

*Pacific
Journal of
Mathematics*

SOME NEW CANONICAL FORMS FOR POLYNOMIALS

BRUCE REZNICK

Volume 266 No. 1

November 2013

SOME NEW CANONICAL FORMS FOR POLYNOMIALS

BRUCE REZNICK

We give some new canonical representations for forms over \mathbb{C} . For example, a general binary quartic form can be written as the square of a quadratic form plus the fourth power of a linear form. A general cubic form in (x_1, \dots, x_n) can be written uniquely as a sum of the cubes of linear forms $l_{ij}(x_i, \dots, x_j)$, $1 \leq i \leq j \leq n$. A general ternary quartic form is the sum of the square of a quadratic form and three fourth powers of linear forms. The methods are classical and elementary.

1. Introduction and overview

Introduction. Let $H_d(\mathbb{C}^n)$ denote the vector space of complex forms of degree d in n variables, or n -ary d -ics forms; it has dimension $N(n, d) := \binom{n+d-1}{d}$. One of the major accomplishments of 19th century algebra was the discovery of canonical forms for certain classes of n -ary d -ics, especially as the sum of d -th powers of linear forms. By a *canonical form* we mean a polynomial $F(t; x)$ in two sets of variables, $t \in \mathbb{C}^{N(n,d)}$ and $x \in \mathbb{C}^n$, with the property that, for general $p \in H_d(\mathbb{C}^n)$, there exists t so that $p(x) = F(t; x)$. Put another way, the set $\{F(t; x) : t \in \mathbb{C}^{N(n,d)}\}$ is a Zariski open set in $H_d(\mathbb{C}^n)$.

In this paper, we present some new canonical forms, whose main novelty is that they involve intermediate powers of forms of higher degree, or forms with a restricted set of monomials. (These variations have been suggested by Hilbert's study of ternary quartics [1888], which led to his 17th problem, as well as by a remarkable theorem of B. Reichstein [1987] on cubic forms.) These expressions are less susceptible to apolarity arguments than the traditional canonical forms, and lead naturally to (mostly open) enumeration questions.

To take a simple, yet familiar, example,

$$(1-1) \quad F(t_1, t_2, t_3; x, y) = (t_1x + t_2y)^2 + (t_3y)^2$$

is a canonical form for binary quadratic forms. By the usual completion of squares, $p(x, y) = ax^2 + 2bxy + cy^2$ can be put into (1-1) for $t_1 = \sqrt{a}$, $t_2 = b/t_1$ and

MSC2000: primary 11E76, 14N15; secondary 11E25, 11P05, 15A72.

Keywords: canonical forms, sums of powers of polynomials.

$t_3^2 = c - t_2^2$. Many of the examples in this paper can be viewed as imperfect attempts to generalize (1-1).

Sylvester [1851a; 1851b] presented a family of canonical forms for binary forms in all degrees.

Theorem 1.1 (Sylvester's theorem). (i) *A general binary form p of odd degree $2s - 1$ can be written as*

$$(1-2) \quad p(x, y) = \sum_{j=1}^s (\alpha_j x + \beta_j y)^{2s-1}.$$

(ii) *A general binary form p of even degree $2s$ can be written as*

$$(1-3) \quad p(x, y) = \lambda x^{2s} + \sum_{j=1}^s (\alpha_j x + \beta_j y)^{2s}$$

for some $\lambda \in \mathbb{C}$.

The somewhat unsatisfactory nature of the asymmetric summand in (1-3) has been the inspiration for other canonical forms for binary forms of even degree.

Another familiar canonical form is the generalization of (1-1) into the upper-triangular expression for quadratic forms, found by repeated completion of the square:

Theorem 1.2. *A general quadratic form $p \in H_2(\mathbb{C}^n)$ can be written as*

$$(1-4) \quad p(x_1, \dots, x_n) = \sum_{k=1}^n (t_{k,k}x_k + t_{k,k+1}x_{k+1} + \dots + t_{k,n}x_n)^2, \quad t_{k,l} \in \mathbb{C}.$$

The expression in (1-4) is unique, up to the signs of the linear forms.

There are two ways to verify that a candidate expression $F(t; x)$ is, in fact, a canonical form. One is the classical nonconstructive method based on the existence of a point at which the Jacobian matrix has full rank. (See Corollary 2.3, and see Theorem 3.2 for the apolar version.) Lasker [1904] attributes the underlying idea to Kronecker and Lüroth—see [Wakeford 1920, p. 208].

Ideally, however, a canonical form can be derived constructively, and the number of different representations can thereby be determined. The convention in this paper will be that two representations are the same if they are equal, up to a permutation of like summands and with the identification of f^k and $(\zeta f)^k$ when $\zeta^k = 1$. The representation in (1-2) is unique in this sense, even though there are $s! \cdot (2s - 1)^s$ different $2s$ -tuples $(\alpha_1, \beta_1, \dots, \alpha_s, \beta_s)$ for which (1-2) is valid.

In addition to Theorem 1.1, another motivational example for this paper is a remarkable canonical form for cubic forms found in [Reichstein 1987], which can be thought of as a “completion of the cube”.

Theorem 1.3 (Reichstein). *A general cubic $p \in H_3(\mathbb{C}^n)$ can be written uniquely as*

$$(1-5) \quad p(x_1, \dots, x_n) = \sum_{k=1}^n l_k^3(x_1, \dots, x_n) + q(x_3, \dots, x_n),$$

where $l_k \in H_1(\mathbb{C}^n)$ and $q \in H_3(\mathbb{C}^{n-2})$.

This is a canonical form, provided q is viewed as a t -linear combination of the monomials in (x_3, \dots, x_n) ; since $N(n, 3) = n^2 + N(n - 2, 3)$, the constant count is right. Iteration (see (6-1)) gives p as a sum of roughly $n^2/4$ cubes. The minimum from constant-counting, which is justified by the Alexander–Hirschowitz theorem [1995], is roughly $n^2/6$. We give Reichstein’s constructive proof of Theorem 1.3 in Section 6.

Here are some representative examples of the new canonical forms in this paper.

Theorem 1.4. *A general cubic form $p \in H_3(\mathbb{C}^n)$ has a unique representation*

$$(1-6) \quad p(x_1, \dots, x_n) = \sum_{1 \leq i < j \leq n} (t_{\{i,j\},i}x_i + \dots + t_{\{i,j\},j}x_j)^3,$$

where $t_{\{i,j\},k} \in \mathbb{C}$.

Theorem 1.5. *A general binary sextic $p \in H_6(\mathbb{C}^2)$ can be written as $p(x, y) = f^2(x, y) + g^3(x, y)$, where $f \in H_3(\mathbb{C}^2)$ is a cubic form and $g \in H_2(\mathbb{C}^2)$ is a quadratic form.*

Theorem 1.4 has a constructive proof. Theorem 1.5 is, in fact, a very special case of much deeper recent results of Várilly-Alvarado [2008, especially Theorem 1.2 and Remark 4.5; 2011, Section 1.2]). We include it because our proof, in the next section, is very short.

Theorems 1.1 and 1.5 are both special cases of a more general class of canonical forms for $H_d(\mathbb{C}^2)$, which is a corollary of [Ehrenborg and Rota 1993, Theorem 4.4] (see Theorem 3.4), but not worked out explicitly there.

Theorem 1.6. *Suppose $d \geq 1$, $\{l_j : 1 \leq j \leq m\}$ is a fixed set of pairwise non-proportional linear forms, and suppose $e_k \mid d$, $d > e_1 \geq \dots \geq e_r$, $1 \leq k \leq r$, and*

$$(1-7) \quad m + \sum_{k=1}^r (e_k + 1) = d + 1.$$

Then a general binary d -ic form $p \in H_d(\mathbb{C}^2)$ can be written as

$$(1-8) \quad p(x, y) = \sum_{j=1}^m t_j l_j^d(x, y) + \sum_{k=1}^r f_k^{d/e_k}(x, y),$$

where $t_j \in \mathbb{C}$ and $\deg f_k = e_k$.

The condition $e_k < d$ excludes the vacuous case $m = 0, r = 1, e_1 = d$. If each $e_k = 1$ and $r = \lfloor (d+1)/2 \rfloor$, then $m = d+1 - 2\lfloor (d+1)/2 \rfloor \in \{0, 1\}$ and Theorem 1.6 becomes Theorem 1.1; Theorem 1.5 is Theorem 1.6 in the special case $d = 6, m = 0, r = 2, e_1 = 3, e_2 = 2$. As an example of a canonical form that is unlikely to find a constructive proof: for a general $p \in H_{84}(\mathbb{C}^2)$, there exist $f \in H_{42}(\mathbb{C}^2), g \in H_{28}(\mathbb{C}^2)$ and $h \in H_{12}(\mathbb{C}^2)$ so that $p = f^2 + g^3 + h^7$.

By taking $d = 2s, e_1 = 2, e_2 = \dots = e_{s-1} = 1$ and $m = 0$, in Theorem 1.6, we obtain an alternative to the dangling term λx^{2s} in (1-3).

Corollary 1.7. *A general binary form p of even degree $2s$ can be written as*

$$(1-9) \quad p(x, y) = (\alpha_0 x^2 + \beta_0 xy + \gamma_0 y^2)^s + \sum_{j=1}^{s-1} (\alpha_j x + \beta_j y)^{2s}.$$

A different generalization of Theorem 1.1 focuses on the number of summands.

Theorem 1.8. *A general binary form of degree uv can be written as a sum of $\lceil (uv+1)/(u+1) \rceil$ v -th powers of binary forms of degree u .*

Cayley proved that, after an invertible linear change of variables $(x, y) \mapsto (X, Y)$, a general binary quartic can be written as $X^4 + 6\lambda X^2 Y^2 + Y^4$. There are two natural ways to generalize this to higher even degree, and, almost 100 years ago, Wakeford [1913; 1920] did both.

Theorem 1.9 (Wakeford’s theorem). *After an invertible linear change of variables, a general $p \in H_d(\mathbb{C}^n)$ can be written so that the coefficient of each x_i^d is 1 and the coefficient of each $x_i^{d-1} x_j$ is 0.*

There are $N(n, d) - n^2$ unmentioned monomials above, and, when combined with the n^2 coefficients in the change of variables, the constant count is correct for a canonical form. Wakeford was also interested in knowing *which* sets of $n(n-1)$ monomials can be eliminated by a change of variables, and we are able to settle this for binary forms in Theorem 2.4. (Theorem 1.9 was independently discovered in [Guazzone 1975], as an attempt to generalize the canonical form $X^3 + Y^3 + Z^3 + 6\lambda XYZ$ for $H_3(\mathbb{C}^3)$. Babbage [1976] subsequently observed that this can be proved by the Lasker–Wakeford theorem, without noting that Wakeford [1920] had already done so.)

The second generalization of $X^4 + 6\lambda X^2 Y^2 + Y^4$ will not be pursued here; see [Ehrenborg and Rota 1993, Corollary 4.11]. A canonical form for binary forms of even degree $2s$ is given by

$$(1-10) \quad \sum_{k=1}^s l_k^{2s}(x, y) + \lambda \prod_{k=1}^s l_k^2(x, y), \quad l_k(x, y) = \alpha_k x + \beta_k y.$$

This construction is due to Sylvester [1851b] for $2s = 4, 8$. His methods failed for $2s = 6$, but Wakeford [1913] was able to prove it. The full version of (1-10) is proved in [Wakeford 1920, p. 408], where he notes that “the number of ways this reduction can be performed is interesting”, citing “3, 8, 5” for $2s = 4, 6, 8$.

The nontrivial study of canonical forms was initiated by Clebsch’s discovery [1861] (see, e.g., [Geramita 1996, pp. 50–51; Reznick 1992a, pp. 59–60]) that, despite the fact that $N(3, 4) = 5 \times N(3, 1)$, a general ternary quartic cannot be written as a sum of five fourth powers of linear forms. This was early evidence that constant-counting can fail. But $N(3, 4)$ is also equal to $1 \times N(3, 2) + 3 \times N(3, 1)$, and ternary quartics *do* satisfy an alternative canonical form as a mixed sum of powers.

Theorem 1.10. *A general ternary quartic $p \in H_4(\mathbb{C}^3)$ can be written as*

$$(1-11) \quad p(x_1, x_2, x_3) = q^2(x_1, x_2, x_3) + \sum_{k=1}^3 l_k^4(x_1, x_2, x_3),$$

where $q \in H_2(\mathbb{C}^3)$ and $l_k \in H_1(\mathbb{C}^3)$.

As an alternative generalization of canonical forms, one might also consider polynomial maps $F : S \mapsto H_d(\mathbb{C}^n)$, where S is an N -dimensional subspace of some \mathbb{C}^M . In the simplest case, for binary quadratic forms, observe that the coefficient of x^2 in

$$(1-12) \quad (t_1x + t_2y)^2 + (it_1x + t_3y)^2$$

is 0, so (1-12) is not canonical. This is essentially the only kind of exception.

Theorem 1.11. *Suppose $(c_1, c_2, c_3, c_4) \in \mathbb{C}^4$, and it is not true that $c_3 = \epsilon c_1$ and $c_4 = \epsilon c_2$ for $\epsilon \in \{\pm i\}$. Then, for general $p \in H_2(\mathbb{C}^2)$, there exists $(t_1, t_2, t_3, t_4) \in \mathbb{C}^4$ satisfying*

$$\sum_{j=1}^4 c_j t_j = 0$$

and such that

$$(1-13) \quad p(x, y) = (t_1x + t_2y)^2 + (t_3x + t_4y)^2.$$

In the exceptional case, there exists (x_0, y_0) so that, for all feasible choices of t_j , $p(x_0, y_0) = 0$.

Another alternative version of (1-3) is the following conjecture, which can be verified up to degree 8.

Conjecture 1.12. *A general binary form p of even degree $2s$ can be written as*

$$(1-14) \quad p(x, y) = \sum_{j=1}^{s+1} (\alpha_j x + \beta_j y)^{2s}, \quad \text{where } \sum_{j=1}^{s+1} (\alpha_j + \beta_j) = 0.$$

Outline. Here is an outline of the paper. In Section 2, we introduce notation and definitions. The definition of canonical form is the classical one and roughly parallels that in [Ehrenborg and Rota 1993], an important updating of this subject about 20 years ago. Our point of view is considerably more elementary in many respects than [Ehrenborg and Rota 1993], but uses the traditional criterion: a polynomial map $F : \mathbb{C}^N \mapsto H_d(\mathbb{C}^n)$ is a *canonical form* if a general $p \in H_d(\mathbb{C}^n)$ is in the range; this occurs if and only if there is at least one point $u \in \mathbb{C}^N$ so that $\{\partial F / \partial t_j(u)\}$ spans $H_d(\mathbb{C}^n)$. (See Corollary 2.3.) This leads to immediate nonconstructive proofs of Theorems 1.2, 1.5, 1.9 and 1.10, and a somewhat more complicated proof of Theorem 2.4, which answers Wakeford’s question about missing monomials for binary forms.

In Section 3, we discuss classical apolarity and its implications for canonical forms. (Apolarity methods become more complicated when a component of a canonical form comes from a restricted set of monomials.) A generalization of the classical fundamental theorem of apolarity from [Reznick 1996] allows us to identify a class of bases for $H_d(\mathbb{C}^n)$ which give a nonconstructive proof of Theorem 1.6, and hence Theorem 1.1. A similar argument yields the proof of Theorem 1.8. We also present Sylvester’s algorithm, Theorem 3.8, allowing for a constructive proof of Theorem 1.1. We conclude with a brief summary of connections with the theorems of Alexander and Hirschowitz and recent work on the rank of forms.

In Section 4 we discuss some special cases of Theorem 1.6. Sylvester’s algorithm is used in constructive proof of Theorem 1.6 when $e_k \equiv 1$, in which case the representation is unique. We give some other constructive proofs for $d \leq 4$, and present numerical evidence regarding the number of representations in Corollary 1.7 and a few other cases. Using elementary number theory, we show that, for each r , there are only finitely many canonical forms (1-8) with $m = 0$, and, up to degree N , there are $N + \mathcal{O}(N^{1/2})$ such canonical forms in which the e_k are equal.

Section 5 discusses some familiar results on sums of two squares of binary forms and canonical representations of quadratic forms as a sum of squares of linear forms. This includes a constructive proof of Theorem 1.2, which provides the groundwork for the proof of Theorem 1.4. We also give a short proof of a canonical form which illustrates the classical result that a general ternary quartic is the sum of three squares of quadratic forms.

In Section 6, we turn to forms in more than two variables and low degree, give constructive proofs of Theorems 1.3 and 1.4, as well as the noncanonical Theorem 6.2, which shows that *every* cubic in $H_3(\mathbb{C}^n)$ is a sum of at most $n(n+1)/2$ cubes of linear forms. Theorem 1.3 can be “lifted” to an ungainly canonical form for quartics as a sum of fourth powers (see Corollary 6.3), but not further to quintics. Number theoretic considerations rule out a Reichstein-type canonical form for quartics in 12 variables; see Proposition 6.4 for other instances of this phenomenon.

In Section 7, we offer a preliminary discussion of canonical forms in which the domain of a polynomial map $F : \mathbb{C}^M \mapsto H_d(\mathbb{C}^n)$ is restricted to an N -dimensional subspace of \mathbb{C}^M , of which Theorem 1.11 and Conjecture 1.12 are examples.

2. Basic definitions, and proofs of Theorems 1.2, 1.5, 1.9 and 1.10

Let $\mathcal{I}(n, d)$ denote the index set of monomials in $H_d(\mathbb{C}^n)$:

$$(2-1) \quad \mathcal{I}(n, d) = \left\{ (i_1, \dots, i_n) : 0 \leq i_k \in \mathbb{Z}, \sum_k i_k = d \right\}.$$

Let $x^i = x_1^{i_1} \cdots x_n^{i_n}$ and $c(i) = d! / (\prod i_k!)$ denote the multinomial coefficient. If $p \in H_d(\mathbb{C}^n)$, then we write

$$(2-2) \quad p(x_1, \dots, x_n) = \sum_{i \in \mathcal{I}(n, d)} c(i) a(p; i) x^i, \quad a(p; i) \in \mathbb{C}.$$

We say that two forms are *distinct* if they are nonproportional, and a set of forms is *honest* if the forms are pairwise distinct. For later reference, recall Biermann's theorem; see [Reznick 1992a, p. 31].

Theorem 2.1 (Biermann's theorem). *If $p \in H_d(\mathbb{C}^n)$ and $p \neq 0$, then there exists $i \in \mathcal{I}(n, d)$ so that $p(i) \neq 0$.*

The easy verification of whether a formula is a canonical form for $H_d(\mathbb{C}^n)$ relies on a crucial alternative. A self-contained accessible proof is in [Ehrenborg and Rota 1993, Theorem 2.4], for which Ehrenborg and Rota thank M. Artin and A. Mattuck. For further discussion of the underlying algebraic geometry, see Section 9.5 in [Cox et al. 2007].

Theorem 2.2. *Suppose $M \geq N$ and $F : \mathbb{C}^M \rightarrow \mathbb{C}^N$ is a polynomial map; that is,*

$$F(t_1, \dots, t_M) = (f_1(t_1, \dots, t_M), \dots, f_N(t_1, \dots, t_M)),$$

where each $f_j \in \mathbb{C}[t_1, \dots, t_M]$. Then either (i) or (ii) holds:

- (i) *The N polynomials $\{f_j : 1 \leq j \leq N\}$ are algebraically dependent and $F(\mathbb{C}^M)$ lies in some nontrivial variety $\{P = 0\}$ in \mathbb{C}^N .*
- (ii) *The N polynomials $\{f_j : 1 \leq j \leq N\}$ are algebraically independent and $F(\mathbb{C}^M)$ is dense in \mathbb{C}^N .*

The second case occurs if and only there is a point $u \in \mathbb{C}^M$ at which the Jacobian matrix $[\partial f_i / \partial t_j(u)]$ has full rank.

When $M = N = N(n, d)$, we may interpret such an F as a map from \mathbb{C}^N to $H_d(\mathbb{C}^n)$ by indexing $\mathcal{I}(n, d)$ as $\{i(k) : 1 \leq k \leq N\}$ and making the interpretation in

an abuse of notation that

$$(2-3) \quad F(t; x) = \sum_{k=1}^N c(i(k)) f_k(t_1, \dots, t_N) x^{i(k)}.$$

Definition. A canonical form for $H_d(\mathbb{C}^n)$ is any polynomial map $F : \mathbb{C}^{N(n,d)} \mapsto H_d(\mathbb{C}^n)$ in which F satisfies Theorem 2.2(ii).

That is, F is a canonical form if and only if $N = N(n, d)$ and, for a general $p \in H_d(\mathbb{C}^n)$, there exists $t \in \mathbb{C}^N$ so that $p(x) = F(t; x)$. The significance of this choice of N is that it is the smallest possible value. In the rare cases where F is surjective, we say that the canonical form is *universal*.

By translating the definitions and using (2-1) and (2-3), we obtain an immediate corollary of Theorem 2.2:

Corollary 2.3. *The polynomial map $F : \mathbb{C}^N \mapsto H_d(\mathbb{C}^n)$ is a canonical form if and only if there exists $u \in \mathbb{C}^n$ so that $\{\partial F / \partial t_j(u)\}$ spans $H_d(\mathbb{C}^n)$.*

We shall let $J := J(F; u)$ denote the span of the forms $\{\partial F / \partial t_j(u)\}$. In any particular case, the determination of whether $J = H_d(\mathbb{C}^n)$ amounts to the computation of the determinant of an $N(n, d) \times N(n, d)$ matrix. As much as possible in this paper, we give proofs which can be checked by hand, by making a judicious choice of u and ordering of the monomials in $H_d(\mathbb{C}^n)$, showing sequentially that they all lie in J .

Classically, the use of the term ‘‘canonical form’’ has been limited to cases in which $F(t; x)$ has a natural interpretation as a combination of forms in $H_d(\mathbb{C}^n)$, such as a sum of powers of linear forms, or as a result of a linear change of variables. It seems odd that canonical forms are perceived as rare, since a ‘‘general’’ polynomial map from $\mathbb{C}^N \mapsto H_d(\mathbb{C}^n)$ is a canonical form. (This is an observation which goes back at least to [Richmond 1902].) For example, if $\{f_j(x)\}$ is a basis for $H_d(\mathbb{C}^n)$, then

$$(2-4) \quad F(t; x) = \sum_{j=1}^N t_j f_j(x)$$

should be (but usually isn’t) considered a canonical form. In particular, (2-2) with $f_j(x) = c(i_j)x^{i_j}$ is itself a canonical form.

The following computation will occur repeatedly. If $es = d$, then

$$(2-5) \quad g = \sum_{i_j \in \mathcal{I}(n,e)} t_j x^{i_j} \implies \frac{\partial g^s}{\partial t_j} = s x^{i_j} g^{s-1}.$$

If g is specialized to be a monomial, then all these partials will also be monomials.

Nonconstructive proof of Theorem 1.2. Given (1-4), let

$$l_k(x) = \sum_{m=k}^n t_{k,m} x_m, \quad F(x) = \sum_{k=1}^n l_k^2(x).$$

Then $\partial F / \partial t_{k,m} = 2x_m l_k$. Set $t_{k,m} = \delta_{k,m}$, so that $l_k = x_k$ and $\partial F / \partial t_{k,m} = 2x_k x_m$. Since $1 \leq k \leq m \leq n$, all monomials from $H_2(\mathbb{C}^n)$ appear in J . □

Nonconstructive proof of Theorem 1.5. Suppose

$$(2-6) \quad p(x, y) = f^2(x, y) + g^3(x, y),$$

with

$$f(x, y) = t_1 x^3 + t_2 x^2 y + t_3 x y^2 + t_4 y^3, \quad g(x, y) = t_5 x^2 + t_6 x y + t_7 y^2.$$

Then, by (2-5), the partials with respect to the t_j are

$$2x^3 f, 2x^2 y f, 2x y^2 f, 2y^3 f; \quad 3x^2 g^2, 3x y g^2, 3y^2 g^2.$$

Upon specializing at $f = x^3, g = y^2$, these become

$$2x^6, 2x^5 y, 2x^4 y^2, 2x^3 y^3; \quad 3x^2 y^4, 3x y^5, 3y^6.$$

It is then evident that $J = H_6(\mathbb{C}^2)$. □

Nonconstructive proof of Theorem 1.9. Let $\mathcal{L} \subset \mathcal{F}(n, d)$ consist of all n -tuples except the permutations of $(d, 0, \dots, 0)$ and $(d - 1, 1, \dots, 0)$ and let $X_i = \sum_{j=1}^n \alpha_{ij} x_j$. The assertion is that, with the $(N(n, d) - n - \binom{n}{2}) + n^2 = N(n, d)$ parameters t_l and α_{ij} ,

$$(2-7) \quad \sum_{i=1}^n X_i^d + \sum_{l \in \mathcal{L}} t_l X_1^{l_1} \cdots X_n^{l_n}$$

is a canonical form. Evaluate the partials at the point where $X_i = x_i$ and $t_l = 0$: they are $dx_j x_i^{d-1}$ (for α_{ij}) and x^l (for t_l). Taking $1 \leq i, j \leq n$ and $l \in \mathcal{L}$, we see that J contains all monomials in $H_d(\mathbb{C}^n)$. □

As a special case (used later in Theorem 4.6), we obtain the familiar result that, after appropriate linear changes of variable, a general binary quartic may be written as $x^4 + 6\lambda x^2 y^2 + y^4$. It is classically known (see [Elliott 1913, Section 211]) that the choice of λ is not unique: in fact, after appropriate linear changes of variable, $x^4 + 6\lambda x^2 y^2 + y^4$ can be written as $x^4 + 6\mu x^2 y^2 + y^4$ for $\mu \in \{\pm\lambda, \pm(1 - \lambda)/(1 + 3\lambda), \pm(1 + \lambda)/(1 - 3\lambda)\}$.

Wakeford asserts that Theorem 1.9 is also true with $x_i^{d-1} x_j$ replaced by $x_i^{d-r} x_j^r$ (evidently when $r \neq d/2$), but his proof seems sketchy. He also gives necessary conditions for sets of $n(n - 1)$ monomials which may be omitted, and these are hard

to follow as well. Below, we answer his question in the binary case: in the only two excluded cases below, (2-8) has a square factor, and so cannot be canonical.

Theorem 2.4. *Let $\mathcal{B} = (m_1, m_2, n_1, n_2)$ be four distinct integers in $\{0, \dots, d\}$ so that $\{m_1, m_2\} \neq \{0, 1\}, \{d-1, d\}$. Then, after an invertible linear change of variable, a general binary form p of degree d can be written as*

$$(2-8) \quad p(x, y) = x^{d-n_1}y^{n_1} + x^{d-n_2}y^{n_2} + \sum_{k \notin \mathcal{B}} t_k x^{d-k}y^k \quad \text{for some } \{t_k\} \subset \mathbb{C}.$$

Proof. Writing $(x, y) \mapsto (\alpha_1 x + \alpha_2 y, \alpha_3 x + \alpha_4 y) := (X, Y)$, we have

$$(2-9) \quad F = X^{d-n_1}Y^{n_1} + X^{d-n_2}Y^{n_2} + \sum_{k \notin \mathcal{B}} t_k X^{d-k}Y^k.$$

Evaluate the partials of (2-9) at $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (1, 0, 0, 1)$ (so $X = x, Y = y$) and $t_k = 1$ (note the difference with the previous proof, in which $t_k = 0$). The $d-3$ partials with respect to the t_k are simply $x^{d-k}y^k$, $k \notin \mathcal{B}$, so these are in J . Further,

$$(2-10) \quad \frac{\partial F}{\partial \alpha_1} = \sum_{j \neq m_1, m_2} (d-j)x^{d-j}y^j, \quad \frac{\partial F}{\partial \alpha_4} = \sum_{j \neq m_1, m_2} jx^{d-j}y^j.$$

Since most monomials used in (2-10) are already in J , it follows that J also contains

$$(2-11) \quad (d-n_1)x^{d-n_1}y^{n_1} + (d-n_2)x^{d-n_2}y^{n_2}, \quad n_1x^{d-n_1}y^{n_1} + n_2x^{d-n_2}y^{n_2},$$

and since $(d-n_1)n_2 \neq (d-n_2)n_1$, (2-11) implies that $x^{d-n_j}y^{n_j} \in J$ for $j = 1, 2$. To this point, we have shown that J contains all monomials from $H_d(\mathbb{C}^2)$ except for $x^{d-m_j}y^{m_j}$, where $m_1 < m_2$. The two remaining partial derivatives are

$$(2-12) \quad \frac{\partial F}{\partial \alpha_2} = \sum_{j \neq m_1, m_2} (d-j)x^{d-j-1}y^{j+1}, \quad \frac{\partial F}{\partial \alpha_3} = \sum_{j \neq m_1, m_2} jx^{d-j+1}y^{j-1},$$

and so J contains as well the forms in (2-12) of the shape $c_1x^{d-m_1}y^{m_1} + c_2x^{d-m_2}y^{m_2}$. We need to distinguish a number of cases. If $m_1 = 0, m_2 = d$, then these forms are y^d, x^d . If $m_1 = 0$ and $2 \leq m_2 \leq d-1$, then these forms are $(d-m_2)x^{d-m_2}y^{m_2}$ and $x^d + (m_2+1)x^{d-m_2}y^{m_2}$, and similarly when $1 \leq m_1 \leq d-2$ and $m_2 = d$. (Recall that we have excluded the cases $(m_1, m_2) = (0, 1)$ and $(d-1, d)$). In the remaining cases, $1 \leq m_1 < m_2 \leq d-1$. If $m_2 = m_1 + 1$, then these forms are $(d-(m_1-1))x^{d-m_1}y^{m_1}$ and $(m_2+1)x^{d-m_2}y^{m_2}$. Finally, if $m_2 > m_1 + 1$, then all four terms appear, and the forms are

$$(2-13) \quad (d-m_1+1)x^{d-m_1}y^{m_1} + (d-m_2+1)x^{d-m_2}y^{m_2}, \\ (m_1+1)x^{d-m_1}y^{m_1} + (m_2+1)x^{d-m_2}y^{m_2}.$$

In each of the cases, linear combinations of the forms produce the missing monomials, so $J = H_d(\mathbb{C}^2)$. \square

Remark. By writing $p(x, y) = \prod_k(x + \alpha_k y)$, it follows from Theorem 1.9 that, for a general set of d complex numbers α_k , there exists a Möbius transformation T so that

$$(2-14) \quad \sum_{k=1}^d T(\alpha_k) = 0, \quad \sum_{k=1}^d T\left(\frac{1}{\alpha_k}\right) = 0, \quad \prod_{k=1}^d T(\alpha_k) = 1.$$

Nonconstructive proof of Theorem 1.10. Write (1-11) as $F(x; t)$, where

$$q(x_1, x_2, x_3) = t_1 x_1^2 + t_2 x_2^2 + t_3 x_3^2 + t_4 x_1 x_2 + t_5 x_1 x_3 + t_6 x_2 x_3,$$

$$l_k(x_1, x_2, x_3) = t_{k1} x_1 + t_{k2} x_2 + t_{k3} x_3.$$

Evaluate the partials at $q = x_1 x_2 + x_1 x_3 + x_2 x_3$ and $(l_1, l_2, l_3) = (x_1, x_2, x_3)$. Then $\partial F / \partial t_{kl} = 4x_l x_k^3$, so $x_i^4, x_i^3 x_j \in J$; since $\partial F / \partial t_1 = 2x_1^2 q = 2x_1^2(x_1 x_2 + x_1 x_3 + x_2 x_3)$, it follows that $x_1^2 x_2 x_3 \in J$. Similarly, by considering $\partial F / \partial t_2$ and $\partial F / \partial t_3$, it follows that $x_1 x_2^2 x_3, x_1 x_2 x_3^2$ are in J . Finally, $\partial F / \partial t_4 = 2x_1 x_2 q = 2x_1 x_2(x_1 x_2 + x_1 x_3 + x_2 x_3)$, and so now $x_1^2 x_2^2 \in J$. Similarly, by considering $\partial F / \partial t_5$ and $\partial F / \partial t_6$, it follows that $x_1^2 x_3^2, x_2^2 x_3^2$ are also in J , and this accounts for all monomials in $H_4(\mathbb{C}^3)$. \square

Other applications of Corollary 2.3 to canonical forms can be found in [Ehrenborg and Rota 1993], including interpretations of the older results in [Richmond 1902; Turnbull 1960, pp. 265–269].

3. Apolarity and proofs of Theorems 1.1, 1.6 and 1.8

Using the notation of (2-1) and (2-2), for $p, q \in H_d(\mathbb{C}^n)$, define the following bilinear form:

$$(3-1) \quad [p, q] = \sum_{i \in \mathcal{J}(n, d)} c(i) a(p; i) a(q; i).$$

Recall two basic notations. For $\alpha \in \mathbb{C}^n$, define $(\alpha \cdot)^d \in H_d(\mathbb{C}^n)$ by

$$(3-2) \quad (\alpha \cdot)^d(x) = (\alpha \cdot x)^d = \left(\sum_{j=1}^n \alpha_j x_j \right)^d = \sum_{i \in \mathcal{J}(n, d)} c(i) \alpha^i x^i.$$

Define the differential operator $f(D)$ for $f \in H_e(\mathbb{C}^n)$ in the usual way by

$$(3-3) \quad f(D) = \sum_{i \in \mathcal{J}(n, e)} c(i) a(f; i) \left(\frac{\partial}{\partial x_1} \right)^{i_1} \cdots \left(\frac{\partial}{\partial x_n} \right)^{i_n}.$$

It follows immediately that, for $\alpha \in \mathbb{C}^n$,

$$(3-4) \quad [p, (\alpha \cdot)^d] = \sum_{i \in \mathcal{J}(n, d)} c(i) a(p; i) \alpha^i = p(\alpha).$$

If $i \neq j \in \mathcal{J}(n, d)$, then $i_k > j_k$ for some k , so $D^i x^j = 0$; otherwise $D^i x^i = \prod_k (i_k)! = d!/c(i)$. Suppose $p, q \in H_d(\mathbb{C}^n)$. Bilinearity and (3-3) imply the classical result that

$$\begin{aligned}
 (3-5) \quad p(D)q &= \sum_{i \in \mathcal{J}(n, d)} c(i)a(p; i)D^i \left(\sum_{j \in \mathcal{J}(n, d)} c(j)a(q; j)x^j \right) \\
 &= \sum_{i \in \mathcal{J}(n, d)} \sum_{j \in \mathcal{J}(n, d)} c(i)c(j)a(p; i)a(q; j)D^i x^j \\
 &= \sum_{i \in \mathcal{J}(n, d)} c(i)c(i)a(p; i)a(q; i)D^i x^i \\
 &= \sum_{i \in \mathcal{J}(n, d)} c(i)^2 a(p; i)a(q; i) \frac{d!}{c(i)} = d! [p, q] = d! [q, p] = q(D)p.
 \end{aligned}$$

Definition. If $p \in H_d(\mathbb{C}^n)$ and $q \in H_e(\mathbb{C}^n)$, then p and q are *apolar* if $p(D)q = q(D)p = 0$.

Note that, if $d = e$, then p and q are apolar if and only if $[p, q] = 0$, and, if $d > e$, say, then the equation $p(D)q = 0$ is automatic, so only $q(D)p = 0$ need be checked. By (3-4), p is apolar to $(\alpha \cdot)^d$ if and only if $p(\alpha) = 0$.

The following lemma is both essential and trivial.

Lemma 3.1. *Suppose $X = \text{span}(\{h_j\}) \subseteq H_d(\mathbb{C}^n)$. Then $X = H_d(\mathbb{C}^n)$ if and only if there is no $0 \neq p \in H_d(\mathbb{C}^n)$ which is apolar to each of the h_j .*

From this point of view, Theorem 3.2 is a direct consequence of Corollary 2.3:

Theorem 3.2 (Lasker–Wakeford). *If $F : \mathbb{C}^N \rightarrow H_d(\mathbb{C}^n)$, then F is a canonical form if and only if there is a point u so that there is no nonzero form $q \in H_d(\mathbb{C}^n)$ which is apolar to all N forms $\{\partial F / \partial t_k(u)\}$.*

The attribution “Lasker–Wakeford” (for [Lasker 1904; Wakeford 1920]) is taken from [Turnbull 1960]: H. W. Turnbull (1885–1961) was one of the last practicing invariant theorists who had been trained in the pre-Hilbert approach; see [Fisher 1967, pp. 231–232]. (His text [Turnbull 1960] is a Rosetta Stone for understanding the 19th century approach to algebra in more modern terminology.) Turnbull referred to Theorem 3.2 as “paradoxical and very curious”. E. Lasker (1868–1941) received his Ph.D. under M. Noether at Göttingen in 1902. He is probably better known for being the world chess champion for 27 years (1894–1921), spanning the life of E. K. Wakeford (1894–1916). J. H. Grace, Wakeford’s professor at Oxford, edited the second half of his thesis into [Wakeford 1920] and also wrote a memorial article [Grace 1918] for him in 1918:

“He [EKW] was slightly wounded early in 1916, and soon after coming home was busy again with Canonical Forms. . . [H]e discovered a paper

of Hilbert’s which contained the very theorem he had long been in want of — first vaguely, and later quite definitely. This was in March; April found him, full of the most joyous and reverential admiration for the great German master, working away in fearful haste to finish the dissertation. . . He returned to the front in June and was killed in July. . . He only needed a chance, and he never got it.”

The following properties are easily established; see, e.g., [Reznick 1992a; 1996] for proofs.

Theorem 3.3. (i) *If $e \leq d$ and $f \in H_e(\mathbb{C}^n)$, $g \in H_{d-e}(\mathbb{C}^n)$ and $p \in H_d(\mathbb{C}^n)$, then*

$$(3-6) \quad d! [fg, p] = (fg)(D)p = f(D)g(D)p = e! [f, g(D)p].$$

Thus, p is apolar to every multiple of g in $H_d(\mathbb{C}^n)$ if and only if p and g are apolar.

(ii) *If $p \in H_d(\mathbb{C}^n)$, then $(1/d)\partial p/\partial x_j(\alpha) = [p, x_j(\alpha \cdot)^{d-1}]$. Thus, p is apolar to $(\alpha \cdot)^{d-1}$ if and only if p is singular at α . More generally, p is apolar to $(\alpha \cdot)^{d-e}$ if and only if p vanishes to e -th order at α .*

(iii) *If $e \leq d$ and $g \in H_{d-e}(\mathbb{C}^n)$, then $g(D)(\alpha \cdot)^d = (d!/e!)g(\alpha)(\alpha \cdot)^e$.*

Suppose $F(t; x)$ contains h^s as a summand, where $h(x) = \sum_{l \in \mathcal{F}(n,e)} t_l x^l$, and suppose that no t_l occurs elsewhere in $F(t; x)$. If p is apolar to each partial of F , then it will be apolar to $\partial F/\partial t_l = s x^l h^{s-1}$ by (2-5). Since this is true for every $l \in \mathcal{F}(n, e)$, it follows from (i) that p is apolar to h^{s-1} . It is critical to note that this observation requires that each of the monomials of degree e appear in h , and does not apply if h is defined as a sum from a restricted set of monomials.

We are now able to give a short proof of the “second main theorem on apolarity” from [Ehrenborg and Rota 1993], which was not concerned with preserving the constant count.

Theorem 3.4. *Suppose $j_l = (j_{l,1}, \dots, j_{l,m})$, $1 \leq l \leq r$, are m -tuples of nonnegative integers, and suppose positive integers d_k , $1 \leq k \leq m$, and d are chosen so that*

$$(3-7) \quad u_l := d - \sum_{k=1}^m j_{l,k} d_k \geq 0$$

for each l . Fix forms $q_l \in H_{u_l}(\mathbb{C}^n)$ and, for $f_k \in H_{d_k}(\mathbb{C}^n)$, define

$$(3-8) \quad F(f_1, \dots, f_m) = \sum_{l=1}^r q_l(x) f_1^{j_{l,1}} \cdots f_m^{j_{l,m}}.$$

Let $F_j := \partial F/\partial f_j$. Then a general $p \in H_d(\mathbb{C}^n)$ can be written as (3-8) if and only if there exists a specific $\bar{f} = (\bar{f}_k)$ so that no nonzero $p \in H_d(\mathbb{C}^n)$ is apolar to

each $F_j(\bar{f})$, $1 \leq j \leq m$. If, in addition,

$$(3-9) \quad \sum_{k=1}^m N(n, d_k) = N(n, d),$$

then (3-8) is a canonical form.

Proof. Let

$$(3-10) \quad f_j(x) = \sum_{i_v \in \mathcal{G}(n, d_j)} t_{j,v} x^{i_v}.$$

By Theorem 2.2, (3-7) and Lemma 3.1, (3-8) represents general $p \in H_d(\mathbb{C}^n)$ if and only if there is some \bar{f} so that there is no nonzero form in $p \in H_d(\mathbb{C}^n)$ which is apolar to each $\partial F / \partial t_{j,v}(\bar{f}) = d_k x^{i_v} F_j(\bar{f})$, or, by Theorem 3.3(i), to each $F_j(\bar{f})$. The constant count is checked by (3-9). \square

By Theorem 3.3(ii) and Theorem 3.4,

$$F = \sum_{k=1}^r (\alpha_k \cdot)^d$$

is a canonical form if and only if there exist r points $\bar{\alpha}_k \in \mathbb{C}^n$ at which no nonzero form $p \in H_d(\mathbb{C}^n)$ is singular. This result is classical, and goes back to [Clebsch 1861]; see also [Ehrenborg and Rota 1993, Theorem 4.2]. A particularly deep result of Alexander and Hirschowitz [1995] states that a general form in $H_d(\mathbb{C}^n)$, $d \geq 3$, may be written as a sum of $\lceil N(n, d)/n \rceil$ d -th powers of linear forms, except when $(n, d) = (5, 3), (3, 4), (4, 4), (5, 4)$, when an extra summand is needed. (For much more on this, see [Geramita 1996, Lecture 7; Iarrobino and Kanev 1999, Corollary 1.62; Landsberg 2012, Chapter 15; Ranestad and Schreyer 2000, Theorem 0.2]; for a brief exposition of the proof, see [Landsberg 2012, Chapter 15].) These references also discuss the exceptional examples, which were all known in the 19th century. The expression of forms as a sum of powers of forms is currently a very active area of interest; see the references above as well as [Carlini et al. 2012; Fröberg et al. 2012; Landsberg and Teitler 2010].

The fundamental theorem of apolarity (see [Reznick 1996] for a history) states that, if f is irreducible and $p \in H_d(\mathbb{C}^n)$, then f and p are apolar if and only if p can be written as a sum of terms of the form $(\alpha_j \cdot)^d$, where $f(\alpha_j) = 0$. This was generalized as follows:

Theorem 3.5 [Reznick 1996, Theorem 4.1]. *Suppose $q \in H_e(\mathbb{C}^n)$ factors as $\prod_{j=1}^r q_j^{m_j}$ into a product of powers of distinct irreducible factors and suppose $p \in H_d(\mathbb{C}^n)$. Then $q(D)p = 0$ if and only if there exist $\alpha_{jk} \subset \{q_j(\alpha) = 0\}$ and*

$\phi_{jk} \in H_{m_j-1}(\mathbb{C}^n)$ such that

$$p = \sum_{j=1}^r \sum_{k=1}^{n_j} \phi_{jk}(\alpha_{jk} \cdot)^{d-(m_j-1)}.$$

The application of apolarity to binary forms is particularly simple, because zeros correspond to factors. If $e = d + 1$, then $q(D)p = 0$ for every $p \in H_d(\mathbb{C}^n)$, and we obtain the following result, also found in [Ehrenborg and Rota 1993, Theorem 4.5].

Corollary 3.6. *Suppose $\{\alpha_j x + \beta_j y : 1 \leq j \leq r\}$ is honest and suppose $\sum_{j=1}^r m_j = d + 1$. Then the following set is a basis for $H_d(\mathbb{C}^2)$:*

$$(3-11) \quad \mathcal{S} = \{x^k y^{m_j-1-k} (\beta_j x - \alpha_j y)^{d-m_j+1} : 0 \leq k \leq m_j - 1, 1 \leq m_j \leq r\}.$$

Proof. If p is apolar to each term in (3-11), then $(\alpha_j x + \beta_j y)^{m_j} \mid p$ by Theorem 3.3(ii). Thus $p = 0$ by degree considerations, and \mathcal{S} has $d + 1$ elements, so it is a basis. \square

If each $m_j = 1$, then Corollary 3.6 states that an honest set $\mathcal{S} = \{(\alpha_j x + \beta_j y)^d\}$ of $d + 1$ forms is a basis for $H_d(\mathbb{C}^2)$. This is easily proved directly, since the representation of \mathcal{S} with respect to the basis $\{ \binom{d}{j} x^{d-j} y^j \}, [\alpha_j^{d-k} \beta_j^k]$, has Vandermonde determinant

$$(3-12) \quad \prod_{1 \leq i < j \leq n} (\alpha_i \beta_j - \alpha_j \beta_i).$$

Each product in (3-12) is nonzero because $\{(\alpha_j x + \beta_j y)^d\}$ is honest. One implication of this independence is found in [Reznick 2013, Corollary 4.3].

Lemma 3.7. *If $p(x, y) \in H_d(\mathbb{C}^2)$ has two honest representations*

$$(3-13) \quad p(x, y) = \sum_{i=1}^m (\alpha_i x + \beta_i y)^d = \sum_{j=1}^n (\gamma_j x + \delta_j y)^d$$

and $m + n \leq d + 1$, then the representations are permutations of each other.

Proof. If (3-13) holds, then $\{(\alpha_i x + \beta_i y)^d, (\gamma_j x + \delta_j y)^d\}$ is linearly dependent, which is impossible unless the dependence is trivial. \square

It follows immediately from Lemma 3.7 that the representations (1-2) and (1-3), if they exist for p , are unique. When $n \geq 3$, the linear dependence of a set $\{(\alpha_j \cdot)^d\}$ depends on the geometry of the points as well as the number (see the discussion of Serret’s theorem in [Reznick 1992a, p. 29].) Even for powers of binary forms of degree $e \geq 2$, there are singular cases. It is not hard to show that a *general* set of $(2k + 1)$ k -th powers of quadratic forms is linearly independent; however, for example, $(x^2 - y^2)^2 + (2xy)^2 = (x^2 + y^2)^2$. For much more on this, see [Reznick \geq 2013].

Nonconstructive proof of Theorem 1.6. For $1 \leq k \leq r$, write

$$f_k(x, y) = \sum_{l=0}^{e_k} t_{k,l} x^{e_k-l} y^l.$$

By Corollary 2.3 and (2-5), (1-8) is a canonical form in the variables $\{t_j, t_{k,l}\}$ provided there is a point at which the partials

$$\{l_j^d, 1 \leq j \leq m\} \cup \{x^{e_k-l} y^l f_k^{d/e_k-1}, 1 \leq l \leq e_k, 1 \leq k \leq r\}$$

span $H_d(\mathbb{C}^2)$. Let $f_k = \tilde{l}_k^{e_k}$, where $\{l_1, \dots, l_m, \tilde{l}_1, \dots, \tilde{l}_r\}$ is chosen to be honest. Then, by (1-7), the desired assertion follows immediately from Corollary 3.6. \square

Nonconstructive proof of Theorem 1.8. Write $uv + 1 = r(u + 1) + s$. If $s = 0$, then Theorem 1.8 is simply a special case of Theorem 1.6 with $m = 0$, $d = uv$ and $e_k \equiv u$. Otherwise, $1 \leq s \leq u$, so that $r + 1 = \lceil (uv + 1)/(u + 1) \rceil$. Let

$$F(\{\alpha_{ij}\}) = \sum_{i=1}^{r+1} f_i^v(x, y), \quad f_i(x, y) = \sum_{j=0}^u \alpha_{ij} x^{u-j} y^j.$$

This is *not* a canonical form, as there are too many constants. As before,

$$\frac{\partial F}{\partial \alpha_{ij}} = v x^{u-j} y^j f_i^{v-1}.$$

We now specialize to $f_i(x, y) = (ix - y)^u$ and use the apolarity argument to show that $J = H_{uv}(\mathbb{C}^2)$. Suppose $q \in H_{uv}(\mathbb{C}^2)$ is apolar to each partial. Then, by Theorem 3.3, it is apolar to $f_i^{v-1} = (ix - y)^{u(v-1)}$, and so q vanishes to u -th order at $(i, -1)$ for $1 \leq i \leq r + 1$. It follows that q is a multiple of $\prod_{i=1}^{r+1} (x + iy)^{u+1}$, and so $q = 0$ by degree considerations.

It is an exercise to show that F can be converted to an canonical form by requiring, say, that f_{r+1} only contain monomials $x^{u-j} y^j$ for $0 \leq j \leq s - 1$. \square

We present now Sylvester’s algorithm. For modern discussions of this, along with Gundelfinger’s generalization [1887], which is not included here, see [Kung and Rota 1984, Section 5; Kung 1986; 1987; 1990; Reznick 1996; 2013].

Theorem 3.8 (Sylvester’s algorithm). *Let*

$$p(x, y) = \sum_{j=0}^d \binom{d}{j} a_j x^{d-j} y^j$$

be a given binary form and suppose $\{\alpha_j x + \beta_j y\}$ is honest. Let

$$h(x, y) = \sum_{t=0}^r c_t x^{r-t} y^t = \prod_{j=1}^r (\beta_j x - \alpha_j y).$$

Then there exist $\lambda_k \in \mathbb{C}$ so that

$$p(x, y) = \sum_{k=1}^r \lambda_k (\alpha_k x + \beta_k y)^d$$

if and only if

$$(3-14) \quad \begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r+1} & \cdots & a_d \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Theorem 3.8 can be put in the context of our previous discussion. Let $A_r(p)$ denote the $(d - r + 1) \times (r + 1)$ Hankel matrix on the left-hand side of (3-14). If $h(D) = \prod_{j=1}^r (\beta_j \partial/\partial x - \alpha_j \partial/\partial y)$, then a direct computation shows that

$$(3-15) \quad h(D)p = \sum_{m=0}^{d-r} \frac{d!}{(d-r-m)!m!} \left(\sum_{i=0}^{d-r} a_{i+m} c_i \right) x^{d-r-m} y^m.$$

It follows from (3-15) that the coefficients of $h(D)p$ are thus, up to multiple, the rows of the matrix product, so (3-14) is equivalent to $h(D)p = 0$. In this way, Theorem 3.8 follows from Theorem 3.5. Sylvester’s algorithm can also be visualized as seeking constant-coefficient linear recurrences satisfied by $\{a_k\}$ and looking for the shortest one whose characteristic equation has distinct roots; this is the proof given in [Reznick 2013]. In this case, Gundelfinger’s results handle the case when the roots are not distinct.

Constructive proof of Theorem 1.1. Suppose $d = 2s - 1$ is odd. The matrix $A_s(p)$ is $s \times (s + 1)$ and has a nontrivial null-vector. The corresponding h (which can be given in terms of the coefficients of p) has distinct factors unless its discriminant vanishes. Thus, for general $p \in H_{2s-1}(\mathbb{C}^2)$, Theorem 3.8 gives p as a sum of s $(2s - 1)$ -st powers of linear forms.

If $d = 2s$, the matrix $A_s(p)$ is square, and if p is a sum of s $2s$ -th powers, then $\det A_s(p) = 0$. Conversely, if $\det A_s(p) = 0$ and the corresponding h has distinct factors (which is generally true), then p is a sum of s $2s$ -th powers. If M_1 and M_2 are two square matrices and $\text{rank}(M_2) = k$, then $\det(M_1 + \lambda M_2)$ is a polynomial in λ of degree k . In particular, if $q = (\alpha x + \beta y)^{2s}$, then $\text{rank}(H_s(q)) = 1$. Thus, in general, there is a unique value of λ and some matrix M so that $0 = \det A_s(p - \lambda(\alpha x + \beta y)^{2s}) = \det A_s(p) - \lambda \det M$. (When $\alpha x + \beta y = x$, M is the $(1, 1)$ -cofactor of $A_s(p)$.) In the special case $\alpha x + \beta y = x$, this proves Theorem 1.1(ii). The same argument shows that, for general $q \in H_{2s}(\mathbb{C}^2)$, there exist $s + 1$ values of λ so that $p - \lambda q$ is a sum of s $2s$ -th powers. \square

Sylvester [1870] recalled his discovery of this algorithm and its consequences.

“I discovered and developed the whole theory of canonical binary forms for odd degrees, and, as far as yet made out, for even degrees too, at one evening sitting, with a decanter of port wine to sustain nature’s flagging energies, in a back office in Lincoln’s Inn Fields. The work was done, and well done, but at the usual cost of racking thought — a brain on fire, and feet feeling, or feelingless, as if plunged in an ice-pail. *That night we slept no more.*”

Example 3.1. This example of Sylvester’s algorithm will be used in Example 4.1. Let

$$\begin{aligned} p(x, y) &= 2x^3 + 3x^2y - 21xy^2 - 41y^3 \\ &= \binom{3}{0} \cdot 2 x^3 + \binom{3}{1} \cdot 1 x^2y + \binom{3}{2} \cdot (-7) xy^2 + \binom{3}{3} \cdot (-41) y^3. \end{aligned}$$

Since

$$\begin{pmatrix} 2 & 1 & -7 \\ 1 & -7 & -41 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ -5 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

we have $h(x, y) = 6x^2 - 5xy + y^2 = (2x - y)(3x - y)$. It now follows that $p(x, y) = \lambda_1(x + 2y)^3 + \lambda_2(x + 3y)^3$, and a simple computation shows that $\lambda_1 = 5$, $\lambda_2 = -3$.

Lemma 3.1, when applied to Theorem 2.1, yields the following corollary.

Corollary 3.9. *A basis for $H_d(\mathbb{C}^n)$ is given by $\{(i \cdot)^d : i \in \mathcal{F}(n, d)\}$.*

This in turn gives a very weak version of the Alexander–Hirschowitz theorem:

Corollary 3.10. *A general form in $H_d(\mathbb{C}^n)$ is a sum of*

$$N(n, d - 1) = \frac{nd}{n + d - 1} \cdot \frac{1}{n} N(n, d)$$

d -th powers of linear forms.

Proof. Consider the sum

$$\sum_{l=1}^{N(n, d-1)} (t_{l,1}x_1 + \cdots + t_{l,n}x_n)^d,$$

and apply Corollary 2.3 with t_l specialized to $i_l \in \mathcal{F}(n, d - 1)$. Then J contains $x_k(i_l \cdot)^{d-1}$ for each k, l and hence $x_k H_{d-1}(\mathbb{C}^n) \subseteq J$ for each k , so $J = H_d(\mathbb{C}^n)$. \square

4. Examples of binary canonical forms and the proof of Theorem 1.6

This section is devoted to special cases of Theorem 1.6. First, in the special case $e_k = 1$, we give a constructive proof showing uniqueness, which gives a kind of interpolation between Sylvester’s theorem and the representations of $H_d(\mathbb{C}^2)$ by (2-4) with a fixed basis consisting of d -th powers, as in Corollary 3.6.

Corollary 4.1. *Suppose $d \geq 1$, and $\{l_j(x, y) = \alpha_j x + \beta_j y\}$ is a fixed honest set of $m = d + 1 - 2r$ linear forms. Then a general binary d -ic form $p \in H_d(\mathbb{C}^2)$ can be written uniquely as*

$$(4-1) \quad p(x, y) = \sum_{j=1}^m t_j l_j(x, y)^d + \sum_{k=1}^r (t_{k1}x + t_{k2}y)^d$$

for suitable $t_{k1}, t_{k2} \in \mathbb{C}$.

Proof. Let

$$f(x, y) = \prod_{j=1}^m (\beta_j x - \alpha_j y).$$

Then $f(D)p$ has degree $d - m = 2r - 1$ and by Theorem 3.8 generally has a unique representation as a sum of r $(2r - 1)$ -st powers of linear forms, say

$$(4-2) \quad f(D)p = \sum_{k=1}^r (u_{k1}x + u_{k2}y)^{2r-1}.$$

Further, it is generally true that $f(u_{k1}, u_{k2}) \neq 0$. Let

$$(4-3) \quad q(x, y) = \frac{(2r - 1)!}{d!} \sum_{k=1}^r \frac{(u_{k1}x + u_{k2}y)^d}{f(u_{k1}, u_{k2})}.$$

It follows from Theorem 3.3(iii), (4-2) and (4-3) that $f(D)p = f(D)q$. Since f has distinct factors, it then follows from Theorem 3.8 that there exist $t_j \in \mathbb{C}$ so that

$$p(x, y) - q(x, y) = \sum_{j=1}^m t_j (\alpha_j x + \beta_j y)^d.$$

Conversely, suppose p has two different representations:

$$(4-4) \quad \sum_{j=1}^m t_j l_j^d(x, y) + \sum_{k=1}^r (t_{k1}x + t_{k2}y)^d = \sum_{j=1}^m \tilde{t}_j l_j^d(x, y) + \sum_{k=1}^r (\tilde{t}_{k1}x + \tilde{t}_{k2}y)^d.$$

By combining the first sum on each side, (4-4) becomes a linear dependence with $m + 2r = d + 1$ summands, which by Lemma 3.7 must be trivial; thus, the representations in (4-4) are essentially the same. □

Example 4.1. Let $l_1(x, y) = x + y$ and $l_2(x, y) = -x + 3y$ and let

$$p(x, y) = -x^5 + 15x^4y - 170x^3y^2 + 390x^2y^3 - 505x^2y^3 + 483y^5.$$

In an application of the last proof, $f(x, y) = (x - y)(3x + y) = 3x^2 - 2xy - y^2$, and

$$3\frac{\partial^2 p}{\partial x^2} - 2\frac{\partial^2 p}{\partial x \partial y} - \frac{\partial^2 p}{\partial y^2} = 160x^3 + 240x^2y - 1680xy^2 - 3280y^3.$$

Example 3.1 implies that this expression equals $400(x + 2y)^3 - 240(x + 3y)^3$. Since $f(1, 2) = -5$ and $f(1, 3) = -12$, it follows that

$$\begin{aligned} p(x, y) &= \frac{3! \cdot 400}{5! \cdot (-5)}(x + 2y)^5 + \frac{3! \cdot (-240)}{5! \cdot (-12)}(x + 3y)^5 + t_1(x + y)^5 + t_2(-x + 3y)^5 \\ &= -4(x + 2y)^5 + (x + 3y)^5 + t_1(x + y)^5 + t_2(-x + 3y)^5 \end{aligned}$$

and it can be readily be computed that $t_1 = \frac{7}{2}$ and $t_2 = \frac{3}{2}$.

If each $e_k = 2$ in Theorem 1.6 and m is as small as possible, then we obtain an analogue of Sylvester's theorem for forms of even degree.

Corollary 4.2. (i) *A general binary form of degree $d = 6s$ can be written as*

$$(4-5) \quad \lambda x^{6s} + \sum_{j=1}^{2s} (\alpha_j x^2 + \beta_j xy + \gamma_j y^2)^{3s} \quad \text{for some } \lambda \in \mathbb{C}.$$

(ii) *A general binary form of degree $d = 6s + 2$ can be written as*

$$(4-6) \quad \sum_{j=1}^{2s+1} (\alpha_j x^2 + \beta_j xy + \gamma_j y^2)^{3s+1}.$$

(iii) *A general binary form of degree $d = 6s + 4$ can be written as*

$$(4-7) \quad \lambda_1 x^{6s+4} + \lambda_2 y^{6s+4} + \sum_{j=1}^{2s+1} (\alpha_j x^2 + \beta_j xy + \gamma_j y^2)^{3s+2} \quad \text{for some } \lambda_i \in \mathbb{C}.$$

We have not been able to find an analogue to Sylvester's algorithm for determining the representations (4-5), (4-6), (4-7) in Corollary 4.2. In the linear case, $(\alpha x + \beta y)^d$ is killed by $\beta \partial / \partial x - \alpha \partial / \partial y$, and two operators of this shape commute. Although each $(\alpha x^2 + 2\beta xy + \gamma y^2)^d$ is killed by the nonconstant-coefficient $(\beta x + \gamma y) \partial / \partial x - (\alpha x + \beta y) \partial / \partial y$, two operators of this kind do not usually commute. The smallest constant-coefficient differential operator which kills $(\alpha x^2 + 2\beta xy + \gamma y^2)^d$ has degree $d + 1$; the product of any two of these would kill every form of degree $2d$ and so provide no information.

Let us say that (1-8) is a *neat* canonical form if $m = 0$, and of *Sylvester-type* if it is neat and if $e_k = e$ for $1 \leq k \leq r$. Counting the numbers of neat and Sylvester-type canonical forms leads to some number theory. The first lemma is standard.

Lemma 4.3. *Given $0 < p/q \in \mathbb{Q}$ and $0 < n \in \mathbb{N}$, there exist only finitely many choices of $m_j \in \mathbb{Z}$, $0 < m_1 \leq m_2 \cdots \leq m_n$, such that $p/q = \sum_{j=1}^n 1/m_j$.*

Proof. If $n = 2$, then $p/q > 1/m_1 \geq p/(2q)$ implies that there are finitely many integral choices for m_1 , each of which determines $m_2 = (p/q - 1/m_1)^{-1}$. Supposing the lemma valid for $n - 1$, we have $p/q > 1/m_1 \geq p/(nq)$, and each choice of m_1 implies the equation $p/q - 1/m_1 = \sum_{j=2}^n 1/m_j$. This has finitely many solutions by the induction hypothesis. □

Theorem 4.4. *For a fixed value of r , there are only finitely many neat canonical forms (1-8) with r summands.*

Proof. Suppose $m = 0$ in Theorem 1.6. Write $d = e_k m_k$; then, by (1-7),

$$d + 1 = \sum_{k=1}^r \left(\frac{d}{m_k} + 1 \right),$$

which implies

$$(4-8) \quad 1 = \sum_{k=1}^r \frac{1}{m_k} + \frac{r-1}{d} = \sum_{k=1}^r \frac{1}{m_k} + \sum_{l=1}^{r-1} \frac{1}{d}.$$

Now apply Lemma 4.3 with $p/q = 1$ and $n = 2r - 1$: there are only finitely many expressions of 1 as a sum of $2r - 1$ unit fractions, of which only a subset satisfy the additional restrictions of (4-8). □

It is not hard to work out that, for $r = 2$, there are three neat canonical forms: $(d, e_1, e_2) = (3, 1, 1)$, $(4, 2, 1)$ and $(6, 3, 2)$. The first is Theorem 1.1(i) with $d = 3$, the second is Corollary 1.7 with $d = 4$ (see Theorem 4.6 below), and the third is Theorem 1.5. When $r = 3$, there are twenty-two neat canonical forms.

Let $s(d)$ denote the number of neat Sylvester-type canonical forms of degree d . Suppose $e_k = e$ for all k in one of these. Then $e \mid d$ and, by (1-7), $r(e + 1) = d + 1$, so $(e + 1) \mid (d + 1)$. Since $d \equiv 0 \pmod{e}$ and $d \equiv -1 \pmod{(e + 1)}$, it follows from the Chinese remainder theorem that $d \equiv e \pmod{e(e + 1)}$; that is, $d = e + ue(e + 1)$, $u \geq 1$, so that $e < \sqrt{d}$.

Theorem 4.5. *Let $S(N) := \sum_{d=1}^N s(d)$. Then $S(N) = N + \mathcal{O}(N^{1/2})$, $\sup_d s(d) = \infty$.*

Proof. The generating function for the sequence $(s(d))$ is

$$(4-9) \quad \sum_{n=1}^{\infty} s(d)x^d = \sum_{e=1}^{\infty} \sum_{u=1}^{\infty} x^{e+ue(e+1)} = \sum_{e=1}^{\infty} \frac{x^{e^2+2e}}{1 - x^{e^2+e}} = \sum_{N=e}^{\infty} \left\lfloor \frac{N - e}{e^2 + e} \right\rfloor X^N.$$

Let $T = \lfloor N^{1/2} \rfloor$. It follows from (4-9) that

$$(4-10) \quad S(N) = \sum_{n=1}^N s_n = \sum_{e=1}^{\infty} \left\lfloor \frac{N-e}{e^2+e} \right\rfloor = \sum_{e=1}^T \left\lfloor \frac{N-e}{e^2+e} \right\rfloor.$$

Thus, using the telescoping sum for $\sum \frac{1}{e(e+1)}$, (4-10) implies that

$$(4-11) \quad \begin{aligned} S(N) &\leq \sum_{e=1}^T \frac{N-e}{e^2+e} = N \sum_{e=1}^T \frac{1}{e^2+e} - \sum_{e=1}^T \frac{1}{e+1} \\ &\leq N \left(1 - \frac{1}{T+1} \right) - \log T + \mathcal{O}(1) = N - N^{1/2} + \mathcal{O}(\log N). \end{aligned}$$

The lower bound is the same, minus T , so (4-11) implies that $S(N) = N + \mathcal{O}(N^{1/2})$.

Now, $s(d)$ counts the number of $e < d$ so that e divides d and $e + 1$ divides $d + 1$. If $d = 2^r - 1$, then $e + 1 \mid 2^r$ implies that $e + 1 = 2^t$ for some $t < r$. But $2^t - 1 \mid 2^r - 1$ if and only if $t \mid r$; hence $s(2^r - 1) = d(r) - 1$, where $d(n)$ denotes the divisor function. In particular, $s(2^{2^t} - 1) = t$, so the sequence $(s(d))$ is unbounded. More generally, if $e \mid d$ and $e + 1 \mid d + 1$, then $e \mid d^2 + 2d$ and $e + 1 \mid d^2 + 2d + 1$, and since $e = d$ contributes to the count in $s(d^2 + 2d)$ but not in $s(d)$, $s(d^2 + 2d) \geq s(d) + 1$. \square

Half of the neat Sylvester forms come from Theorem 1.1(i), another sixth come from Corollary 4.2(ii), etc. The smallest d for which $s(d) = 2$ is $d = 15$: $(e, r) = (1, 8), (3, 4)$, so a general binary form of degree 15 is a sum of eight linear forms to the 15th power, or four cubics to the 5th power. Mathematica computations show that the smallest d for which $s(d) = 3$ is $d = 99$: $(e, r) = (1, 50), (3, 25), (9, 10)$. For $d < 10^7$, the largest value of $s(d)$ is $s(7316000) = 12$. Note that $2^{2^{13}} - 1 = 2^{4096} - 1 \approx 1.04 \times 10^{1233}$, so the examples given in the proof are not likely to describe the fastest growth. We conjecture as well that $\{s(d)\}$ has an underlying distribution.

If the degree d is prime, then Corollary 4.1 accounts for all canonical forms in Theorem 1.6. The smallest d which is not covered by Corollary 4.1 is then $d = 4$, and there are two such cases, one of which is neat: $e_1 = 2, e_2 = 1, m = 0$ and $e_1 = 2, m = 2$. Both can be discussed constructively.

Theorem 4.6. *A general binary quartic $p \in H_4(\mathbb{C}^2)$ can be written as*

$$(4-12) \quad p(x, y) = (t_1x^2 + t_2xy + t_3y^2)^2 + (t_4x + t_5y)^4$$

in six different ways. Further, the set of possible values for $\{t_5/t_4\}$ is the image of the set $\{0, \infty, 1, -1, i, -i\}$ under a Möbius transformation.

Proof. By Theorem 2.4, if p is a general binary quartic, then there exist c_i, λ so that $p(c_1x + c_2y, c_3x + c_4y) = p_\lambda(x, y) := x^4 + 6\lambda x^2y^2 + y^4$. If (4-12) holds for p_λ ,

then

$$(4-13) \quad \begin{aligned} 1 &= t_1^2 + t_4^4, & 0 &= 2t_1t_2 + 4t_4^3t_5, & 6\lambda &= 2t_1t_3 + t_2^2 + 6t_4^2t_5^2, \\ & & 0 &= 2t_2t_3 + 4t_4t_5^3, & 1 &= t_3^2 + t_5^4. \end{aligned}$$

First suppose that $t_4 = 0$. Then (4-13) implies that $1 = t_1^2$ and $0 = 2t_1t_2$, so $t_1 = 1$ (without loss of generality) and $t_2 = 0$. The remaining equations imply that $t_3 = 3\lambda$ and $t_5^4 = 1 - 9\lambda^2$. A similar argument works if $t_5 = 0$, giving two representations:

$$(4-14) \quad p_\lambda(x, y) = (x^2 + 3\lambda y^2)^2 + (1 - 9\lambda^2)y^4 = (3\lambda x^2 + y^2)^2 + (1 - 9\lambda^2)x^4.$$

Now suppose $t_4t_5 \neq 0$, so $t_1t_2t_3 \neq 0$ and we get successively

$$\frac{t_3}{t_1} = \frac{-2t_2t_3}{-2t_1t_2} = \frac{4t_4t_5^3}{4t_4^3t_5} = \frac{t_5^2}{t_4^2} \implies \frac{1 - t_3^2}{1 - t_1^2} = \frac{t_5^4}{t_4^4} = \frac{t_3^2}{t_1^2} \implies t_1^2 = t_3^2.$$

It follows that $t_5 = i^k t_4$ and $t_3 = (-1)^k t_1$, and (4-13) can be completely solved:

$$t_4^4 = 1 - t_1^2, \quad t_2 = 2i^k(t_1 - t_1^{-1}), \quad 2 + 6(-1)^k\lambda = 4t_1^{-2}.$$

After some massaging of the algebra, this gives four representations:

$$(4-15) \quad p_\lambda(x, y) = \left(\frac{(-1)^k 2}{3\lambda + (-1)^k} \right) (x^2 - i^{3k}(3\lambda - (-1)^k)xy + (-1)^k y^2)^2 + \left(\frac{3\lambda - (-1)^k}{3\lambda + (-1)^k} \right) (x + i^k y)^4, \quad k = 0, 1, 2, 3.$$

In order to find the six representations of p as (4-12), we start with the six representations of p_λ given in (4-14) and (4-15), in which $t_4x + t_5y$ is a multiple of one of the six linear forms $x, y, x + i^k y$. Apply the inverse of the map $(x, y) \mapsto (c_1x + c_2y, c_3x + c_4y)$, which takes $t_4x + t_5y$ to a multiple of $t_4(c_4x - c_2y) + t_5(-c_3x + c_1y)$: $t_5/t_4 \mapsto G(t_5/t_4)$, where $G(z) = (c_1z - c_2)/(c_4 - c_3z)$. \square

Theorem 4.7. *Given two fixed nonproportional binary linear forms l_1, l_2 , a general binary quartic in $H_4(\mathbb{C}^2)$ has two representations as*

$$(4-16) \quad p(x, y) = (t_1x^2 + t_2xy + t_3y^2)^2 + t_4l_1(x, y)^4 + t_5l_2(x, y)^4.$$

Proof. Given p, l_1, l_2 , make an invertible linear change of variable taking $(l_1, l_2) \mapsto (x, y)$, and suppose $p(x, y) \mapsto q(x, y) = \sum_i a_i x^{4-i} y^i$. Then q has the shape (4-16) if and only if the coefficients of x^3y, x^2y^2, xy^3 in $(t_1x^2 + t_2xy + t_3y^2)^2$ and q agree. Thus, we seek to solve the system

$$(4-17) \quad a_1 = 2t_1t_2, \quad a_2 = 2t_1t_3 + t_2^2, \quad a_3 = 2t_2t_3.$$

But (4-17) implies $a_1t_2^2 - 2a_2t_1t_2 + 2a_3t_1^2 = 0$; hence, in general, there are exactly two values of β so that $t_2 = \beta t_1$; in each case, $t_1^2 = a_1/(2\beta)$. The two choices of

d	e_1, \dots, e_r	m	$F(d; e)$	Source
any	$1^{\lfloor (d+1)/2 \rfloor}$	0 or 1	1	Theorem 1.1
any	1^r	$d + 1 - 2r$	1	Corollary 4.1
4	2, 1	0	6	Theorem 4.6
4	2	2	2	Theorem 4.7
6	3, 2	0	40	[Várilly-Alvarado 2008; 2011]
6	$2, 1^2$	0	22	Experiment
6	3, 1	1	14	Experiment
6	2^2	1	9	Experiment
6	2, 1	2	12	Experiment
6	3	3	5	Experiment
6	2	4	5	Experiment
8	$2, 1^3$	0	62	Experiment
10	$2, 1^4$	0	147	Experiment
12	$2, 1^5$	0	308	Experiment
$2s$	$2, 1^{s-1}$	0	$2\binom{s+3}{5} + \binom{s+2}{3}$	Conjecture

Table 1. Proved and conjectural values of $F(d; e)$.

sign for t_1 lead to the same square, and $t_3 = (a_1/a_3)t_1$, so (4-17) has these two solutions. □

In the case of Theorem 1.6 let $F(d; e_1, \dots, e_r)$ denote the number of different representations that a general $p \in H_d(\mathbb{C}^2)$ has, by our convention. We present in Table 1 a complete list of proved or conjectural values when $d \leq 6$, reflecting numerical experiments on Mathematica. (Recall that, if d is prime, then Corollary 4.1 presents all possible canonical forms of this type.) The conjectural value of $F(2s; 2, 1^{s-1})$ is suggested by the given data for $2 \leq s \leq 6$ and [OEIS 2013, A081282].

Várilly-Alvarado [2008; 2011] constructs explicitly all 240 representations of $x^6 + y^6$ as $f^2 + g^3$; he considers forms multiplied by roots of unity as different, which explains the appearance of $240/(2 \cdot 3)$ in the table above. This is also proved to be the number of representations for a general sextic.

To describe the experiments for $F(2s; 2, 1^{s-1})$ more precisely, we generate a form

$$p(x, y) = \sum_{k=0}^{2s} \binom{2s}{k} a_k x^{2s-k} y^k,$$

where $a_k = t + iu$ for random integers t, u in $[-100, 100]$. In case $s = 1$, we assume a change of variables so that the fixed linear forms are x^d or y^d ; for $s \geq 2$ we choose additional linear forms with random coefficients. Let $h(x, y) = Ux^2 + Vxy + Wy^2$ for variables (U, V, W) and let $q(x, y) = p(x, y) - h^s(x, y)$, and apply Sylvester's

algorithm to q . That is, we construct the $(s+2) \times s$ matrix $A_{s-1}(q)$, with polynomial entries in (U, V, W) of degree s and require that it have rank less than s . This is done by counting the number of (U, V, W) which are common zeros of all $s \times s$ minors. This number is divided by s to account for $h^s = (\zeta_s^k h)^s$. As a back of the envelope calculation, one might take the first $s - 1$ rows of A_{s-1} and use the cofactors to compute a nontrivial null-vector. Ignoring possible cancellation, the components would be polynomials of degree $s(s - 1)$ in (U, V, W) . Taking the dot product with the last three rows of A_{s-1} gives three polynomials of degree s^2 . Ignoring cancellations and multiplicity, there should be $(s^2)^3$ common zeros, and dividing by s gives an upper bound for $F(2s; 2, 1^{s-1})$ of s^5 . The conjectural value is asymptotically $\frac{1}{60}s^5$, which shows the same order of growth.

5. Quadratic forms and the proof of Theorem 1.2

We begin this section with a constructive proof of Theorem 1.2 which will serve as a template for constructive proofs involving cubic forms.

Constructive Proof of Theorem 1.2. Suppose $p \in H_2(\mathbb{C}^n)$, and, specifically,

$$p(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j.$$

Then $\partial p / \partial x_1 = 2 \sum_{j=1}^n a_{1j}x_j$. Since $a_{11} \neq 0$ in general, we can define

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \frac{1}{a_{11}} \left(\sum_{j=1}^n a_{1j}x_j \right)^2.$$

Observe that $\partial q / \partial x_1 = 0$, so $q = q(x_2, \dots, x_n)$. Iterating this argument gives the construction. There is only one linear form $\pm l$ so that $\partial p / \partial x_1 = 2l \partial l / \partial x_1$, so the representation is unique. □

Constant-counting for sums of squares is complicated by the action of the orthogonal group on a sum of t squares. If $M \in \text{Mat}_t(\mathbb{C})$ and $MM^t = I$, then

$$\sum_{i=1}^t f_i^2 = \sum_{i=1}^t \left(\sum_{j=1}^t m_{ij} f_j \right)^2.$$

When $t = 2$, choose $\theta \in \mathbb{C}$ and let $e^{i\theta} = \cos \theta + i \sin \theta := (u, v)$, so that

(5-1)
$$f^2 + g^2 = (uf - vg)^2 + (vf + ug)^2.$$

This means that we may safely remove one monomial from one of the summands.

Theorem 5.1. *A general binary form $p \in H_{2s}(\mathbb{C}^2)$ can be written as*

$$(5-2) \quad \left(\sum_{k=0}^s t_k x^{s-k} y^k \right)^2 + \left(\sum_{k=1}^s t_{s+k} x^{s-k} y^k \right)^2$$

in $\binom{2s-1}{s}$ different ways.

Proof. The nonconstructive proof is a simple application of Corollary 2.3. Writing (5-2) as $f^2 + g^2$ gives the partials with respect to the t_j as

$$\{2x^{s-k}y^k f, 0 \leq k \leq s\} \cup \{2x^{s-k}y^k g, 1 \leq k \leq s\};$$

specializing to $f = x^s$ and $g = y^s$ above gives all monomials in $H_{2s}(\mathbb{C}^2)$.

The more obvious expression

$$(5-3) \quad p(x, y) = f^2(x, y) + g^2(x, y), \quad g, h \in H_s(\mathbb{C}^2)$$

is not a canonical form, because $2(s+1) > 2s+1$. However, every sum of two squares can be formally factored, and these behave nicely with respect to (5-1):

$$\begin{aligned} f^2 + g^2 &= (f + ig)(f - ig) \iff \\ (uf + vg)^2 + (vf - ug)^2 &= (e^{i\theta}(f + ig))(e^{-i\theta}(f - ig)). \end{aligned}$$

Suppose $p(1, 0) = a_0 \neq 0$ (true for general p) and (5-3) holds, where $f(1, 0) = \rho$ and $g(1, 0) = \tau$. Then $\rho^2 + \tau^2 = a_0$, so that $\tau/\rho \neq \pm i$ and the coefficient of x^s in $vf + ug$ will be $v\rho + u\tau = \sin\theta\rho + \cos\theta\tau$, which is zero exactly when $\tan\theta = -\tau/\rho$. Thus, for precisely one value of $\tan\theta$, the right-hand side of (5-1) will be in the form (5-2). This determines a pair $(\pm u, \pm v)$; however, the squares in (5-2) will be the same.

In other words, each distinct factorization of p (up to multiple) as a product of two s -ic forms, when combined with the orthogonal action of (5-1), yields exactly one representation as (5-2). A general $p \in H_{2s}(\mathbb{C}^2)$ is a product of $2s$ distinct linear factors; these can be organized into an unordered pair of products of s distinct linear factors in $\frac{1}{2}\binom{2s}{s} = \binom{2s-1}{s}$ ways. \square

The “lost” degree of freedom in a sum of squares never arises in Theorem 1.6 because $2(d/2 + 1) > d + 1$. The missing monomial x^s in the second summand of (5-2) may be replaced by any specified monomial $x^{s-k_0}y^{k_0}$ by a similar argument.

Another classical result is that a general ternary quartic is a sum of three squares of quadratic forms, generally in 63 different ways up to the action of the orthogonal group (see [Powers et al. 2004]). Hilbert [1888] proved that *every* positive semidefinite $p \in H_4(\mathbb{R}^3)$ is a sum of three squares from $H_2(\mathbb{R}^3)$. He then showed that there exist psd forms in $H_6(\mathbb{R}^3)$ and $H_4(\mathbb{R}^4)$ which are not sums of squares in $H_3(\mathbb{R}^3)$ and $H_2(\mathbb{R}^4)$, respectively, which ultimately led to his 17th problem. (See [Reichstein 1987] for much more on this subject.)

A constructive discussion of Hilbert's theorem on $p \in H_4(\mathbb{R}^3)$ has recently been given in [Powers and Reznick 2000; Powers et al. 2004; Pfister and Scheiderer 2012; Plaumann et al. 2011]. A nonconstructive proof (without the count) can easily be given.

Theorem 5.2. *A general ternary quartic $p \in H_4(\mathbb{C}^3)$ can be written as $p = q_1^2 + q_2^2 + q_3^2$, where $q_j \in H_2(\mathbb{C}^3)$.*

Proof. We take q_i so that the monomial x^2 only appears in q_1 and the monomial y^2 only appears in q_1 and q_2 , and so the number of coefficients in the q_j is $6+5+4=15$. Taking the partials where $(q_1, q_2, q_3) = (x^2, y^2, z^2)$ shows that J contains $2x^2\{x^2, y^2, z^2, xy, xz, yz\}$, $2y^2\{y^2, z^2, xy, xz, yz\}$ and $2z^2\{z^2, xy, xz, yz\}$, and so is equal to $H_4(\mathbb{C}^3)$. \square

Since $3\binom{m+2}{2} - 3 < \binom{2m+1}{2}$ for $m \geq 3$, this result does not generalize to ternary forms of higher even degree.

The situation is somewhat simpler over \mathbb{R} . A real version of Theorem 5.1 appears in [Reznick 2000]. If p is real and positive definite and $p = f^2 + g^2$, where f and g are also real, then the factors of p consist of s conjugate pairs. In the factorization $p = (f + ig)(f - ig)$, the pairs must be split between the conjugate factors, and if p has distinct factors, this can be done in 2^{s-1} different ways. A real generalization of Theorem 5.2 appears in [Choi et al. 1995, Corollary 2.12]. Suppose a real psd form $p \in H_{2s}(\mathbb{R}^n)$ is a sum of t squares and $x^{\beta_i} \in H_s(\mathbb{R}^n)$, $1 \leq i \leq t$, is given. Then there is a representation $p = \sum_{j=1}^t g_j^2$, in which x^{β_i} does not occur in g_j for $j > i$. This argument can also be applied to a *general* sum of t squares over \mathbb{C} , but it no longer applies to all forms. For example, if $xy = (ax + by)^2 + (cx + dy)^2$, then $abcd \neq 0$.

6. Cubic forms and proofs of Theorems 1.3 and 1.4

In this section, we present three representations for forms in $H_3(\mathbb{C}^n)$ as a sum of cubes of linear forms. The first two are canonical; the third isn't, but it represents *all* cubics, not just general cubics.

We begin with Theorem 1.3, which first appeared in [Reichstein 1987]. At the time of this writing, that paper has had no citations in MathSciNet. (It was discussed in [Reznick 1992b] and, from there, in [Comon and Mourrain 1996]. The former was never submitted for publication and the latter appeared in an unindexed journal.) The original presentation and proof in [Reichstein 1987] were given for trilinear forms (see Section 2); the theorem is applied to cubic forms there mainly in the examples.

By iterating (1-5), we obtain a canonical form for $p \in H_3(\mathbb{C}^n)$; see [Reichstein 1987, p. 98].

Corollary 6.1. *A general n -ary cubic $p \in H_3(\mathbb{C}^n)$ can be written uniquely as*

$$(6-1) \quad p(x_1, \dots, x_n) = \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} \sum_{k=1}^{n-2m} (t_{m,1+2m}^{\{k\}} x_{1+2m} + \dots + t_{m,n}^{\{k\}} x_n)^3$$

for some $t_{m,j}^{\{k\}} \in \mathbb{C}$.

This gives p as a sum of $n + (n-2) + \dots = \lfloor (n+1)^2/4 \rfloor$ cubes. Recall that, by Alexander–Hirschowitz, for $n \neq 5$, a general cubic form in n variables can be written as a sum of $\lceil (n+1)(n+2)/6 \rceil$ cubes. Thus (6-1) is a canonical form which represents a general cubic as a sum of about 50% more cubes than the true minimum; this is due to the large number of linear forms with restricted sets of variables.

Reichstein’s proof of Theorem 1.3 requires the well-known “generalized eigenvalue problem” for pairs of symmetric matrices, as interpreted for quadratic forms: if a general pair of quadratic forms $f, g \in H_2(\mathbb{C}^n)$ is given, then there exist n linearly independent forms $L_i(x) = \sum_{j=1}^n \alpha_{ij} x_j$ and $c_i \in \mathbb{C}$ so that

$$(6-2) \quad f = \sum_{i=1}^n L_i^2, \quad g = \sum_{i=1}^n c_i L_i^2.$$

If M_f, M_g are the matrices associated to f, g , then the c_i are the n roots of the determinantal equation $\det(M_g - \lambda M_f) = 0$, which are generally distinct, so the L_i are uniquely determined up to multiple. We may also assume that the coefficients α_{ij} of the linear forms are generally nonzero; cf. Corollary 6.3.

Proof of Theorem 1.3. For general $p \in H_3(\mathbb{C}^n)$, we simultaneously diagonalize $f = \partial p / \partial x_1$ and $g = \partial p / \partial x_2$ as in (6-2). Since mixed partials are equal,

$$(6-3) \quad \frac{\partial f}{\partial x_2} = \frac{\partial g}{\partial x_1} = \sum_{i=1}^n 2\alpha_{i2} L_i = \sum_{i=1}^n 2c_i \alpha_{i1} L_i,$$

and since the L_i are linearly independent, (6-3) implies that $\alpha_{i2} = c_i \alpha_{i1}$.

It is generally true that $\alpha_{i1} \neq 0$. Let

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \sum_{i=1}^n \frac{1}{3\alpha_{i1}} L_i^3.$$

It follows that

$$\frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \sum_{i=1}^n \frac{3\alpha_{i1}}{3\alpha_{i1}} L_i^2 = \frac{\partial p}{\partial x_1} - \sum_{i=1}^n L_i^2 = 0,$$

$$\frac{\partial q}{\partial x_2} = \frac{\partial p}{\partial x_2} - \sum_{i=1}^n \frac{3\alpha_{i2}}{3\alpha_{i1}} L_i^2 = \frac{\partial p}{\partial x_2} - \sum_{i=1}^n c_i L_i^2 = 0.$$

Since $\partial q/\partial x_1 = \partial q/\partial x_2 = 0$, we have $q = q(x_3, \dots, x_n)$.

For uniqueness, suppose (1-5) holds and $l_k(x_1, \dots, x_n) = \sum_j \beta_{kj} x_j$. Then

$$\frac{\partial p}{\partial x_1} = \sum_{k=1}^n 3\beta_{k1} l_k^2, \quad \frac{\partial p}{\partial x_2} = \sum_{k=1}^n 3\beta_{k2} l_k^2.$$

Thus, after a scaling, $\partial p/\partial x_1$ and $\partial p/\partial x_2$ have already been simultaneously diagonalized (as in (6-2)), and the l_k are, up to multiples, a rearrangement of the L_k . \square

We now give a constructive proof of Theorem 1.4, which gives a different canonical form for $H_3(\mathbb{C}^n)$ requiring even more cubes.

Proof of Theorem 1.4. The constant-counting makes this a potential canonical form: the variables are $t_{\{i,j\},k}$ with $1 \leq i \leq j \leq k \leq n$, and there are $\binom{n+2}{3} = N(n, 3)$ such triples (i, j, k) . Given $p \in H_3(\mathbb{C}^n)$, $\partial p/\partial x_n$ is a quadratic form, so we can generally complete the square by Theorem 1.2:

$$\frac{\partial p}{\partial x_n} = \sum_{j=1}^n (t_{jj} x_j + \dots + t_{jn} x_n)^2.$$

Then $t_{jn} \neq 0$ for general p and if we let

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \sum_{j=1}^n \frac{1}{3t_{jn}} (t_{jj} x_j + \dots + t_{jn} x_n)^3,$$

then $\partial q/\partial x_n = 0$, so $q = q(x_1, \dots, x_{n-1})$. Iterate this construction to get (1-6).

Uniqueness follows by working backwards. If (1-6) holds for a cubic p , then it gives $\partial p/\partial x_n$ in its (unique) upper-triangular diagonalization. This can be integrated with respect to x_n and subtracted from p , giving a cubic $q(x_1, \dots, x_{n-1})$. Again, iterate. \square

It is not hard to give nonconstructive proofs of Theorems 1.3 and 1.4 using Corollary 2.3. These are left for the reader.

We first presented this next construction in [Reznick 1992b]; an outline of the proof can be found in [Comon and Mourrain 1996]. This is not a canonical form, but is included here because it gives an absolute upper bound for the length of cubic forms.

Theorem 6.2. *If $p \in H_3(\mathbb{C}^n)$, then there exists an invertible linear change of variables $y_j = \sum \lambda_{jk}x_k$ and n linear forms l_j so that, for some $q \in H_3(\mathbb{C}^{n-1})$,*

$$(6-4) \quad p(x_1, \dots, x_n) = \sum_{j=1}^n l_j^3(x_1, \dots, x_n) + q(y_2, \dots, y_n).$$

Thus every cubic in n variables is a sum of at most $\binom{n+1}{2}$ cubes of linear forms.

Proof. Define linear forms $l_{j,m}(y)$ for $1 \leq j \leq m+1$ by

$$(6-5) \quad l_{j,m}(y_1, \dots, y_n) = y_j + \alpha \sum_{j=1}^m y_j, \quad 1 \leq j \leq m$$

and

$$(6-6) \quad l_{m+1,m}(y_1, \dots, y_n) = -(1+m\alpha) \sum_{j=1}^m y_j, \quad \alpha = \frac{-(m+1) + \sqrt{m+1}}{m(m+1)}.$$

Then it can be easily checked that

$$(6-7) \quad \sum_{j=1}^{m+1} l_{j,m}(y) = 0 \quad \text{and} \quad \sum_{j=1}^{m+1} l_{j,m}^2(y) = \sum_{k=1}^m y_k^2.$$

Suppose $0 \neq p \in H_3(\mathbb{C}^n)$. Use Biermann's theorem to find a point u where $p(u) \neq 0$, and, after an invertible linear change of variables, taking $\{x_j\} \mapsto \{u_j\}$, we may assume that $p(1, 0, \dots, 0) = 1$ and so

$$(6-8) \quad p = u_1^3 + 3h_1(u_2, \dots, u_n)u_1^2 + 3h_2(u_2, \dots, u_n)u_1 + h_3(u_2, \dots, u_n),$$

where $\deg h_j = j$. Now let $u_1 = y_1 - h_1(u_2, \dots, u_n)$ to clear the quadratic term, so

$$(6-9) \quad p = y_1^3 + 3y_1\tilde{h}_2(u_2, \dots, u_n) + \tilde{h}_3(u_2, \dots, u_n),$$

where again $\deg \tilde{h}_j = j$. Diagonalize $\tilde{h}_2(u_2, \dots, u_n)$ as a quadratic form into $y_2^2 + \dots + y_r^2$, where $r \leq n$, and make the accompanying change of variables. We now have

$$(6-10) \quad p = y_1^3 + 3y_1(y_2^2 + \dots + y_r^2) + k_3(y_2, \dots, y_n), \quad r \leq n,$$

where $\deg k_3 = 3$. Finally, using (6-5) and (6-7), we construct g , a sum of $r \leq n$ cubes:

$$\begin{aligned}
 (6-11) \quad g(y_1, \dots, y_n) &:= \frac{1}{r} \sum_{j=1}^r (y_1 + \sqrt{r} \cdot l_{j,r-1}(y_2, \dots, y_r))^3 \\
 &= \frac{1}{r} \sum_{j=1}^r y_1^3 + \frac{3}{\sqrt{r}} \sum_{j=1}^r y_1^2 l_{j,r-1} + 3 \sum_{j=1}^r y_1 l_{j,r-1}^2 + \sqrt{r} \sum_{j=1}^r l_{j,r-1}^3 \\
 &= y_1^3 + 3y_1(y_2^2 + \dots + y_r^2) + \sqrt{r} \sum_{j=1}^r l_{j,r-1}^3(y_2, \dots, y_r).
 \end{aligned}$$

Then

$$q := p - g$$

is a cubic form in (y_2, \dots, y_n) as in (6-4). Iteration of this argument shows that any cubic $p \in H_3(\mathbb{C}^n)$ is a sum of at most $n(n+1)/2$ cubes. \square

Equation (1-5) can be extended to a canonical form for quartics as a sum of fourth powers of linear forms. Note that x_n appears in each summand of (6-1), with, generally, a nonzero coefficient.

Corollary 6.3. *For general $p \in H_4(\mathbb{C}^n)$, there exist $l_k \in H_1(\mathbb{C}^n)$ and $q \in H_4(\mathbb{C}^{n-1})$ so that, with $a(n) = \lfloor (n+1)^2/4 \rfloor$,*

$$p(x_1, \dots, x_n) = \sum_{k=1}^{a(n)} l_k(x_1, \dots, x_n)^4 + q(x_1, \dots, x_{n-1}).$$

As a consequence, a general $p \in H_4(\mathbb{C}^n)$ can be written as

$$p(x_1, \dots, x_n) = \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} \sum_{r=1+2m}^n \sum_{k=1}^{r-2m} (t_{m,r,1+2m}^{\{k\}} x_{1+2m} + \dots + t_{m,r,r}^{\{k\}} x_r)^4.$$

Proof. By Theorem 1.3 and (6-1), for general $p \in H_4(\mathbb{C}^n)$, we can write

$$\begin{aligned}
 (6-12) \quad \frac{\partial p}{\partial x_n} &= \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} \sum_{k=1}^{n-2m} (t_{m,1+2m}^{\{k\}} x_{1+2m} + \dots + t_{m,n}^{\{k\}} x_n)^3 \\
 &=: \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} \sum_{k=1}^{n-2m} (l_m^{\{k\}}(x))^3.
 \end{aligned}$$

As before, if $q = p - \sum_{k,m} 1/(4t_{m,n}^{\{k\}}) l_{k,m}^4$, then $\partial q / \partial x_n = 0$, so $q = q(x_1, \dots, x_{n-1})$. Repeat as before. There are $N(n, 3)$ coefficients in (6-12), and, since $N(n, 3) + N(n-1, 4) = N(n, 4)$, the count is correct for a canonical form. \square

Note that there is no variable which appears in each linear form in (6-12), so the argument can't be extended to quintics. For the same reason, Theorem 1.4 does not extend to quartics. By combining Theorems 1.3 and 6.3, we obtain canonical forms as a sum of powers of linear forms in the four exceptional cases of Alexander–Hirschowitz, of course at the expense of the number of summands. With regard to ternary quartics and Theorem 1.10, Corollary 6.3 becomes the following canonical form for $H_4(\mathbb{C}^3)$ as a sum of seven fourth powers:

$$\sum_{k=1}^3 (t_{k1}x_1 + t_{k2}x_2 + t_{k3}x_3)^4 + t_{10}x_3^4 + \sum_{l=1}^2 (u_{l1}x_1 + u_{l2}x_2)^4 + u_5x_1^4.$$

There is an arithmetic obstruction to a “Reichstein-type” canonical form for quartics, that is, one in which each linear form is allowed to involve each variable. If

$$(6-13) \quad p(x_1, \dots, x_n) = \sum_{k=1}^r (\alpha_{k1}x_1 + \dots + \alpha_{kn}x_n)^4 + q(x_1, \dots, x_m)$$

were a canonical form for some n , then we would have $N(n, 4) = rn + N(m, 4)$. However, for $n = 12$, there does not exist $m < 12$ so that $12 \mid \binom{15}{4} - \binom{m+3}{4}$, so no such canonical form can exist. More generally, let

$$(6-14) \quad A_d = \left\{ n : n \nmid \binom{n+d-1}{d} - \binom{m+d-1}{d} \text{ for } 0 \leq m < n \right\}$$

denote the set of n for which this argument rules out Reichstein-type canonical forms. We present without proof a number of results about A_d . Note that there is no obstacle for (6-13) in prime degree, such as $d = 2, 3$.

Proposition 6.4. (i) *If $3 \nmid k$, then $n = 2^{2k} \cdot 3 \in A_4$.*

(ii) *If $p \equiv 1 \pmod{144}$ is prime, then $12p \in A_4$.*

(iii) *If p is prime, then $p \mid \binom{n+p-1}{p} - \binom{n}{p}$; hence $A_p = \emptyset$ for prime p .*

(iv) *The smallest elements of $A_6, A_8, A_{10}, A_{12}, A_{14}$ and A_{15} are 10, 1792, 6, 242, 338 and 273, respectively. If A_9 or A_{16} are nonempty, then their smallest elements are at least 10^5 .*

7. Subspace canonical forms and the proof of Theorem 1.11

One natural generalization of the definition of canonical forms is to consider maps $F : X \mapsto H_d(\mathbb{C}^n)$ where $X \subset \mathbb{C}^M$ is an $N(n, d)$ -dimensional subspace of \mathbb{C}^M . (Similar ideas can be found in [Wakeford 1920], though his approach is different from ours.) These can be analyzed in the simplest nontrivial case: $M = 4$, $N(2, 2) = 3$.

Proof of Theorem 1.11. Assume that some $c_j \neq 0$. Without loss of generality, we may assume that $c_4 \neq 0$ and divide through by c_4 so that the equation is $t_4 = a_1t_1 + a_2t_2 + a_3t_3$, where $a_i = -c_i/c_4$ for $i = 1, 2, 3$. Then (1-13) can be parametrized as a map from $\mathbb{C}^3 \mapsto H_2(\mathbb{C}^2)$ as

$$(7-1) \quad F(t; x) = (t_1x + t_2y)^2 + (t_3x + (a_1t_1 + a_2t_2 + a_3t_3)y)^2.$$

The partials with respect to the t_j are

$$(7-2) \quad \begin{aligned} &2x(t_1x + t_2y) + 2a_1y(t_3x + (a_1t_1 + a_2t_2 + a_3t_3)y), \\ &2y(t_1x + t_2y) + 2a_2y(t_3x + (a_1t_1 + a_2t_2 + a_3t_3)y), \\ &2(x + a_3y)(t_3x + (a_1t_1 + a_2t_2 + a_3t_3)y). \end{aligned}$$

Now, (7-1) is a canonical form if and only if there exists a choice of t_i so that the three quadratics in (7-2) span $H_2(\mathbb{C}^2)$. A computation shows that the determinant of the forms in (7-2) with respect to the basis $\{x^2, xy, y^2\}$ is the cubic

$$(7-3) \quad -8((a_1a_2 - a_3)t_1 + (1 + a_2^2)t_2 + (a_2a_3 + a_1)t_3)(a_1t_1^2 + a_2t_1t_2 + a_3t_1t_3 - t_2t_3).$$

The second factor in (7-3) always has the term $-t_2t_3$ and so never vanishes; hence this determinant is not identically zero (and (7-1) is a canonical form), unless

$$(7-4) \quad a_1a_2 - a_3 = 1 + a_2^2 = a_2a_3 + a_1 = 0.$$

In the exceptional case where (7-4) holds, then $a_2 = \epsilon$, where $\epsilon = \pm i$, and $a_3 = \epsilon a_1$. Evaluating (7-1) at $(x, y) = (a_1, \epsilon)$ yields

$$\begin{aligned} &(a_1t_1 + \epsilon t_2)^2 + (a_1t_3 + \epsilon a_1t_1 + \epsilon^2 t_2 + \epsilon^2 a_1t_3)^2 \\ &= (a_1t_1 + \epsilon t_2)^2 + ((1 + \epsilon^2)a_1t_3 + \epsilon a_1t_1 + \epsilon^2 t_2)^2 = (a_1t_1 + \epsilon t_2)^2 + \epsilon^2 (a_1t_1 + \epsilon t_2)^2 = 0, \end{aligned}$$

as claimed. \square

It would be interesting to know how Theorem 1.11 generalizes to higher degrees.

Conjecture 1.12 is true for degree 2 by Theorem 1.11. We have verified it for even degrees up to eight by Corollary 2.3 applied to random choices for α_j, β_j in (1-14). We hold some hope that generalizations such as Conjecture 1.12 will have applications in more than two variables as well.

Acknowledgements

The greatest debt of the author is due Richard Ehrenborg and Gian-Carlo Rota for writing [Ehrenborg and Rota 1993]. Thanks to Dave Anderson and Julianna Tymoczko for organizing the Special Session on Geometric Commutative Algebra and Applications at the March 2011 AMS Sectional Meeting in Iowa City, to Lek-Heng Lim for organizing the Minisymposium on Tensor Rank at the October

2011 SIAM Conference on Applied Algebraic Geometry in Raleigh and to Eugene Mukhin for organizing ALGECOM5 in Indianapolis in October 2011. Invitations to speak at these conferences provided an opportunity to present preliminary versions of this material.

I am indebted to Tony Várilly-Alvarado for extremely helpful correspondence. I also want to thank T. Y. Lam, Hal Schenck, Bernd Sturmfels and Doron Zeilberger for their assistance. Special thanks go to the Center for Advanced Study at UIUC, where the author was an Associate in the Fall 2011 semester, and thereby free of teaching responsibilities. Finally, I want to thank the referee for a careful reading of the manuscript and for making many useful suggestions.

References

- [Alexander and Hirschowitz 1995] J. Alexander and A. Hirschowitz, “Polynomial interpolation in several variables”, *J. Algebraic Geom.* **4**:2 (1995), 201–222. MR 96f:14065 Zbl 0829.14002
- [Babbage 1976] D. W. Babbage, ““Canonical forms for general polynomials” (J. London Math. Soc. (2) **10**:3 (1975), 271–280)—a note on a paper by Stefano Guazzzone”, *J. London Math. Soc.* (2) **13**:3 (1976), 564. MR 54 #5264 Zbl 0338.13007
- [Carlini et al. 2012] E. Carlini, M. V. Catalisano, and A. V. Geramita, “The solution to the Waring problem for monomials and the sum of coprime monomials”, *J. Algebra* **370** (2012), 5–14. MR 2966824 Zbl 06162665
- [Choi et al. 1995] M. D. Choi, T. Y. Lam, and B. Reznick, “Sums of squares of real polynomials”, pp. 103–126 in *K-theory and algebraic geometry: connections with quadratic forms and division algebras, II* (Santa Barbara, CA, 1992), edited by B. Jacob and A. Rosenberg, Proc. Sympos. Pure Math. **58**, Amer. Math. Soc., Providence, RI, 1995. MR 96f:11058 Zbl 0821.11028
- [Clebsch 1861] A. Clebsch, “Ueber Curven vierter Ordnung”, *J. Reine Angew. Math.* **59** (1861), 125–145. Zbl 059.1559cj
- [Comon and Mourrain 1996] P. Comon and B. Mourrain, “Decomposition of quantics in sums of powers of linear forms”, *Signal Process.* **53**:2–3 (1996), 93–107. Zbl 0875.94079
- [Cox et al. 2007] D. Cox, J. Little, and D. O’Shea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, 3rd ed., Springer, New York, 2007. MR 2007h:13036 Zbl 1118.13001
- [Ehrenborg and Rota 1993] R. Ehrenborg and G.-C. Rota, “Apolarity and canonical forms for homogeneous polynomials”, *Eur. J. Combin.* **14**:3 (1993), 157–181. MR 94e:15062 Zbl 0784.05019
- [Elliott 1913] E. B. Elliott, *An introduction to algebra of quantics*, 2nd ed., Clarendon, Oxford, 1913. Reprinted by Chelsea, New York, 1964. JFM 44.0155.05
- [Fisher 1967] C. S. Fisher, “The last invariant theorists: a sociological study of the collective biographies of mathematical specialists”, *Eur. J. Sociol.* **8**:2 (1967), 216–244.
- [Fröberg et al. 2012] R. Fröberg, G. Ottaviani, and B. Shapiro, “On the Waring problem for polynomial rings”, *Proc. Natl. Acad. Sci. USA* **109**:15 (2012), 5600–5602. MR 2935563 Zbl 06133962
- [Geramita 1996] A. V. Geramita, “Inverse systems of fat points: Waring’s problem, secant varieties of Veronese varieties and parameter spaces for Gorenstein ideals”, pp. 2–114 in *Curves Seminar at Queen’s, 10* (Kingston, ON, 1995), edited by A. V. Geramita, Queen’s Papers in Pure and Appl. Math. **102**, Queen’s Univ., Kingston, ON, 1996. MR 97h:13012 Zbl 0864.14031

- [Grace 1918] J. H. Grace, “Nachruf auf Edward Kingsley Wakeford (1894–1916)”, *Proc. Lond. Math. Soc.* (2) **16** (1918), 54–57. JFM 46.0035.22
- [Guazzone 1975] S. Guazzone, “Canonical forms for general polynomials”, *J. London Math. Soc.* (2) **10** (1975), 271–280. MR 52 #3183 Zbl 0307.14014
- [Gundelfinger 1887] S. Gundelfinger, “Zur Theorie der binären Formen”, *J. Reine Angew. Math.* **100** (1887), 413–424. JFM 19.0109.01
- [Hilbert 1888] D. Hilbert, “Ueber die Darstellung definiter Formen als Summe von Formenquadraten”, *Math. Ann.* **32**:3 (1888), 342–350. MR 1510517 JFM 20.0198.02
- [Iarrobino and Kanev 1999] A. Iarrobino and V. Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Math. **1721**, Springer, Berlin, 1999. MR 2001d:14056 Zbl 0942.14026
- [Kung 1986] J. P. S. Kung, “Gundelfinger’s theorem on binary forms”, *Stud. Appl. Math.* **75**:2 (1986), 163–169. MR 87m:11020 Zbl 0607.10016
- [Kung 1987] J. P. S. Kung, “Canonical forms for binary forms of even degree”, pp. 52–61 in *Invariant theory* (West Chester, PA, 1985), edited by S. S. Koh, Lecture Notes in Math. **1278**, Springer, Berlin, 1987. MR 89h:15037 Zbl 0647.15019
- [Kung 1990] J. P. S. Kung, “Canonical forms of binary forms: variations on a theme of Sylvester”, pp. 46–58 in *Invariant theory and tableaux* (Minneapolis, MN, 1988), edited by D. Stanton, IMA Vol. Math. Appl. **19**, Springer, New York, 1990. MR 91b:11046 Zbl 0708.15013
- [Kung and Rota 1984] J. P. S. Kung and G.-C. Rota, “The invariant theory of binary forms”, *Bull. Amer. Math. Soc. (N.S.)* **10**:1 (1984), 27–85. MR 85g:05002 Zbl 0577.15020
- [Landsberg 2012] J. M. Landsberg, *Tensors: geometry and applications*, Grad. Stud. Math. **128**, Amer. Math. Soc., Providence, RI, 2012. MR 2865915 Zbl 1238.15013
- [Landsberg and Teitler 2010] J. M. Landsberg and Z. Teitler, “On the ranks and border ranks of symmetric tensors”, *Found. Comput. Math.* **10**:3 (2010), 339–366. MR 2011d:14095 Zbl 1196.15024
- [Lasker 1904] E. Lasker, “Zur Theorie der kanonischen Formen”, *Math. Ann.* **58**:3 (1904), 434–440. MR 1511244 JFM 35.0122.01
- [OEIS 2013] OEIS Foundation, “Online Encyclopedia of Integer Sequences”, 2013, Available at <http://oeis.org>.
- [Pfister and Scheiderer 2012] A. Pfister and C. Scheiderer, “An elementary proof of Hilbert’s theorem on ternary quartics”, *J. Algebra* **371** (2012), 1–25. MR 2975385 Zbl 06166647
- [Plaumann et al. 2011] D. Plaumann, B. Sturmfels, and C. Vinzant, “Quartic curves and their bitangents”, *J. Symbolic Comput.* **46**:6 (2011), 712–733. MR 2012e:14065 Zbl 1214.14049
- [Powers and Reznick 2000] V. Powers and B. Reznick, “Notes towards a constructive proof of Hilbert’s theorem on ternary quartics”, pp. 209–227 in *Quadratic forms and their applications* (Dublin, 1999), edited by E. Bayer-Fluckiger et al., Contemp. Math. **272**, Amer. Math. Soc., Providence, RI, 2000. MR 2001h:11049 Zbl 1026.11044
- [Powers et al. 2004] V. Powers, B. Reznick, C. Scheiderer, and F. Sottile, “A new approach to Hilbert’s theorem on ternary quartics”, *C. R. Math. Acad. Sci. Paris* **339**:9 (2004), 617–620. MR 2005i:11051 Zbl 1061.11016
- [Ranestad and Schreyer 2000] K. Ranestad and F.-O. Schreyer, “Varieties of sums of powers”, *J. Reine Angew. Math.* **525** (2000), 147–181. MR 2001m:14009 Zbl 1078.14506
- [Reichstein 1987] B. Reichstein, “On expressing a cubic form as a sum of cubes of linear forms”, *Linear Algebra Appl.* **86** (1987), 91–122. MR 88e:11022 Zbl 0609.15015

- [Reznick 1992a] B. Reznick, “Sums of even powers of real linear forms”, *Mem. Amer. Math. Soc.* **96**:463 (1992). MR 93h:11043 Zbl 0762.11019
- [Reznick 1992b] B. Reznick, “Sums of powers of complex linear forms”, preprint, 1992.
- [Reznick 1996] B. Reznick, “Homogeneous polynomial solutions to constant coefficient PDE’s”, *Adv. Math.* **117**:2 (1996), 179–192. MR 97a:12006 Zbl 0858.12004
- [Reznick 2000] B. Reznick, “Some concrete aspects of Hilbert’s 17th problem”, pp. 251–272 in *Real algebraic geometry and ordered structures* (Baton Rouge, LA, 1996), edited by C. N. Delzell and J. J. Madden, Contemp. Math. **253**, Amer. Math. Soc., Providence, RI, 2000. MR 2001i:11042 Zbl 0972.11021
- [Reznick 2013] B. Reznick, “On the length of binary forms”, in *Quadratic and higher degree forms*, edited by K. Alladi et al., Dev. Math. **31**, Springer, New York, 2013. arXiv 1007.5485
- [Reznick \geq 2013] B. Reznick, “Linear dependence among powers of quadratic forms”. In preparation.
- [Richmond 1902] H. W. Richmond, “On canonical forms”, *Quart. J. Pure Appl. Math.* **33** (1902), 331–340. JFM 33.0127.02
- [Sylvester 1851a] J. J. Sylvester, *An essay on canonical forms: supplement to a sketch of a memoir on elimination, transformation and canonical forms*, George Bell, London, 1851. Reprinted as Paper 34, pp. 203–216 in *The collected mathematical papers of James Joseph Sylvester, I: 1837–1853*, Cambridge University Press, 2012.
- [Sylvester 1851b] J. J. Sylvester, “On a remarkable discovery in the theory of canonical forms and of hyperdeterminants”, *Philos. Mag.* (4) **2**:12 (1851), 391–410. Reprinted as Paper 41, pp. 265–283 in *The collected mathematical papers of James Joseph Sylvester, I: 1837–1853*, Cambridge University Press, 2012.
- [Sylvester 1870] J. J. Sylvester, “Additional notes to Prof. Sylvester’s Exeter British Association address”, pp. 140–143 in *The laws of verse*, Longmans, London, 1870. Reprinted in Appendix, pp. 714–719 in *The collected mathematical papers of James Joseph Sylvester, II: 1857–1873*, Cambridge University Press, 2012.
- [Turnbull 1960] H. W. Turnbull, *The theory of determinants, matrices, and invariants*, 3rd ed., Dover, New York, 1960. MR 24 #A123 Zbl 0103.00702
- [Várilly-Alvarado 2008] A. Várilly-Alvarado, “Weak approximation on del Pezzo surfaces of degree 1”, *Adv. Math.* **219**:6 (2008), 2123–2145. MR 2009j:14045 Zbl 1156.14017
- [Várilly-Alvarado 2011] A. Várilly-Alvarado, “Density of rational points on isotrivial rational elliptic surfaces”, *Algebra Number Theory* **5**:5 (2011), 659–690. MR 2889751 Zbl 06054517
- [Wakeford 1913] E. K. Wakeford, “A canonical form of the binary sextic”, *Mess. Math.* **4** (1913), 25–28. JFM 44.0142.01
- [Wakeford 1920] E. K. Wakeford, “On canonical forms”, *Proc. London Math. Soc.* (2) **18**:1 (1920), 403–410. MR 1576066 JFM 47.0880.01

Received March 28, 2012. Revised June 17, 2013.

BRUCE REZNICK
DEPARTMENT OF MATHEMATICS AND CENTER FOR ADVANCED STUDY
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
1409 W. GREEN STREET
327 ALTGELD HALL
URBANA, IL 61801-2975
UNITED STATES
reznick@math.uiuc.edu

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

EDITORS

V. S. Varadarajan (Managing Editor)
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
pacific@math.ucla.edu

Paul Balmer
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
balmer@math.ucla.edu

Don Blasius
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
blasius@math.ucla.edu

Vyjayanthi Chari
Department of Mathematics
University of California
Riverside, CA 92521-0135
chari@math.ucr.edu

Daryl Cooper
Department of Mathematics
University of California
Santa Barbara, CA 93106-3080
cooper@math.ucsb.edu

Robert Finn
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
finn@math.stanford.edu

Kefeng Liu
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
liu@math.ucla.edu

Jiang-Hua Lu
Department of Mathematics
The University of Hong Kong
Pokfulam Rd., Hong Kong
jhlu@maths.hku.hk

Sorin Popa
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
popa@math.ucla.edu

Jie Qing
Department of Mathematics
University of California
Santa Cruz, CA 95064
qing@cats.ucsc.edu

Paul Yang
Department of Mathematics
Princeton University
Princeton NJ 08544-1000
yang@math.princeton.edu

PRODUCTION

Silvio Levy, Scientific Editor, production@msp.org

SUPPORTING INSTITUTIONS

ACADEMIA SINICA, TAIPEI
CALIFORNIA INST. OF TECHNOLOGY
INST. DE MATEMÁTICA PURA E APLICADA
KEIO UNIVERSITY
MATH. SCIENCES RESEARCH INSTITUTE
NEW MEXICO STATE UNIV.
OREGON STATE UNIV.

STANFORD UNIVERSITY
UNIV. OF BRITISH COLUMBIA
UNIV. OF CALIFORNIA, BERKELEY
UNIV. OF CALIFORNIA, DAVIS
UNIV. OF CALIFORNIA, LOS ANGELES
UNIV. OF CALIFORNIA, RIVERSIDE
UNIV. OF CALIFORNIA, SAN DIEGO
UNIV. OF CALIF., SANTA BARBARA

UNIV. OF CALIF., SANTA CRUZ
UNIV. OF MONTANA
UNIV. OF OREGON
UNIV. OF SOUTHERN CALIFORNIA
UNIV. OF UTAH
UNIV. OF WASHINGTON
WASHINGTON STATE UNIVERSITY

These supporting institutions contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its contents or policies.

See inside back cover or msp.org/pjm for submission instructions.

The subscription price for 2013 is US \$400/year for the electronic version, and \$485/year for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2013 Mathematical Sciences Publishers

PACIFIC JOURNAL OF MATHEMATICS

Volume 266 No. 1 November 2013

Multiplicity of solutions to the Yamabe problem on collapsing Riemannian submersions	1
RENATO G. BETTIOL and PAOLO PICCIONE	
Rank gradient of small covers	23
DARLAN GIRÃO	
Nonrationality of nodal quartic threefolds	31
KYUSIK HONG	
Supertropical linear algebra	43
ZUR IZHAKIAN, MANFRED KNEBUSCH and LOUIS ROWEN	
Isometry groups among topological groups	77
PIOTR NIEMIEC	
Singularities and Liouville theorems for some special conformal Hessian equations	117
QIANZHONG OU	
Attaching handles to Delaunay nodoids	129
FRANK PACARD and HAROLD ROSENBERG	
Some new canonical forms for polynomials	185
BRUCE REZNICK	
Applications of the deformation formula of holomorphic one-forms	221
QUANTING ZHAO and SHENG RAO	



0030-8730(201311)266:1;1-5