

# ON THE LENGTH OF BINARY FORMS

BRUCE REZNICK

ABSTRACT. The  $K$ -length of a form  $f$  in  $K[x_1, \dots, x_n]$ ,  $K \subset \mathbb{C}$ , is the smallest number of  $d$ -th powers of linear forms of which  $f$  is a  $K$ -linear combination. We present many results, old and new, about  $K$ -length, mainly for  $n = 2$ , and often about the length of the same form over different fields. For example, the  $K$ -length of  $3x^5 - 20x^3y^2 + 10xy^4$  is three for  $K = \mathbb{Q}(\sqrt{-1})$ , four for  $K = \mathbb{Q}(\sqrt{-2})$  and five for  $K = \mathbb{R}$ .

## 1. INTRODUCTION AND OVERVIEW

Suppose  $f(x_1, \dots, x_n)$  is a form of degree  $d$  with coefficients in a field  $K \subseteq \mathbb{C}$ . The  $K$ -length of  $f$ ,  $L_K(f)$ , is the smallest  $r$  for which there exist  $\lambda_j, \alpha_{jk} \in K$  so that

$$(1.1) \quad f(x_1, \dots, x_n) = \sum_{j=1}^r \lambda_j (\alpha_{j1}x_1 + \dots + \alpha_{jn}x_n)^d.$$

In this paper, we consider the  $K$ -length of a fixed form  $f$  as  $K$  varies; this is apparently an open question in the literature, even for binary forms ( $n = 2$ ). Sylvester [53, 54] explained how to compute  $L_{\mathbb{C}}(f)$  for binary forms in 1851 and gave a lower bound for  $L_{\mathbb{R}}(f)$  for binary forms in 1864. Except for a few remarks, we shall restrict our attention to binary forms.

It is trivially true that  $L_K(f) = 1$  for linear  $f$  and for  $d = 2$ ,  $L_K(f)$  equals the rank of  $f$ : a representation over  $K$  can be found by completing the square, and this length cannot be shortened by enlarging the field. Accordingly, we shall also assume that  $d \geq 3$ .

When  $K = \mathbb{C}$ , the  $\lambda_j$ 's in (1.1) are superfluous. The computation of  $L_{\mathbb{C}}(f)$  is a huge, venerable and active subject, and difficult when  $n \geq 3$ . The interested reader is directed to [2, 7, 8, 14, 17, 18, 22, 25, 28, 34, 44, 45, 46] as representative recent works. Even for small  $n, d \geq 3$ , there are still many open questions. Landsberg and Teitler [34] complete a classification of  $L_{\mathbb{C}}(f)$  for ternary cubics  $f$  and also discuss  $L_{\mathbb{C}}(x_1x_2 \cdots x_n)$ , among other topics. Historically, much attention has centered on the  $\mathbb{C}$ -length of a *general* form of degree  $d$ . In 1995, Alexander and Hirschowitz [1] (see also [5, 36]) established that for  $n, d \geq 3$ , this length is  $\lceil \frac{1}{n} \binom{n+d-1}{n-1} \rceil$ , the constant-counting value, with the four exceptions known since the 19th century –

---

*Date:* July 29, 2012.

*2000 Mathematics Subject Classification.* Primary: 11E76, 11P05, 14N10.

*Key words and phrases.* length, sums of powers, binary forms, Sylvester.

$(n, d) = (3, 5), (4, 3), (4, 4), (4, 5)$  – in which the length is  $\lceil \frac{1}{n} \binom{n+d-1}{n-1} \rceil + 1$ . There has been a recent series of papers studying  $L_{\mathbb{R}}(f)$  [4, 9, 15]; these study the length in a greater depth than we do here.

Two central examples illustrate the phenomenon of multiple lengths over different fields.

*Example 1.1.* Suppose  $f(x, y) = (x + \sqrt{2}y)^d + (x - \sqrt{2}y)^d \in \mathbb{Q}[x, y]$ . Then  $L_K(f)$  is 2 (if  $\sqrt{2} \in K$ ) and  $d$  (otherwise). This example first appeared in [47, p.137]. (See Theorem 4.6 for a generalization.)

*Example 1.2.* If  $\phi(x, y) = 3x^5 - 20x^3y^2 + 10xy^4$ , then  $L_K(\phi) = 3$  if and only if  $\sqrt{-1} \in K$ ,  $L_K(\phi) = 4$  for  $K = \mathbb{Q}(\sqrt{-2}), \mathbb{Q}(\sqrt{-3}), \mathbb{Q}(\sqrt{-5}), \mathbb{Q}(\sqrt{-6})$  (at least) and  $L_{\mathbb{R}}(\phi) = 5$ . (We give proofs of these assertions in Examples 2.1 and 3.1.)

The following simple definitions and remarks apply in the obvious way to forms in  $n \geq 3$  variables, but for simplicity are given for binary forms. A representation such as (1.1) is called *K-minimal* if  $r = L_K(f)$ . Two linear forms are called *distinct* if they (or their  $d$ -th powers) are not proportional. A representation is *honest* if the summands are pairwise distinct. Any minimal representation is honest. Two honest representations are *different* if the ordered sets of summands are not rearrangements of each other; we shall not distinguish between  $\ell^d$  and  $(\zeta_d^k \ell)^d$  where  $\zeta_d = e^{2\pi i/d}$ . If  $g$  is obtained from  $f$  by an invertible linear change of variables over  $K$ , then  $L_K(f) = L_K(g)$ .

Given a form  $f \in \mathbb{C}[x, y]$ , we let  $E_f$  denote the field generated by the coefficients of  $f$  over  $\mathbb{C}$ ;  $L_K(f)$  is defined for fields  $K$  satisfying  $E_f \subseteq K \subseteq \mathbb{C}$ . The following implication is immediate:

$$(1.2) \quad K_1 \subset K_2 \implies L_{K_1}(f) \geq L_{K_2}(f).$$

Strict inequality in (1.2) is possible, as shown by the two examples. Finally, we define the *cabinet* of  $f$ ,  $\mathcal{C}(f)$ , to be the set of all possible lengths for  $f$ .

There is a natural alternative definition of length in which sums of powers are considered without coefficients. This makes no difference when  $K = \mathbb{C}$  or  $K = \mathbb{R}$  and  $d$  is odd, but in other cases, a form might not even be a sum of powers. For example,  $\sqrt{2}$  is not totally positive in  $K = \mathbb{Q}(\sqrt{2})$ , so  $\sqrt{2} \cdot x^2$  is not a sum of squares in  $K[x]$ , and  $x^4 + \lambda x^2 y^2 + y^4$  is a sum of fourth powers of real linear forms if and only if  $0 \leq \lambda \leq 6$ . This alternative definition was studied by Ellison [19] in the special cases  $K = \mathbb{C}, \mathbb{R}, \mathbb{Q}$ . Newman and Slater [39] do not restrict to homogeneous polynomials. They write  $x$  as a sum of  $d$   $d$ -th powers of linear polynomials; by substitution, any polynomial is a sum of at most  $d$   $d$ -th powers of polynomials. They also show that the minimum number of  $d$ -th powers in this formulation is  $\geq \sqrt{d}$ . Because of the degrees of the summands, these methods do not homogenize to forms. Mordell [37] showed that a polynomial that is a sum of cubes of linear forms over  $\mathbb{Z}$  is also a sum of at most eight such cubes. More generally, if  $R$  is a commutative ring, then its *d-Pythagoras number*,  $P_d(R)$ , is the smallest integer  $k$  so that any sum of  $d$ -th powers

in  $R$  is a sum of  $k$   $d$ -th powers. Helmke [25] uses both definitions for length for forms, but is mainly concerned with the alternative definition in the case when  $K$  is an algebraically closed (or a real closed) field of characteristic zero, not necessarily a subset of  $\mathbb{C}$ . This subject is closely related to Hilbert's 17th Problem; see [10, 12, 11]. In [47], a principal object of study is  $Q_{n,2k}$ , the (closed convex) cone of real forms which are a sum of  $2k$ -th powers of real linear forms.

We now outline the remainder of the paper.

In section two, we give a self-contained proof of Sylvester's 1851 Theorem (Theorem 2.1). Although originally given over  $\mathbb{C}$ , it adapts easily to any  $K \subset \mathbb{C}$  (Corollary 2.2). If  $f$  is a binary form, then  $L_K(f) \leq r$  if and only if a certain subspace of the binary forms of degree  $r$  (a subspace determined by  $f$ ) contains a form that splits into distinct factors over  $K$ . We illustrate the algorithm by proving the assertions of lengths 3 and 4 for  $\phi$  in Example 1.2.

In section three, we prove (Theorem 3.2) a homogenized version of Sylvester's 1864 Theorem (Theorem 3.1), which implies that if real  $f$  has  $r$  linear factors over  $\mathbb{R}$  (counting multiplicity), then  $L_{\mathbb{R}}(f) \geq r$ . As far as we can tell, Sylvester did not connect his two theorems: perhaps because he presented the second one for non-homogeneous polynomials in one variable.

We apply these theorems and some other simple observations in sections four and five. We first show that if  $L_{\mathbb{C}}(f) = 1$ , then  $L_{E_f}(f) = 1$  as well (Theorem 4.1). Any set of  $d+1$   $d$ -th powers of pairwise distinct linear forms is linearly independent (Theorem 4.2). It follows quickly that if  $f(x, y)$  has two different honest representations of length  $r$  and  $s$ , then  $r + s \geq d + 2$  (Corollary 4.3), and so if  $L_{E_f}(f) = r \leq \frac{d+1}{2}$ , then the representation over  $E_f$  is the unique minimal  $\mathbb{C}$ -representation (Corollary 4.4). We show that Example 1.1 gives a template for forms  $f$  satisfying  $L_{\mathbb{C}}(f) = 2 < L_{E_f}(f)$  (see Theorem 4.6), and give two generalizations which provide other types of constructions (Corollaries 4.7 and 4.8) of forms with multiple lengths. We apply Sylvester's 1851 Theorem to give an easy proof of the known result that  $L_{\mathbb{C}}(f) \leq d$  (Theorem 4.9) and a slightly trickier proof of the probably-known result that  $L_K(f) \leq d$  as well (Theorem 4.10). Theorem 4.10 combines with Theorem 3.2 into Corollary 4.11: if  $f \in \mathbb{R}[x, y]$  is a product of  $d$  linear factors and not a  $d$ -th power, then  $L_{\mathbb{R}}(f) = d$ . Conjecture 4.12 asserts that  $f \in \mathbb{R}[x, y]$  is a product of  $d$  linear factors if and only if  $L_{\mathbb{R}}(f) = d$ . This conjecture has recently been proven by Comon-Ottaviani-Causa-Re [15, 9] when the factors of  $f$  are distinct.

In Corollary 5.1, we discuss the various possible cabinets when  $d = 3, 4$ ; and give examples for each one not already ruled out. We then completely classify binary cubics; the key point of Theorem 5.2 is that if the cubic  $f$  has no repeated factors, then  $L_k(f) = 2$  if and only if  $E_f(\sqrt{-3\Delta(f)}) \subseteq K$ ; this significance of the discriminant  $\Delta(f)$  can already be found in Salmon [52, §167]. This proves Conjecture 4.12 for  $d = 3$ . In Theorem 5.3, we show that Conjecture 4.12 also holds for  $d = 4$ . We then show (Theorem 5.4) that  $L_{\mathbb{C}}(f) = d$  if and only if there are distinct linear forms  $\ell, \ell'$  so that  $f = \ell^{d-1}\ell'$ . (One direction of this result is well-known; the other has recently been

proved by Białyński-Birula and Schinzel [2].) The minimal representations of  $x^k y^k$  are parameterized (Theorem 5.5), and in Corollary 5.6, we show that  $L_K((x^2 + y^2)^k) \geq k + 1$ , with equality if and only if  $\tan \frac{\pi}{k+1} \in K$ . In particular,  $L_{\mathbb{Q}}((x^2 + y^2)^2) = 4$ . Theorem 5.7 shows that  $L_{\mathbb{Q}}(x^4 + 6\lambda x^2 y^2 + y^4) = 3$  if and only if a certain quartic diophantine equation over  $\mathbb{Z}$  has a non-zero solution.

Section six lists some open questions.

We would like to express our appreciation to the organizers of the *Higher Degree Forms* conference in Gainesville in May 2009 for offering the opportunities to speak on these topics, and to write this article for its Proceedings. We also thank Mike Bennett, Tony Geramita, Giorgio Ottaviani, Joe Rotman and Zach Teitler for helpful conversations and correspondence.

## 2. SYLVESTER'S 1851 THEOREM

Modern proofs of Theorem 2.1 can be found in the work of Kung and Rota: [33, §5], with further discussion in [30, 31, 32, 49]. We present here a very elementary proof showing the connection with constant coefficient linear recurrences, in the hopes that this remarkable theorem might become better known to the modern reader.

**Theorem 2.1** (Sylvester). *Suppose*

$$(2.1) \quad f(x, y) = \sum_{j=0}^d \binom{d}{j} a_j x^{d-j} y^j$$

and suppose

$$(2.2) \quad h(x, y) = \sum_{t=0}^r c_t x^{r-t} y^t = \prod_{j=1}^r (-\beta_j x + \alpha_j y)$$

is a product of pairwise distinct linear factors. Then there exist  $\lambda_k \in \mathbb{C}$  so that

$$(2.3) \quad f(x, y) = \sum_{k=1}^r \lambda_k (\alpha_k x + \beta_k y)^d$$

if and only if

$$(2.4) \quad \begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r+1} & \cdots & a_d \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix};$$

that is, if and only if

$$(2.5) \quad \sum_{t=0}^r a_{\ell+t} c_t = 0, \quad \ell = 0, 1, \dots, d-r.$$

*Proof.* First suppose that (2.3) holds. Then for  $0 \leq j \leq d$ ,

$$\begin{aligned} a_j &= \sum_{k=1}^r \lambda_k \alpha_k^{d-j} \beta_k^j \implies \sum_{t=0}^r a_{\ell+t} c_t = \sum_{k=1}^r \sum_{t=0}^r \lambda_k \alpha_k^{d-\ell-t} \beta_k^{\ell+t} c_t \\ &= \sum_{k=1}^r \lambda_k \alpha_k^{d-\ell-r} \beta_k^\ell \sum_{t=0}^r \alpha_k^{r-t} \beta_k^t c_t = \sum_{k=1}^r \lambda_k \alpha_k^{d-\ell-r} \beta_k^\ell h(\alpha_k, \beta_k) = 0. \end{aligned}$$

Now suppose that (2.4) holds and suppose first that  $c_r \neq 0$ . We may assume without loss of generality that  $c_r = 1$  and that  $\alpha_j = 1$  in (2.2), so that the  $\beta_j$ 's are distinct. Define the *infinite* sequence  $(\tilde{a}_j)$ ,  $j \geq 0$ , by:

$$(2.6) \quad \tilde{a}_j = a_j \quad \text{if } 0 \leq j \leq r-1; \quad \tilde{a}_{r+\ell} = - \sum_{t=0}^{r-1} \tilde{a}_{t+\ell} c_t \quad \text{for } \ell \geq 0.$$

This sequence satisfies the recurrence of (2.5), so that

$$(2.7) \quad \tilde{a}_j = a_j \quad \text{for } j \leq d.$$

Since  $|\tilde{a}_j| \leq \gamma \cdot M^j$  for suitable  $\gamma, M$  by induction, the generating function

$$\Phi(T) = \sum_{j=0}^{\infty} \tilde{a}_j T^j$$

converges in a neighborhood of 0. We have

$$\left( \sum_{t=0}^r c_{r-t} T^t \right) \Phi(T) = \sum_{n=0}^{r-1} \left( \sum_{j=0}^n c_{r-(n-j)} \tilde{a}_j \right) T^n + \sum_{n=r}^{\infty} \left( \sum_{t=0}^r c_{r-t} \tilde{a}_{n-t} \right) T^n.$$

It follows from (2.6) that the second sum vanishes, and hence  $\Phi(T)$  is a rational function with denominator

$$\sum_{t=0}^r c_{r-t} T^t = h(T, 1) = \prod_{j=1}^r (1 - \beta_j T).$$

By the method of partial fractions, there exist  $\lambda_k \in \mathbb{C}$  so that

$$(2.8) \quad \sum_{j=0}^{\infty} \tilde{a}_j T^j = \Phi(T) = \sum_{k=1}^r \frac{\lambda_k}{1 - \beta_k T} \implies \tilde{a}_j = \sum_{k=1}^r \lambda_k \beta_k^j.$$

A comparison of (2.8) and (2.7) with (2.1) shows that

$$(2.9) \quad f(x, y) = \sum_{j=0}^d \binom{d}{j} a_j x^{d-j} y^j = \sum_{k=1}^r \lambda_k \sum_{j=0}^d \binom{d}{j} \beta_k^j x^{d-j} y^j = \sum_{k=1}^r \lambda_k (x + \beta_k y)^d,$$

as claimed in (2.3).

If  $c_r = 0$ , then  $c_{r-1} \neq 0$ , because  $h$  has distinct factors. We may proceed as before, replacing  $r$  by  $r - 1$  and taking  $c_{r-1} = 1$ , so that (2.2) becomes

$$(2.10) \quad h(x, y) = \sum_{t=0}^{r-1} c_t x^{r-t} y^t = x \prod_{j=1}^{r-1} (y - \beta_j x).$$

Since  $c_r = 0$ , the system (2.4) can be rewritten as

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{r-1} \\ a_1 & a_2 & \cdots & a_r \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r+1} & \cdots & a_{d-1} \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{r-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We may now argue as before, except that (2.7) becomes

$$(2.11) \quad \tilde{a}_j = a_j \quad \text{for } j \leq d-1, \quad a_d = \tilde{a}_d + \lambda_r$$

for some  $\lambda_r$ , and (2.9) becomes

$$(2.12) \quad \begin{aligned} f(x, y) &= \sum_{j=0}^d \binom{d}{j} a_j x^{d-j} y^j = \lambda_r y^d + \sum_{k=1}^{r-1} \lambda_k \sum_{j=0}^d \binom{d}{j} \beta_k^j x^{d-j} y^j \\ &= \lambda_r y^d + \sum_{k=1}^{r-1} \lambda_k (x + \beta_k y)^d. \end{aligned}$$

By (2.10), (2.12) meets the description of (2.3), completing the proof.  $\square$

The  $(d-r+1) \times (r+1)$  Hankel matrix in (2.4) will be denoted  $H_r(f)$ . If  $(f, h)$  satisfy the criterion of this theorem, we shall say that  $h$  is a *Sylvester form for  $f$* . If the only Sylvester forms of degree  $r$  are  $\lambda h$  for  $\lambda \in \mathbb{C}$ , we say that  $h$  is the *unique* Sylvester form for  $f$ . Any polynomial multiple of a Sylvester form that has no repeated factors is also a Sylvester form, since there is no requirement that  $\lambda_k \neq 0$  in (2.3). If  $f$  has a unique Sylvester form of degree  $r$ , then  $L_{\mathbb{C}}(f) = r$  and  $L_K(f) \geq r$ .

The proof of Theorem 2.1 in [49] is based on apolarity. If  $f$  and  $h$  are given by (2.1) and (2.2), and  $h(D) = \prod_{j=1}^r (-\beta_j \frac{\partial}{\partial x} + \alpha_j \frac{\partial}{\partial y})$ , then

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

Thus, (2.4) is equivalent to  $h(D)f = 0$ . One can then argue that each linear factor in  $h(D)$  kills a different summand, and dimension counting takes care of the rest. In particular, if  $\deg h > d$ , then  $h(D)f = 0$  automatically, and this implies that  $L_{\mathbb{C}}(f) \leq d+1$ . Theorem 4.2 gives another explanation of this fact.

If  $h$  has repeated factors, a condition of interest in [30, 31, 32, 33, 49], then Gundelfinger's Theorem [23], first proved in 1886, shows that a factor  $(-\beta x + \alpha y)^\ell$  of  $h$  corresponds to a summand  $q(x, y)(\alpha x + \beta y)^{d+1-\ell}$  in  $f$ , where  $q$  is an arbitrary form

of degree  $\ell - 1$ . (We are not interested in such summands when  $\ell > 1$ . For more discussion of this case, see [49].)

If  $d = 2s - 1$  and  $r = s$ , then  $H_s(f)$  is  $s \times (s + 1)$  and has a non-trivial null-vector; for a general  $f$ , the resulting form  $h$  has distinct factors, and so is a unique Sylvester form. (The coefficients of  $h$ , and its discriminant, are polynomials in the coefficients of  $f$ .) This is how Sylvester proved that a general binary form of degree  $2s - 1$  is a sum of  $s$  powers of linear forms over  $\mathbb{C}$ , and the minimal representation is unique.

If  $d = 2s$  and  $r = s$ , then  $H_s(f)$  is square;  $\det(H_s(f))$  is the *catalecticant* of  $f$ . (For more on the term ‘‘catalecticant’’, see [47, pp.49-50] and [22, pp.104-105].) In general, there exists  $\lambda$  so that the catalecticant of  $f(x, y) - \lambda x^{2s}$  vanishes, and the resulting non-trivial null vector is generally a Sylvester form (no repeated factors). Thus, a general binary form of degree  $2s$  is a sum of  $\lambda x^{2s}$  plus  $s$  powers of linear forms over  $\mathbb{C}$ .

Sylvester’s Theorem can also be adapted to compute  $K$ -length when  $K \subsetneq \mathbb{C}$ , with the understanding that a Sylvester form of minimal degree might not split over  $K$ .

**Corollary 2.2.** *Given  $f \in K[x, y]$ ,  $L_K(f)$  is the minimal degree of a Sylvester form for  $f$  which splits completely over  $K$ .*

*Proof.* If (2.3) is a minimal representation for  $f$  over  $K$ , where  $\lambda_k, \alpha_k, \beta_k \in K$ , then  $h(x, y) \in K[x, y]$  splits over  $K$  by (2.2). Conversely, if  $h$  is a Sylvester form for  $f$  satisfying (2.2) with  $\alpha_k, \beta_k \in K$ , then (2.3) holds for some  $\lambda_k \in \mathbb{C}$ . This is equivalent to saying that the linear system

$$(2.13) \quad a_j = \sum_{k=1}^r \alpha_k^{d-j} \beta_k^j X_k, \quad (0 \leq j \leq d)$$

has a solution  $\{X_k = \lambda_k\}$  over  $\mathbb{C}$ . Since  $a_j, \alpha_k^{d-j} \beta_k^j \in K$  and (2.13) has a solution over  $\mathbb{C}$ , it also has a solution over  $K$ . Thus,  $f$  has a  $K$ -representation of length  $r$ .  $\square$

*Example 2.1* (Continuing Example 1.2). Note that

$$\begin{aligned} \phi(x, y) &= 3x^5 - 20x^3y^2 + 10xy^4 = \binom{5}{0} \cdot 3 x^5 + \binom{5}{1} \cdot 0 x^4y \\ &+ \binom{5}{2} \cdot (-2) x^3y^2 + \binom{5}{3} \cdot 0 x^2y^3 + \binom{5}{4} \cdot 2 xy^4 + \binom{5}{5} \cdot 0 y^5. \end{aligned}$$

Since

$$\begin{pmatrix} 3 & 0 & -2 & 0 \\ 0 & -2 & 0 & 2 \\ -2 & 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff (c_0, c_1, c_2, c_3) = r(0, 1, 0, 1),$$

$\phi$  has a *unique* Sylvester form of degree 3:  $h(x, y) = y(x^2 + y^2) = y(y - ix)(y + ix)$ . Accordingly, there exist  $\lambda_k \in \mathbb{C}$  so that

$$\phi(x, y) = \lambda_1 x^5 + \lambda_2 (x + iy)^5 + \lambda_3 (x - iy)^5.$$

Indeed,  $\lambda_1 = \lambda_2 = \lambda_3 = 1$ , as may be checked. It follows that  $L_K(\phi) = 3$  if and only if  $i \in K$ . (There is no representation of length two.)

To find representations for  $\phi$  of length 4, we consider (2.4) for  $\phi$  with  $r = 4$ :

$$\begin{aligned} H_4(\phi) \cdot (c_0, c_1, c_2, c_3, c_4)^t &= (0, 0)^t \iff 3c_0 - 2c_2 + 2c_4 = -2c_1 + 2c_3 = 0 \\ &\iff (c_0, c_1, c_2, c_3, c_4) = r_1(2, 0, 3, 0, 0) + r_2(0, 1, 0, 1, 0) + r_3(0, 0, 1, 0, 1), \end{aligned}$$

hence  $h(x, y) = r_1x^2(2x^2 + 3y^2) + y(x^2 + y^2)(r_2x + r_3y)$ . Given a field  $K$ , it is unclear whether there exist  $\{r_\ell\}$  so that  $h$  splits into distinct factors over  $K$ . We have found such  $\{r_\ell\}$  for small imaginary quadratic fields.

The choice  $(r_1, r_2, r_3) = (1, 0, 2)$  gives  $h(x, y) = (2x^2 + y^2)(x^2 + 2y^2)$  and

$$24\phi(x, y) = 4(x + \sqrt{-2}y)^5 + 4(x - \sqrt{-2}y)^5 + (2x + \sqrt{-2}y)^5 + (2x - \sqrt{-2}y)^5.$$

Similarly,  $(r_1, r_2, r_3) = (2, 0, 9)$  and  $(2, 0, -5)$  give  $h(x, y) = (x^2 + 3y^2)(4x^2 + 3y^2)$  and  $(x^2 - y^2)(4x^2 + 5y^2)$ , leading to representations for  $\phi$  of length 4 over  $\mathbb{Q}(\sqrt{-3})$  and  $\mathbb{Q}(\sqrt{-5})$ . The simplest such representation we have found for  $\mathbb{Q}(\sqrt{-6})$  uses  $(r_1, r_2, r_3) = (8450, 0, -104544)$  and

$$h(x, y) = (5x + 12y)(5x - 12y)(6 \cdot 13^2x^2 + 33^2y^2).$$

We conjecture that  $L_{\mathbb{Q}(\sqrt{-m})}(\phi) = 4$  for all non-square  $m \geq 2$ . In Example 3.1, we shall show that there is no choice of  $(r_1, r_2, r_3)$  for which  $h$  splits into distinct factors over any subfield of  $\mathbb{R}$ .

### 3. SYLVESTER'S 1864 THEOREM

Theorem 3.1 was discovered by Sylvester [55] in 1864 while proving Isaac Newton's conjectural variation on Descartes' Rule of Signs, see [27, 56]. This theorem appeared in Pólya-Szegő [42, Ch.5, Prob.79], and has been used by Pólya and Schoenberg [41] and Karlin [29, p.466]. The (dehomogenized) version proved in [42] is:

**Theorem 3.1** (Sylvester). *Suppose  $0 \neq \lambda_k$  for all  $k$  and  $\gamma_1 < \dots < \gamma_r$ ,  $r \geq 2$ , are real numbers such that*

$$Q(t) = \sum_{k=1}^r \lambda_k (t - \gamma_k)^d$$

*does not vanish identically. Suppose the sequence  $(\lambda_1, \dots, \lambda_r, (-1)^d \lambda_1)$  has  $C$  changes of sign and  $Q$  has  $Z$  zeros, counting multiplicity. Then  $Z \leq C$ .*

We shall prove an equivalent version which exploits the homogeneity of  $f$  to avoid discussion of zeros at infinity in the proof. (The equivalence is discussed in [50].)

**Theorem 3.2.** *Suppose  $f(x, y)$  is a non-zero real form of degree  $d$  with  $\tau$  real linear factors (counting multiplicity) and*

$$(3.1) \quad f(x, y) = \sum_{j=1}^r \lambda_j (\cos \theta_j x + \sin \theta_j y)^d$$



where  $-\frac{\pi}{2} < \theta_1 < \dots < \theta_r \leq \frac{\pi}{2}$ ,  $r \geq 2$  and  $\lambda_j \neq 0$ . If there are  $\sigma$  sign changes in the tuple  $(\lambda_1, \lambda_2, \dots, \lambda_r, (-1)^d \lambda_1)$ , then  $\tau \leq \sigma$ . In particular,  $\tau \leq r$ .

*Example 3.1* (Examples 1.2 and 2.1 concluded). Since

$$\phi(x, y) = 3x(x^2 - \frac{10-\sqrt{70}}{3}y^2)(x^2 - \frac{10+\sqrt{70}}{3}y^2)$$

is a product of five linear factors over  $\mathbb{R}$ ,  $L_{\mathbb{R}}(\phi) \geq 5$ . The representation

$$6\phi(x, y) = 36x^5 - 10(x+y)^5 - 10(x-y)^5 + (x+2y)^5 + (x-2y)^5$$

over  $\mathbb{Q}$  implies that  $L_{\mathbb{R}}(\phi) = 5$ . It will follow from Theorem 4.10 that  $\mathcal{C}(\phi) = \{3, 4, 5\}$ .

*Proof of Theorem 3.2.* We first rewrite (3.1):

$$(3.2) \quad \begin{aligned} 2f(x, y) &= \sum_{j=1}^r \lambda_j (\cos \theta_j x + \sin \theta_j y)^d + \\ &\sum_{j=1}^r (-1)^d \lambda_d (\cos(\theta_j + \pi)x + \sin(\theta_j + \pi)y)^d \end{aligned}$$

View the sequence  $(\lambda_1, \lambda_2, \dots, \lambda_r, (-1)^d \lambda_1, (-1)^d \lambda_2, \dots, (-1)^d \lambda_r, \lambda_1)$  cyclically, identifying the first and last term. There are  $2\sigma$  pairs of consecutive terms with a negative product. This count is independent of the starting point, so if we make any invertible change of variables  $(x, y) \mapsto (\cos \theta x + \sin \theta y, -\sin \theta x + \cos \theta y)$  in (3.1) (which doesn't affect  $\tau$ , and which "dials" the angles by  $\theta$ ), and reorder the "main" angles to  $(-\frac{\pi}{2}, \frac{\pi}{2}]$ , the value of  $\sigma$  is unchanged. We may therefore assume that neither  $x$  nor  $y$  divide  $f$ , that  $x^d$  and  $y^d$  are not summands in (3.2) (i.e.,  $\theta_j$  is not a multiple of  $\frac{\pi}{2}$ ), and that if there is a sign change in  $(\lambda_1, \lambda_2, \dots, \lambda_r)$ , then  $\theta_u < 0 < \theta_{u+1}$  implies  $\lambda_u \lambda_{u+1} < 0$ . Under these hypotheses, we may safely dehomogenize  $f$  by setting either  $x = 1$  or  $y = 1$  and avoid zeros at infinity and know that  $\tau$  is the number of zeros of the resulting polynomial. The rest of the proof generally follows [42].

Let  $\bar{\sigma}$  denote the number of sign changes in  $(\lambda_1, \lambda_2, \dots, \lambda_r)$ . We induct on  $\bar{\sigma}$ . The base case is  $\bar{\sigma} = 0$  (and  $\lambda_j > 0$  without loss of generality). If  $d$  is even, then  $\sigma = 0$  and

$$f(x, y) = \sum_{j=1}^r \lambda_j (\cos \theta_j x + \sin \theta_j y)^d$$

is definite, so  $\tau = 0$ . If  $d$  is odd, then  $\sigma = 1$ . Let  $g(t) = f(t, 1)$ , so that

$$g'(t) = \sum_{j=1}^r d(\lambda_j \cos \theta_j) (\cos \theta_j t + \sin \theta_j)^{d-1}.$$

Since  $d - 1$  is even,  $\cos \theta_j > 0$  and  $\lambda_j > 0$ ,  $g'$  is definite and  $g' \neq 0$ . Rolle's Theorem implies that  $g$  has at most one zero; that is,  $\tau \leq 1 = \sigma$ .

Suppose the theorem is valid for  $\bar{\sigma} = m \geq 0$  and suppose that  $\bar{\sigma} = m + 1$  in (3.1). Now let  $h(t) = f(1, t)$ . We have

$$h'(t) = \sum_{j=1}^r d(\lambda_j \sin \theta_j) (\cos \theta_j + \sin \theta_j t)^{d-1}.$$

Note that  $h'(t) = q(1, t)$ , where

$$q(x, y) = \sum_{j=1}^r d(\lambda_j \sin \theta_j) (\cos \theta_j x + \sin \theta_j y)^{d-1}.$$

Since  $\bar{\sigma} \geq 1$ ,  $\theta_u < 0 < \theta_{u+1}$  implies that  $\lambda_u \lambda_{u+1} < 0$ , so that the number of sign changes in  $(d\lambda_1 \sin \theta_1, d\lambda_2 \sin \theta_2, \dots, d\lambda_r \sin \theta_r)$  is  $m$ , as the sign change at the  $u$ -th consecutive pair has been removed, and no other possible sign changes are introduced. The induction hypothesis implies that  $q(x, y)$  has at most  $m$  linear factors, hence  $q(1, t) = h'(t)$  has  $\leq m$  zeros (counting multiplicity) and Rolle's Theorem implies that  $h$  has  $\leq m + 1$  zeros, completing the induction.  $\square$

#### 4. APPLICATIONS TO FORMS OF GENERAL DEGREE

We begin with a folklore result: the vector space of complex forms  $f$  in  $n$  variables of degree  $d$  is spanned by the set of linear forms taken to the  $d$ -th power. It follows from a 1903 theorem of Biermann (see [47, Prop.2.11] or [51] for a proof) that a canonical set of the ‘‘correct’’ number of  $d$ -th powers over  $\mathbb{Z}$  forms a basis:

$$(4.1) \quad \{(i_1 x_1 + \dots + i_n x_n)^d : 0 \leq i_k \in \mathbb{Z}, i_1 + \dots + i_n = d\}.$$

If  $f \in K[x_1, \dots, x_n]$ , then  $f$  is a  $K$ -linear combination of these forms and so  $L_K(f) \leq \binom{n+d-1}{n-1}$ . We show below (Theorems 4.10, 5.4) that when  $n = 2$ , the bound for  $L_K(f)$  can be improved from  $d + 1$  to  $d$ , but this is best possible.

The first two results are presented explicitly for completeness.

**Theorem 4.1.** *If  $f \in K[x, y]$ , then  $L_K(f) = 1$  if and only if  $L_{\mathbb{C}}(f) = 1$ .*

*Proof.* One direction is immediate from (1.2). For the other, suppose  $f(x, y) = (\alpha x + \beta y)^d$  with  $\alpha, \beta \in \mathbb{C}$ . If  $\alpha = 0$ , then  $f(x, y) = \beta^d y^d$ , with  $\beta^d \in K$ . If  $\alpha \neq 0$ , then  $f(x, y) = \alpha^d (x + (\beta/\alpha)y)^d$ . Since the coefficients of  $x^d$  and  $dx^{d-1}y$  in  $f$  are  $\alpha^d$  and  $\alpha^{d-1}\beta$ , it follows that  $\alpha^d$  and  $\beta/\alpha = (\alpha^{d-1}\beta)/\alpha^d$  are both in  $K$ .  $\square$

**Theorem 4.2.** *Any set  $\{(\alpha_j x + \beta_j y)^d : 0 \leq j \leq d\}$  of pairwise distinct  $d$ -th powers is linearly independent and spans the binary forms of degree  $d$ .*

*Proof.* The matrix of this set with respect to the basis  $\binom{d}{i} x^{d-i} y^i$  is  $[\alpha_j^{d-i} \beta_j^i]$ , whose determinant is Vandermonde:

$$\prod_{0 \leq j < k \leq d} \begin{vmatrix} \alpha_j & \beta_j \\ \alpha_k & \beta_k \end{vmatrix}.$$

This determinant is a product of non-zero terms by hypothesis.  $\square$

By considering the difference of two representations of a given form, we obtain an immediate corollary about different representations of the same form. Trivial counterexamples, formed by splitting summands, occur in non-honest representations.

**Corollary 4.3.** *If  $f$  has two different honest representations:*

$$(4.2) \quad f(x, y) = \sum_{i=1}^s \lambda_i (\alpha_i x + \beta_i y)^d = \sum_{j=1}^t \mu_j (\gamma_j x + \delta_j y)^d,$$

then  $s + t \geq d + 2$ . If  $s + t = d + 2$  in (4.2), then the combined set of linear forms,  $\{\alpha_i x + \beta_i y, \gamma_j x + \delta_j y\}$ , is pairwise distinct.

The next result collects some consequences of Corollary 4.3.

**Corollary 4.4.** *Let  $E = E_f$ .*

- (1) *If  $L_E(f) = r \leq \frac{d}{2} + 1$ , then  $L_{\mathbb{C}}(f) = r$ , so  $\mathcal{C}(f) = \{r\}$ .*
- (2) *If, further,  $L_E(f) = r \leq \frac{d}{2} + \frac{1}{2}$ , then  $f$  has a unique  $\mathbb{C}$ -minimal representation.*
- (3) *If  $d = 2s - 1$  and  $H_s(f)$  has full rank,  $f$  has a unique Sylvester form  $h$  of degree  $s$  and  $E_f \subseteq K$ , then  $L_K(f) \geq s$ , with equality if and only if  $h$  splits in  $K$ .*

*Proof.* We take the parts in turn.

- (1) A different representation of  $f$  over  $\mathbb{C}$  must have length  $\geq d + 2 - r \geq \frac{d}{2} + 1 \geq r$  by Corollary 4.3, and so  $L_{\mathbb{C}}(f) = r$ .
- (2) If  $r \leq \frac{d}{2} + \frac{1}{2}$ , then any other representation has length  $\geq \frac{d}{2} + \frac{3}{2} > r$ , and so cannot be minimal.
- (3) If  $d = 2s - 1$  and  $r = s$ , then the last case applies, so  $f$  has a unique  $\mathbb{C}$ -minimal representation, and by Corollary 2.2, this representation can be expressed in  $K$  if and only if the Sylvester form splits over  $K$ . □

We now give some more explicit constructions of forms with multiple lengths. We first need a lemma about cubics.

**Lemma 4.5.** *If  $f$  is a cubic given by (2.1) and  $H_2(f) = \begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \end{pmatrix}$  has rank  $\leq 1$ , then  $f$  is a cube.*

*Proof.* If  $a_0 = 0$ , then  $a_1 = 0$ , so  $a_2 = 0$  and  $f$  is a cube. If  $a_0 \neq 0$ , then  $a_2 = a_1^2/a_0$  and  $a_3 = a_1 a_2/a_0 = a_1^3/a_0^2$  and  $f(x, y) = a_0(x + \frac{a_1}{a_0}y)^3$  is again a cube. □

**Theorem 4.6.** *Suppose  $d \geq 3$  and there exist  $\alpha_i, \beta_i \in \mathbb{C}$  so that*

$$(4.3) \quad f(x, y) = \sum_{i=0}^d \binom{d}{i} a_i x^{d-i} y^i = (\alpha_1 x + \beta_1 y)^d + (\alpha_2 x + \beta_2 y)^d \in K[x, y].$$

*If (4.3) is honest and  $L_K(f) > 2$ , then there exists  $u \in K$  with  $\sqrt{u} \notin K$  so that  $L_{K(\sqrt{u})}(f) = 2$ . The summands in (4.3) are conjugates of each other in  $K(\sqrt{u})$ .*

*Proof.* First observe that if  $\alpha_2 = 0$ , then  $\alpha_2\beta_1 \neq \alpha_1\beta_2$  implies that  $\alpha_1 \neq 0$ . But then  $a_0 = \alpha_1^d \neq 0$  and  $a_1 = \alpha_1^{d-1}\beta_1$  imply that  $\alpha_1^d, \beta_1/\alpha_1 \in K$  as in Theorem 4.1, and so

$$f(x, y) - \alpha_1^d(x + (\beta_1/\alpha_1)y)^d = (\beta_2y)^d = \beta_2^d y^d \in K[x, y].$$

This contradicts  $L_K(f) > 2$ , so  $\alpha_2 \neq 0$ ; similarly,  $\alpha_1 \neq 0$ . Let  $\lambda_i = \alpha_i^d$  and  $\gamma_i = \beta_i/\alpha_i$  for  $i = 1, 2$ , so  $\lambda_1\lambda_2 \neq 0$  and  $\gamma_1 \neq \gamma_2$ . We have

$$f(x, y) = \lambda_1(x + \gamma_1y)^d + \lambda_2(x + \gamma_2y)^d \implies a_i = \lambda_1\gamma_1^i + \lambda_2\gamma_2^i.$$

Now let

$$g(x, y) = \lambda_1(x + \gamma_1y)^3 + \lambda_2(x + \gamma_2y)^3 = a_0x^3 + 3a_1x^2y + 3a_2xy^2 + a_3y^3.$$

Since  $\lambda_i \neq 0$  and (4.3) is honest, Corollary 4.3 implies that  $L_{\mathbb{C}}(g) = 2$ , so  $H_2(g)$  has full rank by Lemma 4.5. It can be checked directly that

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \end{pmatrix} \cdot \begin{pmatrix} \gamma_1\gamma_2 \\ -(\gamma_1 + \gamma_2) \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and this gives  $h(x, y) = (y - \gamma_1x)(y - \gamma_2x)$  as the unique Sylvester form for  $g$ . Since  $H_2(g)$  has entries in  $K$  and hence has a null vector in  $K$ , we must have  $h \in K[x, y]$ . By hypothesis,  $h$  does not split over  $K$ ; it must do so over  $K(\sqrt{u})$ , where  $u = (\gamma_1 - \gamma_2)^2 = (\gamma_1 + \gamma_2)^2 - 4\gamma_1\gamma_2 \in K$ . Moreover, if  $\sigma$  denotes conjugation with respect to  $\sqrt{u}$ , then  $\gamma_2 = \sigma(\gamma_1)$  and since  $\lambda_1 + \lambda_2 \in K$ ,  $\lambda_2 = \sigma(\lambda_1)$  as well. Note that  $\lambda_i = \alpha_i^d$  and  $\gamma_i = \beta_i/\alpha_i \in K(\sqrt{u})$ , but this is not necessarily true for  $\alpha_i$  and  $\beta_i$  themselves.  $\square$

**Corollary 4.7.** *Suppose  $g \in E[x, y]$  does not split over  $E$ , but factors into distinct linear factors  $g(x, y) = \prod_{j=1}^r (x + \alpha_jy)$  over an extension field  $K$  of  $E$ . If  $d > 2r - 1$ , then for each  $\ell \geq 0$ ,*

$$f_\ell(x, y) = \sum_{j=1}^r \alpha_j^\ell (x + \alpha_jy)^d \in E[x, y],$$

and  $L_K(f_\ell) = r < d + 2 - r \leq L_E(f_\ell)$ .

*Proof.* The coefficient of  $\binom{d}{k}x^{d-k}y^k$  in  $f_\ell$  is  $\sum_{j=1}^r \alpha_j^{\ell+k}$ . Each such power-sum belongs to  $E$  by Newton's Theorem on Symmetric Forms. If  $\alpha_s \notin E$  (which must hold for at least one  $\alpha_s \neq 0$ ), then  $\alpha_s^\ell (x + \alpha_sy)^d \notin E[x, y]$ . Apply Corollary 4.3.  $\square$

**Corollary 4.8.** *Suppose  $K$  is an extension field of  $E_f$ ,  $r \leq \frac{d+1}{2}$ , and*

$$f(x, y) = \sum_{i=1}^r \lambda_i(\alpha_ix + \beta_iy)^d$$

with  $\lambda_i, \alpha_i, \beta_i \in K$ . Then every automorphism of  $K$  which fixes  $E_f$  permutes the summands of the representation of  $f$ .

*Proof.* We interpret  $\sigma(\lambda(\alpha x + \beta y)^d) = \sigma(\lambda)(\sigma(\alpha)x + \sigma(\beta)y)^d$ . Since  $\sigma(f) = f$ , the action of  $\sigma$  is to give another representation of  $f$ . Corollary 4.4(2) implies that this is the same representation, perhaps reordered.  $\square$

**Theorem 4.9.** *If  $f \in K[x, y]$ , then  $L_{\mathbb{C}}(f) \leq \deg f$ .*

*Proof.* By a change of variables, which does not affect the length, we may assume that neither  $x$  nor  $y$  divide  $f$ , hence  $a_0 a_d \neq 0$  and  $h = a_d x^d - a_0 y^d$  is a Sylvester form which splits over  $\mathbb{C}$ .  $\square$

We have been unable to find an “original” citation for Theorem 4.9. It appears as an exercise in Harris [24, Ex.11.35], with the (dehomogenized) maximal length occurring at  $x^{d-1}(x+1)$  (see Theorem 5.4). Landsberg and Teitler [34, Cor. 5.2] prove that  $L_{\mathbb{C}}(f) \leq \binom{n+d-1}{n-1} - (n-1)$ , which reduces to Theorem 4.9 for  $n=2$ . The proof of Theorem 4.9 will not apply to  $L_K(f)$  for  $K \neq \mathbb{C}$ , because  $a_d x^d - a_0 y^d$  usually does not split over  $K$ . A more careful argument is required, constructing an explicit Sylvester form of degree  $d$  for  $f$  which splits over  $K$ .

**Theorem 4.10.** *If  $f \in K[x, y]$ , then  $L_K(f) \leq \deg f$ .*

*Proof.* Write  $f$  as in (2.1). If  $f$  is identically zero, there is nothing to prove. Otherwise, we may assume that  $f(1, 0) = a_0 \neq 0$  after a change of variables if necessary. By Corollary 2.2, it suffices to find  $h(x, y) = \sum_{k=0}^d c_k x^{d-k} y^k$  which splits into distinct linear factors over  $K$  and satisfies  $\sum_{k=0}^d a_k c_k = 0$ .

Let  $e_0 = 1$  and  $e_k(t_1, \dots, t_{d-1})$  denote the usual  $k$ -th elementary symmetric functions. We make a number of definitions:

$$h_0(t_1, \dots, t_{d-1}; x, y) := \sum_{k=0}^{d-1} e_k(t_1, \dots, t_{d-1}) x^{d-1-k} y^k = \prod_{j=1}^{d-1} (x + t_j y),$$

$$\beta(t_1, \dots, t_{d-1}) := - \sum_{k=0}^{d-1} a_k e_k(t_1, \dots, t_{d-1}),$$

$$\alpha(t_1, \dots, t_{d-1}) := \sum_{k=0}^{d-1} a_{k+1} e_k(t_1, \dots, t_{d-1}),$$

$$\Phi(t_1, \dots, t_{d-1}) := \prod_{j=1}^{d-1} (\alpha(t_1, \dots, t_{d-1}) t_j - \beta(t_1, \dots, t_{d-1})),$$

$$\Psi(t_1, \dots, t_{d-1}) := \Phi(t_1, \dots, t_{d-1}) \times \prod_{1 \leq i < j \leq d-1} (t_i - t_j).$$

Then  $\beta(0, \dots, 0) = -a_0 e_0 = -a_0 \neq 0$ , so  $\Phi(0, \dots, 0) = a_0^{d-1} \neq 0$  and  $\Phi$  is not the zero polynomial, and thus neither is  $\Psi$ . Choose  $\gamma_j \in K$ ,  $1 \leq j \leq d-1$ , so that  $\Psi(\gamma_1, \dots, \gamma_{d-1}) \neq 0$ . It follows that the  $\gamma_j$ 's are distinct, and  $\alpha \gamma_j \neq \beta$ , where

$\alpha = \alpha(\gamma_1, \dots, \gamma_{d-1})$  and  $\beta = \beta(\gamma_1, \dots, \gamma_{d-1})$ . Let  $\tilde{e}_k = e_k(\gamma_1, \dots, \gamma_{d-1})$ . We claim that

$$\begin{aligned} h(x, y) &= \sum_{i=0}^d c_i x^{d-1} y^i := (\alpha x + \beta y) h_0(\gamma_1, \dots, \gamma_{d-1}; x, y) = (\alpha x + \beta y) \prod_{j=1}^{d-1} (x + \gamma_j y) \\ &= (\alpha x + \beta y) \sum_{k=0}^{d-1} \tilde{e}_k x^{d-1-k} y^k = \alpha \tilde{e}_0 x^d + \sum_{k=1}^{d-1} (\alpha \tilde{e}_k + \beta \tilde{e}_{k-1}) x^{d-k} y^k + \beta \tilde{e}_{d-1} y^d \end{aligned}$$

is a Sylvester form for  $f$ . Note that the  $\gamma_j$ 's are distinct and  $\alpha \gamma_j \neq \beta$ ,  $1 \leq j \leq d-1$ , so that  $h$  is a product of distinct linear factors. Finally,

$$\begin{aligned} \sum_{k=0}^d a_k c_k &= \alpha \tilde{e}_0 a_0 + \sum_{k=1}^{d-1} (\alpha \tilde{e}_k + \beta \tilde{e}_{k-1}) a_k + \beta \tilde{e}_{d-1} a_d = \\ &= \alpha \sum_{k=0}^{d-1} \tilde{e}_k a_k + \beta \sum_{k=0}^{d-1} \tilde{e}_k a_{k+1} = \alpha(-\beta) + \beta\alpha = 0. \end{aligned}$$

This completes the proof.  $\square$

**Corollary 4.11.** *If  $f$  is a product of  $d$  real linear forms and not a  $d$ -th power, then  $L_{\mathbb{R}}(f) = d$ .*

*Proof.* Write  $f$  as a sum of  $L_{\mathbb{R}}(f) = r \leq d$   $d$ -th powers and rescale into the shape (3.1). Taking  $\tau = d$  in Theorem 3.2, we see that  $d \leq \sigma \leq r$ .  $\square$

**Conjecture 4.12.** *If  $f \in \mathbb{R}[x, y]$  is a form of degree  $d \geq 3$ , then  $L_{\mathbb{R}}(f) = d$  if and only if  $f$  is a product of  $d$  linear forms.*

We shall see in Theorems 5.2 and 5.3 that this conjecture is true for  $d = 3, 4$ .

After a preprint of this paper was distributed, Giorgio Ottaviani pointed out that in the case that the roots of  $f$  are distinct, Conjecture 4.12 has been proved very recently by Comon and Ottaviani [15] and by Causa and Re [9].

## 5. APPLICATIONS TO FORMS OF PARTICULAR DEGREE

Corollary 4.3 and Theorem 4.10 impose some immediate restrictions on the possible cabinets of a form of degree  $d$ .

**Corollary 5.1.** *Suppose  $\deg f = d$ .*

- (1) *If  $L_{\mathbb{C}}(f) = r$ , then  $\mathcal{C}(f) \subseteq \{r\} \cup \{d - i : 0 \leq i \leq r - 2\}$ .*
- (2) *If  $L_{\mathbb{C}}(f) = 2$ , then  $\mathcal{C}(f)$  is either  $\{2\}$  or  $\{2, d\}$ .*
- (3) *If  $f$  has  $k$  different lengths, then  $d \geq 2k - 1$ .*
- (4) *If  $f$  is cubic, then  $\mathcal{C}(f) = \{1\}, \{2\}, \{3\}$  or  $\{2, 3\}$ .*
- (5) *If  $f$  is quartic, then  $\mathcal{C}(f) = \{1\}, \{2\}, \{3\}, \{4\}, \{2, 4\}$  or  $\{3, 4\}$ .*

We now completely classify  $L_K(f)$  when  $f$  is a binary cubic.

**Theorem 5.2.** *Suppose  $f(x, y) \in E_f[x, y]$  is a cubic form with discriminant  $\Delta$  and suppose  $E_f \subseteq K \subseteq \mathbb{C}$ .*

- (1) *If  $f$  is a cube, then  $L_K(f) = 1$  and  $\mathcal{C}(f) = \{1\}$ .*
- (2) *If  $f$  has a repeated linear factor, but is not a cube, then  $L_K(f) = 3$  and  $\mathcal{C}(f) = \{3\}$ .*
- (3) *If  $f$  does not have a repeated factor, then  $L_K(p) = 2$  if  $\sqrt{-3\Delta} \in K$  and  $L_K(p) = 3$  otherwise, so either  $\mathcal{C}(f) = \{2\}$  or  $\mathcal{C}(f) = \{2, 3\}$ .*

*Proof.* The first case follows from Theorem 4.1. In the second case, after an invertible linear change of variables, we may assume that  $f(x, y) = 3x^2y$ , and apply Theorem 2.1 to test for representations of length 2. But

$$(5.1) \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies c_0 = c_1 = 0,$$

so  $h$  has repeated factors. Hence  $L_K(x^2y) \geq 3$  and by Theorem 4.10,  $L_K(x^2y) = 3$ .

Finally, suppose

$$f(x, y) = a_0x^3 + 3a_1x^2y + 3a_2xy^2 + a_3y^3 = \prod_{j=1}^3 (r_jx + s_jy)$$

does not have repeated factors, so that

$$0 \neq \Delta(f) = \prod_{j < k} (r_js_k - r_k s_j)^2,$$

and consider the system:

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

By Lemma 4.5, this system has rank 2; the unique Sylvester form is

$$h(x, y) = (a_1a_3 - a_2^2)x^2 + (a_1a_2 - a_0a_3)xy + (a_0a_2 - a_1^2)y^2,$$

which happens to be the Hessian of  $f$ . Since  $h \in E_f[x, y] \subseteq K[x, y]$ , it splits over  $K$  if and only if its discriminant is a square in  $K$ . A computation shows that

$$(a_1a_2 - a_0a_3)^2 - 4(a_1a_3 - a_2^2)(a_0a_2 - a_1^2) = -\frac{\Delta(f)}{27} = -\frac{3\Delta(f)}{9^2}.$$

Thus,  $L_K(f) = 2$  if and only if  $\sqrt{-3\Delta(f)} \in K$ . If  $h$  does not split over  $F$ , then  $L_F(f) = 3$  by Theorem 4.10.  $\square$

In particular,  $x^3$ ,  $x^3 + y^3$ ,  $x^2y$  and  $(x + iy)^3 + (x - iy)^3$  have the cabinets enumerated in Corollary 5.1(4). If  $f$  has three distinct real linear factors, then  $\Delta(f) > 0$ , so  $\sqrt{-3\Delta(f)} \notin \mathbb{R}$  and  $L_{\mathbb{R}}(f) = 3$ . If  $f$  is real and has one real and two conjugate complex linear factors, then  $\Delta(f) < 0$ , so  $L_{\mathbb{R}}(f) = 2$ . Counting repeated roots, we

see that if  $f$  is a real cubic, and not a cube, then  $L_{\mathbb{R}}(f) = 3$  if and only if it has three real factors, thus proving Conjecture 4.12 when  $d = 3$ .

*Example 5.1.* We find all representations of  $3x^2y$  of length 3. Note that

$$H_3(f) \cdot (c_0, c_1, c_2, c_3)^t = (0) \iff c_1 = 0 \iff h(x, y) = c_0x^3 + c_2xy^2 + c_3y^3.$$

If  $c_0 = 0$ , then  $y^2 \mid h$ , which is to be avoided, so we scale and assume  $c_0 = 1$ . We can parameterize the Sylvester forms as  $h(x, y) = (x - ay)(x - by)(x + (a + b)y)$  with  $a, b, -(a + b)$  distinct. This leads to an easily checked general formula

$$(5.2) \quad \begin{aligned} & 3(a - b)(a + 2b)(2a + b)x^2y = \\ & (a + 2b)(ax + y)^3 - (2a + b)(bx + y)^3 + (a - b)(-(a + b)x + y)^3. \end{aligned}$$

Białynicki-Birula and Schinzel [2, Lemma 7.1] give the general formula for  $dx^{d-1}y$  as a sum of  $d$   $d$ -th powers of linear forms.

**Theorem 5.3.** *If  $f$  is a real quartic form, then  $L_{\mathbb{R}}(f) = 4$  if and only if  $f$  is a product of four linear factors.*

*Proof.* Factor  $\pm f$  as a product of  $k$  positive definite quadratic forms and  $4 - 2k$  linear forms. If  $k = 0$ , then Corollary 4.11 implies that  $L_{\mathbb{R}}(f) = 4$ . We must show that if  $k = 1$  or  $k = 2$ , then  $f$  has a representation over  $\mathbb{R}$  as a sum of  $\leq 3$  fourth powers.

If  $k = 2$ , then  $f$  is positive definite and by [43, Thm.6], after an invertible linear change of variables,  $f(x, y) = x^4 + 6\lambda x^2y^2 + y^4$ , with  $6\lambda \in (-2, 2]$ . (This is also proved in [51].) If  $r \neq 1$ , then

$$(5.3) \quad \begin{aligned} & (rx + y)^4 + (x + ry)^4 - (r^3 + r)(x + y)^4 \\ & = (r - 1)^2(r^2 + r + 1) \left( x^4 - \left( \frac{6r}{r^2 + r + 1} \right) x^2y^2 + y^4 \right). \end{aligned}$$

Let  $\phi(r) = -\frac{6r}{r^2 + r + 1}$ . Then  $\phi(-2 + \sqrt{3}) = 2$  and  $\phi(1) = -2$ , and since  $\phi$  is continuous, it maps  $[-2 + \sqrt{3}, 1)$  onto  $(-2, 2]$ , and (5.3) shows that  $L_{\mathbb{R}}(f) \leq 3$ .

If  $k = 1$ , there are two cases, depending on whether the linear factors are distinct. Suppose that after a linear change,  $f(x, y) = x^2h(x, y)$ , where  $h$  is positive definite, and so for some  $\lambda > 0$  and linear  $\ell$ ,  $h(x, y) = \lambda x^2 + \ell^2$ . After another linear change,

$$(5.4) \quad f(x, y) = x^2(2x^2 + 12y^2) = (x + y)^4 + (x - y)^4 - 2y^4,$$

and (5.4) shows that  $L_{\mathbb{R}}(f) \leq 3$ .

If the linear factors are distinct, then after a linear change,

$$f(x, y) = xy(ax^2 + 2bxy + cy^2)$$

where  $a > 0, c > 0, b^2 < ac$ . After a scaling,  $f(x, y) = xy(x^2 + dxy + y^2)$ ,  $|d| < 2$ , and by taking  $\pm f(x, \pm y)$ , we may assume  $d \in [0, 2)$ . If  $r \neq 1$ , then

$$(5.5) \quad \begin{aligned} & (r^4 + 1)(x + y)^4 - (rx + y)^4 - (x + ry)^4 \\ & = 4(r - 1)^2(r^2 + r + 1) \left( x^3y + \left( \frac{3(1+r)^2}{2(r^2+r+1)} \right) x^2y^2 + xy^3 \right). \end{aligned}$$



Let  $\psi(r) = \frac{3(1+r)^2}{2(r^2+r+1)}$ . Since  $\psi(-1) = 0$ ,  $\psi(1) = 2$  and  $\psi$  is continuous, it maps  $[-1, 1)$  onto  $[0, 2)$ , and (5.5) shows that  $L_{\mathbb{R}}(f) \leq 3$ .  $\square$

The next result may be very old;  $L_{\mathbb{C}}(x^{d-1}y) = d$  seems well known, but the only reference we have seen for the converse is the very recent [2, Corollary 3]. Białynicki-Birula and Schinzel also classify all binary  $p$  with  $\deg p = d$  and  $L_{\mathbb{C}}(p) = d - k$  for  $1 \leq k \leq 3$  and sufficiently large  $d$ . Landsberg and Teitler [34, Cor.4.5] and Boij, Carlini and Geramita [4] have both shown that  $L_{\mathbb{C}}(x^a y^b) = \max(a + 1, b + 1)$  if  $a, b \geq 1$ .

**Theorem 5.4.** *If  $d \geq 3$ , then  $L_{\mathbb{C}}(f) = d$  if and only if there are two distinct linear forms  $\ell$  and  $\ell'$  so that  $f = \ell^{d-1}\ell'$ .*

*Proof.* If  $f = \ell^{d-1}\ell'$ , then after an invertible linear change, we may assume that  $f(x, y) = dx^{d-1}y$ . If  $L_{\mathbb{C}}(dx^{d-1}y) \leq d - 1$ , then  $f$  would have a Sylvester form of degree  $d - 1$ . But then, as in (5.1), (2.4) becomes

$$\begin{pmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{d-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies c_0 = c_1 = 0,$$

so  $h(x, y) = \sum_{t=0}^{d-1} c_t x^{d-1-t} y^t$  does not have distinct factors. Thus,  $L_{\mathbb{C}}(dx^{d-1}y) = d$ .

Conversely, suppose  $L_{\mathbb{C}}(f) = d$ . Factor  $f = \prod \ell_j^{m_j}$  as a product of pairwise distinct linear forms, with  $\sum m_j = d$ ,  $m_1 \geq m_2 \geq \cdots \geq m_s \geq 1$ , and  $s > 1$  (otherwise,  $L_{\mathbb{C}}(f) = 1$ .) Make an invertible linear change taking  $(\ell_1, \ell_2) \mapsto (x, y)$ , and call the new form  $g$ ;  $L_{\mathbb{C}}(g) = d$  as well. If  $g(x, y) = \sum_{\ell=0}^d \binom{d}{\ell} b_{\ell} x^{d-\ell} y^{\ell}$ , then  $b_0 = b_d = 0$ . By hypothesis, there does not exist a Sylvester form of degree  $d - 1$  for  $g$ . Consider Theorem 2.1 in this case. We have

$$\begin{pmatrix} 0 & b_1 & \cdots & b_{d-2} & b_{d-1} \\ b_1 & b_2 & \cdots & b_{d-1} & 0 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{d-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

If  $m_1 \geq m_2 \geq 2$ , then  $x^2, y^2 \mid g(x, y)$  and  $b_1 = b_{d-1} = 0$  and  $x^{d-1} - y^{d-1}$  is a Sylvester form of degree  $d - 1$  for  $f$ . Thus  $m_2 = 1$  and so  $y^2$  does not divide  $g$  and  $b_1 \neq 0$ . Let  $q(t) = \sum_{i=0}^{d-2} b_{i+1} t^i$  (note the absence of binomial coefficients!) and suppose  $q$  is not the constant polynomial. Then there exists  $t_0$  so that  $q(t_0) = 0$ . Since  $q(0) = b_1$ ,  $t_0 \neq 0$ . We have

$$\begin{pmatrix} 0 & b_1 & \cdots & b_{d-2} & b_{d-1} \\ b_1 & b_2 & \cdots & b_{d-1} & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ t_0 \\ \vdots \\ t_0^{d-1} \end{pmatrix} = \begin{pmatrix} t_0 q(t_0) \\ q(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since

$$h(x, y) = \sum_{i=0}^{d-1} t_0^i x^{d-1-i} y^i = \frac{x^d - t_0^d y^d}{x - t_0 y} = \prod_{k=1}^{d-1} (x - \zeta_{d-1}^k t_0 y)$$

has distinct linear factors, it is a Sylvester form for  $g$ , and  $L_{\mathbb{C}}(g) \leq d - 1$ . This contradiction implies that  $q$  has no zeros, so  $q(t) = b_1$  must be a constant. It follows that  $g(x, y) = db_1 x^{d-1} y$ , as promised.  $\square$

By Corollaries 4.4 and 5.1, instances of the first five cabinets in Corollary 5.1(5) are:  $x^4$ ,  $x^4 + y^4$ ,  $x^4 + y^4 + (x + y)^4$ ,  $x^3 y$  and  $(x + iy)^4 + (x - iy)^4$ . It will follow from the next results that  $\mathcal{C}((x^2 + y^2)^2) = \{3, 4\}$ .

**Theorem 5.5.** *If  $d = 2k$  and  $f(x, y) = \binom{2k}{k} x^k y^k$ , then  $L_{\mathbb{C}}(f) = k + 1$ . The minimal  $\mathbb{C}$ -representations of  $f$  are given by*

$$(5.6) \quad (k + 1) \binom{2k}{k} x^k y^k = \sum_{j=0}^k (\zeta_{2k+2}^j w x + \zeta_{2k+2}^{-j} w^{-1} y)^{2k}, \quad 0 \neq w \in \mathbb{C}.$$

*Proof.* We first evaluate the right-hand side of (5.6) by expanding the powers:

$$(5.7) \quad \begin{aligned} \sum_{j=0}^k (\zeta_{2k+2}^j w x + \zeta_{2k+2}^{-j} w^{-1} y)^{2k} &= \sum_{j=0}^k \sum_{t=0}^{2k} \binom{2k}{t} \zeta_{2k+2}^{j(2k-t)-jt} w^{(2k-t)-t} x^{2k-t} y^t \\ &= \sum_{t=0}^{2k} \binom{2k}{t} w^{2k-2t} x^{2k-t} y^t \left( \sum_{j=0}^k \zeta_{k+1}^{j(k-t)} \right). \end{aligned}$$

But  $\sum_{j=0}^{m-1} \zeta_m^{rj} = 0$  unless  $m \mid r$ , in which case it equals  $m$ . Since the only multiple of  $k + 1$  in the set  $\{k - t : 0 \leq t \leq 2k\}$  occurs for  $t = k$ , (5.7) reduces to the left-hand side of (5.6). We now show that these are *all* the minimal  $\mathbb{C}$ -representations of  $f$ .

Since  $H_k(x^k y^k)$  has 1's on the NE-SW diagonal, it is non-singular, so  $L_{\mathbb{C}}(x^k y^k) > k$ , and  $L_{\mathbb{C}}(x^k y^k) = k + 1$  by (5.6). By Corollary 4.3, any minimal  $\mathbb{C}$ -representation *not* given by (5.6) can only use powers of forms which are distinct from *any*  $w x + w^{-1} y$ . If  $ab = c^2 \neq 0$ , then  $ax + by$  is a multiple of  $\frac{a}{c}x + \frac{c}{a}y$ . This leaves only  $x^{2k}$  and  $y^{2k}$ , and there is no linear combination of these giving  $x^k y^k$ .  $\square$

The representations in (5.6) arise because the null-vectors of  $H_{k+1}(x^k y^k)$  can only be  $(c_0, 0, \dots, 0, c_{k+1})^t$  and  $c_0 x^{k+1} + c_{k+1} y^{k+1}$  is a Sylvester form when  $c_0 c_{k+1} \neq 0$ .

**Corollary 5.6.** *For  $k \geq 2$ ,  $L_{\mathbb{C}}((x^2 + y^2)^k) = k + 1$ , and  $L_K((x^2 + y^2)^k) = k + 1$  iff  $\tan \frac{\pi}{k+1} \in K$ . The  $\mathbb{C}$ -minimal representations of  $(x^2 + y^2)^k$  are given by*

$$(5.8) \quad \binom{2k}{k} (x^2 + y^2)^k = \frac{1}{k+1} \sum_{j=0}^k \left( \cos\left(\frac{j\pi}{k+1} + \theta\right)x + \sin\left(\frac{j\pi}{k+1} + \theta\right)y \right)^{2k}, \quad \theta \in \mathbb{C}.$$

*Proof.* The invertible map  $(x, y) \mapsto (x - iy, x + iy)$  takes  $x^k y^k$  into  $(x^2 + y^2)^k$ . Setting  $0 \neq w = e^{i\theta}$  in (5.6) gives (5.8). If  $\tan \alpha \neq 0$ , then

$$(\cos \alpha x + \sin \alpha y)^{2k} = \cos^{2k} \alpha \cdot (x + \tan \alpha y)^{2k} = (1 + \tan^2 \alpha)^{-k} (x + \tan \alpha y)^{2k}.$$

Thus, if  $\tan \alpha \in K$ , then  $(\cos \alpha x + \sin \alpha y)^{2k} \in K[x, y]$ . Further, if  $\cos \alpha = 0$ , then  $(\cos \alpha x + \sin \alpha y)^{2k} = y^{2k} \in K[x, y]$ . Conversely, if  $(\cos \alpha x + \sin \alpha y)^{2k} \in K[x, y]$  and  $\cos \alpha \neq 0$ , then the ratio of the coefficients of  $x^{2k-1}y$  and  $x^{2k}$  equals  $2k \tan \alpha$ , which must be in  $K$ . It follows that  $L_K((x^2 + y^2)^k) = k + 1$  if and only if there exists  $\theta \in \mathbb{C}$  so that for each  $0 \leq j \leq k$ , either  $\cos(\frac{j\pi}{k+1} + \theta) = 0$  or  $\tan(\frac{j\pi}{k+1} + \theta) \in K$ . Since  $\tan \alpha, \tan \beta \in K$  imply  $\tan(\alpha - \beta) \in K$  and  $k \geq 2$ , we see that (5.8) is a representation over  $K$  if and only if  $\tan \frac{\pi}{k+1} \in K$ .  $\square$

In particular, since  $\tan \frac{\pi}{3} = \sqrt{3} \notin \mathbb{Q}$ ,  $L_{\mathbb{Q}}((x^2 + y^2)^2) > 3$  and so must equal 4. Thus,  $\mathcal{C}((x^2 + y^2)^2) = \{3, 4\}$ , as promised. Since  $\tan \frac{\pi}{m}$  is irrational for  $m \geq 5$  (see e.g. [40, Cor.3.12]), it follows that  $L_{\mathbb{Q}}((x^2 + y^2)^k) = k + 1$  only for  $k = 1, 3$ .

It is worth remarking that  $x^k y^k$  is a highly singular complex form, as is  $(x^2 + y^2)^k$ . However, as a *real* form,  $(x^2 + y^2)^k$  is interior to the real convex cone  $Q_{2,2k}$ . For real  $\theta$ , the formula in (5.8) goes back at least to Friedman [21] in 1957. It was shown in [47] that all minimal *real* representations of  $(x^2 + y^2)^k$  have this shape. There is an equivalence between representations of  $(x^2 + y^2)^k$  as a real sum of  $2k$ -th powers and quadrature formulas on the circle – see [47]. In this sense, (5.8) can be traced back to Mehler [35] in 1864.

A real representation (1.1) of  $(\sum x_i^2)^k$  (with positive real coefficients  $\lambda_j$ ) is called a Hilbert Identity; Hilbert [26, 20] used such representations with rational coefficients to solve Waring's problem. Hilbert Identities have been important in studying quadrature problems on  $S^{n-1}$ , the Delsarte-Goethals-Seidel theory of spherical designs in combinatorics and for embedding questions in Banach spaces [47, Ch.8,9], as well as for explicit computations in Hilbert's 17th problem [48]. It can be shown that any such representation requires at least  $\binom{n+k-1}{n-1}$  summands, and this bound also applies if negative coefficients  $\lambda_j$  are allowed. It is not known whether allowing negative coefficients can reduce the total number of summands. However, Blekherman [3] has recently constructed  $f \in Q_{6,4}$  which has a smaller length if one allows negative  $\lambda_j$  in a real representation. When  $(\sum x_i^2)^k$  is a sum of exactly  $\binom{n+k-1}{n-1}$   $2k$ -th powers, the coordinates of minimal representations can be used to produce tight spherical designs. Such representations exist when  $n = 2$ ,  $2k = 2$ ,  $(n, 2k) = (3, 4)$ ,  $(n, 2k) = (u^2 - 2, 4)$  ( $u = 3, 5$ ),  $(n, 2k) = (3v^2 - 4, 6)$  ( $v = 2, 3$ ),  $(n, 2k) = (24, 10)$ . It has been proved that they do not exist otherwise, unless possibly  $(n, 2k) = (u^2 - 2, 4)$  for some odd integer  $u \geq 7$  or  $(n, 2k) = (3v^2 - 4, 6)$  for some integer  $v \geq 4$ . These questions have been largely open for thirty years. It is also not known whether there exist  $(k, n)$  so that  $L_{\mathbb{R}}((\sum x_i^2)^k) > L_{\mathbb{C}}((\sum x_i^2)^k)$ , although this cannot happen for  $n = 2$ . For that matter, it is not known whether there exists any  $f \in Q_{n,d}$  so that  $L_{\mathbb{R}}(f) > L_{\mathbb{C}}(f)$ .

We conclude this section with a related question: if  $f_\lambda(x, y) = x^4 + 6\lambda x^2 y^2 + y^4$  for  $\lambda \in \mathbb{Q}$ , what is  $L_{\mathbb{Q}}(f_\lambda)$ ? If  $\lambda \leq -\frac{1}{3}$ , then  $f_\lambda$  has four real factors, so  $L_{\mathbb{Q}}(f_\lambda) = 4$ . Since  $\det H_2(f_\lambda) = \lambda - \lambda^3$ ,  $L_{\mathbb{C}}(f_\lambda) = 2$  for  $\lambda = 0, 1, -1$ . The formula

$$(x^4 + 6\lambda x^2 y^2 + y^4) = \frac{\lambda}{2} ((x+y)^4 + (x-y)^4) + (1-\lambda)(x^4 + y^4).$$

shows that  $L_{\mathbb{Q}}(f_0) = L_{\mathbb{Q}}(f_1) = 2$ ;  $2f_{-1}(x, y) = (x+iy)^4 + (x-iy)^4$  has  $\mathbb{Q}$ -length 4.

**Theorem 5.7.** *Suppose  $\lambda = \frac{a}{b} \in \mathbb{Q}$ ,  $\lambda^3 \neq \lambda$ . Then  $L_{\mathbb{Q}}(x^4 + 6\lambda x^2 y^2 + y^4) = 3$  if and only if there exist integers  $(m, n) \neq (0, 0)$  so that*

$$(5.9) \quad \Gamma(a, b, m, n) = 4a^3 b m^4 + (b^4 - 6a^2 b^2 - 3a^4) m^2 n^2 + 4a^3 b n^4$$

is a non-zero square.

*Proof.* By Corollary 2.2, such a representation occurs if and only if there is a cubic  $h(x, y) = \sum_{i=0}^3 c_i x^{3-i} y^i$  which splits over  $\mathbb{Q}$  and satisfies

$$(5.10) \quad c_0 + \lambda c_2 = \lambda c_1 + c_3 = 0.$$

Assume that  $h(x, y) = (mx + ny)g(x, y)$ ,  $(m, n) \neq (0, 0)$  with  $m, n \in \mathbb{Z}$ . If  $g(x, y) = rx^2 + sxy + ty^2$ , then  $c_0 = mr$ ,  $c_1 = ms + nr$ ,  $c_2 = mt + ns$ ,  $c_3 = nt$  and (5.10) becomes

$$(5.11) \quad \begin{pmatrix} m & \lambda n & \lambda m \\ \lambda n & \lambda m & n \end{pmatrix} \cdot \begin{pmatrix} r \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

If  $m = 0$ , then the general solution to (5.11) is  $(r, s, t) = (r, 0, -\lambda r)$  and  $rx^2 - \lambda ry^2$  splits over  $\mathbb{Q}$  into distinct factors iff  $\lambda$  is a non-zero square; that is, iff  $ab$  is a square, and similarly if  $n = 0$ . Otherwise, the system has full rank since  $\lambda^2 \neq 1$  and any solution is a multiple of

$$(5.12) \quad rx^2 + sxy + ty^2 = (\lambda n^2 - \lambda^2 m^2)x^2 + (\lambda^2 - 1)mnxy + (\lambda m^2 - \lambda^2 n^2)y^2.$$

The quadratic in (5.12) splits over  $\mathbb{Q}$  into distinct factors iff its discriminant

$$(5.13) \quad 4\lambda^3 m^4 + (1 - 6\lambda^2 - 3\lambda^4)m^2 n^2 + 4\lambda^3 n^4 = b^{-4}\Gamma(a, b, m, n)$$

is a non-zero square in  $\mathbb{Q}$ . □

In particular, we have the following identities:  $\Gamma(u^2, v^2, v, u) = (u^5 v - uv^5)^2$  and  $\Gamma(uv, u^2 - uv + v^2, 1, 1) = (u - v)^6 (u + v)^2$ , hence  $L_{\mathbb{Q}}(f_\lambda) = 3$  for  $\lambda = \tau^2$  and  $\lambda = \frac{\tau}{\tau^2 - \tau + 1}$ , where  $\tau = \frac{u}{v} \in \mathbb{Q}$ ,  $\tau \neq \pm 1$ . These show that  $L_{\mathbb{Q}}(f_\lambda) = 3$  for a dense set of rationals in  $[-\frac{1}{3}, \infty)$ . These families do not exhaust the possibilities. If  $\lambda = \frac{38}{3}$ , so  $f_\lambda(x, y) = x^4 + 76x^2 y^2 + y^4$ , then  $\lambda$  is expressible neither as  $\tau^2$  nor  $\frac{\tau}{\tau^2 - \tau + 1}$  for  $\tau \in \mathbb{Q}$ , but  $\Gamma(38, 3, 2, 19) = 276906^2$ .

We mention two negative cases: if  $\lambda = \frac{1}{3}$ ,  $\Gamma(1, 3, m, n) = 12(m^2 + n^2)^2$ , which is never a square, giving another proof that  $L_{\mathbb{Q}}((x^2 + y^2)^2) = 4$ . If  $\lambda = \frac{1}{2}$ , then

$$\Gamma(1, 2, m, n) = 8m^4 - 11m^2 n^2 + 8n^4 = \frac{27}{4}(m^2 - n^2)^2 + \frac{5}{4}(m^2 + n^2)^2,$$

hence if  $L_{\mathbb{Q}}(x^4 + 3x^2y^2 + y^4) = 3$ , then there is a solution to the Diophantine equation  $27X^2 + 5Y^2 = Z^2$ . A simple descent shows that this has no non-zero solutions: working mod 5, we see that  $2X^2 = Z^2$ ; since 2 is not a quadratic residue mod 5, it follows that  $5 \mid X, Z$ , and these imply that  $5 \mid Y$  as well. Thus,  $L_{\mathbb{Q}}(x^4 + 3x^2y^2 + y^4) = 4$ .

Solutions of the Diophantine equation  $Am^4 + Bm^2n^2 + Cn^4 = r^2$  were first studied by Euler; see [16, pp.634-639] and [38, pp.16-29] for more on this topic. This equation has not yet been completely solved; see [6, 13]. We hope to return to the analysis of (5.9) in a future publication.

## 6. OPEN QUESTIONS

We are confident that Conjecture 4.12 can be completely settled. This raises the question of whether there exist other fields besides  $\mathbb{C}$  (and possibly  $\mathbb{R}$ ) for which there is a simple description of  $\{f : L_K(f) = \deg f\}$ .

Which cabinets are possible for binary forms? Are there other restrictions beyond Corollary 5.1(1)? How many different lengths are possible? If  $|\mathcal{C}(f)| \geq 4$ , then  $d \geq 7$ . Can anything more be said about forms in  $n \geq 3$  variables?

Can  $f$  have more than one, but a finite number, of  $K$ -minimal representations, where  $K$  is not necessarily equal to  $E_f$ ? Theorem 5.7 might be a way to find such examples.

Length is generic over  $\mathbb{C}$ , but not over  $\mathbb{R}$ . For  $d = 2r$ , the  $\mathbb{R}$ -length of a real form is always  $2r$  in a small neighborhood of  $\prod_{j=1}^d (x - jy)$ , but the  $\mathbb{R}$ -length is always  $r + 1$  in a small neighborhood of  $(x^2 + y^2)^r$ , by [47]. Which combinations of degrees and lengths have interior? Does the parity of  $d$  matter? This question is explored in much greater detail in [15].

## REFERENCES

- [1] J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Algebraic Geom., **4** (1995), 201-222, MR1311347 (96f:14065).
- [2] A. Białynicki-Birula and A. Schinzel, *Extreme binary forms*, Acta Arith. **142** (2010), 219–249, MR2606966 (2011d:12001).
- [3] G. Blekherman, personal correspondence.
- [4] M. Boij, E. Carlini, A. Geramita, *Monomials as sums of powers: the real binary case*, Proc. Amer. Math. Soc. **139** (2011), 3039-3043, arXiv:1005.3050, MR2811260 (2012e:11070).
- [5] M. Brambilla and G. Ottaviani, *On the Alexander-Hirschowitz theorem*, J. Pure Appl. Algebra, **212** (2008), 1229-1251, arXiv:math/0701409, MR2387598 (2008m:14104).
- [6] E. Brown,  $x^4 + dx^2y^2 + y^4 = z^2$ : some cases with only trivial solutions—and a solution Euler missed, Glasgow Math. J., **31** (1989), 297–307, MR1021805 (91d:11026).
- [7] E. Carlini, *Varieties of simultaneous sums of power for binary forms*, Matematiche (Catania), **57** (2002), 83–97, arXiv:math.AG/0202050, MR2075735 (2005d:11058).
- [8] E. Carlini and J. Chipalkatti, *On Waring's problem for several algebraic forms*, Comment. Math. Helv., **78** (2003), 494–517, arXiv:math.AG/0112110, MR1998391 (2005b:14097).
- [9] A. Causa and R. Re, *On the maximum rank of a real binary form*, Annali di Matematica Pura et Applicata (4) **190** (2011), no. 1, 5559, arXiv:1006.5127, MR2747463 (2011k:12001).

- [10] M. D. Choi, Z. D. Dai, T. Y. Lam and B. Reznick, *The Pythagoras number of some affine algebras and local algebras*, J. Reine Angew. Math., **336** (1982), 45–82, MR0671321 (84f:12012).
- [11] M. D. Choi, T. Y. Lam, A. Prestel and B. Reznick, *Sums of  $2m$ th powers of rational functions in one variable over real closed fields*, Math. Z., **221** (1996), 93–112, MR1369464 (96k:12003).
- [12] M. D. Choi, T. Y. Lam and B. Reznick, *Sums of squares of real polynomials,  $K$ -theory and algebraic geometry: connections with quadratic forms and division algebras* (Santa Barbara, CA, 1992), 103–126, Proc. Sympos. Pure Math., **58**, Part 2, Amer. Math. Soc., Providence, RI, 1995, MR1327293 (96f:11058).
- [13] J. H. E. Cohn, *On the Diophantine equation  $z^2 = x^4 + Dx^2y^2 + y^4$* , Glasgow Math. J., **36** (1994), 283–285, MR1295501 (95k:11035).
- [14] P. Comon and B. Mourrain, *Decomposition of quantics in sums of powers of linear forms*, Signal Processing, **53** (1996), 93–107.
- [15] P. Comon and G. Ottaviani, *On the typical rank of real binary forms*, arXiv:0909.4865.
- [16] L. E. Dickson, *History of the Theory of Numbers, vol II: Diophantine Analysis*, Carnegie Institute, Washington 1920, reprinted by Chelsea, New York, 1971, MR0245500 (39 #6807b).
- [17] I. Dolgachev and V. Kanev, *Polar covariants of plane cubics and quartics*, Adv. Math., **98** (1993), 216–301, MR1213725 (94g:14029).
- [18] R. Ehrenborg and G.-C. Rota, *Apolarity and canonical forms for homogeneous polynomials*, European J. Combin., **14** (1993), 157–181, MR1215329 (94e:15062).
- [19] W. J. Ellison, *A ‘Waring’s problem’ for homogeneous forms*, Proc. Cambridge Philos. Soc., **65** (1969), 663–672, MR0237450 (38 #5732).
- [20] W. J. Ellison, *Waring’s problem*, Amer. Math. Monthly, **78** (1971), 10–36, MR0414510 (54 #2611).
- [21] A. Friedman, *Mean-values and polyharmonic polynomials*, Michigan Math. J., **4** (1957), 67–74, MR0084045 (18,799b).
- [22] A. Geramita, *Inverse systems of fat points: Waring’s problem, secant varieties of Veronese varieties and parameter spaces for Gorenstein ideals*, The Curves Seminar at Queen’s, Vol. X (Kingston, ON, 1995), 2–114, Queen’s Papers in Pure and Appl. Math., **102**, Queen’s Univ., Kingston, ON, 1996, MR1381732 (97h:13012).
- [23] S. Gundelfinger, *Zur Theorie der binären Formen*, J. Reine Angew. Math., **100** (1886), 413–424.
- [24] J. Harris, *Algebraic geometry. A first course*, Graduate Texts in Mathematics, **133**, Springer-Verlag, New York, 1992, MR1182558 (93j:14001).
- [25] U. Helmke, *Waring’s problem for binary forms*, J. Pure Appl. Algebra, **80** (1992), 29–45, MR1167385 (93e:11057).
- [26] D. Hilbert, *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl  $n$ -ter Potenzen (Waringsches Problem)*, Math. Ann., **67** (1909), 281–300, Ges. Abh. 1, 510–527, Springer, Berlin, 1932, reprinted by Chelsea, New York, 1981.
- [27] P. Holgate, *Studies in the history of probability and statistics. XLI. Waring and Sylvester on random algebraic equations*, Biometrika, **73** (1986), 228–231, MR0836453 (87m:01026).
- [28] A. Iarrobino and V. Kanev, *Power Sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, **1721** (1999), MR1735271 (2001d:14056).
- [29] S. Karlin, *Total Positivity, vol. 1*, Stanford University Press, Stanford, 1968, MR0230102 (37 #5667).
- [30] J. P. S. Kung, *Gundelfinger’s theorem on binary forms*, Stud. Appl. Math., **75** (1986), 163–169, MR0859177 (87m:11020).
- [31] J. P. S. Kung, *Canonical forms for binary forms of even degree*, in *Invariant theory*, Lecture Notes in Mathematics, **1278**, 52–61, Springer, Berlin, 1987, MR0924165 (89h:15037).

- [32] J. P. S. Kung, *Canonical forms of binary forms: variations on a theme of Sylvester*, in *Invariant theory and tableaux (Minnesota, MN, 1988)*, 46–58, IMA Vol. Math. Appl., **19**, Springer, New York, 1990, MR1035488 (91b:11046).
- [33] J. P. S. Kung and G.-C. Rota, *The invariant theory of binary forms*, Bull. Amer. Math. Soc. (N. S.), **10** (1984), 27–85, MR0722856 (85g:05002).
- [34] J. M. Landsberg and Z. Teitler, *On the ranks and border ranks of symmetric tensors*, Found. Comp. Math., **10** (2010), 339–366, arXiv:0901.0487, MR2628829 (2011d:14095).
- [35] G. Mehler, *Bemerkungen zur Theorie der mechanischen Quadraturen*, J. Reine Angew. Math., **63**, (1864), 152–157.
- [36] R. Miranda, *Linear systems of plane curves*, Notices Amer. Math. Soc., **46** (1999), 192–202, MR1673756 (99m:14012).
- [37] L. J. Mordell, *Binary cubic forms expressed as a sum of seven cubes of linear forms*, J. London. Math. Soc., **42** (1967), 646–651, MR0249355 (40 #2600).
- [38] L. J. Mordell, *Diophantine equations*, Academic Press, London-New York 1969, MR0249355 (40 #2600).
- [39] D. J. Newman and M. Slater, *Waring’s problem for the ring of polynomials*, J. Number Theory, **11** (1979), 477–487, MR0544895 (80m:10016).
- [40] I. Niven, *Irrational numbers*, Carus Mathematical Monographs, No. 11. Math. Assoc. Amer., New York, 1956, MR0080123 (18,195c).
- [41] G. Pólya and I. J. Schoenberg, *Remarks on de la Vallée Poussin means and convex conformal maps of the circle*, Pacific J. Math., **8** (1958), 295–234, MR0100753 (20 #7181).
- [42] G. Pólya and G. Szegő, *Problems and theorems in analysis, II*, Springer-Verlag, New York-Heidelberg 1976, MR0465631 (57 #5529).
- [43] V. Powers and B. Reznick, *Notes towards a constructive proof of Hilbert’s theorem on ternary quartics*, Quadratic forms and their applications (Dublin, 1999), 209–227, Contemp. Math., **272**, Amer. Math. Soc., Providence, RI, 2000, MR1803369 (2001h:11049).
- [44] K. Ranestad and F.-O. Schreyer, *Varieties of sums of powers*, J. Reine Angew. Math., **525** (2000), 147–181, MR1780430 (2001m:14009).
- [45] B. Reichstein, *On expressing a cubic form as a sum of cubes of linear forms*, Linear Algebra Appl., **86** (1987), 91–122, MR0870934 (88e:11022).
- [46] B. Reichstein, *On Waring’s problem for cubic forms*, Linear Algebra Appl., **160** (1992), 1–61, MR1137842 (93b:11048).
- [47] B. Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc., **96** (1992), no. 463, MR1096187 (93h:11043).
- [48] B. Reznick, *Uniform denominators in Hilbert’s Seventeenth Problem*, Math. Z., **220** (1995), 75–97, MR1347159 (96e:11056).
- [49] B. Reznick, *Homogeneous polynomial solutions to constant coefficient PDE’s*, Adv. Math., **117** (1996), 179–192, MR1371648 (97a:12006).
- [50] B. Reznick, *Laws of inertia in higher degree binary forms*, Proc. Amer. Math. Soc., **138** (2010), 815–826, arXiv:0906.5559, MR26566547 (2011e:11074).
- [51] B. Reznick, *Blenders*, in *Notions of Positivity and the Geometry of Polynomials* (P. Bränden, M. Passare, M. Putinar, editors), Trends in Math., Birkhuser, Basel, 2011, pp. 345–373, arXiv:1008.4533v1.
- [52] G. Salmon, *Lesson introductory to the modern higher algebra*, fifth edition, Chelsea, New York, 1964.
- [53] J.J. Sylvester, *An Essay on Canonical Forms, Supplement to a Sketch of a Memoir on Elimination, Transformation and Canonical Forms*, originally published by George Bell, Fleet Street, London, 1851; Paper 34 in *Mathematical Papers*, Vol. 1, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1904.

- [54] J. J. Sylvester, *On a remarkable discovery in the theory of canonical forms and of hyperdeterminants*, originally in *Philosophical Magazine*, vol. 2, 1851; Paper 42 in *Mathematical Papers*, Vol. 1, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1904.
- [55] J. J. Sylvester, *On an elementary proof and demonstration of Sir Isaac Newton's hitherto undemonstrated rule for the discovery of imaginary roots*, *Proc. Lond. Math. Soc.* **1** (1865/1866), 1–16; Paper 84 in *Mathematical Papers*, Vol.2, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1908.
- [56] J.-C. Yakoubsohn, *On Newton's rule and Sylvester's theorems*, *J. Pure Appl. Algebra*, **65** (1990), 293–309, MR1072286 (91j:12002).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA,  
IL 61801

*E-mail address:* reznick@math.uiuc.edu