \rangle = 0$$

for all forms $h(\mathbf{x})$ of degree $p - r$. If $r \geq p$ then $f(\mathbf{x})$ is apolar to $g(\mathbf{u})$ whenever

$$\langle h(\mathbf{u}) \cdot g(\mathbf{u}) \mid f(\mathbf{x}) \rangle = 0$$

for all dual forms $h(\mathbf{u})$ of degree $r - p$.

DEFINITION 3.2. For $\mathbf{c} = (c_1, \ldots, c_q) \in \mathbb{C}^q$, define the polarization operator $D_{\mathbf{c},\mathbf{x}}$ as

$$D_{\mathbf{c},\mathbf{x}} = c_1 \frac{\partial}{\partial x_1} + \ldots + c_q \frac{\partial}{\partial x_q}.$$

PROPOSITION 3.3. *The polarization operator $D_{\mathbf{c},\mathbf{x}}$ is a derivation: that is,*
(1) *$D_{\mathbf{c},\mathbf{x}}$ is linear,*
(2) *$D_{\mathbf{c},\mathbf{x}}(1) = 0$, and*
(3) *$D_{\mathbf{c},\mathbf{x}}(f(\mathbf{x}) \cdot g(\mathbf{x})) = D_{\mathbf{c},\mathbf{x}}(f(\mathbf{x})) \cdot g(\mathbf{x}) + f(\mathbf{x}) \cdot D_{\mathbf{c},\mathbf{x}}(g(\mathbf{x}))$.*
Moreover, $D_{\mathbf{c},\mathbf{x}}$ takes a p-form into a $(p-1)$-form.

Notice that for $\mathbf{d} = (d_1, \ldots, d_q) \in \mathbb{C}^q$, we have

$$\begin{aligned} D_{\mathbf{c},\mathbf{x}}(\mathbf{d} \mid \mathbf{x})^p &= \sum_{i=1}^{q} c_i \frac{\partial}{\partial x_i} (d_1 x_1 + \ldots + d_q x_q)^p \\ &= \sum_{i=1}^{q} c_i d_i p (d_1 x_1 + \ldots + d_q x_q)^{p-1} \\ &= p(\mathbf{d} \mid \mathbf{c})(\mathbf{d} \mid \mathbf{x})^{p-1}. \end{aligned}$$

PROPOSITION 3.4. *Let $f(\mathbf{x})$ be a form of degree $p-1$, and let $g(\mathbf{u})$ be a dual form of degree p. Then*

$$\langle g(\mathbf{u}) \mid (\mathbf{c} \mid \mathbf{x}) f(\mathbf{x}) \rangle = \langle D_{\mathbf{c},\mathbf{u}} g(\mathbf{u}) \mid f(\mathbf{x}) \rangle.$$

PROOF. Notice that the above identity is linear in $f(\mathbf{x})$, $g(\mathbf{u})$ and $\mathbf{c}$. Hence we need only to prove it for $f(\mathbf{x}) = \mathbf{x}^{\mathbf{i}}$, $g(\mathbf{u}) = \mathbf{u}^{\mathbf{j}}$ and $\mathbf{c} = \mathbf{e}_r$, the rth unit vector (that is, $e_r = (\delta_{r,1}, \ldots, \delta_{r,q})$). Hence

$$\begin{aligned} \langle \mathbf{u}^{\mathbf{j}} \mid (\mathbf{e}_r \mid \mathbf{x}) \mathbf{x}^{\mathbf{i}} \rangle &= \langle \mathbf{u}^{\mathbf{j}} \mid x_r \mathbf{x}^{\mathbf{i}} \rangle \\ &= \langle \mathbf{u}^{\mathbf{j}} \mid \mathbf{x}^{\mathbf{i}+\mathbf{e}_r} \rangle \\ &= \mathbf{j}!\, \delta_{\mathbf{j},\mathbf{i}+\mathbf{e}_r} \\ &= j_r \mathbf{i}!\, \delta_{\mathbf{j}-\mathbf{e}_r,\mathbf{i}} \\ &= \langle j_r \mathbf{u}^{\mathbf{i}-\mathbf{e}_r} \mid \mathbf{x}^{\mathbf{i}} \rangle \\ &= \langle D_{\mathbf{e}_r,\mathbf{u}} \mathbf{u}^{\mathbf{j}} \mid \mathbf{x}^{\mathbf{i}} \rangle, \end{aligned}$$

and the proof is complete. □

COROLLARY 3.2. *Let $f(\mathbf{x})$ be a form of degree r, and let $g(\mathbf{u})$ be a dual form of degree p, where $r \leq p$. Then $f(\mathbf{x})$ and $g(\mathbf{u})$ are apolar iff*

$$\langle D_{\mathbf{c}_{p-r},\mathbf{u}} \cdots D_{\mathbf{c}_1,\mathbf{u}} g(\mathbf{u}) \mid f(\mathbf{x}) \rangle = 0$$

for all $\mathbf{c}_1, \ldots, \mathbf{c}_{p-r} \in \mathbb{C}^q$.

PROOF. Notice that V_{p-r} is linearly spanned by polynomials on the form $(\mathbf{c}_1 \mid \mathbf{x}) \cdots (\mathbf{c}_{p-r} \mid \mathbf{x})$. By Proposition 3.4 we have

$$\langle g(\mathbf{u}) \mid (\mathbf{c}_1 \mid \mathbf{x}) \cdots (\mathbf{c}_{p-r} \mid \mathbf{x}) f(\mathbf{x}) \rangle = \langle D_{\mathbf{c}_{p-r},\mathbf{u}} \cdots D_{\mathbf{c}_1,\mathbf{u}} g(\mathbf{u}) \mid \mathrm{f}(\boldsymbol{x}) \rangle$$

and the result follows. □

4. CANONICAL FORMS

4.1. *First main theorem on apolarity*

DEFINITION 4.1. We shall say that a statement holds for a generic q-ary p-ic $f(\mathbf{x})$ whenever there exists a dense subset in the space of all q-ary p-ics (in the Euclidean topology) such the statement holds for all q-ary p-ics in such a dense set.

Our objective is the study of canonical forms for homogeneous polynomials in q variables. We begin with Waring's problem; that is, the problem of expressing a q-ary p-ic as a sum of powers of linear forms.

THEOREM 4.1. *A generic q-ary p-ic $f(\mathbf{x})$ can be written in the form*

$$f(\mathbf{x}) = (\mathbf{a} \mid \mathbf{x})^p + (\mathbf{b} \mid \mathbf{x})^p + \ldots + (\mathbf{c} \mid \mathbf{x})^p$$

iff there exist vectors $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ such that there is no non-zero dual q-ary p-ic $g(\mathbf{u})$ apolar to all the forms $(\mathbf{a}' \mid \mathbf{x})^{p-1}$, $(\mathbf{b}' \mid \mathbf{x})^{p-1}, \ldots, (\mathbf{c}' \mid \mathbf{x})^{p-1}$.

PROOF. Assume that a generic q-ary p-ic $f(\mathbf{x}) = \sum_{\mathbf{i}} e_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$ can be written in the above form: that is, assume that, for a generic $f(\mathbf{x})$,

$$f(\mathbf{x}) = (\mathbf{a} \mid \mathbf{x})^p + (\mathbf{b} \mid \mathbf{x})^p + \ldots + (\mathbf{c} \mid \mathbf{x})^p. \tag{8}$$

Say that the set $\{\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}\}$ is of size n. On the right-hand side we have nq arbitrary coefficients $a_1, \ldots, a_q, b_1, \ldots, c_q$. On the left-hand side we have $\left\langle {q \atop p} \right\rangle$ arbitrary coefficients $e_{\mathbf{i}}$, where $\mathbf{i} = (i_1, \ldots, i_q)$ is a multi-index and $i_1 + \cdots + i_q = p$. Comparing the number of coefficients on both sides we obtain the inequality

$$\left\langle {q \atop p} \right\rangle \leqslant nq.$$

Comparing the coefficients on both sides of equation (8) we obtain $\left\langle {q \atop p} \right\rangle$ identities to be satisfied:

$$e_{\mathbf{i}} = \phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}).$$

Note that $\phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c})$ is a polynomial in nq variables.

We can write identity (8) as

$$f(\mathbf{x}) = \sum_{\mathbf{i}:\, |\mathbf{i}| = p} \phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}) \mathbf{x}^{\mathbf{i}}.$$

The $\left\langle {q \atop p} \right\rangle$ polynomials $\phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c})$ define a map $\boldsymbol{\Phi}$ from $\mathbb{C}^{nq}$ to $\mathbb{C}^{\left\langle {q \atop p} \right\rangle}$ by

$$\boldsymbol{\Phi}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}) = (\phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}))_{\mathbf{i}:\, |\mathbf{i}| = p}$$

where the co-ordinates of $\mathbb{C}^{\left\langle {q \atop p} \right\rangle}$ are indexed by the multi-index $\mathbf{i} = (i_1, \ldots, i_q)$, with $i_1 + \cdots + i_q = p$. Since we assume that we can write a generic $f(\mathbf{x})$ in the form described by equation (8), the range of the map $\boldsymbol{\Phi}$ is dense. By Theorem 2.4 we conclude that the $\left\langle {q \atop p} \right\rangle$ polynomials $\phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c})$ are algebraically independent.

Since $\left\langle {q \atop p} \right\rangle \leqslant nq$, Theorem 2.3 tells us that the statement that $\phi_{\mathbf{i}}$ are algebraically independent is equivalent to the statement that the matrix

$$\left(\frac{\partial \phi_{\mathbf{i}}}{\partial a_1}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial a_q}, \frac{\partial \phi_{\mathbf{i}}}{\partial b_1}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial c_q} \right)_{\mathbf{i}:\, |\mathbf{i}| = p}, \tag{9}$$

has full rank, where the rows are indexed by the non-negative integer vector $\mathbf{i}$.

We can then find values for $\mathbf{a} = (a_1, \ldots, a_q)$, $\mathbf{b} = (b_1, \ldots, b_q), \ldots, \mathbf{c} = (c_1, \ldots, c_q)$ such that the matrix (9) has full rank. Denote these actual values by $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$.

Hence the columns of the matrix

$$\left(\frac{\partial \phi_{\mathbf{i}}}{\partial a_1'}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial a_q'}, \frac{\partial \phi_{\mathbf{i}}}{\partial b_1'}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial c_q'} \right)_{\mathbf{i}:\, |\mathbf{i}| = p} \tag{10}$$

span the the space $\mathbb{C}^{\langle q \rangle}$. Recall that the co-ordinates of the space $\mathbb{C}^{\langle q \rangle}$ are indexed by the multi-index $\mathbf{i}$ such that $|\mathbf{i}| = p$. But the space $\mathbb{C}^{\langle q \rangle}$ is isomorphic to linear space V_p, by the natural isomorphism

$$(e_{\mathbf{i}})_{\mathbf{i}: |\mathbf{i}|=p} \to \sum_{\mathbf{i}: |\mathbf{i}|=p} e_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}.$$

Under this isomorphism, the vector $(\partial \phi_{\mathbf{i}} / \partial a_1)_{\mathbf{i}}$ is mapped into a q-ary p-ic:

$$\left(\frac{\partial \phi_{\mathbf{i}}}{\partial a_1'} \right)_{\mathbf{i}: |\mathbf{i}|=p} \to \sum_{\mathbf{i}: |\mathbf{i}|=p} \frac{\partial \phi_{\mathbf{i}}}{\partial a_1'} \mathbf{x}^{\mathbf{i}} = \frac{\partial f(\mathbf{x})}{\partial a_1'},$$

and similarly for $a_2', \ldots, c_q'$.

Because of this isomorphism, the matrix (10) has full rank iff the nq q-ary p-ics

$$\frac{\partial f(\mathbf{x})}{\partial a_1'}, \ldots, \frac{\partial f(\mathbf{x})}{\partial a_q'}, \frac{\partial f(\mathbf{x})}{\partial b_1'}, \ldots, \frac{\partial f(\mathbf{x})}{\partial c_q'}$$

span the entire space V_p.

However, note that

$$\partial f(\mathbf{x}) / \partial a_1' = p(\mathbf{a}' \mid \mathbf{x})^{p-1} x_1$$

and similarly for $a_2', \ldots, c_q'$. Hence the q-ary p-ics $(\mathbf{a}' \mid \mathbf{x})^{p-1} x_1, \ldots, (\mathbf{a}' \mid \mathbf{x})^{p-1} x_q$, $(\mathbf{b}' \mid \mathbf{x})^{p-1} x_1, \ldots, (\mathbf{c}' \mid \mathbf{x})^{p-1} x_q$ span the space V_p.

Recall from linear algebra that a set of elements in a linear space spans iff the zero functional is the only linear functional that sends such a set to 0. Hence the only linear functional $L: V_p \to \mathbb{C}$ such that

$$L((\mathbf{a}' \mid \mathbf{x})^{p-1} x_1) = \cdots = L((\mathbf{c}' \mid \mathbf{x})^{p-1} x_q) = 0$$

is the zero functional. But the space V_p^* is a representation of the dual space of V_p, as observed in Section 3.1.

Thus, if any dual q-ary p-ic $g(\mathbf{u})$ satisfies

$$\langle g(\mathbf{u}) \mid (\mathbf{a}' \mid \mathbf{x})^{p-1} x_1 \rangle = \cdots = \langle g(\mathbf{u}) \mid (\mathbf{c}' \mid \mathbf{x})^{p-1} x_q \rangle = 0$$

then $g(\mathbf{u}) = 0$.

The conditions that

$$\langle g(\mathbf{u}) \mid (\mathbf{a}' \mid \mathbf{x})^{p-1} x_1 \rangle = \langle g(\mathbf{u}) \mid (\mathbf{a}' \mid \mathbf{x})^{p-1} x_2 \rangle = \cdots = \langle g(\mathbf{u}) \mid (\mathbf{a}' \mid \mathbf{x})^{p-1} x_q \rangle = 0$$

are precisely the conditions that $g(\mathbf{u})$ be apolar to $(\mathbf{a}' \mid \mathbf{x})^{p-1}$. Hence if there exists a dual q-ary p-ic $g(\mathbf{u})$ so that $g(\mathbf{u})$ is apolar to each $(\mathbf{a}' \mid \mathbf{x})^{p-1}$, $(\mathbf{b}' \mid \mathbf{x})^{p-1}, \ldots, (\mathbf{c}' \mid \mathbf{x})^{p-1}$, then $g(\mathbf{u})$ is zero. This proves half the theorem.

To prove the other half, notice that we did prove a sequence of equivalent statements in the above argument. Assume that there exist vectors $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ such that there is no non-zero dual q-ary p-ic $g(\mathbf{u})$ apolar to all the forms $(\mathbf{a}' \mid \mathbf{x})^{p-1}$, $(\mathbf{b}' \mid \mathbf{x})^{p-1}, \ldots, (\mathbf{c}' \mid \mathbf{x})^{p-1}$. As proved above, the assumption is equivalent to the assumption that the matrix

$$\left(\frac{\partial \phi_{\mathbf{i}}}{\partial a_1'}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial a_q'}, \frac{\partial \phi_{\mathbf{i}}}{\partial b_1'}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial c_q'} \right)_{\mathbf{i}: |\mathbf{i}|=p} \tag{11}$$

has full rank, where the rows are indexed by a multi-index $\mathbf{i}$, such that $|\mathbf{i}| = p$.

The assumption that the matrix (11) has full rank in turn implies that there is a non-zero minor of size $\langle {q \atop p} \rangle$. This minor has to remain non-zero when, instead of

specific values $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$, we use variables $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}$. Thus the matrix

$$\left(\frac{\partial \phi_{\mathbf{i}}}{\partial a_1}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial a_q}, \frac{\partial \phi_{\mathbf{i}}}{\partial b_1}, \ldots, \frac{\partial \phi_{\mathbf{i}}}{\partial c_q}\right)_{\mathbf{i}:\,|\mathbf{i}|=p} \tag{12}$$

has a non-zero minor of size $\langle {q \atop p} \rangle$, and the matrix (12) has full rank.

Since the matrix (12) has full rank, by Theorem 2.3 the $\langle {q \atop p} \rangle$ polynomials $\phi_{\mathbf{i}}$ are algebraically independent. By Theorem 2.4 we conclude that the range of the map

$$\boldsymbol{\Phi}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}) = (\phi_{\mathbf{i}}, (\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}))_{\mathbf{i}:\,|\mathbf{i}|=p}$$

is dense in the space $\mathbb{C}^{\langle {q \atop p} \rangle}$, where the co-ordinates of the space $\mathbb{C}^{\langle {q \atop p} \rangle}$ are indexed by a non-negative integer vector $\mathbf{i}$, such that $|\mathbf{i}| = p$.

Recall that

$$f(\mathbf{x}) = \sum_{\mathbf{i}:\,|\mathbf{i}|=p} \phi_{\mathbf{i}}(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c})\mathbf{x}^{\mathbf{i}}.$$

This identity can be viewed as a map from $\mathbb{C}^{nq}$ to V_p. By the naturally isomorphism between the linear spaces V_p and $\mathbb{C}^{\langle {q \atop p} \rangle}$, we know that the range of this map is dense in V_p. Thus we have proven that a generic $f(\mathbf{x}) \in V_p$ can be written in the form

$$f(\mathbf{x}) = (\mathbf{a} \mid \mathbf{x})^p + (\mathbf{b} \mid \mathbf{x})^p + \ldots + (\mathbf{c} \mid \mathbf{x})^p. \qquad \square$$

4.2. *Applications of the first main theorem.* We begin by giving an exeedingly simple proof of a classical result.

COROLLARY 4.1. *A generic q-ary quadratic form can be written as the sum of q squares.*

PROOF. Choose the vectors $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ in the statement of Theorem 4.1 to be the unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q$. Suppose that there exists a dual q-ary quadratic form $g(\mathbf{u})$ apolar to each of the linear forms $(\mathbf{e}_i \mid \mathbf{x}) = x_i$ for $i = 1, \ldots, q$. Let

$$g(\mathbf{u}) = \sum_{1 \leqslant k \leqslant l \leqslant q} c_{k,l} u_k u_l.$$

If $g(\mathbf{u})$ is to be apolar to x_i, then $g(\mathbf{u})$ is also apolar to $x_j x_i$ for all j. If $i \leqslant j$, then

$$\begin{aligned} 0 &= \langle g(\mathbf{u}) \mid x_i x_j \rangle \\ &= \left\langle \sum_{1 \leqslant k \leqslant l \leqslant q} c_{k,l} u_k u_l \,\middle|\, x_i x_j \right\rangle \\ &= \sum_{1 \leqslant k \leqslant l \leqslant q} c_{k,l} \langle u_k u_l \mid x_i x_j \rangle \\ &= c_{i,j} \langle u_i u_j \mid x_i x_j \rangle. \end{aligned}$$

But this implies $c_{i,j} = 0$ for all $i \leqslant j$. Hence $g(\mathbf{u}) = 0$, and the conclusion follows immediately from Theorem 4.1. $\square$

COROLLARY 4.2. *A generic quaternary cubic can be written as the sum of 5 cubes.*

PROOF. Here, $q = 4$, $p = 3$ and $n = 5$. Choose the vectors $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ to be $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$. Assume that a dual quaternary cubic $g(\mathbf{u})$ exists which is apolar to all $(\mathbf{e}_1 \mid \mathbf{x})^2$, $(\mathbf{e}_2 \mid \mathbf{x})^2$, $(\mathbf{e}_3 \mid \mathbf{x})^2$, $(\mathbf{e}_4 \mid \mathbf{x})^2$ and $(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4 \mid \mathbf{x})^2$. Note that $(\mathbf{e}_i \mid \mathbf{x})^2 = x_i^2$. If $g(\mathbf{u})$ is to be apolar to x_i^2, then $g(\mathbf{u})$ is also apolar to $x_j x_i^2$ for all j. Any term of $g(\mathbf{u})$ containing a square term, say $u_j u_i^2$ must have coefficient equal to 0,

because $g(\mathbf{u})$ is apolar $x_j x_i^2$. Hence we can write $g(\mathbf{u})$ in the form

$$g(\mathbf{u}) = c_1 u_2 u_3 u_4 + c_2 u_1 u_3 u_4 + c_3 u_1 u_2 u_4 + c_4 u_1 u_2 u_3.$$

But $g(\mathbf{u})$ is also apolar to $x_j(x_1 + x_2 + x_3 + x_4)^2$. This gives the equations $c_1 + c_2 + c_3 + c_4 - c_j = 0$, the only solution of which is $c_j = 0$. Thus, $g(\mathbf{u}) = 0$, and the conclusion follows immediately from Theorem 4.1. □

COROLLARY 4.3. *A generic ternary quintic can be written as the sum of 7 fifth powers.*

PROOF. Here $q = 3$, $p = 5$ and $n = 7$. Choose the vectors $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ to be $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$. Assume that a dual ternary quintic $g(\mathbf{u})$ exists which is apolar to $(\mathbf{e}_1 \mid \mathbf{x})^4$, $(\mathbf{e}_2 \mid \mathbf{x})^4$, $(\mathbf{e}_3 \mid \mathbf{x})^4$, $(\mathbf{e}_1 + \mathbf{e}_2 \mid \mathbf{x})^4$, $(\mathbf{e}_1 + \mathbf{e}_3 \mid \mathbf{x})^4$, $(\mathbf{e}_2 + \mathbf{e}_3 \mid \mathbf{x})^4$ and $(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 \mid \mathbf{x})^4$. Since $g(\mathbf{u})$ is apolar to $(\mathbf{e}_i \mid \mathbf{x})^4 = x_i^4$, the coefficients of $u_j u_i^4$ in $g(\mathbf{u})$ are zero. Now $g(\mathbf{u})$ is apolar to $x_i(x_i + x_j)^4$ and to $x_j(x_i + x_j)^4$, hence the coefficients of $u_i^2 u_j^3$ and $u_i^3 u_j^2$ are zero. We can therefore write $g(\mathbf{u})$ in the form

$$g(\mathbf{u}) = c_{3,1,1} u_1^3 u_2 u_3 + c_{1,3,1} u_1 u_2^3 u_3 + c_{1,1,3} u_1 u_2 u_3^3 + c_{2,2,1} u_1^2 u_2^2 u_3 + c_{2,1,2} u_1^2 u_2 u_3^2 + c_{1,2,2} u_1 u_2^2 u_3^2.$$

Apolarity with $x_3(x_1 + x_2)^4$ gives the equation:

$$6 \cdot 4c_{3,1,1} + 4 \cdot 6c_{2,2,1} + 6 \cdot 4c_{1,3,1} = 0.$$

Apolarity with $x_3(x_1 + x_2 + x_3)^4$ gives the equation:

$$6 \cdot 12c_{1,1,3} + 4 \cdot 12c_{2,1,2} + 4 \cdot 12c_{1,2,2} + 6 \cdot 4c_{3,1,1} + 4 \cdot 6c_{2,2,1} + 6 \cdot 4c_{1,3,1} = 0.$$

Thus we have

$$c_{3,1,1} + c_{2,2,1} + c_{1,3,1} = 0, \qquad 6c_{1,1,3} + 4c_{2,1,2} + 4c_{1,2,2} = 0.$$

By symmetry we obtain 4 more linear equations. Let us solve this linear system of 6 equations and 6 unknowns. From

$$\begin{aligned} 0 &= 6c_{1,1,3} + 4c_{2,1,2} + 4c_{1,2,2} \\ &= 6c_{1,1,3} + 4(-c_{1,1,3} - c_{3,1,1}) + 4(-c_{1,1,3} - c_{1,3,1}) \\ &= -2c_{1,1,3} - 4c_{3,1,1} - 4c_{1,3,1} \end{aligned}$$

it is easy to see that $c_{1,1,3} = c_{1,3,1} = c_{3,1,1} = 0$, which implies that the other coefficients are also 0. Hence $g(\mathbf{u}) = 0$, and the conclusion follows immediately from Theorem 4.1. □

COROLLARY 4.4 (Clebsch). *The generic ternary quartic cannot in general be written as the sum of 5 fourth powers.*

PROOF. Here $q = 3$, $p = 4$ and $n = 5$. By Theorem 4.1, all we need to show is that, given $\mathbf{a}', \mathbf{b}', \mathbf{c}', \mathbf{d}', \mathbf{e}' \in \mathbb{C}^3$, there is a non-trivial dual ternary quartic $g(\mathbf{u})$, which is apolar to $(\mathbf{a}' \mid \mathbf{x})^3$, $(\mathbf{b}' \mid \mathbf{x})^3$, $(\mathbf{c}' \mid \mathbf{x})^3$, $(\mathbf{d}' \mid \mathbf{x})^3$ and $(\mathbf{e}' \mid \mathbf{x})^3$.

Case 1. Among the vectors $\mathbf{a}', \mathbf{b}', \mathbf{c}', \mathbf{d}'$ and $\mathbf{e}'$, no three are linearly independent. Change variables, so that all five vectors will have their last co-ordinate equal to 0. If so, then the dual ternary quartic $g(\mathbf{u}) = u_3^4$ is apolar to each of the five cubes $(\mathbf{a}' \mid \mathbf{x})^3, \ldots, (\mathbf{e}' \mid \mathbf{x})^3$.

Case 2. The three vectors $\mathbf{c}', \mathbf{d}'$ and $\mathbf{e}'$ are linearly independent. Change variables, so that each of these three vectors is one of the unit vectors. Having made this change

of variables, we need to show that for all $\mathbf{a}', \mathbf{b}' \in \mathbb{C}^3$ there is a non-trivial ternary quartic $g(\mathbf{u})$ which is apolar to $(\mathbf{a}' \mid \mathbf{x})^3$, $(\mathbf{b}' \mid \mathbf{x})^3$, x_1^3, x_2^3 and x_3^3. The condition that $g(\mathbf{u})$ be apolar to $x_i x_i^3$ forces the coefficient of $u_i u_i^3$ in $g(\mathbf{u})$ to vanish. Thus we can write $g(\mathbf{u})$ in the form:

$$g(\mathbf{u}) = c_{2,2,0}u_1^2u_2^2 + c_{2,1,1}u_1^2u_2u_3 + c_{2,0,2}u_1^2u_3^2 + c_{1,2,1}u_1u_2^2u_3 + c_{1,1,2}u_1u_2u_3^2 + c_{0,2,2}u_2^2u_3^2.$$

The conditions that $g(\mathbf{u})$ be apolar to $x_1(\mathbf{a}' \mid \mathbf{x})^3$, $x_2(\mathbf{a}' \mid \mathbf{x})^3$, $x_3(\mathbf{a}' \mid \mathbf{x})^3$, $x_1(\mathbf{b}' \mid \mathbf{x})^3$, $x_2(\mathbf{b}' \mid \mathbf{x})^3$ and $x_3(\mathbf{b}' \mid \mathbf{x})^3$ lead to the following system of linear equation:

$$\begin{pmatrix} 3a_1a_2^2 & 6a_1a_2a_3 & 3a_1a_3^2 & 3a_2^2a_3 & 3a_2a_3^2 & 0 \\ 3a_1^2a_2 & 3a_1^2a_3 & 0 & 6a_1a_2a_3 & 3a_1a_3^2 & 3a_2a_3^2 \\ 0 & 3a_1^2a_2 & 3a_1^2a_3 & 3a_1a_2^2 & 6a_1a_2a_3 & 3a_2^2a_3 \\ 3b_1b_2^2 & 6b_1b_2b_3 & 3b_1b_3^2 & 3b_2^2b_3 & 3b_2b_3^2 & 0 \\ 3b_1^2b_2 & 3b_1^2b_3 & 0 & 6b_1b_2b_3 & 3b_1b_3^2 & 3b_2b_3^2 \\ 0 & 3b_1^2b_2 & 3b_1^2b_3 & 3b_1b_2^2 & 6b_1b_2b_3 & 3b_2^2b_3 \end{pmatrix} \cdot \begin{pmatrix} 4c_{2,2,0} \\ 2c_{2,1,1} \\ 4c_{2,0,2} \\ 2c_{1,2,1} \\ 2c_{1,1,2} \\ 4c_{0,2,2} \end{pmatrix} = \mathbf{0}.$$

The determinant of the above matrix equals zero. Hence the matrix is singular and the system has a non-trivial solution. Thus there exists a non-trivial dual ternary quartic $g(\mathbf{u})$.

In both cases we have shown that there exists a non-trivial dual ternary quartic $g(\mathbf{u})$; hence the conclusion by Theorem 4.1. □

COROLLARY 4.5. *The generic quinary cubic cannot in general be written as the sum of 7 cubes.*

PROOF. Here $q = 5$, $p = 3$ and $n = 7$. We proceed as in the previous corollary. We show that for all $\mathbf{a}', \mathbf{b}' \in \mathbb{C}^5$ there is a non-zero dual quinary cubic $g(\mathbf{u})$ apolar to $(\mathbf{a}' \mid \mathbf{x})^2$, $(\mathbf{b}' \mid \mathbf{x})^2$, $x_1^2, \ldots, x_5^2$. But again we see that these apolar conditions force $g(\mathbf{u})$ to be of the form:

$$\begin{aligned} g(\mathbf{u}) = {} & c_{1,1,1,0,0}u_1u_2u_3 + c_{1,1,0,1,0}u_1u_2u_4 + c_{1,1,0,0,1}u_1u_2u_5 + c_{1,0,1,1,0}u_1u_3u_4 \\ & + c_{1,0,1,0,1}u_1u_3u_5 + c_{1,0,0,1,1}u_1u_4u_5 + c_{0,1,1,1,0}u_2u_3u_4 + c_{0,1,1,0,1}u_2u_3u_5 \\ & + c_{0,1,0,1,1}u_2u_4u_5 + c_{0,0,1,1,1}u_3u_4u_5. \end{aligned}$$

The conditions that $g(\mathbf{u})$ be apolar to $(\mathbf{a}' \mid \mathbf{x})^2$ and $(\mathbf{b}' \mid \mathbf{x})^2$ lead to a linear system of 10 equations and 10 unknowns. The matrix of the system is written below. To see that such a system has a non-trivial solution, one verifies that the determinant of the matrix vanishes, as indeed it does:

$$\det \begin{pmatrix} a_2a_3 & a_2a_4 & a_2a_5 & a_3a_4 & a_3a_5 & a_4a_5 & 0 & 0 & 0 & 0 \\ a_1a_3 & a_1a_4 & a_1a_5 & 0 & 0 & 0 & a_3a_4 & a_3a_5 & a_4a_5 & 0 \\ a_1a_2 & 0 & 0 & a_1a_4 & a_1a_5 & 0 & a_2a_4 & a_2a_5 & 0 & a_4a_5 \\ 0 & a_1a_2 & 0 & a_1a_3 & 0 & a_1a_5 & a_2a_3 & 0 & a_2a_5 & a_3a_5 \\ 0 & 0 & a_1a_2 & 0 & a_1a_3 & a_1a_4 & 0 & a_2a_3 & a_2a_4 & a_3a_4 \\ b_2b_3 & b_2b_4 & b_2b_5 & b_3b_4 & b_3b_5 & b_4b_5 & 0 & 0 & 0 & 0 \\ b_1b_3 & b_1b_4 & b_1b_5 & 0 & 0 & 0 & b_3b_4 & b_3b_5 & b_4b_5 & 0 \\ b_1b_2 & 0 & 0 & b_1b_4 & b_1b_5 & 0 & b_2b_4 & b_2b_5 & 0 & b_4b_5 \\ 0 & b_1b_2 & 0 & b_1b_3 & 0 & b_1b_5 & b_2b_3 & 0 & b_2b_5 & b_3b_5 \\ 0 & 0 & b_1b_2 & 0 & b_1b_3 & b_1b_4 & 0 & b_2b_3 & b_2b_4 & b_3b_4 \end{pmatrix} = 0. \quad \square$$

DEFINITION 4.2. A k-fold point, $\mathbf{a}$, of q-ary p-ic $f(x)$, is a non-zero q-tuple $(a_1, \ldots, a_q)$ such that, for all $i = 0, 1, \ldots, k-1$,

$$[D_{\mathbf{c}_1,\mathbf{x}} \cdots D_{\mathbf{c}_i,\mathbf{x}} f(\mathbf{x})]_{\mathbf{x}=a} = 0$$

for all $\mathbf{c}_1, \ldots, \mathbf{c}_i \in \mathbb{C}^q$.

If $k = 1$ then such a point is called a *simple point.* When $k = 2$ the point is said to be a *double point.* Notice that a k-fold point of a binary form is a factor of the form, where the factor is the k-power of a linear form.

PROPOSITION 4.1. *Let $g(\mathbf{u})$ be a dual q-ary p-ic, and let $1 \leqslant k \leqslant p$. Then $g(\mathbf{u})$ is apolar to $(\mathbf{a} \mid \mathbf{x})^{p+1-k}$ iff $\mathbf{a}$ is k-fold point of $g(\mathbf{u})$.*

PROOF. Assume that $g(\boldsymbol{u}) \in V_p^*$ is apolar to $(\mathbf{a} \mid \mathbf{x})^{p+1-k}$. Let $0 \leqslant i \leqslant k-1$. Then $g(\mathbf{u})$ is apolar to $(\mathbf{a} \mid \mathbf{x})^{k-1-i} \cdot (\mathbf{a} \mid \mathbf{x})^{p+1-k} = (\mathbf{a} \mid \mathbf{x})^{p-i}$ and, by Propositions 3.4 and 3.1,

$$\begin{aligned} 0 &= \langle g(\mathbf{u}) \mid (\mathbf{c}_1 \mid \mathbf{x}) \cdots (\mathbf{c}_i \mid \mathbf{x})(\mathbf{a} \mid \mathbf{x})^{p-i} \rangle \\ &= \langle D_{\mathbf{c}_i,\mathbf{u}} \cdots D_{\mathbf{c}_1,\mathbf{u}} g(\mathbf{u}) \mid (\mathbf{a} \mid \mathbf{x})^{p-i} \rangle \\ &= (p-i)! \cdot [D_{\mathbf{c}_i,\mathbf{u}} \cdots D_{\mathbf{c}_1,\mathbf{u}} g(\mathbf{u})]_{\mathbf{u}=\mathbf{a}}. \end{aligned}$$

Hence we conclude that $\mathbf{a}$ is a k-fold point on $g(\mathbf{u})$.

By tracing the above string of equalities in the reverse direction, we prove the other direction of the proposition. □

By combining Theorem 4.1 and Proposition 4.1 we can state the first main theorem on apolarity, as follows:

THEOREM 4.2. *A generic q-ary p-ic $f(\mathbf{x})$ can be written in the form*

$$f(\mathbf{x}) = (\mathbf{a} \mid \mathbf{x})^p + (\mathbf{b} \mid \mathbf{x})^p + \ldots + (\mathbf{c} \mid \mathbf{x})^p$$

iff there exist $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ so that there does not exist a non-zero dual q-ary p-ic $g(\mathbf{u})$ which has $\mathbf{a}', \mathbf{b}', \ldots, \mathbf{c}'$ as double points.

COROLLARY 4.6 (Clebsch). *The generic ternary quartic cannot in general be written as the sum of 5 fourth powers.*

PROOF. Here $q = 3$, $p = 4$ and $n = 5$. We must check that for all vectors $\mathbf{a}', \mathbf{b}', \mathbf{c}', \mathbf{d}', \mathbf{e}' \in \mathbb{C}$ that there exists a non-trivial dual ternary quartic $g(\mathbf{u})$, which has $\mathbf{a}', \mathbf{b}', \mathbf{c}', \mathbf{d}'$ and $\mathbf{e}'$ as double points.

The vectors $\mathbf{a}', \ldots, \mathbf{e}'$ are the homogeneous co-ordinates of points in the projective plane. Through 5 points in the projective plane there is at least one conic section; in other words, there exists a non-zero ternary quadratic $h(\mathbf{u})$, such that $h(\mathbf{a}') = \cdots = h(\mathbf{e}') = 0$. Let $g(\mathbf{u}) = h(\mathbf{u})^2$. It is easy that $\mathbf{a}', \ldots, \mathbf{e}'$ are double points of $g(\mathbf{u})$, and the desired result is obtained. □

THEOREM 4.3 (Sylvester). *A generic binary $(2j-1)$-ic form can be written as a sum of the $(2j-1)$st powers of j linear forms.*

PROOF. Choose j distinct linear forms. These j linear forms correspond to j points in the plane. Assume that there is a non-zero dual binary form of degree $2j-1$, which has all these points as double points. Each double point is to a square factor of the binary

form. Hence we have j square factors of the binary form $g(\mathbf{u})$. But the product of j square factors has degree $2j$, which is larger than the degree $2j-1$ of $g(\mathbf{u})$. Now use the statement of Theorem 4.2 to obtain the desired conclusion. □

4.3. *Second main theorem on apolarity.* Let $d_1, \ldots, d_s$ be non-negative integers and let J be a set of multi-indices; that is, $J \subseteq \mathbb{N}^s$. For each $\mathbf{j} \in J$ let $t_{\mathbf{j}}(\mathbf{x})$ be a homogeneous polynomial in the variables $x_1, \ldots, x_q$. Assume that for all $\mathbf{j} \in J$

$$j_1 d_1 + \ldots + j_s d_s + \deg(t_{\mathbf{j}}(\mathbf{x})) = p.$$

The problem we solve is the following: When can one write a generic q-ary p-ic $f(\mathbf{x})$ in the form

$$f(\mathbf{x}) = \sum_{\mathbf{j} \in J} t_{\mathbf{j}}(\mathbf{x}) h_1(\mathbf{x})^{j_1} \cdots h_s(\mathbf{x})^{j_s} = \sum_{\mathbf{j} \in J} t_{\mathbf{j}}(\mathbf{x}) h^{\mathbf{j}}(\mathbf{x}),$$

where h_i is a homogeneous polynomial of degree d_i for $i = 1, \ldots, s$?

THEOREM 4.4. *A generic q-ary p-ic $f(\mathbf{x})$ can be written in the form $f(\mathbf{x}) = \sum_{\mathbf{j} \in J} t_{\mathbf{j}}(\mathbf{x}) h^{\mathbf{j}}(\mathbf{x})$ for some $h_1, \ldots, h_s$ iff there exist $h_1(\mathbf{x}), \ldots, h_s(\mathbf{x})$ so that there is no non-zero dual q-ary p-ic which is apolar to all the forms $\partial f / \partial h_i$, $1 \leqslant i \leqslant s$.*

PROOF. Call the coefficients of the polynomials h_i *parameters*. We denote a parameter by *par* and we let **Par** be the set of parameters. Notice that

$$|\mathbf{Par}| = \left\langle \begin{matrix} q \\ d_1 \end{matrix} \right\rangle + \ldots + \left\langle \begin{matrix} q \\ d_s \end{matrix} \right\rangle.$$

Assume that a generic q-ary p-ic $f(\mathbf{x}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$ can be written in the above form. By counting coefficients on the left-hand side and parameters on the right-hand side, we obtain the inequality

$$\left\langle \begin{matrix} q \\ p \end{matrix} \right\rangle \leqslant |\mathbf{Par}| = \left\langle \begin{matrix} q \\ d_1 \end{matrix} \right\rangle + \ldots + \left\langle \begin{matrix} q \\ d_s \end{matrix} \right\rangle.$$

Expand $\sum_{\mathbf{j} \in J} t_{\mathbf{j}}(\mathbf{x}) h^{\mathbf{j}}(\mathbf{x})$ into a polynomial in $\mathbf{x}$; that is, write

$$f(\mathbf{x}) = \sum_{\mathbf{i}:\, |\mathbf{i}| = p} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} = \sum_{\mathbf{j} \in J} t_{\mathbf{j}}(\mathbf{x}) h^{\mathbf{j}}(\mathbf{x}) = \sum_{\mathbf{i}:\, |\mathbf{i}| = p} \phi_{\mathbf{i}}(par's) \mathbf{x}^{\mathbf{i}}.$$

We obtain $\left\langle \begin{matrix} q \\ p \end{matrix} \right\rangle$ identities

$$a_{\mathbf{i}} = \phi_{\mathbf{i}}(par's);$$

that is, we view the coefficients of $f(\mathbf{x})$ as polynomials in *par's*.

Consider the map

$$\boldsymbol{\Phi}: \mathbb{C}^{\mathbf{Par}} \to \mathbb{C}^{\left\langle \begin{matrix} q \\ p \end{matrix} \right\rangle}$$

defined by

$$\Phi(par's) = (\phi_{\mathbf{i}}(par's))_{\mathbf{i}:\, |\mathbf{i}| = p},$$

where the co-ordinates of $\mathbb{C}^{\left\langle \begin{matrix} q \\ p \end{matrix} \right\rangle}$ are indexed by $\mathbf{i}$, with $|\mathbf{i}| = p$, and the co-ordinates of $\mathbb{C}^{\mathbf{Par}}$ are indexed by the set **Par**.

The assumption is that the range of the map $\boldsymbol{\Phi}$ are dense in $\mathbb{C}^{\left\langle \begin{matrix} q \\ p \end{matrix} \right\rangle}$. By Theorem 2.4 we infer that the $\left\langle \begin{matrix} q \\ p \end{matrix} \right\rangle$ polynomials $\phi_{\mathbf{i}}$ are algebraically independent. By Theorem 2.3 we further infer that the matrix

$$(\partial \phi_{\mathbf{i}} / \partial par)_{\mathbf{i}:\, |\mathbf{i}| = p,\, par \in \mathbf{Par}} \tag{13}$$

has full rank, where the rows are indexed by $\mathbf{i}$ and the columns by the set **Par** of parameters.

Since the matrix (13) has full rank, we can choose values for the parameters such that the matrix (13) still has full rank. Choosing values of the parameters is the same as choosing homogeneous polynomials $h_i(\mathbf{x})$ of degree d_i for $i = 1, \ldots, s$.

Because the matrix (13) has full rank, the columns of the matrix span the linear space $\mathbb{C}^{\binom{q}{p}}$. But such a space is naturally isomorphic to V_p. In particular,

$$\left(\frac{\partial \phi_{\mathbf{i}}}{\partial par}\right)_{\mathbf{i}:\,|\mathbf{i}|=p} \to \sum_{\mathbf{i}:\,|\mathbf{i}|=p} \frac{\partial \phi_{\mathbf{i}}}{\partial par}\mathbf{x}^{\mathbf{i}} = \frac{\partial f}{\partial par}.$$

Thus, the q-ary p-ics $\partial f/\partial par$ span the linear space V_p.

Hence there is no non-zero dual q-ary p-ic $g(\mathbf{u})$ such that

$$\left\langle g(\mathbf{u}) \,\middle|\, \frac{\partial f}{\partial par} \right\rangle = 0$$

for all parameters $par \in \mathbf{Par}$.

Now each of the parameters will only occur in one of the homogeneous polynomials $h_1(\mathbf{x}), \ldots, h_s(\mathbf{x})$. Say that par occurs in $h_i(\mathbf{x})$. Because $\partial h_i/\partial par = \mathbf{x}^{\mathbf{k}}$ for some $\mathbf{k}$ such that $|\mathbf{k}| = d_i$. Hence by the chain rule we conclude that

$$\frac{\partial f}{\partial par} = \frac{\partial f}{\partial h_i}\frac{\partial h_i}{\partial par} = \frac{\partial f}{\partial h_i}\mathbf{x}^{\mathbf{k}}.$$

As par ranges over all parameters in h_i, the multi-index $\mathbf{k}$ will range over all multi-indices such that $|\mathbf{k}| = d_i$. Thus the condition that

$$\left\langle g(\mathbf{u}) \,\middle|\, \frac{\partial f}{\partial h_i}\mathbf{x}^{\mathbf{k}} \right\rangle = 0$$

is equivalent to saying that $g(\mathbf{u})$ is apolar to $\partial f/\partial h_i$. Hence we have proven that there is no non-zero $g \in V_p^*$ which is apolar to all the forms $\partial f/\partial h_i$, for $i = 1, \ldots, s$. This argument proves half the theorem.

To prove the second part, all we need to do is trace the equivalences above in the opposite direction, much as we did in the first main theorem. □

4.4. *Applications of the second main theorem*

COROLLARY 4.7. *A generic ternary quartic can be written in the form $h_1 \cdot h_2 + h_3^2$, where h_1, h_2 and h_3 are ternary quadratics.*

PROOF. We are to prove that the ternary quartic canonical form $f(\mathbf{x}) = h_1 h_2 + h_3^2$. To this end we check that there is no non-zero dual ternary quartic $g(\mathbf{u})$ apolar to the three quadratic forms h_2, h_1 and $2h_3$, for an appropriate choice of h_1, h_2 and h_3. Choose $h_1 = x_1^2$, $h_2 = x_2^2$ and $h_3 = x_3^2$. By the pigeonhole principle, every term $\mathbf{u}^{\mathbf{i}}$ of such a $g(\mathbf{u})$ must contain one of the factors u_1^2, u_2^2 or u_3^2. Hence, $\langle g(\mathbf{u}) \,|\, \mathbf{x}^{\mathbf{i}} \rangle = 0$ for all $\mathbf{i}$ by the apolarity condition. Hence $g(\mathbf{u}) = 0$, and so we can write $f(\mathbf{x})$ in the above form. □

COROLLARY 4.8. *A generic ternary cubic can be written in the form $h_1^3 + h_2^3 + h_3^3 + c \cdot h_1 h_2 h_3$, where h_1, h_2 and h_3 are ternary linear forms, and c is suitable constant.*

PROOF. The canonical form we propose is

$$f(\mathbf{x}) = h_1^3 + h_2^3 + h_3^3 + h_0 h_1 h_2 h_3,$$

where h_0 is a form of degree 0. Hence we must check that there is no non-zero dual

ternary cubic $g(\mathbf{u})$ apolar to each of the forms

$$h_1h_2h_2, \qquad 3h_1^2+h_0h_2h_3, \qquad 3h_2^2+h_0h_1h_3, \qquad 3h_3^2+h_0h_1h_2$$

for at least one appropriate choice of h_0, h_1, h_2 and h_3. Choose $h_0=0$ and $h_i=x_i$ for $i=1, 2, 3$. Since $g(\mathbf{u})$ is apolar to x_i^2 any term in $g(\mathbf{u})$ containing u_i^2 vanishes. The only non-zero term in $g(\mathbf{u})$ is therefore $u_1u_2u_3$. But this term will also disappear, because $g(\mathbf{u})$ is apolar to $x_1x_2x_3$. Hence $g(\mathbf{u})=0$, as desired. □

COROLLARY 4.9. *Let $q=2j$. Then a generic q-ary quadratic can be written in the form $h_1h_2+h_3h_4+\cdots+h_{q-1}h_q$, where the h_i's are linear forms.*

PROOF. We check that there is no non-zero dual q-ary quadratic, $g(\mathbf{u})$, apolar to $h_1, h_2, \ldots, h_q$. Choose $h_i=x_i$. Then h_i will make all terms of $g(\mathbf{u})$ containing u_i vanish. Hence all terms of $g(\mathbf{u})$ vanish. □

COROLLARY 4.10. *A generic quaternary cubic can be written in the form $h_1h_2h_3+h_4h_5h_6$, where the h_i's are linear forms.*

PROOF. We check that there is no non-zero dual quaternary cubic, $g(\mathbf{u})$ apolar to all the forms h_1h_2, h_1h_3, h_2h_3, h_4h_5, h_4h_6 and h_5h_6 for a suitable choice of $h_1, \ldots, h_6$. First let $h_1=x_1$, $h_2=x_2$ and $h_3=x_3$. This forces all terms containing two different u_i's, $i=1, 2, 3$, to vanish. The terms left are of the form $u_i^ju_4^{3-j}$, where $i=1, 2, 3$ and $j=0, 1, 2, 3$. Now let $h_4=x_4$, $h_5=x_1+x_2+x_3$ and $h_6=x_1+x_2+x_3+x_4$. The condition that $g(\mathbf{u})$ is apolar to $x_ih_4h_5$ forces the coefficient of $u_i^2u_4$ to vanish. Since $g(\mathbf{u})$ is apolar to $h_4h_6-h_4h_5=x_4^2$, the coefficient of u_4^3 and $u_iu_4^2$ also vanish. Finally, $g(\mathbf{u})$ is apolar to $h_5h_6-h_5h_4=(x_1+x_2+x_3)^2$ makes x_i^3 vanish. By the main theorem, the above form is canonical. □

COROLLARY 4.11. *Let $f(\mathbf{x})$ be a generic binary form of even degree $p\geq 4$. Then $f(\mathbf{x})$ can be written in the form $h_1^p+h_2^p+\ldots+h_j^p+c\cdot h_1^2h_2^2\cdots h_j^2$, where $h_1, h_2, \ldots, h_j$ are binary linear forms, c is a suitable constant and $p=2j$.*

PROOF. The proposed canonical form is $h_1^p+\ldots+h_j^p+h_0h_1^2\cdots h_j^2$. Expression in the form $\partial f/\partial h_i$ are the following:

$$h_1^2\cdots h_j^2, \qquad ph_1^{p-1}+2h_0h_1h_2^2\cdots h_j^2, \qquad \ldots, \qquad ph_j^{p-1}+2h_0h_1^2\cdots h_{j-1}^2h_j.$$

Set $h_0=0$, and let $h_1, \ldots, h_j$ be distinct linear forms. Let us see if we can find dual binary form $g(\mathbf{u})$ of degree p under these conditions. Such a $g(\mathbf{u})$ is apolar to $h_1^{p-1}, \ldots, h_j^{p-1}$; hence by Proposition 4.1 $g(\mathbf{u})$ must have j double points. But because $g(\mathbf{u})$ is binary, these double points are squares of linear factors of $g(\mathbf{u})$. Hence the only possibility is that $g(\mathbf{u})$ be of the form $h_1^2\cdots h_j^2$. But $g(\mathbf{u})$ should also be apolar to $h_1^2\cdots h_j^2$, so $g=0$. By the use of the second main theorem, we are done. □

PROPOSITION 4.2 (Grace). *Let $r>2$ and let λ_i be an integer so that $0<\lambda_i\leq p$ for $i=1, \ldots, r$. Assume that $\sum_{i=1}^r\lambda_i=(r-1)(p+1)$. Let $l_1, \ldots, l_r$ be pairwise independent binary linear forms. Then for every binary p-ic $f(\mathbf{x})$ there exist unique binary forms $h_i(\mathbf{x})$, i = 1, . . . , r, so that*

$$f(\mathbf{x})=\sum_{i=1}^{r}l_i(\mathbf{x})^{\lambda_i}h_i(\mathbf{x}).$$

PROOF. By the main theorem, we need only prove that there is no non-zero dual q-ary p-ic $g(\mathbf{u})$ apolar to $\partial f/\partial h_i = l_i(\mathbf{x})^{\lambda_i}$ for $i = 1, \ldots, r$. But the linear form $l_i(\mathbf{x})$ can be written as $(\mathbf{a}^{(i)} \mid \mathbf{x})$ for some point $\mathbf{a}^{(i)}$. The condition that $g(\mathbf{u})$ be apolar to $(\mathbf{a}^{(i)} \mid \mathbf{x})^{\lambda_i}$ implies by Proposition 4.1 that $\mathbf{a}^{(i)}$ is a $(p+1-\lambda_i)$-fold point of $g(\mathbf{u})$. But since $g(\mathbf{u})$ is a binary form, we have the factor $(\mathbf{a}^{(i)} \mid \mathbf{u})^{p+1-\lambda_i}$ of $g(\mathbf{u})$, where the factor has degree $p+1-\lambda_i$. Summing the degrees of the different factors of $g(\mathbf{u})$ we find that

$$\sum_{i=1}^{r} (p+1-\lambda_i) = r(p+1) - (r-1)(p+1) = p+1 > p,$$

and this contradicts the assumption that degree of $g(\mathbf{u})$ is p. Hence the result follows for generic $f(\mathbf{x})$.

The maps ϕ_i (in the notation of the proof of the main theorem) are linear and their range is dense. Thus the map $\boldsymbol{\Phi}$ is surjective, and the dimensions of the domain and the range are equal. Hence the linear map is bijective, and the result follows without exceptions. □

COROLLARY 4.12 (Jordan's Lemma). *Let λ, μ and ν be positive integers so that $\lambda + \mu + \nu = 2p + 2$. If $x + y + z = 0$, then every homogeneous polynomial $S(x, y, z)$ of degree p can be uniquely written as*

$$S(x, y, z) = x^{\lambda} P(x, y, z) + y^{\mu} Q(x, y, z) + z^{\nu} R(x, y, z),$$

where P, Q and R are homogeneous of degrees $p - \lambda$, $p - \mu$ and $p - \nu$.

4.5. *Binary forms.* We are now going to focus on binary forms; that is, on the case $q = 2$. This will lead us to a generalization of Sylvester's theorem, which will deal with the non-generic cases.

Notice that for the space of binary forms of degree p one has $\dim(V_p) = \binom{2}{p} = p + 1$.

In this section let $f(\mathbf{x})$ be a binary p-ic, and $g(\mathbf{u})$ be a binary r-ic, where $r \leqslant p$. Define

$$f^{\perp} = \{g \in V_r^* : g \text{ is apolar to } f\},$$

$$g^{\perp} = \{f \in V_p : f \text{ is apolar to } g\}.$$

We see that $f^{\perp}$ and $g^{\perp}$ are linear spaces.

PROPOSITION 4.3. *We have* $\dim(f^{\perp}) \geqslant 2r - p$ *and* $\dim(g^{\perp}) = r$.

PROOF. The condition that $f(\mathbf{x})$ is apolar to $g(\mathbf{u})$ can be written as the set of linear equations

$$\langle g(\mathbf{u}) u_1^{p-r} \mid f(\mathbf{x}) \rangle = \langle g(\mathbf{u}) u_1^{p-r-1} u_2 \mid f(\mathbf{x}) \rangle = \ldots = \langle g(\mathbf{u}) u_2^{p-r} \mid f(\mathbf{x}) \rangle = 0.$$

So altogether there are $p - r + 1$ linear conditions.

Given $f(\mathbf{x})$, $\dim(f^{\perp}) \geqslant \dim(V_r^*) - (p - r + 1) = (r + 1) - (p - r + 1) = 2r - p$.

Given $g(\mathbf{u})$, we claim that the above $p - r + 1$ linear conditions are independent. Assume that there is a linear dependency among them. Thus there is linear dependency among $g(\mathbf{u})u_1^{p-r}$, $g(\mathbf{u})u_1^{p-r-1}u_2, \ldots, g(\mathbf{u})u_2^{p-r}$; that is, $u_1^{p-r}, u_1^{p-r-1}u_2, \ldots, u_2^{p-r}$ are linear dependent, a contradiction. Hence the $p - r + 1$ linear conditions are independent. Thus $\dim(g^{\perp}) = \dim(V_p) - (p - r + 1) = (p + 1) - (p - r + 1) = r$. □

For a vector $\mathbf{a} = (a_1, a_2)$ define $\mathbf{a}^* = (a_2, -a_1)$. Notice that $(\mathbf{a}^* \mid \mathbf{a}) = 0$.

PROPOSITION 4.4. *Assume that* $(\mathbf{a} \mid \mathbf{u})$ *is a linear factor of multiplicity* k *of* $g(\mathbf{u})$; *that is,* $g(\mathbf{u}) = (\mathbf{a} \mid \mathbf{u})^k h_2(\mathbf{u})$. *Then the form* $(\mathbf{a}^* \mid \mathbf{x})^{p-k+1} h(\mathbf{x})$ *is apolar to* $g(\mathbf{u})$, *where* h *is an arbitrary form of degree* $k-1$.

PROOF. We would like to show that

$$\langle (\mathbf{a} \mid \mathbf{u})^k h_2(\mathbf{u}) \mid (\mathbf{a}^* \mid \mathbf{x})^{p-k+1} h(\mathbf{x}) \rangle = 0.$$

Factorize $h_2(\mathbf{u})$; that is, write $h_2(\mathbf{u}) = (\mathbf{c}_1 \mid \mathbf{u}) \cdots (\mathbf{c}_{p-k} \mid \mathbf{u})$. Now

$$\begin{aligned}\langle (\mathbf{c}_1 \mid \mathbf{u}) \cdots (\mathbf{c}_{p-k} \mid \mathbf{u})(\mathbf{a} \mid \mathbf{u})^k \mid (\mathbf{a}^* \mid \mathbf{x})^{p-k+1} h(\mathbf{x}) \rangle \\ &= \langle (\mathbf{a} \mid \mathbf{u})^k \mid D_{\mathbf{c}_1,\mathbf{x}} \cdots D_{\mathbf{c}_{p-k},\mathbf{x}} (\mathbf{a}^* \mid \mathbf{x})^{p-k+1} h(\mathbf{x}) \rangle \\ &= \langle (\mathbf{a} \mid \mathbf{u})^k \mid (\mathbf{a}^* \mid \mathbf{x}) \cdot H(\mathbf{x}) \rangle \\ &= k! \cdot (\mathbf{a}^* \mid \mathbf{a}) \cdot H(\mathbf{a}) = 0,\end{aligned}$$

where $(\mathbf{a}^* \mid \mathbf{x}) \cdot H(\mathbf{x}) = D_{\mathbf{c}_1,\mathbf{x}} \cdots D_{\mathbf{c}_{p-k},\mathbf{x}} (\mathbf{a}^* \mid \mathbf{x})^{p-k+1} h(\mathbf{x})$. Here we have used Propositions 3.4 and 3.1. □

Recall from linear algebra that if W is a subspace of V_p, the orthogonal space $W^\perp$ which is a subspace of V_p^*, is defined as

$$W^\perp = \{h(\mathbf{u}) \in V_p^* : \forall k(\mathbf{x}) \in W \langle h(\mathbf{u}) \mid k(\mathbf{x}) \rangle = 0\}.$$

We have

$$\dim(W) + \dim(W^\perp) = \dim(V_p) = p + 1, \qquad (W_1 \oplus W_2)^\perp = W_1^\perp \cap W_2^\perp. \tag{14}$$

PROPOSITION 4.5. *Let* $p \geq r$, *let* $g \in V_r^*$, *and let* $g(\mathbf{u}) = \prod_{i=1}^m (\mathbf{a}^{(i)} \mid \mathbf{u})^{k_i}$, *where* $\mathbf{a}^{(i)}$ *and* $\mathbf{a}^{(l)}$ *are pairwise linearly independent. Then the space* $g^\perp$ *is spanned by the forms:* $h_i(\mathbf{x})((\mathbf{a}^{(i)})^* \mid \mathbf{x})^{p-k_i+1}$ *for* $i = 1, \ldots, m$, *where* $h_i(\mathbf{x})$ *is an arbitrary form of degree* $k_i - 1$.

PROOF (due to Giudici). Let

$$W_i = \{h_i(\mathbf{x})((\mathbf{a}^{(i)})^* \mid \mathbf{x})^{p-k_i+1} : h_i \in V_{k_i-1}\}$$

for $i = 1, \ldots, m$. Clearly, $\dim(W_i) = \dim(V_{k_i-1}) = k_i$.

Now, $\dim(W_i^\perp) = p + 1 - \dim(W_i) = p + 1 - k_i$. Every element in the form $h(\mathbf{u}) \cdot (\mathbf{a}^{(i)} \mid \mathbf{u})^{k_i}$, where $h \in V_{p-k_i}^*$, belongs to $W_i^\perp$ by Proposition 4.4. Comparing dimensions we infer that

$$W_i^\perp = \{h_i(\mathbf{u}) \cdot (\mathbf{a}^{(i)} \mid \mathbf{u})^{k_i} : h_i \in V_{p-k_i}^*\}.$$

Now, by the identity (14) we see that

$$\begin{aligned}\left(\bigoplus_{i=1}^m W_i\right)^\perp &= \bigcap_{i=1}^m W_i^\perp \\ &= \{h(\mathbf{u}) \cdot g(\mathbf{u}) : h(\mathbf{u}) \in V_{p-r}^*\}.\end{aligned}$$

Hence

$$\begin{aligned}\dim\left(\bigoplus_{i=1}^m W_i\right) &= p + 1 - \dim\left(\left(\bigoplus_{i=1}^m W_i\right)^\perp\right) \\ &= p + 1 - \dim(\{h(\mathbf{u}) \cdot g(\mathbf{u}) : h \in V_{p-r}^*\}) \\ &= p + 1 - (p - r + 1) = r.\end{aligned}$$

This proves that $\bigoplus_{i=1}^{m} W_i$ has dimension r. But $g^{\perp} \supseteq \bigoplus_{i=1}^{m} W_i$ and we know that $\dim(g^{\perp}) = r$. Hence $g^{\perp} = \bigoplus_{i=1}^{m} W_i$.

We have thus shown that the subspace $g^{\perp}$ is spanned by the polynomials $h_i(\mathbf{x})((\mathbf{a}^{(i)})^* \mid \mathbf{x})^{p-k_i+1}$, where $h_i(\mathbf{x})$ have degree $k_i - 1$ and $i = 1, \ldots, m$. □

THEOREM 4.5. *Every binary form $f(\mathbf{x})$ of degree $(2r-1)$ can be written in the form*

$$\sum_{i=1}^{m} h_i(\mathbf{x}) \cdot l_i(\mathbf{x})^{2r-k_i},$$

where l_i are pairwise independent linear forms, h_i are suitable binary forms of degree $(k_i - 1)$, and $\sum_{i=1}^{m} k_i = r$.

PROOF. Let $p = 2r - 1$. By Proposition 4.3, $\dim(f^{\perp}) \geq 2r - p = 1$. Hence there exists a dual binary r-form $g(\mathbf{u})$ in $f^{\perp}$. Factorize $g(\mathbf{u})$; that is, write $g(\mathbf{u})$ in the form $g(\mathbf{u}) = \prod_{i=1}^{m} (\mathbf{a}^{(i)} \mid \mathbf{u})^{k_i}$. Let $l_i(\mathbf{x}) = ((\mathbf{a}^{(i)})^* \mid \mathbf{x})$. However, by the previous proposition we know that a spanning set for $g^{\perp}$ is $h_i \cdot l^{p-k_i+1}$, for $i = 1, \ldots, m$ and $h_i \in V_{k_i-1}$. Hence we can write $f(\mathbf{x})$ in the above form. □

THEOREM 4.6 (Sylvester). *A generic binary form $f(\mathbf{x})$ of degree $(2r-1)$ can be written as a sum of r $(2r-1)$-powers.*

PROOF. Apply the previous theorem. Notice that in the generic case we can assume that $g \in f^{\perp}$ does not have any double roots. Hence each h_i is a constant, and the left-hand side is a sum of r $(2r-1)$-powers. □

ACKNOWLEDGMENT

The research of Gian-Carlo Rota was supported by National Science Foundation Grant MCS-8104855.

REFERENCES

1. A. Björner and L. Lovász, Psuedomodular lattices and continuous matroids, *Acta Scient. math. Szeged*, **51** (1987), 295–308.
2. R. Ehrenborg and G.-C. Rota, *Lectures on Polynomials*, to appear (1992).
3. J. H. Grace and A. Young, *The Algebra of Invariants*, Chelsea, New York, 1903.
4. W. Gröber, *Moderne Algebraische Geometrie*, Springer-Verlag, Vienna, 1949.
5. A. W. Ingleton, Representations of matroids, in: *Combinatorial Mathematics and its Applications*, D. J. A. Welsh (ed.), Academic Press, London, 1971.
6. E. Lasker, Zur Theorie der kanonische Formen, *Math. Ann.*, **58** (1904), 434–440.
7. B. Lindström, A Desarguesian theorem for algebraic combinatorial geometries, *Combinatorica*, **5** (1985), 237–239.
8. S. MacLane, A lattice formulation for transcendence degrees and p-bases, *Duke Math. J.*, **4** (1938), 455–468.
9. F. Meyer, *Apolarität und Ratinale Curven*, Franz Fues, Tübingen, 1883.
10. B. Reznick, *Sums of Even Powers of Real Linear Forms, Mem. Am. Math. Soc.*, **96**, no. 463, March 1992.
11. H. W. Richmond, On canonical forms, *Q. J. Math.*, **33** (1902), 331–340.
12. B. Segre, *Prodromi di Geometria Algebrica*, Cremonese, Roma, 1971.

13. H. W. Turnbull, *The Theory of Determinants, Matrices and Invariants*, Dover, New York, 1960.
14. E. K. Wakeford, On canonical forms, *Proc. Lond. Math. Soc.*, 2, **18** (1918–1919), 403–410.

Received 30 January 1992 and accepted in revised form 20 September 1992

RICHARD EHRENBORG AND GIAN-CARLO ROTA
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, U.S.A.